

Detection and Classification of Intestinal Parasites with Bayesian-Optimized Model

Haifa Hamza^{1*}, Kamarul Hawari Ghazali^{2*}, Abubakar Ahmad³

Faculty of Electrical Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan,
Pahang, Malaysia^{1,2}

Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia³

Abstract—Automated detection of intestinal parasites in medical imaging enhances diagnostic efficiency and reduces human error. This study evaluates object detection techniques using Faster R-CNN with different backbone architectures such as ResNet, RetinaNet, ResNext and YOLOv8 series for detecting *Ascaris lumbricoides* and *Trichuris trichiura* in microscopic images. A dataset of 2000 images was split into training (1500), validation (300), and testing (200). Results show Faster R-CNN with RetinaNet achieves the highest Average Precision (AP) across varying Intersection over Union (IoU) thresholds, making it robust in feature extraction. However, YOLOv8 excels in real-time detection, with YOLOv8n (nano) providing the best trade-off between accuracy and computational efficiency. Bayesian Optimization further improves YOLOv8n, achieving an AP of 99.6% and an Average Recall (AR) of 99.7%, surpassing two-stage architectures. This study highlights the potential of deep learning for automated parasite detection, reducing reliance on manual microscopy. Future research should explore transformer-based models, self-supervised learning, and mobile deployment for real-world clinical applications.

Keywords—Intestinal parasites; faster region convolutional neural network; You Look Only Once (YOLOv8); Bayesian Optimization; medical imaging; object detection

I. INTRODUCTION

Intestinal parasitic infections significantly impact global public health, particularly in low-resource and developing regions [1]. Among the most prevalent species are *Ascaris lumbricoides* and *Trichuris trichiura*, which together infect hundreds of millions of individuals worldwide and contribute to malnutrition, cognitive impairments, and socioeconomic challenges [2], [3]. Accurate and timely diagnosis of these infections is essential for effective treatment, surveillance, and public health intervention strategies. Traditional diagnosis via manual microscopic examination, although widely used, is fraught with limitations such as labor intensiveness, inter-observer variability, and significant dependency on expert knowledge [4], [5]. These constraints often lead to delayed diagnoses or misclassification, undermining effective disease control. As such, there is a pressing need for automated, robust, and scalable diagnostic tools that can reliably identify parasite eggs across varying image conditions.

Recent advancements in machine learning (ML) and deep learning (DL) have demonstrated promising capabilities in automating visual diagnostic tasks. However, ML techniques

frequently rely on handcrafted features and struggle with image variability and segmentation challenges. Meanwhile, DL approaches such as CNNs and U-Nets offer improved performance through hierarchical feature extraction, demanding substantial computational resources and large annotated datasets. These resources are often unavailable in the very settings most affected by parasitic diseases [6].

A. Research Problem and Objectives

The central research problem is the lack of real-time, high-accuracy parasite detection tools suitable for resource-constrained clinical settings. Existing models either compromise on computational efficiency or fall short on precision in complex image environments [7]. This study aims to overcome this limitation by identifying and optimizing an object detection architecture that provides a reliable trade-off between accuracy and processing speed.

This research addresses these challenges by proposing a novel, optimized diagnostic solution based on the YOLOv8n model, which is a part of a single-stage object detection framework known for real-time efficiency and accuracy. The core innovation lies in using Bayesian Optimization to fine-tune YOLOv8n's hyperparameters, enabling the model to deliver state-of-the-art accuracy (AP of 99.6%) and recall (AR of 99.7%) with minimal computational overhead.

The study begins with a literature review covering traditional ML and recent DL techniques in parasite detection, highlighting their respective strengths and shortcomings. The methodology section details the dataset, model architecture, and evaluation metrics used in the study. Results are presented comparing various model performances, with a particular focus on the improved YOLOv8n model. Finally, the discussion emphasizes the practical implications of the findings and proposes directions for future research, including the integration of transformer models and mobile deployment

B. Significance and Contributions

This work makes significant contributions to the field of biomedical imaging and parasitology by:

- Systematically comparing both two-stage (Faster R- NN) and single-stage (YOLOv8 series) object detection models across standard benchmarks.
- Demonstrating the superior performance of YOLOv8n for real-time detection in low-resource settings.

- Introducing a Bayesian Optimization framework that enhances the model's performance through intelligent hyperparameter tuning.
- Presenting a detection pipeline that can feasibly be deployed in clinical workflows, thereby reducing diagnostic delays and improving healthcare outcomes.

The findings contribute to advancing automated parasite detection, paving the way for real-time, scalable, and resource-efficient diagnostic solutions in medical and environmental applications.

II. LITERATURE REVIEW

A. Machine Learning

Machine learning techniques have been instrumental in solving some of the challenges in detecting and classifying intestinal parasites. These include Support Vector Machines, BoVW, and Laplacian SVM, among others, which have achieved success in automating parasite classification, enhancing energy efficiency, and solving the out-of-distribution problem in parasite-egg detection [8], [9], [10], [11]. The combination of BoVW with SVM achieved considerable accuracy on classification for various reptilian parasites from stool images, whereas SoftMax thresholds are used for feature selection to deal effectively with out-of-distribution (OO-Do detection).

Most of the ML methods inherently suffer from issues of limited labelled data, manually crafted feature extraction, and dealing with high-dimensional image data despite their successful performance; hence, there is an ever-rising need to develop an automatic feature-learning technique and handle variability in image quality.

Besides, many of these ML models suffer in general from the problem of segmentation, which makes them easily lose their performance on new unseen datasets, and was presented in [9] by Ren et al. The work has thus recently shifted more toward deep learning methods because they have been seen to have the capabilities for high-level feature learning; in addition, deeper learning will be able to capture and model more complex image information data, hence overcoming so much weaknesses related to more conventional ML techniques.

B. Deep Learning

Deep learning emerged as a revolutionary methodology to solve complex problems in the detection, classification, and segmentation of parasites. Models with CNNs, YOLO architectures, and transfer learning strategies have delivered exceptional performance in application scenarios that demand high accuracy and automated feature extraction. For instance, YOLOv5, CNN, have achieved considerable success in the detection of protozoan cysts and helminth eggs and malaria parasites with accuracies mostly greater than 95% in [12], [13]. Besides, deep learning models like U-Net have achieved detection accuracies as high as 99.8% in detecting human intestinal parasites [14].

However, this is not to say that there are no limitations in deep learning. In particular, these include dependencies on large and diverse datasets, high computational costs, and sensitivity to

variations in image quality. Suwannaphong et al., in [15], for instance, recorded a drop in performance upon using low-resolution images from USB microscopes. In addition, some approaches cannot classify morphologically similar types of parasites easily [16]. Some challenges identified include: integrating clinical real-world data sets, improving model architectural robustness, and employing hybrid models to leverage strengths from the different machine learning and deep learning models.

In this work, efforts are made in optimisation of models for resource-constrained settings to enhance generalisability to unseen data. Transfer learning is a method in which pre-trained models are used, especially with small-sized datasets, in order to perform better. For parasite detection, this technique has often been used due to its limited and low-quality dataset [17]. Therefore, transfer learning leverage knowledge from larger and higher quality datasets to enhance feature selection with much better accuracy. Several works, such as [18] and [15], have shown success in using transfer learning methods to improve the accuracy of parasite detection models.

C. Hybrid and Ensemble Learning

Some intractable problems in parasite detection are being tried to be overcome by the hybrid and ensemble learning methods combining the powers of ML and DL. Among these techniques, some methods like VGG16 along with SVM and some ensemble approaches, such as ResNet50 with DenseNet201, have outperformed all previous works related to intestinal and blood parasite classification. For example, Bhuayan and Islam in [19], reported 97.92% accuracy using a hybrid model for detecting protozoa and helminth eggs. Ensembles of CNNs and traditional ML classifiers have also performed well in addressing variability in feature extraction and boosting the accuracy over multi-class tasks in works such as [20] and [21].

Although ensemble methods tend to give higher accuracy, there is usually an added problem of computational complexity and high training times, a process that was noted by Butploy et al. in [22]. Therefore, hybrid models are computationally expensive and in some cases constitute a major source of concern, especially within resource-poor clinical areas.

These challenges further raise the need for refined research on ensemble methods to reduce computational demands and involve sophisticated optimization techniques, such as quantum learning or lightweight models. This analysis underlines the movement from traditional ML techniques to advanced DL and hybrid methods. This reflects the unruffled effort that has gone into overcoming the challenges of parasite detection to improve upon the accuracy, efficiency, and scalability of the approach.

III. METHODOLOGY

The methodology outlines the systematic approach undertaken to evaluate the performance of state-of-the-art object detection models in detecting intestinal parasites. This section describes the dataset, the preprocessing steps employed, the models used, and the evaluation metrics applied. The goal is to assess and compare the effectiveness of the models in classifying and detecting two classes of parasites, *Ascaris*

lumbricoides and *Trichuris trichiura*, using robust and reproducible experimental protocols.

A. Dataset Description

A dataset of 2000 microscopic images was used, comprising two classes of intestinal parasites: *Ascaris lumbricoides* and *Trichuris trichiura*. The dataset was divided into; 1500 images as training set, 300 images as validation set and 200 images as testing set. Each image was pre-processed to ensure uniform dimensions and enhanced contrast for optimal model input.

B. Models Evaluated

This research attempts to optimize the best-performing models among the established baseline models. Object detection for parasite identification, specifically *Ascaris lumbricoides* and *Trichuris trichiura*, requires a balance between detection accuracy and inference speed. Object detection architectures fall into two categories: two-stage and single-stage models. Two-stage architectures, such as Faster R-CNN (FRCNN), excel in precision but often suffer from higher computational costs. Conversely, single-stage architectures, such as the YOLOv8 series, prioritize real-time detection with competitive accuracy. Ensemble learning leverages multiple models trained on the same dataset, combining their predictions to enhance precision and reduce variability. The individual models trained include:

1) *Two-Stage architectures*: Faster R-CNN (FRCNN) is a well-established two-stage detection framework that provides high detection accuracy by first generating region proposals and then refining predictions. To enhance performance, several backbone architectures and frameworks have been integrated with FRCNN:

a) *Faster RCNN with ResNet Backbone*: Utilizes ResNet for feature extraction, known for its accuracy and efficiency in hierarchical feature learning. Both ResNet_50_FPN and ResNet_101_FPN were used in the experiment.

b) *Faster RCNN with RetinaNet Backbone*: Incorporates RetinaNet's focal loss function to address class imbalance, ensuring precise detection of small and irregularly shaped objects.

c) *FRCNN with ResNeXt backbone*: ResNeXt's grouped convolution structure was used to improve feature representation and classification.

These configurations provide robust detection performance but may introduce computational overhead, limiting real-time applications in field environments.

2) *Single-Stage architecture*: The YOLOv8 series offers a single-stage alternative with five model sizes: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large) and YOLOv8x (extra-large), balancing accuracy and efficiency. Single-stage models eliminate the region proposal step, allowing for faster inference while maintaining high detection precision. In this study, YOLOv8n performed better than other variants with the parasite datasets used for this study. The optimal trade-off between speed and accuracy, made it even more suitable for real-time parasite detection with limited hardware resources. YOLOv8n's advantages include:

- Efficient feature extraction using CSPDarkNet backbone.
- Improved object localization through anchor-free detection.
- Optimized performance on edge devices for real-world applications.

The different architectures are summarized in Fig. 1.

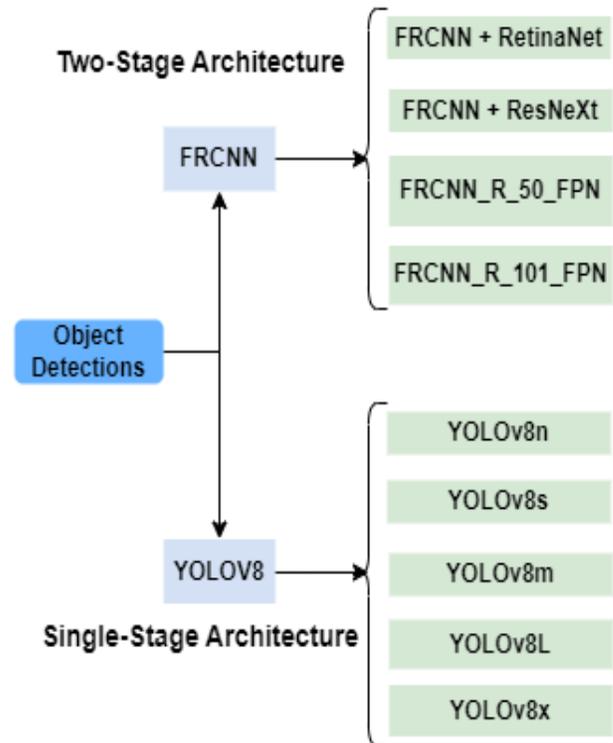


Fig. 1. The different Object detection models utilized in intestinal parasite detection and classification tasks.

3) *Selection of YOLOv8n for optimal performance*: A comprehensive evaluation was conducted by training both two-stage and single-stage object detection architectures on the same dataset to identify the most effective model for parasite detection. The two-stage Faster R-CNN (FRCNN) framework was tested with multiple backbone architectures, including ResNet_50, ResNet_101, RetinaNet and ResNeXt, each offering high detection accuracy but at the cost of increased computational complexity.

In contrast, the single-stage YOLOv8 series, comprising five model sizes (YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x), provided a real-time alternative with improved inference speed and detection precision. Among these, YOLOv8n (nano) was selected as the best-performing model, offering an optimal balance between accuracy and efficiency, making it particularly well-suited for real-time parasite detection in resource-constrained environments. The results of the experiments are presented in the 'Results and Discussion' section of this study.

4) The structure of the proposed optimized YOLOv8n: YOLOv8 represents a significant advancement in real-time object detection, introducing a refined architectural design that enhances accuracy, efficiency, and adaptability over its predecessors. At its core (Backbone), YOLOv8 adopts a CSPDarkNet-inspired backbone, incorporating an advanced Spatial Pyramid Pooling (SPP) module and CSPLayer blocks, which improve gradient flow and reduce computational redundancy, thereby enhancing feature extraction. The core operation involves splitting feature maps and performing transformation separately before merging, as formulated in Eq. (1):

$$X' = F(X_1, \theta) \oplus X_2 \quad (1)$$

where,

- X' is the input feature map, and X_1, X_2 are the split feature maps,
- $F(\cdot, \theta)$ represents the transformation function (e.g, convolution, activation and normalization).
- \oplus denotes concatenation.

In addition, Spatial Pyramid Pooling (SPP) enhances receptive field aggregation by applying multi-scale max pooling as shown in Eq. (2):

$$SPP(X) = \bigoplus_{i=1}^N \max_{r_i}(X) \quad (2)$$

where,

- N represent the number of pooling scales,
- r_i denotes the pooling kernel size at scale i
- $\max_{r_i}(\cdot)$ applies max pooling over a region of size r_i

These components collectively improve feature representation by capturing both fine and coarse-grained spatial structures.

At its neck, the model optimized Path Aggregation Network (PAN) to facilitate multi-scale feature fusion, ensuring the effective integration of fine-grained and high-level semantic information critical for detecting intricate structures such as parasites. Feature fusion in PAN is mathematically expressed in Eq. (3):

$$F_{out} = W_1 * U(F_{in}) + W_2 * D(F_{in}) \quad (3)$$

where,

- F_{in} is the input feature map,
- $U(U(F_{in}))$ and $D(D(F_{in}))$ represent up sampling and down sampling functions, respectively,
- W_1, W_2 are learnable weight parameters
- $*$ denotes convolution

This hierarchical fusion ensures better retention of spatial and contextual information across different scales.

A key departure from previous YOLO variants is the introduction of an adaptive decoupled head, which independently processes classification and regression tasks, improving both localization accuracy and confidence calibration. The classification confidence score and bounding box regression are modelled as shown in Eq. (4) and Eq. (5) respectively:

- **Classification confidence:** The probability of object presence in an anchor-free paradigm is computed using a sigmoid activation in Eq. (4):

$$P(c|X) = \frac{1}{1+e^{-z}} \quad (4)$$

where, z is the output of the classification branch before activation.

- **Bounding box regression:** The predicted bounding box coordinates (x, y, w, h) are obtained using Eq. (5):

$$\hat{x} = x_a + S_x \sigma(x) \quad (5)$$

$$\hat{y} = y_a + S_y \sigma(y)$$

$$\hat{w} = w_a e^{S_w \omega}$$

$$\hat{h} = h_a e^{S_h h}$$

where,

- (x_a, y_a, w_a, h_a) are the anchor box parameters
- $\sigma(\cdot)$ Is the sigmoid function ensuring localization stability.
- S_x, S_y, S_w, S_h are scaling factors learned during training

The decoupling of classification and regression enables YOLOv8 to achieve higher precision and faster convergence compared to prior versions. Furthermore, YOLOv8 transitions to an anchor-free detection paradigm, eliminating the reliance on predefined anchor boxes that characterized earlier versions [23], [24]. This innovation streamlines the detection process, improves generalization, and reduces computational complexity, making it highly effective for parasite detection where object variability is high. The model further enhances performance through an advanced post-processing pipeline, incorporating adaptive non-maximum suppression (NMS) to minimize false positives while maintaining high recall rates.

Additionally, improved loss functions such as IoU Loss and Distribution Focal Loss (DFL) enable superior bounding box regression and confidence estimation. These advancements collectively yield a highly efficient model with reduced inference latency, making YOLOv8 particularly well-suited for real-time and resource-constrained applications in medical and biological imaging. By integrating these state-of-the-art improvements, YOLOv8 establishes itself as a robust framework for precision-driven detection tasks, offering superior speed and accuracy while preserving computational efficiency, making it an optimal choice for high-impact applications such as automated parasite detection [25].

To conclude this section, it is obvious to note that the mathematical formalization of YOLOv8's architectural components underscores its computational efficiency, multi-

scale feature aggregation, and enhanced detection accuracy. The CSPDarkNet backbone facilitates efficient feature extraction, the PAN neck strengthens multi-scale feature fusion, and the decoupled detection head optimizes classification and localization, collectively ensuring state-of-the-art performance in real-time object detection, including applications such as parasite detection in biomedical imaging.

5) *Hyperparameter tuning using Bayesian Optimization:* Bayesian Optimization (BO) has emerged as a superior hyperparameter tuning strategy for deep learning models, particularly in optimizing YOLOv8n for parasite detection, where achieving high precision with minimal computational overhead is critical. Unlike conventional grid search [26], which exhaustively evaluates all possible hyperparameter combinations, or random search [27], which blindly samples the search space, Bayesian Optimization constructs a probabilistic model of the objective function using Gaussian Processes (GPs) or Tree-structured Parzen Estimators (TPE). By iteratively refining this surrogate model and leveraging an acquisition function, such as Expected Improvement (EI), Upper Confidence Bound (UCB), or Probability of Improvement (PI)—Bayesian Optimization dynamically selects the most promising hyperparameter configurations, balancing exploration and exploitation [28].

This adaptive learning process significantly reduces the number of training iterations required to reach an optimal solution while ensuring improved detection performance.

Additionally, Bayesian Optimization mitigates the inefficiencies of traditional methods by intelligently guiding the search space, preventing the combinatorial explosion characteristic of grid search and outperforming the stochastic nature of random search. This results in enhanced sample efficiency, faster convergence, and improved generalization capabilities of YOLOv8n in parasite detection tasks. By integrating Bayesian Optimization into the hyperparameter tuning process, the model achieves superior object detection accuracy with reduced computational costs, making it an ideal choice for real-time and resource-constrained applications in biomedical imaging and parasitology.

Hyperparameter optimization is a critical factor in enhancing the performance of deep learning models for parasite detection, particularly when leveraging Bayesian Optimization to refine the YOLOv8n architecture. By defining a well-structured search space, Bayesian Optimization efficiently navigates the trade-offs between convergence speed, generalization, and computational efficiency. Fig. 2 outlines the Bayesian-Optimized algorithm with YOLOv8n.

The initial learning rate (lr0), constrained within the range of $1e-4$ to $1e-2$ and sampled using a log-uniform prior, governs the magnitude of weight updates, ensuring a balance between rapid convergence and model stability. Momentum, ranging from 0.1 to 1.0, modulates the persistence of past gradients in stochastic gradient descent (SGD), mitigating oscillations and improving convergence stability, particularly in complex parasite detection tasks with highly variable morphological structures.

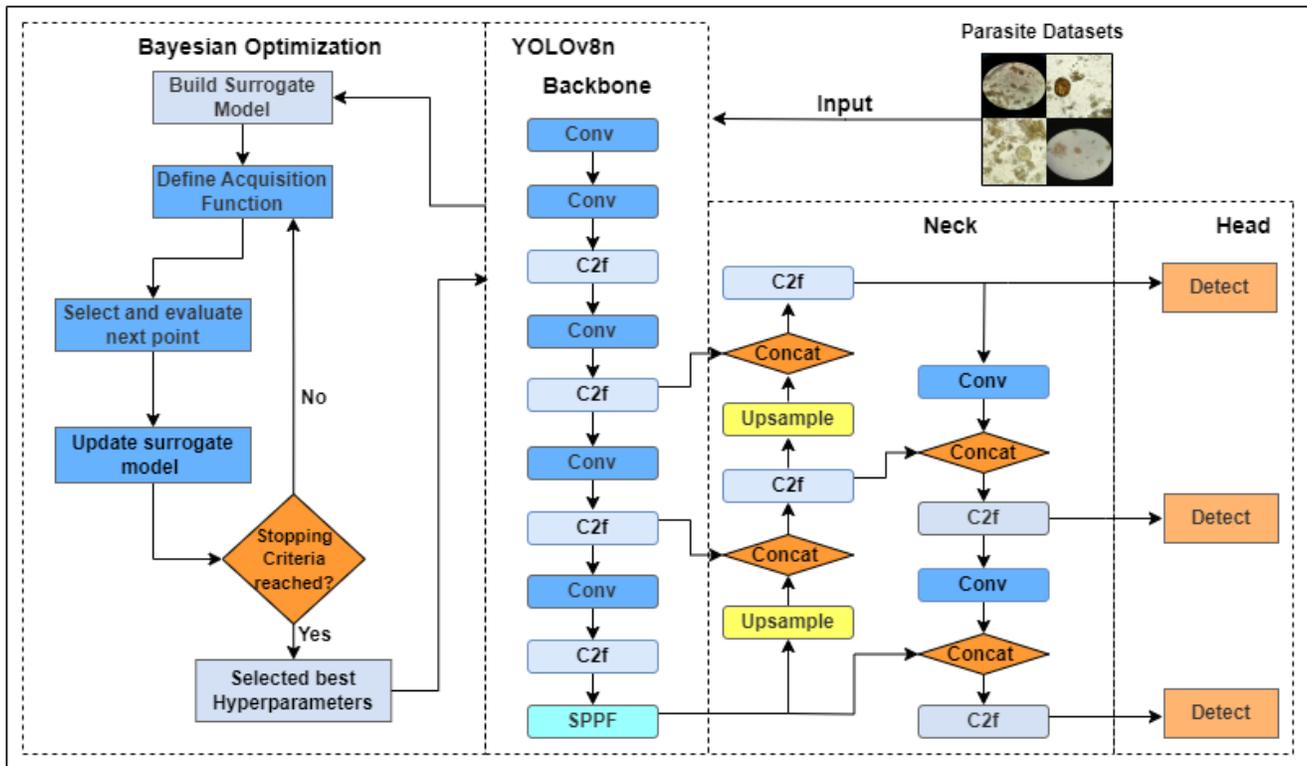


Fig. 2. Performance trends of Average Precision and Recall across varying IoU thresholds, highlighting the consistency and accuracy of detection models in intestinal parasite classification tasks.

The Weight decay (weight_decay), bounded between 0.0 and 0.0005, functions as an L2 regularization term, constraining excessive parameter growth to prevent overfitting and enhance model generalization on unseen parasitic instances. The batch size (batch), selected within the range of 4 to 32, directly impacts gradient estimation, where smaller batches offer improved generalization at the cost of higher variance, while larger batches provide smoother updates but demand greater computational resources.

Additionally, the number of training epochs (epochs), varying from 10 to 1000, determines the duration of model training, requiring careful optimization to balance learning progression with computational efficiency, thereby avoiding underfitting or excessive overfitting. By leveraging Bayesian

Optimization to systematically explore these hyperparameters, YOLOv8n achieves superior detection accuracy while minimizing computational overhead, ensuring robust performance in real-time parasite detection applications.

This intelligent search process dynamically adapts hyperparameter selection based on model performance metrics such as mean Average Precision (mAP) and Intersection over Union (IoU), ultimately facilitating a highly efficient and precise detection framework tailored for biomedical imaging and parasitology research. Table I summarizes the hyperparameter ranges.

The proposed architecture for the Bayesian-Optimized YOLOv8n model is summarized in Table II and the proposed algorithm is presented in Fig. 3.

TABLE I. RANGES FOR HYPERPARAMETER TUNING

Hyperparameter	Abbreviation	Range
Learning rate	lr0	(1e-4, 1e-2)
Momentum	Momentum	(0.1, 1.0)
Weight decay	weight_decay	(0.0, 0.0005)
Batch size	batch	(4, 32)
Number of epochs	epochs	(10, 1000)

TABLE II. THE PROPOSED ARCHITECTURE OF THE PROPOSED OPTIMIZED YOLOV8N

Layer	Output Shape	Filter Size	Number of Filters	Stride	Padding	Activation
0	-1	[3, 16, 3, 2]	16	2	1	ReLU
1	-1	[16, 32, 3, 2]	32	2	1	ReLU
2	-1	[32, 32, 1, True]	64	2	1	ReLU
4	-1	[64, 64, 2, True]	64	2	0	ReLU
5	-1	[64, 128, 3, 2]	128	2	1	ReLU
6	-1	[128, 128, 2, True]	128	2	0	ReLU
7	-1	[128, 256, 3, 2]	256	2	1	ReLU
8	-1	[256, 256, 1, True]	256	1	0	ReLU
9	-1	[256, 256, 5]	256	1	2	ReLU
10	-1	[None, 2, 'nearest']	-	-	-	-
11	[-1, 6]	[1]	-	-	-	-
12	-1	[384, 128, 1]	128	1	0	ReLU
13	-1	[None, 2, 'nearest']	-	-	-	-
14	[-1, 4]	[1]	-	-	-	-
15	-1	[192, 64, 1]	64	1	0	ReLU
16	-1	[64,64,3,2]	64	2	1	ReLU
17	[-1,12]	[1]	-	-	-	-
18	-1	[192,128,1]	128	1	0	ReLU
19	-1	[128,128,3,2]	128	2	1	ReLU
20	[-1,9]	[1]	-	-	-	-
21	-1	[384,256,1]	256	1	0	ReLU
22	[15,18,21]	[2, [64,128,256]]	-	-	-	-

Algorithm: Bayesian Optimization for YOLOv8 Hyperparameter Tuning

Input: Pre-trained YOLOv8 model MMM , Training dataset D_{train} , validation dataset D_{val} , Hyperparameter search space $H=\{lr_0, \mu, wd, B, E\}$, Number of optimization iterations N_{calls} , Random seed s for reproducibility

Output: Optimal hyperparameter set $H^*=\{lr_0^*, \mu^*, wd^*, B^*, E\}$ maximizing validation mean Average Precision (mAP@0.5).

Initialize Parameters:

1 *Set* $D_{train}, D_{val}, D_{test}, P_{results}$
2 *Set* experiment name N_{exp}
3 *Load* pre-trained YOLOv8 model MMM

Define Hyperparameter Search Space

4 **Define** the search space H as follows:
5 $lr_0 \sim \text{LogUniform}(10^{-4}, 10^{-2})$
6 $M \sim \text{Uniform}(0.1, 1.0)$
7 $Wd \sim \text{Uniform}(0.0, 0.0005)$
8 $B \in \{4, 8, 16, 32\}$
9 $E \in \{10, 20, \dots, 1000\}$

Define Objective Function

10 *Given* hyperparameter set H_i , extract batch size B and number of epochs E .
11 *Train* the YOLO model M using:
12 *Dataset:* D_{train} , and Hyperparameter H_i
13 *Perform* model validation on D_{val}
14 *Compute* mAP_{0.5} (Mean Average Precision at IOU threshold 0.5)
15 *Return* – mAP_{0.5} as the objective function value to minimize

Perform Bayesian Optimization

16 **Initialize** Gaussian Process Optimization (GPO) with prior search space H
17 *Set* number of function (e.g, Expected improvement or upper confidence bound)
18 **For** $i = 1$ to N_{calls}
19 | Sample a new hyperparameter set H_i from the search space
20 | Evaluate the objective function using Steps 11-16.
21 | Update the Gaussian Process model with new results
22 **end**
23 *Store* the best hyperparameter set H^*

Output Best Hyperparameters

24 *Extract* optimal values $H^*=\{lr_0^*, \mu^*, wd^*, B^*, E^*\}$
25 *Print* the best hyperparameter values found
26 *Initial Learning Rate* lr_0^*
27 *Momentum* μ^*
28 *Weight Decay* wd^*
29 *Batch Size* B^*
30 *Number of Epochs* E^*

Fig. 3. The Proposed algorithm for the Bayesian-Optimized YOLOv8n model.

C. Evaluation Metrics

The models were evaluated using the following metrics:

Average Precision (AP) at varying IoU thresholds (0.50:0.95). Measures the area under the precision-recall curve, indicating the model's accuracy in detecting objects at varying IoU thresholds as depicted in Eq. (6).

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (6)$$

where, AP is Average precision for class i and

n is the number of IoU thresholds evaluated (e.g., IoU = 0.5, 0.55, ..., 0.95 in 0.05 increments).

Average Recall (AR) across IoU thresholds. Represents the average recall across all IoU thresholds, reflecting the model's ability to detect true positive objects consistently as presented in Eq. (7).

$$mAR = \frac{1}{n} \sum_{k=1}^n AR_k \quad (7)$$

where, R_k is the recall at the k -th IoU threshold and n is the number of IoU thresholds considered.

IV. RESULTS AND DISCUSSION

This section provides a comprehensive analysis of the performance of different object detection architectures evaluated in this study. The assessment focuses on key performance metrics, including Average Precision (AP) and Average Recall (AR) at different Intersection over Union (IoU) thresholds. By comparing the effectiveness of various detection architectures, this section highlights their respective strengths and limitations in detecting *Ascaris lumbricoides* and *Trichuris trichiura*, ultimately informing the selection of robust and scalable diagnostic models.

A. Performance of Faster R-CNN with ResNet Backbone

The Faster R-CNN with ResNet-50 and Feature Pyramid Network (FPN) demonstrated competitive performance, achieving an AP of 85.8% at IoU 0.50:0.95, with a significant increase to 99.6% at IoU 0.50. The model maintained a relatively high AR of 99.6%, indicating strong recall capabilities in detecting true positive instances. However, a noticeable limitation was observed at stricter IoU thresholds, where precision declined, suggesting potential difficulties in accurately localizing objects at higher overlap requirements. This behaviour aligns with previous findings [29], where ResNet-based architectures prioritize robust feature extraction but may struggle in fine-grained localization due to their fixed receptive fields.

The ResNet-101 FPN variant exhibited slightly lower precision compared to ResNet-50, with an AP of 85.5% at IoU 0.50:0.95. Although it maintained a stable AR of 88.9%, it did not provide significant improvements over its shallower counterpart. The marginal performance difference suggests that deeper feature hierarchies introduced by ResNet-101 did not contribute meaningfully to detection accuracy, likely due to diminishing returns in feature extraction depth.

B. Performance Faster R-CNN with ResNeXt Backbone

The ResNeXt-50 backbone offered a moderate improvement over ResNet-based architectures, achieving AP 87.4% at IoU 0.50:0.95, with slightly lower AP values than RetinaNet but outperforming ResNet-50 and ResNet-101. The model maintained an AR of 99.1%, indicating that it effectively captures diverse object instances, leading to a high detection recall. The grouped convolutions in ResNeXt likely contributed to enhanced feature aggregation and spatial sensitivity, allowing the model to detect a broader range of object scales with better contextual understanding. While ResNeXt's performance suggests an improvement in multi-scale feature representation, the relatively small AP gain over ResNet-50 indicates that for this specific detection task, its additional computational complexity does not necessarily translate into a proportionate improvement in detection accuracy.

C. Influence of RetinaNet as a Backbone for Faster R-CNN

The integration of RetinaNet as a backbone for Faster R-CNN led to substantial improvements in detection performance, achieving an AP of 91.1% at IoU 0.50:0.95 and reaching 99.9% AP at both IoU 0.50 and 0.75. The model consistently maintained a high AR of 93.8%, demonstrating exceptional reliability in detecting positive instances across varying IoU thresholds. The superior performance can be

attributed to RetinaNet's balanced handling of foreground and background samples, as its Focal Loss formulation effectively mitigates the imbalance between easily detected and hard-to-detect instances.

The marked increase in AP and AR values indicates that incorporating RetinaNet as a feature extractor enhances feature refinement and region proposal quality, leading to higher detection confidence and better localization accuracy. This underscores RetinaNet's superior feature representation capabilities, particularly in challenging detection tasks involving subtle object variations or occlusions.

D. Comparative Analysis (Two-stage Architecture)

A comparative overview of the evaluated Faster R-CNN models is provided in Table III, summarizing their AP and AR scores at varying IoU thresholds:

TABLE III. AVERAGE PRECISION (AP) OVER DIFFERENT THRESHOLD

Baseline FRCNN at different Threshold (IoU)				
Models	AP	AR	AP @ 50	AP @ 50-95
F-RCNN ResNet_50_FPN +	0.858	0.996	0.996	0.858
F-RCNN ResNet_101_FPN +	0.855	0.889	0.889	0.855
F-RCNN + ResNeXt-50	0.874	0.991	0.991	0.874
F-RCNN + RetinaNet	0.911	0.938	0.999	0.911

From the results, Faster R-CNN with RetinaNet emerges as the most effective architecture, offering the highest AP (91.1%) and AR (93.8%) across varying IoU thresholds. This suggests that RetinaNet's enhanced feature refinement and balanced detection capability make it well-suited for the accurate identification of *Ascaris lumbricoides* and *Trichuris trichiura*.

In contrast, ResNet-50 and ResNet-101 demonstrated similar performance, with ResNeXt offering a slight improvement over ResNet-based variants but falling short of RetinaNet's superior AP and AR scores. While ResNeXt enhances feature learning through grouped convolutions, its computational trade-offs may not justify its minor accuracy gains.

E. Performance of Single-Stage YOLOv8 Architectures

In contrast to the two-stage Faster R-CNN models, single-stage architectures such as YOLOv8 offer a streamlined detection pipeline, eliminating the region proposal step and directly predicting object locations and classifications in a single forward pass. This approach is particularly advantageous for real-time applications where inference speed is critical, such as in automated parasitic detection in medical diagnostics.

The performance of YOLOv8 models was assessed across five different variants, ranging from the smallest YOLOv8n (nano) to the largest YOLOv8x (extra-large), with results presented in Table IV. Among the YOLOv8 variants, YOLOv8n (nano) achieved the highest overall precision, with an AP@50-95 of 93.8%, marginally surpassing YOLOv8x (extra-large) and YOLOv8m (medium), which also scored 93.8%. The YOLOv8l (large) and YOLOv8s (small) models exhibited slightly lower mAP@50-95 (93.6%), indicating that model scaling has

minimal impact on detection accuracy at standard IoU thresholds. Notably, YOLOv8n (nano) achieved the highest recall (mAR = 99.5%), outperforming its larger counterparts. This suggests that even with a reduced parameter count, YOLOv8n maintains strong object detection capabilities, making it an efficient choice for resource-constrained environments. Table IV summarized the results.

TABLE IV. ACCURACY-PERFORMANCE TRADE-OFFS ACROSS YOLOV8 VARIANTS

Baseline YOLOv8 at different Threshold (IoU)				
Models	AP	AR	AP@50	mAP@50-95
YOLOv8x	0.964	0.990	0.993	0.949
YOLOv8l	0.976	0.982	0.993	0.936
YOLOv8m	0.993	0.986	0.995	0.938
YOLOv8s	0.982	0.990	0.991	0.936
YOLOv8n	0.994	0.995	0.994	0.938

Despite being the most computationally intensive model, YOLOv8x did not yield a significant accuracy advantage, achieving a mAP@50-95 of 94.9%, only slightly higher than its smaller counterparts. In contrast, YOLOv8n (nano) emerged as a highly competitive alternative, offering comparable accuracy while delivering superior inference efficiency. This makes YOLOv8n particularly well-suited for embedded medical imaging systems and real-time diagnostic applications, where computational efficiency is paramount.

F. Performance Analysis of Optimized YOLOv8n Model

To further enhance the detection accuracy and efficiency of YOLOv8n, Bayesian Optimization was employed to determine the optimal hyperparameter configuration. This optimization approach efficiently explores the hyperparameter space, balancing the trade-offs between accuracy and computational cost. The key hyperparameters tuned and their respective search ranges are presented in Table II.

The optimized YOLOv8n model achieved an AP of 0.996 and an AR of 0.997, demonstrating near-perfect object detection capability. The exceptionally high recall (0.997) ensures that nearly all instances of *Ascaris lumbricoides* and *Trichuris trichiura* are accurately identified, significantly reducing false negatives and enhancing detection reliability. Compared to the baseline YOLOv8n (AP@50-95 = 0.938), the optimized model achieved an improved AP@50-95 of 0.947, reflecting greater accuracy across varying IoU thresholds. Additionally, AP@50 remained consistently high at 0.995, confirming that the model maintains robust detection performance even under more lenient overlap conditions.

The Bayesian-Optimized training configuration enhanced accuracy without imposing significant computational overhead, making it an ideal choice for real-time diagnostic applications. The fine-tuned learning rate, momentum, and weight decay likely contributed to improved convergence and reduced overfitting, ensuring greater generalizability across diverse detection scenarios. Fig. 4 depicts the training and validation metrics of the optimized model.

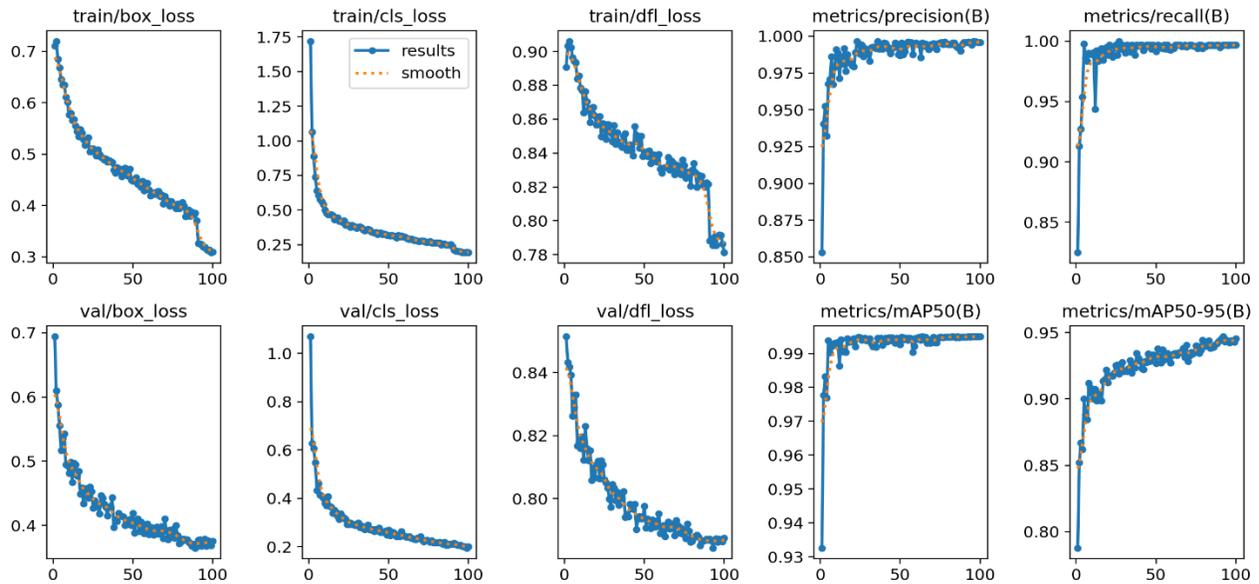


Fig. 4. Training and Validation metrics for optimized YOLOv8n model.

The qualitative detection results in Fig. 5 showcase the model's ability to accurately localize and classify parasite eggs in microscopy images, with predicted bounding boxes and confidence scores reflecting high detection reliability. The Precision-Recall Curve in Fig. 6 further validates the model's robustness, achieving a mean average precision (mAP@0.5) of 0.995 for both *Ascaris lumbricoides* and *Trichuris trichiura*,

highlighting its near-perfect classification capability. The F1-Confidence Curve in Fig. 7, demonstrates the model's optimal F1 score of 1.00 at a confidence threshold of 0.740, indicating a well-calibrated balance between precision and recall. These findings underscore the model's efficacy in automated parasite detection, with significant potential for deployment in diagnostic and epidemiological applications.

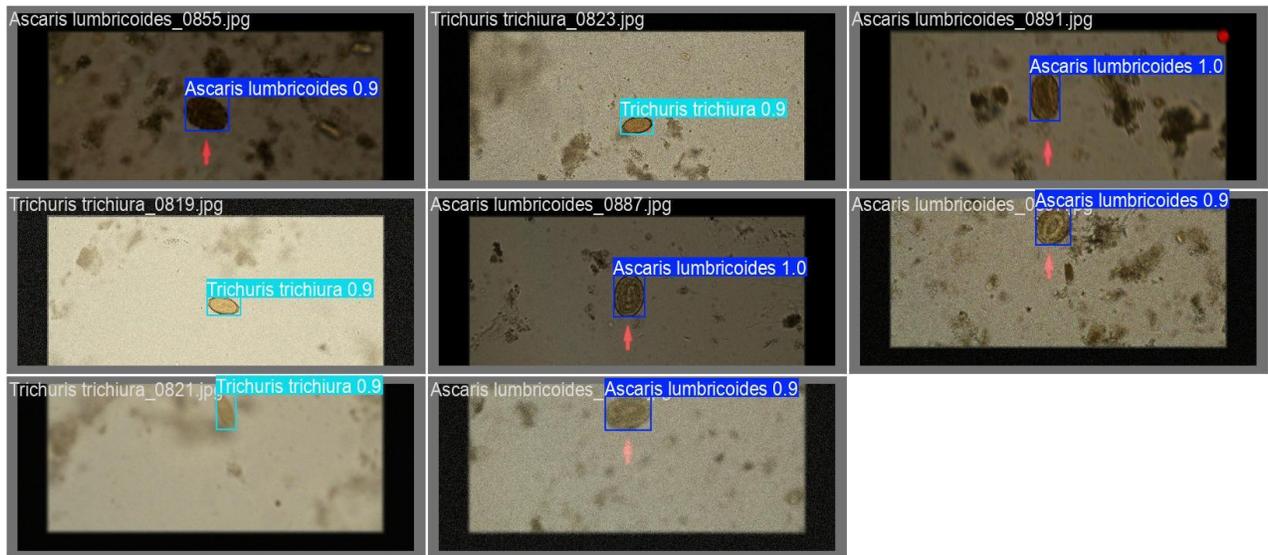


Fig. 5. Detection results of parasite eggs using optimized YOLOv8n: predicted bounding boxes with confidence scores.

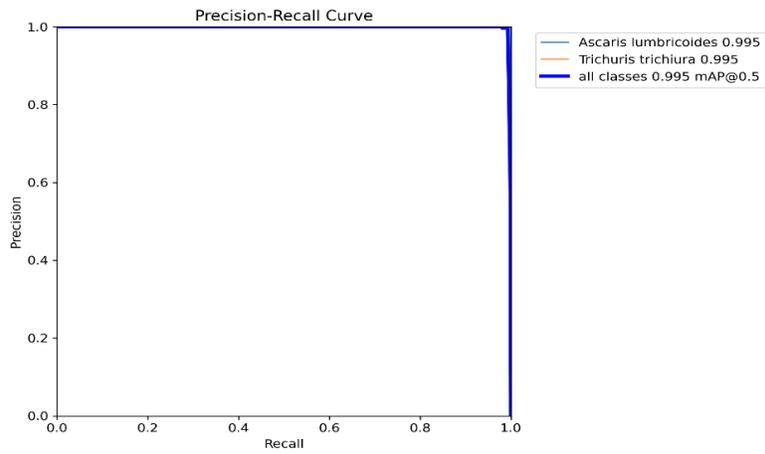


Fig. 6. Precision-Recall Curve for parasite detection: achieving 0.995 mAP@0.5 for all classes.

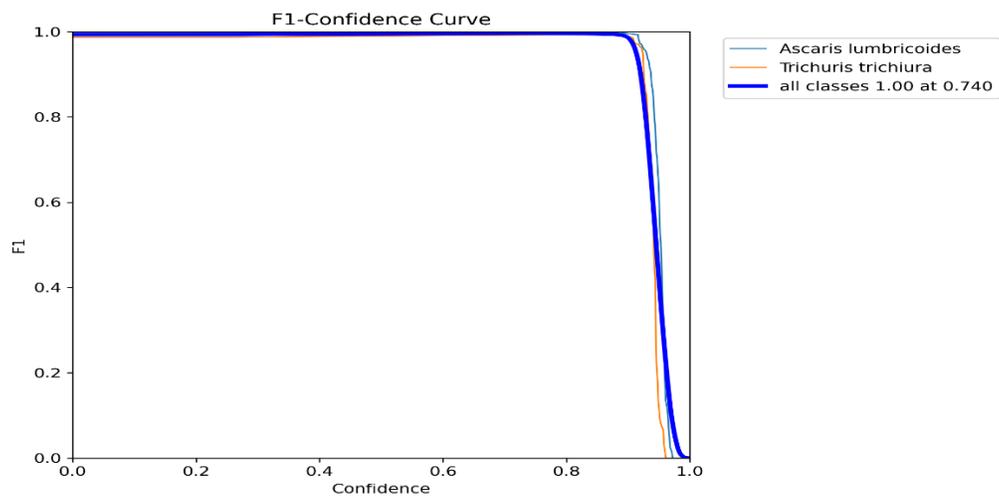


Fig. 7. F1-Confidence Curve for parasite detection: Optimal F1 score of 1.00 at 0.740 confidence threshold.

The optimized YOLOv8n achieved exceptional precision and recall, with results summarized in Table V.

TABLE V. AVERAGE PRECISION (AP) OVER DIFFERENT THRESHOLD

Baselines with Optimized YOLOv8n at different Threshold (IoU)				
Models	AP	AR	AP @50	AP @ 50 - 95
F-RCNN + RetinaNet	0.911	0.999	0.999	0.911
Baseline YOLOv8n	0.994	0.995	0.994	0.938
Optimized YOLOv8n	0.996	0.997	0.995	0.947

The Bayesian-Optimized YOLOv8n demonstrates superior accuracy and efficiency, making it a powerful and practical model for real-time medical diagnostics. Compared to FRCNN with RetinaNet backbone and YOLO counterpart, including the larger YOLOv8 models, it delivers state-of-the-art precision (AP = 0.996) and recall (AR = 0.997) while maintaining its lightweight structure. This highlights the critical role of hyperparameter tuning in enhancing deep learning models for high-stakes applications such as parasitic infection detection. The following graph in Fig. 8 visualizes the trends of mean average precision for the different higher-performing models.

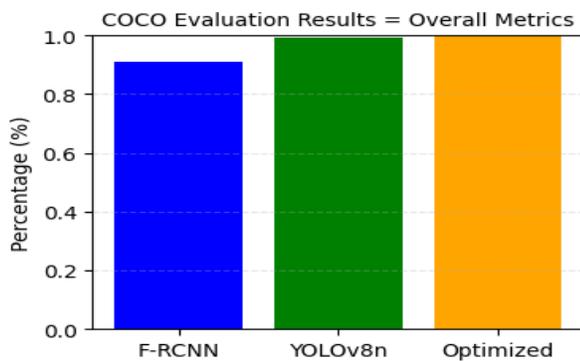


Fig. 8. Performance comparison of detection models (Faster R-CNN with RetinaNet, YOLOv8n and Optimized YOLOv8n) highlighting differences in average precision for intestinal parasite detection.

V. CONCLUSION

This study underscores the significant potential of advanced object detection models in automating intestinal parasite detection. The evaluation of various detection models highlights the optimized YOLOv8n as the best-performing model, achieving the highest AP (0.996), AR (0.997), and AP@50-95 (0.947). Compared to the baseline YOLOv8n, the optimized version demonstrates superior precision and recall, ensuring more accurate and reliable detection across varying IoU thresholds. Furthermore, it outperforms the Faster R-CNN with RetinaNet, which, despite maintaining high recall (0.999), lags in overall precision (AP@50-95 = 0.911).

The Bayesian-Optimized YOLOv8n strikes an optimal balance between detection accuracy and computational efficiency, making it the ideal choice for real-time, high-precision medical diagnostics. Its lightweight architecture, coupled with enhanced performance, positions it as the most

viable model for scalable and resource-efficient deployment in automated parasitic detection systems.

Future research can explore transformer-based enhancements like Swin Transformer to improve feature representation and localization. Self-supervised learning and domain adaptation could further refine performance in real-world clinical settings. Additionally, optimizing the model for edge AI and mobile deployment will enhance scalability for global healthcare applications.

ACKNOWLEDGMENT

This research is supported and funded by the University Malaysia Pahang Al-Sultan Abdullah, Malaysia, under the RDU242743 International Matching Grant.

REFERENCES

- [1] WHO, "Soil-transmitted helminth infections - Fact sheets available at," World Health Organization. Accessed: Jul. 28, 2024. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/soil-transmitted-helminth-infections>
- [2] WHO, "Soil-transmitted helminth infections," World Health Organization. Accessed: Mar. 07, 2025. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/soil-transmitted-helminth-infections>
- [3] WHO, "Schistosomiasis and soil-transmitted helminthiasis: progress report, 2023 = Schistosomiase et géohelminthiasis: rapport de situation, 2023," Weekly Epidemiological Record = Relevé épidémiologique hebdomadaire, vol. 99, no. 48, pp. 707–717, 2024, doi: 10.6665/379641.
- [4] N. Butploy, W. Kanarkard, and P. Maleewong Intapan, "Deep Learning Approach for Ascaris lumbricoides Parasite Egg Classification," J Parasitol Res, vol. 2021, 2021, doi: 10.1155/2021/6648038.
- [5] A. Nuhu Ahmad, M. R. Anis Farhan, A. R. Mohd Faizul, and A. Ahmad, "Distributed Denial of Service Attack Detection in IoT Networks using Deep Learning and Feature Fusion_ A Review," Mesopotamian Journal of Cyber Security, Feb. 2024.
- [6] S. Kumar, H. Vardhan, S. Priya, and A. Kumar, "Malaria detection using Deep Convolution Neural Network," Mar. 2023, [Online]. Available: <http://arxiv.org/abs/2303.03397>
- [7] S. Boit and R. Patil, "An Efficient Deep Learning Approach for Malaria Parasite Detection in Microscopic Images," Diagnostics, vol. 14, no. 23, Dec. 2024, doi: 10.3390/diagnostics14232738.
- [8] N. Penpong, Y. Wanna, C. Kamjanlard, A. Techasen, and T. Intharath, "Attacking the out-of-domain problem of a parasite egg detection in-the-wild," Heliyon, vol. 10, no. 4, Feb. 2024, doi: 10.1016/j.heliyon.2024.e26153.
- [9] Y. Ren, H. Jiang, H. Zhu, Y. Tian, and J. Hu, "A Semi-supervised Classification Method of Parasites Using Contrastive Learning," IEEEJ Transactions on Electrical and Electronic Engineering, vol. 17, no. 3, pp. 445–453, Mar. 2022, doi: 10.1002/tee.23525.
- [10] G. Yanasual, C. Parra, F. Grijalva, D. Benítez, N. Pérez, and V. Párraga-Villamar, "Using Bag-of-Visual Words to Classify Intestinal Parasites in Reptiles from Stool Images," in ECTM 2023 - 2023 IEEE 7th Ecuador Technical Chapters Meeting, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ETCM58927.2023.10309023.
- [11] S. M. Tanko, M. Sani, and A. Ahmad, "Enhancing Bacteria Classification using Image Processing and Convolutional Neural Network," Journal of Basics and Applied Sciences Research, vol. 2, no. 1, pp. 156–161, Mar. 2024, doi: 10.33003/jobasr-2024-v2i1-42.
- [12] J. Amin, M. A. Anjum, A. Sharif, M. Raza, S. Kadry, and Y. Nam, "Malaria parasite detection using a quantum-convolutional network," Computers, Materials and Continua, vol. 70, no. 3, pp. 6023–6039, 2022, doi: 10.32604/cmc.2022.019115.
- [13] K. Satish, A. Tasleem, A. Gulfam, A. C. Anis, K. Salahuddin, and A. M. A. Mohamed, "An Efficient and Effective Framework for Intestinal Parasite Egg Detection Using YOLOv5," Diagnostics, vol. 13, no. 18,

- 2023, Accessed: May 27, 2024. [Online]. Available: <https://www.mdpi.com/2075-4418/13/18/2978>
- [14] I. O. Libouga, L. Bitjoka, D. L. L. Gwet, O. Boukar, and A. M. N. Niôga, "A supervised U-Net based color image semantic segmentation for detection & classification of human intestinal parasites," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 2, Jan. 2022, doi: 10.1016/j.prime.2022.100069.
- [15] T. Suwannaphong, S. Chavana, S. Tongsom, D. Palasuwan, T. H. Chalidabhongse, and N. Anantrasirichai, "Parasitic Egg Detection and Classification in Low-Cost Microscopic Images Using Transfer Learning," *SN Comput Sci*, vol. 5, no. 1, Jan. 2024, doi: 10.1007/s42979-023-02406-8.
- [16] C. Zhang et al., "Deep learning for microscopic examination of protozoan parasites," *Comput Struct Biotechnol J*, vol. 20, pp. 1036–1043, 2022, doi: 10.1016/j.csbj.2022.02.005.
- [17] V. Mezhuyev and F. Pérez-Rodríguez, "METAMODELLING APPROACH AND SOFTWARE TOOLS FOR PHYSICAL MODELLING AND SIMULATION," *International Journal of Computer Systems & Software Engineering*, vol. 1, no. 1, pp. 1–13, Feb. 2015, doi: 10.15282/ijsecs.1.2015.1.0001.
- [18] N. H. Shabrina, S. Indarti, R. A. Lika, and R. Maharani, "A COMPARATIVE ANALYSIS OF CONVOLUTIONAL NEURAL NETWORKS APPROACHES FOR PHYTOPARASITIC NEMATODE IDENTIFICATION," *Communications in Mathematical Biology and Neuroscience*, vol. 2023, 2023, doi: 10.28919/cmbn/7993.
- [19] M. Bhuiyan and M. S. Islam, "A new ensemble learning approach to detect malaria from microscopic red blood cell images," *Sensors International*, vol. 4, p. 100209, Nov. 2023, doi: 10.1016/j.sintl.2022.100209.
- [20] S. A. Ali, D. R. Singamsetty, and P. Kumar, "AN INNOVATIVE ENSEMBLE LEARNING METHODOLOGY FOR THE IDENTIFICATION OF MALARIA USING MICROSCOPIC RED BLOOD CELL IMAGES," *J Theor Appl Inf Technol*, vol. 29, no. 4, 2024, [Online]. Available: www.jatit.org
- [21] D. Osaku, C. F. Cuba, C. T. N. Suzuki, J. F. Gomes, and A. X. Falcão, "Automated diagnosis of intestinal parasites: A new hybrid approach and its benefits," *Comput Biol Med*, vol. 123, p. 103917, Aug. 2020, doi: 10.1016/j.compbimed.2020.103917.
- [22] N. Butploy, W. Kanarkard, P. M. Intapan, and O. Sanpool, "An Approach for Egg Parasite Classification Based on Ensemble Deep Learning," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 27, no. 6, pp. 1113–1121, 2023, doi: 10.20965/jaciii.2023.p1113.
- [23] J. Terven, D.-M. Cordova-Esparza, and J.-A. Romero-Gonzalez, "Comprehensive Review of YOLO Architectures in Computer Vision," *Mach Learn Knowl Extr*, no. 4, Nov. 2023, doi: 10.3390/make5040083.
- [24] J. Wang et al., "Road defect detection based on improved YOLOv8s Model," *Sci Rep*, vol. 14, Jul. 2024, doi: 10.1038/s41598-024-67953-3.
- [25] H. Cho, Y. Kim, E. Lee, D. Choi, Y. Lee, and W. Rhee, "Basic Enhancement Strategies When Using Bayesian Optimization for Hyperparameter Tuning of Deep Neural Networks," *IEEE Access*, vol. 8, pp. 52588–52608, 2020, doi: 10.1109/ACCESS.2020.2981072.
- [26] H. Alibrahim and S. A. Ludwig, "Hyperparameter Optimization: Comparing Genetic Algorithm against Grid Search and Bayesian Optimization," in *2021 IEEE Congress on Evolutionary Computation, CEC 2021 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 1551–1559. doi: 10.1109/CEC45853.2021.9504761.
- [27] L. Villalobos-Arias, C. Quesada-López, J. Guevara-Coto, A. Martínez, and M. Jenkins, "Evaluating hyper-parameter tuning using random search in support vector machines for software effort estimation," in *PROMISE 2020 - Proceedings of the 16th ACM International Conference on Predictive Models and Data Analytics in Software Engineering*, Co-located with ESEC/FSE 2020, Association for Computing Machinery, Inc, Nov. 2020, pp. 31–40. doi: 10.1145/3416508.3417121.
- [28] J. Wu, X. Y. Chen, H. Zhang, L. D. Xiong, H. Lei, and S. H. Deng, "Hyperparameter optimization for machine learning models based on Bayesian optimization," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, Mar. 2019, doi: 10.11989/JEST.1674-862X.80904120.
- [29] S. Kumar, T. Arif, G. Ahamad, A. A. Chaudhary, A. M. , M. Ali, and A. Islam, "Improving Faster R-CNN generalization for intestinal parasite," *Discover Applied Sciences*, May 2024.