



Efficient Microplastic Detection in Edible Food Images Using the Enhanced YOLOv8 with CBAM Architecture

Neha Harde^{1*}, Tulasi B¹

¹ CHRIST (Deemed to be University), Bangalore, Karnataka, INDIA.

* Correspondence: Neha Harde (Email ID: hardeneha.vijay@res.christuniversity.in)

Abstract

Microplastic (MP) pollution in edible food products has become a substantial global health concern that requires easy, accurate, and automated detection methods. Majorly used detection methods, such as FTIR or manual microscopic analysis, are slow and labor-intensive and not suitable for high-throughput screening. To address this limitation, this study proposes an advanced deep learning-based detection framework based on the YOLOv8 architecture with the integration of the Convolutional Block Attention Module (CBAM). A custom dataset was prepared and annotated with optical microscope (OM) images of edible food containing different types of MP, such as fragments, threads, and beads. The experimental results validate that the proposed model is better in small-scale MP detection in terms of precision, recall, and mean Average Precision (mAP). Experimental evaluation demonstrates the effectiveness of the model, achieving an overall precision of 98.32%, a recall of 98.30%, and an mAP of 98.81% across all classes. The model performed best for thread detection with 0.985 mAP, followed by fragments (0.984 mAP) and beads (0.862 mAP). After CBAM integration, the inference process became computationally efficient (an average of 8.1 ms per image of processing speed). The model reported reaching a small deployment size of 52 MB, comprising 218 layers and 25.8M parameters, and was inferred at 78.7 GFLOPs, which is computationally efficient enough to enable MP detection and is best suited for deployment in quality control, laboratory analysis, and environmental screening applications.

Keywords: Microplastic, Food contamination, Deep learning, Automated detection, Detection, Imaging

1. Introduction

Microplastics (MP) are defined as plastic particles less than 5mm, which are mainly generated by improper disposal and degradation of plastic waste. Degradation is the main source of MP generation, and the factors responsible for degradation are UV, weathering, photodegradation, mechanical, biodegradation, and hydrolysis [1]. MP presence has been observed in numerous environments, including seawater [2] [3], land [4] [5], the atmosphere [6] [7], drinking water [8], food packaging material [9] & edible food [10]. MP enters edible food from polluted water bodies, contaminated agricultural soils, airborne particles, plastic packaging, & food carrying item. Irrigating contaminated water or using plastic-laden fertilizers may worsen agricultural contamination. Once in the soil, MPs can be taken up by crops or animals and subsequently enter the human food chain [11]. Processed foods, packaged food items, and even drinking water have been reported to contain substantial amounts of MPs. This further emphasizes that the problem is not limited to food but stretches across the global food supply chain [12].

A growing number of studies have focused on the development of contamination detection techniques in food. With the advancement of AI/ML, the exploration of food contamination has become a new research path [13] [14]. Most MP in edible food objects are tiny, covered with organic media, and tightly packed. This makes it more challenging for detectors to detect MP [10] [15] [16]. Measuring the number of MPs ingested through food is difficult, as the concentrations in the raw ingredients and cooked products are not only highly uncertain but often very low. Extensive sample pretreatment is required to isolate the MP from the samples. The range of solid and liquid food matrices and their varied physical and chemical natures have made detection difficult and prevented the establishment of universal analytical protocols [17]. However, there is still scope for performance improvements compared to state-of-the-art detectors. Thus, there is a growing demand for low-cost, simple detection methods, coupled with automated classifiers, to monitor MPs in food systems.

The challenges illustrate that detection of MPs in food is not just a problem of contamination identification but also a problem of image analysis with the need to discriminate the small size, irregular shape, and low contrast of MPs from the complex and diverse visual information of the foods themselves. This has encouraged the use of AI- and DL-based solutions to automate the detection of MP and enhance the accuracy of particle classification and localization.

To overcome these detection difficulties, recent progress in the area of AI has led to the development of automatic methods for MP detection and analysis. CNN has been extensively used for classification of MP due to its excellent feature extraction ability and object detection systems like YOLO and Faster R-CNN has made it possible to localize and classify individual particles. Recently, the introduction of attention mechanisms, transformer-based networks and lightweight detection architectures has been proposed to enhance the detection of small irregularly shaped MPs in complex imaging scenarios when their size, shape and contrast are low. In recent years, several works have shown the benefits of adding attention modules to DL models in terms of feature discrimination and decreasing the number of false detections, especially for hard MP morphologies [18] [19] [16] [20] [21] [22] [23]. Although numerous developments have been reported, currently available methods are mostly for environmental samples like seawater, sediments, and wastewater, while relatively few studies have been undertaken on food matrices that are consumed. Additionally, the detection of very small and obscure MP particles is still a tricky challenge and requires the development of more accurate and efficient detection models.

Recently, deep learning (DL) models used for object detection have gained significant importance in identifying MPs owing to their accuracy in processing complex visual information. Among these models, YOLO (You Only Look Once) has become the fastest real-time detection framework with excellent speed in inference and feature extraction performance. However, MP particles are mostly irregularly shaped and small, and they can scarcely be inferred normally with the conventional YOLO architecture owing to a lack of contrast. Therefore, to improve their sensitivity for detection, integrated CBAM-based approaches have enabled the model to focus on and differentiate relevant spatial and channel information from others, thus enabling the recognition of MPs influenced by noise [24]. Therefore, the two-part YOLO-CBAM combination is an efficient solution for precise MP detection, with great potential for environmental surveillance and food safety applications.

To address these limitations, this study introduces an improved YOLOv8 architecture integrated with CBAM to assist an optical microscopy measurement framework for automated MP detection. The CBAM mechanism is strategically embedded within the key backbone layers and the final neck stage, which enables the detector to selectively emphasize meaningful MP features while suppressing irrelevant background information. The experimental evaluation showed that the proposed YOLOv8 CBAM framework led to higher precision, recall, and overall detection accuracy compared to the baseline YOLOv8 model, confirming its applicability as a reliable, scalable, and computationally inexpensive tool for the automated detection and monitoring of MPs in food safety applications. Here again, this attention-based enhancement strengthens the model's ability to detect very small, low-contrast, and visually ambiguous particles that are often present in edible food matrices.

2. Methodology

This section presents the proposed research methodology adopted in the study.

2.1. Research Overview

In this paper, we focus on detecting MPs in food products. It begins by generating specialized datasets for accurate & precise detection, then augments the data to enhance training variety, model incorporation, and assessment metrics. Figure 1 Illustrates the architecture used in this study. In this study, an AI framework was proposed to detect MP in edible food products, such as sugar, salt, milk, juices, and honey. Firstly, MPs were separated from food samples by chemical treatment and sedimentation, and optical microscopy was used to capture images. The microscopy images acquired are used as input to YOLOv8-CBAM. A new algorithm has been developed to automatically detect and classify MP particle morphological classes. This workflow is the same as Optical microscopy. The layer receives the prediction results from the AI. The latter part (deep learning layer) provides the prediction results to the AI. The computation measurement enhancement layer is provided by the model. The model uses microscopy images to automatically convert to a framework. This alleviates manual inspection requirements and enables consistent and scalable MP analysis. The particles detected can also be counted and further quantified.

In a further attempt to make the model more robust, data augmentation (e.g., rotation, resizing, brightness) was applied prior to training. The well-known object detection models, YOLOv5 and YOLOv8, are combined with a CBAM to improve object detection accuracy and feature representation, and are evaluated and compared in the current study to improve model performance. Standard evaluation metrics were used to measure the model performance, and the end products were the quantitative performance analysis and visual validation of the identified MPs. The methodology has prospective applications in automated, efficient, scalable, and reliable food safety detection platforms.

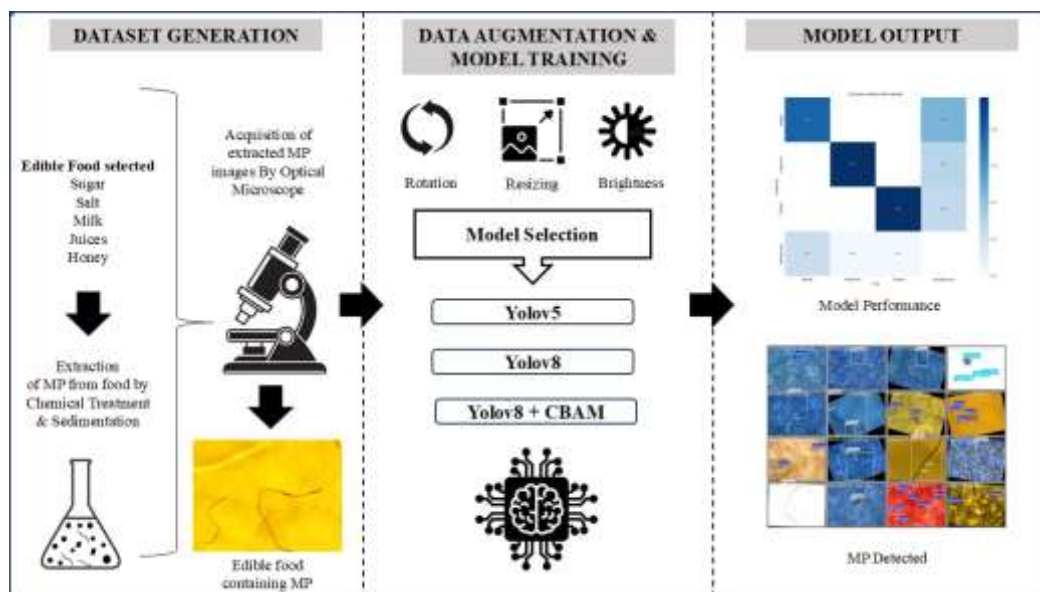


Figure 1. Flowchart of methodology used in the study

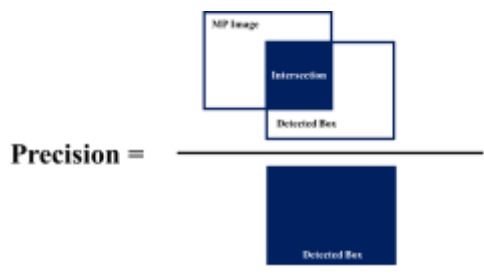
2.2. MP Detection model

A modified version of YOLOv8, a one-stage detector, was used to perform object detection with the help of the Ultralytics framework (v8.1.0) written in Python and PyTorch. CBAM was augmented on the baseline YOLOv8 backbone, and a PAN-FPN neck was added to incorporate joint channel and spatial attention with the goal of making the network focus on discriminative MP details and downplay the background objects. The CBAM-YOLOv8 architecture was trained end-to-end on the manually merged dataset of MP images to 640 x 640 pixels and then annotated with the respective class labels for beads, fragments, and threads. It was trained on Google Colab, using GPU acceleration, using the default YOLOv8 loss (a combination of CIOU-based box regression, objectness, and classification losses) and conventional data augmentations provided by the Ultralytics implementation. The optimal checkpoint was picked by its performance in validation in producing qualitative detection output on held-out microscopic images. YOLOv5, YOLOv8, and YOLOv8+CBAM have been chosen in particular due to their high accuracy, speed of inference, and moderate complexity, and YOLOv8 is located in the intermediate range of the available model sizes. The decision provides optimal MP detection performance in environmental microscopy while maintaining an acceptable level of computational efficiency for use in resource-constrained environments. Comparative analysis established better mAP and F1-scores of the CBAM-enhanced one on the validation set.

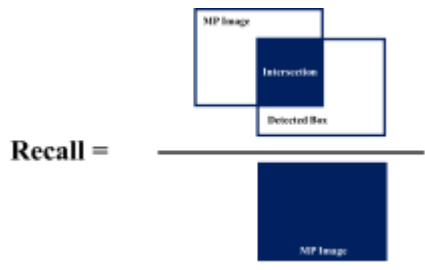
2.3. Performance metric

In order to estimate the model performance, standard YOLOv8 metrics, such as precision (P), recall (R), F1-score, and mean Average Precision (mAP50) at intersection over union (IoU) thresholds of 0.50 (mAP50) and 0.50:0.95 (mAP50-95), were used. This CBAM-YOLOv8 measure is implemented across the MP type identification, that is, beads, fragments, and threads, to show how it improves the ability with respect to false positives and localization accuracy in the validation set. Precision is used to determine the balance between the proportion of positive detections, and Recall is used to evaluate the fraction of MPs identified, and the F1-score offers the harmonic mean of the two to evaluate the balance. mAP50 uses the quality of detection at 50 percent intersection over union, and mAP50-95 is a more stringent evaluation adopting multi-scale criteria in 10 levels of intersection over union between 0.50 and 0.95. The performance metrics, together with confusion matrices, also indicated the strengths of the model in MP morphologies.

Terms like precision (P) and recall (R) metrics provide insight into the accuracy of the model in terms of identifying objects within the image. P is an expression of what the model got right in identifying objects that were correct against all other things it said were positive. R indicates how well the model is able to identify all relevant ground truth bounding boxes and is the ratio of true positives to the total number of all ground truth boxes. The equations for P and R are given as equations 1 and 2, respectively.



$$\text{Precision} = \frac{\text{Intersection}}{\text{Detected Box}} \dots\dots\dots(1)$$



$$\text{Recall} = \frac{\text{Intersection}}{\text{MP Image}} \dots\dots\dots(2)$$

A precision-recall curve was constructed to visualize the correlation of precision and recall for different thresholds. This curve also shows the maximum threshold of optimization of both metrics and is a potential tool for optimizing an object detection model. The AP is then calculated as a summary of the precision-recall curve, indicating the mean of all precisions. The area under the AP curve represents the weighted mean of the precision at each threshold whilst the weight shows how much the recall has increased with respect to the previous threshold. The formula of AP is as follows:

$$AP = \sum_{k=0}^{k=n-1} [\text{Recalls}(k) - \text{Recalls}(k+1)] \times \text{Precisions}(k)$$

As illustrated in the equation below, the mAP is determined by averaging the AP across all classes,

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k$$

AP_k = the AP of class k, n = the number of classes

3. Dataset

This section represents details of the dataset preparation for this study.

3.1. Data Acquisition

A unique custom MP image dataset was generated & selected to mitigate the lack of publicly available real-time datasets. This dataset was generated from edible food products. A wide range of commonly used food ingredients was used as source materials in this collection, including milk, sugar, salt, honey, juices, and rock salt. The establishment of this custom dataset is crucial since it helps bridge a research gap by solidifying an avenue toward the assessment of DL models on real-world and relevant samples. To build a MP classification architecture in edible food, a sample dataset must first be generated, which includes MP separation, MP image acquisition, and image processing. In this study, food samples were procured from local supermarkets in Bengaluru, India. Edible sea salt, rock salt, milk, and sugar were selected on the basis of their daily consumption by human beings. In this manner, a number of brands were selected for study (four each of milk, sea salt, and sugar; two each of rock salt and honey; three samples of tetra pack juices; and one of toned milk). MP extraction procedure was carried out by the test reported by several authors [25] [26] [27] [28] [29] [30]. The details of the experimental procedure are shown in Figure 2.



Figure 2 Scientific procedure for MP extraction from edible food

The MPs extraction process has remained the same as reported in our previous work [18]. This meant that the background of the OM images captured was relatively uniform and the amount of matrix interference was significantly lower. Thus, the proposed model is able to learn the morphology in MP as the primary factor, instead of the food matrix-specific texture. However, matrix-wise annotation and cross-matrix validation would give a more thorough evaluation of model generalization and would be seen as important avenues for future research.

3.2. Dataset generation and Preprocessing

The extracted MP sample was placed on a glass slide, and images were captured at each magnification using a Lica DM4 B OM. Sample images of the extracted MPs with an OM are shown in Figure 3. The images of the MPs were enhanced by changing the brightness, contrast, and color balance to ensure that the recognition by the model images was accurate.



Figure 3. Example of original MPs images extracted from edible foods

The extracted MP samples were kept on clean glass slides, and then optical microscopy pictures were taken with a Leica DM4 B optical microscope at 5×, 10×, 20×, and 50× magnification levels. The same sample handling and image-acquisition pathway was used for the collected food samples to keep the dataset aligned in a consistent way. After extracting the particles, the particles were evenly distributed over the slide, and before recording any image, the microscope was focused on the correct particle plane. The images were taken at the selected magnifications, which would mean that MPs with a variety of particle sizes and irregular shapes would be captured in the images. Following, brightness, contrast, and color balance adjustments were performed during image preprocessing to deliver a better visual representation of the removed particles. The aim of these edits was primarily to increase the visibility of the particles, rather than the morphological properties of the MP particles. There are several factors that can cause variations in optical microscopy-based MP measurements, such as the way the particles settle. Their orientation on the glass slide, the intensity of light, the stability of the microscope focus, the size of the particles, and even the contrast of the image. These can change the appearance of the particles, particularly for lower contrast MP or smaller particles, where the outlines may be less prominent or more blurred. Those types of variations were minimized by using the identical sample preparation, magnifications, and by maintaining the same image preprocessing.

3.3. Data Augmentation

Data augmentation plays an important role, especially when real-time datasets are limited, by enhancing robustness, variation in MPs, and increasing the effectiveness of the dataset to improve model accuracy. The Albumentation library was used to perform various geometric and color transformations.

Table 1 displays the data-augmentation techniques applied for the preprocessing of MP images. The initial

sample size of 400 images was divided at the sample level into training (320 images), validation (40 images), and test (40 images) sample sets in an 80:10:10 proportion. Data augmentation was only applied to the training set, which was expanded from 320 to 3,973 images, whereas the validation and test sets were not augmented. The obtained final augmented dataset has been divided into three classes of MPs, including Beads, Fragment, and Thread, with an almost equal distribution of classes. In particular, the dataset contains 1,325 Beads images, 1,324 Fragment images, and 1,324 Thread images that minimize the imbalance in classes and bias towards model training.

MP Images dataset augmentation parameters are given in Table 2. The custom merged dataset was annotated with bounding boxes for three MP classes (beads, fragments, and threads) using the LabelImg tool and split into train/validation sets in YOLO format (images and labels directories).

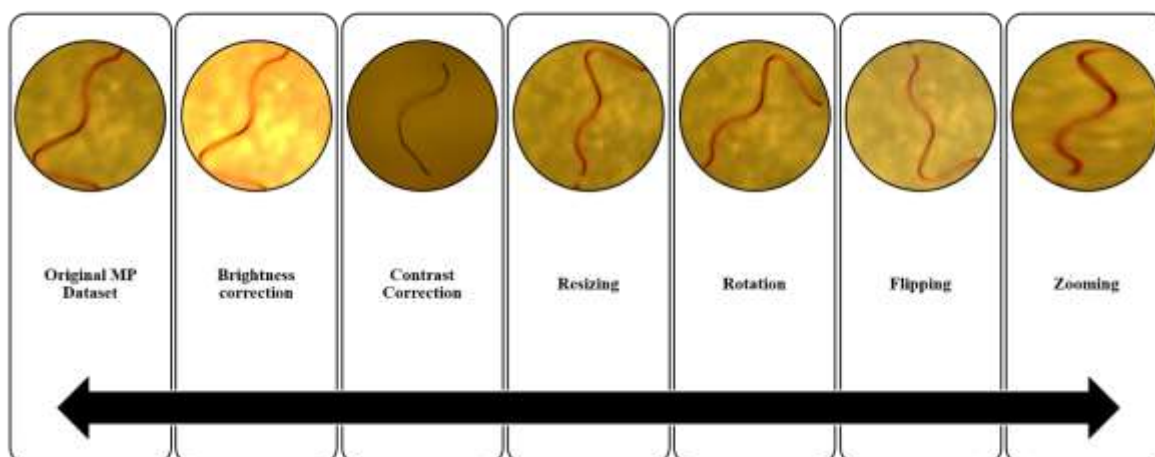
Table 1 MP Images dataset augmentation parameters.

Augmentation	Description
Horizontal Flip	Applied with $p = 0.5$ to generate mirror images, improving robustness to particle orientation.
Vertical Flip	Applied with $p = 0.5$ to simulate inverted orientations of MP particles.
Random 90° Rotation	Randomly rotates images by 90° ($p = 0.5$) to account for orientation variability.
Shift, Scale, and Rotate	Random transformations ($p = 0.5$) involving image translation, scaling, and slight rotations to mimic natural distortions.
Random Brightness/Contrast	Adjusts image brightness and contrast ($p = 0.5$) to address variations in illumination conditions.
Blur	Introduces a blur effect ($p = 0.3$) to replicate out-of-focus or low-quality imaging scenarios.
Zooming	Random zoom-in/zoom-out to enhance scale invariance of particle detection.
Resizing	For consistent model input, all photos were scaled to 640×640 pixels.

In addition, augmentation techniques, such as zooming and flipping, help to boost the model's effectiveness in handling various forms of alteration, as shown in Figure 4. Figure 4 (B) summarizes the spatial and scale distribution of annotated MP instances by showing the distribution of sample amounts across pollutant classes, normalized bounding box size distribution, X-Y spatial distribution of label centroids, and histogram of aspect ratios.

The bar plot indicates that beads are the most frequent class, followed by fragments and threads, revealing a moderate inter-class imbalance, while the scatter plots demonstrate that bounding-box centers cover the image plane and that most objects are relatively small, with width and height concentrated at low normalized values and only a few larger instances along the diagonal. As shown in Figure 4 (C) The complementary pairwise plot of normalized x , y and further illustrates the joint and marginal distributions of these parameters, with heatmap-like panels showing broadly distributed object centers but width and height skewed toward small values, and triangular patterns suggesting that larger boxes occur in a narrower positional range, whereas smaller boxes span the full field of view. Together, these visualizations highlight good spatial coverage but a pronounced bias toward small-sized instances, as well as potential local clustering of annotations in regions of higher density.

A)



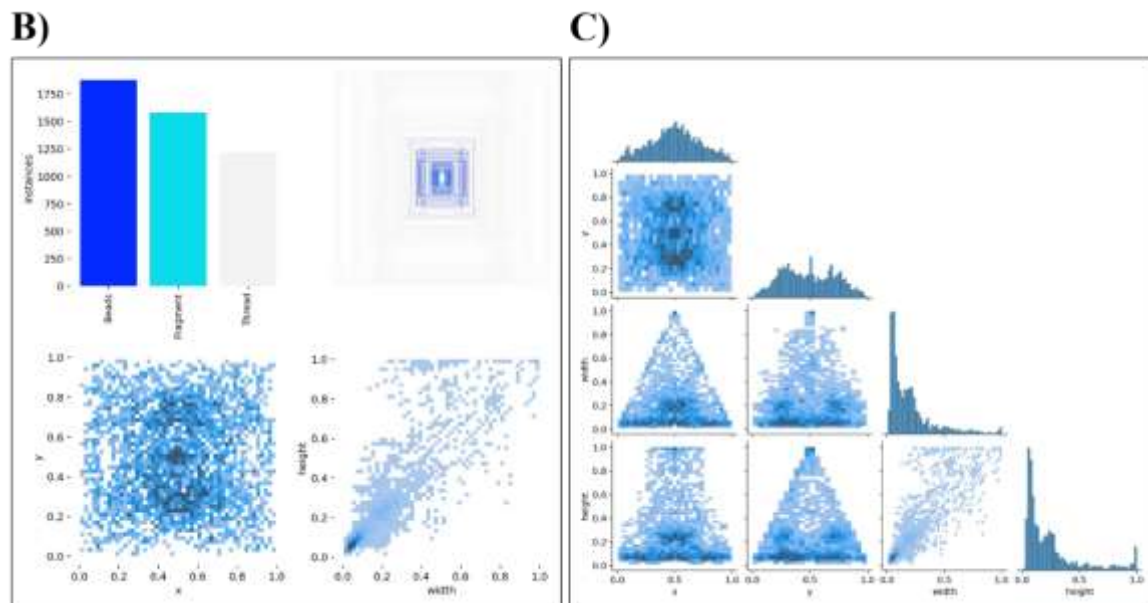


Figure 4. Comprehensive analysis diagram of image dataset (A) data enhancement, (B) label annotation visualization, and (C) division.

4. Experiments

4.1. YOLOv5

YOLOv5 was employed as a strong convolutional baseline owing to its proven balance between accuracy and real-time inference in many applied computer vision tasks. The network adopts an anchor-based detection head and C3 modules in the backbone, combining residual and cross-stage fractional networks to enhance the gradient flow while maintaining a lightweight model. In this study, YOLOv5 was trained on the same MP dataset and under identical hyperparameter settings (input resolution, batch size, optimizer, and training schedule), providing a reference for both accuracy and computational cost against which newer architectures could be compared.

4.2. YOLOv8

YOLOv8 was chosen as the second baseline model because it represents a more advanced model in the YOLO series, incorporating architectural and optimization advances over YOLOv5. In particular, the C3 modules are removed in YOLOv8 in favor of C2f modules; an anchor-free, decoupled detection head is added; and new loss functions, such as Distribution Focal Loss and Task-Aligned Label Assignment, are employed. In experimental conditions, YOLOv8-based models tend to have a higher mAP value compared to YOLOv5 models of a similar or equal size, while also in many cases having a similar inference speed.

4.3. YOLOv8+ CBAM

The standard YOLOv8 model faces a big challenge in the detection of MPs in edible food because, usually, small particles with low contrast blend in with a very complex background. Consequently, the model's feature representations are often quite weak, leading to many MPs being missed.

In addition, features from different scales may not be weighted optimally, which can limit precision-recall performance for fine, fragment-like structures at the model's detection head. To mitigate these issues, the model was augmented with a CBAM in the feature extraction stages, so that channel and spatial attention reweight feature maps toward informative object cues and away from irrelevant background responses. By inserting CBAM blocks into the selected backbone and neck layers while leaving the anchor-free decoupled detection head unchanged, the YOLOv8+CBAM variant preserves the inference efficiency while providing stronger attention-guided feature refinement, leading to improved precision-recall behavior and higher per-class average precision for beads, fragments, and threads in complex backgrounds. The details of YOLOv8 with CBAM architecture are shown in Figure 5.

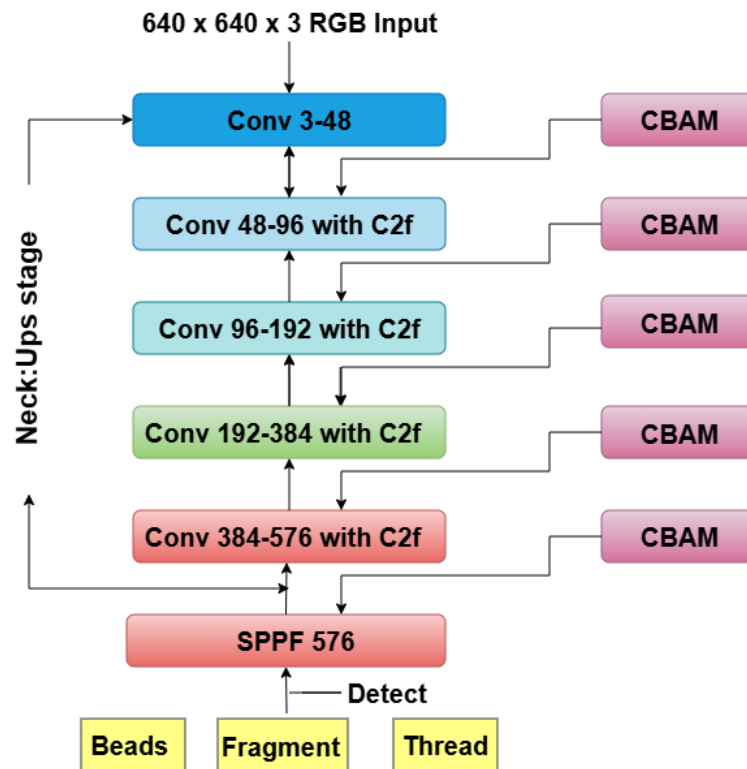


Figure 5. YOLOv8 with CBAM architecture

5. Methodological Approach for Comparison

To systematically assess the performance of the enhanced YOLOv8 model for MP external form detection in edible food samples, we first reproduced and benchmarked YOLOv5 and YOLOv8 on the same dataset. To substantiate the efficacy of the proposed algorithms and maintain experimental rigor, the YOLOv5, YOLOv8, and the YOLOv8+CBAM models were trained and evaluated below identical experimental conditions using the same hardware platform and software configurations, as detailed in

Table 2. Due to the light weight and accuracy, the proposed YOLOv8-CBAM detector is a promising model for MP screening. 218 layers and 25.84 million parameters with 78.7 GFLOPs allow the model to be computationally amenable enough to capture rich features of small targets. A model size of 52 MB with an inference speed of 8.1 ms/image ensures deployment on contemporary GPUs, which is crucial in high-throughput imaging setups. Training for 50 epochs in 1.781 h to reach an mAP50 of 0.944 demonstrates efficient convergence and strong detection performance on the target dataset. The proposed model YOLOv8-CBAM is tested on Google Colab. It has an NVIDIA Tesla T4 GPU. The mean of the average inference latency is 8.1. A real-time inference capability was achieved by running at around 123.5 FPS. With the tested GPU setup. CPU-only inference, embedded deployment (e.g., NVIDIA Jetson), and quantization of models were not assessed in this study and are mentioned as key directions for future work.

Table 2. Model Specs

Attribute	Value
Architecture	YOLOv8 + CBAM
Total Layers	218
Parameters	25,841,497
FLOPs	78.7 GFLOPs
Model Size (best.pt)	52 MB
Training Time	1.781 hours
Epochs	50
Inference Speed	8.1 ms/image
mAP50	0.944

5.1. Cross-Validation & Statistical Testing Setup

To ensure the robustness and generalization capability of the proposed YOLOv8 + CBAM framework, a 5-fold cross-validation strategy was implemented using the KFold method from Scikit-learn with shuffle=True and random state=42. In each fold, the dataset was divided into training and validation subsets: four folds were used for training and one for validation. After training, the best model weights from each fold were

evaluated on the corresponding validation subset. Performance metrics, including precision, recall, F1-score, accuracy, and confusion matrix, were computed using Scikit-Learn's classification report and confusion matrix functions. Macro-average metrics across all folds were calculated to measure model consistency and reliability. To statistically validate the performance improvement of the proposed model, one-sample t-tests and Wilcoxon signed-rank tests were conducted against baseline performance values. Additionally, per-class statistical analysis and Cohen's d effect size calculations were performed to quantify the significance and magnitude of improvement achieved by the YOLOv8 + CBAM framework.

5.2. Integrated Software Systems Development

For ease of operation of the MP detection system, a custom GUI was designed. Gradio, an open-source framework for interactive online applications that uses straightforward browser-based interfaces to facilitate the quick deployment of machine learning models, was used for this purpose. Researchers can easily import sample images for MP detection, and model outputs, such as segmentation masks or projected classes, can be quickly observed. Therefore, it is highly beneficial to confirm the findings and promote collaboration in MP research. The use of the GUI is straightforward and involves very few steps, starting from uploading images, followed by detection, where a YOLOv8-based detection model is applied to detect the MP particles within the image. In addition, there is an option to delete or reset. Figure 6 shows a typical image of MP displayed within the GUI. Figure 6a shows a typical image of MP displayed in the GUI. Figure 6b shows the bounding boxes produced by the YOLOv8 model around each MP particle.

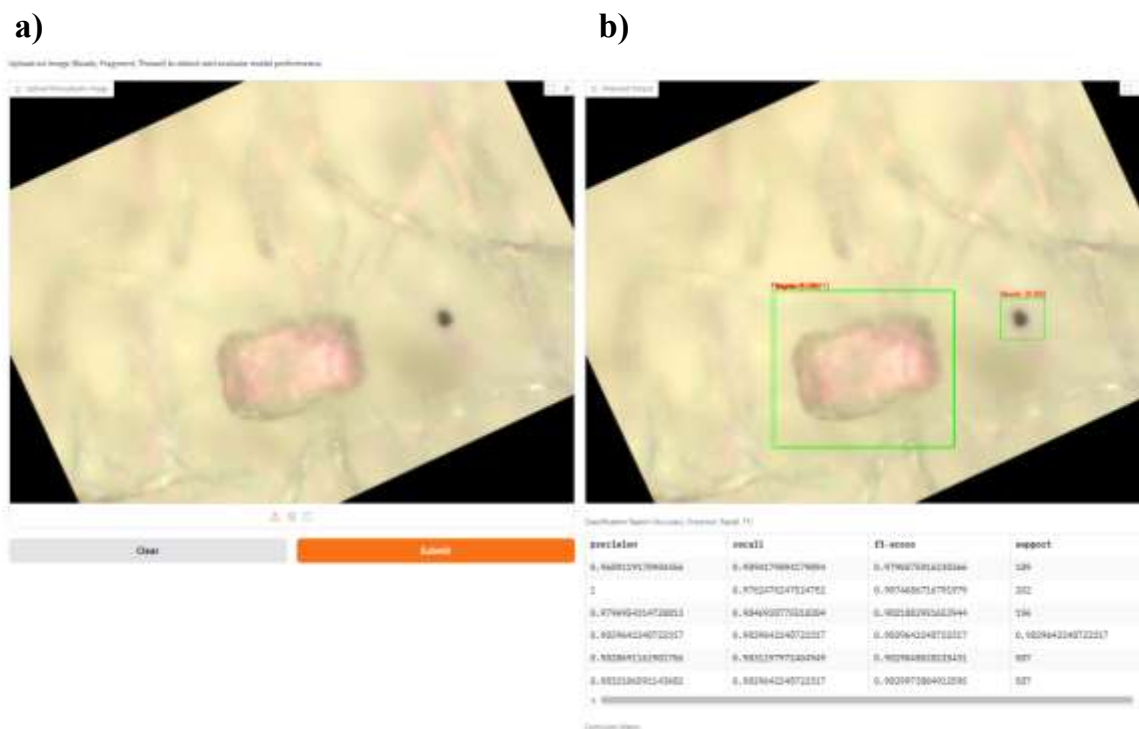


Figure 6. Workflow visualization: (a) raw MP image, (b) YOLOv8 detection results with bounding boxes.

6. Results and Discussion

In the present study, the most comprehensive evaluation of all possible aspects of performance was conducted, along with comparisons between the proposed YOLOv8-CBAM model and the baseline YOLOv8. The proposed detection framework was assessed for this research work using a range of performance metrics in terms of efficiency, accuracy, and robustness. Comparative outcomes give a broad overview of the model's proficiency in comprehensively understanding and distinguishing MP features in complex edible food matrices. A systematic comparison with other approaches demonstrated that the proposed model has better detection capabilities and generalization, thus proving its high reliability and practicality for MP analysis. YOLOv8 is considered a well-established baseline one-stage object detector with superior feature extraction and real-time performance. However, its performance will be found to be influenced by more complicated backgrounds, as well as in detecting smaller or low-contrast MP in edible food matrices. To address these issues, a CBAM was integrated into the YOLOv8 architecture to improve its efficiency in feature representation. Convolutional feature maps are employed by YOLOv8. All spatial and channel information is treated uniformly. Unlike YOLOv8-CBAM, which uses channel attention.

6.1. Model training results analysis

The YOLOv8 model was trained for 50 epochs, with a batch size of 16, and an input image resolution of 640×640 pixels. Also, the YOLOv8 is integrated with CBAM—that is, Convolutional Block Attention Module for enhanced feature learning. The CBAM is intended to augment the channel-wise and spatial feature attention to better concentrate on MP characteristics that could distinguish these particles based on their shapes and textures. Overall, this method offers a decent compromise between speed and accuracy, thereby enabling better recognition of MP varieties among such complex features.

Using the CBAM increased the performance of both models across all virtually excellent thresholds for thread class. The precision improved from 0.935 to 0.979, the Recall improved from 0.962 to 0.984, and mAP@50 went from 0.978 to 0.985, while the same mAP@50-95 barely changed from 0.798 to 0.802. The very small improvements in the mAP@50–95 metric are indicative of improved detection stability for thin, long fiber-like MP. Thus, from the reading results, it is clear that the Thread class achieves the most suitable baseline performance between the two models, wherein the CBAM further improves the model. This improvement in mAP@50–95 indicates an increased stability in detection for thin and elongated fiber-like MP. Overall, the CBAM enhances the detection performance of YOLOv8 at the channel-wise and spatial attention levels, thereby enhancing feature representation and suppressing background interference. The improvements are most pronounced for the irregularly shaped morphologies of MP, which show the suitability of attention-enhanced architectures for reliable MP detection under complex environmental conditions.

Table 3 The classification description shows precision, recall, and F1-score for Beads, Fragments, and Threads using the proposed YOLOv8-CBAM-based MP detection model against YOLOv8, providing a class-wise assessment of the baseline YOLOv8 model and the proposed YOLOv8-CBAM for MP detection. For the Beads class, the proposed model achieves a high precision (0.968) and recall (0.989) compared with YOLOv8 0.918 and 0.962, respectively, indicating better object discrimination and fewer false detections. The reduction in AP50 and AP50–95 for the Beads class may be associated with the relatively small size and near-circular morphology of the particles, which can make precise bounding-box localization more sensitive to IoU-based evaluation. This effect becomes more pronounced for AP50–95, which evaluates detection performance across progressively stricter IoU thresholds up to 0.95. According to YOLOv8-CBAM, there was an observed improvement in the performance of all metrics for the fragment class. The precision improved from 0.875 to 1.00, and the recall increased from 0.686 to 0.975, which proves that the CBAM is successful in capturing the irregular and heterogeneous geometries of fragments. Accordingly, mAP@50 improved from 0.862 to 0.984, and mAP@50-95 improved from 0.588 to 0.691, which shows an enhanced consistency of localization under relaxed and strict conditions of IoU.

Table 3 Class-wise performance comparison of YOLOv8 and YOLOv8+CBAM for MP Detection

	Precision		Recall		mAP50		mAP50-95	
	YOLOv8	YOLOv8+CBAM	YOLOv8	YOLOv8+CBAM	YOLOv8	YOLOv8+CBAM	YOLOv8	YOLOv8+CBAM
Beads	0.918	0.968	0.962	0.989	0.973	0.862	0.68	0.604
Fragment	0.875	1.00	0.686	0.975	0.862	0.984	0.588	0.691
Thread	0.935	0.979	0.962	0.984	0.978	0.985	0.798	0.802

6.2. Performance evaluation

Figure 7 Illustrate the Precision-Recall (PR) curves for evaluating the proposed YOLOv8 model for detecting three forms of MP beads: fragments and threads. On the one hand, precision is defined as the percentage of properly classified MP particles among all classified particles; on the other hand, recall is defined as the percentage of true particles that were actually detected. The model maintains very low false positives and high positive detections across a wide range of recalls for inputs containing all three classes. Achieving the highest average precision (AP) of 0.98, the Fragment class stands in first place, while the Thread class registered an AP of 0.96, and the Beads class followed with an AP of 0.95. From the above results, it can be determined that the model performs effectively in detecting irregular and elongated MP structures due to the enhanced feature provided by the CBAM-based attention mechanism.

In comparison, across almost the entire confidence range, the precision offered by YOLOv8 + CBAM exceeds that of YOLOv8 by a considerable margin, as shown in Figure 8. Moreover, the CBAM-enhanced model maintains near-perfect precision (>0.95) at moderate confidence thresholds, indicating effective false-positive suppression. There is substantial improvement specifically for the fragment and thread classes, which further confirms the benefits of attention-based feature refinement. As recall approaches unity,

precision for all classes declines rapidly. Such an outcome is expected, as pushing the model to detect every potential object increases the likelihood of false alarms, especially in a complex microscopic background. Nevertheless, the overall shape of the curves remains quite close to the top right-hand of the plot, indicating that the classes are well-separated, and the detection abilities are good. Maintaining large areas under the PR curves shows that the integration of the CBAM with the YOLOv8 is efficient for MP detection, particularly under conditions of small size, low contrast, and visually cluttered backgrounds. The YOLOv8+CBAM recall curve shows that recall values are sustained at a higher score until increasing confidence thresholds than the baseline model. Hence, the CBAM module assists in maintaining true detections with more stringent confidence filtering, particularly for MP structures with a high level of visual complexity.

The precision-recall curves attest to the performance improvement in detection capability by the YOLOv8 model when integrated with the CBAM embedded into it. The CBAM augmented model shows an approximate value of 0.944 for the area under the PR curve, based on mAP@0.5, compared to 0.938 obtained by the base YOLOv8 model. The incremental improvements are consistent across classes, especially on thread-like MPs, which are inherently hard to detect, being very thin and elongated, and having fairly low contrast with their background. Another aspect of the confidence-F1 analysis is that, at the best confidence thresholds, YOLOv8+CBAM peaks at an F1-score of around 0.90, slightly higher than a baseline YOLOv8 with an F1-score of about 0.89. This is an improvement towards a better balance between precision and recall, thereby demonstrating the framework's higher reliability with CBAM in realistic detection environments. Overall, the incorporation of CBAM helps the model emphasize better highlighting of MP structures and reduce background noise. Precision, recall, mAP, and F1-score improvements can always be observed consistently for the classes analyzed. In this case, YOLOv8+CBAM proves to be the most robust and effective method for MP detection, as compared to the typical YOLOv8 architecture.

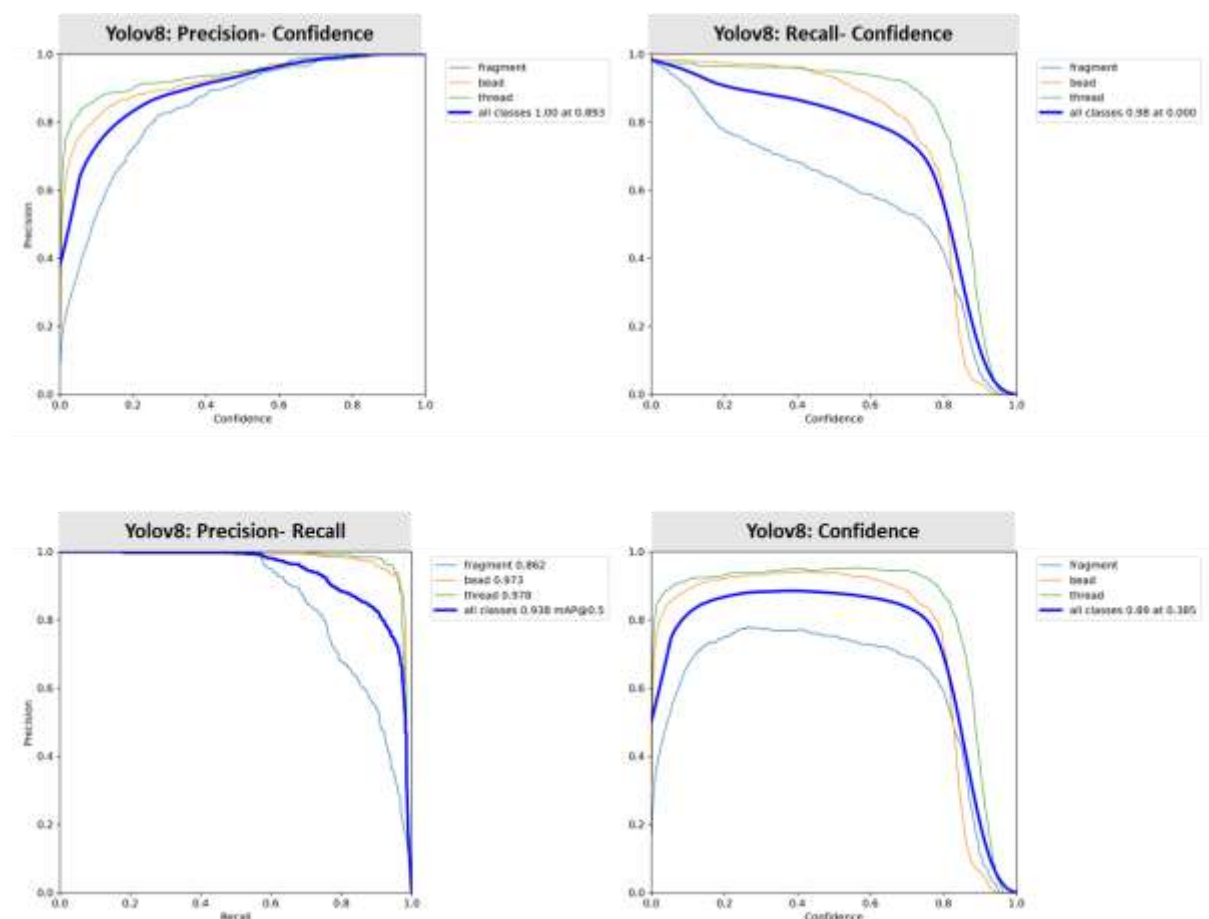


Figure 7. Performance evaluation curves for YOLOv8: Precision, recall, and balanced scores F1 changes.

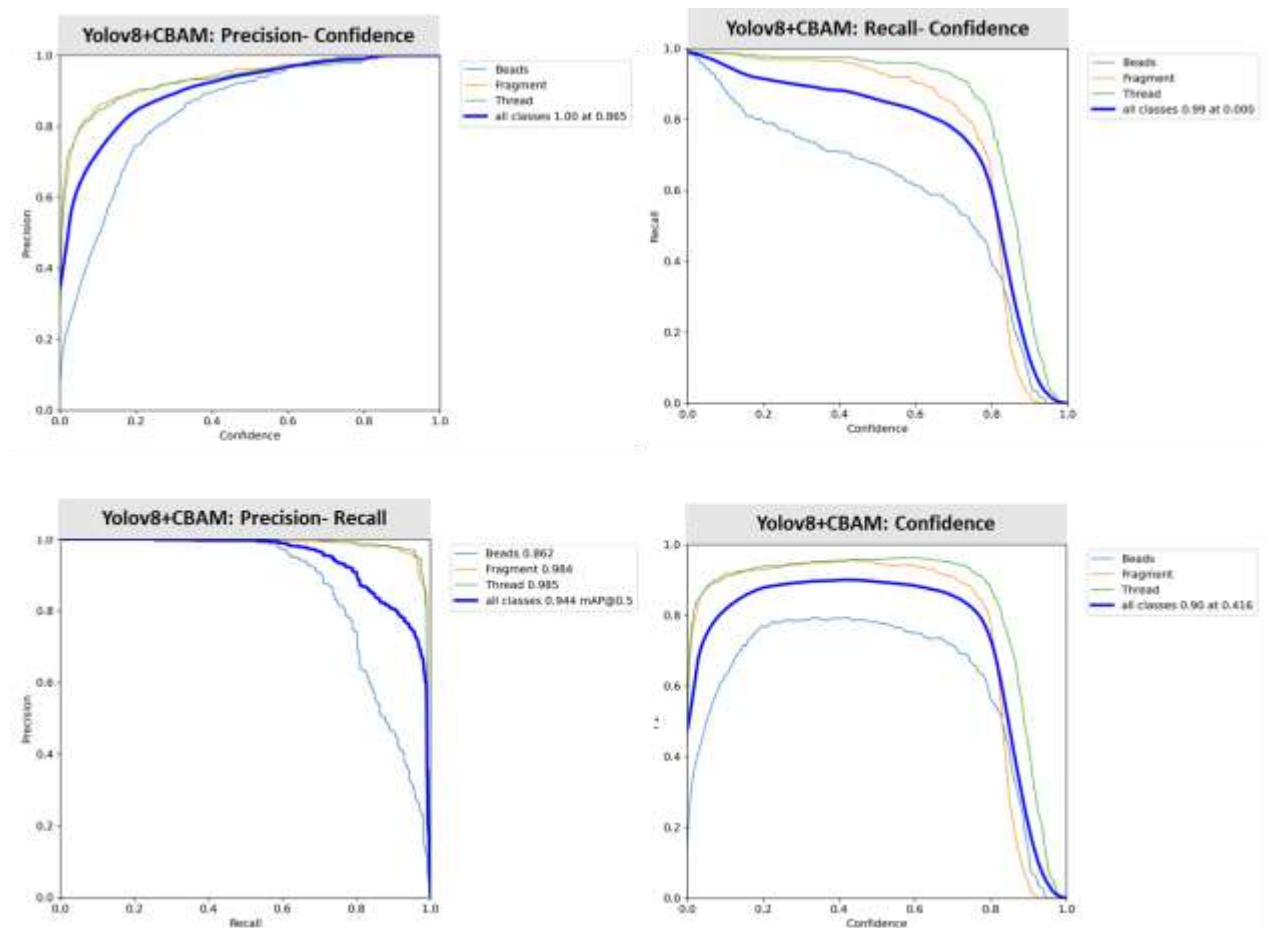


Figure 8. Performance evaluation curves for YOLOv8+CBAM: Precision, recall, and balanced scores F1 changes.

6.2.1. Analysis of the Precision-Recall (PR) Curve

The PR curves reflect the degree of success achieved by the YOLOv8-CBAM model in detecting different types of MPs, alongside its overall performance. The model achieved high average precision on the Fragments (AP = 0.984) and Threads (AP = 0.985) classes, suggesting that it can fairly well and confidently detect across a wide choice of recall levels. The performance of the Beads class is poorer (AP = 0.862) due to a larger degree of variability in the appearance of beads, as they also partially overlap with features of the background. Combined curve from all classes yields an mAP@0.5 of 0.944, indicating a strong level of the overall detection capability. The curves lie near the upper-right side of the figure, indicating that excellent precision remains attainable at higher levels of recall. The outcomes from this show that CBAM-integrated YOLOv8 did exceptionally well in the detection of different MP shapes, even under very tough imaging conditions.

6.2.2. Analysis of the F1–Confidence Curve

The F1 confidence curves show the detection confidence thresholds used by the YOLOv8-CBAM model across various types of MPs. The F1 score for Fragments and Threads remains consistently above most confidence levels, indicating high detection reliability and a fair precision-recall trade-off for the model. The Beads class maintains a lower F1 score than other classes that differ but exhibit greater visual diversity, with certain features overlapping with the background. The inter-class aggregating curve peaks at an F1 score of nearly 0.90, a confidence range of about 0.416. As a result, F1 scores drop beyond this point due to reduced recall at higher confidence levels. All in all, this research shows that it gives the model great overall performance, supporting the selected confidence level for detecting MPs in real scenarios.

6.3. Convergence analysis

This paper compared the performance of the YOLOv8-CBAM model with that of YOLOv8 in order to evaluate the convergence performance of the model on a self-built dataset. The study was aimed at examining the convergence characteristics of the loss functions of these models in training and validation, and used four important metrics, namely, precision, recall, mAP@0.5, and mAP@ [0.5:0.95], to overall evaluate and compare the convergence behavior of the various models.

Performance training and validation results from the 50 epochs of the YOLOv8-based model are shown in

Figure 9. In this evaluation on the convergence of the model for loss functions like box loss, classification loss, as well as the distribution focal loss, with the principal detectors of precision, recall, and average precision (mAP). During the conduction of the training phase of the model, the loss of the bounding box (train/box loss), classification loss (train/cls loss), as well as the distribution focal loss (train/df loss) depict a consistent decline on the graph. The steady decline shows stable optimizations, and gradient instability is absent. Similarly, it indicates this for the validation losses (val/box loss, val/cls loss, and val/df loss), following the same trend with training curves. The maximum potential of generalization is replicated by a minimum gap between both training and validation losses, thereby avoiding overfitting in the model. The precision and recall show a gradual upward trend as training epochs increase, reaching high and stable values toward the last epochs. This indicates an enhanced capacity of the model to appropriately categorize MP particles, avoiding wrong detections and missed objects. Along the same lines, we see improvement in mAP@0.5 and mAP@0.5:0.95 training scores, stabilizing at much higher values. This attests to reliable detection accuracies for various IoU thresholds and the strength of the model in detecting both coarse and fine-grained MP features. This validates that the model is proficient of detecting both coarse- & fine-grained MP particles with stable convergence.

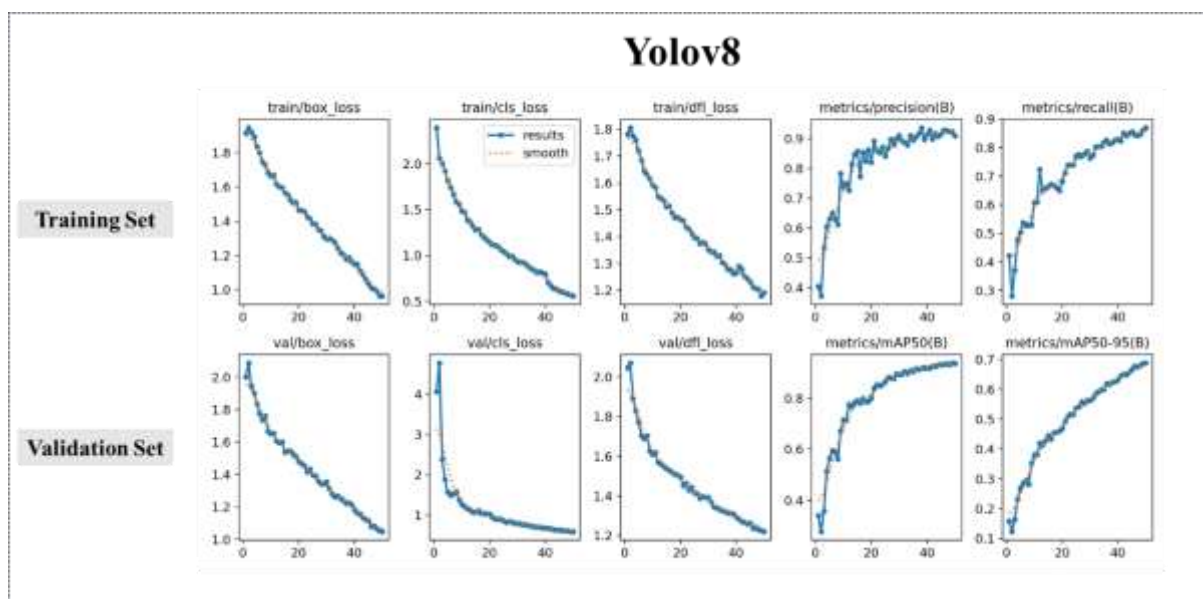


Figure 9. Training and Validation Performance Curves of the YOLOv8 Model

The results of training and validation curves derived from the second experimental setup of the YOLOv8 model are shown in Figure 10. The trends observed in the loss functions and evaluation metrics reinforce the robustness and stability of the proposed detection framework. The training losses for box loss, classification loss, and distribution focal loss exhibit a steady and monotonic decay throughout the training epochs, indicating good learning of object boundaries, class separation, and bounding box regressions, which led to improved performance in localization and classification over time.

For the validation set, the box loss and DFL loss decrease smoothly and closely track the training curves, reflecting strong generalization capability. The validation classification loss starts off with a high value at the beginning of training; however, its value stabilizes shortly afterward and keeps decreasing thereafter.

This is mainly attributable to class imbalance and the learning of features in the early stages, which was resolved as the model began to learn to distinguish MPs' features. Both precision and recall curves increase sharply during the first two epochs, and their values settle at higher points. This is the way the ever-growing capability of the model to locate MP particles correctly is being depicted, as the number of false positives and false negatives is being minimized. At the same time, mAP@0.5 and mAP@0.5:0.95 have shown gradual enhancement and later balance, signifying good detection performances across various IOU thresholds. Steady metrics prove the reliable capacity of the presented model for the recognition of abundant and subtle instances of MPs. Trends observed to date indicate that, during the course of this work, the YOLOv8 model achieves stabilization with good generalization and high detection accuracy, thereby corroborating its applicability to MP detection and quantitative analyses.

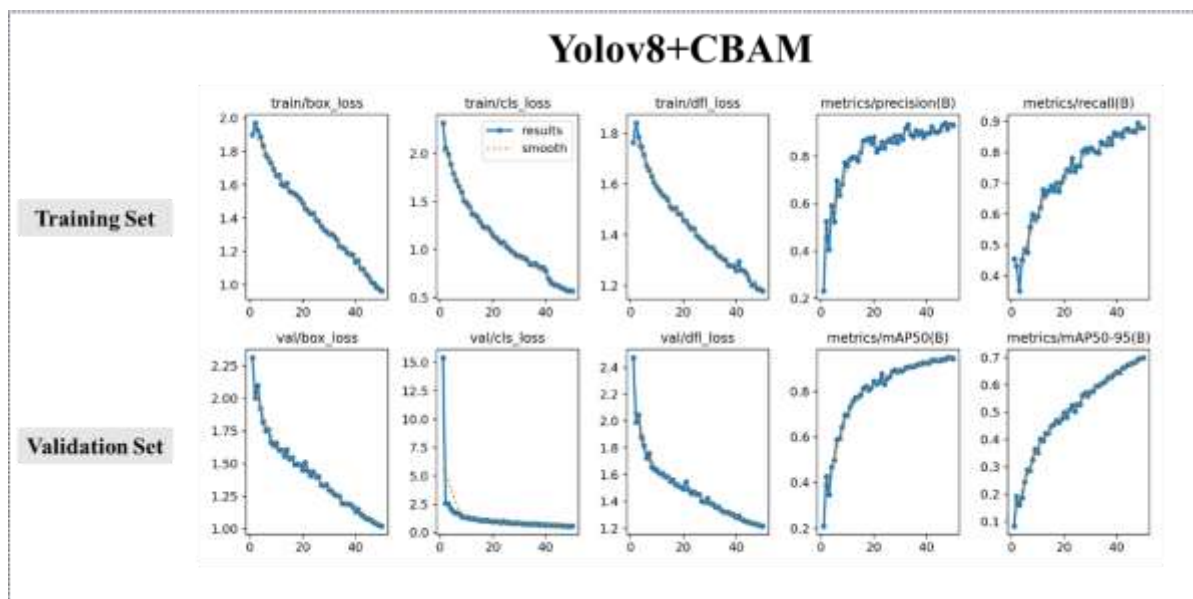


Figure 10. Training and Validation Performance Curves of the YOLOv8+CBAM Model

6.4. Model performance analysis

6.4.1. Summary of Classification Performance

The results show that the model performs uniformly superior across all MP classes, with an average accuracy of 98.3%. The precision, recall, and F1-score of all classes are greater than 0.97, which is assuring regarding the distinction among Beads, Fragments, and Threads. Although the precision of Fragments is the highest, there seems to be a slight improvement in the recall of Beads and Threads, which can recognize round as well as elongated MPs. The closeness of the macro and weighted scores indicates no significant bias among the classes.

Figure 11 (a) presents the unnormalized confusion matrix of the YOLOv8 model. Similarly, Figure 11 (b) presents the normalized confusion matrix of the YOLOv8 model. The above confusion matrices calculate classification accuracy over four classes: fragment, bead, thread, and background.

The diagonal elements in the un-normalized confusion matrix have large values for MPs, which signify robust classification strength. The proposed system is able to correctly classify 409 fragments, 456 beads, and 332 threads. Notably, there is also a considerable amount of confusion between MPs and other classes, and background samples are misclassified into MPs like fragments (97), beads (56), and threads (28), which signifies the difficulty in MPs having subtle features compared to background textures in edible food samples. The MPs are also sometimes erroneously classified into background, specifically fragments (130), which signifies the difficulty in identifying low-contrast MP particles in the background.

The normalized confusion matrix thus helps in analyzing the predicted behavior of each class. It turns out that the model performs really well and yields high normalized accuracy values at about 0.97 for bead and thread classes, indicating very good discrimination on these two very different MP samples. However, on the fragment class, it achieves a relatively low normalized accuracy value of 0.76, which indicates greater confusion with background images because of its irregular boundary. This includes background misclassification that is still significant in that it classifies about 54% of all background samples as fragments. This indicates that the corresponding thickness of food samples and MP fragments also had a similar texture effect.

The confusion matrix analysis shows the accuracy and efficiency of the YOLOv8 algorithm in bead and thread-bead MP classification, while for fragments, accuracy is affected by background interference. This emphasizes the need for the incorporation of superior feature discrimination techniques in MP classification in the edible food environment.

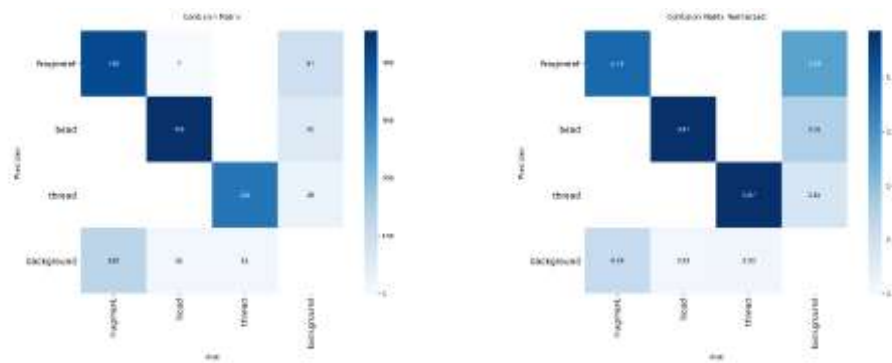


Figure 11. Confusion matrices for YOLOv8: (a) unnormalized confusion matrix and (b) normalized confusion matrix for MP classification.

Figure 12 gives the confusion matrix analysis of the projected YOLOv8 + CBAM model (a) unnormalized and (b) normalized matrices to classify the categories of MP, including Beads, Fragments, Threads, and Background. The unnormalized confusion matrix reveals a high degree of diagonal dominance with 274 Beads, 300 Fragments, and 267 Threads successfully classified as they are MP particles, whereas it is inaccurate between background and MP classes due to the similarities between MP and background observed in optical microscopy images. The normalized confusion matrix also supports the strength of the model with high class-wise accuracies of 78 and 97 percent of both Beads and both Fragments and Threads, respectively, with few classes mixing themselves. The vast majority of mistakes are explained by the fact that the background areas were mistakenly detected as MPs, especially Beads, and can be explained by differences in textures and illumination. In general, the findings indicate that the addition of the CBAM attention mechanism greatly improves feature discrimination in YOLOv8, resulting in high classification rates and low inter-class confusion, confirming the model's suitability for MP detection and classification.

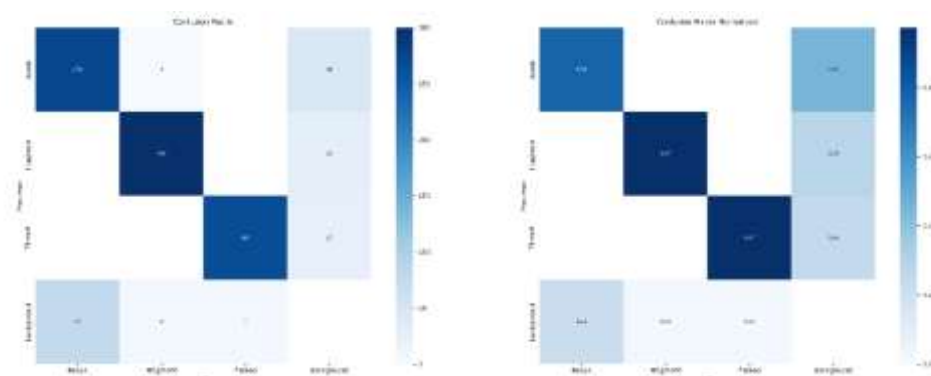


Figure 12. Confusion Matrix YOLOv8+CBAM a) unnormalized b) normalized

The normalized confusion matrix indicates that each of the three categories of MPs, Beads, Fragments, and Threads, is correctly categorized quite well. Fragments and Threads are also well differentiated in the model with an approximate of 0.97 hits. It is once more that beads perform well with 0.78 correct predictions, but a little bewilderment with the background class (0.22). Types of MPs-beads (0.45), fragments (0.29), and threads (0.26) are mostly misclassified as the background class, which means that some textures are shared between the properties of the background and MPs. The model is known to identify the morphologies of MPs with high reliability. The difference between background and particle classes is, however, far less significant.

Qualitative detection results of the proposed YOLOv8 + CBAM model on optical microscopy images of MPs were qualitatively demonstrated in Figure 13. This model can identify and classify various particles, such as beads, threads, and fragments, in different background textures, colors, and illumination, and thus, has the potential to be applied in other computational imaging fields like ultrasound and magnetic resonance imaging. The sub-images exhibit the bounding boxes containing the labels of all the classes forecasted by the model, which is why it is very strong to recognize small, overlapping, and elongated MP objects. Beads are always detected as small and round objects, threads are correctly detected as very thin and long fibers, even when some curves are present or thin ones are weak, and Fragments are correctly localized as irregular forms and the lack of contrast with the background. The findings suggest that the CBAM attention system improves spatial and channel-wise representation of features, which allows the model to pay attention to salient regions of MP and ignore background noise. On the whole, this eye-to-brain analysis

proves the efficiency and the accuracy of the offered method in the case of the actual MP detection and classification under the conditions of complicated microscopy.

YOLOv8-CBAM is capable of delivering high and consistent confidence scores on all four MP classes across different intensities of illumination and scene complexity conditions, and almost no confidence deterioration is observed in difficult configurations (such as trichome overlap, dense particle clustering, low illumination, etc.). Interestingly, when detecting multiple targets (composites) of interest, the proposed model with the greatest confidence scores will perform the most accurate localization of single MP targets. These findings provide evidence that the combination of the CBAM attention mechanism improves feature discrimination and focus on space, resulting in high localization of the object, fewer false positives, and fewer false repetitions. YOLOv8-CBAM is proven to be a powerful and stable method of MP detection.

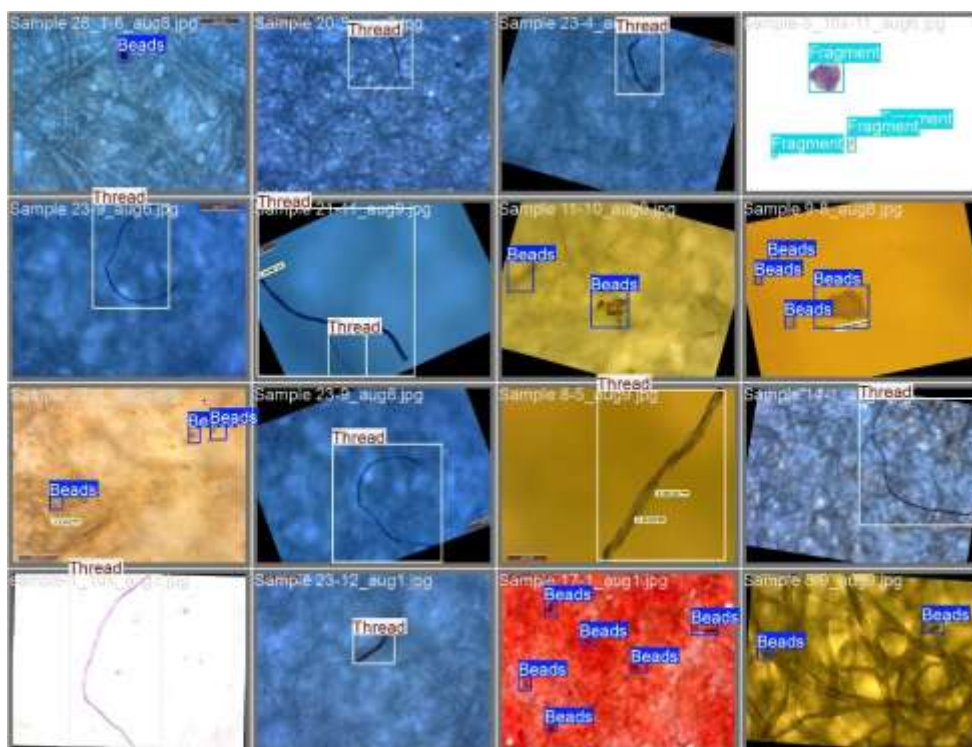


Figure 13. Detection results of YOLOv8+CBAM

6.5. Qualitative Detection Outcomes of YOLOv8-CBAM

Moreover, it is also depicted in Figure 14, the qualitative outcomes of the use of our proposed YOLOv8 with CBAM on some examples of MP use. These findings reveal that the methodology has succeeded in identifying and detecting a number of classes of MPs, including beads, fragments, and threads in microscopic images with variations in background, color, position, and concentration of particles

It is worth mentioning here that long, slender thread-like particles are constantly being detected, even if they happen to be very thin with irregular curves. Beads and fragments are also being detected quite well in busy or textured backgrounds. The robustness of the model is further demonstrated in the detection of clusters of particles in a single image. These qualitative results suggest that the CBAM-enhanced YOLOv8 has successfully focused on essential spatial and channel information, thereby enabling it to reliably and accurately detect MPs, even under extreme imaging conditions.



Figure 14. GUI results for MP detection in edible food image

Using YOLOv8-CBAM, a graphical user interface (GUI) for finding MPs. This demonstration shows was developed situation shows how a web-based GUI using Gradio was created, which served to allow any person to use the recommended YOLOv8-CBAM model to find MPs interactively. Users are allowed to upload microscopy images with visible MPs and use the GUI for real-time detection and sorting. An actual input comes in through the left side, while the output of the model lives on the right. Detected MP particles are shown with boxed markings and class names, for instance, "beads," "fragments," or "threads." This side-by-side view makes it easy to compare the raw inputs with detection results. The GUI also provides quantitative performance indicators - precision, recall, F1-score, and support for each class, which provide instant feedback on the level of success of the detection process.

To conclude, the GUI developed using Gradio can be easily applied in real life because it is easier and faster to use and does not require too much knowledge of technology to analyze MPs in the environment. The ease of access to gadgets makes it a quick tool in laboratories to check if there are MPs present.

6.6. Comparative example performance discussion

For the comparative performance analysis, YOLOv8-CBAM (Proposed), YOLOv8, and YOLOv5 were implemented and evaluated on the proposed dataset.

Table 4 showed that the proposed YOLOv8-CBAM model exhibits stable, high detection rates across all MP types in edible food samples, achieving an overall mAP of 98.30. The deployment of the CBAM enables the model to prioritize fine MP features and reduce its focus on complex food textures, i.e., fibrous matrices, surface irregularities, and background noise, typical of edible materials. This attention-directed feature refinement can greatly improve fine-grained MP structures such as fibers, fragments, thin films, and plastic sequins. In edible food matrices, repetitive textures and heterogeneous backgrounds often cause misalignment of candidate boxes and suppress recall for small MP targets. Altogether, the findings prove that the implementation of CBAM in YOLOv8 significantly enhances the level of feature discrimination and robustness, and the suggested model would be the most appropriate to use when it comes to detecting MP in the edible food samples with great precision and reliability.

Table 4 YOLOv8+CBAM performance comparison

Model	AP (Threads)	AP (Fragments)	AP (Plastic Beads)	mAP
YOLOv8-CBAM	98.42	98.95	99.21	98.81
YOLOv8	95.28	95.35	95.22	95.13
YOLOv5	95.46	95.21	95.34	95.16

As shown in Table 5 determine that YOLOv8-CBAM achieves the best performance across all four evaluation metrics—Precision, Recall, mAP, and F1-score, with values of 98.32 %, 98.30 %, 98.30 %, and 98.31 %, respectively—significantly outperforming all eleven comparative models. The proposed YOLOv8-CBAM model achieves a large improvement in performance compared to the baseline YOLOv8, based on all the external validation measures. Namely, YOLOv8-CBAM shows improvements in the form of +6.83% in precision, +5.95% in recall, +3.26% in mAP, and +6.39% in F1-score, indicating its higher performance in identifying MPs in complex food matrices, which is edible. These advancements are mainly explained by the

incorporation of the CBAM, which could provide great channel-wise and spatial features refinement that enabled the model to concentrate on minute MP features and reduce background food textures. YOLOv8-CBAM provides a more reasonable trade-off between the detection performance, architectural simplicity, and robustness, and is hence very suitable for practical, effective MP detection of edible food items in the real world. When compared with the baseline YOLOv8, the suggested YOLOv8-CBAM model has improved by +6.83% in precision, +5.95% in recall, +3.26% in mAP, and +6.39% in F1-score, which validates the relevance of CBAM in improving the performance of MP detection in edible food settings.

Table 5 External validation performance and Δ -improvement comparison of YOLOv8-CBAM against YOLOv8

Model	Precision (%)	Recall (%)	mAP (%)	F1 (%)	Δ Precision	Δ Recall	Δ mAP	Δ F1
YOLOv8-CBAM (Proposed)	98.32	98.30	98.30	98.31	+6.83	+5.95	+3.26	+6.39
YOLOv8	91.49	92.35	95.04	91.92	0.00	0.00	0.00	0.00
YOLOv5	93.48	90.68	95.28	92.06	+1.99	-1.67	+0.24	+0.14

The proposed YOLOv8 + CBAM framework demonstrated strong and consistent performance across all five folds during cross-validation as shown in **Error! Reference source not found.** The model achieved a mean accuracy of 98.40%, mean precision of 98.35%, mean recall of 98.42%, and mean macro F1-score of 98.38%, with very low standard deviations (± 0.003), indicating high robustness and stable generalization across different data partitions. Class-wise evaluation further confirmed the effectiveness of the proposed model. The Thread class achieved the highest F1-score of 98.78%, followed by Fragment (98.45%) and Beads (97.90%). The high recall value of 99.36% for the Thread category indicates excellent detection capability with minimal false negatives. Similarly, the Fragment class achieved the highest precision, 99.57%, indicating accurate classification with very few false positives. The low standard deviation in class-wise F1-scores confirms consistent performance across all folds. Statistical significance testing was conducted to validate the reliability of the obtained results as shown in Table 7. The one-sample t-test showed highly significant improvements over the baseline performance value (F1 = 0.80 and Accuracy = 0.80), with p-values less than 0.05. The proposed model achieved t-statistics of 123.33 for F1-score and 129.48 for accuracy, confirming statistically significant performance improvements. Per-class t-tests also demonstrated significant improvements for Beads, Fragment, and Thread categories as shown in Table 8.

Table 6 Average Classification Performance Across 5 Folds

Class	Precision	Recall	F1-Score	Std. Dev (F1)
Beads	0.9725	0.9856	0.9790	0.0050
Fragment	0.9957	0.9735	0.9845	0.0044
Thread	0.9822	0.9936	0.9878	0.0043
Macro Avg	0.9835	0.9842	0.9838	0.0030
Weighted Avg	0.9842	0.9840	0.9840	0.0028

Table 7 Statistical Significance Analysis of YOLOv8 + CBAM

Statistical Test	Metric	Value
One-Sample t-test	Mean F1-Score	0.9838 $\pm$ 0.0030
	t-statistic	123.3375
	p-value	< 0.000001
One-Sample t-test	Mean Accuracy	0.9840 $\pm$ 0.0028
	t-statistic	129.4854
	p-value	< 0.000001

Table 8 Per-Class Statistical Analysis (F1-Score)

Class	Mean F1 $\pm$ Std. Dev	t-statistic	p-value
Beads	0.9790 $\pm$ 0.0050	72.0846	< 0.000001
Fragment	0.9845 $\pm$ 0.0044	84.4929	< 0.000001
Thread	0.9878 $\pm$ 0.0043	88.1682	< 0.000001

The integration of DL with optical microscopy extends the measurement range of optical microscopy to the microscopic world. The ability of traditional microscopy-based MP analysis. In optical microscopy, they can be directly visualized from the extracted results. Subsequent identification and counting can rely on, however, particles; Many on manual inspection and can be time-consuming if a large amount of images have to be analyzed. The proposed YOLOv8-CBAM This computational bottleneck is dealt with by the framework

itself automatically: Identification and localization of MP particles in microscopy images. The use Given the above, CBAM further enhances the feature representation by highlighting The informative spatial and channel level features are especially, Appropriate for low- and small-contrast particles. Thus, the contribution One of the proposed framework is not restricted to classification accuracy However, the challenge remains of converting observations made using optical microscopy into Automated, reproducible particle-level measurements. This integration can enhance the efficiency, consistency, and scalability of MP assessment using microscopy in food samples.

7. Conclusion

This paper proposes a DL architecture to detect MP in images of edible food, based on an improved YOLOv8 variant and CBAM. YOLOv8-CBAM is more accurate and robust for detection than YOLOv8 because it uses channel and spatial attention, especially for detecting small and visually complex MP particles. The experimental findings demonstrate that YOLOv8-CBAM consistently outperforms the baseline YOLOv8 in terms of precision, recall, and mAP. The enhancement is also noticeable in small-sized and irregularly shaped MPs, where attention mechanisms control fine-grained details. Higher values on higher mAP50.5 and mAP50.5:0.95 indicate greater localization and strength across varying IoU thresholds. Both models are shown to converge stably, but YOLOv8-CBAM has more continuous validation loss curves and less performance variability. This means that it generalizes better and overfits less, particularly with class-imbalanced or visually noisy datasets. CBAM integration leads to a small increment in the computational complexity and inference time. Despite this, the performance gain achieved in detection accuracy outweighs the additional overhead, making YOLOv8-CBAM suitable for applications where accuracy is prioritized over minimal latency. The outcome of this work suggests that MP in edible food samples can be analyzed using attention-guided object detection, which offers a promising direction for addressing complex MP contamination in edible food samples. With the YOLOv8+CBAM model, a scalable and computationally efficient solution was given, which can be employed in applications pertaining to automated food safety monitoring and environmental assessments.

Author Contributions

Neha Harde: Conceptualization; Analysis; Methodology; Data Curation; Investigation; Writing - Original Draft; Writing - Review & Editing; Resources.

Tulasi B: Conceptualization; Methodology; Data Curation; Investigation; Writing - Original Draft; Writing - Review & Editing; Visualization; Supervision; Resources.

Funding

The authors declare that no external funding was received for the conduct of this research or the preparation of this article.

Data Availability

The dataset used in this study is available upon reasonable request from the corresponding author.

Declarations

Ethics approval

This study did not involve human participants or animals. The research was based on MP samples and image datasets; hence, ethical approval was not required.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Consent to Participate declaration

Not applicable

Clinical Trial Registration

Not Applicable, as this study does not involve a clinical trial

Declaration on the Use of Generative AI and AI-assisted Technologies

The authors declare that no generative AI or AI-assisted technologies were used in the writing, analysis, or preparation of this manuscript.

References

- [1] Y. Gao et al., "A comprehensive review of microplastic aging: Laboratory simulations, physicochemical properties, adsorption mechanisms, and environmental impacts," *Sci. Total Environ.*, vol. 957, p. 177427, Dec. 2024, doi: 10.1016/j.scitotenv.2024.177427.
- [2] S. Manzo and S. Schiavo, "Physical and chemical threats posed by micro(nano)plastic to sea urchins," *Sci. Total Environ.*, vol. 808, p. 152105, Feb. 2022, doi: 10.1016/j.scitotenv.2021.152105.
- [3] S. Acarer Arat, "An overview of microplastic in marine waters: Sources, abundance, characteristics and negative effects on various marine organisms," *Desalination Water Treat.*, vol. 317, p. 100138, Jan. 2024, doi: 10.1016/j.dwt.2024.100138.
- [4] R. Ullah, M. T.-K. Tsui, A. Chow, H. Chen, C. Williams, and A. Ligaba-Osena, "Micro(nano)plastic pollution in terrestrial ecosystem: emphasis on impacts of polystyrene on soil biota, plants, animals, and humans," *Environ. Monit. Assess.*, vol. 195, no. 1, p. 252, Dec. 2022, doi: 10.1007/s10661-022-10769-3.
- [5] Q. Kang, K. Zhang, S. C. Dekker, and J. Mao, "Microplastics in soils: A comprehensive review," *Sci. Total Environ.*, vol. 960, p. 178298, Jan. 2025, doi: 10.1016/j.scitotenv.2024.178298.
- [6] J. Gasperi et al., "Microplastics in air: Are we breathing it in?," *Micro Nanoplastics* Ed. Dr Teresa AP Rocha-St., vol. 1, pp. 1–5, Feb. 2018, doi: 10.1016/j.coesh.2017.10.002.
- [7] F. Carrieria, C. Di Fiore, and P. Avino, "Occurrence of Microplastics in the Atmosphere: An Overview on Sources, Analytical Challenges, and Human Health Effects," *Atmosphere*, vol. 15, no. 7, p. 863, 2024, doi: 10.3390/atmos15070863.
- [8] L. Yang, S. Kang, X. Luo, and Z. Wang, "Microplastics in drinking water: A review on methods, occurrence, sources, and potential risks assessment," *Environ. Pollut.*, vol. 348, p. 123857, May 2024, doi: 10.1016/j.envpol.2024.123857.
- [9] E. B. Jadhav, M. S. Sankhla, R. A. Bhat, and D. S. Bhagat, "Microplastics from food packaging: An overview of human consumption, health threats, and alternative solutions," *Environ. Nanotechnol. Monit. Manag.*, vol. 16, p. 100608, Dec. 2021, doi: 10.1016/j.enmm.2021.100608.
- [10] M. A. Mir et al., "Microplastics in food products: Prevalence, artificial intelligence based detection, and potential health impacts on humans," *Emerg. Contam.*, vol. 11, no. 2, p. 100477, Jun. 2025, doi: 10.1016/j.emcon.2025.100477.
- [11] S. Gündoğdu, C. Çevik, Y. Terzi, K. Gedik, F. Büyükdeveci, and R. Ç. Öztürk, "Microplastics in Turkish coastal lagoons: Unveiling the hidden threat to wetland ecosystems," *Environ. Pollut.*, vol. 375, p. 126351, Jun. 2025, doi: 10.1016/j.envpol.2025.126351.
- [12] Md. I. Muhib, Md. K. Uddin, Md. M. Rahman, and G. Malafaia, "Occurrence of microplastics in tap and bottled water, and food packaging: A narrative review on current knowledge," *Sci. Total Environ.*, vol. 865, p. 161274, Mar. 2023, doi: 10.1016/j.scitotenv.2022.161274.
- [13] P. Balakrishnan et al., "Artificial intelligence for food safety: From predictive models to real-world safeguards," *Trends Food Sci. Technol.*, vol. 163, p. 105153, Sep. 2025, doi: 10.1016/j.tifs.2025.105153.
- [14] H. Yang, W. Jiao, L. Zouyi, H. Diao, and S. Xia, "Artificial intelligence in the food industry: innovations and applications," *Discov. Artif. Intell.*, vol. 5, no. 1, p. 60, May 2025, doi: 10.1007/s44163-025-00296-8.
- [15] A. Miteloudis, "AI for Microplastics Detection in Food Systems: A Conceptual Framework," *Zenodo*, Sep. 2025. doi: 10.5281/zenodo.17169647.
- [16] H. Rawat et al., "Artificial intelligence-driven detection of microplastics in food: A comprehensive review of sources, health risks, detection techniques, and emerging artificial intelligence solutions," *Food Chem. X*, vol. 29, p. 102687, Jul. 2025, doi: 10.1016/j.fochx.2025.102687.
- [17] J.-H. Kwon et al., "Microplastics in Food: A Review on Analytical Methods and Challenges," *Int. J. Environ. Res. Public Health*, vol. 17, no. 18, p. 6710, Sep. 2020, doi: 10.3390/ijerph17186710.
- [18] N. Harde and T. B., "Deep learning based classification of microplastic in edible food using optical microscopy images," *Discov. Artif. Intell.*, vol. 6, no. 1, p. 478, Apr. 2026, doi: 10.1007/s44163-026-01246-8.
- [19] K. Rermborirak, P. Nanuan, P. Komonpan, and S. Sukpancharoen, "Low-cost portable microplastic detection system integrating Nile red fluorescence staining with YOLOv8-based deep learning," *J. Hazard. Mater. Adv.*, vol. 19, p. 100787, Aug. 2025, doi: 10.1016/j.hazadv.2025.100787.
- [20] H. Long, "Hybrid Design of CNN and Vision Transformer: A Review," in *Proceedings of the 2024 7th International Conference on Computer Information Science and Artificial Intelligence*, in CISA '24. New York, NY, USA: Association for Computing Machinery, 2024, pp. 121–127. doi:

- 10.1145/3703187.3703208.
- [21] Y. Li et al., "A Local-Global Fusion Vision Mamba UNet Framework for medical image segmentation," *Eng. Appl. Artif. Intell.*, vol. 169, p. 113987, Apr. 2026, doi: 10.1016/j.engappai.2026.113987.
- [22] B. Singh, R. Das, A. Kumar, S. Goyal, and P. K. Sharma, "Integrating Artificial Intelligence with Analytical Techniques for Enhanced Microplastics Analysis," *Crit. Rev. Anal. Chem.*, pp. 1–22, Mar. 2026, doi: 10.1080/10408347.2026.2647106.
- [23] P. Akhil et al., "AI-driven portable smartphone-based sensor for microplastic detection: a comprehensive review," *Toxicol. Environ. Health Sci.*, vol. 18, no. 1, pp. 103–124, Mar. 2026, doi: 10.1007/s13530-025-00294-4.
- [24] M. He et al., "Research on hybrid microplastic recognition method based on dual-branch convolutional neural network combined with attention mechanism," *Microchem. J.*, vol. 218, p. 115131, Nov. 2025, doi: 10.1016/j.microc.2025.115131.
- [25] S. Ravikumar, A. Jeyameenakshi, M. Syed Ali, and K. S. Ebenezer, "Assessment of microplastics in edible salts from solar salt pans and commercial salts," *Total Environ. Res. Themes*, vol. 6, p. 100032, Jun. 2023, doi: 10.1016/j.totert.2023.100032.
- [26] A. Vidyasakar et al., "Microplastic contamination in edible sea salt from the largest salt-producing states of India," *Mar. Pollut. Bull.*, vol. 171, p. 112728, Oct. 2021, doi: 10.1016/j.marpolbul.2021.112728.
- [27] P. Akkajit, S. Thongnonghin, S. Sriraksa, and S. Pumsri, "Preliminary Study of Distribution and Quantity of Plastic-debris on Beaches Along the Coast at Phuket Province," *Appl. Environ. Res.*, pp. 54–62, Jul. 2019, doi: 10.35762/AER.2019.41.2.5.
- [28] R. Kiruba et al., "Identification of Microplastics as Emerging Contaminant in Branded Milk of Tamil Nadu State, India," *Asian J. Biol. Life Sci.*, vol. 11, no. 1, pp. 181–187, May 2022, doi: 10.5530/ajbls.2022.11.25.
- [29] K. Kadam-Czapska et al., "Isolation and identification of microplastics in infant formulas – A potential health risk for children," *Food Chem.*, vol. 440, p. 138246, May 2024, doi: 10.1016/j.foodchem.2023.138246.
- [30] Q. Zhang et al., "Microplastics in infant milk powder," *Environ. Pollut. Barking Essex 1987*, vol. 323, p. 121225, Apr. 2023, doi: 10.1016/j.envpol.2023.121225.