

Development of an Image-Based Identification System for Freshwater Diatoms Using Computer Vision Techniques

John O-Neil Geronimo^{1,*}, Eldrin DLR. Arguelles^{2*}, Concepcion Khan¹, Maria Art Antonette Clariño¹, Val Randolph Madrid¹

¹Institute of Computer Science, University of the Philippines Los Baños, 4031, Laguna, Philippines; ²Philippine National Collection of Microorganisms, National Institute of Molecular Biology and Biotechnology (BIOTECH), University of the Philippines Los Baños, 4031, Laguna, Philippines

Received: October 17, 2025; Revised: December 21, 2025; Accepted: January 16, 2026

Abstract

Diatoms are widely used indicators of environmental condition because of their rapid and sensitive responses to changes in water quality. However, routine identification and enumeration using morpho-taxonomic and molecular approaches remain labor-intensive and time-consuming. This study evaluates the performance of a YOLO-based object detection framework for automated detection, localization, and genus-level classification of freshwater diatoms from microscopic images. Using annotated image datasets, model performance was assessed at the cell level, with detection and classification accuracy quantified using mean average precision (mAP). YOLOv4 achieved a mAP of 56.64% for diatom detection and localization. In multi-class classification experiments involving 16 genera, YOLOv4 outperformed YOLOv7, yielding a higher mAP of 19.21%. Performance analysis revealed that class imbalance and limited cell counts negatively affected detection accuracy in several genera. Restricting the dataset to eight genera with higher cell representation improved performance, with YOLOv4 trained on raw images reaching a mAP of 41.98%. In terms of efficiency, the automated system processed and quantified diatom cells from approximately 100 images within minutes, compared with several days required for manual analysis of a smaller image set. These results demonstrate that YOLO-based models can support rapid, cell-level diatom analysis, with performance strongly influenced by dataset composition, and highlight their potential as complementary tools for ecological monitoring workflows.

Keywords: Digital image processing, Freshwater diatoms, Philippines, YOLO (You Only Look Once)

1. Introduction

Diatoms, unicellular phototrophs with approximately 100,000 extant species worldwide, play an important role in biomass production and have various applications in industrial, cell biological, and molecular fields (Bozarth *et al.*, 2009; Wang and Seibert, 2017). With a reproduction cycle of around two weeks, diatoms have one of the shortest generation times among all species. Due to their rapid responsiveness and high sensitivity to various factors including changes in nutrient concentration, organic pollution, and aquatic productivity, diatoms help researchers monitor early warning signs of increasing pollution and track the success of habitat restoration efforts within ecosystems (Rimet, 2012; Masouras *et al.*, 2021; Arguelles, 2022). Their robustness and potential as water quality indicators have even prompted the European Water Framework Directive to advocate for and require the use of bioindicators, including benthic diatoms, for assessing stream ecosystems (Masouras *et al.*, 2021). The crucial role of diatoms as bioindicators underscores the need for effective methods to identify and categorize the diverse species present in aquatic environments.

In the Philippines, numerous studies have documented the diversity, ecological significance, and bioindicator value of diatoms across marine, freshwater, and extreme environments. Investigations have reported diverse assemblages from the West Philippine Basin (Shen *et al.*, 2017), toxin-producing *Pseudo-nitzschia* species in San Pedro Bay (Yap-Dejeto *et al.*, 2013), epiphytic communities associated with seagrasses (Yap-Dejeto and Baqueros, 2015), thermophilic taxa from hot springs (Martinez-Goss *et al.*, 2019), and pollution-tolerant species in urban river systems such as the Marikina River (Martinez-Goss *et al.*, 2023). Additional studies have highlighted diatom epiphytes in Laguna de Bay and other freshwater systems (Arguelles *et al.*, 2019). While these works underscore the ecological relevance of diatoms for environmental assessment, they also highlight the labor-intensive and time-consuming nature of traditional morpho-taxonomic identification, which requires extensive expertise and manual microscopy (Arguelles and Monsalud, 2017; Arguelles *et al.*, 2018; Arguelles, 2021; Arguelles, 2023).

To address these limitations, automated image-based approaches using digital image processing and machine learning have gained increasing attention. Convolutional neural networks (CNNs) and object detection frameworks,

* Corresponding author. e-mail: jvgeronimo@up.edu.ph; edarguelles@up.edu.ph

including YOLO-based architectures, have been successfully applied to diatom, phytoplankton, and microalgal detection and classification in international studies, demonstrating varying levels of accuracy depending on dataset size, taxonomic complexity, and image quality (Dimitrovski *et al.*, 2012; Wang *et al.*, 2021; Liu *et al.*, 2022; Syed *et al.*, 2024). These studies have established YOLO and related CNN-based methods as effective tools for microscopic organism recognition, particularly for high-throughput detection and counting tasks. However, reported performance metrics such as mean average precision, precision, and recall vary widely across studies, reflecting differences in training data composition, class balance, and annotation strategies (Kloster *et al.*, 2014; Pedraza *et al.*, 2017; Libreros *et al.*, 2019; Siddik and Bostan, 2019; Arguelles, 2020; Geronimo *et al.*, 2023).

In this context, the present study evaluates the applicability of YOLO-based object detection models (YOLOv4 and YOLOv7) for automated detection, counting, and genus-level classification of freshwater diatoms using locally generated microscopic image datasets. Rather than introducing YOLO as a novel methodology, this work focuses on assessing model performance under conditions relevant to Philippine laboratories, including dataset size, class imbalance, and taxonomic complexity. Specifically, the study examines detection and classification performance across 16 diatom genera and explores how reducing class complexity influences mean average precision at the cell level. By comparing model performance and processing efficiency with conventional manual approaches, this study aims to clarify the strengths and limitations of YOLO-based systems as complementary tools for routine diatom-based environmental monitoring.

2. Materials and Methods

2.1. Sample Collection, Preservation, and Slide Preparation

Benthic diatom samples (10–15 mL) were collected from freshwater sampling sites using sterile Tarson tubes (Arguelles, 2019). Samples were stored in insulated coolers during transport and processed within 24 h of collection. For microscopic examination, 3 mL subsamples were preserved in 4% formaldehyde to maintain cellular integrity. Preserved samples were gently homogenized prior to slide preparation to ensure even distribution of diatom cells. Temporary wet mounts were prepared for light microscopy, following standard diatom examination protocols, without chemical digestion to preserve natural assemblage structure.

2.2. Image acquisition

Microscopic images were captured using an Olympus CX31 binocular light microscope equipped with an Infinity X digital camera at 1000× magnification (oil immersion). Images contained heterogeneous diatom assemblages exhibiting variation in size, morphology, and abundance across taxa. A total of 237 images were collected across three independent sampling batches to reflect realistic variability in environmental samples.

The first batch (108 images) was used exclusively for model training. The second and third batches, comprising

84 and 45 images respectively (129 images total), were reserved for model testing. Genus-level representation varied across batches, with the first batch dominated by *Aulacoseira*, *Encyonema*, *Cymbella*, and *Fragilaria*, while *Achnanthisidium* was more prevalent in the latter batches. Several genera (e.g., *Cocconeis*, *Cyclotella*, *Rhoicosphenia*, *Surirella*, *Achnanthes*, *Planothidium*) were sparsely represented in both training and testing datasets.

2.3. Dataset Annotation and Composition

All images were manually annotated using LabelImg software. Annotation was performed by one trained diatom taxonomist with extensive experience in freshwater diatom identification. Each visible diatom cell was enclosed in a bounding box and labeled at the genus level. Annotation was performed at the cell level, rather than per image, and exported in YOLO-compatible format.

Overlapping or clustered cells were annotated as separate objects where cell boundaries were visually distinguishable. Ambiguous cells (those with unclear morphological features due to orientation, overlap, or partial visibility) were excluded from annotation to minimize labeling uncertainty. Given that a single expert performed annotation, formal inter-rater validation was not conducted; however, annotations were reviewed iteratively to ensure internal consistency.

Table 1 summarizes the number and proportion of annotated diatom cells per genus in the training and testing datasets. Class imbalance was evident, with several genera represented by fewer than 10 annotated cells. This imbalance was retained to reflect real-world environmental sample composition rather than artificially balancing the dataset.

Table 1. Number of cells found per diatom genus in the training and testing datasets.

Class	Training (no. of cells)	Training (%)	Testing (no. of cells)	Testing (%)
<i>Aulacoseira</i>	198	20.62%	21	2.88%
<i>Fragilaria</i>	189	19.68%	35	4.80%
<i>Encyonema</i>	115	11.97%	149	20.46%
<i>Cymbella</i>	201	20.93%	82	11.26%
<i>Cocconeis</i>	5	0.52%	16	2.19%
<i>Synedra</i>	79	8.22%	43	5.90%
<i>Gomphonema</i>	13	1.35%	49	6.73%
<i>Achnanthes</i>	7	0.72%	0	0.00%
<i>Nitzschia</i>	45	4.68%	87	11.95%
<i>Planothidium</i>	1	0.10%	1	0.13%
<i>Sellaphora</i>	15	1.56%	0	0.00%
<i>Surirella</i>	6	0.62%	0	0.00%
<i>Navicula</i>	73	7.60%	33	4.53%
<i>Cyclotella</i>	10	1.04%	9	1.23%
<i>Rhoicosphenia</i>	1	0.10%	0	0.00%
<i>Achnanthisidium</i>	2	0.20%	203	27.88%

2.4. Image Pre-processing

Image pre-processing was applied as a controlled experimental factor rather than a default step. Two pre-processing techniques were evaluated: bilateral filtering and contrast-limited adaptive histogram equalization (CLAHE). Bilateral filtering was used to reduce background noise and small debris while preserving cell

edges, which is particularly important for microscopy images derived from preserved environmental samples. CLAHE was applied to enhance local contrast and improve edge definition under variable illumination conditions.

Pre-processing was applied consistently to both training and testing datasets within the designated experimental setups. Raw (unprocessed) images were also retained to evaluate whether pre-processing significantly influenced detection and classification performance.

2.5. Experimental setup

The experimental workflow was structured in two sequential stages to systematically evaluate the feasibility and performance of YOLO-based object detection for freshwater diatom analysis (Table 2). In the first stage, YOLOv4 was applied exclusively to raw microscopic images to assess its capability to detect and localize individual diatom cells under minimally processed conditions. This preliminary stage served to confirm whether the YOLO object detection framework could reliably identify microscopic diatom structures within complex environmental images containing debris and heterogeneous cell assemblages.

In the second stage, classification performance was evaluated using four experimental setups that combined two YOLO architectures (YOLOv4 and YOLOv7) with two image conditions (raw and pre-processed). This design enabled direct comparison of (i) model architecture effects and (ii) the influence of image pre-processing on genus-level classification accuracy. All experimental setups used the same training and testing datasets to ensure comparability of results.

Table 2. List of experimental setups for the classification of 16 diatom genera.

Setup no.	Pre-processing	Classifier
1	Bilateral Filtering and Contrast-Limited Adaptive Histogram Equalization (CLAHE)	YOLOv4
2	None	YOLOv4
3	Bilateral Filtering and Contrast-Limited Adaptive Histogram Equalization (CLAHE)	YOLOv7
4	None	YOLOv7

For the initial detection experiment, a binary classification task was performed to distinguish diatom cells from non-diatom objects in microscopic images. A total of 237 expert-annotated images were used to train a YOLOv4 model using raw images without additional pre-processing. Model evaluation was conducted using an independent test set consisting of 27 images. Performance was assessed using an Intersection over Union (IoU) threshold of 0.50, following standard object detection evaluation protocols.

2.6. Model Training Configuration

A total of four object detection models were trained: YOLOv4 using raw images, YOLOv4 using pre-processed images, YOLOv7 using raw images, and YOLOv7 using pre-processed images. Model training was conducted on a workstation running Windows 11, equipped with an Intel

Core i7-8700 processor and an NVIDIA GeForce GTX 1070 graphics processing unit.

For all models, the learning rate was fixed at 0.001, and the maximum number of training iterations was set to 32,000 to ensure sufficient convergence. Input image dimensions and other hyperparameters followed the default settings recommended by the official YOLOv4 and YOLOv7 implementations. No additional data augmentation techniques, such as rotation or flipping, were applied beyond the defined pre-processing steps, in order to isolate the effects of image enhancement on model performance.

Training time differed markedly between architectures. YOLOv4 models required approximately 12 h to complete training, whereas YOLOv7 models completed training in approximately 30 min, reflecting architectural optimizations that reduce computational complexity.

2.7. Cell Detection, Classification, and Counting

Model inference was performed exclusively on the independent testing dataset. During inference, each trained model generated bounding boxes and corresponding genus-level predictions for individual diatom cells present in each image. Detection outputs were filtered based on confidence thresholds defined by the YOLO framework to reduce false-positive predictions.

Detected diatom cells were counted on a cell-by-cell basis, and the number of cells per genus was recorded for each image. These cell counts formed the basis for both quantitative comparison across experimental setups and subsequent performance evaluation. This approach allowed assessment of the models' ability to simultaneously detect, classify, and enumerate diatom cells in complex microscopic images.

2.8. Performance Evaluation

Model performance was evaluated at the cell level, rather than at the image level, to accurately reflect the object detection nature of the task. Mean average precision (mAP) was used as the primary performance metric and was computed by comparing predicted bounding boxes and genus labels against expert-annotated data. A prediction was considered correct when both the localization and classification criteria were satisfied.

Diatom curated annotations generated by the taxonomist served as the reference standard for evaluation. Performance metrics therefore reflect the system's ability to correctly detect, localize, and classify individual diatom cells across multiple genera, providing a robust measure of applicability for automated diatom-based environmental monitoring.

3. Results and Discussion

The study investigated the use of YOLO architecture for detecting, identifying, and counting local freshwater diatoms in the Philippines. The system successfully recognized and categorized 16 genera of freshwater diatoms exhibiting variations in color, size, and shape but with room for improvement. This technological advancement holds potential significance for diatom taxonomists, demonstrating the feasibility of a computational recognition approach for automated identification of selected freshwater diatoms.

The YOLOv4 model achieved a mean average precision (mAP) of 56.64% at an IoU threshold of 0.50 for the binary diatom detection task. The model obtained a precision of 71%, recall of 48%, and an F1-score of 57%. The relatively high precision indicates that the model was effective at correctly identifying diatom cells when detections were made. However, the lower recall suggests that a substantial proportion of expert-labeled diatom cells were not detected, which negatively influenced the overall mAP value. This result highlights the challenge of detecting small and morphologically diverse diatom cells in complex microscopic images using raw data alone.

The system automated the detection, classification, and counting of diatom cells. The second experimental setup specifically assessed its performance in classifying and enumerating cells from 16 freshwater diatom genera. Out of the 237 images prepared for the second setup, a total of 129 images were reserved for the testing phase and were used to evaluate the performance of the system. The remaining 108 images were already used to train both the YOLOv4 and the YOLOv7 models in both the raw and pre-processed datasets.

Table 3. Mean average precision (mAP), precision, and recall of each experimental setup.

Classifier	mAP	Precision	Recall
YOLOv4 on pre-processed dataset	18.32	40%	25%
YOLOv4 on raw dataset	19.21	41%	25%
YOLOv7 on pre-processed dataset	13.3%	54.2%	28.3%
YOLOv7 on raw dataset	2.88%	2.63%	3.31%

The diatom detection results were compared with the prepared labels of the test dataset to evaluate the accuracy of the system in detecting and classifying diatoms in an image. An IoU threshold of 0.50 was set for the confidence scores generated by the trained YOLO models. This means that the detections made by the models need to match at least 50% of the prepared labels in the test dataset. It was observed that YOLOv4, trained on a pre-processed image dataset, successfully detected diatoms with a mean average precision of 18.32%. Additionally, it demonstrated a precision of 40%, recall of 25%, and an F1-score of 31%. Conversely, YOLOv4, trained on a raw image dataset, achieved a mean average precision of 19.21%. Furthermore, it generated a precision of 41%, recall of 25%, and an F1-score of 31%. The YOLOv7 trained on a pre-processed dataset, exhibited a mean average precision of 13.3%, precision of 54.2%, and recall of 28.3%. Meanwhile, the YOLOv7 model trained on the raw dataset yielded a mean average precision of 2.88%, precision of 2.63%, and recall of 3.31%.

Interestingly, YOLOv7, despite being a more recent architecture, underperformed compared to YOLOv4 on this dataset. Several factors may have contributed to this outcome. First, YOLOv7 is sensitive to small datasets and class imbalance, and the present dataset contained several underrepresented genera (≤ 15 cells per genus), which likely hampered the model's ability to learn reliable features. Second, training instability due to the complexity of the YOLOv7 architecture may have further reduced performance, particularly for minority classes. Third, the morphological similarity among certain diatom genera (e.g., *Cymbella* vs. *Encyonema*) adds to the challenge of

generalizing from limited samples. These observations suggest that YOLOv7 may require larger and more balanced datasets to outperform YOLOv4 in diatom classification tasks.

Table 3 shows the mean average precision, precision, and recall scores for each model. The YOLOv4 model, without additional preprocessing, achieved the highest mean average precision score of 19.21% among all models. This indicates that the YOLOv4 model, trained on the raw dataset, can detect more cells and therefore count more cells compared to other setups. However, it still fails to detect some existing cells in the image, as indicated by its low recall score. On the other hand, the YOLOv7 model, trained on processed images, detects fewer cells, as shown by its 13.3% mAP, but excels in distinguishing among cells, as demonstrated by its 54.2% precision score.

To contextualize the performance of the proposed system, it is important to compare it with prior international work on diatom detection and CNN-based image analysis. Previous studies applying deep learning to diatom classification have reported mAP values ranging from 60% to 85% for well-curated datasets with larger sample sizes per class (Valova *et al.*, 2020). Segmentation-based approaches have similarly achieved high accuracy in detecting diatoms but often require extensive manual annotation and are sensitive to overlapping cells. Compared to these studies, the present system shows moderate performance (maximum mAP of 19.21% for YOLOv4 on raw images), which reflects the challenges of working with naturally distributed, morphologically diverse diatoms under limited and imbalanced dataset conditions. This comparison highlights the contribution of the current study in demonstrating the feasibility of automated recognition for local Philippine freshwater diatoms and underscores areas for improvement, particularly dataset balancing and increasing cell representation per genus.

Looking closer at the performance of the YOLOv7 model on the raw dataset through the confusion matrix, *Aulacoseira* produced the highest accuracy of 70%. This is followed by *Fragilaria*, *Encyonema*, and *Cymbella* at 5%, 1%, and 10%, respectively. This shows how the model can only detect and classify these particular genera while failing on others. On the other hand, the YOLOv7 model on the preprocessed dataset achieved better results than the raw dataset. *Aulacoseira* still boasts the highest accuracy of 82%, followed by *Fragilaria* and *Encyonema* at 7%, *Synedra* at 3%, and *Cymbella* at 28%. The PR-curve, F1-curve, P-curve, and R-curve from both YOLOv7 models also showed better scores for *Aulacoseira*, *Cymbella*, and *Encyonema*, both of which represented at least 20% of the total number of cells for training.

The YOLOv4 models were also able to get significantly higher accuracy precision scores for classes *Aulacoseira*, *Encyonema*, and *Cymbella*. In YOLOv4 trained on the raw image dataset, the model achieved 75.90% accuracy precision score for *Aulacoseira*, 46.40% accuracy precision score for *Encyonema*, and 40.18% accuracy precision score for *Cymbella*. Notably, *Cyclotella* also showed a high accuracy precision score of 40.74%. This can be attributed to the unique circular shape of this particular genus compared to the rest of the classes. In the YOLOv4 model trained on the preprocessed image dataset, the model achieved accuracy precision scores of 67.70%

for *Aulacoseira*, 52.06% for *Encyonema*, and 41.08% for *Cymbella*. Similarly, *Cyclotella* also exhibited a high accuracy precision score of 49.85%. The numbers also indicate a slight increase in the number of false positives in the YOLOv4 model for the preprocessed dataset, affecting its overall mean accuracy precision compared to the raw dataset. This demonstrates how the YOLOv4 model trained on the raw dataset outperforms the one trained on the preprocessed dataset.

Based on the listed results, diatoms under the genera *Aulacoseira*, *Encyonema*, *Cymbella*, *Synedra*, *Navicula*, and *Cyclotella* have generally higher average precision on all four models trained for this study. This is directly correlated with the number of labelled cells under these genera which were used to train the YOLO models. The remaining genera, especially the *Achnanthes*, *Planothidium*, *Sellaphora*, *Surirella*, and *Rhoicosphenia* demonstrate very low precision scores. This discrepancy may be attributed to the lower number of cells from these genera in the dataset, which is consistently less than or equal to 15. Given that CNNs (Convolutional Neural Network) rely on analyzing and learning patterns to make accurate predictions, the underrepresentation of certain classes significantly impacted the system's performance. Figure 1 illustrates the substantial discrepancy between the actual cell labels and the predicted labels by the YOLOv7 model for the raw dataset. It is noticeable that genera with larger samples in the training dataset, such as *Aulacoseira*, *Encyonema*, and *Cymbella*, are detected more frequently than genera with fewer samples, like *Achnanthidium*, even though the latter is more prevalent in the test dataset. *Achnanthidium*, for example, has only about two (2) cell samples in the training dataset. This pattern was consistently observed across the other YOLO setups which further emphasizes the significance of a well-balanced and diverse training dataset for achieving optimal performance across various diatom genera. However, it is crucial to note that well-represented classes were still accurately detected and identified, as seen in the example of *Aulacoseira* cells within the blue bounding boxes. This underscores the system's potential to reduce the burden of classification and counting on diatom experts.

It is also noticeable that the performance of YOLOv4 and YOLOv7 differed when subjected to an additional pre-processing step. YOLOv4 achieved a mean average precision of 18.32% when trained on pre-processed images, whereas the model trained on raw images yielded a higher mean average precision of 19.21%. For YOLOv7, the model trained on raw images demonstrated a mean average precision of 2.88%, while the model trained on processed images exhibited a substantially higher mean average precision of 13.3%. This underscores the impact of pre-processing techniques on the performance of both YOLO models and emphasizes the importance of considering such factors in optimizing their effectiveness.

The impact of pre-processing on model performance differed between the two architectures. For YOLOv7, pre-processing significantly improved mAP (from 2.88% to 13.3%), suggesting that image enhancement-such as background normalization, contrast adjustment, or noise reduction-helped the deeper network better extract features from small and morphologically diverse cells. Conversely, YOLOv4 exhibited a slight decrease in mAP after pre-processing (from 19.21% to 18.32%), which may be due to

the model's simpler architecture already being well-suited to raw image patterns, and pre-processing potentially altering subtle textures or colors that the network relied on for detection. These results underscore the importance of tailoring pre-processing steps to the architecture being used and highlight that pre-processing is not universally beneficial.

Building on the insights gained from the analysis of the 16-class experimental setup, which evaluated the performance of YOLOv4 and YOLOv7, assessed the impact of image pre-processing, and highlighted the significance of the number of cells required to represent a diatom genus during training, a subsequent experiment focused on the top eight diatom genera. The genera *Aulacoseira*, *Fragilaria*, *Encyonema*, *Cymbella*, *Synedra*, *Nitzschia*, *Navicula*, and *Achnanthidium* were selected due to their high cell abundance (Table 4 and 5). The images used for both training and testing were also divided to achieve a close to 80-20 division of the cell numbers per genus for training and testing, respectively. In this additional experiment, the highest-performing model from previous analyses, YOLOv4 trained on a raw image dataset, was employed. The key distinction in this experiment was the exclusion of the other eight classes characterized by a lower cell count.

Table 4. Number of cells found per classification in the eight-class experimental setup.

Class	Training (no of cells)	Testing (no of cells)
<i>Aulacoseira</i>	179	40
<i>Fragilaria</i>	186	38
<i>Encyonema</i>	215	49
<i>Cymbella</i>	198	85
<i>Synedra</i>	81	41
<i>Nitzschia</i>	111	21
<i>Navicula</i>	88	18
<i>Achnanthidium</i>	196	9

Table 5. Average precision (AP), true positive (TP), and false positive (FP) of each class in the eight-class experimental setup using the YOLOv4 model trained on raw image dataset.

Class	AP	TP	FP
<i>Aulacoseira</i>	71.61%	33	29
<i>Fragilaria</i>	22.89%	11	17
<i>Encyonema</i>	49.15%	33	46
<i>Cymbella</i>	59.08%	51	26
<i>Synedra</i>	26.54%	11	7
<i>Nitzschia</i>	41.70%	11	12
<i>Navicula</i>	51.52%	9	4
<i>Achnanthidium</i>	13.33%	3	12



Figure 1. Comparison between the actual labels (top) and the predicted labels (bottom) of the YOLOv7 for the raw dataset.

The system demonstrated significant improvement, as indicated by its mean average precision score of 41.98%, a substantial increase compared to the 19.21% mAP achieved in the 16-class experiment. Additionally, the system achieved a precision of 51%, recall of 54%, and an F1-score of 53%. This underscores the importance of a large number of diatom cells per genus needed to effectively train a YOLO model, enabling not only the detection but also the classification of diatoms in a digital image.

Figure 2 displays the PR-curve for the one-class, 16-class, and eight-class experimental setups across all classes. PR-curves were used to gauge the model's performance in the positive class, which is the focus of this study – i.e., to detect and classify diatoms in a digital image. In the eight-class PR curve, precision values range from 0.25 to 0.63, with the highest precision at 0.63. This indicates that when the model predicts a positive class, it is correct approximately 63% of the time. The recall values range from 0.44 to 0.65, with the highest recall being 0.65. This implies that the model can correctly identify 65% of the actual positive instances. Comparing this with the PR curves for the one-class experiment and 16-class experiment shows the added complexity of the problem when introducing additional or reducing the number of diatom genera for classification.

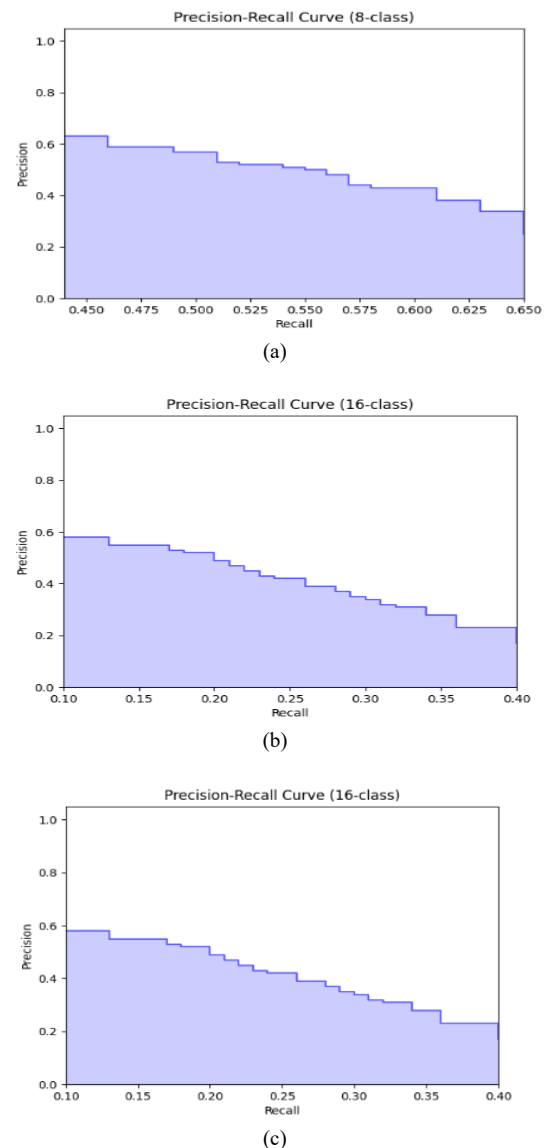


Figure 2. PR-curves for all classes of the (a) one-class, (b) 16-class, and (c) eight-class experimental setup.

Closer inspection of these scores also revealed the impact of imbalanced datasets on the average precision scores of each genus. Imbalanced datasets tend to exhibit a bias toward the majority classes. In extreme cases, the model may completely ignore the minority classes (Valova *et al.*, 2020). In this study, it can be observed that a higher number of false positives is present in most of the majority classes, especially *Aulacoseira*, *Encyonema*, and *Cymbella*. This indicates that a significant portion of the false detections made by the model falls under these classifications.

The decrease in the mAP score of the system can be attributed to another factor: the misidentification of certain genera due to similarities in their morphology. For instance, both *Cymbella* and *Encyonema* exhibit a boat-like or crescent shape in the images used for this study. The primary distinction lies in the appearance of their central lines, with *Encyonema* having a straight line and *Cymbella* featuring a curved one. In some instances, the system erroneously identified *Cymbella* as *Encyonema*, highlighting the challenge posed by their close resemblance. Another diatom genus, *Synedra*, and

Fragilaria, also faced a similar issue due to their close resemblance to each other. While *Fragilaria* typically has a rectangular shape compared to the pointed ends of *Synedra*, both exhibit elongated structures that may confuse the model.

Another factor contributing to the decrease in the mAP score of the system is the varying orientation of cells in the images, for example the *Cymbella* cell in a side view, unlike the typical front view seen in most images. This difference in orientation can lead to confusion in the pattern learned by the YOLOv4 model, ultimately impacting its performance on test images. *Fragilaria* faced a similar issue due to the variability of its samples across the datasets. In numerous instances, *Fragilaria* cells were found in groups, often overlapping with each other. However, there were also instances where a single cell appeared without any group, and instances where their sizes differed. This variability could potentially affect the system's performance in identifying cells under this genus. To enhance the model's ability to recognize the various configurations present in each image, more samples for training would be needed.

The automated approach offers a significant time advantage compared to traditional methods. The system requires only seconds to minutes for identifying and counting diatoms in 100 images, whereas the conventional method typically necessitates at least two days for 50 images. It is crucial to acknowledge that during the identification process, subtle distinctions in the cellular characteristics, such as shape, diameter, and size, of diatom species can be time-consuming and susceptible to human bias without the assistance of computer-aided automation procedures. To address this challenge, future studies employing this system could explore alternative pre-processing techniques that further emphasize the features of these cells. Improving the visibility of cellular edges and enhancing color in the images has the potential to enhance the classification outcomes of the system. It is also recommended to integrate CNNs with complementary techniques tailored specifically for diatom morphology analysis. Moreover, exposing the system to a diverse range of images taken under various conditions, including different lighting conditions, microscope magnifications, and capturing devices, can enhance the model's adaptability in identifying and recognizing diatoms in digital images.

The proposed automated approach demonstrates that reliable diatom identification can be achieved without removing the organic cover by leveraging visual features extracted from light microscopy images, such as overall cell shape, size, and geometric outlines, which remain discernible even when fine frustule details are partially obscured. Under these conditions, the system is most effective at the genus level, as major morphometric characteristics are sufficient for robust classification, whereas species-level identification remains constrained by subtle intra-genus morphological differences, cell orientation, and overlapping individuals inherent to light microscopy. Consequently, the method is positioned as a rapid, high-throughput screening tool that substantially reduces manual workload for diatomists, while acknowledging that species-level resolution would require higher-resolution imaging or complementary validation approaches. The use of 4% formalin as a preservative,

although not ideal for preserving delicate frustule features over long periods, was justified by its practicality and effectiveness in maintaining overall cell integrity for timely image acquisition; moreover, the reliance on general morphological patterns rather than ultrastructural details mitigates the impact of minor preservative-induced distortions. In comparison with earlier and contemporary automated identification frameworks, the present methodology advances the field by applying modern YOLO-based deep learning to enable simultaneous detection, classification, and counting of multiple freshwater diatom genera within complex images, achieving substantial gains in processing speed and scalability. While traditional and SEM-based methods remain indispensable for high-precision species-level taxonomy, the proposed approach provides a practical and efficient complement for large-scale, genus-level monitoring and ecological assessments.

The study aimed to mirror the natural process undertaken by diatom experts during microscopic slide inspections. Consequently, the distribution of each genus in the images could not be controlled, reflecting real-world scenarios. To mitigate this limitation, future efforts should consider introducing new datasets to the model where other classes are more prominently featured. This approach would enable a more comprehensive training experience, allowing the models to adapt to a wider range of scenarios and potentially leading to substantial improvements in their mean average precision scores. Fine-tuning the models with diverse datasets could enhance their performance and make them more robust in handling variations in diatom distribution in digital images.

The inclusion of additional diatom genera into the model will enhance the identification system for freshwater diatom species and contribute to its overall robustness. Integrating the developed identification system into a mobile application would offer a more convenient, flexible, and faster means of recognizing and identifying freshwater diatoms. It is recommended that future research concentrate on establishing a database management system for all images of Philippine freshwater diatom species. This proposed system should be accessible via the Internet to encourage ongoing indexing and tagging of new images. Furthermore, online databases can facilitate swift access and retrieval of historical data.

4. Conclusion

The study applied YOLO architecture to detect, identify, and count freshwater diatoms in the Philippines, achieving significant progress in computational recognition. Using YOLOv4, the initial detection phase reached a mean average precision (mAP) of 56.64%. In the classification phase, YOLOv4 without pre-processing performed best in detecting 16 genera, with a mAP of 19.21%, precision of 41%, recall of 25%, and an F1-score of 31%. Focusing on the eight most abundant genera in a third setup improved the mAP to 41.98%. Challenges such as dataset imbalance, morphological similarities, and cell orientation affected precision, but the system drastically outpaced manual analysis, processing 100 images in minutes. While currently limited to 16 genera, the study demonstrates machine vision's potential to automate and enhance diatom identification. Future work includes

diversifying the dataset, refining pre-processing, integrating CNNs, expanding genera coverage, developing a mobile app, and building an online Philippine diatom database to advance taxonomy and ecological monitoring.

Conflict Of Interests

The authors have no conflict of interests to declare.

Acknowledgment

The authors acknowledge the Institute of Computer Science and the National Institute of Molecular Biology and Biotechnology (BIOTECH-UPLB) for the technical and financial support provided for the study. Lastly, the authors are grateful for the constructive comments and suggestions of the journal editors and reviewers.

References

- Arguelles EDLR. 2019. Morphotaxonomic study of algal epiphytes from *Ipomoea aquatica* Forssk. (Convolvulaceae) found in Laguna de Bay (Philippines). *Pertanika J. Trop. Agric. Sci.*, **42**(2): 817–832.
- Arguelles EDLR. 2020. Evaluation of nutritional composition and in vitro antioxidant and antibacterial activities of *Codium intricatum* Okamura from Ilocos Norte (Philippines). *Jordan J. Biol. Sci.*, **13**: 375–382.
- Arguelles EDLR. 2021. Unseen microalgal diversity – phytotelmata in *Neoregelia* spp. L. B. Smith (Bromeliaceae) from florists wholesalers in Laguna, Philippines. *Philipp. J. Sci.*, **150**(1): 123–137. <https://doi.org/10.56899/150.01.10>
- Arguelles EDLR. 2022. Preliminary studies on the potential antioxidant and antidiabetic activities of *Sargassum polycystum* C. Agardh (Phaeophyceae, Ochrophyta). *Jordan J. Biol. Sci.*, **15**(3): 449–456.
- Arguelles EDLR. 2023. New distribution and rare records of microalgae from aquatic microcosms of *Guzmania lingulata* (L.) Mez (Bromeliaceae). *Philipp. J. Sci.*, **152**(2): 687–701. <https://doi.org/10.56899/152.02.15>
- Arguelles EDLR and Monsalud RG. 2017. Morphotaxonomy and diversity of terrestrial microalgae and cyanobacteria in biological crusts of soil from paddy fields of Los Baños, Laguna (Philippines). *Philipp. J. Syst. Biol.*, **11**(2): 25–36. <https://doi.org/10.26757/pjsb.2017b11016>
- Arguelles EDLR, Laurena AC, Monsalud RG, and Martinez-Goss MR. 2018. Fatty acid profile and fuel-derived physico-chemical properties of biodiesel obtained from an indigenous green microalga, *Desmodesmus* sp. (I-AU1), as potential source of renewable lipid and high-quality biodiesel. *J. Appl. Phycol.*, **30**(1): 411–419. <https://doi.org/10.1007/s10811-017-1264-6>
- Bozarth A, Maier UG, and Zauner S. 2009. Diatoms in biotechnology: Modern tools and applications. *Appl. Microbiol. Biotechnol.*, **82**(2): 195–201. <https://doi.org/10.1007/s00253-008-1804-8>
- Dimitrovski I, Kocev D, Loskovska S, and Džeroski S. 2012. Hierarchical classification of diatom images using ensembles of predictive clustering trees. *Ecol. Inform.*, **7**(1): 19–29. <https://doi.org/10.1016/j.ecoinf.2011.09.001>
- Geronimo JONV, Arguelles EDLR, and Abriol-Santos KJM. 2023. Automated classification and identification system for freshwater algae using convolutional neural networks. *Philipp. J. Sci.*, **152**(1): 325–335. <https://doi.org/10.56899/152.01.25>
- Kloster M, Kauer G, and Beszteri B. 2014. SHERPA: An image segmentation and outline feature extraction tool for diatoms and other objects. *BMC Bioinform.*, **15**: 218. <https://doi.org/10.1186/1471-2105-15-218>
- Libreros J, Bueno G, Trujillo M, and Ospina M. 2019. Automated identification and classification of diatoms from water resources. *Lecture Notes in Computer Science*, 11401 LNCS. https://doi.org/10.1007/978-3-030-13469-3_58
- Liu D, Wang P, Cheng Y and Bi H. 2022. An improved algae-YOLO model based on deep learning for object detection of ocean microalgae considering aquacultural lightweight deployment. *Front. Mar. Sci.*, **9**:1070638. <https://doi.org/10.3389/fmars.2022.1070638>.
- Lo JH, Lin LK, and Hung CC. 2023. Real-time personal protective equipment compliance detection based on deep learning algorithm. *Sustainability*, **15**(1): 391. <https://doi.org/10.3390/su15010391>
- Martinez-Goss MR, Manlapas J, and Arguelles EDLR. 2019. Cyanobacteria and diatoms in the cyanobacterial mats in a natural saltwater hot spring in Coron, Palawan, Philippines. *Philipp. Sci. Lett.*, **12**: 11–32.
- Martinez-Goss MR, Ohtsuka T, Inoue H, Arguelles EDLR, Ikeya T, Peralta E, Papa R, and Okuda N. 2023. *Gomphonema* species (Bacillariophyceae) from Marikina River, Rizal (Luzon), Philippines. *Philipp. J. Sci.*, **152**(5): 1653–1676. <https://doi.org/10.56899/152.05.11>
- Masouras A, Karaouzas I, Dimitriou E, Tsirtsis G, and Smeti E. 2021. Benthic diatoms in river biomonitoring – Present and future perspectives within the water framework directive. *Water*, **13**(4): 478. <https://doi.org/10.3390/w13040478>
- Siddik O and Bostan A. 2019. *Spirulina* detection using deep learning approach. *Br. J. Comput. Netw. Inf. Technol.*, **2**(1): 34–44.
- Syed T, Krujatz F, Ihadjadene Y, Mühlstädt G, Hamed H, Mädler J, and Urbas L. 2024. A review on machine learning approaches for microalgae cultivation systems. *Comput. Biol. Med.*, **172**: 108248. <https://doi.org/10.1016/j.compbiomed.2024.108248>
- Pedraza A, Bueno G, Deniz O, Cristóbal G, Blanco S, and Borrego-Ramos M. 2017. Automated diatom classification (Part B): A deep learning approach. *Appl. Sci.*, **7**(5): 460. <https://doi.org/10.3390/app7050460>
- Rimet F. 2012. Recent views on river pollution and diatoms. *Hydrobiologia*, **683**(1): 1–24. <https://doi.org/10.1007/s10750-011-0949-0>
- Shen L, Chen M, Lan B, Qi H, Zhang A, Lan D, and Fang Q. 2017. Diatom distribution as an environmental indicator in surface sediments of the West Philippine Basin. *Chin. J. Oceanol. Limnol.*, **35**(2): 431–443. <https://doi.org/10.1007/s00343-016-5306-8>
- Valova I, Harris C, Mai T, and Gueorguieva N. 2020. Optimization of convolutional neural networks for imbalanced set classification. *Procedia Comput. Sci.*, **176**: 660–669. <https://doi.org/10.1016/j.procs.2020.09.038>
- Wang Y, Ju P, Wang S, Su J, Zhai W, and Wu C. 2021. Identification of living and dead microalgae cells with digital holography and verified in the East China Sea. *Mar. Pollut. Bull.*, **163**: 111927. <https://doi.org/10.1016/j.marpolbul.2020.111927>
- Wang JK and Seibert M. 2017. Prospects for commercial production of diatoms. *Biotechnol. Biofuels*, **10**(1): 16. <https://doi.org/10.1186/s13068-017-0699-y>
- Yap-Dejeto LG and Baqueros ARS. 2015. Epiphytic diatom composition on the leaf blades of *Enhalus acoroides* in Almeria, Biliran, Philippines during the Northeast Monsoon of 2012–2013. *Philipp. Sci.*, **52**: 56–72.
- Yap-Dejeto LG, Omura T, Cinco GF, Cobacha MM, and Fukuyo Y. 2013. Species account of marine diatoms of the genus *Pseudonitzschia* in San Pedro Bay, Philippines. *Philipp. J. Sci.*, **142**(1): 27–37.