

IDUNAS	NATURAL & APPLIED SCIENCES JOURNAL	2026 Vol. 9 No. 1 (27-38)
--------	---------------------------------------	------------------------------------

Comparative Evaluation of YOLO Architectures with Optimized Preprocessing for Multi-Class Microalgae Detection

Research Article

Görkem Duman ^{1*} , Başak Esin Köktürk Güzel^{1,2} 

¹ İzmir Demokrasi University, Faculty of Engineering, Department of Electrical and Electronics Engineering, İzmir, Türkiye.

² İzmir Demokrasi University, Artificial Intelligence and Data Analytics Research and Application Center, İzmir, Türkiye.

Author E-mails:

Gorkemdumann7@gmail.com

Basak.guzel@idu.edu.tr

G. Duman ORCID ID: 0009-0002-7661-3762

B.E. Köktürk Güzel ORCID ID: 0000-0002-9429-1149

*Correspondence to: Görkem Duman, İzmir Demokrasi University, Department of Electrical and Electronics Engineering, İzmir, Türkiye.

DOI: 10.38061/idunas.1941287

Received: 30.04.2026; Accepted: 26.05.2026

Abstract

Accurate detection and counting of microalgae species are essential for biomass monitoring, contamination control, and smart cultivation management in biotechnology applications. However, automated analysis of microscopy images remains challenging because of dense cell distributions, small object sizes, low-contrast boundaries, and inter-class visual similarity. This study presents a two-stage deep learning framework for multi-class microalgae detection using recent YOLO architectures. In the first stage, five preprocessing pipelines were comparatively evaluated using YOLOv8s to determine an effective image enhancement strategy; LAB color space conversion combined with CLAHE and sharpening visibly improved cell boundary definition. In the second stage, the selected preprocessing configuration was fixed and four lightweight detectors (YOLOv8s, YOLOv9s, YOLOv10s, and YOLOv11s) were benchmarked under identical training conditions. YOLOv9s achieved the best overall performance with the highest precision (0.867), F1-score (0.824), and mAP50 (0.889). The findings indicate that detector architecture and preprocessing strategy jointly influence microscopic algae detection performance, and that model recency alone does not guarantee superior results.

Keywords: Microalgae detection; Deep learning; YOLO; Object detection; Microscopy analysis

1. INTRODUCTION

Microalgae are photosynthetic microorganisms of growing scientific and industrial interest. Their rapid growth rates and rich biochemical composition make them valuable for renewable energy, pharmaceuticals, aquaculture, wastewater treatment, and carbon sequestration [1]. Commercially relevant species such as *Chlorella vulgaris*, *Dunaliella salina*, *Haematococcus pluvialis*, and *Porphyridium cruentum* produce high-value metabolites, pigments, and nutritional compounds that are increasingly exploited in biotechnology pipelines [2].

Accurate real-time monitoring of microalgae populations is indispensable for biomass estimation, contamination control, productivity optimization, and species-specific cultivation management. Conventionally, such monitoring relies on manual microscopy-based counting, which is labor-intensive, time-consuming, and strongly dependent on operator expertise. In large-scale industrial bioreactors, manual inspection is neither scalable nor standardizable, creating a clear need for automated solutions [3]. Recent advances in computer vision and deep learning have demonstrated strong potential for automated biological image analysis. Object detection models that simultaneously localize and classify multiple targets in a single image are particularly well suited to dense microscopy scenes. Within this category, the YOLO (You Only Look Once) family has become a dominant framework owing to its single-stage architecture, high inference speed, and strong accuracy across diverse domains [4]. Subsequent YOLO generations — YOLOv5 through YOLOv11 — have progressively refined feature extraction, neck design, and loss functions, each claiming performance improvements over its predecessor.

Despite this progress, automated microalgae detection from microscopy images remains a difficult open problem. Raw microscopy images typically exhibit low global contrast, uneven illumination, and weak cell boundary definition, which degrade both localization accuracy and inter-class discrimination. Algal cells are translucent, their outlines blur into the background, and closely packed cells of similar species become nearly indistinguishable. These conditions cannot be overcome by feeding raw image data directly into an object detector. A preprocessing stage that restores contrast and sharpens boundaries is therefore a necessary first step rather than an optional refinement [5]. The visual impact of these challenges, and the improvements achieved by the proposed pipeline are illustrated in Figure 1 in Section 3.3.

Yet, a second separate question must also be answered: Once the preprocessing strategy is fixed, which YOLO architecture performs best for dense multi-class microalgae detection? Prior studies have evaluated individual or custom-modified architectures [6-8], but systematic comparisons of recent vanilla YOLO generations under identical preprocessing and training conditions are scarce. Without such a controlled benchmark, it is impossible to attribute observed performance differences to architecture rather than confounding experimental variables.

This study addresses both questions through a deliberate two-stage experimental design. In Stage 1, we isolate the effect of preprocessing by evaluating five enhancement pipelines — RGB baseline, LAB, LAB+CLAHE, LAB+CLAHE+Blur, and LAB+CLAHE+Sharpening — using YOLOv8s as a fixed reference detector. The preprocessing question must be resolved first because feeding a different image distribution to each architecture would confound the architecture comparison in Stage 2. In Stage 2, with preprocessing fixed to the configuration identified in Stage 1, we benchmark YOLOv8s, YOLOv9s, YOLOv10s, and YOLOv11s under identical training conditions, enabling a fair attribution of performance differences to architectural choices alone. The remainder of this paper is structured as follows: Section 2 reviews related work on YOLO-based microalgae detection and preprocessing strategies; Section 3 describes the dataset, preprocessing pipelines, model architectures, and evaluation protocol; Section 4

presents and discusses experimental results; and Section 5 concludes with practical implications and future directions.

2. LITERATURE REVIEW

2.1. YOLO-Based Object Detection in Biological Microscopy

The YOLO framework has been progressively adopted for microorganism detection since its success in natural image benchmarks. Early work by Park et al. (2021) applied YOLOv3 and YOLOv4 to algae cell detection, achieving a mAP of 81% while demonstrating that lightweight model variants could reduce inference time by up to ten-fold without severe accuracy loss [9]. Abdullah et al. (2022) tested YOLOv3, YOLOv4, and YOLOv5 on a custom algae dataset across four species, with YOLOv5 achieving the highest mAP of 90.1%. These early benchmarks established YOLO as a viable baseline but also revealed that standard configurations often struggle with dense, small-scale biological objects [10].

More recent work has sought architectural modifications to address these limitations. Liu et al. (2022) proposed that Algae-YOLO, a lightweight YOLOv5 variant that replaces the standard backbone with ShuffleNetV2 and integrates an ECA attention mechanism, achieves detection accuracy comparable to the full YOLOv5s while substantially reducing parameter count, a critical advantage for laboratory deployment. Yan et al. (2023) designed a YOLOx-based pipeline for multi-scale microalgae comprising *Chlorella*, *Scenedesmus*, and *Spirulina*, incorporating Focal and DIoU Loss to address class imbalance; their model reached a precision of 95.93% and a recall of 93.48% on a purpose-built multi-morphology dataset [6]. Duan et al. (2024) introduced the RWD-YOLO model — an improved YOLOv5s integrating a Receptive Field Enhancement module, Wise-IoU v3, and a Dynamic detection head — achieving a mAP₅₀ of 93.2% and mAP₅₀₋₉₅ of 65.1%, representing improvements of 3.7 and 5.7 percentage points over baseline YOLOv5s [8]. Most recently, a Feature-Enhanced YOLOv7 variant (FE-YOLO) incorporating CAGS attention and SIoU loss was evaluated on a seven-species dataset of 6,300 images, yielding further gains in both precision and recall [11].

Collectively, these studies demonstrate a clear trend: Performance on microalgae detection improves when architectures are adapted to handle small, densely packed objects. However, they share a common limitation. Each study introduces architectural modifications alongside changes in dataset, preprocessing, and hyperparameters, making cross-study comparisons unreliable. The presented work deliberately avoids architectural modification and instead evaluates unmodified, publicly available YOLO variants under a single controlled experimental setting, enabling conclusions that generalize more directly to practitioner model selection.

2.2. The Role of Preprocessing in Microscopy-Based Detection

Raw microscopy images of microalgae present several optical challenges that are not typically encountered in natural image datasets such as globally low contrast due to the translucent nature of algal cells, uneven illumination from microscope light sources, and weak boundary definition between cells of adjacent species. These conditions are not addressed by standard deep learning data augmentation pipelines which assume that the input already contains adequate contrast and feature separability.

The LAB color space separates luminance (L) from chromatic information (a, b), making it possible to enhance brightness-related features without distorting color. Applying Contrast Limited Adaptive Histogram Equalization (CLAHE) to the L channel — rather than globally to the RGB image — improves local contrast in low-brightness regions while suppressing noise amplification via the clip limit parameter [5]. This LAB+CLAHE approach has been validated across multiple biomedical imaging domains. In deep

learning-based segmentation of temporomandibular joint MRI, CLAHE preprocessing significantly improved Dice similarity coefficient, sensitivity, and positive predictive value compared with raw image training, particularly when scanner-induced contrast variability was presented [12]. Similarly, in retinal fundus image analysis, CLAHE correction was reported as essential for preventing contrast-related false positives in geographic atrophy segmentation models. In a broader object detection context, Wang et al. (2023) demonstrated that a CLAHE-based preprocessing chain applied to the LAB luminance channel consistently improved object detection performance in low-contrast and shadow-affected scenes, supporting its applicability beyond medical imaging [13].

Sharpening filters applied after contrast enhancement can further recover high-frequency edge information that is compressed during histogram equalization. The combination of LAB conversion, CLAHE, and sharpening thus forms a theoretically motivated pipeline for microscopy preprocessing. However, to our knowledge, no prior study has systematically evaluated this exact pipeline against its component-wise alternatives in the context of multi-class microalgae detection with a standardized YOLO baseline. This gap motivates the first stage of our experimental design.

2.3. Comparative YOLO Benchmarks and Research Gaps

While microalgae detection literature has grown substantially, direct comparisons between unmodified recent YOLO generations on a shared dataset remain rare. Most published studies optimize a specific architecture for a specific dataset, making it impossible to disentangle the contribution of architectural advances from dataset-specific tuning. The VisAlgae 2023 challenge introduced a standardized multi-class algae dataset and evaluation protocol, providing the necessary controlled setting for a fair architecture comparison — yet comparative benchmarks across YOLOv8 through YOLOv11 on this dataset have not been published [14]. To the best of our knowledge, this study is among the first to provide a controlled benchmark of four consecutive YOLO generations under a unified preprocessing and training configuration on the VisAlgae dataset.

3. MATERIAL AND METHODS

3.1. Dataset

The experiments were conducted on the High-Throughput Algae Cell Detection dataset introduced by Sun et al. [14], which provides microscopy images annotated for six algae species: *Chlorella*, *Dunaliella salina*, *Effrenium*, *Haematococcus*, *Platymonas*, and *Porphyridium*. The dataset encompasses substantial visual diversity in terms of cell density, species co-occurrence, and imaging conditions, making it a challenging and ecologically representative benchmark for multi-class detection.

A total of 1,000 labeled images were used. The dataset was partitioned into 560 training images (56%), 140 validation images (14%), and 300 test images (30%). The validation subset was obtained by randomly sampling 20% of the training data using a fixed random seed of 42, ensuring reproducibility. The test set was held out completely during all training, model selection, and preprocessing evaluation steps, preventing any information leakage into reported performance estimates. The relatively large test proportion was chosen deliberately to enable statistically stable class-wise counting analysis across six species.

3.2. Two-Stage Experimental Framework

The central methodological contribution of this study has a two-stage design, which decouples the preprocessing optimization problem from the architecture selection problem. Conflating the two, as it is

common in the literature, makes it impossible to determine whether observed performance differences stem from the model or from the data preparation. By resolving preprocessing in Stage 1 before touching the architecture variable, we ensure that Stage 2 comparisons reflect architectural differences alone.

Stage 1: Preprocessing Optimization. YOLOv8s was chosen as the reference detector for Stage 1 because it represents a well-established, computationally efficient baseline with stable training behavior. Using a fixed, widely validated detector allowed us to attribute any performance difference across preprocessing conditions exclusively to image enhancement quality. Five configurations were evaluated: (i) RGB — raw images with no enhancement; (ii) LAB — color space conversion only; (iii) LAB+CLAHE — local contrast enhancement applied to the luminance channel; (iv) LAB+CLAHE+Blur — Gaussian smoothing applied after contrast enhancement to suppress noise; and (v) LAB+CLAHE+Sharpening — an unsharp mask applied after contrast enhancement to recover edge detail.

Stage 2: Architecture Benchmarking. The preprocessing configuration selected at the end of Stage 1 was frozen. Four lightweight YOLO architectures — YOLOv8s, YOLOv9s, YOLOv10s, and YOLOv11s — were then trained under identical hyperparameters on the preprocessed dataset. The use of the small ("s") variant across all generations ensures that computational scale is held constant, isolating the effect of architectural design choices.

3.3. Preprocessing Pipeline Details

As illustrated in Figure 1, the raw microscopy images in this dataset exhibit globally low contrast and indistinct cell boundaries — a consequence of the translucent nature of algal cells and the diffuse illumination typical of bright-field microscopy. These properties directly harm the detector: low contrast reduces the gradient signal available during feature extraction, and blurred boundaries cause regression uncertainty in bounding-box prediction.

To address these issues, the LAB+CLAHE+Sharpening pipeline was constructed as follows. First, the RGB image was converted to LAB color space, separating luminance from chrominance and enabling contrast enhancement without distorting the color information that aids inter-class discrimination. Second, CLAHE was applied to the L channel with a clip limit designed to prevent noise amplification in already-bright regions while boosting local contrast in low-intensity areas corresponding to cell interiors and boundaries. Third, an unsharp masking filter was applied to the enhanced luminance channel to amplify high-frequency edge information, making cell membranes more salient to the convolutional feature extractor. The final image was reconverted to RGB for compatibility with standard YOLO input pipelines. Figure 1 shows the visual effect of each stage of this pipeline on a representative image from the dataset, illustrating the progressive improvement in cell boundary definition and inter-cell contrast.

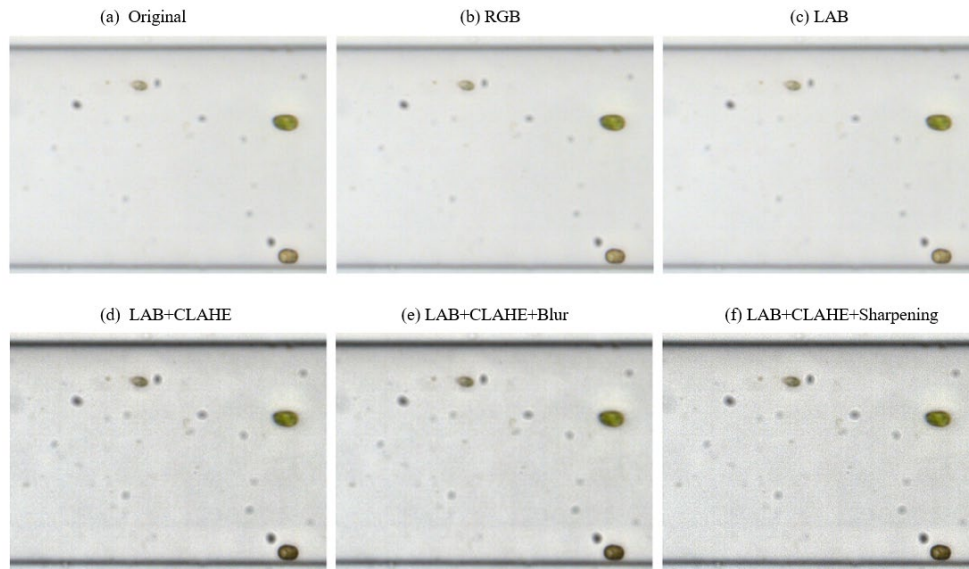


Figure 1. Visual comparison of evaluated preprocessing pipelines. From left to right: (a,b) RGB baseline, (c) LAB, (d) LAB+CLAHE, (e) LAB+CLAHE+Blur, (f) LAB+CLAHE+Sharpening.

3.4. Model Architectures

Four lightweight YOLO variants spanning three years of architectural development were selected for benchmarking. YOLOv8s, released by Ultralytics in 2023, introduced an anchor-free detection head and a redesigned C2f bottleneck module. YOLOv9s, introduced in 2024, added a Programmable Gradient Information (PGI) mechanism and a Generalized Efficient Layer Aggregation Network (GELAN) backbone, designed to preserve gradient information more effectively during training [15]. YOLOv10s, also from 2024, proposed a dual-label assignment strategy that eliminates the need for non-maximum suppression at inference time, reducing latency. YOLOv11s, released in late 2024, further refined the attention and aggregation mechanisms with the stated goal of improved small-object detection. All four models were used in their publicly available, unmodified configurations.

3.5. Training Configuration and Data Augmentation

To ensure a fair comparison across architectures, all models were trained with identical hyperparameters: 50 epochs, batch size 8, input resolution 640×640 pixels, and the Ultralytics default optimizer and augmentation settings. Experiments were conducted on Kaggle Notebooks using dual NVIDIA Tesla T4 GPUs. The default augmentation pipeline includes mosaic augmentation, random horizontal and vertical flipping, HSV color perturbation, random scaling, and random translation. No custom augmentation was added, as doing so would introduce an additional confounding variable.

3.6. Evaluation Metrics

Models were evaluated on the held-out test set using precision, recall, F1-score, mAP50, and mAP50-95. Precision measures the fraction of predicted detections that are correct; recall measures the fraction of ground-truth objects that are successfully detected; F1-score combines both into a single balanced metric; mAP50 is the mean Average Precision at IoU threshold 0.5; and mAP50-95 averages mAP over thresholds from 0.5 to 0.95 in steps of 0.05, providing a stricter measure of localization quality. In addition to these detection metrics, a class-wise counting analysis was performed by aggregating predicted bounding boxes per species across all test images and comparing aggregate predicted counts to ground-

truth counts. This analysis captures reliable population estimation which is a practically important capability that aggregate mAP scores do not fully reflect.

4. RESULTS AND DISCUSSION

4.1. Stage 1: Effect of Preprocessing on Detection Performance

Table 1 summarizes the detection metrics obtained by YOLOv8s under each of the five preprocessing configurations. The raw RGB baseline already achieves a competitive mAP50 of 0.9439, reflecting the reasonable inherent quality of the dataset. LAB color space conversion alone slightly reduces performance (mAP50 = 0.9364), suggesting that the color separation benefit is only realized when contrast enhancement is also applied. Adding CLAHE to LAB recovers and marginally exceeds the RGB baseline (mAP50 = 0.9440, mAP50-95 = 0.7304), confirming that local contrast enhancement provides a measurable benefit for cell detection in this domain —consistent with findings reported in biomedical imaging contexts where CLAHE has been shown to improve model robustness to contrast variability [12,13].

Introducing Gaussian blur after CLAHE (LAB+CLAHE+Blur) reduces both mAP50 (0.9400) and mAP50-95 (0.7281) relative to LAB+CLAHE, indicating that noise suppression comes at the cost of suppressing the edge information that the detector needs to localize small cells. Adding sharpening instead of blur (LAB+CLAHE+Sharpening) produces the lowest aggregate scores in this group (mAP50 = 0.9300, mAP50-95 = 0.7116). This result might initially suggest that sharpening is counterproductive, but a closer inspection of visual outputs (Figure 1) tells a different story. As visible in the rightmost panel of Figure 1, the LAB+CLAHE+Sharpening image displays markedly clearer cell boundaries and stronger contrast between adjacent cells than any other configuration, particularly in regions of high cell density.

This observation highlights a key methodological point: in dense microscopic object detection, but aggregate quantitative metrics do not tell the full story. A preprocessing pipeline that produces visually superior boundary definition and inter-cell contrast may support better downstream generalization even when it does not achieve the highest mAP on the validation set. Although the LAB+CLAHE+Sharpening pipeline did not produce the highest aggregate validation mAP, qualitative inspection revealed substantially clearer cell boundaries and improved visual separability in dense overlapping regions. Since microalgae monitoring requires not only box-level detection accuracy but also reliable object delineation and counting in crowded scenes, this configuration was selected for Stage 2 as a practically motivated preprocessing strategy.

Table 1. Performance comparison of preprocessing pipelines using YOLOv8s.

Preprocessing Pipeline	Precision	Recall	mAP50	mAP50-95
RGB	0.9298	0.9140	0.9439	0.7284
LAB	0.8923	0.9100	0.9364	0.7232
LAB + CLAHE	0.9187	0.9100	0.9440	0.7304
LAB + CLAHE + Blur	0.9158	0.9000	0.9400	0.7281
LAB + CLAHE + Sharpening	0.9000	0.8900	0.9300	0.7116

4.2. Stage 2: Comparative Evaluation of YOLO Architectures

Table 2 presents the detection performance of all four architectures under the LAB+CLAHE+Sharpening preprocessing pipeline on the held-out test set. The results reveal a clear performance ordering with substantial differences between the best and worst performing models.

YOLOv9s achieves the strongest results across all five metrics: precision 0.867, recall 0.783, F1-score 0.824, mAP50 0.889, and mAP50-95 0.618. These figures compare favorably with the state of the art on related datasets. Duan et al. [8], using a heavily modified YOLOv5s (RWD-YOLO) on a water-quality algae dataset, reported a mAP50 of 93.2% with a higher absolute value, but achieved on a different dataset with substantial architectural engineering. Yan et al. [6], using improved YOLOx on a three-species dataset, reached 95.93% precision and 93.48% recall on single-class scale-adapted images. That our unmodified YOLOv9s achieves 86.7% precision and 88.9% mAP50 on a six-class, densely packed dataset without any architectural modification is therefore a meaningful result. It suggests that the architectural improvements introduced in YOLOv9, particularly the PGI mechanism, which preserves gradient information across layers, provide genuine benefits for dense multi-class detection tasks without requiring task-specific customization.

YOLOv10s ranks second with precision 0.841, recall 0.778, F1-score 0.806, and mAP50 0.871. The performance gap between YOLOv9s and YOLOv10s is modest, and YOLOv10s has a practical advantage: its NMS-free inference head reduces post-processing latency, making it potentially preferable in real-time laboratory monitoring scenarios where throughput is critical.

YOLOv11s, despite being the most recently released architecture, ranks third overall with mAP50 0.814 and F1-score 0.747. Its recall (0.699) is notably lower than the other models, suggesting a tendency to miss detections in difficult regions which have likely densely packed cells where boundary separability is limited. This finding is consistent with a broader observation in the object detection literature: architectural advances targeting natural image benchmarks do not automatically transfer to domain-specific challenges such as dense microscopic imaging. Liu et al. [7] made a similar observation when comparing standard YOLOv5 configurations on low-magnification algae images, noting that backbone complexity alone did not determine performance in this domain.

YOLOv8s produces the weakest results (precision 0.709, mAP50 0.732), lagging behind all three later architectures. Given that YOLOv8s was also used as the preprocessing baseline in Stage 1, where it achieved mAP50 values of 0.92–0.94 under various preprocessing conditions, the performance drop observed here reflects the additional difficulty of the held-out test set compared with the validation set, as well as the relative disadvantage of the older architecture under difficult imaging conditions. The higher scores observed in Stage 1 relative to Stage 2 are attributable to the use of the validation subset for preprocessing selection, whereas Stage 2 reports performance on the fully held-out test set, which contains more challenging dense-scene samples and therefore provides a stricter estimate of generalization performance.

Table 2. Comparative performance of YOLO architectures using the LAB+CLAHE+Sharpening preprocessing pipeline

Model	Precision	Recall	F1-Score	mAP50	mAP50-95
YOLOv8s	0.709	0.681	0.695	0.732	0.521
YOLOv9s	0.867	0.783	0.824	0.889	0.618
YOLOv10s	0.841	0.778	0.806	0.871	0.607
YOLOv11s	0.802	0.699	0.747	0.814	0.563

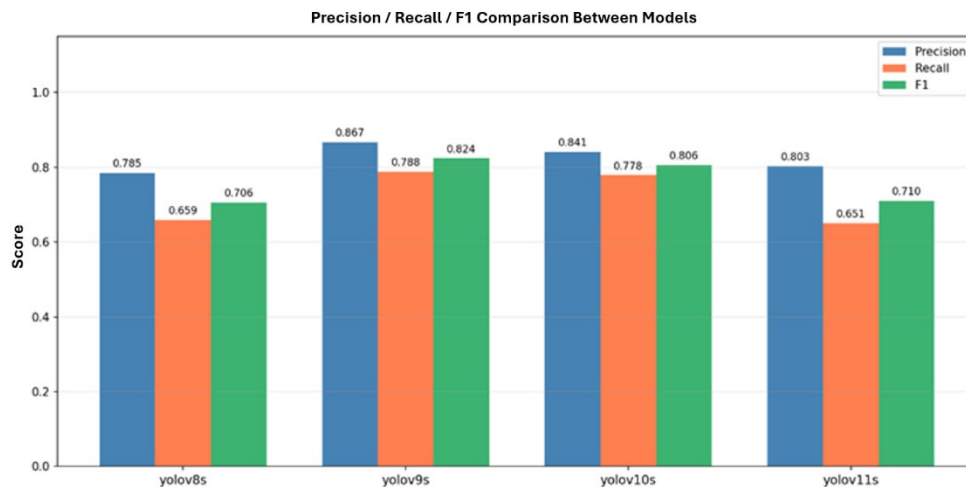


Figure 2. Comparative precision, recall, and F1-score of YOLOv8s, YOLOv9s, YOLOv10s, and YOLOv11s on the test set.

Figure 2 visualizes the precision, recall, and F1-score profiles of all four models. YOLOv9s exhibits the most balanced profile, with neither precision nor recall disproportionately dominating, which is a desirable property for multi-class monitoring systems where both false positives and false negatives carry operational costs. YOLOv11s, by contrast, shows a recall deficit that widens the gap between its precision and recall, indicating systematic under-detection.

4.3. Class-wise Counting Analysis

Table 3 reports aggregated predicted and ground-truth cell counts for each algae species across the 300 test images. This analysis reveals patterns that are not visible in the aggregate mAP scores and that are directly relevant to the practical use case of population monitoring.

YOLOv9s and YOLOv10s consistently produce count estimates closest to ground truth. For *Porphyridium*, the species with the largest ground-truth count (790 cells), YOLOv9s predicts 784 and YOLOv10s predicts 810, errors of less than 3% in both cases. For *Platymonas*, YOLOv9s predicts 322 against 328 ground truth — an error below 2%. These results suggest that the architectural strengths of YOLOv9s and YOLOv10s translate into reliable population estimation, not just accurate bounding-box localization.

The most pronounced counting failure occurs for *Chlorella*, where YOLOv8s dramatically over-detects (1406 predicted versus 972 ground truth — a 45% overcount). This over-detection pattern is characteristic of models with weaker spatial disambiguation: when cells overlap or cluster tightly, the detector produces multiple low-confidence predictions that collectively survive thresholding. YOLOv11s shows the opposite pattern, a systematic under-detection across multiple species, including *Dunaliella salina* (230 vs. 400 ground truth) and *Porphyridium* (640 vs. 790). The low recall of YOLOv11s observed in Table 2 manifests directly as population underestimates in these ecologically important species.

These class-wise asymmetries underscore the argument that aggregating mAP alone is an insufficient performance criterion for biological monitoring applications. A practitioner deploying a model to estimate species abundance needs assurance that the model does not systematically over- or under-count specific species — a requirement that class-wise counting analysis directly addresses.

Table 3. Ground-truth and predicted algae cell count for each YOLO model

Algae Class	Ground Truth	YOLOv8s	YOLOv9s	YOLOv10s	YOLOv11s
<i>Chlorella</i>	972	1406	844	924	1018
<i>Dunaliella salina</i>	400	294	310	330	230
<i>Effrenium</i>	556	514	522	552	498
<i>Haematococcus</i>	208	176	180	196	170
<i>Platymonas</i>	328	248	322	310	250
<i>Porphyridium</i>	790	510	784	810	640

4.4. Synthesis and Practical Implications

Taken together, the two-stage results paint a coherent picture. The preprocessing stage revealed that image enhancement meaningfully affects detection quality, but the relationship between quantitative metrics and practical utility is nuanced: the pipeline with the best visual output (LAB+CLAHE+Sharpening) did not achieve the highest mAP50 on the validation set. This is a cautionary finding for practitioners who rely exclusively on validation metrics to select preprocessing strategies.

The architecture benchmark showed that YOLOv9s provides the best overall trade-off for dense multi-class microalgae detection among the four candidates. Importantly, neither the most recent architecture (YOLOv11s) nor the most computationally efficient (YOLOv8s) achieved this balance, an outcome that challenges the common assumption that architectural recency correlates with downstream performance on specialized tasks. This finding aligns with observations from Liu et al. [7] and Yan et al. [6], who both found that architectural modifications needed to be specifically tailored to the properties of microscopic algae images to be effective.

From an application perspective, YOLOv9s is the recommended choice for systems requiring balanced precision and recall, such as contamination screening or multi-species biomass monitoring. YOLOv10s is a strong alternative when real-time inference latency is prioritized, given its NMS-free design. Both models achieve class-wise counting accuracy that would support automated population estimation in photobioreactor monitoring and laboratory cultivation management systems.

Moreover, several limitations of this study should be acknowledged. All experiments were conducted on a single dataset; cross-dataset validation is needed to confirm generalizability. Only lightweight ("s") model variants were evaluated; medium and large variants may show different ordering. Architectural evaluation was restricted to unmodified YOLO configurations — domain-adapted modifications, as explored by Duan et al. [8] and Yan et al. [6], may further improve performance. Future work should also investigate transformer-based detectors, instance segmentation approaches for overlapping cells, and adaptive preprocessing pipelines that adjust enhancement parameters to local image statistics.

5. CONCLUSION

This study presented a controlled two-stage framework for automated multi-class microalgae detection that separates the preprocessing optimization problem from the architecture selection problem, a methodological distinction that is often overlooked in the literature. By resolving preprocessing before introducing architectural variation, we obtained results that attribute performance differences cleanly to their source.

Stage 1 demonstrated that LAB color space conversion combined with CLAHE and unsharp mask sharpening improves the visual quality and practical utility of microscopy images for dense cell detection, even when this pipeline does not maximize aggregate validation mAP. This finding highlights the importance of qualitative evaluation and class-wise analysis as complements to standard detection metrics, which is a lesson that extends beyond microalgae detection to any domain where aggregate mAP may mask class-specific or boundary-level weaknesses.

Stage 2 showed that YOLOv9s achieves the strongest and most balanced detection performance among four recent lightweight YOLO generations by outperforming the older YOLOv8s, the newer YOLOv11s, and the competing YOLOv10s on most metrics. The finding that the most recent architecture does not perform best is a practically important result: it warns against the common practice of defaulting to the latest model version without empirical validation on the target domain. YOLOv10s is identified as a compelling alternative when inference latency is a priority, while YOLOv11s shows a systematic recall deficit that should be considered before deployment.

Class-wise counting analysis confirmed that the performance ranking from mAP metrics translates to population estimation accuracy: YOLOv9s and YOLOv10s both produce species-level count estimates within a few percent of ground truth for most classes, while YOLOv8s over-detects *Chlorella* by 45% and YOLOv11s systematically under-counts multiple species. These results directly support the use of these models in automated biomass monitoring, contamination screening, and smart photobioreactor management.

Future research should extend this benchmark to medium and large YOLO variants, validate the selected pipeline on additional microscopy datasets from different imaging systems, and explore domain-adapted architectures, segmentation-based cell delineation, and cross-dataset transfer learning as avenues for further performance improvement.

REFERENCES

1. Borowitzka, M. A. (2013). High-value products from microalgae—their development and commercialisation. *Journal of applied phycology*, 25(3), 743-756.
2. Khan, M. I., Shin, J. H., & Kim, J. D. (2018). The promising future of microalgae: current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. *Microbial cell factories*, 17(1), 36.
3. Hlavova, M., Turoczy, Z., & Bisova, K. (2015). Improving microalgae for biotechnology—From genetics to synthetic biology. *Biotechnology advances*, 33(6), 1194-1203.
4. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

5. Gonzalez, R.C. and Woods, R.E. (2018) Digital Image Processing. 4th Edition, Pearson Education, New York, 1022 p.
6. Yan, H., Peng, X., Chen, C., Xia, A., Huang, Y., Zhu, X., ... & Liao, Q. (2023). YOLOx model-based object detection for microalgal bioprocess. *Algal Research*, 74, 103178.
7. Liu, D., Wang, P., Cheng, Y., & Bi, H. (2022). An improved algae-YOLO model based on deep learning for object detection of ocean microalgae considering aquacultural lightweight deployment. *Frontiers in Marine Science*, 9, 1070638.
8. Duan, Z., Xie, T., Wang, L., Chen, Y., & Wu, J. (2024). Microalgae detection based on improved YOLOv5. *IET image processing*, 18(10), 2602-2613.
9. Park, J., Baek, J., You, K., Nam, S.W., & Kim, J. (2021). Microalgae detection using a deep learning object detection algorithm, YOLOv3. *J. Korean Soc. Water Environ.*, 37, 275–285.
10. Abdullah, Ali, S., Khan, Z., Hussain, A., Athar, A., & Kim, H. C. (2022). Computer vision based deep learning approach for the detection and classification of algae species using microscopic images. *Water*, 14(14), 2219.
11. Ding, G., Shi, Y., Liu, Z., Wang, Y., Yao, Z., Zhou, D., ... & Li, Y. (2025). Fe-yolo: an efficient deep learning model based on feature-enhanced yolov7 for microalgae identification and detection. *Biomimetics*, 10(1), 62.
12. Yoshimi, Y., Mine, Y., Ito, S., Takeda, S., Okazaki, S., Nakamoto, T., ... & Tanimoto, K. (2024). Image preprocessing with contrast-limited adaptive histogram equalization improves the segmentation performance of deep learning for the articular disk of the temporomandibular joint on magnetic resonance images. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, 138(1), 128-141.
13. Wang, T. S., Kim, G. T., Kim, M., & Jang, J. (2023). Contrast enhancement-based preprocessing process to improve deep learning object task performance and results. *Applied Sciences*, 13(19), 10760.
14. Sun, M., Jiang, J., Yang, Z., Kong, S., Qi, J., Shang, J., Luo, S., Sun, W., Wang, T., Wang, Y., Wang, Q., Dai, T., Chen, T., Zhang, J., Zhang, X., He, Y., Fu, P., Guan, Q., Zhou, S., ... Yan, H. (2025). VisAlgae 2023: A Dataset and Challenge for Algae Detection in Microscopy Images. *arXiv preprint arXiv:2505.20687*.
15. Wang, C. Y., Yeh, I. H., & Mark Liao, H. Y. (2024, September). Yolov9: Learning what you want to learn using programmable gradient information. In *European conference on computer vision* (pp. 1-21). Cham: Springer Nature Switzerland.