Computer Vision for Underwater Biomass Estimation in Aquafarming: Intelligent Imaging Systems, AI-Based Modeling, and Precision Aquaculture Monitoring

Dr. Jean Luc Moreau
School of Agriculture and Environment, University of Western Australia (Agri-robotics in CEA), Australia

* Corresponding Author: **Dr. Jean Luc Moreau**

Article Info

P-ISSN: 3051-3421
E-ISSN: 3051-343X
Volume: 05
Issue: 01
Received: 18-02-2024
Accepted: 20-03-2024
Published: 22-04-2024
Page No: 40-48

Abstract
Accurate biomass estimation is critical for optimizing feed management, growth monitoring, and harvest planning in commercial aquafarming operations. Traditional methods involving manual sampling and physical measurements are labor-intensive, stressful to aquatic organisms, and provide only sparse temporal data. This article presents a comprehensive review of computer vision-based approaches for non-invasive underwater biomass estimation in controlled aquaculture environments. We examine the development of intelligent imaging systems designed to operate in challenging submerged conditions, including stereoscopic and monocular camera configurations, specialized lighting systems, and adaptive optical designs. The integration of artificial intelligence, particularly deep learning architectures for object detection, segmentation, and regression-based biomass prediction, has enabled automated analysis of fish size, weight, and population density from continuous video streams. Key challenges include underwater image degradation from turbidity, variable illumination, occlusion in dense populations, and the need for species-specific calibration models. Applications span cage-based marine aquaculture, recirculating aquaculture systems, and offshore farming installations, where real-time biomass monitoring supports precision feeding strategies and early detection of growth anomalies. Future developments will focus on improved generalization across species and environmental conditions, integration with digital twin frameworks for holistic farm management, and deployment of autonomous underwater vehicles for large-scale monitoring. Computer vision-based biomass estimation represents a transformative technology enabling sustainable intensification and data-driven decision-making in modern aquafarming.

Keywords: Underwater computer vision, Biomass estimation, Precision aquaculture, Artificial intelligence, Fish size and weight estimation, Smart aquafarming

Introduction

Biomass estimation constitutes a fundamental operational requirement in commercial aquafarming, directly influencing feeding protocols, stock management decisions, and economic performance [1, 2]. Precise knowledge of total biomass and individual fish size distributions enables farmers to optimize feed conversion ratios, predict harvest timing, and detect growth abnormalities that may indicate disease or suboptimal environmental conditions [3]. In intensive aquaculture systems, where production efficiency and resource utilization are paramount, even modest improvements in biomass estimation accuracy can yield substantial economic benefits through reduced feed waste and improved growth rates [4].

Conventional biomass estimation techniques rely predominantly on periodic manual sampling, whereby a subset of the population is netted, anesthetized, measured, and extrapolated to estimate total stock biomass [5]. These invasive procedures induce physiological stress responses in fish, potentially compromising immune function and growth performance for extended periods following handling [6]. Furthermore, manual sampling provides only discrete temporal snapshots of biomass, limiting the ability to detect short-term growth variations or respond dynamically to changing farm conditions [7]. The labor requirements

The labor requirements and operational risks associated with frequent sampling create practical constraints on monitoring frequency, particularly in offshore cage systems where access may be weather-dependent [8].

The inherent limitations of traditional methodologies have motivated the development of automated, non-invasive biomass estimation solutions. Computer vision-based approaches offer the potential for continuous monitoring without physical contact, generating high-resolution temporal data on growth dynamics while eliminating stress-related impacts on fish welfare [9, 10]. By deploying underwater imaging systems within production environments, farmers can acquire real-time insights into biomass accumulation rates, size distribution heterogeneity, and individual growth trajectories [11]. These capabilities align with the broader objectives of precision aquaculture, which seeks to apply sensor technologies and data-driven decision support systems to optimize production outcomes while minimizing environmental footprint [12].

Recent advances in artificial intelligence, particularly deep learning methodologies for image analysis, have dramatically enhanced the feasibility of automated underwater biomass estimation [13]. Convolutional neural networks and related architectures can learn complex visual patterns directly from annotated training data, enabling robust detection and measurement of fish even under challenging conditions of turbidity, variable illumination, and partial occlusion [14, 15]. The convergence of improved computational performance, specialized underwater imaging hardware, and sophisticated machine learning algorithms has positioned computer vision as a transformative technology for next-generation aquafarming operations [16].

2. Overview of Biomass Estimation Techniques in Aquaculture

Traditional Sampling-Based Methods

Classical biomass estimation protocols in aquaculture have historically depended on physical sampling and extrapolation techniques [17]. The standard approach involves capturing a representative subset of individuals, typically 1-5% of the population, using nets or specialized harvesting equipment [18]. Captured fish are anesthetized to facilitate safe handling, then individually weighed and measured for length using manual or electronic scales and measuring boards [19]. Statistical methods, often assuming log-normal or Gaussian length distributions, are applied to extrapolate sample measurements to total population biomass [20].

The accuracy of sampling-based estimation is fundamentally constrained by sample size and representativeness. Small samples may fail to capture the full range of size variation within heterogeneous populations, leading to systematic bias in biomass predictions [21]. Behavioral factors, including differential catchability of size classes and avoidance responses to sampling gear, can further compromise sample representativeness [22]. The stress imposed by capture, handling, and anesthesia triggers cortisol elevation and metabolic disruption that may suppress feeding and growth for days following sampling events [23]. In production systems operating near carrying capacity, such disturbances can have measurable impacts on cumulative growth performance over production cycles [24].

Acoustic and Sensor-Based Approaches

Acoustic technologies have been investigated as alternatives to physical sampling for biomass assessment in aquaculture [25]. Echo sounders and split-beam sonar systems can detect fish targets and estimate abundance within defined water volumes [26]. By combining acoustic backscatter intensity with length-target strength relationships, it is theoretically possible to infer biomass from acoustic signatures [27]. However, practical application in aquaculture environments faces substantial challenges. High fish densities typical of intensive farming create overlapping acoustic returns that complicate individual target discrimination [28]. Near-boundary effects from cage nets and tank walls introduce multipath reflections and clutter that degrade signal quality [29]. Species-specific calibration of acoustic responses is required, and existing models may not generalize across different sizes, orientations, and environmental conditions [30].

Additional sensor modalities including flow meters for feed consumption monitoring and oxygen sensors for metabolic demand estimation have been proposed as indirect biomass indicators [31]. While these approaches can provide useful complementary information, they lack the spatial resolution and direct measurement capabilities needed for precise biomass quantification [32]. The inherent relationship between metabolic rate and biomass is influenced by numerous confounding factors including temperature, activity level, and feeding state, limiting the reliability of metabolism-based biomass inference [33].

Transition Toward Computer Vision-Based Estimation

The emergence of computer vision as a viable biomass estimation technology reflects several converging technological trends. Advances in digital imaging sensors have delivered cameras with improved low-light performance, higher resolution, and enhanced dynamic range, all critical for underwater deployment [34]. Simultaneously, the proliferation of machine learning frameworks and training datasets has enabled the development of sophisticated image analysis algorithms capable of robust object detection and measurement under variable conditions [35]. The decreasing cost of computational hardware has made real-time video processing economically feasible even for continuous monitoring applications.

Computer vision systems offer several distinctive advantages relative to traditional and acoustic methods. Visual imaging provides direct geometric information about individual fish, enabling measurement of length, body depth, and volumetric dimensions without reliance on empirical backscatter models. The ability to process individual targets within crowded environments addresses a fundamental limitation of acoustic approaches in high-density aquaculture settings. Non-invasive operation eliminates stress responses associated with physical sampling, while continuous data acquisition captures temporal growth dynamics that discrete sampling inevitably misses. These capabilities position computer vision as a foundational technology for implementing precision aquaculture management strategies.

3. Underwater Imaging Systems and Hardware Design

Camera Configurations and Optics

The design of underwater imaging systems for biomass estimation must address the unique optical properties of aquatic environments. Water absorption and scattering attenuate light exponentially with distance, with wavelength-dependent attenuation coefficients that preferentially remove red wavelengths in most natural waters. This necessitates careful consideration of imaging distance, with most practical systems operating at ranges below 5 meters to maintain adequate image quality. Camera selection criteria include sensor resolution, frame rate, dynamic range, and spectral sensitivity characteristics matched to the illumination spectrum and target species.

Stereoscopic camera configurations employing two or more synchronized cameras have been widely adopted for biomass estimation applications requiring accurate three-dimensional measurements. By capturing simultaneous images from different viewpoints, stereo vision systems enable triangulation-based depth estimation and reconstruction of object geometry. The baseline separation between cameras, focal length, and convergence angle represent critical design parameters that determine measurement precision and working volume. Wider baselines improve depth resolution but reduce the overlapping field of view, requiring optimization based on specific deployment scenarios.

Monocular vision systems using single cameras offer advantages in terms of cost, complexity, and ease of deployment, but sacrifice direct depth measurement capability. These systems typically employ known reference objects, such as calibrated markers or cage structure elements, to establish scale and enable length estimation from apparent image size. Alternatively, assumption-based approaches may infer fish distance from position within the image frame, though such methods introduce additional uncertainty. Recent developments in monocular depth estimation using deep learning have shown promise for inferring scene geometry from single images, potentially expanding the capabilities of single-camera systems.

Lighting, Spectral, and Depth Considerations

Artificial illumination is typically required for underwater imaging in aquaculture, both to compensate for depth-related light attenuation and to enable operation during low-light periods. Light-emitting diode (LED) arrays have emerged as the dominant illumination technology, offering advantages in energy efficiency, spectral tunability, compact form factor, and thermal management relative to traditional halogen or metal halide sources. The illumination geometry, including light position relative to the camera and target, significantly influences image quality through its effects on contrast, shadowing, and specular reflections.

Spectral considerations in underwater imaging extend beyond simple illumination intensity. The selective absorption of longer wavelengths in water creates a blue-shifted color appearance that complicates color-based segmentation and species identification. Narrow-band or tunable LED illumination can be designed to maximize contrast for specific species based on their pigmentation and reflectance properties. Polarization imaging techniques have been investigated to reduce backscatter and enhance visibility in turbid water, though practical implementation complexity has limited widespread adoption.

Depth-related pressure effects impose constraints on housing design for cameras and electronics deployed in deeper aquaculture installations. Pressure-resistant housings must maintain optical quality through transparent viewports while protecting sensitive components from water ingress. Pressure compensation systems or rigid pressure vessels are employed depending on deployment depth and cost constraints. Biofouling accumulation on optical surfaces represents an additional challenge requiring periodic cleaning or implementation of antifouling coatings and mechanical cleaning mechanisms.

Fixed and Mobile Deployment Platforms

Fixed installation architectures mount cameras and illumination at permanent positions within aquaculture infrastructure, providing continuous monitoring of defined spatial regions. In cage-based systems, cameras are typically deployed on cage walls or mounted on specialized frames suspended within the cage volume. Tank-based recirculating systems may integrate cameras into viewing ports or position them above the water surface using polarization filters to eliminate surface reflections. Fixed systems offer advantages in terms of consistent viewing geometry, simplified calibration, and straightforward integration with facility power and data infrastructure.

Mobile platforms, including remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs), enable coverage of larger spatial areas and flexible positioning for optimal viewing angles. ROV-based systems are typically tethered, receiving power and transmitting video through an umbilical connection to surface control stations. While tethers complicate operation in complex cage environments with potential entanglement risks, they eliminate battery capacity constraints and enable real-time operator supervision. Autonomous platforms offer untethered operation but face challenges in underwater navigation, collision avoidance, and energy management.

Hybrid approaches combining fixed cameras with motorized pan-tilt-zoom mechanisms provide intermediate capabilities, enabling coverage of larger volumes without the complexity of fully mobile platforms. These systems can track individual fish or school movements, maintaining optimal framing and focus as targets move through the environment. The trade-offs between spatial coverage, system complexity, cost, and operational reliability drive deployment strategy selection for specific aquaculture contexts.

4. Computer Vision-Based Biomass Estimation Methods

Fish Detection and Segmentation

The foundational step in vision-based biomass estimation involves detecting and localizing individual fish within acquired images. Traditional computer vision approaches employed background subtraction, edge detection, and template matching techniques to identify fish-shaped objects. These methods relied on hand-crafted features designed to capture characteristic visual patterns such as elongated contours, specific aspect ratios, and texture properties. While computationally efficient, classical approaches struggled with robustness to variations in pose, illumination, and environmental conditions.

The advent of deep learning, particularly convolutional neural networks (CNNs), has revolutionized fish detection performance. Object detection architectures including Faster R-CNN, YOLO (You Only Look Once), and SSD (Single Shot Detector) have been successfully adapted for underwater fish detection, achieving detection rates exceeding 95% under favorable conditions. These networks learn hierarchical feature representations directly from training data, capturing complex patterns that distinguish fish from background elements and other objects. Transfer learning strategies, where networks pre-trained on large terrestrial image datasets are fine-tuned for underwater fish imagery, have proven particularly effective given the limited availability of annotated aquaculture datasets.

Segmentation methods that delineate precise pixel-level boundaries of detected fish provide enhanced geometric information for biomass estimation. Semantic segmentation architectures such as U-Net and DeepLab assign each pixel to predefined classes, enabling extraction of accurate fish silhouettes even in the presence of partial occlusion. Instance segmentation approaches including Mask R-CNN extend this capability to distinguish individual fish in crowded scenes, assigning unique identities to overlapping instances. The quality of segmentation directly impacts downstream measurement accuracy, making it a critical component of the overall estimation pipeline.

Length, Volume, and Biomass Estimation Models

Once fish are detected and segmented, geometric measurements must be extracted and converted to biomass estimates. Length measurement represents the most direct approach, typically quantifying standard length (snout to caudal peduncle), fork length (snout to fork of caudal fin), or total length (snout to caudal fin tip) depending on species conventions. In calibrated imaging systems with known camera parameters and target distance, length can be computed directly from pixel measurements using perspective projection equations. For uncalibrated systems, reference objects of known size enable scale determination.

Conversion from length to weight typically employs allometric relationships of the form W = aL^b, where W represents weight, L denotes length, and a and b are species-specific coefficients determined empirically. These length-weight relationships assume isometric or near-isometric growth, which may not hold across all size ranges or under variable nutritional conditions. More sophisticated approaches incorporate additional morphometric measurements including body depth, girth, and cross-sectional area to improve weight prediction accuracy.

Volumetric estimation methods reconstruct three-dimensional fish geometry from multi-view imagery or depth data, enabling direct calculation of body volume. Assuming constant tissue density, volume provides a more direct proxy for biomass than length-based allometry. Photogrammetric reconstruction techniques triangulate corresponding points across multiple views to generate 3D surface models from which volume is computed. Depth cameras based on structured light or time-of-flight principles can acquire volumetric data directly, though their performance in turbid underwater conditions requires careful evaluation. Machine learning approaches that directly regress weight from image features, bypassing explicit length measurement, have shown promising results in recent studies.

Multi-View and Monocular Vision Approaches

Multi-view imaging systems leverage multiple cameras to capture simultaneous perspectives of fish targets, enabling robust 3D reconstruction and measurement. Stereo vision configurations with two cameras represent the minimal multi-view setup, providing sufficient geometric constraints for depth estimation through triangulation. The epipolar geometry relating corresponding points across stereo image pairs is characterized by the fundamental matrix, which encodes the relative position and orientation of the two cameras. Feature matching algorithms identify corresponding points, and disparity (the difference in image-plane position of corresponding points) is inversely proportional to depth.

Challenges in stereo matching for underwater fish imagery include the lack of distinctive texture on smooth fish bodies, occlusion from overlapping individuals, and the dynamic nature of swimming targets. Dense stereo reconstruction algorithms that estimate disparity for every pixel can generate detailed 3D models but are computationally intensive. Sparse matching approaches that focus on distinctive keypoints are more efficient but may miss relevant geometric features. Deep learning-based stereo matching networks trained on fish-specific datasets have demonstrated improved robustness relative to traditional algorithms.

Monocular approaches extract biomass information from single-camera imagery, relying on assumptions or learned models to resolve the inherent depth ambiguity. Geometric constraints such as fish position relative to a known bottom surface or cage boundary can provide distance cues. Statistical models learned from training data with ground-truth depth information can infer probable depth from image appearance features. Recent advances in monocular depth estimation using encoder-decoder CNN architectures have achieved impressive results on general imagery, though performance on underwater fish requires domain-specific training. The reduced hardware complexity and cost of monocular systems make them attractive for large-scale deployment despite their inherent measurement uncertainty.

5. Artificial Intelligence and Learning Models

Deep Learning Architectures for Underwater Imagery

The application of deep learning to underwater fish imagery must address several domain-specific challenges that distinguish aquatic environments from terrestrial imaging contexts. Underwater images characteristically exhibit reduced contrast due to light scattering, color distortion from wavelength-dependent absorption, and spatial variability in illumination from surface waves and suspended particles. These degradations can significantly impair the performance of models trained on clear terrestrial images.

Domain adaptation techniques have been developed to improve deep learning performance on underwater imagery. Image enhancement preprocessing, including histogram equalization, white balancing, and dehazing algorithms, can normalize underwater images to more closely resemble terrestrial distributions. Generative adversarial networks (GANs) have been employed to synthesize realistic underwater training data by applying learned transformations to clear images, expanding available training datasets. Alternatively, underwater-specific augmentation strategies that simulate typical degradations during training improve model robustness to real-world variability.

Architecture selection significantly influences detection and measurement performance. Two-stage detectors such as Faster R-CNN generate region proposals followed by refined classification and localization, achieving high accuracy at the cost of computational complexity. Single-stage detectors including YOLOv5 and its successors perform detection in a single forward pass, enabling real-time processing on commodity hardware. The choice between accuracy and speed depends on application requirements, with continuous monitoring systems favoring efficient architectures while high-precision applications may justify additional computational cost.

Attention mechanisms and transformer architectures represent recent developments showing promise for underwater fish detection. Self-attention layers enable models to focus on relevant image regions while suppressing distracting background elements. Vision transformers that process images as sequences of patches have demonstrated competitive performance with CNN architectures on various computer vision tasks. Hybrid architectures combining convolutional feature extraction with transformer-based reasoning may offer advantages for complex underwater scenes with variable fish densities and environmental conditions.

Regression and Estimation Models

Direct regression from image features to biomass represents an alternative to the traditional detect-measure-convert pipeline. Deep learning regression architectures employ CNN encoders to extract high-level visual representations, followed by fully connected layers that predict continuous biomass values. These end-to-end approaches can potentially learn complex relationships between visual appearance and biomass that bypass explicit geometric measurement.

Multi-task learning frameworks that jointly optimize detection, segmentation, and regression objectives have shown improved performance relative to single-task models. By sharing learned representations across related tasks, multi-task networks leverage complementary supervision signals and improve generalization. For example, a network might simultaneously predict bounding boxes for detection, segmentation masks for precise localization, and biomass values for estimation, with all tasks benefiting from shared feature extraction.

Uncertainty quantification in biomass predictions is critical for operational decision-making but often neglected in deterministic deep learning models. Bayesian neural networks and ensemble methods can provide prediction uncertainty estimates that indicate confidence in individual measurements . Monte Carlo dropout, where random dropout is applied during inference to generate prediction distributions, offers a practical approach to uncertainty estimation without substantial architectural modifications. Calibrated uncertainty estimates enable farmers to weight automated measurements appropriately in management decisions and identify cases requiring manual verification.

Robustness to Turbidity and Occlusion

Turbidity from suspended particles, phytoplankton blooms, and sediment resuspension represents one of the most significant challenges for underwater computer vision in aquaculture. Scattering and absorption by suspended matter degrade image contrast and limit effective imaging range. Fish become partially or completely obscured when turbidity exceeds moderate levels, directly impacting detection rates.

Computational approaches to turbidity mitigation include image restoration algorithms that estimate and remove scattering effects. Dark channel prior methods exploit the statistical observation that natural images typically contain pixels with very low intensity in at least one color channel, enabling estimation of transmission maps that characterize turbidity effects. Learning-based restoration using CNNs trained on paired clear and degraded images can achieve superior performance to physics-based methods. However, restoration quality degrades in severe turbidity where information loss is fundamental rather than merely obscured. Alternative strategies focus on learning robust features directly from turbid imagery rather than attempting restoration. Training datasets that include turbid examples teach models to extract relevant information despite degraded visibility. Multi-spectral imaging with wavelengths selected to minimize scattering for specific water types can improve penetration and contrast. Active imaging approaches including range-gated systems that time illumination pulses to reject scattered light show promise but increase system complexity.

Occlusion from overlapping fish in dense populations creates ambiguity in individual detection and measurement. Tracking algorithms that maintain identities across frames can leverage temporal information to infer complete fish geometry even when individuals are partially occluded in single frames. Multi-view systems can observe occluded regions from alternative viewpoints where they may be visible. Deep learning segmentation architectures with recurrent connections can learn to infer occluded portions of fish bodies based on visible portions and learned shape priors.

6. Applications in Aquafarming Systems

Cage-Based Aquaculture

Marine cage aquaculture, where fish are grown in netted enclosures in coastal waters, represents a dominant production system for species including salmon, sea bass, and sea bream. Cage volumes can exceed 10,000 cubic meters with stocking densities approaching carrying capacity, creating challenging environments for biomass monitoring. Computer vision systems deployed in cages must withstand dynamic conditions including currents, waves, and marine fouling while providing coverage of volumes too large for comprehensive manual sampling.

Fixed camera installations on cage walls provide continuous monitoring of specific cage sections, capturing fish that swim through the field of view. Statistical sampling theory is applied to extrapolate observations from monitored volumes to total cage biomass, accounting for non-uniform fish distribution and behavioral patterns. Multiple camera positions distributed throughout the cage improve coverage and reduce extrapolation uncertainty. Integration with feeding systems enables targeted monitoring during feeding periods when fish are most active and evenly distributed.

The ability to track growth rates at cage-level resolution supports optimized feeding strategies that account for biomass accumulation dynamics. Traditional feeding tables based on expected growth rates and temperature may under- or over-deliver feed if actual growth deviates from predictions. Vision-based biomass monitoring enables adaptive feeding models that adjust daily rations in response to measured growth, improving feed conversion efficiency and reducing waste . Early detection of reduced growth rates can indicate disease onset or environmental stress, triggering investigation and intervention before widespread impacts occur.

Recirculating Aquaculture Systems (RAS)

Recirculating aquaculture systems maintain fish in controlled indoor environments where water is continuously filtered and reused, enabling intensive production with minimal water consumption. RAS tanks typically range from a few cubic meters to hundreds of cubic meters, with viewing conditions that can be more controlled than marine cages. The enclosed nature of RAS facilities facilitates camera installation and maintenance while enabling integration with comprehensive sensor networks for water quality and system performance monitoring.

Computer vision for RAS biomass estimation benefits from relatively stable imaging conditions including consistent artificial lighting, reduced turbidity from filtration systems, and fixed tank geometry. However, high stocking densities typical of intensive RAS create severe occlusion challenges, with individuals frequently overlapping in camera views. Temporal tracking across video sequences helps resolve individual identities and enables measurement of temporarily unoccluded fish.

The integration of vision-based biomass data with other RAS operational parameters enables sophisticated control strategies. Feeding automation can respond to both measured biomass and observed feeding behavior, delivering feed in response to demonstrated appetite rather than predetermined schedules. Oxygen supplementation and biofilter sizing can be dynamically adjusted based on actual biomass rather than conservative design assumptions. The closed-loop nature of RAS makes it an ideal testbed for precision aquaculture technologies, with vision systems serving as key information sources for decision support.

Indoor and Offshore Aquafarming

Land-based indoor aquaculture facilities represent an emerging production model that combines RAS technology with industrial-scale production in controlled warehouse environments. These facilities enable production near consumption centers, reducing transportation costs and improving product freshness. The highly controlled nature of indoor environments allows deployment of sophisticated monitoring systems including multi-camera vision networks providing comprehensive spatial coverage.

Computer vision in indoor aquaculture can leverage controlled artificial lighting optimized for both fish welfare and imaging quality. Specialized lighting designs including near-infrared illumination enable passive imaging that minimizes behavioral disturbance while maintaining image quality for automated analysis. The availability of facility infrastructure for power, networking, and environmental control simplifies system integration relative to remote offshore deployments.

Offshore aquaculture in exposed oceanic environments presents extreme challenges for vision system deployment. Deep-water cage systems may operate in areas with significant currents, wave action, and limited accessibility for maintenance. Ruggedized camera housings designed for subsea oil and gas applications have been adapted for offshore aquaculture monitoring. Autonomous surface vessels equipped with underwater imaging systems can provide periodic monitoring without requiring dedicated infrastructure at cage sites. The economic scale of offshore operations justifies investment in advanced monitoring technologies that improve operational efficiency and reduce risk.

7. Challenges and Future Perspectives

Environmental Variability and Data Limitations

The diversity of aquaculture species, production systems, and environmental conditions creates significant challenges for developing generalizable computer vision solutions. Visual appearance varies dramatically across species, from streamlined pelagic fish to benthic species with cryptic coloration. Size ranges span orders of magnitude from larval stages measuring millimeters to market-size individuals exceeding one meter. Each species and life stage may require customized imaging configurations and analysis algorithms.

Training deep learning models requires large annotated datasets that are expensive and time-consuming to generate. Manual annotation of fish in crowded underwater imagery is particularly challenging, requiring expert knowledge to distinguish individuals and assign labels consistently. The lack of large-scale public datasets for aquaculture applications limits research progress and forces practitioners to develop proprietary datasets. Collaborative efforts to create standardized benchmark datasets with diverse species and conditions would accelerate the field.

Environmental variability within and across production facilities introduces additional challenges for model generalization . Water quality parameters including turbidity, color, and suspended particle composition vary seasonally and geographically. Lighting conditions change with depth, time of day, and weather in outdoor systems. Models trained on data from specific conditions may fail when deployed in different environments, requiring either comprehensive training datasets spanning operational variability or adaptation techniques that enable models to generalize .

Model Generalization and Scalability

Achieving robust performance across the full range of operational conditions encountered in commercial aquaculture requires models that generalize beyond their training distributions. Domain randomization during training, where models are exposed to artificially varied conditions including lighting, turbidity, and background appearance, can improve robustness. Few-shot learning approaches that enable rapid adaptation to new species or environments with minimal labeled data show promise for reducing data requirements.

Scalability challenges arise when deploying vision systems across multiple facilities or expanding from research prototypes to production-scale implementations. Cloud-based processing architectures enable centralized model deployment and updates but require reliable network connectivity and introduce latency. Edge computing approaches that perform inference locally on cameras or dedicated processors reduce bandwidth and latency but complicate model updates and version management. Hybrid architectures that perform initial processing locally with cloud-based aggregation and refinement may offer practical compromises.

The computational cost of processing continuous video streams from multiple cameras can become prohibitive at scale. Efficient inference strategies including model compression, quantization, and knowledge distillation reduce computational requirements while maintaining acceptable accuracy.

Adaptive frame rate processing that triggers detailed analysis only when fish are present conserves resources relative to blanket processing of all frames. The trade-off between monitoring coverage, temporal resolution, and computational cost must be optimized for specific operational requirements and budget constraints.

Integration with Digital Twins and Farm Management Systems

The true value of vision-based biomass monitoring is realized through integration into comprehensive farm management systems that support data-driven decision-making. Digital twin concepts, where real-time sensor data feeds computational models that mirror and predict physical system behavior, represent a promising framework for precision aquaculture. Vision-derived biomass estimates combine with water quality sensors, feeding records, and environmental forecasts to enable predictive modeling of growth, health, and production outcomes.

Machine learning models trained on historical production data can identify patterns linking biomass growth rates to management interventions and environmental conditions. These insights enable optimization of feeding strategies, stocking densities, and harvest timing to maximize economic performance while maintaining sustainability. Anomaly detection algorithms that identify deviations from expected growth patterns can trigger alerts for investigation, enabling early intervention for disease or environmental problems.

Interoperability standards and data exchange protocols are needed to enable integration across heterogeneous systems from multiple vendors. The development of open APIs and standard data formats for aquaculture sensor data would facilitate the creation of integrated management platforms. Blockchain-based traceability systems that record production data could incorporate vision-based biomass measurements as verifiable inputs supporting product certification and consumer transparency.

The ethical considerations surrounding data ownership, privacy, and competitive advantage in shared data platforms require careful consideration. Farmers may be reluctant to share detailed production data with third parties or competitors even when aggregation and anonymization could enable valuable industry-wide insights. Establishing trust and demonstrating value through pilot implementations will be critical for achieving widespread adoption of integrated digital aquaculture platforms.

8. Tables

Table 1: Comparison of traditional, acoustic, and computer vision-based biomass estimation methods

Method	Measurement Principle	Advantages	Limitations	Typical Accuracy
Manual sampling	Physical capture, weighing	Direct measurement, established protocols	Labor-intensive, invasive, discrete data	±5-10% (sampling error)
Net sampling	Population subset extrapolation	Simple implementation	Stress responses, low temporal resolution	±10-20% (small samples)
Echo sounder	Acoustic backscatter intensity	Non-invasive, large volume coverage	Limited resolution in dense populations	±15-30% (calibration dependent)
Split-beam sonar	Target strength analysis	Individual target tracking	Complex near-boundary effects	±20-40% (environmental sensitivity)
Stereoscopic vision	3D reconstruction from images	Direct geometry, continuous monitoring	Hardware complexity, calibration requirements	±3-8% (favorable conditions)
Monocular vision	Single-camera measurement with scale references	Lower cost, simpler deployment	Depth ambiguity, reduced precision	±5-15% (with calibration)
Deep learning regression	End-to-end image-to-biomass prediction	Learns complex patterns, automated	Requires large training datasets	±4-10% (species-specific models)

Table 2: Advantages, limitations, and technical challenges of underwater vision-based biomass estimation systems

Aspect	Advantages	Limitations	Technical Challenges
Data acquisition	Continuous non-invasive monitoring, high temporal resolution, no stress to fish	Limited imaging range (typically <5m), environmental sensitivity	Underwater optical degradation, wavelength-dependent attenuation, backscatter from particles
Fish detection	Automated processing, handles high densities, individual tracking possible	Occlusion in crowded populations, similar appearance to background	Robust segmentation in turbid water, distinguishing overlapping individuals, real-time processing requirements
Measurement accuracy	Direct geometric measurement, 3D reconstruction capability, species-agnostic geometry	Requires calibration, perspective distortions, partial visibility	Accurate depth estimation, handling fish orientation variability, calibration maintenance
Environmental robustness	Operates in diverse conditions with appropriate design	Performance degrades with turbidity, variable illumination affects consistency	Turbidity compensation algorithms, adaptive illumination control, biofouling prevention on optics
Computational requirements	Leverages modern GPU acceleration, edge computing feasible	High processing load for multi-camera systems, bandwidth for video streaming	Efficient deep learning inference, real-time analysis, model optimization for deployment
System integration	Compatible with existing infrastructure, scalable deployment	Requires power and networking at deployment sites, maintenance access needed	Weatherproofing and pressure resistance, reliable underwater connectors, remote diagnostics
Cost and scalability	Decreasing hardware costs, cloud processing reduces local requirements	Initial investment for quality systems, species-specific model training	Multi-site deployment management, model generalization across facilities, long-term operational costs

9. Figure

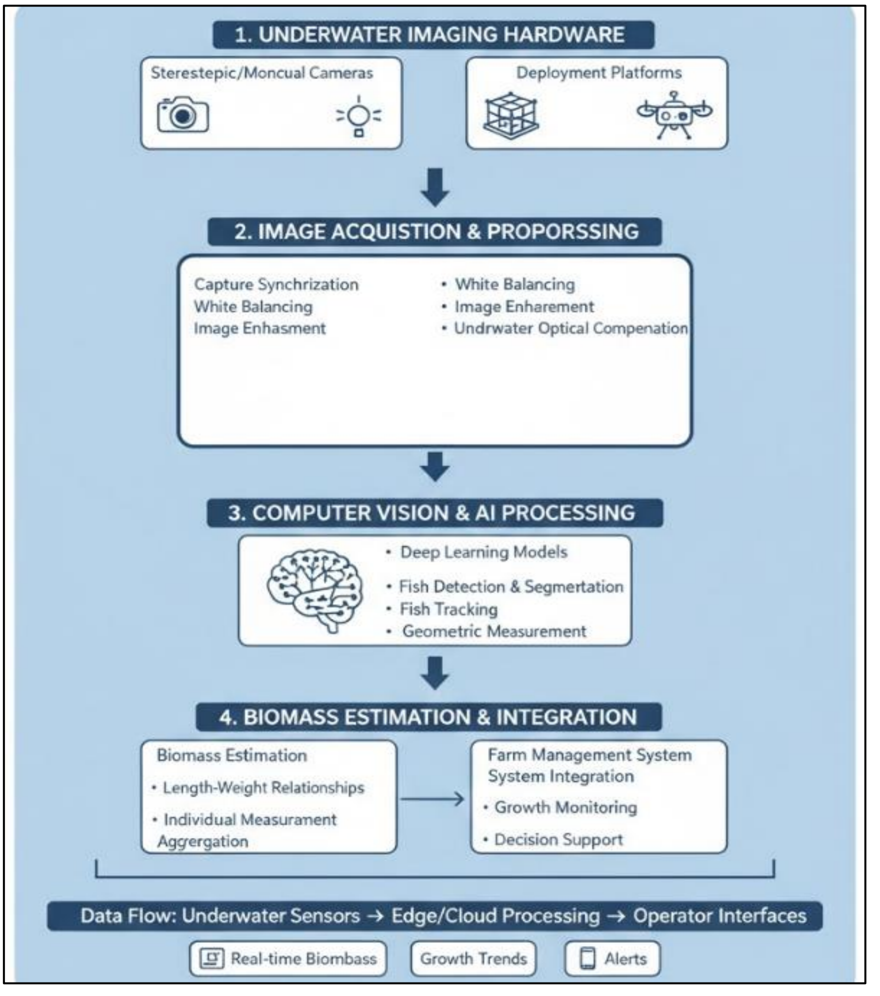

Fig 1: System architecture of computer vision-based underwater biomass estimation in aquafarming

10. Conclusion

Computer vision-based technologies for underwater biomass estimation have emerged as transformative tools enabling precision aquaculture management. The convergence of advanced imaging sensors, deep learning algorithms, and decreasing computational costs has made continuous, non-invasive monitoring of fish populations technically and economically viable. Stereoscopic and monocular imaging systems deployed in diverse aquaculture environments provide real-time data on fish size distributions and growth dynamics, eliminating the stress and operational constraints of traditional manual sampling.

Deep learning architectures adapted for underwater conditions achieve robust detection and measurement performance despite challenges from turbidity, occlusion, and variable illumination. The progression from hand-crafted feature extraction to learned representations has dramatically improved system capabilities, while ongoing research addresses remaining challenges in model generalization and environmental robustness. Applications across cage-based marine systems, land-based recirculating facilities, and emerging offshore operations demonstrate the versatility of vision-based approaches.

The contribution of automated biomass monitoring to sustainable aquaculture intensification is substantial. Optimized feeding strategies enabled by accurate biomass data reduce feed waste and minimize environmental impacts from nutrient loading. Early detection of growth anomalies supports proactive health management, reducing disease-related losses and antimicrobial use. The operational efficiency gains from automated monitoring improve economic viability while reducing the physical labor requirements of traditional aquaculture.

Future developments will focus on several key areas. Improved algorithms that maintain performance across species, life stages, and environmental conditions will reduce the need for system-specific customization. Integration with broader digital twin frameworks will enable holistic farm management that considers biomass monitoring alongside water quality, feeding, health, and market factors. Autonomous underwater platforms will expand spatial coverage and enable monitoring of large offshore installations. Collaborative dataset development and model sharing within the research community will accelerate progress toward robust, generalizable solutions.

The trajectory toward data-driven precision aquaculture is clear, with computer vision-based biomass estimation serving as a foundational enabling technology. As systems mature from research demonstrations to commercial deployments, the aquaculture industry will increasingly rely on automated monitoring to meet growing global demand for seafood while maintaining environmental sustainability and animal welfare standards. Continued research, commercial development, and practical deployment will cement the role of intelligent underwater imaging as an indispensable component of modern aquafarming operations.

11. References

1. Balaban MO, Chombeau M, Cırban D, Gümüş B. Prediction of the weight of Alaskan pollock using image analysis. J Food Sci. 2010;75(8):E552–E556.
2. Føre M, Frank K, Norton T, Svendsen E, Alfredsen JA, Dempster T, *et al*. Precision fish farming: A new framework to improve production in aquaculture. Biosyst Eng. 2018;173:176–193.
3. Atoum Y, Srivastava S, Liu X. Automatic feeding control for dense aquaculture fish tanks. IEEE Signal Process Mag. 2015;32(5):103–111.
4. Muñoz-Benavent P, Andreu-García G, Valiente-González JM, Atienza-Vanacloig V, Puig-Pons V, Espinosa V. Enhanced fish bending model for automatic tuna sizing using computer vision. Comput Electron Agric. 2018;150:52–61.
5. Godø OR, Pennington M, Vølstad JH. Effect of tow duration on length composition of trawl catches. Fish Res. 1990;9(2):165–179.
6. Iversen M, Finstad B, McKinley RS, Eliassen RA. The efficacy of metomidate, clove oil, Aqui-S and Benzoak as anaesthetics in Atlantic salmon (Salmo salar L.) smolts, and their potential stress-reducing capacity. Aquaculture. 2003;221(1–4):549–566.
7. Zion B. The use of computer vision technologies in aquaculture - A review. Comput Electron Agric. 2012;88:125–132.
8. Klebert P, Lader P, Gansel L, Oppedal F. Hydrodynamic interactions on net panel and aquaculture fish cages: A review. Ocean Eng. 2013;58:260–274.
9. Lines JA, Tillett RD, Ross LG, Chan D, Hockaday S, McFarlane NJB. An automatic image-based system for estimating the mass of free-swimming fish. Comput Electron Agric. 2001;31(2):151–168.
10. Costa C, Loy A, Cataudella S, Davis D, Scardi M. Extracting fish size using dual underwater cameras. Aquacult Eng. 2006;35(3):218–227.
11. Petrell RJ, Shi X, Ward RK, Naiberg A, Savage CR. Determining fish size and swimming speed in cages and tanks using simple video techniques. Aquacult Eng. 1997;16(1–2):63–84.
12. Føre M, Alfredsen JA, Gronungsater A. Development of two telemetry-based systems for monitoring the feeding behaviour of Atlantic salmon (Salmo salar L.) in aquaculture sea-cages. Comput Electron Agric. 2011;76(2):240–251.
13. Li D, Hao Y, Duan Y. Nonintrusive methods for biomass estimation in aquaculture with emphasis on fish: a review. Rev Aquac. 2020;12(3):1390–1411.
14. Rauf HT, Lali MI, Zahoor S, Shah SZH, Rehman AU, Bukhari SAC. Visual features based automated identification of fish species using deep convolutional neural networks. Comput Electron Agric. 2019;167:105075.
15. Salman A, Jalal A, Shafait F, Mian A, Shortis M, Seager J, *et al*. Fish species classification in unconstrained underwater environments based on deep learning. Limnol Oceanogr Methods. 2016;14(9):570–585.
16. Marini S, Fanelli E, Sbragaglia V, Azzurro E, Del Rio Fernandez J, Aguzzi J. Tracking fish abundance by underwater image recognition. Sci Rep. 2018;8:13748.
17. Miranda A, Marçalo A, Marques TA. Weighing the cost: Sample size requirements for precision aquaculture weight estimation. Aquacult Eng. 2020;90:102095.
18. Buentello A, Jirsa D, Barrows FT, Drawbridge M. Minimizing handling stress in white seabass: efficacy of anesthetic agents and comparison of blood stress indicators. N Am J Aquac. 2016;78(1):48–53.
19. Saberioon M, Gholizadeh A, Cisar P, Pautsina A, Urban J. Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues. Rev Aquac. 2017;9(4):369–387.
20. Froese R. Cube law, condition factor and weight-length relationships: history, meta-analysis and recommendations. J Appl Ichthyol. 2006;22(4):241–253.
21. Sánchez-Carnero N, Rodríguez-Pérez D, Couñago E, Freire J. Species identification and biomass estimation of fish schools in the Aegean Sea using remotely operated vehicles. Hydrobiologia. 2016;772(1):231–245.
22. Handegard NO, Michalsen K, Tjøstheim D. Avoidance behaviour in cod (Gadus morhua) to a bottom-trawling vessel. Aquat Living Resour. 2003;16(3):265–270.
23. Acerete L, Balasch JC, Espinosa E, Josa A, Tort L. Physiological responses in Eurasian perch (Perca fluviatilis, L.) subjected to stress by transport and handling. Aquaculture. 2004;237(1–4):167–178.
24. Barton BA. Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids. Integr Comp Biol. 2002;42(3):517–525.
25. Balk H, Lindem T. Improved fish detection in data from split-beam sonar. Aquat Living Resour. 2000;13(5):297–303.
26. Simmonds J, MacLennan DN. Fisheries Acoustics: Theory and Practice. 2nd ed. Oxford: Blackwell Science; 2005.
27. Ona E. An expanded target-strength relationship for herring. ICES J Mar Sci. 2003;60(3):493–499.
28. Manik HM, Furusawa M, Amakasu K. Observation of fish school dynamics using an ultrasonic multi-beam sonar and an echo-sounder. Jpn J Appl Phys. 2006;45(5S):4782.
29. Fässler SMM, Fernandes PG, Semple SI, Brierley AS. Depth-dependent swimbladder compression in herring Clupea harengus observed using magnetic resonance imaging. J Fish Biol. 2009;74(2):296–303.
30. Korneliussen RJ, Heggelund Y, Eliassen IK, Johansen GO. Acoustic species identification of schooling fish. ICES J Mar Sci. 2009;66(6):1111–1118.
31. Alanärä A, Brännäs E. Dominance in demand-feeding behaviour in Arctic charr and rainbow trout: the effect of stocking density. J Fish Biol. 1996;48(2):242–254.
32. Noble C, Gismervik K, Iversen MH, Kolarevic J, Nilsson J, Stien LH, *et al*. Welfare Indicators for Farmed Atlantic Salmon: Tools for Assessing Fish Welfare. Tromsø: Nofima; 2018.
33. Chabot D, Steffensen JF, Farrell AP. The determination of standard metabolic rate in fishes. J Fish Biol. 2016;88(1):81–121.
34. Jaffe JS. Computer modeling and the design of optimal underwater imaging systems. IEEE J Ocean Eng. 1990;15(2):101–111.
35. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84–90.