

Article

A Training-Free Selective-Processing Workflow for In Situ Marine Particle Fields Using Parallel Phase-Shifting Digital Holography

Xinran Liu ^{1,2}  and Haoran Meng ^{1,*}

¹ Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

² University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: menghaoran@ciomp.ac.cn

Abstract

In situ marine particle-field observation by parallel phase-shifting digital holography (PPSDH) produces long image sequences under real deployment conditions, but exhaustive full-frame reconstruction and segmentation are computationally expensive when many frames are low contrast and particle-like targets occupy sparse regions. This paper presents a training-free two-stage selective-processing workflow for a 9521-frame coastal South China Sea PPSDH campaign. Stage 1 uses an amplitude-derived contrast metric as a campaign-specific pruning rule to form a retained-frame pool, and Stage 2 combines coarse reconstruction, candidate filtering, valid-field gating, and ROI merging for ROI-restricted reconstruction and segmentation. Stage 1 retained 6970 frames, corresponding to 73.2% of the full sequence. On a balanced 120-frame benchmark, Stage 2 achieved a spatial-support reduction ratio of $49.9\% \pm 12.0\%$, and the complete workflow provided a 5.66-fold end-to-end speedup relative to a matched full-frame baseline. The efficiency gain was accompanied by a measurable fidelity cost, with a baseline-matched correspondence rate of 0.612 and a count-based yield gap of 0.287, mainly associated with small or weak targets within the selected ROI support. These results show that the proposed workflow can support computation-aware review of real marine PPSDH particle fields by efficiently prioritizing informative frames and particle-like regions for downstream visual assessment.

Keywords: digital holographic microscopy; in situ marine imaging; marine particles; parallel phase-shifting digital holography; computational efficiency; frame-level screening; selective processing; region of interest; training-free workflow



Academic Editor: Alen Soldo

Received: 27 April 2026

Revised: 14 May 2026

Accepted: 29 May 2026

Published: 31 May 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

1. Introduction

In situ optical imaging provides a direct way to record suspended material in seawater, including detrital aggregates, mineral particles, bubbles, marine snow fragments, and plankton-sized biological objects. Digital holographic microscopy (DHM) is attractive for marine particle-field observation because a single hologram can encode a depth-resolved water volume that can be numerically reconstructed after acquisition. Early underwater holography demonstrated the ability to record particle-containing water volumes rather than only a single focal plane [1–3]. Subsequent digital and submersible holographic systems, including in-line marine particulate sensors, HoloCam, eHoloCam, and more recent underwater holographic instruments, extended this capability to in situ imaging of marine particles and micrometre-scale targets [4–9]. These developments make DHM suitable for

long-duration particle-field acquisition, but they also create a practical processing bottleneck: each recorded hologram may require numerical propagation over many candidate depths before particle-level measurements can be obtained.

Parallel phase-shifting digital holography (PPSDH) provides a useful route for field holography because four phase-shifted intensity samples are recorded within a single exposure and combined to recover the complex optical field analytically. This reduces the ambiguity associated with conventional in-line holography and provides a complex-valued input for subsequent screening and reconstruction. The fibre-coupled PPSDH instrument used in this study acquired 9521 raw frames during a coastal South China Sea campaign in September 2021, corresponding to approximately 95 GB of 16-bit data over about 40 min. Although such acquisition rates are practical for field deployment, exhaustive full-frame reconstruction and segmentation remain computationally expensive, especially when many frames have low contrast and particle-like targets occupy only sparse regions of the illuminated field.

This study addresses the processing problem rather than the taxonomic interpretation problem. Two characteristics of real marine PPSDH sequences motivate selective processing. First, low-contrast frames affected by illumination variation, bubbles, sediment, or window artefacts are less suitable for rapid high-confidence processing, even though some may still contain weak recoverable structures after offline reconstruction. Second, in retained frames, useful particle-like content is spatially sparse, so full-frame propagation and segmentation spend substantial time on regions without visible particle content. A practical workflow should therefore reduce unnecessary reconstruction while preserving interpretable particle-field information for later inspection.

Previous studies have addressed related issues from different perspectives, including post-reconstruction preprocessing of holographic particle images [10], sparse-target processing in underwater digital in-line holography [11], quantitative benchmarking of holographic size and concentration measurements [12], platform-integrated marine holographic monitoring [13,14], and learned detection or reconstruction methods for holographic data [15–17]. These studies indicate that scalable interpretation is a central challenge in aquatic holography. However, reported performance values are difficult to compare directly because of differences in imaging geometry, reconstruction strategy, target definition, hardware, and evaluation protocol. For this reason, the present study uses a matched full-frame baseline under the same implementation and evaluates retained-frame fraction, spatial-support reduction, selective/full-frame correspondence, particle-yield gap, and end-to-end processing speedup.

This paper presents a training-free two-stage selective-processing workflow for real marine PPSDH particle-field data. Stage 1 applies an amplitude-derived contrast metric to the demultiplexed complex field and defines a campaign-specific low-contrast pruning rule. Stage 2 combines coarse reconstruction, candidate filtering, valid-field gating, and spatial merging to form compact regions of interest for ROI-restricted reconstruction and segmentation. The workflow is evaluated on the full 9521-frame campaign and on a stratified 120-frame benchmark using a matched full-frame baseline.

The main contributions of this work are threefold. First, it defines a transparent, training-free low-contrast pruning rule for real PPSDH field sequences. Second, it introduces a merged-ROI processing strategy that reduces spatial support before downstream reconstruction and segmentation. Third, it quantifies the efficiency–fidelity trade-off against a matched full-frame baseline, including retained-frame precision, spatial-support reduction, processing speedup, particle correspondence, and yield gap. The results provide an interpretable preprocessing route for prioritising informative frames and particle-like regions in large-scale in situ holographic datasets.

The remainder of this paper is organised as follows. Section 2 summarises the marine PPSDH instrument and in situ dataset. Section 3 defines the two-stage workflow, notation, evaluation metrics, and statistical analyses. Section 4 reports the Stage 1 operating outcome, visual audit, Stage 2 spatial-support reduction, timing benchmark, and downstream particle summaries. Section 5 discusses the engineering rationale, fidelity costs, limitations, and campaign-specific scope of the evidence. Section 6 concludes the paper.

2. Marine PPSDH Instrument and In Situ Dataset

2.1. Parallel Phase-Shifting Principle via Polarisation Multiplexing

The PPSDH instrument uses a fibre-coupled dual-arm optical layout in which the laser output is split into object and reference beams by a fibre-based polarisation beam splitter. The two beams are delivered through polarisation-maintaining fibres, emitted by fibre collimators, and recombined before the polarisation-imaging sensor. After circular polarisation conversion, the micro-polariser array records four phase-shifted interference samples in a single exposure. Thus, the optical implementation is compact and Mach-Zehnder-like, while the recorded hologram corresponds to near-coaxial object-reference interference at the detector plane.

The polarisation camera contains a repeating 2×2 micro-polariser super-pixel with analyser angles of 0° , 45° , 90° , and 135° . After demultiplexing, the four phase-shifted intensity samples I_0 – I_3 are combined to recover the complex field

$$X(x, y) = \frac{1}{4}[(I_2 - I_0) + i(I_3 - I_1)], \quad (1)$$

which is the common input to Stage 1 screening and Stage 2 ROI-restricted processing. The demultiplexed field has 1224×1024 complex samples at an effective sampling interval of $6.90 \mu\text{m}$, derived from the 2448×2048 camera with a $3.45 \mu\text{m}$ native pixel pitch. A more detailed description of the single-exposure polarisation-multiplexing principle is provided in Supplementary Note S2. Figure 1 summarises the relationship between the micro-polariser sampling pattern and complex-field recovery.

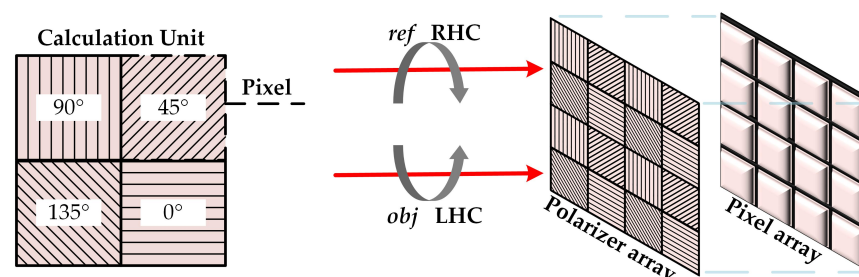


Figure 1. Principle of single-exposure parallel phase-shifting by polarisation multiplexing. The 2×2 micro-polariser super-pixel records four phase-shifted intensity samples, I_0 – I_3 , from which the demultiplexed complex field X is recovered and used as the common input to the selective-processing workflow.

2.2. Optical Module Design

Conventional parallel phase-shifting digital holography often uses free-space Mach-Zehnder layouts [18]. Our field instrument preserves the object-reference separation principle in a fibre-coupled form while producing a near-coaxial interference field at the sensor. This distinction separates the optical implementation from the recording geometry. For in situ marine use, the main constraint is practical: bulky free-space polarisation optics are difficult to package and stabilise under vibration and thermal drift. Fibre coupling reduces footprint and improves stability.

Figure 2 shows the resulting optical layout. A single fibre-pigtailed semiconductor laser emitting at $\lambda = 785$ nm drives the system. The randomly polarised output enters a polarisation-maintaining fibre-based polarising beam splitter. The splitter produces two orthogonal linearly polarised outputs. The *s*-polarised arm serves as the object beam and is routed through the sample volume, while the *p*-polarised arm serves as the reference beam and bypasses the sample. Both arms terminate in adjustable fibre collimators and are recombined by an unpolarised beam-splitting prism. A quarter-wave plate at 45° converts the combined beam into left- and right-handed circularly polarised waves. This state satisfies the interference condition of the micro-polariser array (Section 2.1). The polarisation-imaging camera uses a Sony IMX250MZR complementary metal–oxide–semiconductor (CMOS) image sensor (Sony Semiconductor Solutions Corporation, Atsugi, Japan) with 2448×2048 native pixels at a $3.45 \mu\text{m}$ native pitch. Demultiplexing via Equation (1) reduces this to 1224×1024 complex-field samples at an equivalent pitch of $s = 2 \times 3.45 = 6.90 \mu\text{m}$.

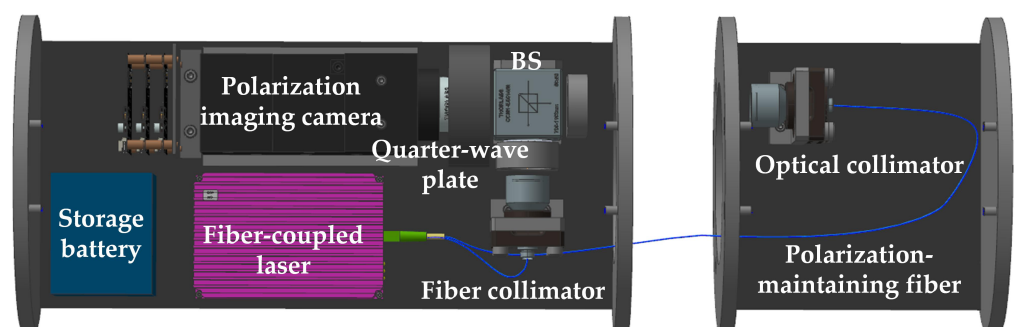


Figure 2. Fibre-coupled PPSDH instrument layout used for the in situ campaign. The object and reference arms are housed in two titanium cylinders and recombined before the polarisation-imaging camera, producing the demultiplexed complex field used by the selective-processing workflow.

Figure 3 documents the marine-engineering context: the assembled submersible PPSDH instrument, a water-immersion sealing test, and the sea-trial deployment.

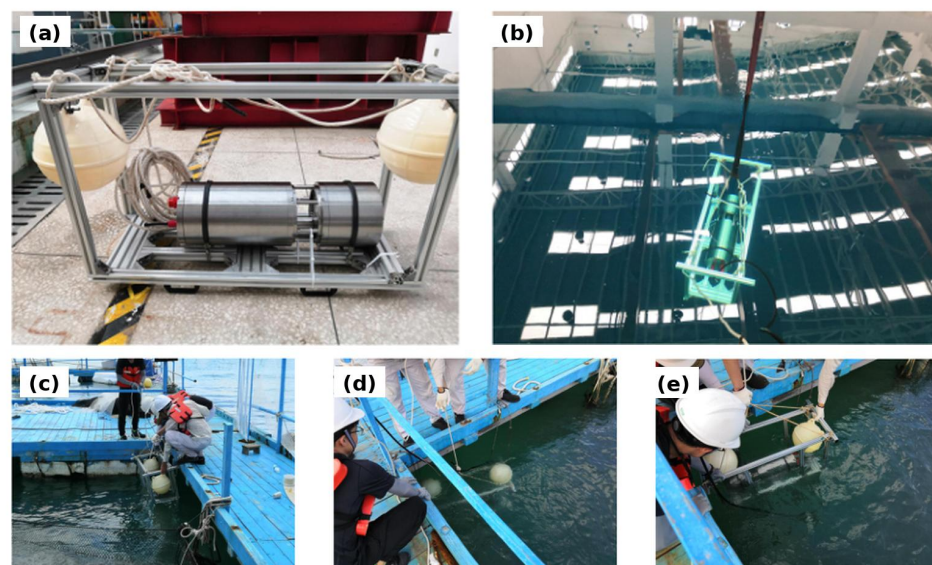


Figure 3. Submersible PPSDH system and field-deployment context: (a) assembled instrument mounted on the support frame; (b) water-immersion sealing test; (c) preparation before deployment from the sea-trial platform; (d) lowering of the instrument into seawater; and (e) instrument during initial deployment. These photographs document the engineering and deployment context of the dataset analysed in this study.

2.3. Practical Resolution and Exposure-Time Bound

The sample is imaged without magnifying optics, so the practical lateral resolution is governed by the finite numerical aperture subtended by the detector, the sample-to-detector distance, and underwater optical disturbance. A desktop 1951 USAF resolution target (Thorlabs, Inc., Newton, NJ, USA) was used to verify that the system resolved $13.92\text{ }\mu\text{m}$ line widths in air and approximately $17.54\text{ }\mu\text{m}$ in water. We therefore use approximately $20\text{ }\mu\text{m}$ as the practical underwater lower-size scale for downstream segmentation in this campaign. The derivation, USAF validation image, and air/water comparison are provided in Supplementary Note S3 and Figure S2.

The exposure time was 2 ms. A simple displacement estimate, $\Delta x = vt_{\text{exp}}$, gives a motion blur of about $2\text{--}6\text{ }\mu\text{m}$ for particles moving at $1\text{--}3\text{ mm s}^{-1}$, which is below one demultiplexed pixel ($6.90\text{ }\mu\text{m}$) and below the adopted practical underwater resolution scale. Faster organisms or turbulence-driven particles may still blur; no synchronized velocity record was available. This estimate only bounds the exposure setting used for the present campaign.

2.4. Deployment Context and Dataset

The instrument is enclosed in two coaxial titanium pressure cylinders of 200 mm outer diameter and 600 mm combined length, joined by six titanium pins with set-screw locking that preserve optical-axis collimation under pressure cycling. The optical beams cross the intervening seawater through two 10 mm-thick sapphire windows whose clear apertures define the useful imaging volume. A watertight rear connector carries the Gigabit Ethernet link used for camera control and data retrieval. The instrument runs on an internal battery for more than ten hours without external power. These hardware details define the field context of the campaign; the primary analytical evidence begins with the two-stage workflow in Section 3.

The dataset analysed in this paper was acquired during a September 2021 coastal deployment in the South China Sea off the coast of Sanya, Hainan Province, China (18.188° N , 109.483° E) at an operating depth of approximately 5 m. No synchronized measurements of water temperature, salinity, chlorophyll-*a*, turbidity, or community-composition information were recorded with this engineering deployment. The dataset is therefore used to evaluate processing behaviour on real marine PPSDH frames, not to infer local community structure. To minimise instrument self-induced turbulence we captured data only during the stable suspended phase and excluded the descent and ascent transients from the recording. A total of $N_{\text{total}} = 9521$ raw frames were acquired and saved across two storage partitions. Each raw frame is a single-exposure 2448×2048 -pixel capture of the polarisation-multiplexed hologram, stored as 16-bit Tagged Image File Format (TIFF). The raw storage footprint is approximately 10 MB per frame and 95 GB for the full campaign. After demultiplexing (Equation (1)), each frame yields a 1224×1024 complex-amplitude-field array at an equivalent pitch of $s = 6.90\text{ }\mu\text{m}$. All workflow operations in Section 3 are defined on this format.

A summary of the instrument and campaign parameters used throughout this paper is given in Table 1. Section 2 establishes the acquisition context, data format, and resolution floor used in the analysis below.

Table 1. Optical, deployment, and dataset parameters defining the PPSDH campaign analysed in this study.

Parameter	Value
Wavelength λ	785 nm
Laser nominal output power	20 mW (specified range 10–50 mW)
Laser linewidth	10–20 MHz
Camera	Sony IMX250MZR polarisation-imaging CMOS
Native pixel array	2448 $\times$ 2048
Native pixel pitch	3.45 μm
Complex-field pitch after demultiplexing s	6.90 μm
Complex-field array size	1224 $\times$ 1024
Maximum frame rate	50 frames per second (fps)
Integration time (this campaign)	2 ms
Effective acquisition frame rate (this campaign)	~ 4 Hz
Pressure housing outer diameter $\times$ length	200 $\times$ 600 mm
Window material, thickness	sapphire, 10 mm
Lateral resolution in air (desktop)	~ 14 μm
Practical underwater lateral resolution	~ 20 μm
Operating sample-to-detector distance range	151–230 mm
Campaign date	September 2021
Campaign location	South China Sea, off Sanya, Hainan
Coordinates	18.188° N, 109.483° E
Operating depth	~ 5 m
Total raw frames acquired N_{total}	9521
Cumulative acquisition duration	~ 40 min
Raw storage footprint	~ 95 GB

3. Two-Stage Selective-Processing Workflow

3.1. Workflow Overview and Design Rationale

Processing long in situ recordings at 2448 $\times$ 2048 px and up to 50 fps is dominated by two avoidable costs: low-contrast or unstable frames, and sparse particle-like targets within otherwise useful frames. Full-frame angular-spectrum propagation and autofocus search over $\mathcal{O}(10^2)$ candidate depths can therefore spend most of the compute budget on regions without visible particle content.

We organise the workflow into two rule-based stages that operate at different spatial scales (Figure 4):

- Stage 1 (frame-level low-contrast pruning). A scalar statistic computed directly on the demultiplexed complex field X forms a retained-frame pool before reconstruction. This stage is training-free, requires no labelled training data, and is parallel across frames. It is a low-contrast pruning rule for fast review, not a detector that proves target absence in rejected frames.
- Stage 2 (region-level selective processing). For each retained frame, a coarse single-plane reconstruction is used to locate candidate particle-like targets. Candidate regions are filtered, gated by the valid illuminated field, and grouped into padded rectangular merged ROIs. These ROIs are then used as the spatial supports for downstream reconstruction and segmentation.

The two stages are deliberately kept separate rather than jointly optimised, so that illumination, turbidity, or target-density changes can be re-checked at frame and region levels separately. We use merged ROI for the padded rectangular supports carried forward after Stage 2; the notation Ω_k is reserved for formulas, and X denotes the demultiplexed complex amplitude field throughout.

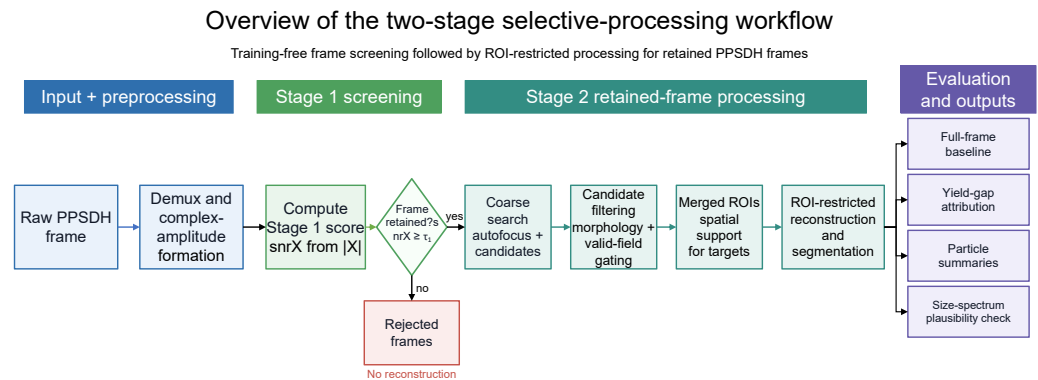


Figure 4. Overview of the two-stage selective-processing workflow. Stage 1 forms a campaign-specific retained-frame pool from the demultiplexed PPSDH complex field, and Stage 2 converts retained frames into merged ROIs through coarse reconstruction, candidate filtering, valid-field gating, and spatial merging. The retained-frame pool and the stratified evaluation set used in the benchmark analyses are defined in Sections 3.3.2 and 3.3.3.

3.2. Notation Used in the Workflow

Table 2 lists the main symbols used in Sections 3–4.

Table 2. Main notation used in the two-stage workflow.

Symbol	Meaning
$X \in \mathbb{C}^{H \times W}$	Demultiplexed complex amplitude field recovered from the four PPSDH intensity samples.
H, W	Height and width of X in pixels.
snr_X	Normalised amplitude-contrast statistic used for Stage 1.
τ_1, τ_2	Campaign-specific contrast thresholds used for frame screening and stratification.
A_z	Reconstructed complex field at propagation distance z .
z^*	Selected in-focus reconstruction distance for a frame or ROI.
V	Per-frame valid-field mask derived from the illuminated reconstructed field.
τ_c	Per-frame candidate threshold applied to the background-suppressed amplitude image.
Ω_k	Padded rectangular merged ROI used as a downstream processing support.
K	Number of merged ROIs in a retained frame.
A_p, c_p	Area and circularity of segmented particle-like target p .
ρ	Spatial-support reduction ratio, i.e., the fraction of the full reconstructed field excluded from downstream processing.

3.3. Stage 1: Frame-Level Quality Screening

3.3.1. Screening Statistic

Let $X \in \mathbb{C}^{H \times W}$ denote the complex amplitude field obtained from the parallel phase-shifting demultiplexing step described in Section 2.1 (Equation (1)). We define the per-frame screening statistic as the normalised peak-to-peak contrast of the amplitude map $|X|$:

$$\text{snr}_X = \frac{\max_{(x,y)} |X(x,y)| - \min_{(x,y)} |X(x,y)|}{\frac{1}{HW} \sum_{(x,y)} |X(x,y)|}. \quad (2)$$

The statistic snr_X is suitable at the frame level for three reasons. First, it is computed directly on the demultiplexed complex amplitude field and requires no reconstruction, so its cost is a single pass over $HW \approx 1.25 \times 10^6$ pixels. Second, mean normalisation reduces the effect of global illumination drift. Third, in-focus or near-focus particles produce local $|X|$ excursions, whereas defocused bubble clouds, saturated sediment plumes, or laser instability produce flatter $|X|$ profiles; the score therefore captures an amplitude-dominant

contrast distinction. PPSDH also recovers phase information. The official workflow uses the amplitude-derived statistic because it is simple, fast, and was the frozen operating point for all downstream Stage 2 analyses. As an auxiliary sensitivity check, we evaluated a wrapped phase-gradient score on the same 180-frame visual-audit set (Section 4.1); it is not used to replace the official operating point. Percentile- and median absolute deviation (MAD)-based amplitude variants are examined in Section 4.4 as extensions rather than components of the current workflow.

3.3.2. Threshold Selection and Retention Outcome

Figure 5 shows the distribution of snr_X over all $N_{\text{total}} = 9521$ frames acquired during the September 2021 campaign. The distribution contains a low-contrast population at small snr_X , a broad mid-contrast body, and a high-contrast tail. We set the Stage 1 operating threshold at $\tau_1 = 11.0$ and retain a frame if $\text{snr}_X \geq \tau_1$. A second threshold, $\tau_2 = 18.0$, is used only for stratification of retained frames. Both values are campaign-specific operating points for the September 2021 dataset and should be re-estimated or visually audited before transfer to other seasons, depths, or optical settings.

Applying τ_1 to the full campaign retains $N_{\text{ret}} = 6970$ frames (73.2%) and rejects 2551 frames (26.8%). The rejected frames are best described as low-contrast frames under this operating point, not as frames proven to contain no recoverable targets. This distinction is evaluated explicitly by the offline visual audit in Section 4.1.

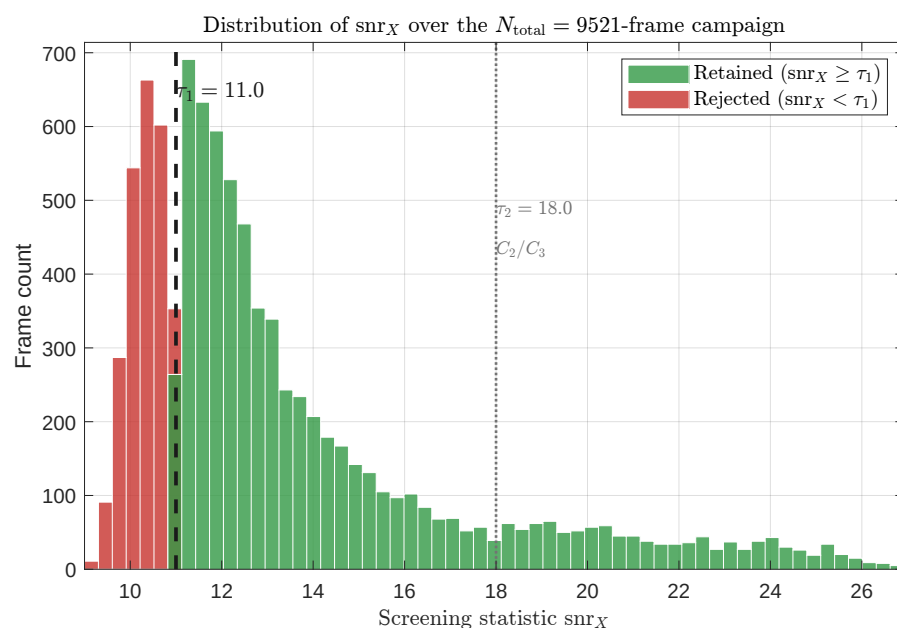


Figure 5. Campaign-wide distribution of the Stage 1 statistic snr_X over all 9521 frames. The campaign-specific operating threshold $\tau_1 = 11.0$ forms the retained-frame pool, while $\tau_2 = 18.0$ is used only to stratify the retained frames into mid- and high-contrast subsets.

3.3.3. Three-Axis Stratified Sampling for Controlled Evaluation

To evaluate Stage 2 under a balanced distribution of scene difficulty, we organise the retained frames along three axes:

- Axis A—image complexity. Mean spatial gradient magnitude $|\nabla I|$ of the amplitude image $|X|$, stratified at the 33rd and 67th percentiles into strata A_1 (simple), A_2 (moderate), and A_3 (complex).
- Axis C—image contrast. The screening statistic snr_X itself, stratified at $\tau_1 = 11.0$ and $\tau_2 = 18.0$ into C_1 (rejected by Stage 1), C_2 (mid-contrast, $11 \leq \text{snr}_X < 18$), and C_3 (high-contrast, $\text{snr}_X \geq 18$).

- Axis D—target density. Post hoc particle count per frame, populated after Stage 2 and segmentation; used for diagnostic reporting rather than sampling.

Because C_1 is removed by Stage 1, the effective $A \times C$ stratification contains six non-empty cells: $A_1C_2, A_1C_3, A_2C_2, A_2C_3, A_3C_2, A_3C_3$. From each cell we draw $n_s = 20$ frames uniformly at random, yielding a balanced evaluation set of $N_{\text{eval}} = 120$ frames. All quantitative results reported in Section 4 are computed on this set.

3.4. Stage 2: Region-Level Selective Processing

3.4.1. Complex-Field Reconstruction with Off-Axis +1 Order Filtering

For each frame that enters Stage 2, we first extract the object-side +1 diffraction order from the demultiplexed complex amplitude field X . Although parallel phase-shifting demultiplexing (Section 2.1) nominally cancels the zero-order and conjugate-image components, small residual alignment errors in the fibre-coupled collimators leave a low-amplitude carrier at spectrum coordinates $(d_r, d_c) \approx (54, 151)$ px, reproducibly located in the right half-plane. The +1 order is therefore isolated in the frequency domain as follows:

$$\tilde{X} = \mathcal{F}_{\text{shift}}\{X\}, \quad \tilde{X}_{+1} = \mathcal{W}_{65}(\tilde{X}), \quad (3)$$

where $\mathcal{F}_{\text{shift}}$ is the centred two-dimensional Fourier transform and $\mathcal{W}_{65}$ is a rectangular crop with half-width 65 px (about 130×130 px in total), centred on the +1-order peak. The peak is located by a constrained search in columns with index greater than $N_c/2 + 100$. The filtered spectrum $\tilde{X}_{+1}$ is re-centred and propagated to a candidate reconstruction distance z using the angular-spectrum transfer function [19]

$$H_z(f_x, f_y) = \exp\left(i \frac{2\pi z}{\lambda} \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2}\right), \quad (4)$$

yielding the reconstructed complex amplitude $A_z(x, y) = \mathcal{F}_{\text{shift}}^{-1}\{\tilde{X}_{+1} \cdot H_z\}$. The 65-px half-width was chosen empirically as the smallest support that preserved object-side spectral content above the residual noise floor without admitting excessive zero-order leakage.

The in-focus reconstruction distance, denoted z^* , is identified by maximising the Laplacian-energy focus measure

$$\mathcal{L}(z) = \sum_{(x,y)} \left[\nabla^2 \left(\frac{|A_z(x,y)|}{\max_{x',y'} |A_z(x',y')|} \right) \right]^2, \quad (5)$$

where ∇^2 is the discrete Laplacian operator. Normalising the amplitude by its per-frame maximum before applying ∇^2 makes $\mathcal{L}$ less sensitive to global illumination variation. The search uses coarse ($\Delta z = 5$ mm) and fine ($\Delta z = 1$ mm) steps over $z \in [151, 230]$ mm, set by the sapphire window spacing and rear detector plane. The coarse maximum centres the fine search window (± 10 mm), giving a final axial sampling granularity of 1 mm.

3.4.2. Per-Frame Valid Field and Effective Field of View

Because the illuminated footprint occupies only part of the detector, weak edge regions are prone to spurious candidates after reconstruction. We therefore defined a per-frame valid field from the normalised reconstructed amplitude at z^* . Let

$$U_n(x, y) = \frac{|A_{z^*}(x, y)| - \min |A_{z^*}|}{\max |A_{z^*}| - \min |A_{z^*}|}, \quad (6)$$

where the extrema are taken over the current frame. The valid field was then obtained as

$$V = \text{Erode}_{20}(\text{Fill}(\mathbf{1}[U_n > 0.15])), \quad (7)$$

where Fill denotes hole filling and Erode₂₀ denotes morphological erosion with a disc structuring element of radius 20 px. Across the 120-frame evaluation set, the valid field occupied $34.6\% \pm 1.8\%$ of the full sensor area on average. This fraction is the illuminated, window-defined effective field, not an arbitrary removal of two thirds of the detector. The beam footprint and usable aperture define the dominant support, while the 20 px erosion conservatively reduces edge-induced false candidates. Candidate formation was first performed on the full reconstructed field; the valid field then acted as a centroid gate.

3.4.3. Candidate Detection and Morphological Filtering

At the estimated in-focus plane z^* , candidate regions were extracted from a background-suppressed amplitude image. A Gaussian background ($\sigma_{bg} = 150$ px) was estimated from U_n and subtracted to form

$$U_{fg}(x, y) = U_n(x, y) - (G_{\sigma_{bg}} * U_n)(x, y). \quad (8)$$

Because particle-like targets appear darker than the local background in this reconstruction convention, only the negative residual was retained,

$$U_{inv}(x, y) = \max(-U_{fg}(x, y), 0), \quad (9)$$

and then re-normalised to $[0, 1]$ before thresholding. Candidate pixels were identified at the per-frame threshold

$$\tau_c = \text{mean}(U_{inv}) + 3.0 \text{std}(U_{inv}). \quad (10)$$

Connected components smaller than 30 px^2 were removed, followed by a disk-closing operation. Each surviving component was retained as a candidate only if it satisfied all of the following conditions:

- area $A_b \in [30, 3000] \text{ px}^2$;
- circularity $c_b = 4\pi A_b / P_b^2 \geq 0.3$ with P_b the perimeter;
- mean of the major- and minor-axis lengths in the range $[5, 80] \text{ px}$;
- centroid within the per-frame valid field V .

Across the evaluation set, the candidate-count cascade decreases from 160.7 ± 51.2 at the raw thresholding stage to 154.3 ± 46.8 after morphological filtering and to 100.3 ± 28.4 after valid-field gating.

3.4.4. Spatial Merging into Reconstruction Regions

Candidate blobs that survive filtering can be fragments of the same target or particles closer than the practical autofocus neighbourhood. Reconstructing each candidate separately would add redundant computation and discontinuous ROI tilings. We therefore merge candidates into padded rectangular merged ROIs using a deterministic centroid-proximity graph. Two candidate centroids are linked when their distance is less than $d_{\text{merge}} = 80 \text{ px}$; each connected component forms one group. The enclosing axis-aligned rectangle is padded by $p = 30 \text{ px}$ on each side to preserve local diffraction context. The resulting merged ROIs are denoted $\{\Omega_k\}_{k=1}^K$ for a frame with K retained merged ROIs.

The spatial-support reduction ratio of a frame is

$$\rho = 1 - \frac{\sum_{k=1}^K |\Omega_k|}{HW}, \quad (11)$$

where $H \times W$ is the full reconstructed frame size and $|\Omega_k|$ denotes the area of the padded rectangular ROI. Across the 120-frame evaluation set, $\rho = 49.9\% \pm 12.0\%$, with a per-frame region count of 30.6 ± 3.3 . These merged ROIs define the spatial supports carried forward for the selective-processing analysis.

3.5. Downstream Particle Segmentation

Within the Stage 2 merged ROIs, particles are delineated using the same background-subtraction framework as candidate detection (Equation (8)), with parameters retuned for the smaller search space. The threshold is lowered from 3σ to 2.5σ , and the admissible area range is widened to $A_p \in [20, 5000] \text{ px}^2$ with circularity $c_p \geq 0.2$. For each segmented particle p we report:

$$d_{\text{eq}}^{(p)} = 2\sqrt{A_p/\pi} \cdot s, \quad c_p = \frac{4\pi A_p}{P_p^2}, \quad (12)$$

where $s = 6.9 \text{ } \mu\text{m}/\text{px}$ is the equivalent pixel pitch of the demultiplexed field (Section 2.1), together with the in-focus depth z^* inherited from the parent region.

3.6. Evaluation Metrics

Stage 1 is evaluated by the retention rate $R_1 = N_{\text{ret}}/N_{\text{total}}$, the balance of the resulting $A \times C$ stratification, and an offline visual audit of a balanced 180-frame set. In that audit, each frame was reconstructed offline and assigned a binary visual-usefulness label indicating whether the preview contained at least one visually interpretable particle-like target suitable for downstream processing. These labels were used only for post hoc evaluation of the Stage 1 operating point; they were not used to train the workflow or tune the threshold.

Stage 2 is evaluated, per frame, by: (i) the spatial-support reduction ratio ρ (Equation (11)); (ii) the filter-cascade attrition $n_{\text{raw}} \rightarrow n_{\text{morph}} \rightarrow n_{\text{fov}} \rightarrow K$; (iii) the recovered in-focus distance z^* ; and (iv) the end-to-end processing time. Timing is decomposed into screening, per-frame valid-field generation, coarse reconstruction plus candidate detection, merging, and downstream selective processing. The spatial-support reduction ratio measures the fraction of the full reconstructed field excluded from downstream processing. It is not a codec-based data-compression ratio.

Downstream segmentation is evaluated by particle count N_p , particle density $\rho_p = N_p/|V| \cdot s^{-2}$ (in mm^{-2}), equivalent diameter, circularity, and the particle size spectrum. The spectrum exponent is fitted as $n(d) \propto d^{-\mu}$ in log-log coordinates over $50 \leq d \leq 300 \text{ } \mu\text{m}$ using ordinary least squares on logarithmically binned counts, following standard practice in marine particle-spectrum analysis [20,21]. The fitted exponent μ is used as an observation-oriented plausibility check for particle summaries.

As an efficiency baseline, we report the cost of full-frame downstream processing over the same 120 frames without Stage 2 gating (Section 4.3). The manuscript uses two internal baseline families to keep the claim boundary explicit. The efficiency comparison evaluates (i) exhaustive all-frame full-frame processing, (ii) Stage 1-only screening followed by full-frame reconstruction of retained frames, and (iii) the proposed two-stage workflow. A matched-retention robustness baseline compares the official M_0 score with the MAD-based alternative M_3 . These baselines bound the claims rather than exhaust the space of alternatives. Unless otherwise stated, the parameter set in Sections 3.3 and 3.4 was fixed before the main 120-frame evaluation and was not tuned by stratum.

3.7. Statistical Analysis

All summary statistics are reported as mean $\pm$ standard deviation unless otherwise stated. Bootstrap confidence intervals were computed by resampling frame-level observations with replacement. Spearman's rank correlation was used for monotonic associations between spatial-support reduction, candidate count, particle density, and related frame-level variables, because these quantities were not assumed to follow a normal distribution. Mann–Whitney tests were used for two-sample comparisons such as matched versus missed particle sizes. Kolmogorov–Smirnov tests were used as distributional checks. When multiple tests or correlations were evaluated within the same figure, table, or analysis family, we controlled the family-wise error rate using the Holm–Bonferroni procedure and used adjusted p values for inferential statements. Comparisons not used to support a thresholded inferential claim are treated as descriptive sensitivity checks. In particular, the auxiliary screening-metric analyses bound dependence on the chosen Stage 1 score; they do not select a new operating point by significance testing.

3.8. Auxiliary Robustness Evaluation Protocols

Because the Stage 1 score is a pragmatic engineering choice rather than a unique optimum, we added a compact robustness check alongside the main 120-frame evaluation. The current screening statistic was written as

$$M_0 \equiv \text{snr}_X = \frac{\max |X| - \min |X|}{\text{mean}(|X|)}, \quad (13)$$

and recomputed from the raw demultiplexed frames to verify exact consistency with the official Stage 1 values. On a stratified auxiliary sample of 499 frames, M_0 was then compared with three robust amplitude-based alternatives,

$$M_1 = \frac{p_{99}(|X|) - p_1(|X|)}{\text{median}(|X|)}, \quad M_2 = \frac{p_{75}(|X|) - p_{25}(|X|)}{\text{median}(|X|)}, \quad M_3 = \frac{\text{MAD}(|X|)}{\text{median}(|X|)}, \quad (14)$$

as well as two proxy features (M_4 : mean gradient magnitude; M_5 : entropy). Agreement was assessed at matched retention and within the boundary window $M_0 \in [9.5, 12.5]$.

An auxiliary matched-retention check was used to test whether the most plausible robust alternative (M_3) changed downstream behaviour. In brief, balanced pools with 20 frames from each $A \times C$ cell ($N = 180$ per pool) were sampled over five random seeds; within each pool, the M_3 threshold was adjusted to match the M_0 retention rate, and the same Stage 2 statistics were recomputed. The full seed-wise procedure and metric-delta results are reported in Supplementary Note S1 and Figure S1. Throughout the manuscript, M_0 remains the official Stage 1 score because it generated the stored campaign-wide snr_X values, defines the 9521-frame operating point and main 120-frame evaluation, and coincides with a visually separable low-contrast mode at $\tau_1 = 11.0$. We therefore treat M_3 as a limited sensitivity comparator, not as a replacement screening rule in this manuscript.

4. Results

This section follows evidential priority. We first report Stage 1 screening, the offline visual audit, Stage 2 spatial-support reduction, and the matched full-frame benchmark that defines the main timing result. We then quantify the efficiency–fidelity trade-off through low-correspondence diagnostics, report the auxiliary M_0/M_3 Stage 1 robustness analyses, and examine downstream particle statistics as observation-oriented plausibility checks.

Unless otherwise stated, all quantitative results are computed on the balanced 120-frame evaluation set defined in Section 3.3.3 (20 frames per stratum, six strata). All timings were measured on an Intel Core i5-13600K workstation at 3.5 GHz with 64 GB

of memory, running MATLAB R2023b in single-threaded mode. All reported speedups are relative comparisons under this fixed reference implementation. Absolute wall-clock values would shift under parallel or graphics-processing-unit (GPU) execution, but the ordering of the efficiency ladder is the relevant quantity here.

4.1. Stage 1 Screening Performance and Offline Visual Audit

The $N_{\text{total}} = 9521$ raw frames span a broad range of scene difficulty. We used the 33/67 percentile cuts on $|\nabla I|$ and the fixed thresholds $\tau_1 = 11.0$, $\tau_2 = 18.0$ on snr_X (Section 3.3.3). This partitioned the retained frames into six cells, $A_1C_2, \dots, A_3C_3$, with per-cell populations of 585–2108 frames. Uniform random sampling of $n_s = 20$ frames per cell yielded the final $N_{\text{eval}} = 120$ -frame evaluation set used for the main Stage 2 benchmark.

Applying the Stage 1 operating rule $\text{snr}_X \geq \tau_1 = 11.0$ retains $N_{\text{ret}} = 6970$ frames out of 9521 for a retention rate of $R_1 = 73.2\%$ and a rejection rate of 26.8%. The computational cost is small: each frame requires approximately 18 ms to load, demultiplex, and compute snr_X on the reference workstation, so full-campaign screening completes in under three minutes on a single thread and is trivially parallelisable.

To test the interpretation of this threshold, we performed an offline visual audit on a balanced 180-frame set drawn from the nine $A \times C$ cells (20 frames per cell). Each frame was reconstructed offline and assigned a binary visual-usefulness label: label 1 indicates at least one visually interpretable particle-like target suitable for downstream processing, and label 0 indicates no reliable target after author review. The resulting confusion matrix is summarised in Table 3.

Table 3. Offline reconstruction-based visual audit of the Stage 1 operating threshold on the balanced 180-frame set. The audit evaluates frame-level usefulness, not a species-identification audit. Accuracy is reported descriptively because no visually negative C1 frames occurred in this balanced audit.

Quantity	Value
Audited frames	180
Stage-1-retained frames in audit	120
Stage-1 rejected low-contrast frames in audit	60
Visually usable frames	177
Visually rejected frames	3
True positives (retained and visually usable)	117
False positives (retained but visually rejected)	3
False negatives (rejected but visually usable)	60
True negatives	0
Retained-frame precision	0.975
Recall under visual-usefulness definition	0.661
Accuracy	0.650

This audit clarifies Stage 1. The retained-frame pool has high visual precision: 117 of 120 retained audit frames contained visually interpretable particle-like targets. However, all 60 rejected C_1 audit frames also contained weak or visually recoverable structures after offline reconstruction. Stage 1 is therefore not a universal target-absence detector. It is a low-contrast pruning rule that defines a high-confidence retained-frame pool while accepting a measurable loss of weak low-contrast targets.

As a sensitivity check, we evaluated a wrapped phase-gradient score on the same 180-frame visual-audit set. At the same retained-frame count of 120, it gave a retained-frame precision of 1.000 and recall of 0.678 under the author-checked visual-usefulness labels, compared with 0.975 and 0.661 for the official amplitude-derived decision. However, frame-level decision agreement was only 0.567:81 frames were retained by both rules, 39 only by the phase-gradient score, 39 only by the amplitude-derived rule, and 21 by

neither rule. We therefore treat phase-derived screening as an informative future extension, not as a replacement for the frozen amplitude-derived Stage 1 operating point used for downstream evaluation in this revision.

4.2. Stage 2 ROI Selection and Spatial-Support Reduction

Across the 120-frame evaluation set, Stage 2 yields a per-frame spatial-support reduction ratio (Equation (11)) of $\rho = 49.9\% \pm 12.0\%$ (mean $\pm$ standard deviation), with the corresponding mean region count $K = 30.6 \pm 3.3$ per frame. The filter cascade (Section 3.4.3) reduces the per-frame candidate count from 160.7 ± 51.2 (raw), to 154.3 ± 46.8 (after morphological filtering), to 100.3 ± 28.4 (after per-frame valid-field gating), with spatial merging consolidating those into the ~ 31 final ROIs per frame. The recovered in-focus distance z^* averaged 184.3 ± 24.2 mm and showed two preferred ranges, with peaks near 170 mm and 215 mm (Figure 6, top-left). Because no concurrent density or fluorescence profile was recorded, this bimodality is reported as a reconstruction-depth pattern, not as evidence for two physically separated particle layers.

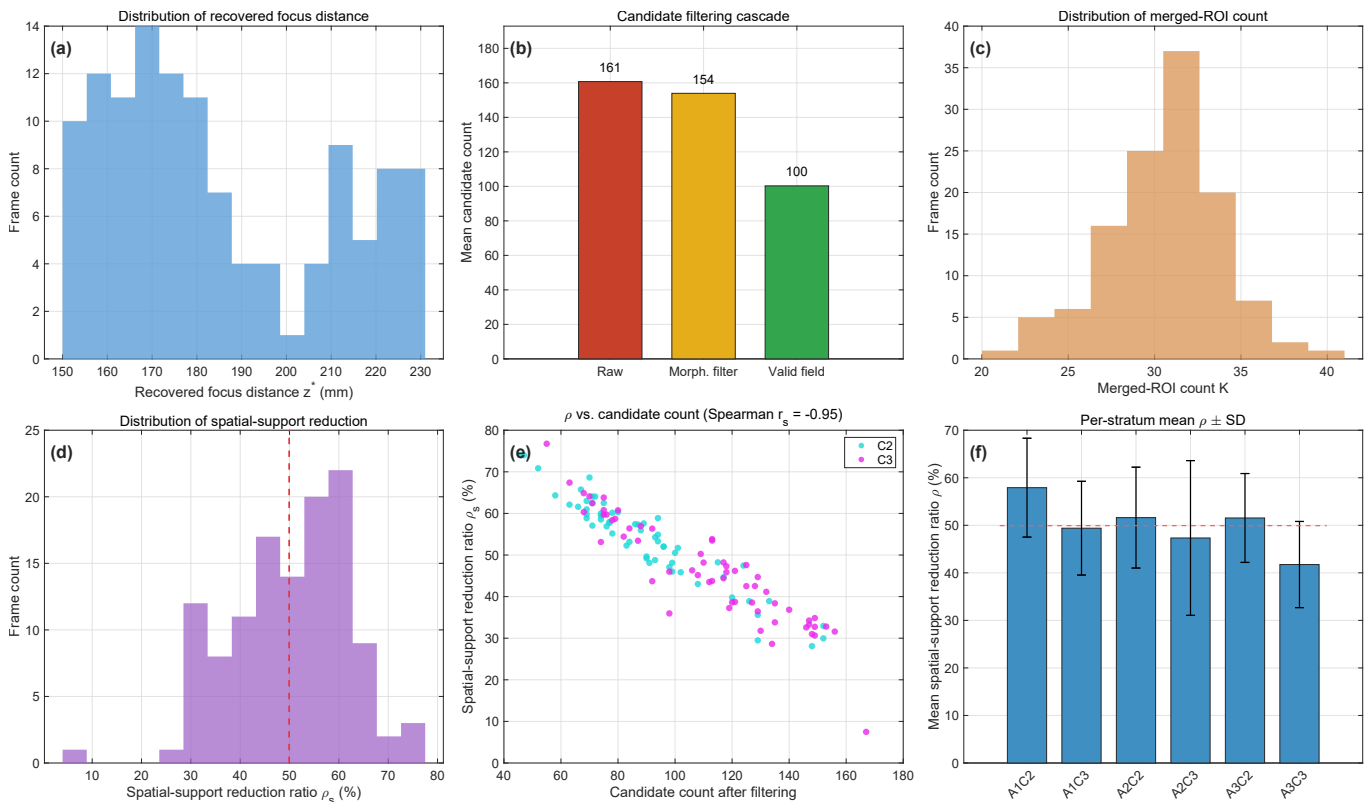


Figure 6. Stage 2 ROI selection and spatial-support reduction on the stratified 120-frame evaluation set. (a) Distribution of the recovered focus distance z^* . (b) Mean candidate count after the raw-detection, morphological-filtering, and valid-field-filtering steps. (c) Distribution of the merged-ROI count K . (d) Distribution of the spatial-support reduction ratio ρ , with the dashed line indicating the overall mean. (e) Relationship between the retained candidate count after filtering and ρ , showing a strong negative Spearman correlation ($r_s = -0.95$). (f) Per-stratum mean ρ with standard deviation; the dashed line marks the overall mean.

Per-stratum statistics are summarised in Table 4. The spatial-support reduction ratio was highest in the sparsest stratum A_1C_2 ($57.9 \pm 10.4\%$) and lowest in the densest stratum A_3C_3 ($41.7 \pm 9.1\%$). Across individual frames, the spatial-support reduction ratio was strongly anti-correlated with the retained candidate count (Spearman $r_s = -0.946$, $p = 2.42 \times 10^{-59}$), indicating that the merged-ROI support expands content-adaptively as scene density increases. A one-way analysis of variance (ANOVA) across the six $A \times C$

strata further supported a significant stratum-level effect on ρ ($p = 7.79 \times 10^{-4}$). Thus, the main evidence for self-adaptivity is the frame-level support-reduction–density relationship; the per-stratum means provide a compact operating-point summary.

The Stage 2 per-frame time for coarse reconstruction, candidate detection, filtering, and merging was 2.7 ± 0.3 s on the reference workstation, excluding subsequent ROI-restricted downstream passes. Variation across frames was dominated by the number of merged regions rather than by raw candidate count. The next subsection converts this region-level reduction into the main wall-clock comparison against a full-frame baseline.

Table 4. Numerical Stage 2 operating points by $A \times C$ stratum on the stratified 120-frame evaluation set. Each stratum contains 20 frames. z^* is the recovered in-focus distance, n_{raw} and n_{fov} are candidate counts before and after the filtering cascade, K is the merged-ROI count, $\bar{\rho}$ is the mean spatial-support reduction ratio, and $\bar{N}_p$ and $\bar{\rho}_p$ are downstream particle-summary quantities.

Stratum	N	$\overline{z^*}$ (mm)	$\overline{n_{\text{raw}}}$	$\overline{n_{\text{fov}}}$	$\bar{K}$	$\bar{\rho}$ (%)	$\bar{N}_p$	$\bar{\rho}_p$ (mm ^{−2})
A_1C_2	20	176.8	123.8	80.5	29.4	57.9	113.5	5.39
A_1C_3	20	186.7	164.9	101.8	30.1	49.4	154.1	7.59
A_2C_2	20	183.3	154.9	95.5	31.1	51.6	146.6	7.18
A_2C_3	20	186.9	174.6	107.5	30.5	47.3	159.3	8.06
A_3C_2	20	180.2	147.1	93.7	31.3	51.5	141.2	6.74
A_3C_3	20	192.0	199.1	122.8	31.3	41.7	188.2	9.43
All	120	184.3	160.7	100.3	30.6	49.9	150.5	7.40

4.3. Comparison with Full-Frame Reconstruction Baseline

To quantify end-to-end efficiency, we ran a full-frame baseline on the same stratified 120-frame evaluation set used for the main Stage 2 analysis. This is the principal full-frame reference, not an exhaustive campaign-wide reprocessing. The baseline follows the same Stage 1 screening, so both configurations operate on the same retained frames, but skips Stage 2 ROI construction. The in-focus plane z^* is re-determined by full-frame Laplacian-energy autofocus, and reconstruction and segmentation are applied to the full-frame support. Both configurations share the same segmentation routine and parameters (Section 3.5); differences therefore reflect only the spatial scope of reconstruction and autofocus.

Table 5 reports the per-frame comparison on the 120-frame evaluation set. For single-pass campaign analysis, the headline result is the end-to-end comparison. Including the Stage 2 upstream cost (2.7 ± 0.3 s per frame, Section 3.4), the two-stage workflow costs approximately 2.88 s per frame, compared with 16.32 ± 3.13 s for the full-frame baseline. Bootstrap resampling of the 120 benchmark frames yields an end-to-end speedup of 5.66 with a 95% confidence interval of [5.47, 5.85].

Once merged ROIs and cached focus depths are available, downstream reconstruction and segmentation take 0.184 ± 0.023 s per frame, whereas the full-frame baseline takes 16.32 ± 3.13 s from the complex amplitude field. The downstream-only ratio is 88.8 (bootstrap 95% CI [85.8, 91.6]). This larger figure is secondary: it represents cached merged-ROI reuse, after the full-frame autofocus cost has been paid once. Once a depth plane and spatial support are given, the reconstruction and segmentation sub-steps are comparable fixed-cost operations.

Intermediate storage also decreases: the two-stage configuration stores merged padded ROI data (2.38 ± 0.23 MB per frame), whereas the baseline stores the full-frame support (3.47 ± 0.15 MB per frame), a $1.46\times$ reduction. The reduction is smaller than the naive area

ratio because each merged ROI includes 30 px padding on all four sides (Section 3.4.4) and storage is accounted on padded rectangles rather than an idealised unpadded union.

For this benchmark, particle yield is the output count under the baseline-matched protocol and is reported separately from the broader observation-summary pool used for the size-spectrum plausibility check in Section 4.6. Under this protocol, the two-stage configuration yielded 115 ± 38 particles per frame, compared with 161 ± 44 under the full-frame baseline. Bootstrap resampling gives a baseline-matched correspondence rate of 0.612 (95% CI [0.576, 0.647]) and a count-based yield gap of 0.287 (95% CI [0.269, 0.306]). Across the 120 benchmark frames, 76 (63.3%) have correspondence ≥ 0.6 , whereas 11 (9.2%) fall below 0.2; all 11 low-correspondence cases occur in C_2 strata.

Table 5. Primary full-frame benchmark on the stratified 120-frame evaluation set. The main single-pass comparison is the end-to-end cost of the proposed two-stage workflow, including Stage 2 ROI generation, relative to the full-frame baseline. Cached downstream-only timing is reported separately as a secondary merged-ROI reuse scenario. Correspondence rate and count-based yield gap quantify the fidelity cost relative to the full-frame baseline.

Configuration	Single-Pass End-to-End Cost	Cached Downstream Time	Storage	Yield and Correspondence
Full-frame baseline	16.32 ± 3.13 s frame ⁻¹ ; $1.0 \times$	16.32 ± 3.13 s frame ⁻¹	3.47 ± 0.15 MB/fr.	161 ± 44 particles frame ⁻¹ ; reference particle set
Two-stage workflow	2.88 s frame ⁻¹ ; $5.66 \times$ CI [5.47, 5.85]	0.184 ± 0.023 s frame ⁻¹	2.38 ± 0.23 MB/fr.	115 ± 38 particles frame ⁻¹ ; correspondence 0.612; yield gap 0.287

End-to-end cost for the two-stage workflow includes Stage 2 ROI generation and the ROI-restricted downstream pass. Cached downstream time denotes the secondary reuse scenario in which merged ROIs and focus depths have already been computed. Confidence intervals were obtained by bootstrap resampling of the 120 benchmark frames.

These numbers show that Stage 2 spatial-support reduction becomes a measured wall-clock gain by amortising or eliminating the autofocus bottleneck. At the retained-frame volume ($N_{\text{ret}} = 6970$), full-campaign single-pass processing scales to approximately 31.6 h under Stage 1-only full-frame processing versus 5.6 h under the two-stage workflow on one reference workstation thread. Extending the accounting to exhaustive all-frame full-frame processing gives approximately 43.2 h for all 9521 frames. The baseline ladder in Table 6 shows that Stage 1-only frame pruning gives a modest 1.36-fold campaign-level reduction, whereas adding Stage 2 gives 7.7-fold reduction relative to exhaustive processing and 5.66-fold relative to Stage 1-only full-frame processing on the retained subset. The point is that the main incremental efficiency gain comes from ROI-restricted downstream processing, not frame rejection alone. If Stage 2 is amortised across repeated downstream passes, such as segmentation-threshold tuning or updated post-processing, the marginal cost falls to 0.184 s per frame, or about 21 min for the full retained set.

To bound the origin of the yield gap in Table 5, we used the full-frame baseline particles as the reference and matched them to the two-stage output using the same 50 μm centroid-radius criterion. Of the 19,310 full-frame baseline particles in the 120-frame benchmark, 11,839 (61.3%) were matched, 5954 (30.8%) were missed despite lying inside the Stage 2 merged-ROI union, and 1517 (7.9%) were missed outside the ROI union. Thus, incomplete ROI coverage explains only a minority of unmatched baseline particles; most were inside the selected support but were not recovered by the two-stage segmentation and matching procedure.

Table 6. Campaign-scale timing comparison derived from the measured 120-frame benchmark. The first two rows are accounting references, whereas the final row is the measured single-pass operating point of the proposed two-stage workflow applied to the retained-frame volume. The table separates the modest benefit of Stage 1 frame rejection from the larger incremental gain produced by Stage 2 ROI-restricted downstream processing.

Operating Point	Frames	Campaign-Scale Time	Relative Speedup	Interpretation
[Derived] Exhaustive all-frame full-frame processing	9521	≈43.2 h	1.0×	Upper-cost reference if every acquired frame were reconstructed full-frame.
[Derived] Stage 1-only screening + retained full-frame reconstruction	6970	≈31.6 h	1.36×	Isolates the benefit of frame rejection before any ROI restriction.
[Measured] Proposed Stage 1 + 2 workflow	6970	≈5.6 h	7.7× vs. exhaustive; 5.66× vs. Stage 1-only	Combines frame pruning with ROI-restricted downstream processing.

The missed baseline particles were also smaller than the matched particles, with median equivalent diameters of 43.3 μm and 63.3 μm , respectively (two-sample Kolmogorov–Smirnov statistic $D = 0.540$; Mann–Whitney test $p < 10^{-6}$). The retained particle set should therefore not be interpreted as a random subsample of the full-frame output. Instead, the yield gap is associated mainly with small or otherwise weak particles within the selected ROI support, with a smaller contribution from incomplete ROI coverage. Figure 7 summarises this attribution analysis visually without changing the primary efficiency evidence in Tables 5 and 6.

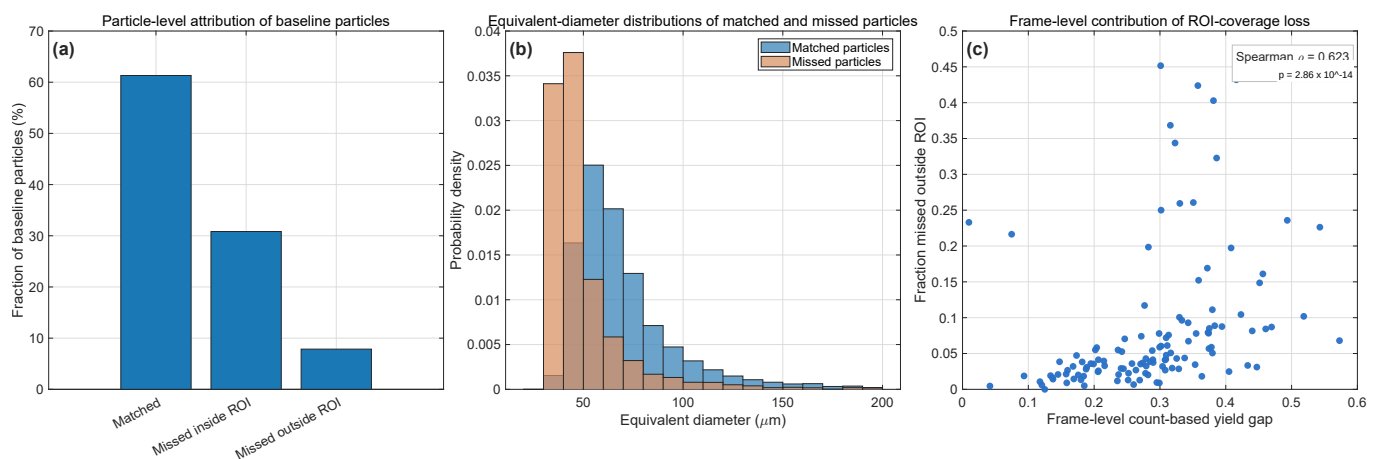


Figure 7. Particle-level attribution of baseline particles in the 120-frame evaluation subset. (a) Fraction of baseline particles categorized as matched, missed inside the ROI, or missed outside the ROI. (b) Equivalent-diameter distributions of matched and missed particles, showing that missed particles are concentrated toward smaller apparent diameters. (c) Frame-level relationship between the count-based yield gap and the fraction of particles missed outside the ROI. The positive Spearman correlation indicates that ROI-coverage loss contributes to the frame-level yield gap. This analysis attributes downstream particle-count differences and does not imply taxonomic particle identification.

The extreme low-correspondence cases represent a related but more severe limiting case. After duplicate rows introduced by metadata joins were removed, the 11 unique frames with correspondence below 0.2 were all in the C_2 strata: six in A_1C_2 , one in A_2C_2 , and four in A_3C_2 . Their median correspondence was 0.096. Among the 1881 missed baseline particles in these frames, 1322 were inside the selected Stage 2 ROI support and 559 were outside it. Thus, these cases are not explained mainly by gross spatial-support exclusion. They are dominated by weak, small, or fragmented targets within the processed support, with a smaller but non-negligible contribution from particles outside the cached ROI support.

Together, these numbers define a measured efficiency–fidelity trade-off. The engineering headline is the 5.66-fold end-to-end speedup, but the lower particle yield, imperfect correspondence, and size-biased missed-particle distribution show that the workflow is a selective observation strategy, not a substitute for full-frame reconstruction when complete recovery is required. Tables 5 and 6 are therefore the primary quantitative evidence; the following analyses bound the screening choice and check downstream plausibility.

With this efficiency baseline established, the remaining Results subsections are secondary: a compact robustness diagnosis of the Stage 1 screening statistic and a check of whether retained outputs support downstream particle summaries.

4.4. Auxiliary Robustness and Matched-Retention Analyses

This subsection is a sensitivity analysis, not a new operating-point search. It first verified exact reproducibility of the official Stage 1 score: the recomputed M_0 matched the stored snr_X exactly (Spearman 1.000, Pearson 1.000, zero mean absolute error and root-mean-square error, and perfect overlap at $\tau_1 = 11.0$).

On the 499-frame stratified auxiliary sample, robust alternatives were not minor perturbations of M_0 . Their Spearman correlations with M_0 were -0.414 (M_1), 0.462 (M_2), and 0.564 (M_3). The corresponding matched-retention decision agreements were 0.511, 0.743, and 0.776. Within the boundary window $M_0 \in [9.5, 12.5]$, at least one alternative flipped the decision label in 206 of 288 frames (71.5%), and 72.5% of frames satisfied $\text{extreme_gap} > 0.5 M_0$. Among the tested amplitude-based variants, $M_3 = \text{MAD}(|X|)/\text{median}(|X|)$ was therefore the most plausible robust comparator. Figure 8 shows the reference-consistency and correlation summaries.

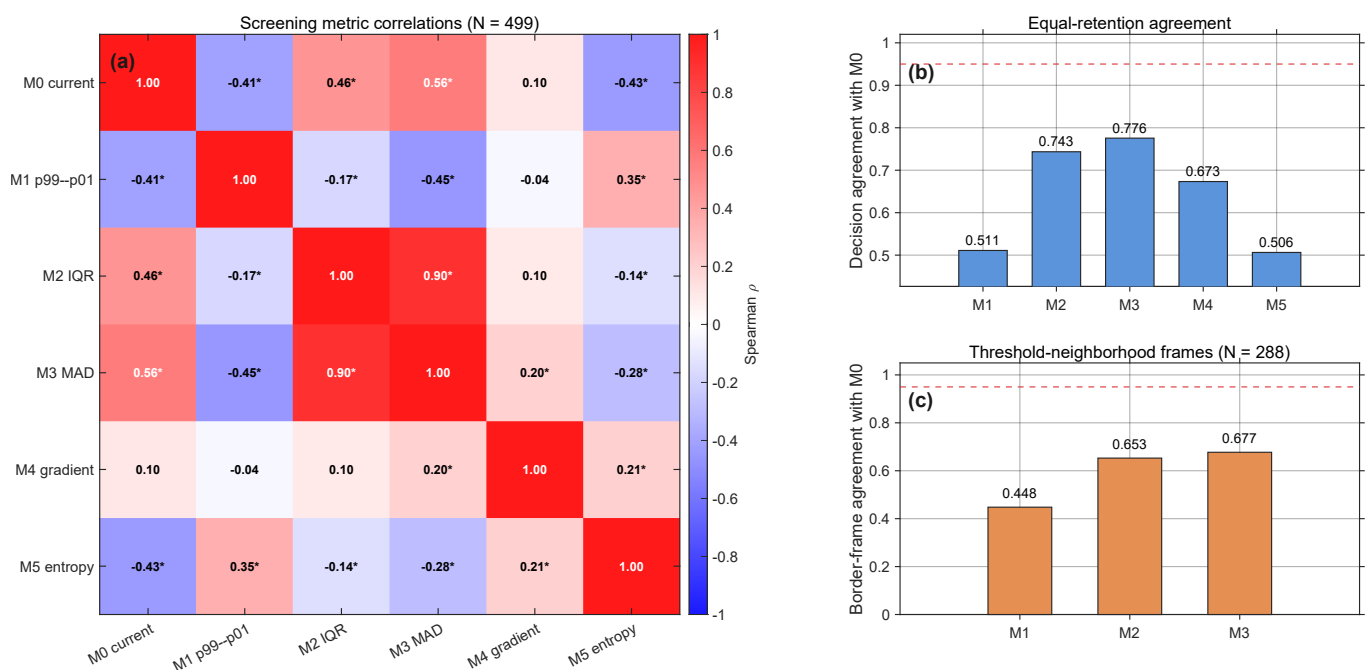


Figure 8. Robustness analysis of alternative Stage 1 screening metrics. Asterisks in panel (a) denote statistically significant Spearman correlations ($p < 0.05$). (a) Spearman correlation matrix among the current metric (M_0) and five alternative metrics: percentile range (M_1), interquartile range (M_2), median absolute deviation (M_3), gradient-based contrast (M_4), and entropy (M_5). Asterisks indicate correlations significant after Holm–Bonferroni correction across 15 pairwise tests at $\alpha = 0.05$. (b) Equal-retention decision agreement with M_0 . (c) Border-frame agreement with M_0 for threshold-neighborhood frames. The dashed line in panels (b,c) indicates the 95% reference level.

We then compared the current $M_0 = (\max |X| - \min |X|) / \text{mean}(|X|)$ criterion with the candidate robust alternative $M_3 = \text{MAD}(|X|) / \text{median}(|X|)$ in a compact multi-seed matched-retention check. Across five random seeds, the average M_0 / M_3 agreement was 0.818 ± 0.017 , and the downstream metric differences remained small. This supports using M_3 as a sensitivity comparator but is not sufficient to redefine the campaign-specific operating point. To keep the main Results focused on the two-stage selective-processing workflow, the full seed-wise comparison, metric-delta values, and interpretation are provided in Supplementary Figure S1 and Note S1.

4.5. Representative Automatically Selected ROI-Restricted Reconstruction Examples

To provide qualitative examples of the local reconstruction output, we added a representative panel from Stage-1-retained campaign frames (Figure 9). For each example, automatic candidate ROI proposals were generated from the reconstructed image response, and one candidate ROI was selected by a fixed candidate-scoring rule for visualization. The selected ROI was then used to form the local crop, ROI-restricted reconstructed amplitude, and background-suppressed target view. These examples illustrate the transition from candidate ROI proposals to local reconstruction outputs; they are not used as a statistical sample, target-boundary annotation, or taxonomic identification.

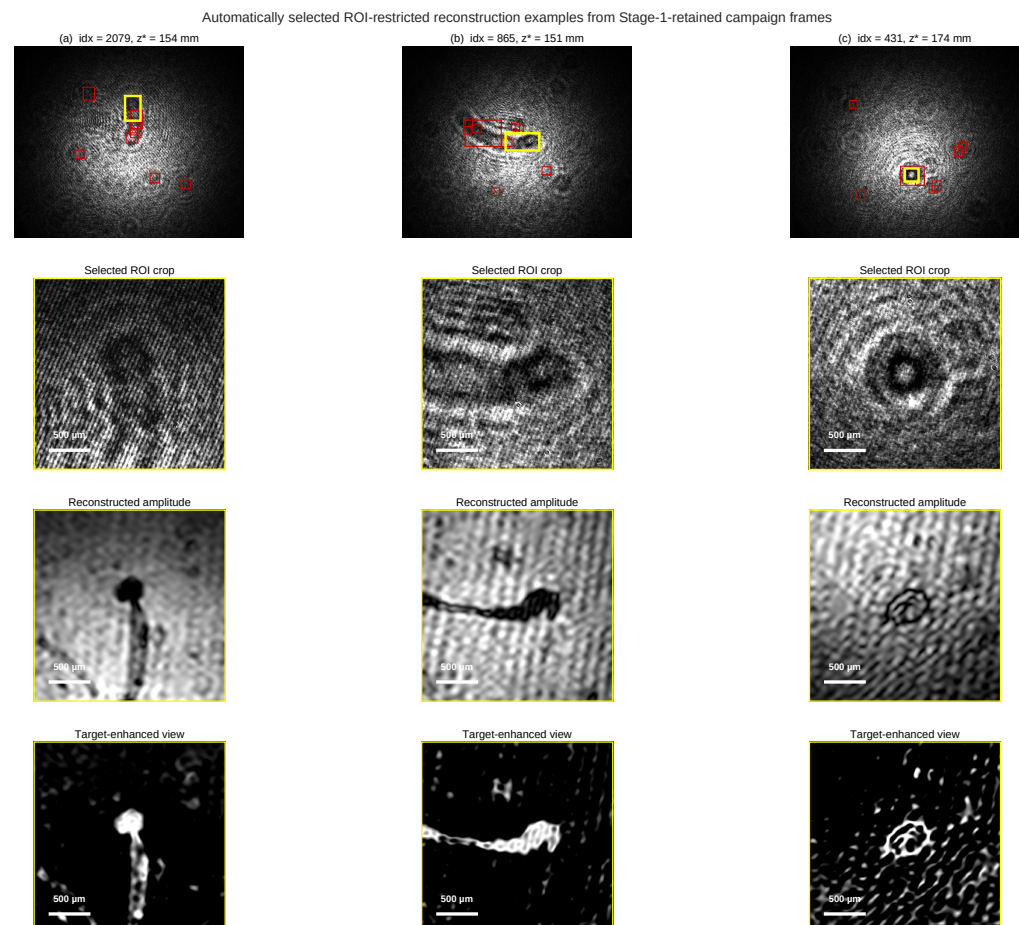


Figure 9. Representative automatically selected ROI-restricted reconstruction examples from Stage-1-retained campaign frames. The first row shows the demultiplexed amplitude frames, where red boxes indicate automatic candidate ROI proposals and the yellow box indicates the candidate ROI automatically selected by the scoring rule. The lower rows show the corresponding selected ROI crop, ROI-restricted reconstructed amplitude, and background-suppressed target view. Scale bars in the local panels correspond to 500 μm . These examples are illustrative particle-like targets and are not taxonomic plankton identifications.

4.6. Downstream Plausibility Checks: Particle Segmentation and Size Spectrum

This subsection asks a narrow downstream question: after the computational pruning quantified above, do the retained ROIs still support interpretable particle summaries? Segmentation within the merged Stage 2 ROIs (Section 3.5) yielded $N_p^{\text{tot}} = 18,057$ particles across the 120-frame evaluation set. This corresponds to 150.5 ± 48.8 particles per frame (range [47, 238]) and $7.40 \pm 2.65 \text{ mm}^{-2}$ referenced to the effective field-of-view area. This pooled count supports the particle-summary and size-spectrum analyses and should not be conflated with the baseline-matched yield metric in Table 5. Per-stratum density spans from 5.39 mm^{-2} in A_1C_2 to 9.43 mm^{-2} in A_3C_3 (Table 4), supporting the interpretation that Axis A stratification on $|\nabla I|$ captures scene-density differences rather than an acquisition artefact.

Pooled across all particles, the equivalent-diameter distribution (Figure 10, top row; Figure 11, top left) is strongly right-skewed, with median $55.6 \mu\text{m}$, mean $65.2 \mu\text{m}$, and interdecile range $38.1\text{--}100.3 \mu\text{m}$. The lower cutoff is consistent with the demultiplexed pixel pitch and segmentation lower-bound area, whereas the upper tail reflects occasional larger segmented objects retained within the merged-ROI support. Mean particle circularity is $\bar{c}_p = 0.794$.

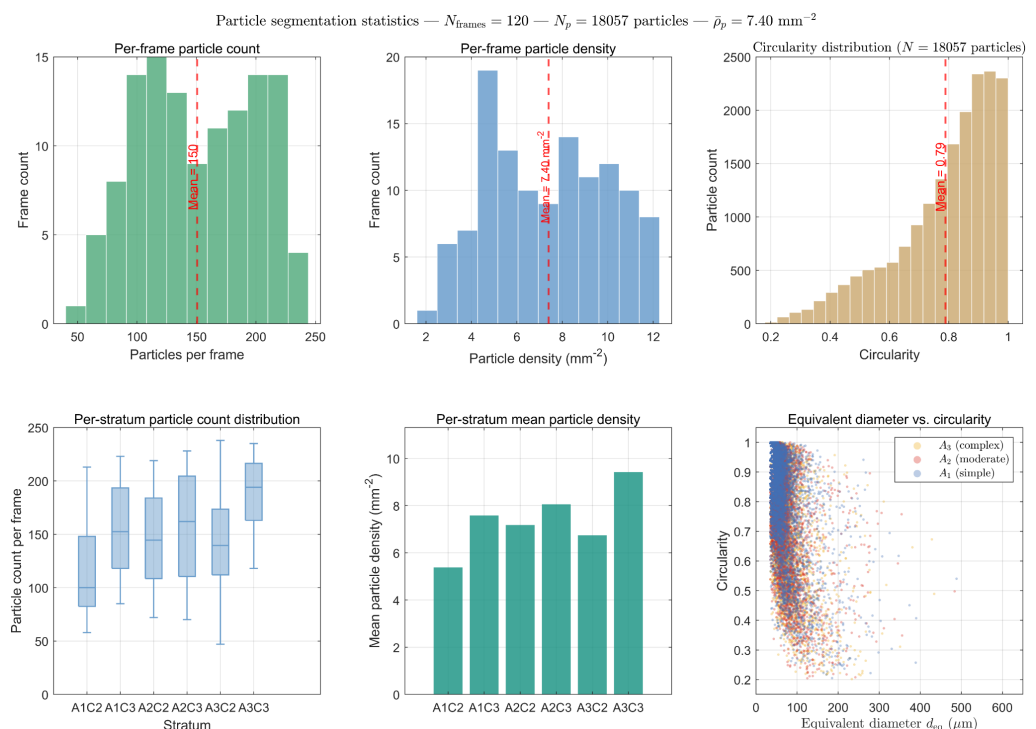


Figure 10. Downstream particle-summary plausibility check for the retained Stage 2 outputs over 120 frames and 18,057 segmented particles. The panels summarise per-frame particle count, particle density, circularity, and stratum-level density variation. These statistics are used to assess whether retained ROIs support interpretable downstream summaries.

Stratification by the two experimental axes produces largely overlapping size distributions (Figure 11, top right and bottom left). The median equivalent diameter shifts by only about $6 \mu\text{m}$ between the sparsest stratum A_1C_2 ($\bar{d}_{\text{eq}} = 61.3 \mu\text{m}$) and the densest stratum A_3C_3 ($54.8 \mu\text{m}$). The sparsest stratum is enriched in larger particles on a per-frame basis. In the present paper this observation is used only descriptively, as a check that the stratification captures a meaningful scene-content gradient rather than as a standalone composition inference.

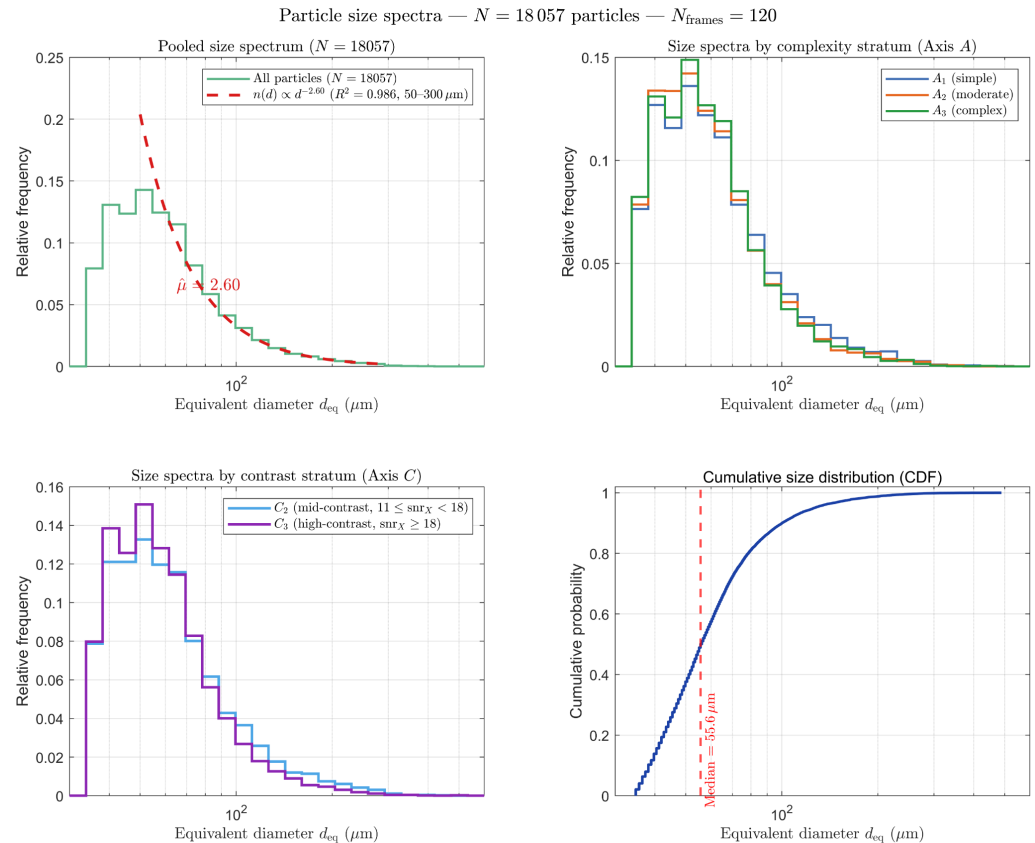


Figure 11. Size-spectrum plausibility check for the retained Stage 2 outputs: pooled equivalent-diameter distribution with the 50–300 μm power-law fit, spectra stratified by axes *A* and *C*, and the cumulative size distribution. The figure assesses particle-summary coherence within the retained outputs.

The particle-number size spectrum, fitted in log–log coordinates over the range $50 \leq d \leq 300 \mu\text{m}$ where bin occupancy exceeds 100 counts per bin (yielding 11 of 15 log-spaced bins retained for the fit), follows a power law

$$n(d) \propto d^{-\mu}, \quad \hat{\mu} = 2.60 \pm 0.10, \quad R^2 = 0.986, \quad (15)$$

with the quoted standard error estimated from the ordinary-least-squares residuals under a homoscedastic log–log model. The high coefficient of determination ($R^2 = 0.986$) indicates a close log–log fit over this binned diameter range.

The fitted exponent $\hat{\mu} = 2.60$ is below the idealised $\mu \approx 3$ expectation for equal-biomass size classes and below broader field values often reported for mixed particle assemblages [20–22]. This offset should not be over-interpreted because the spectrum combines the effects of Stage 1 screening, Stage 2 region merging, and within-region segmentation. Post hoc checks indicate that the value is not a fragile fitting artefact: across stable fit ranges with $d \geq 50 \mu\text{m}$, the exponent remains within $\hat{\mu} \in [2.54, 2.86]$, and area-threshold sweeps leave $\hat{\mu}$ unchanged for $A_{\text{min}} \leq 40 \text{ px}^2$.

The operationally meaningful point is that the retained outputs still form a coherent particle-size summary over the fitted range. This is useful but limited evidence: agreement with a power-law form does not establish absolute counting accuracy, unbiased recovery across the full size range, or equivalence to full-frame reconstruction. The size-spectrum result therefore supports only the narrower claim that retained outputs remain usable for downstream particle-analysis summaries under the present setting.

The representative image panel in Figure 9 is therefore used only to make the workflow output visually explicit. The quantitative argument remains based on the efficiency–

fidelity trade-off, the matched full-frame benchmark, the particle-level yield-gap attribution, and the particle-summary statistics.

5. Discussion

The Results support a specific claim. The workflow reduces processing cost by combining campaign-specific low-contrast pruning with ROI-restricted downstream processing, and the matched full-frame benchmark shows a measurable wall-clock gain. The offline visual audit also shows the cost: rejected low-contrast frames can contain weak recoverable particle-like structures. The workflow should therefore be interpreted as an efficiency–fidelity trade-off for rapid marine PPSDH particle-field processing, not as exhaustive reconstruction, unbiased abundance recovery, or taxon identification. The downstream particle statistics and size spectrum assess whether retained outputs remain interpretable.

5.1. Workflow Design Logic: Why Two Stages, and Why in This Order

The central design choice is to separate frame-level screening from intra-frame region selection. Two asymmetries motivate this order. First, the costs differ. snr_X (Equation (2)) requires one pass over the demultiplexed amplitude field and runs in about 18 ms per frame. Region-level candidate detection requires coarse angular-spectrum reconstruction and background subtraction and runs in about 2.7 s per frame. Second, the information scopes differ: snr_X is a global frame statistic for a high-confidence retained-frame pool, whereas candidate detection localises particle-like content within retained frames. Reversing the order would incur Stage 2 cost on the 26.8% of frames rejected by Stage 1, or would require a separate partial-reconstruction gate before the main reconstruction step. The offline visual audit shows that some rejected low-contrast frames still contain weak recoverable structures; the current ordering is therefore an efficiency-first design choice, not a lossless screening step.

A second reason for decoupling the stages is robustness to deployment drift. In multi-day marine campaigns, illumination, turbidity, and target density can change enough to require re-calibration. The present design exposes two decision points—the screening threshold τ_1 and the morphology-plus-valid-field filter cascade—whose effects can be inspected through cascade attrition counts (Table 4). They are modular in implementation, but not statistically independent: changing τ_1 changes the scene-difficulty mix entering Stage 2 and may require re-checking morphology thresholds. The frozen Stage 1 scores illustrate this sensitivity. In the full 9521-frame table, increasing τ_1 from 11.0 to 18.0 reduces the retained fraction from 0.732 to 0.119. In the 180-frame author-checked audit, the same change reduces recall from 0.661 to 0.328 while keeping retained-frame precision at a similar level. In the reported evaluation, these parameters were fixed before the main 120-frame analysis and were not tuned by stratum. Inspectable decision points were prioritised over a possible gain from tighter coupling.

This separation also clarifies the efficiency claim. The end-to-end baseline comparison in Section 4.3 gives an approximately 5.66-fold wall-clock speedup for the complete two-stage workflow, the primary figure for single-pass campaign analysis. The downstream-only comparison reaches 88.8-fold only after Stage 2 merged ROIs are available, and therefore belongs to a secondary cached-ROI reuse scenario. The main claim is not geometric ROI shrinkage alone, but its conversion into measured computational savings under a full-frame reference implementation.

5.2. Self-Adaptivity via the Support-Reduction–Density Gradient

The primary indicator of self-adaptivity is the emergent support-reduction–density relationship, not a fixed ROI budget. The sparsest stratum A_1C_2 has the highest mean

spatial-support reduction ratio ($57.9 \pm 10.4\%$), whereas the densest stratum A_3C_3 has the lowest ($41.7 \pm 9.1\%$). At the frame level, ρ is strongly anti-correlated with retained candidate count (Spearman $r_s = -0.946$, $p = 2.42 \times 10^{-59}$). In fixed-window pre-processing schemes, such as a uniform stride over the effective field or a pre-declared ROI grid, the reduction ratio is a hyperparameter: aggressive settings truncate dense scenes, whereas conservative settings waste compute in sparse scenes. Here, the ratio is emergent because the morphological-filter-plus-merging step allocates support according to the candidate field within fixed padding and merging rules.

Figure 6e further suggests that this adaptivity is driven mainly by scene content rather than contrast. Points from the mid-contrast C_2 and high-contrast C_3 strata follow the same support-reduction-candidate-count relationship without a systematic vertical offset. Thus, snr_X for Stage 1 screening does not appear to control downstream region allocation, consistent with the intended separation between screening and ROI-restricted processing.

5.3. What the Auxiliary Robustness Analyses Imply About Stage 1

The auxiliary analyses in Section 4.4 do not alter the workflow, but they sharpen its claim boundaries. On the 499-frame diagnostic subset, the current screening statistic $M_0 = (\max |X| - \min |X|) / \text{mean}(|X|)$ is exactly reproducible. It is only moderately aligned with $M_3 = \text{MAD}(|X|) / \text{median}(|X|)$: the two scores have Spearman rank correlation $r = 0.564$ and equal-retention decision agreement of 0.776. Within the ± 1.5 boundary window, at least one alternative flips the M_0 decision in 71.5% of cases. The extreme-gap diagnostic shows that 72.5% of frames satisfy $\text{extreme_gap} > 0.5 M_0$, consistent with a non-negligible influence of unusually high frame maxima.

Across five matched-retention auxiliary-pool tests, replacing M_0 with M_3 changes downstream behaviour only modestly. Retained-subset agreement is 0.818 ± 0.017 , with mean deltas of $\Delta \text{spatial-support reduction} = -0.0046 \pm 0.0068$, $\Delta n_{\text{regions}} = 0.052 \pm 0.258$, $\Delta n_{\text{particles}} = +3.60 \pm 2.69$, and $\Delta \text{density} = +0.165 \pm 0.130 \text{ mm}^{-2}$. These seed-level deltas do not alter the overall conclusion. We therefore retain M_0 as the official campaign-defined operating point and use M_3 only to bound sensitivity to the precise Stage 1 score. This is a campaign-defined reproducibility decision, not a claim that M_0 is statistically optimal for future deployments.

5.4. Observation-Oriented Plausibility Check via the Size Spectrum

Within the present evidence, the recovered size spectrum does not carry the same evidential weight as retention, spatial-support reduction, and baseline speedup. Its role is to check whether upstream pruning still supports a coherent downstream particle statistic. The fitted exponent $\hat{\mu} = 2.60 \pm 0.10$ lies below the idealised $\mu \approx 3$ expectation for equal-biomass size classes and below broader field values often reported for mixed marine particle assemblages [20–22]. The fit itself is tight ($R^2 = 0.986$ over 11 log-spaced bins spanning a factor of six in diameter).

For workflow interpretation, the spectrum shape matters more than the exact exponent. The fit is not fragile. Over stable fit ranges preserving the $50 \mu\text{m}$ lower bound, the estimated exponent remains within $[2.54, 2.86]$. Post hoc sweeps of the minimum retained area leave $\hat{\mu}$ unchanged for $A_{\min} \leq 40 \text{ px}^2$, and volume-equivalent diameter substitutions move the exponent farther from, not closer to, the canonical expectation. These checks do not validate the workflow universally, but they indicate that $\hat{\mu} = 2.60$ is not a trivial artefact of one fit window or one diameter definition.

The absolute value of $\hat{\mu}$ is therefore better interpreted as a property of the sampled water volume than as a workflow performance measure. Consistency with a power law over $50 \leq d \leq 300 \mu\text{m}$ is a necessary but not sufficient plausibility check: it does not rule out

shifts in normalisation, compensating biases near the range limits, or composition-specific effects that would require an independent field reference. The size spectrum therefore supports only the narrower claim that retained outputs remain usable for observation-oriented downstream particle-analysis summaries in the present setting.

5.5. Limitations and Scope of the Present Evidence

Several limitations bound the claims made here.

Single-campaign evidence. The workflow was evaluated on one coastal campaign from September 2021. Although the stratified evaluation set spans a wide range of scene difficulty, it does not capture seasonal or geographic variation. Different water types, especially oligotrophic open-ocean conditions, may require re-calibration of τ_1 and morphology-filter thresholds. Recent marine holographic studies have used platform-integrated biomass estimation or time-series concentration analysis for extended monitoring sequences [13,14]; establishing comparable cross-campaign generality for the present workflow remains future work.

Incomplete and size-biased recovery. The correspondence rate and yield gap relative to the full-frame baseline show that the retained particle set is not an unbiased random subsample of the full-frame output. The particle-level attribution analysis in Section 4.3 makes this limitation explicit: most unmatched baseline particles were inside the Stage 2 ROI union, and unmatched particles were smaller than matched particles. Low-contrast frames, weak small particles, out-of-plane particles, ROI-boundary cases, and frames with disagreement between cached Stage 2 depth and recomputed full-frame focus depth are the most plausible limiting cases. The missed-particle attribution supports interpreting the yield gap as a fidelity cost associated with spatial-support reduction. The workflow is intended for computation-aware selective observation and rapid downstream screening; it should not be used when unbiased abundance estimation or exhaustive full-frame recovery is required.

No independent field reference or environmental co-variables. Particle counts, densities, and size distributions are inferred from the workflow itself. The size-spectrum check is therefore a plausibility indicator, not an external validation. Unlike studies that compare holographic size and concentration estimates against independent reference measurements [12], we do not have co-deployed sampling, microscope counts, or synchronized temperature, salinity, chlorophyll, turbidity, or taxonomic-composition measurements for this campaign. We therefore cannot quantify absolute abundance bias, environmental dependence, or taxonomic composition.

Amplitude-dominant processing. Screening and downstream filtering operate on $|X|$. The recovered phase may contain complementary information. The wrapped phase-gradient comparison on the 180-frame audit indicates that phase-derived screening can be informative, but it selected a substantially different retained-frame set from the frozen amplitude-derived rule. A direct signal-to-noise comparison with a separate single-channel acquisition was not available for this campaign, so the paper does not claim an SNR advantage over a single-channel configuration. More reliable phase–amplitude fusion remains an extension, not a requirement for the present pruning result.

Static thresholds and limited robustness. All operating thresholds are fixed at the values listed in Sections 3.3 and 3.4. Over the present ~ 40 min campaign this assumption is adequate, but longer or more variable deployments may benefit from adaptive thresholding. The same qualification applies to Stage 1: $M_0 = (\max |X| - \min |X|) / \text{mean}(|X|)$ is a campaign-defined operating point, whereas percentile- and MAD-based variants remain relevant extensions. These alternatives affect the Stage 1 screening rule rather than the Stage 2 ROI-generation logic.

Relation to learned detectors. The present workflow targets interpretable pre-processing and computational pruning before any recognition stage. It should be viewed as complementary to, not a replacement for, lightweight learned detectors.

6. Conclusions

This study evaluated a training-free selective-processing workflow for real marine PPSDH particle fields. The workflow first forms a campaign-specific retained-frame pool using a fast amplitude-contrast statistic, and then concentrates downstream processing within merged ROIs inside retained frames. On the 9521-frame coastal campaign, Stage 1 retained 73.2% of frames. An offline 180-frame visual audit showed that the retained-frame pool had high visual precision (0.975), but also that rejected low-contrast frames can contain weak recoverable particle-like structures. Stage 1 is therefore a computational pruning rule, not a universal detector of target absence.

On a stratified 120-frame evaluation set, Stage 2 achieved a mean spatial-support reduction ratio of $49.9\% \pm 12.0\%$. A matched full-frame benchmark on the same evaluation set gave a 5.66-fold end-to-end speedup (95% CI 5.47–5.85) under the single-thread MATLAB reference implementation. The speedup has a quantified fidelity cost: relative to the full-frame reference, the baseline-matched correspondence rate was 0.612 and the count-based yield gap was 0.287. Missed-particle attribution indicated that losses were concentrated among small or weak targets within the selected ROI support.

The evidence supports a campaign-specific engineering workflow for selective processing, not equivalence to exhaustive full-frame reconstruction. The downstream segmentation statistics and recovered size spectrum are best read as observation-oriented particle summaries, not as taxon-level or abundance-complete validation. Future work should test broader water types, seasons, depths, and illumination conditions; develop transfer rules for τ_1 and other campaign-specific thresholds; and evaluate whether phase–amplitude fusion or adaptive ROI generation can reduce the loss of weak low-contrast targets. The study provides a baseline against which future learned or adaptive marine PPSDH processing methods can be benchmarked.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jmse14111030/s1>, Figure S1. Multi-seed robustness comparison between the current Stage-1 screening metric M_0 and the MAD-based alternative M_3 ; Note S1: Multi-seed matched-retention analysis; Note S2: PPSDH polarisation-multiplexing principle. Note S3: Desktop resolution check and exposure-time estimate. Figure S2. Desktop resolution check using a 1951 USAF target.

Author Contributions: Conceptualization, X.L. and H.M.; methodology, X.L.; software, X.L.; validation, X.L.; formal analysis, X.L.; investigation, X.L. and H.M.; resources, H.M.; data curation, X.L.; writing—original draft preparation, X.L.; writing—review and editing, X.L. and H.M.; visualization, X.L.; supervision, H.M.; project administration, H.M.; funding acquisition, H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Strategic Pilot Science and Technology Project (Class A) of the Chinese Academy of Sciences, grant number XDA22020402.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to institutional data-management restrictions and the large size of the raw in situ holographic dataset.

Acknowledgments: The authors thank the Institute of Deep-sea Science and Engineering, Chinese Academy of Sciences, for assistance with the design of the instrument pressure housing and support during the sea trial.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Knox, C. Holographic microscopy as a technique for recording dynamic microscopic subjects. *Science* **1966**, *153*, 989–990. [[CrossRef](#)] [[PubMed](#)]
2. Heflinger, L.O.; Stewart, G.L.; Booth, C.R. Holographic motion pictures of microscopic plankton. *Appl. Opt.* **1978**, *17*, 951–954. [[CrossRef](#)] [[PubMed](#)]
3. Picart, P.; Gross, M.; Marquet, P. Basic fundamentals of digital holography. In *New Techniques in Digital Holography*; Picart, P., Ed.; Wiley-ISTE: London, UK, 2015; pp. 1–66.
4. Owen, R.B.; Zozulya, A.A. In-line digital holographic sensor for monitoring and characterizing marine particulates. *Opt. Eng.* **2000**, *39*, 2187–2197. [[CrossRef](#)]
5. Craig, G.; Alexander, S.J.; Anderson, S.; Hendry, D.C.; Hobson, P.R.; Lampitt, R.S.; Lucas-Leclin, B.; Nareid, H.; Nebrensky, J.J.; Player, M.A.; et al. HoloCam: A subsea holographic camera for recording marine organisms and particles. In *Optical Diagnostics for Industrial Applications*; Society of Photo Optical: Bellingham, WA, USA, 2000; Volume 4076, pp. 111–119.
6. Sun, H.; Hendry, D.C.; Player, M.A.; Watson, J. In situ underwater electronic holographic camera for studies of plankton. *IEEE J. Ocean. Eng.* **2007**, *32*, 373–382.
7. Bochdansky, A.B.; Jericho, M.H.; Herndl, G.J. Development and deployment of a point-source digital inline holographic microscope for the study of plankton and particles to a depth of 6000 m. *Limnol. Oceanogr. Methods* **2013**, *11*, 28–40. [[CrossRef](#)]
8. Dyomin, V.; Davydova, A.; Polovtsev, I.; Olshukov, A.; Kirillov, N.; Davydov, S. Underwater holographic sensor for plankton studies in situ including accompanying measurements. *Sensors* **2021**, *21*, 4863. [[CrossRef](#)] [[PubMed](#)]
9. Liu, Z.; Takahashi, T.; Lindsay, D.; Thevar, T.; Sangekar, M.; Watanabe, H.K.; Burns, N.; Watson, J.; Thornton, B. Digital in-line holography for large-volume analysis of vertical motion of microscale marine plankton and other particles. *IEEE J. Ocean. Eng.* **2021**, *46*, 1248–1260. [[CrossRef](#)]
10. Liu, Z.; Watson, J.; Allen, A. Efficient image preprocessing of digital holograms of marine plankton. *IEEE J. Ocean. Eng.* **2018**, *43*, 83–92. [[CrossRef](#)]
11. Cotter, E.; Fischell, E.; Lavery, A. Computationally efficient processing of in situ underwater digital holograms. *Limnol. Oceanogr. Methods* **2021**, *19*, 476–487. [[CrossRef](#)]
12. Walcutt, N.L.; Knorlein, B.; Cetinic, I.; Ljubecic, Z.; Bosak, S.; Sgouros, T.; Montalbano, A.L.; Neeley, A.; Menden-Deuer, S.; Omand, M.M. Assessment of holographic microscopy for quantifying marine particle size and concentration. *Limnol. Oceanogr. Methods* **2020**, *18*, 516–530. [[CrossRef](#)] [[PubMed](#)]
13. Wang, Y.; Zhang, L.; Ma, W.; Wang, Y.; Niu, W.; Song, Y.; Wang, W. Ocean plankton biomass estimation with a digital holographic underwater glider. *J. Mar. Sci. Eng.* **2022**, *10*, 1258. [[CrossRef](#)]
14. Dyomin, V.; Kurkova, D.; Davydova, A.; Polovtsev, I.; Morgalev, S. Plankton concentration model consistent with natural events and monitoring series of holographic measurements. *J. Mar. Sci. Eng.* **2025**, *13*, 140. [[CrossRef](#)]
15. Guo, B.; Nyman, L.; Nayak, A.R.; Milmore, D.; McFarland, M.; Twardowski, M.S.; Sullivan, J.M.; Yu, J.; Hong, J. Automated plankton classification from holographic imagery with deep convolutional neural networks. *Limnol. Oceanogr. Methods* **2021**, *19*, 21–36. [[CrossRef](#)]
16. Xu, X.; Luo, W.; Ren, Z.; Song, X. Intelligent detection and recognition of marine plankton by digital holography and deep learning. *Sensors* **2025**, *25*, 2325. [[CrossRef](#)] [[PubMed](#)]
17. Hu, W.; Yang, X.; Wang, N.; Zhang, X.; Cui, Y.; Yu, J.; Zheng, H.; Zheng, B. Multi-object reconstruction of plankton digital holograms. *Multimed. Tools Appl.* **2024**, *83*, 51321–51335. [[CrossRef](#)]
18. Awatsuji, Y. Parallel phase-shifting digital holography. In *Multi-Dimensional Imaging*; Wiley: Chichester, UK, 2014; pp. 1–23.
19. Goodman, J.W. *Introduction to Fourier Optics*, 4th ed.; W. H. Freeman and Company: New York, NY, USA, 2017.
20. Sheldon, R.W.; Prakash, A.; Sutcliffe, W.H. The size distribution of particles in the ocean. *Limnol. Oceanogr.* **1972**, *17*, 327–340. [[CrossRef](#)]
21. Stemmann, L.; Boss, E. Plankton and particle size and packaging: From determining optical properties to driving the biological pump. *Annu. Rev. Mar. Sci.* **2012**, *4*, 263–290. [[CrossRef](#)] [[PubMed](#)]
22. Kiørboe, T. *A Mechanistic Approach to Plankton Ecology*; Princeton University Press: Princeton, NJ, USA, 2008.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.