

PAPER • OPEN ACCESS

Snapshot HoloSpec: dispersion-coded 4D feature learning for waterborne particle monitoring

To cite this article: Jingyan Chen *et al* 2026 *J. Phys. Photonics* **8** 015051

View the [article online](#) for updates and enhancements.

You may also like

- [Density Hysteresis and Spin Canting in Rhombohedral Trilayer Graphene](#)
Noelia Fernandez, Tianyi Xu, Ralf Thomas Weitz et al.
- [Green synthesis of *Mentha Spicata*-derived ZnO/Cu₂O hybrid nanomaterials as efficient and stable electrocatalysts for oxygen evolution reaction](#)
Yalamarti Venkat Rajat Rao, Ritu Raj, Abhishek K Soni et al.
- [Data-driven beam diagnostics using betatron radiation](#)
Hossein Saberi, Jasper Burvill, Carys Ann Siddle et al.



PAPER

OPEN ACCESS

RECEIVED
17 October 2025

REVISED
6 January 2026

ACCEPTED FOR PUBLICATION
22 January 2026

PUBLISHED
3 February 2026

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Snapshot HoloSpec: dispersion-coded 4D feature learning for waterborne particle monitoring

Jingyan Chen^{*} , Yaping Zhao[†] , Yuxing Li[†] , Yanmin Zhu[†] , Yuqing Cao[†] and Edmund Y Lam^{*}

Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong, People's Republic of China
^{*} Authors to whom any correspondence should be addressed.

E-mail: u3010336@connect.hku.hk and elam@eee.hku.hk

Keywords: digital holography, deep learning, compressive imaging, microplastic pollutant

Abstract

We present a compact, cost-effective, and highly accurate dispersion-based four-dimensional feature learning (3D spatial and 1D spectral) method for real-time monitoring of water-suspended microparticles, requiring neither sample preparation nor explicit reconstruction. By jointly leveraging digital holography and spectral imaging, the system uses a single prism to compress holograms with transmission spectral information onto an imaging sensor. Coupling an optical model with deep learning, our system recovers the 3D holographic information and class identity for six waterborne particle types, exemplified by microplastics, from a single snapshot. This enables all-in-one reconstruction of both geometry and identity. Both supervised and unsupervised pipelines outperform hologram-only baselines, achieving a macro F_1 of 98.10%. The F_1 score combines precision and recall, offering a comprehensive evaluation of classification performance. With the unsupervised feature-extraction route, our method improves F_1 by 15%, greatly reducing reliance on labeled data. The approach enables long-duration, real-time monitoring in dynamic aquatic environments without sample pretreatment or labeling, and it extends readily from transmission spectroscopy to other spectral modalities.

1. Introduction

Plastic pollution has emerged as one of humanity's most pressing environmental challenges. Global plastic consumption reached 460 million tons in 2019 and is projected to surge to 1.31 billion tons by 2060 [1]. Over time, plastic debris undergoes fragmentation into smaller particles, forming microplastics (MPs, 1 μm –5 mm) and nanoplastics [2]. These particles, alongside natural and anthropogenic particulates such as mineral sediments, organic debris, and aquatic organisms like water fleas, collectively contribute to the complex matrix of water-suspended particles. Within this matrix, MPs are particularly concerning because they degrade slowly, are highly mobile, and have strong adsorption, enabling pervasive distribution across terrestrial, aquatic, and atmospheric ecosystems [3].

MP-biota interactions amplify these risks: MPs adsorb onto cell surfaces, disrupting energy and nutrient transfer [4], while ingestion by organisms ranging from invertebrates to marine mammals impairs physiological functions. Consequently, MP pollution has escalated into a global crisis, demanding urgent advancements in detection methodologies capable of differentiating MPs from co-occurring particles (e.g. mineral grains, organic aggregates, or zooplankton) and quantifying their abundance, morphology, size distribution, and chemical composition. This precise differentiation and classification are critical, as misidentification can lead to significant over- or underestimation of pollution levels. Ultimately, robust data on these parameters is fundamental to understanding the origin, transport routes, and environmental consequences of MP pollution [5].

Conventional spectroscopic techniques, such as Fourier-transform infrared and Raman spectroscopy, remain the gold standard for MP identification [6–8]. However, they suffer from limitations in field applicability, throughput, and multimodal analysis. Laboratory-based approaches require costly equipment, prolonged scanning times, and lack simultaneous spatial-spectral resolution, rendering them

impractical for environmental samples containing thousands of particles [9]. This challenge is especially pronounced in natural waters, where MPs coexist with a diverse array of suspended particulates. Consequently, researchers have highlighted the need for low-cost, high-throughput systems that integrate morphological, spatial, and spectral data to address this complexity, aiming to reduce operational workload, cost, and identification time while maximizing information quality [10–13].

Spectral imaging offers a promising solution by combining digital imaging with spectroscopic analysis to concurrently quantify particle quantity, size, shape, and composition [3, 14, 15]. However, conventional systems trade spatial or spectral resolution for speed, relying on intricate optical components, such as filter arrays and coded apertures, that inflate costs and complexity [16, 17]. Moreover, handling the large volumes of data generated by spectral imaging poses significant challenges. Additionally, traditional intensity-based spectral imaging systems inherently overlook the phase information of incident light, leading to the loss of potentially critical structural and material details of the specimens.

To overcome these limitations, we develop a dispersion-based snapshot hyperspectral imaging system integrated with digital holography. Digital Holography encodes the rich phase information of the object within the intensity pattern of the hologram that contains interference fringes [18, 19]. It enables the simultaneous acquisition of both intensity and phase information using conventional sensors, a capability that proves instrumental for materials monitoring [20–23]. The combination of holography with other techniques, such as polarization imaging [24, 25] or Raman spectroscopy [26], has demonstrated feasibility in multimodal MP identification. However, reliance solely on polarization data and the use of dual sensors with an external spectrometer limit its precision, efficiency, and practicality. It underscores the need for advancements to enhance high-throughput capabilities and enable comprehensive polymer identification.

In this paper, we present a dispersion-based snapshot method combining digital holography with spectroscopy for real-time material monitoring. The system uses a single detector to capture the compressed four-dimensional (4D) signal (3D spatial and 1D spectral) encoded by the simplified optical system called HoloSpec [27]. By analyzing the raw images using a learning-based method, we can obtain 3D spatial information and classification of the sample, enabling the monitoring of material size, number, shape, and chemical composition. This system offers several important advantages, including significant cost savings, enhanced system portability, and the ability to rapidly recover the 3D holographic intensity and class identity of the target object through snapshot computational methods. The system is adaptable for real-time, field use in particle monitoring, and it combines deep learning analysis. This positions it as a novel tool for environmental diagnostics and a practical bridge from the lab to *in situ* applications. By unifying morphological and chemical insights, this platform advances the study of water-suspended particles, prioritizing MPs while accounting for the ecological context of coexisting contaminants.

2. Experiments

2.1. Schematic of snapshot HoloSpec

The HoloSpec setup, as shown in figure 1(a), uses a compact design for multi-dimensional imaging. Two orthogonally aligned light sources, combined by a beam splitter (BS), are directed onto the specimen. The first beam is a 532 nm single-longitudinal-mode continuous-wave laser that forms the lensless in-line digital holography channel [23]; its interference with the unscattered reference wave produces a hologram that encodes the sample's phase information. The second beam originates from a broadband white light (350–800 nm LED) and furnishes a spectral dimension: wavelength-dependent variations in the transmitted intensity provide chemical contrast that augments the classification capability of the system. After passing through the sample, the superimposed optical field is angularly dispersed by a miniature triangular prism (BRP1112, 12.7×12.7 mm), and the resulting spatial-spectral distribution is recorded by a 2D complementary metal-oxide semiconductor (CMOS) sensor (MV-CB120-10UM-C). Figure 1(b) summarizes the proposed workflow: following preprocessing of the raw snapshot, multimodal information of the sample is reconstructed by jointly applying physics-based back-propagation and inference from a trained neural network. The classification prowess of the HoloSpec system arises from its unique dispersive-encoding mechanism coupled with deep learning. The prism maps the sample's broadband transmission spectrum into wavelength dependent spatial shifts on the sensor. Consequently, the material's spectral profile is not recorded as a conventional 1D curve but is embedded as a distinctive 2D spatial-intensity texture. Concurrently, the same snapshot captures the holographic interference pattern, encoding the particle's 3D morphology and refractive-index variations. The neural network is trained to jointly decode these fused cues, learning to associate the composite

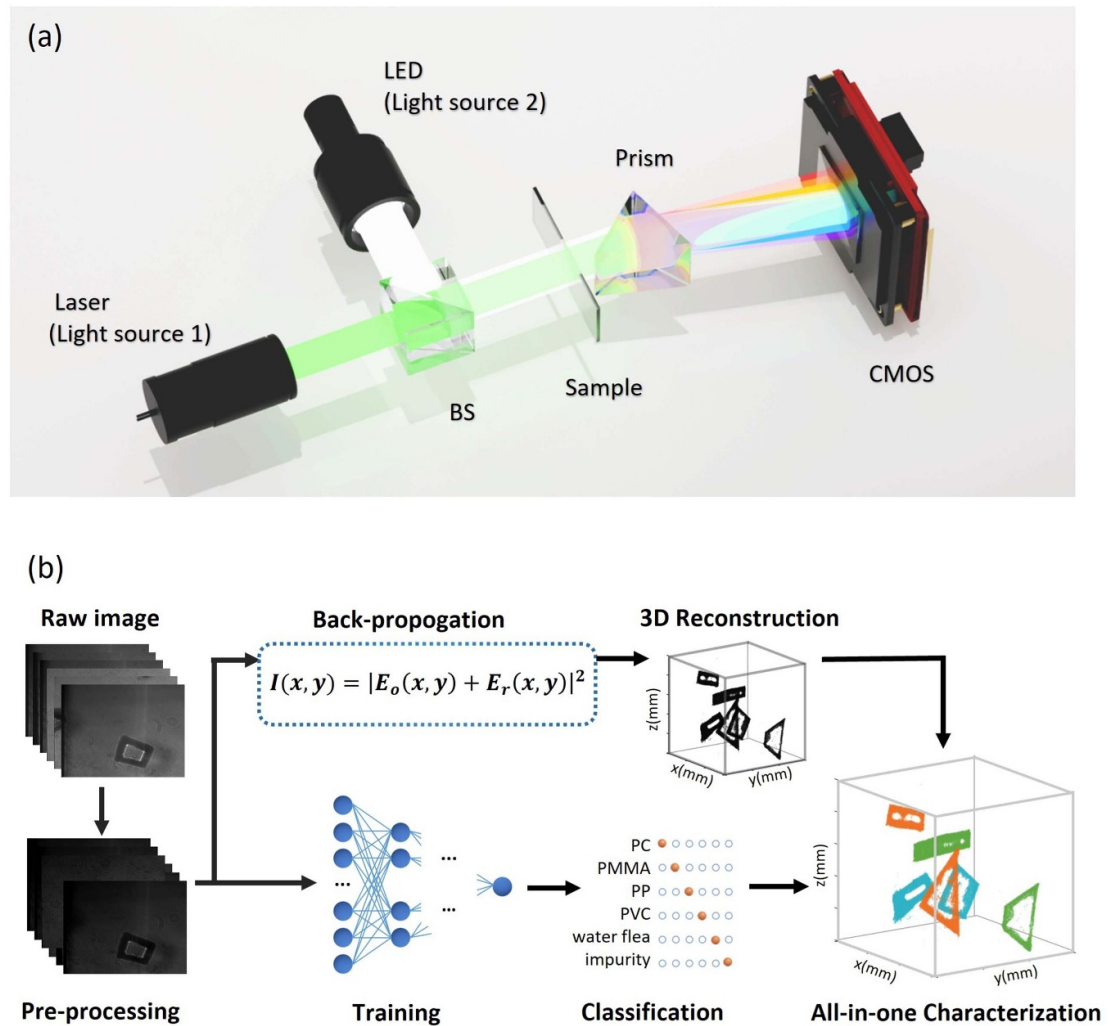


Figure 1. The proposed monitoring method. (a) Experimental setup. A 532 nm single longitudinal mode laser and a broadband white LED light illuminate samples placed on the glass slide at the same time and dispersed by a prism. The dispersion image is captured by a monochromatic CMOS. (b) Schematic diagram of our snapshot HoloSpec system.

patterns that arise from both spectral absorption/scattering and physical structure with specific material identities.

2.2. Samples

Experiments were conducted using well-labeled water fleas, environmental impurities, and MP films, including polymethyl methacrylate (PMMA), polycarbonate (PC), polypropylene (PP), and polyvinyl chloride (PVC). The number of samples prepared for each type is shown in table 1, ranging in size from 1 to 5 mm. The MP films were cut into squares with dimensions ranging from 1 to 5 mm. These MPs represent the most common types of plastics found in the environment, particularly in aquatic ecosystems [28, 29].

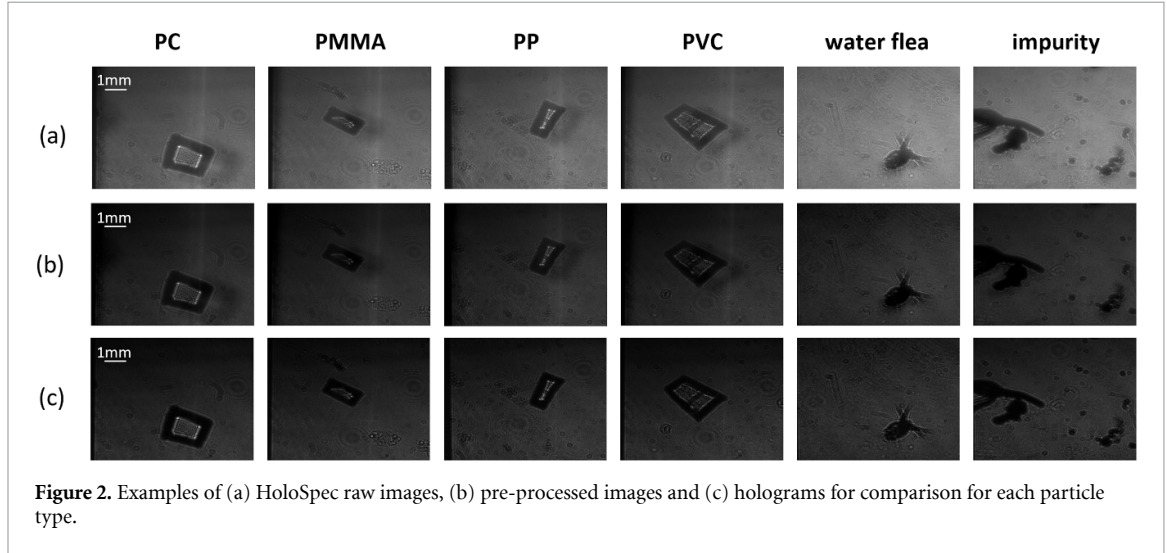
To simulate *in situ* MPs detection in aquatic environments, the experimental samples were supplemented with water fleas and environmental impurities, including gravel, soil particles, and leaf debris. As a highly abundant zooplankton species in freshwater ecosystems, water fleas are widely distributed across various lentic and slow-flowing water bodies [30]. Moreover, their body size range (0.2–5 mm) closely overlaps with that of MPs and environmental impurities.

2.3. Data acquisition and pre-processing

We utilized the described optical system to collect data for all sample types and constructed a well-labeled HoloSpec dataset for neural network training and particle classification. Specifically, a total of 165, 172, 152, 140, 134, and 132 raw images were captured for PC, PMMA, PP, PVC, water flea, and environmental impurity samples, respectively. To enable a comparative evaluation of the proposed

Table 1. Samples used in experiments.

Particle type	Description	Number of samples
Organic carbon	Water flea	20
Impurity	Gravel, soil, leaf	20
Microplastics	PP	30
Microplastics	PC	30
Microplastics	PMMA	30
Microplastics	PVC	30



method, corresponding in-line holographic images (Holo) were also acquired by deactivating the LED light source while maintaining the samples in identical positions.

Representative examples of both HoloSpec and Holo images for each sample type are provided in figures 2(a) and (c) respectively. It can be observed that zooplanktons, such as water fleas, and environmental impurities are relatively easier to classify due to their distinct morphological features and optical transmission properties. However, the four types of MP particles have no significant visual differences, making their classification very challenging.

In the HoloSpec imaging mode, white-light illumination introduces background signals that elevate overall image brightness, thereby reducing the sharpness of interference fringes and impairing image clarity. To address these effects and ensure a fair comparative assessment, all images underwent a pre-processing procedure in which brightness and contrast were normalized within each category. This adjustment harmonizes intensity distributions across images and mitigates illumination-related artifacts, promoting consistency and robustness in downstream analysis. Specifically, each HoloSpec frame was linearly normalized to match the global luminance statistics of its corresponding Holo image pair. We compute the mean and standard deviation of pixel intensities for the source (HoloSpec) and reference (Holo) images, and the brightness and contrast of the source image are adjusted via the formula:

$$I'_s(x, y) = \frac{I_s(x, y) - \mu_s}{\sigma_s + 10^{-8}} \sigma_r + \mu_r, \quad (1)$$

where I_s and I_r represent the pixel intensity arrays of the source images and reference images, while μ and σ are their respective means and standard deviations. The pre-processed images I'_s are used for the later training. Examples of pre-processed images are shown in figure 2(b).

This method ensures robust pre-processing of images, enabling more accurate analysis in environments with variable lighting conditions.

2.4. Joint characterization

This method facilitates the reverse propagation of HoloSpec images containing dispersive holograms using the optical system model in conjunction with the angular spectrum formula. The retrieved optical field information enables the 3D reconstruction of samples, providing a comprehensive characterization of their spatial features.

Additionally, raw images can be directly classified through a neural network, further obtaining the 1D class information. To validate the effectiveness of our method, we employed both supervised and unsupervised learning strategies to classify and analyze water-suspended particles in the HoloSpec and conventional Holo datasets. In this study, the feature extraction strategies for supervised and unsupervised learning were designed based on the specific methodological requirements of each approach. For supervised learning, we trained a multilayer perceptron (MLP) classification model using labeled data to quantitatively evaluate the classification performance across different datasets. For unsupervised learning, approaches such as k -means clustering were used to explore the intrinsic feature distributions within the data, demonstrating the capability of HoloSpec's spectral information to automatically distinguish unknown particles. On a standard consumer-grade GPU (NVIDIA GeForce RTX 4090), the average processing time per image is approximately 1.1 s, enabling near real-time monitoring capabilities in field applications.

2.4.1. Supervised learning

For supervised learning, we used a two-stage pipeline to classify six types of microparticles: PC, PMMA, PP, PVC, impurities, and water fleas, based on the preprocessed HoloSpec and Holo datasets. The pipeline consisted of a fixed deep feature extractor followed by a lightweight classifier.

We employed a ResNet50 [31] model pre-trained on ImageNet as a frozen backbone to extract high-dimensional features. Specifically, the global average pooling layer of ResNet50 was used to generate 2048-dimensional feature representations for each image without any fine-tuning. This approach retained the strong representational power of the pre-trained model while focusing on feature selection for the classification task. The extracted features were subsequently fed into an MLP classifier implemented in scikit-learn, which consisted of two hidden layers with 256 and 128 units, respectively, and utilized ReLU activations. While deep learning has expanded into diverse architectures [32], MLPs remain widely adopted for high-dimensional feature classification due to their interpretability and efficiency [33]. The MLP was trained using the Adam optimizer with a learning rate of 0.001 and early stopping to prevent overfitting. The training process was capped at 50–100 epochs, with a fixed random seed to ensure reproducibility.

Model evaluation was conducted using stratified five-fold cross-validation with shuffled splits. To standardize feature scales and prevent data leakage, a standard scaler was fit on the training partition of each fold and applied to both the training and test data within that fold. For each fold, class probabilities and predictions were generated on the held-out split, and the model with the best performance on the test set was retained. This workflow not only maximized the retention of fine-grained features essential for microparticle classification but also leveraged the regularization mechanisms in the supervised learning framework to mitigate overfitting risks associated with high-dimensional features.

2.4.2. Unsupervised learning

For unsupervised learning, we employed AlexNet [34] and ResNet18 as pre-trained convolutional neural network (CNN) feature extractors, followed by k -means clustering to analyze the intrinsic structure of the data. These architectures were selected to extract low-dimensional features suitable for clustering, considering the sensitivity of clustering algorithms to high-dimensional data. By reducing the feature dimensionality, we mitigated the 'curse of dimensionality', enhanced computational efficiency, and leveraged shallow network features to better reveal the underlying clustering structure of the data.

The AlexNet model was truncated to retain only its convolutional subnetwork, consisting of five convolutional layers interleaved with ReLU activations and max-pooling layers. The final convolutional feature map was processed using adaptive average pooling to compress it into a 1×1 spatial resolution, resulting in a 256-dimensional feature vector for each image. For ResNet18, the network was modified by removing the dense layers, retaining eight residual blocks, each containing two 3×3 convolutional layers with identity shortcut connections. The output of the last residual block, a 512-channel activation map, was subsequently reduced into a 512-dimensional feature vector using global adaptive average pooling.

All images were resized to 512×512 pixels prior to feature extraction. Using the frozen weights of the pre-trained networks, each image was transformed into a high-dimensional feature vector. These feature vectors were then clustered using k -means, with the number of clusters $k = 6$, corresponding to the six target categories: PC, PMMA, PP, PVC, impurities, and water fleas. To quantitatively evaluate the clustering results, the Hungarian algorithm was applied to optimally match the clustering outputs with ground-truth labels.

To ensure reproducibility, a fixed random seed of 42 was used throughout the analysis. Additionally, a batch size of 32 was adopted for processing, balancing computational efficiency and memory usage.

This approach effectively demonstrated the ability of shallow and compact feature representations to uncover the data's inherent clustering structures while maintaining computational feasibility.

2.4.3. Dispersive hologram reconstruction

During holographic reconstruction, the spectrally dispersed signal has a negligible effect. First, the optical power of the 532 nm laser used for holography significantly outweighs that of the broadband LED, ensuring that the recorded interference fringes dominate the intensity pattern. Second, due to the linear dispersion of the prism, the spectral information is smeared into a smooth, low frequency background across the sensor; it does not introduce high spatial frequency modulations that would corrupt the fine fringe structure essential for phase retrieval. Third, the preprocessing step (equation (1)) actively suppresses this uniform background and enhances fringe contrast. Therefore, the dispersive holographic interference fringes captured by the CMOS sensor can be expressed as:

$$I_s(x, y - \delta y) = |E_o(x, y) + E_r(x, y)|^2, \quad (2)$$

where $I_s(x, y)$ represents the intensity distribution of the hologram recorded on the imaging plane with coordinates (x, y) . Here, δy denotes the dispersion displacement caused by prism, E_o is the diffracted wave of the object, while E_r corresponds to the un-diffracted plane reference wave.

After deriving $I_s(x, y - \delta y)$ from the modeled optical system based on $I_s(x, y)$ and pre-processing mentioned above, the hologram can be numerically back-propagated to a specific reconstruction plane with coordinates (x, y, z) , allowing the complex optical field of the target object, $\Gamma(x, y, z)$, to be retrieved. In this paper, we adopt the angular spectrum algorithm [19, 35], a widely-used holographic reconstruction method, to reconstruct the optical field. It is mathematically described as:

$$\Gamma(x, y, z) = \mathcal{F}^{-1} \left\{ \mathcal{F} \{ I_s'(x, y - \delta y) \} \times \exp \left(-i \frac{2\pi z}{\lambda} \sqrt{1 - \lambda^2 f_x^2 - \lambda^2 f_y^2} \right) \right\}, \quad (3)$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Fourier transform and its inverse, respectively, and f_x and f_y are the transverse spatial frequencies. From the retrieved complex optical field $\Gamma(x, y, z)$, both the amplitude $I = |\Gamma|^2$ and the phase $\varphi(x, y, z) = \arg(\Gamma)$ can be obtained. The phase is calculated as $\varphi = \tan^{-1} \frac{\text{Im}(\Gamma)}{\text{Re}(\Gamma)}$, where 'Im' and 'Re' denote the imaginary and real parts, respectively. The phase map directly encodes the optical path length difference induced by the specimen, thereby providing quantitative information about its three dimensional structure and refractive index distribution. In this way, we can reconstruct the 3D spatial information of the specimen.

3. Results and discussion

3.1. Supervised classification

To validate the material recognition capability of our system, we conducted supervised classification experiments on both the HoloSpec and Holo datasets using a ResNet50-based pipeline and performed the same procedure on the Holo dataset for comparison.

Figure 3 compares the results between the two datasets across three metrics: (a) feature-wise significance, (b) intra-class Euclidean distances, and (c) confusion matrices. In figure 3(a), HoloSpec has higher feature significance values in the leading dimensions, suggesting stronger alignment between the extracted ResNet50 features and class labels, which enhances discriminability. The intra-class distance analysis (figure 3(b)) reveals that HoloSpec features are more compact, with lower median distances and tighter interquartile ranges across most classes, though Holo shows marginally shorter extreme tails in some cases. This indicates better intra-class cohesion for HoloSpec. Besides, the confusion matrices in figure 3(c), generated using the ResNet50-MLP pipeline, demonstrate superior performance for HoloSpec, with fewer off-diagonal misclassifications and higher per-class precision and recall. These results suggest that ResNet50 effectively captures a small set of highly discriminative features, enabling rapid overall model convergence.

To further evaluate and quantify the classification performance of the supervised network on different sample types, we calculated the macro F_1 score [36], as well as the precision, recall, and F_1 score for each class in the HoloSpec dataset (as shown in table 2). The same analysis was performed on the Holo dataset for comparison. The results reveal that our system always has a better classification performance, especially for samples with indistinct shape features, such as MP particles. It is noteworthy that the system remains effective for semi transparent or opaque samples, such as water fleas and environmental impurities. Although these specimens may exhibit weak spectral transmission in the visible band, the in-line holographic geometry records their scattered wavefronts, producing interference patterns that

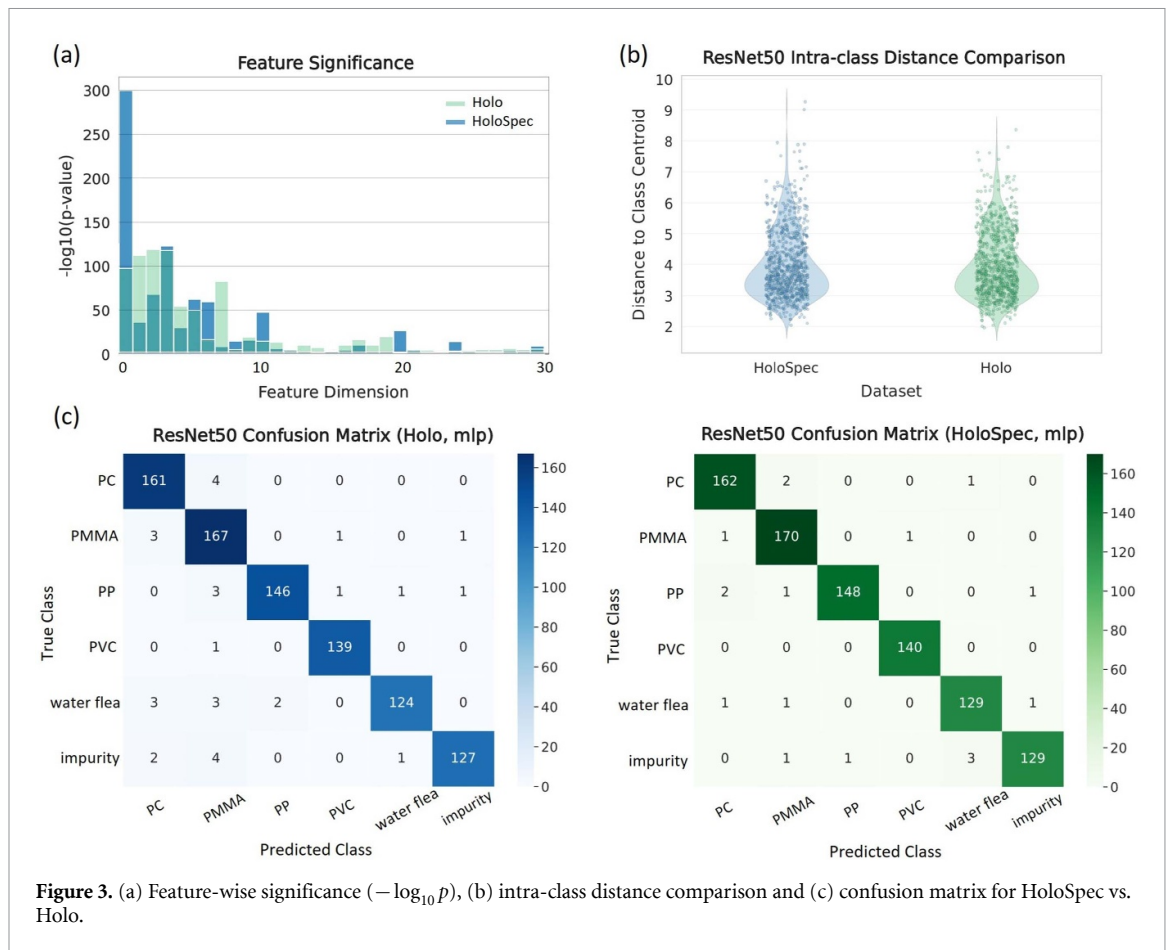


Figure 3. (a) Feature-wise significance ($-\log_{10} p$), (b) intra-class distance comparison and (c) confusion matrix for HoloSpec vs. Holo.

Table 2. Supervised clustering performance (precision, recall, and F_1 score) of ResNet50 for each class on HoloSpec and Holo datasets.

Class	HoloSpec			Holo		
	$\mathcal{P}$	$\mathcal{R}$	F_1	$\mathcal{P}$	$\mathcal{R}$	F_1
PC	0.9759	0.9818	0.9789	0.9527	0.9758	0.9641
PMMA	0.9714	0.9884	0.9798	0.9176	0.9709	0.9435
PP	0.9933	0.9737	0.9834	0.9865	0.9605	0.9733
PVC	0.9929	1.0000	0.9964	0.9858	0.9929	0.9893
impurity	0.9699	0.9699	0.9699	0.9841	0.9394	0.9612
Waterflea	0.9847	0.9699	0.9773	0.9845	0.9478	0.9658
Macro F_1		0.9810			0.9659	

encode distinct morphological features (e.g. sharp edges, surface texture, and shadowing effects). These holographically captured morphological signatures provide robust discriminative information, enabling accurate classification even in the absence of strong spectral contrast. Quantitatively, overall accuracy rises from 96.54% (864/895) on Holo to 98.10% (878/895) on HoloSpec, and macro F_1 improves from 96.59% to 98.10%. Taken together, these converging results indicate that HoloSpec yields more separable representations and stronger downstream classification performance than Holo.

3.2. Unsupervised classification

We further selected two classic well-trained CNN architectures, AlexNet and ResNet18, to perform feature extraction on both datasets, followed by k -means clustering for unsupervised classification. A series of visual analyzes were conducted on the clustering results. The t-SNE projections [37] shown in figure 4(a) reveal that the clusters formed by HoloSpec samples in the embedded space are more distinct and well-separated compared to Holo samples. Further feature-significance analysis (figure 4(b)) indicates that, on the HoloSpec dataset, both CNN models extract a greater number of high-amplitude, statistically significant features, demonstrating that multi-channel inputs facilitate the capture of more

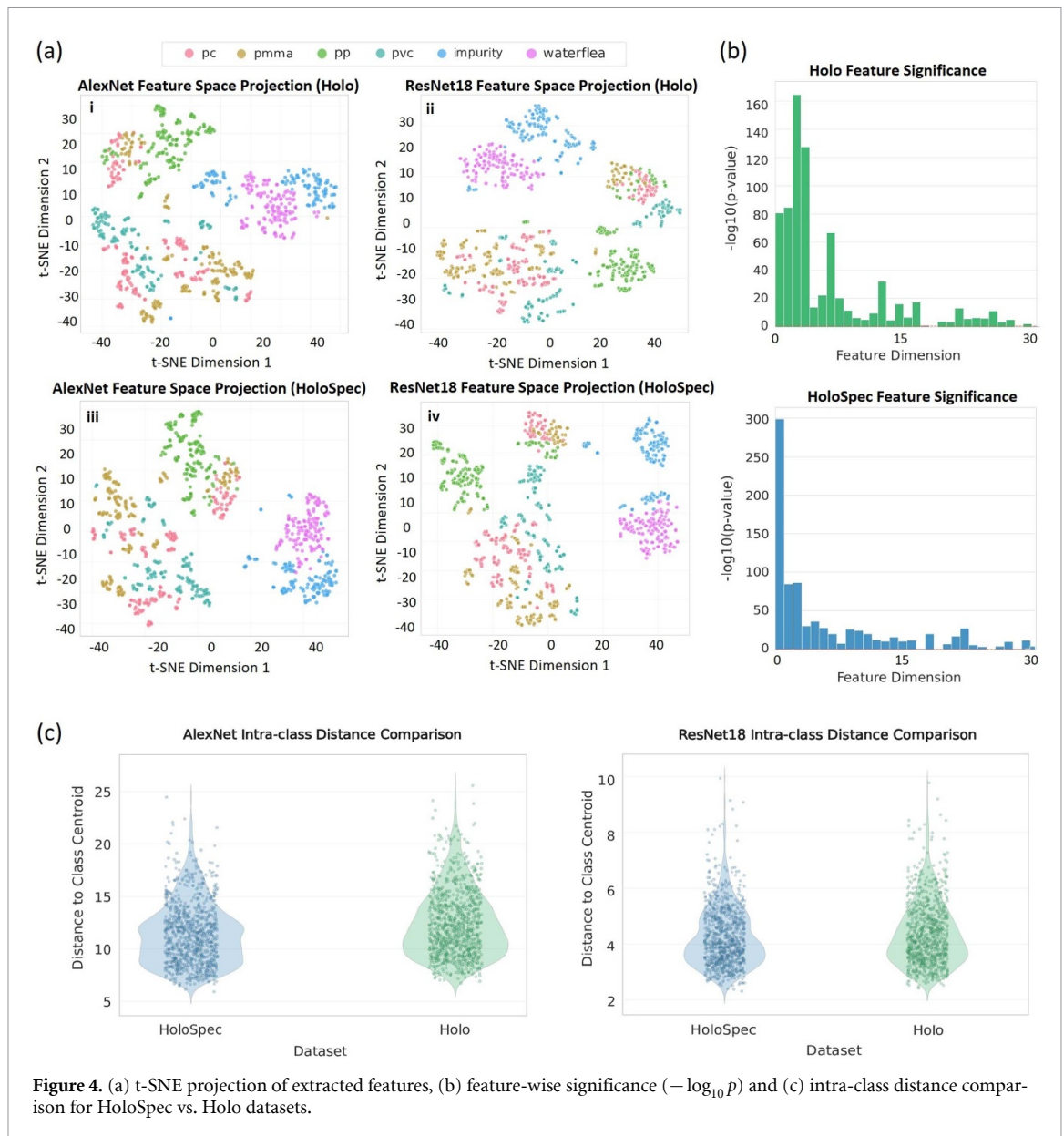


Figure 4. (a) t-SNE projection of extracted features, (b) feature-wise significance ($-\log_{10} p$) and (c) intra-class distance comparison for HoloSpec vs. Holo datasets.

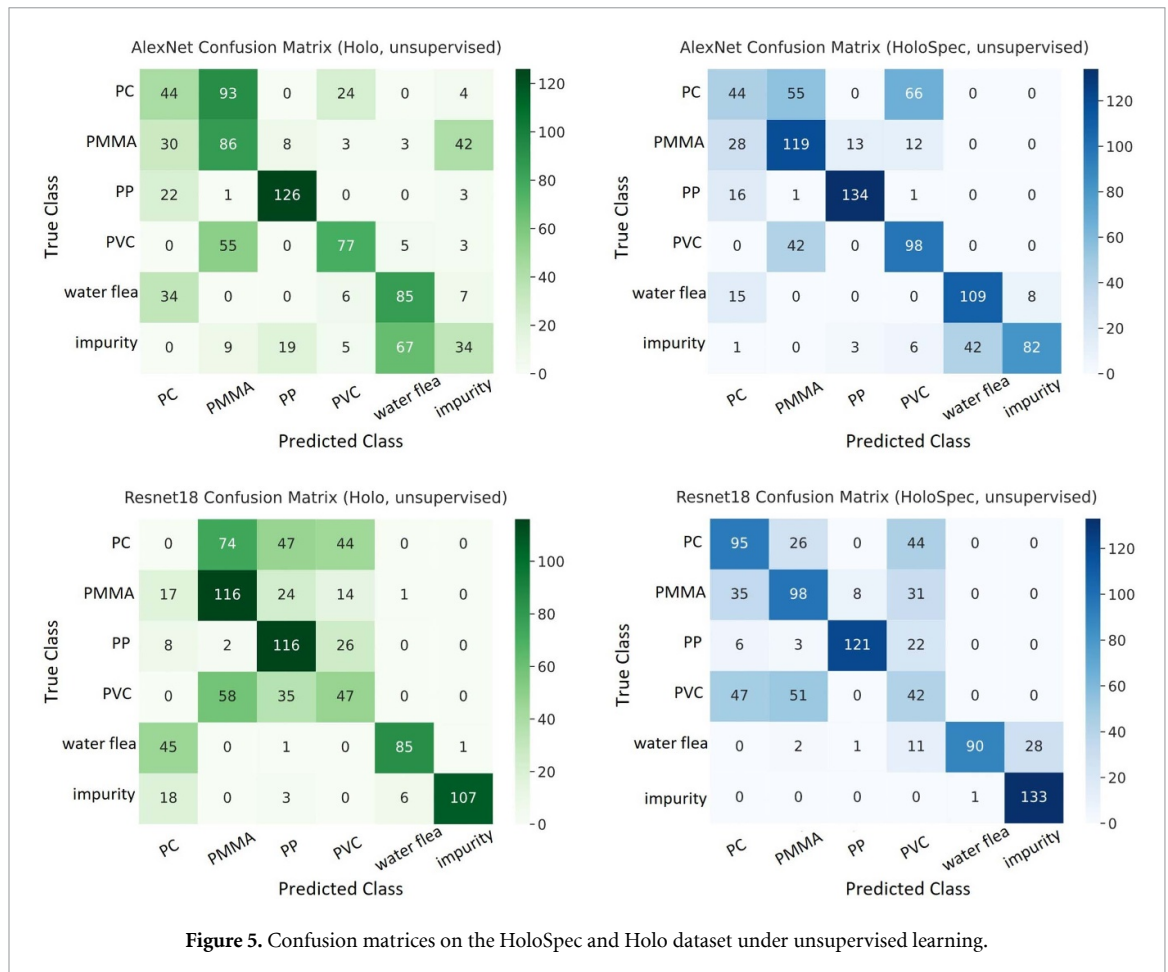
discriminative representations. In contrast, under identical significance thresholds, the Holo dataset yields substantially fewer significant features, highlighting its limitations in feature expressiveness.

The compactness analysis of the feature space further supports this conclusion (figure 4(c)). Across both AlexNet and ResNet-18, the intra-class feature distributions learned from HoloSpec are consistently more compact than those learned from the original Holo dataset. In the violin plots, the median distance to the class centroid for HoloSpec is roughly 10% lower, the inter-quartile range is 15%–20% narrower, and the long-tail extent is markedly reduced, indicating fewer hard or aberrant samples. Taken together, these results show that HoloSpec offers a cleaner class structure and better separability, making it a more suitable dataset for representation-learning tasks.

Tables 3 provides a detailed evaluation of clustering performance metrics. Using the same clustering procedure, the macro F_1 scores for HoloSpec reached 65.62% (AlexNet) and 65.91% (ResNet18), representing improvements of 15.22% and 13.33%, respectively, over the Holo dataset scores of 50.40% and 52.58%. Furthermore, the per-class F_1 scores for all six categories in HoloSpec are consistently higher than those in Holo. The corresponding confusion matrices (figure 5) demonstrate that HoloSpec achieves 15–25 additional correctly classified samples on the diagonal, with significantly fewer misclassifications, further validating its superior performance in material classification. These results conclusively demonstrate that HoloSpec not only generates more discriminative features and effectively compresses intra-class variance but also significantly improves clustering accuracy. It is important to clarify that common MPs do not exhibit sharp, fingerprint like absorption lines in the visible range. Instead,

Table 3. Comparison of F_1 scores for HoloSpec and Holo datasets using AlexNet and ResNet18 features.

	HoloSpec		Holo	
	AlexNet	ResNet18	AlexNet	ResNet18
Macro F_1	0.6562	0.6591	0.5040	0.5258
PC	0.3260	0.5456	0.2980	0.000
PMMA	0.6115	0.5568	0.4133	0.5518
PP	0.8874	0.8532	0.8262	0.6134
PVC	0.6076	0.2885	0.6040	0.3444
Impurity	0.7704	0.8055	0.5822	0.7595
Waterflea	0.7344	0.9017	0.3002	0.8855

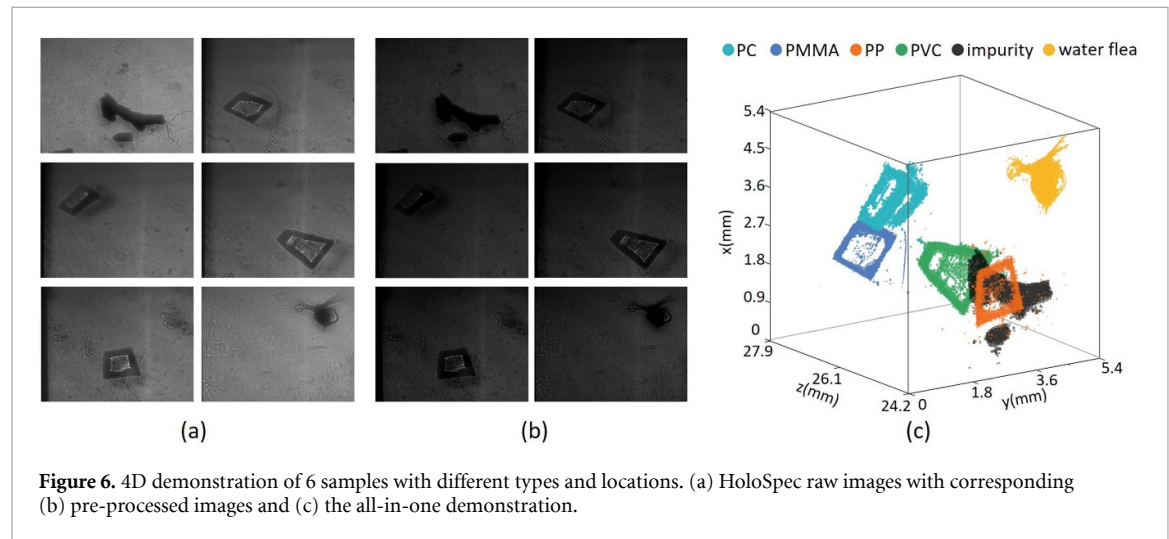


our method leverages their subtle, material specific variations in broadband absorption and scattering profiles. These continuous spectral contrasts are optically encoded into 2D spatial intensity patterns via dispersion. Crucially, this spectral information is analyzed in conjunction with the rich morphological and refractive index data simultaneously captured in the holographic channel. This dual modality fusion provides the comprehensive feature set that enables the system to accurately differentiate between particle types, as demonstrated by its superior performance over Holo baselines.

To provide a more intuitive and quantitative comparison of the performance of the two datasets across supervised and unsupervised learning using three different network architectures, we present the overall accuracy ($\mathcal{A}$), precision ($\mathcal{P}$), recall ($\mathcal{R}$), and F_1 scores in table 4. It can be observed that regardless of the network model used for training and testing, HoloSpec consistently demonstrates superior accuracy, precision, recall, and F_1 scores. Moreover, the metrics are stable, indicating strong robustness. Notably, for unsupervised feature learning, HoloSpec achieves a significant performance improvement exceeding 10%.

Table 4. Classification performances of unsupervised learning Alexnet, unsupervised learning ResNet18, and supervised learning ResNet50.

Methods	AlexNet		ResNet18		ResNet50	
	HoloSpec	Holo	HoloSpec	Holo	HoloSpec	Holo
$\mathcal{A}$	0.6559	0.5050	0.6469	0.5263	0.9810	0.9654
$\mathcal{P}$	0.6722	0.5135	0.6466	0.5409	0.9814	0.9685
$\mathcal{R}$	0.6630	0.5072	0.6516	0.5359	0.9806	0.9645
F_1	0.6562	0.5040	0.6591	0.5258	0.9810	0.9659

**Figure 6.** 4D demonstration of 6 samples with different types and locations. (a) HoloSpec raw images with corresponding (b) pre-processed images and (c) the all-in-one demonstration.

3.3. All-in-one demonstration

To demonstrate the multidimensional monitoring capability of the snapshot HoloSpec system, we performed comprehensive reconstructions of complex objects, encompassing all six sample types used in our experiments. As shown in figure 6(a), the raw monochromatic images include representative particles from PC, PP, PVC, PMMA, impurity, and water flea. Each type of sample was imaged separately, and the resulting datasets were reconstructed together within a unified 4D framework. In this demonstration, we utilized a supervised neural network for classification: the preprocessed images (figure 6(b)) were fed into a trained MLP model, which assigned each image to the most probable category. Subsequently, the angular spectrum method was used to back-propagate the recorded dispersed holograms and retrieve the corresponding three-dimensional spatial distributions. As illustrated in figure 6(c), the system accurately classified all six sample types and resolved their spatial structures within the reconstructed volume. These results highlight the ability of the HoloSpec system to simultaneously achieve high-precision classification and precise 3D localization of diverse microparticles, underscoring its effectiveness for advanced multidimensional sample analysis. This capability enables critical downstream analyzes beyond classification, such as particle localization, size distribution statistics, and tracking in dynamic suspensions, which are essential for comprehensive environmental monitoring.

4. Conclusion

This snapshot HoloSpec system, characterized by its low cost, compact size, and real-time high-throughput imaging capability, overcomes the limitations of conventional spectral scanning and digital holography techniques. By integrating deep learning, it offers a high-performance solution for classifying MPs and other suspended aquatic particles, thereby providing a practical tool for environmental monitoring of MP pollution.

Utilizing a single ordinary glass prism, the system compresses holographic and spectral information of samples into a single monochromatic image and enables label-free 4D monitoring of samples without any pretreatment. Our system leverages subtle differences of samples in broadband absorption and scattering profiles, which are dispersion-encoded into spatial intensity patterns. These patterns, when fused with morphological and refractive index information encoded in the hologram, provide sufficient discriminative power for accurate classification. In comparative evaluations, data acquired from the HoloSpec system consistently outperformed traditional digital holography across all tested neural

networks. Notably, in unsupervised learning settings, the system achieved a significant improvement in the F_1 score of approximately 15%. This underscores its particular advantage in reducing dependence on large labeled datasets. By leveraging the network to decode spectrally-encoded information compressed via the prism, the system enhances the feasibility of deployment in portable devices for real-time MP pollution surveillance. Furthermore, its adaptability supports effective operation in dynamically changing aquatic environments without requiring recurrent data annotation.

The dispersion rate, determined by the prism's material and geometry, defines the spatial density of the encoded spectral information. In this work, we used a fixed prism with a calibrated dispersion. While a different dispersion rate would alter the spatial spectral mapping, the end-to-end trained CNN learns features that are inherently adapted to the specific encoding presented during training. Therefore, the system is robust to the fixed configuration used. For future adaptation using different dispersive elements, the network would require retraining on data acquired with the new optical setup to learn the corresponding feature representation.

Although our experimental validation employed absorption spectroscopy within the visible range, the methodology is readily adaptable to other spectroscopic modalities such as Raman spectroscopy, broadening its application potential. Most importantly, this system provides a robust and scalable approach for monitoring MP contamination in real field conditions, contributing meaningfully to the assessment of pollution sources, pathways, and ecological impacts.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://1drv.ms/f/c/bd28517afbdd485f/EuizNcIfygZOpGLW_JbIEIEBTm8qxtwd7zHNIrjm5bj2lg?e=gEK36s [38].

Acknowledgments

We are grateful to all members of the Imaging Systems Laboratory for their support and valuable suggestions.

Funding

The work was supported by the Research Grants Council of Hong Kong (GRF 17 201 822 and RIF R7003-21).

Roles

J C conceived the project, conducted the experiments, analyzed the data, and drafted the paper. Y Z trained and evaluated network architectures. Y L prepared the sample and reviewed the paper. Y Z prepared the sample and assisted in designing the experiments. Y C trained network architectures and helped with visualization. E Y L conceived and supervised the project, and reviewed and edited the paper.

ORCID iDs

Jingyan Chen  0009-0003-1204-7746
Yaping Zhao  0000-0002-3482-0942
Yuxing Li  0009-0004-8128-1517
Yanmin Zhu  0000-0002-3968-5372
Yuqing Cao  0000-0002-2147-5860
Edmund Y Lam  0000-0001-6268-950X

References

- [1] OECD 2022 Global plastics outlook: Policy scenarios to 2060 (Organisation for Economic Co-operation and Development)
- [2] Frias J P and Nash R 2019 *Mar. Pollut. Bull.* **138** 145–7
- [3] Xu J L, Thomas K V, Luo Z and Gowen A A 2019 *TRAC Trends Anal. Chem.* **119** 115629
- [4] Ye Y, Yu K and Zhao Y 2022 *Sci. Total Environ.* **818** 151851
- [5] Hidalgo-Ruz V, Gutow L, Thompson R C and Thiel M 2012 *Environ. Sci. Technol.* **46** 3060–75
- [6] Iri A H, Shahrah M H, Ali A M, Qadri S A, Erdem T, Ozdur I T and Icoz K 2021 *Environ. Sci. Pollut. Res.* **28** 1–7
- [7] Beichert L, Binhammer Y, Andrade J R, Mevert R, Kniggendorf A K, Roth B and Morgner U 2021 *Opt. Express* **29** 31499–507
- [8] Lei B, Bissonnette J R, Hogan U E, Bec A E, Feng X and Smith R D 2022 *Anal. Chem.* **94** 17011–9

- [9] Primpke S, Lorenz C, Rascher-Friesenhausen R and Gerdt G 2017 *Anal. Methods* **9** 1499–511
- [10] Chaudhry A K and Sachdeva P 2021 *Reg. Stud. Mar. Sci.* **48** 102018
- [11] Neves C V, Gaylarde C C, Neto J A B, Vieira K S, Pierri B, Waite C C, Scott D C and da Fonseca E M 2022 *Water Biol. Secur.* **1** 100080
- [12] Anik A H, Hossain S, Alam M, Sultan M B, Hasnine M T and Rahman M M 2021 *Environ. Nanotechnol., Monito. Manag.* **16** 100530
- [13] Tammina S K, Khan A and Rhim J W 2023 *Chemosphere* **313** 137433
- [14] Serranti S, Palmieri R, Bonifazi G and C  zar A 2018 *Waste Manage.* **76** 117–25
- [15] Chen J, Li Y, Huang J and Lam E Y 2024 Wide-field microplastic identification based on spectrum and deep learning *Imaging Systems and Applications* (Optica Publishing Group) p IM1G–6
- [16] Berns R S, Taplin L A, Nezamabadi M, Mohammadi M and Zhao Y 2005 Spectral imaging using a commercial color-filter array digital camera *Proc. 14th Triennial Meeting The Hague, ICOM Committee for Conservation (Citeseer)* pp 743–50
- [17] Arguello H and Arce G R 2012 *IEEE Trans. Image Process.* **22** 941–54
- [18] Ren Z, Zeng T and Lam E Y 2019 Digital Holographic Imaging via Deep Learning *Computational Optical Sensing and Imaging* (Optica Publishing Group) p CTu3A–4
- [19] Goodman J W 2005 *Introduction to Fourier Optics* (Roberts and Company Publishers)
- [20] Kim M K 2010 *SPIE Rev.* **1** 018005
- [21] Tahara T, Quan X, Otani R, Takaki Y and Matoba O 2018 *Microscopy* **67** 55–67
- [22] Paturzo M, Merola F, Bianco V, Memmolo P, Miccio L, Mandracchia B and Ferraro P 2018 Detection and sorting of microplastics in the marine environment by new imaging tools *Speckle 2018: VII Int. Conf. on Speckle Metrology* vol 10834 (SPIE) pp 186–90
- [23] Zhu Y, Yeung C H and Lam E Y 2021 *J. Phys. Photon.* **3** 024013
- [24] Huang J, Zhu Y, Li Y and Lam E Y 2023 *ACS Photonics* **10** 4483–93
- [25] Zhu Y, Li Y, Huang J and Lam E Y 2024 *Commun. Eng.* **3** 32
- [26] Takahashi T et al 2023 *Opt. Express* **31** 7492–504
- [27] Chen J, Li Y, Zhu Y and Lam E Y *Opto-Electron. Adv.* (in review)
- [28] Andrad   A L and Neal M A 2009 *Phil. Trans. R. Soc. B* **364** 1977–84
- [29] Engler R E 2012 *Environ. Sci. Technol.* **46** 12302–15
- [30] Allan J D 1976 *Am. Naturalist* **110** 165–80
- [31] He K, Zhang X, Ren S and Sun J 2016 *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* pp 770–8
- [32] LeCun Y, Bengio Y and Hinton G 2015 *Nature* **521** 436–44
- [33] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (MIT Press)
- [34] Krizhevsky A, Sutskever I and Hinton G E 2012 *Advances in Neural Information Processing Systems* vol 25
- [35] Huang J, Zhu S, Li Y, Wang C and Lam E Y 2025 *APL Photonics* **10** 036114
- [36] Opitz J and Burst S 2019 arXiv:1911.03347
- [37] Maaten L v d and Hinton G 2008 *J. Mach. Learn. Res.* **9** 2579–605
- [38] Chen J 2025 Dataset of HoloSpec and Holo (available at: https://1drv.ms/f/c/bd28517afbddd485f/EuizNcIfygZOpGLW_JblEIEBTm8qxtwd7zHNiRjm5bj2lg?e=gEK36s)