Simulating Marine Snow Images: Pipeline, Data Set, and Benchmark

Yiqing Huang, Tianshun Han, Haoru Zhao, Yanyan Liang, *Member, IEEE*, Jun Wan, *Senior Member, IEEE*, Sergio Escalera, *Senior Member, IEEE*, and Haiyong Zheng, *Senior Member, IEEE*

***Abstract*—Marine snow causes light spots and veil-like blurs in underwater images due to scattering and degrades their quality. Therefore, removing this type of noise from underwater images is necessary. However, the lack of valid and diverse marine snow image data sets has impeded the development of machine learning-based algorithms. To tackle this issue, we propose a pipeline named *OmniMSI* for generating marine snow images. This pipeline takes into account the diverse morphologies of marine snow particles and their natural distribution patterns. Specifically, *OmniMSI* initiates from collecting large-scale marine snow particle images in natural environments, classifying them by morphology, analyzing their distribution patterns, and thus creating diverse marine snow masks. The pipeline then overlays the masks on real-world underwater images without marine snow and employs the designed post-process to synthesize marine snow images. Leveraging this pipeline, we create a *M*ulti-morphology and *M*ultidistribution *M*arine *S*now image *D*ataset (*MM-MSD*), including 6000 pairs of marine snow images and corresponding reference images. We also present a novel marine snow removal method—*m*arine *s*now-*a*ware *diff*usion, which can accurately estimate the location of marine snow in images and remove it progressively. Based on *MM-MSD*, we conduct a comprehensive study of the state-of-the-art algorithms qualitatively and quantitatively, developing a new benchmark for marine snow removal. The benchmark evaluations demonstrate the performance and shortcomings of existing algorithms, paving the way for further research in marine snow removal. The data set and code will be released.**

Received 16 July 2024; revised 24 December 2024 and 7 April 2025; accepted 27 May 2025. Date of publication 19 August 2025; date of current version 27 October 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62171421 and in part by the TaiShan Scholar Youth Expert Program of Shandong Province under Grant tsqn202306096. *(Yiqing Huang, Tianshun Han, and Haoru Zhao are co-first authors.) (Corresponding author: Haiyong Zheng.)*

Associate Editor: H. Li.

Yiqing Huang, Tianshun Han, and Yanyan Liang are with the Faculty of Innovation Engineering, Macau University of Science and Technology, Taipa 999078, China (e-mail: 3230006115@student.must.edu.mo; 3230002542@student.must.edu.mo; yyliang@must.edu.mo).

Haoru Zhao is with the College of Economic Crime Investigation, Shandong Police College, Jinan 250200, China, and also with the College of Electronic Engineering, Ocean University of China, Qingdao 266404, China (e-mail: zhaohaoru@sdpc.edu.cn).

Jun Wan is with the State Key Laboratory of Multimodal Artificial Intelligence Systems (MAIS), Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190, China, and with the School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS), Beijing 100049, China, and also with the Macau University of Science and Technology, Taipa 999078, China (e-mail: jun.wan@ia.ac.cn).

Sergio Escalera is with the Computer Vision Center, 08193 Barcelona, Spain, and with the Universitat de Barcelona, 08007 Barcelona, Spain, and also with the Aalborg University, 9220 Aalborg, Denmark (e-mail: sergio@maia.ub.es).

Haiyong Zheng is with the College of Electronic Engineering, Ocean University of China, Qingdao 266404, China (e-mail: zhenghaiyong@ouc.edu.cn).

Digital Object Identifier 10.1109/JOE.2025.3590094

***Index Terms*—Image simulation, machine learning, marine snow, marine snow image, marine snow removal.**

I. Introduction

MARINE snow is comprised of multimorphology particles that aggregate from organic and inorganic impurities present in water bodies [1]. As shown in Fig. 1, it typically appears as light spots or veil-like blurs in underwater images [2], a phenomenon collectively called the "marine snow effects" [3], [4]. Underwater images with marine snow usually suffer from noticeable visibility degradation and object occlusion, hindering the development of a range of underwater visual tasks (e.g., underwater object detection [5] and underwater target tracking [6], [7]). Therefore, the research on marine snow removal is crucial for improving the quality of underwater images and the performance of related underwater visual tasks.

Deep learning provides an effective method to remove marine snow from images. However, the performance of such data-dependent methods is affected by the training data sets. This bottleneck arises primarily from the scarcity of valid and diverse marine snow image data sets. As shown in Table I, there are several shortcomings of the existing marine snow image data sets as follows.

1) *Oversimplifying marine snow effects:* Initial attempts [3], [4], [8] to represent the marine snow effects in data sets often simplify them to noise by directly adding noise points to underwater images. Such simplified data sets fail to present the nuanced appearance of the marine snow effects, causing algorithms trained on them to struggle with processing real-world images due to the complexity of these effects.
2) *Underestimating marine snow morphology:* Subsequent research [9], [10] concerns various morphologies of marine snow particles. However, their construction approaches, which model marine snow effects through mathematical or generative models, also struggle to display the complex morphological characteristics of marine snow particles. This is partly because they only use a few real marine snow images for observation and model training, limiting the model's ability to replicate authentic marine snow effects. In addition, these strategies still aim to generate and utilize synthetic marine snow particles, which significantly differ from real-world data.

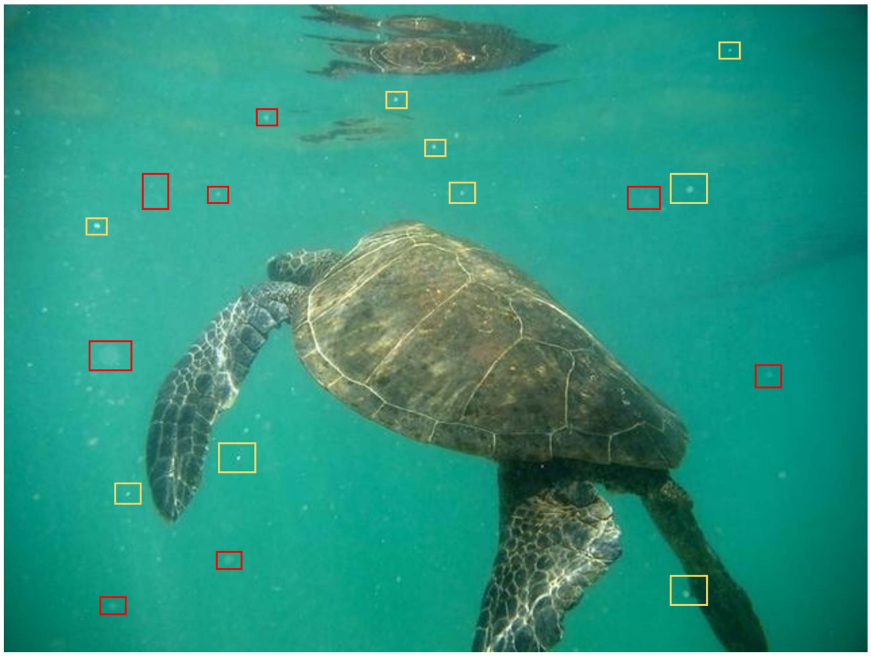

Fig. 1. Samples of marine snow effects in a real-world underwater image. Some light spots are marked in the yellow boxes and some veil-like blurs are marked in the red boxes.

TABLE I
STATISTICS OF EXISTING MARINE SNOW IMAGE DATA SETS, INCLUDING THEIR CONSTRUCTION APPROACH, IMAGE COUNT (#IMG.), PARTICLE MORPHOLOGY COUNT (#MORPH.), DISTRIBUTION PATTERNS COUNT (#DIST.), AND DATA SET AVAILABILITY (#AVAIL.)

Data set	#Img.	#Morph.	#Dist.	#Avail.
Adding Noise Points to Underwater Images				
Banerjee *et al.* [3]	-	-	-	No
Jiang *et al.* [8]	6600	-	-	No
Wang *et al.* [4]	2100	-	-	No
Modeling Marine Snow Effect on Underwater Images				
MSRBD [9]	5400	- (2 types)	1	Yes
Galetto *et al.* [10]	18846	-	-	Yes
Adding Real Marine Snow Particles to Underwater Images				
Hodne *et al.* [11]	3941	-	-	Yes
MMSD [12]	1436	- (5 types)	3	No
Zhao *et al.* [13]	16500	-	-	No
MM-MSD (ours)	6000	>1 M (5 types)	10	Yes

3) *Neglecting marine snow distribution:* Recent efforts [11], [12], [13] improve data sets by employing real-world marine snow particle images, though they still fail to capture their diverse morphologies and distribution patterns in nature. In fact, marine snow has a wide range of morphologies and distribution patterns. These differences depend on factors like ocean location, depth, and season [1]. Therefore, this oversight results in a lack of ecological realism in the data sets. Moreover, lots of existing data sets are not available to the public, leading to data scarcity that impedes the development of marine snow image processing.

The above shortcomings highlight the need for valid and diverse marine snow image data sets. However, constructing such a data set is challenging because it requires not only the representation of diverse marine snow morphologies but also their distribution patterns across ocean areas, depths, and seasons. The distinct vertical and seasonal variations specific to each ocean area require vast data collection and analysis. To resolve these issues, we develop a generic pipeline (called *OmniMSI*) rather than a single data set, which could be applied to broad ocean areas, depths, and seasons. As shown in Fig. 2, our pipeline contains the following four steps:

1) *Raw data preparation:* EcoTaxa[1] offers a large amount of real marine snow particle data, which includes in-situ images, taxonomic information, morphological features, and collection details like time and depth. Our pipeline sets a precedent by fully utilizing this raw data. In addition, the open-source underwater image data sets (such as SUN [14], UIEB [15], and EUVP [16]) provide different underwater scenes for the data set construction.
2) *Morphotype classification and distribution pattern analysis:* Despite the extremely diverse morphological characteristics, marine snow particles share certain similarities, enabling us to classify them into specific morphotypes.

[1][Online]. Available: https://ecotaxa.obs-vlfr.fr

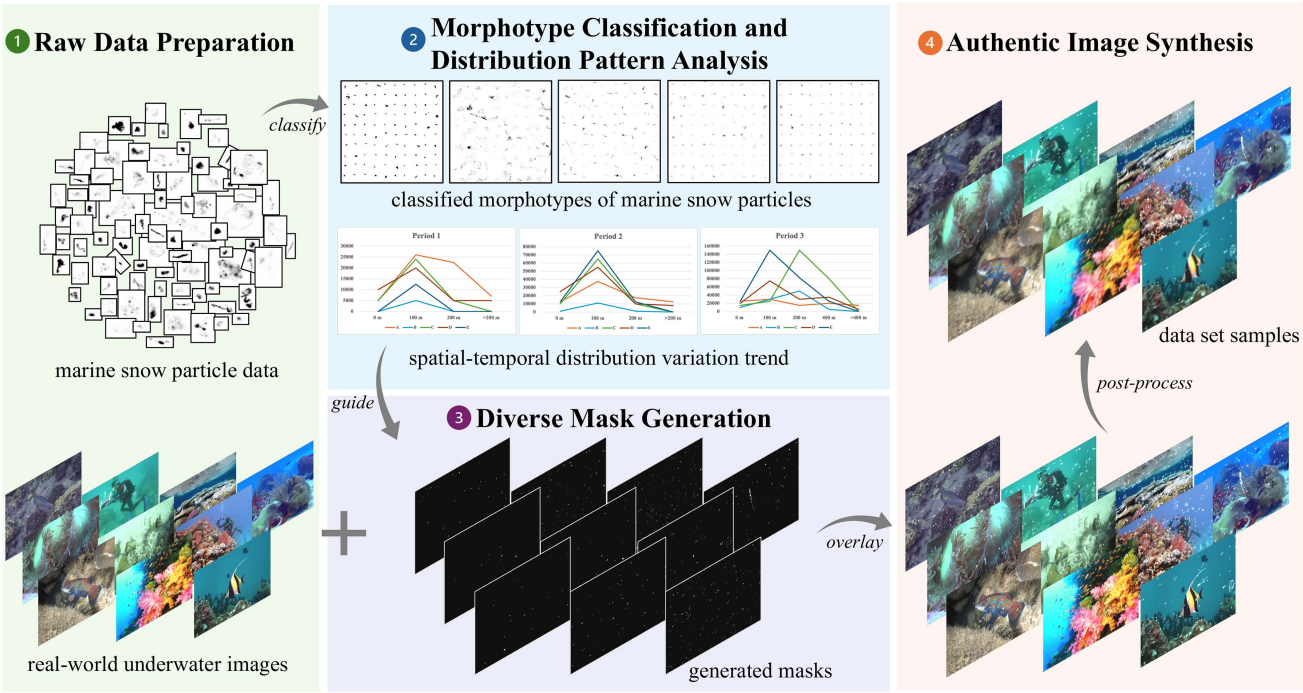

Fig. 2. Overview of our *OmniMSI* pipeline, which contains four steps: ① raw data preparation, ② morphotype clustering and distribution pattern analysis, ③ diverse mask generation, and ④ authentic image synthesis.

Moreover, the abundance of each and all morphotypes varies with ocean area, depth, and season. Therefore, our pipeline leverages these features to first classify marine snow particles via clustering, then analyze the spatio-temporal distribution, and select typical cases as supporting data for subsequent simulations of the real distribution of marine snow.

3) *Diverse mask generation:* Based on the above analysis, our pipeline in this step generates diverse marine snow masks to simulate realistic marine snow effects, and the marine snow particles from different morphotypes are all organized on the masks guided by the distribution patterns.
4) *Authentic image synthesis:* In this step, our pipeline overlays the marine snow masks on real-world underwater images without marine snow. Then it applies the designed post-process, Gaussian blur and dynamic blur, on them to simulate the backscattering effect and motion streaks in the real marine snow scenes.

Utilizing our pipeline, we classify 1 085 506 marine snow particle data into five morphotypes and identify ten distribution patterns to construct our data set. It comprises 6000 pairs of marine snow images and corresponding reference images and is named *M*ultimorphology and *M*ultidistribution *M*arine *S*now image *D*ataset (MM-MSD). Then, we propose a novel marine snow removal method—*M*arine *S*now-*A*ware *Diff*usion (*MSADiff*), which contains two key components: a marine snow estimator (MSE) for estimating multimorphology and multidistribution marine snow particles, and a marine snow remover for removing them and restoring the image details. Based on *MM-MSD*, we conduct a comprehensive study of the state-of-the-art algorithms qualitatively and quantitatively, developing a new benchmark for marine snow removal.

In summary, the main contributions of this article can be summarized as follows.

1) *Pipeline:* To the best of the authors' knowledge, we are the first to propose a pipeline for making marine snow images. Researchers can leverage our pipeline to build diverse and authentic marine snow images for different applications.
2) *Data Set:* We construct a marine snow image data set *MM-MSD* featuring multimorphology and multidistribution for marine snow image processing.
3) *Method:* We design a novel method *MSADiff* trained on *MM-MSD* for marine snow removal, removing marine snow effects while restoring the image details.
4) *Benchmark:* With the *MM-MSD*, we conduct a comprehensive study on state-of-the-art algorithms from qualitative to quantitative evaluations for marine snow removal.

II. Related Work

A. Marine Snow Image Data Sets

In this part, we mainly review the image data sets that can be used for marine snow image processing (see Table I). In the beginning, many marine snow image data sets were constructed by adding noise points to underwater images. Specifically, Banerjee et al. [3] regarded marine snow effects as a form of impulse noise, characterized by tiny sparkling dots, and simulate them as isolated noise points of 2 to 3 pixels in underwater images. Following that, Jiang et al. [8] used photoshop to add spot noise with different densities and orientations to underwater images. Also, inspired by a rain streaks synthesis method [17],

Wang et al. [4] simulated marine snow effects as Gaussian noise on their captured video frames. However, the above data sets only consist of noisy images rather than marine snow images because of the oversimplification of marine snow effects.

With attention and observation of the various morphologies of marine snow particles, researchers try to model marine snow effects on underwater images. As a representative, in MSRB Data Set (MSRBD) [9], the marine snow effects feature two distinct 3-D mathematical models of marine snow particles, where these two particle types appear on each image with a distribution ratio of 7:3. After that, to replicate the complex morphological characteristics, Galetto and Deng[10] generated marine snow particles on underwater images by using a generative adversarial network (GAN). However, these data sets still underestimate marine snow morphology and use fake marine snow particle data to simulate marine snow effects.

Recently, some data sets have been constructed by utilizing real marine snow particle images. Hodne et al. [11] selected marine snow frames with untextured backgrounds and extract marine snow particles from them, then overlay them onto the synthetic underwater data set VAROS [18]. Their published data set is named Snowy-VAROS. Zhao and Li[13] turned to an image-blending method to make the extracted marine snow particles well-integrated into the underwater images. However, they only used ten real marine snow images to serve as raw data. Guo et al. [12] chosed to download some in-situ images of marine snow particles from EcoTaxa, manually defined 3 distribution patterns, and overlaid them onto underwater images. Nevertheless, they dismissed the diverse morphologies and distribution patterns of marine snow particles. To solve the above issues, we consider diverse morphologies and different distribution patterns of marine snow particles and propose a generic synthesizing pipeline.

B. Marine Snow Removal Methods

Banerjee et al. [3] thought the marine snow effects resemble impulse noise caused by power equipment, and proposed an adaptive probabilistic method based on a median filter to remove them. Endorsing marine snow as a significant noise source of underwater images, Farhadifard et al. [19] proposed a supervised median filtering scheme tailored to the specific properties of the bright spots caused by marine snow. After that, Jiang et al. [8] applied deep learning algorithms to the marine snow removal task. They proposed a deep learning network employing GAN with skip connections, self-attention, and spectral normalization to remove marine snow from underwater images. Serving as a sub-task of underwater image enhancement, Wang et al. [4] introduced a local residual learning module for marine snow removal in the high-frequency image layer. Motivated by the success of the attention mechanism, Zhao and Li [13] proposed an attention-guided denoising network to detect and remove marine snow particles under a two-stage training strategy. All of the above methods are designed based on denoising methods, without focusing on the properties of marine snow effects. Therefore, we propose a novel marine snow removal model that can remove marine snow effects clearly. The details will be described in Section IV.

III. Pipeline and Data Set

A. OmniMSI Pipeline

In this section, we detail the proposed pipeline for making marine snow images considering multimorphology and multi-distribution. We call our pipeline *OmniMSI* as it is designed to encapsulate comprehensive marine snow information ("Omni") within the generated marine snow images. It consists of four steps: ① raw data preparation, ② morphological type clustering and distribution pattern analysis, ③ diverse mask generation, and ④ authentic image synthesis. The input of *OmniMSI* is marine snow particle data and real-world underwater images without marine snow, and the output is synthesized marine snow images. Each step is explained as follows.

① *Raw Data Preparation:* Reviewing previous marine snow image data sets in Section II, we identify two fatal flaws in their marine snow particle data: lack of realism [3], [4], [8], [9], [10] and lack of diversity [11], [12], [13]. The former is simulated fake marine snow particles, while the latter fails to capture the diverse morphologies and distribution patterns among different ocean areas, depths, and seasons. Meanwhile, the insufficient diversity of underwater scenes also impacts their ultimate quality and visual representation. Based on these observations, we set two goals for preparing raw data in this step: One is large-scale marine snow particle data with diverse morphology, from various depths and across different seasons. Another is real-world underwater images that comprise various image contents and different degrees of degradation.

The EcoTaxa platform has made a significant contribution towards the first goal. As mentioned in Section I, EcoTaxa is a web platform that comprises millions of images uploaded by researchers worldwide. Projects in EcoTaxa are exhibited on the global map as position markers, with project managers setting access permissions to regulate data sharing. Within these projects, every image is automatically annotated with taxonomic information, morphological characteristics like size and shape, as well as collection details including geographic coordinates, times, and depths. Notably, marine snow particle data in EcoTaxa is classified under "nonliving" except for "bubble," which is the primary focus of our *OmniMSI*. Therefore, the first step is to download a project of interest from EcoTaxa and select the marine snow particle data from it. For the second goal, we suggest making use of published underwater image data sets (like UIEB [15]) to select underwater images without marine snow, prioritizing real and unprocessed ones with diverse visuals that cover a wide range of underwater scenes and degradation features.

② *Morphotype Clustering and Distribution Pattern Analysis:* Although marine snow particles vary widely in their appearances and structures, they have similar morphological characteristics. Therefore, in this step, we leverage principal component analysis (PCA) and the unsupervised K-means clustering algorithm to analyze their morphological characteristics and classify them into different morphotypes.

First, according to the work done by Trudnowska et al. [1], we select the morphological characteristics from three aspects: 1) size, which is most intuitive that related to geometric measures for figuring particles; 2) shape, which reflects the

TABLE II
24 MORPHOLOGICAL CHARACTERISTICS USED FOR MARINE SNOW CLUSTERING, THEIR MORPHOLOGICAL DIMENSIONS, AND THE NORMALIZATION PROCESSES

# Char.	# Desc.	# Dim.	# Norm.
area	surface area of the object in square pixels	size	$\log_{10} trim(x)$
perim.	perimeter of the outside boundary of the object	size	$\log_{10} trim(x)$
major	primary axis length of the best-fit ellipse (secondary axis length named minor)	size	$\log_{10} trim(x)$
feret	maximum Feret diameter within the object	size	$\log_{10} trim(x)$
fractal	fractal dimension of object boundary	size	$trim(x)$
skelarea	surface area of skeleton in pixels	size	$\log_{10} trim(x)$
circ.	circularity of the object — $(4 * \pi * area)/perim.^2$	shape	$trim(x)$
symetrieh	bilateral horizontal symmetry index	shape	$\log_{10} trim(x)$
symetriev	bilateral vertical symmetry index	shape	$\log_{10} trim(x)$
thickr	thickness ratio	shape	$\ln(1 + trim(x))$
elongation	aspect ratio of the object — $major/minor$	shape	$\log_{10} trim(x)$
mean	average grayscale value of all pixels within the object (maximum value named max)	texture	$trim(x)$
mode	modal grey value (most frequently occurring gray value)	texture	$trim(x)$
intden	sum of grayscale values of all pixels within the object — $area * mean$	texture	$\log_{10} trim(x)$
median	median grayscale value within the object	texture	$trim(x)$
slope	slope of the grey level normalized cumulative histogram	texture	$\log_{10} trim(x)$
range	the range from min gray values to the max	texture	$trim(x)$
%area	percentage of object's surface area that is comprises holes	texture	$\ln(1 + trim(x))$
stddev	standard deviation of grayscale values	texture	$trim(x)$
skew	skewness of the grayscale histogram	texture	$trim(x)$
kurt	kurtosis of the grayscale histogram	texture	$trim(x)$
meanpos	balance of pixel intensity distribution — $(max - mean)/range$	texture	$trim(x)$
cv	coefficient of variation of the grayscale histogram — $(stddev/mean) * 100$	texture	$trim(x)$
sr	relative standard deviation of grayscale histogram — $(stddev/range) * 100$	texture	$trim(x)$

In the description column (# Desc.), formulas following the vertical bar are provided for easy understanding. In the normalization procedure column (# Norm.), "trim(x)" means we trim 0.1% extreme values before normalizing.

geometric form, external boundary, or surface; 3) texture, which is mainly related to grayscale values of pixels that reflect the degree of transparency, the underline ones indicate the properties of the grayscale histogram. Table II presents the selected 24 morphological characteristics that meet the above standards, detailed along with their descriptions, associated morphological dimensions, and specific normalization procedures. Then, the finally selected morphological characteristics are trimmed and normalized to ensure that no single one dominates the clustering process due to scale differences.

After that, PCA is used to transform the 24 original features into principal components, which can select more representative characteristics. Furthermore, we apply K-means clustering on the selected principal components to obtain different morphotypes of marine snow particles. The optimal number of clusters (K value) is determined by the Gap Statistic [20] method due to its ability to objectively assess the optimal clustering number by comparing the within-cluster dispersion with the expected value under a null reference distribution. It is important because of the inherent ambiguity and potential visual subtlety of the true cluster count within marine snow particles. In our *MM-MSD*, we finally use the top-4 principal components to cluster 5 morphotypes of marine snow particles for data set construction. Details are displayed in Section III-B as our pipeline mainly focuses on the outline of steps.

In terms of different morphotypes, the abundance of each exhibits variations by depth and seasons. Therefore, we further analyze and define distinct spatial–temporal distribution patterns

across morphotypes, depths, and seasons, by identifying significant feature points—those locations where the abundance changes obviously. Specifically, in a specific season, we group the depth by 10-m intervals. This is predicated upon the fact that the raw data sample depths, as mentioned in [21], were typically measured at every 5 m within the range from 0 to 30 m and at every 10 m from 30 to 100 m. In order to ensure easy reproducibility of our study, we selected the lowest common multiple of these measurement intervals, which is 10 m. Subsequently, we statistics the abundance of each morphotype into a variation trend chart, then point out the depth where obvious compositional changes occur. After that, we record the compositions of each morphotype at these subsequent broader depth intervals and calculate a proportion for each morphotype until all combinations are cataloged. The recorded compositions and the proportions are used to quantify different distribution patterns. Note that while this step is specific to a specific season, the analysis process remains consistent across different seasons.

③ *Diverse Mask Generation:* Based on the above analysis, this step focuses on using marine snow particle images to simulate realistic marine snow effects to generate diverse marine snow masks, guided by the qualified distribution patterns. Particularly, translucent marine snow particles manifest as veil-like blurs in underwater images, while opaque or minimally translucent particles appear as light spots. Therefore, we first preprocess all marine snow particle images to achieve a certain level of transparency using their grayscale values as transparency indicators. Subsequently, guided by the qualified distribution patterns, we set values randomly in the variation ranges to determine the number of marine snow particles on each mask. Lastly, we generate marine snow masks by placing the processed particles on a blank background while randomly adjusting their sizes and rotation angles to increase the diversity of masks, creating realistic visualizations of distribution patterns.

④ *Authentic Image Synthesis:* In this step, we first align the sizes of the marine snow masks and real-world underwater images and overlay the former on the latter by applying this equation:

$$I = M + U(1 - M/255.) \tag{1}$$

where I denotes synthesized marine snow image, M denotes generated mask, and U denotes real-world underwater image.

Our *OmniMSI* also provides two postprocessing technologies to achieve more natural visualizations: 1) Applying Gaussian blur on the coarsely synthesized images to simulate the enhanced backscattering effect attributed to the tiny marine snow particles. 2) Applying a dynamic blur to replicate the relative motion between the capturing device and the movement of the water body. These technologies are instrumental in approximating the actual capture conditions and the distinct marine snow effects, where particles closer to the camera appear clear yet still influenced by the relative motions. It further improves the realism of marine snow images by accurately reflecting the interaction between light and marine snow particles in the underwater environment. Note that each marine snow mask can be applied to one specific underwater image to simulate a particular marine snow scene (one-to-one) or to multiple underwater images to represent a consistent marine snow effect across different scenes (one-to-many).

B. MM-MSD Data Set

In this section, we construct a new marine snow image data set using our *OmniSMI* pipeline. Note that all the selections of parameters are optional. Readers can select as needed according to our following steps. Initially, we select a project located in the Arctic Ocean from EcoTaxa[2] and obtain 1 085 506 marine snow particle data under the "nonliving" taxonomy. As noticed above, we excluded the "predicted," "artifact," and "duplicate" subsets, as they are caused by underwater devices rather than marine snow particles. Then, we employ PAC to transform the 24 selected characteristics into principal components. Table III presents the top-10 principal components and their contributions to variance, providing a reference in selecting the most informative features for further classification. To balance between model simplicity and sufficient data variance retention, we utilize the top-4 components as the primary features that account for over 86% of the variance in the 24 morphological characteristics.

As mentioned above, when determining the optimal number of clusters, we mainly rely on the results of the Gap Statistic, because of its ability to objectively evaluate the optimal clustering number by comparing the within-cluster dispersion with the expected value under a null reference distribution. In this case, with the Top-4 principal components, the optimal value of K is 5. As a result, as shown in Fig. 3, we classify the marine snow particles into five morphotypes. Specifically, the particles in *Type A* particles have a relatively small size and moderately regular shape, with many approximating circular or triangular forms. They are fully opaque, with few surface cavities, and exist in a tight cluster to form a stable aggregate. *Type B* particles are unique in size and transparency. They are quite large, with extremely irregular and randomly extending edges. Their structures are the most sparse among all the morphotypes and look like large flocculent aggregates. Composed of small particles with significant transparency differences, they show a mottled transparency effect. *Type C* particles' linear and elliptical shapes are prominent, making them easily distinguishable. *Type D* particles are mostly transparent but not uniformly, with some hazy regions like gauze. Their loose structure and relatively consistent texture have minimal transparency variation within a particle and are mostly flocculent. *Type E* particles are markedly smaller, with the most regular shape, mostly in a compact cluster. They have low overall transparency and a uniform texture. It is important to note that higher grayscale values correspond to brighter areas in the synthesized images of marine snow. As a result, in our synthesized marine snow images, *Type A* and *Type E* usually appear as light spots, while others typically appear as veil-like blurs. Then, we qualify the abundance of each morphotype across different seasons. As illustrated by the simplified spatio-temporal distribution variation trend in Fig. 2 ②, the concentration of marine snow particles varies to different extents as time progresses. Notably, several common turning

[2][Online]. Available: https://ecotaxa.obs-vlfr.fr/prj/149

TABLE III
TOP-10 PRINCIPAL COMPONENTS (# PC) OF THE 24 MORPHOLOGICAL CHARACTERISTICS, INCLUDING EACH COMPONENT'S EIGENVALUE (# EV), VARIATION CONTRIBUTION (# CONTRI.) AND CUMULATIVE CONTRIBUTION (# CUML.)

# PC	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
# EV	11.80	4.87	2.79	1.34	0.90	0.62	0.60	0.28	0.18	0.15
# Contirb. (%)	49.15	20.28	11.64	5.57	3.75	2.58	2.50	1.18	0.76	0.61
# Cuml. (%)	49.15	69.43	81.07	86.64	90.39	92.97	95.47	96.65	97.41	98.86

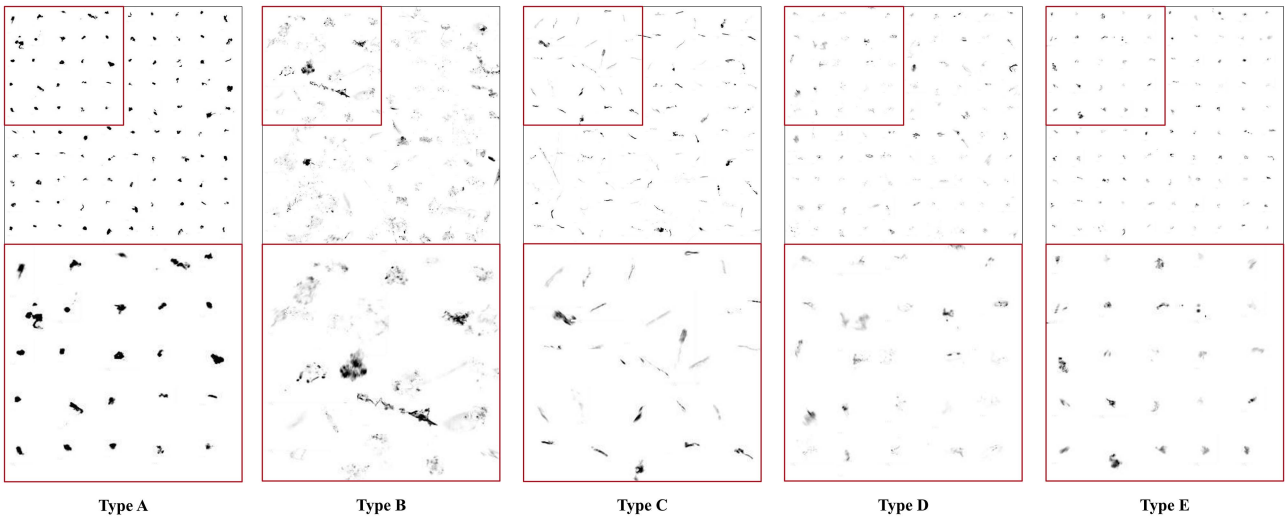

Fig. 3. Samples of five morphotypes of marine snow particles, each distinguished by unique morphological dimensions of size, shape, and texture: Type A: Small-sized with a moderately regular shape and tightly aggregated structure. Type B: Large with an irregular shape, significant transparency variations, and a loose structure. Type C: Characterized by an elliptical shape. Type D: Mostly transparent with a hazy appearance, loose structure, uniform texture, and flocculent-like. Type E: Noticeably small, with a regular and compact clustered shape, low transparency, and uniform texture. Zoom-in views are presented in the red boxes for a more detailed illustration of the morphological differences.

points emerge at specific times. Based on this, we identified ten distribution patterns, with three patterns in period 1, three patterns in period 2, and four patterns in period 3.

For each pattern, we set a minimum threshold of 100 particles per mask to ensure a sufficient amount. We generate 600 masks for each of the ten patterns, totaling 6000 masks across all patterns. Within each pattern, every individual mask maintains a consistent proportion of each morphotype, ensuring uniformity in the representation of marine snow particle distributions. Across different patterns, the overall number of marine snow particles varied significantly, reflecting the realistic abundance from different depth intervals and seasons, thus enriching our data set with diverse morphologies and distribution patterns of marine snow particles. Fig. 4 shows some samples of our data set *MM-MSD*. The first column shows real-world underwater images, the second column shows marine snow masks, and the fourth column shows marine snow images (tagged as "postprocess results"). To intuitively show the performance of our postprocessing for simulating underwater degradation, the figure includes direct overlay results in the third column. As shown, these marine snow particles appear unnatural and do not blend well with the underwater background. Conversely, the postprocessed images are far more natural and look like real-captured ones.

Overall, we select 6000 real-world underwater images without marine snow as reference images from SUN [14], UIEB [15], and EUVP [16] that are widely used in underwater image enhancement research. For each reference image, we apply one marine snow mask on it to synthesize a one-to-one initial marine snow image, then apply the proposed postprocessing to get the refined final image. As a result, our data set consists of 6000 pairs of marine snow images and their reference images, namely, *MM-MSD*.

We divide it into training, validation, and testing sets with a ratio of 8:1:1 for future work. To preserve the balance and utility of our data set, we ensure that the proportion of images under each distribution pattern is consistent across the training, validation, and testing set.

IV. METHODOLOGY IN MASDIFF

Marine snow removal poses a significant challenge due to the diverse morphologies and distribution patterns of marine snow particles. To tackle this issue, we propose a novel marine

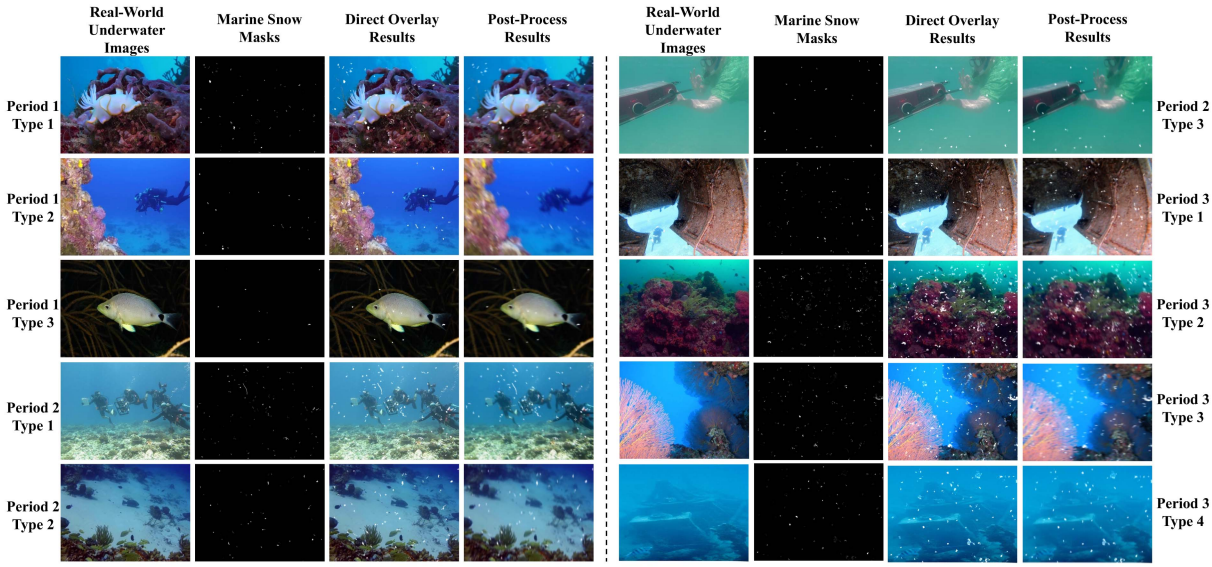

Fig. 4. Samples of our data set *MM-MSD*, including real-world underwater images, marine snow masks, and marine snow images. The direct overlay results are displayed to demonstrate the effect of our postprocessing.

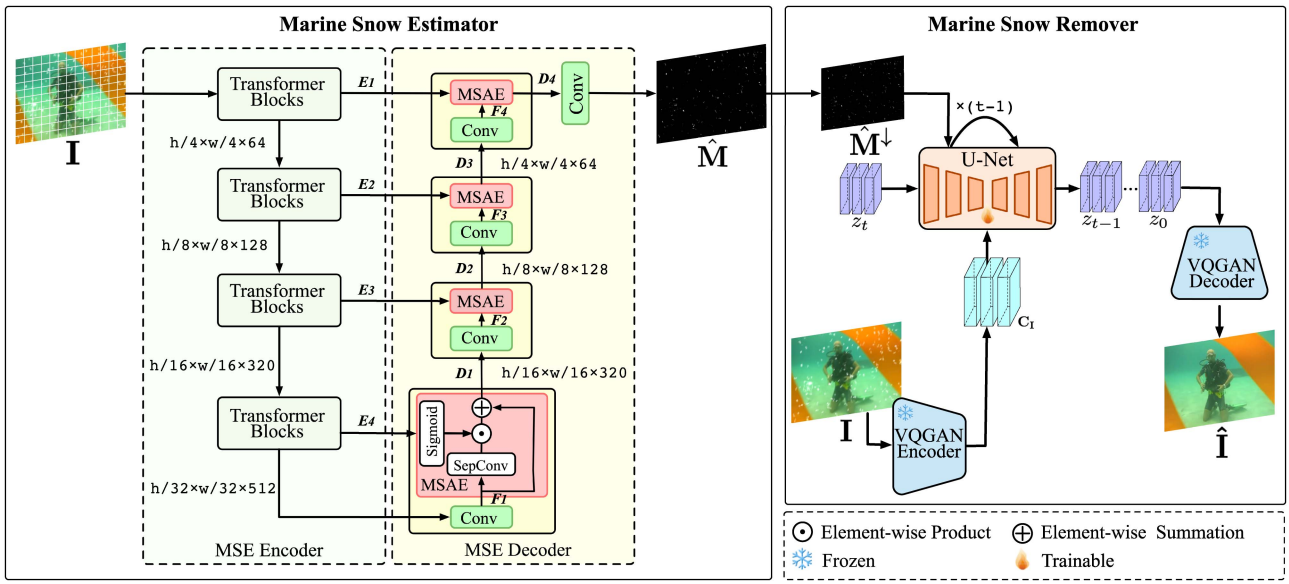

Fig. 5. Overview of the proposed *MSADiff*. *MSADiff* is a two-stage network, which follows a diffusion model-based architecture. The MSE estimates a marine snow mask $\hat{\mathbf{M}}$ from the input marine snow image $\mathbf{I}$, and then the marine snow remover progressively removes marine snow particles and restores the corresponding marine snow-free image using the downsampled marine snow mask $\hat{\mathbf{M}}^{\downarrow}$ and random noise Z_t. E_i denotes the output features of encoder blocks, F_i denotes the output of convolutional layers, and D_i denotes the output of decoder blocks.

snow removal method—MSADiff to progressively estimate and remove marine snow particles.

MSADiff consists of two key components: marine snow estimator and marine snow remover. As shown in Fig. 5, given a marine snow image $\mathbf{I} \in \mathbb{R}^{h\times w\times 3}$ as input, the marine snow estimator estimates a marine snow mask $\hat{\mathbf{M}} \in \mathbb{R}^{h\times w\times 1}$ by thoroughly analyzing the contextual information. Then, the marine snow remover removes the marine snow particles and restores a corresponding marine snow-free image $\hat{\mathbf{I}} \in \mathbb{R}^{h\times w\times 3}$, guided by the downsampled marine snow mask $\hat{\mathbf{M}} \downarrow$. During the training phase, we first train the marine snow estimator based on the marine snow images and corresponding ground truth masks. Once it has a strong capability for marine snow mask estimation, we then freeze its parameters before training the marine snow remover.

A. Marine Snow Estimator

Our marine snow estimator has an encoder-decoder architecture, aiming to perceive the position of marine snow particles for predicting a binary marine snow mask.

MSE encoder: For the encoder, we utilize the PVTv2 [22] instead of traditional Transformer models, as its superior

capability to capture global contextual information across varying scales. The advancement of PVTv2 facilitates the encoder to model the context information (like scene semantics and structure) of the input image, thus enabling focused attention on the diverse morphologies and distribution patterns of marine snow particles. For an input of marine snow image $\mathbf{I} \in \mathbb{R}^{h\times w\times 3}$, the PVTv2 encoder generates multiscale features

$$\{E_i\}_{i=1}^{4} \in \mathbb{R}^{\frac{h}{2^{i+1}}\times\frac{w}{2^{i+1}}\times c_i}(c_i \in \{64, 128, 320, 512\}, \quad i \in \{1, 2, 3, 4\})$$

from its four Transformer blocks. These features are considered to encompass varying levels of edge, texture, and marine snow semantic information.

MSE decoder: We build the decoder based on the encoder, which also contains four blocks. Each block contains two components: a convolutional layer and a *m*arine *s*now-*a*ware *e*nhancement (MSAE) module. The Transformer-based encoder excels at modeling global contextual information but lacks the ability to capture long-range dependencies. Therefore, we utilize the convolutional layer in the decoder to extract local text details like textures and edges through fixed receptive fields. However, directly combining the convolutional layer with the encoder to realize global and local feature extraction for the marine snow estimation may face two challenges: 1) The global features from the Transformer may be too abstract for the convolutional decoder, leading to insufficient representation of local details. 2) The local features from the convolutional neural network (CNN)-based decoder may struggle to integrate global context, leading to suboptimal decoding performance and reduced estimation accuracy. Therefore, we construct the MSAE module to process and enhance the related feature representations.

Algorithmically, the MSAE module utilizes a dot product operation of a Sigmoid layer and a separable convolutional layer (Sigmoid $\odot$ SepConv) with a residual structure to further process the global spatial information for the following convolutional layer. As we know, separable convolution in MobileNetV2 [23] decomposes the standard convolution into two steps: depthwise convolution and pointwise convolution. The depthwise convolution captures spatial information within each channel independently, and the pointwise convolution combines the features across channels, enabling the model to learn rich feature representations. The combination with the Sigmoid activation layer allows for a more refined and selective extraction of features based on the importance weights assigned by the Sigmoid. Then, the residual structure helps in better gradient flow during model training.

As shown in Fig. 5, the multiscale features $\{E_i\}_{i=1}^{4}$ obtained by the encoder are subsequently fed into the decoder, and then, the MSAE integrates both the $\{E_i\}_{i=1}^{4}$ and the convolved features $\{F_i\}_{i=1}^{4}$ to produce the augmented features $\{D_i\}_{i=1}^{4}$. The main process can be formulated as

$$D_i = F_i \oplus \mathrm{Sigmoid}(E_i) \odot \mathrm{SepConv}(F_i) \tag{2}$$

where $\oplus$ is element-wise summation, $\mathrm{Sigmoid}(\cdot)$ is sigmoid activation, $\odot$ is element-wise product, and $\mathrm{SepConv}(\cdot)$ is separable convolution. The multilevel augmented features vary from D_1 with spatial size $(h/16)\times(w/16)$ and channel size 320 to D_4 with spatial size $h\times w$ and channel size 1. With the final augmented features generated by the last convolutional layer, we obtain an estimated binary mask $\hat{\mathbf{M}} \in h\times w\times 1$ that pinpoints the location of the marine snow particles.

B. Marine Snow Remover

Our marine snow remover is designed based on a pretrained latent diffusion model (LDMs) [24], guided by the estimated binary marine snow mask, to optimize its efficiency for marine snow removal. We utilize the remarkable ability of LDMs to generate and reconstruct images with high fidelity, as it is essential for accurately removing the light spots caused by diverse marine snow particles and restoring the underlying details. On the other hand, the model achieves better efficiency and stabilizes the training process by using the pretrained VQGAN [25], which helps to convert the image into a more manageable latent space. This enhanced efficiency is of great significance when dealing with intricate marine snow particles and diverse underwater environments.

As shown in Fig. 5, in the process of removing marine snow, the procedure can be divided into two stages: forward process and reverse process. In the stage of the forward process, we first convert marine snow-free images into encoded latent z with the VQGAN encoder, and then, Gaussian noise with variance $\beta_t \in (0, 1)$ at time t is added to the encoded latent z for producing the noisy latent z_t, which can be formulated as

$$z_t = \sqrt{\overline{\alpha}_t}z + \sqrt{1-\overline{\alpha}_t}\epsilon \tag{3}$$

where $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ with $\mathbf{I}$ denoting identity matrix, $\alpha_t = 1-\beta_t$ and $\overline{\alpha}_t = \Pi_{s=1}^{t}\alpha_s$. When t is large enough, the latent z_t is nearly a standard Gaussian distribution.

During the stage of the reverse process, we use the pretrained VQGAN encoder to encode the marine snow image $\mathbf{I}$ into the latent feature $\mathbf{C_I}$. Subsequently, the estimated mask $\hat{\mathbf{M}}$ is downsampled to generate the reduced-resolution mask $\hat{\mathbf{M}}^{\downarrow}$. The downsampled mask $\hat{\mathbf{M}}^{\downarrow}$ and the latent feature $\mathbf{C_I}$ are utilized as conditioning features to guide and facilitate the removal process appropriately. Subsequently, the denoising network (U-Net) takes the z_t as input conditioned on the latent feature $\mathbf{C_I}$ and the downsampled estimated mask $\hat{\mathbf{M}}^{\downarrow}$ to predict a slightly less noisy feature z_{t-1}. The process is repeated iteratively. When t reaches 0, the U-Net will generate a latent representation z_0, which can be transformed back the pixel space by the VQGAN decoder to yield a removal result $\hat{\mathbf{I}}$ with high-definition texture and faithful snow-free representation.

V. Benchmark

A. Methods and Experimental Setup

Implementation details: All the benchmarking experiments are implemented using PyTorch and the training process is on four RTX 3090 graphics processing units (GPUs). Images used for model training are all resized to 256×256.

MSADiff training and inference: The MSE is trained using binary cross-entropy loss between the ground truth mask and

the estimated mask with a batch size of 4. We use the Adam optimizer [26] with the mean of the gradients $\beta_1 = 0.0$, the variance of the gradients $\beta_2 = 0.9$, and initialize the learning rate $\alpha = 0.0001$. The marine snow remover (utilizing Stable Diffusion 1.7-base [24]) is finetuned on our data set for 70 k iterations with a batch size of 4. We use the Adam optimizer [26] with an initial learning rate of $\alpha = 0.0001$. During inference, we adopt spaced DDPM sampling [27] with 50 timesteps.

Data Sets: The training and testing data sets for the benchmark experiments utilize the proposed *MM-MSD*. Additionally, we conduct a quality assessment among the two published marine snow image data sets (MSRBD [9] and Snowy-VAROS [11]) and our *MM-MSD*. The data set quality assessment is displayed in Section V-C.

1) *MSRBD [9]* is generated by modeling marine snow effects into two mathematical models and employing them on the underwater images. It has 4600 training pairs and 800 testing pairs. Each pair consists of an underwater image and its corresponding marine snow image.
2) *Snowy-VAROS [11]* is generated by extracting real marine snow particles and applied on the synthesized underwater image data set VAROS, comprising 3229 pairs of images in the training set, 342 pairs in the validation set, and 370 pairs in the testing set.
3) *MM-MSD*features five morphotypes and ten distribution patterns of marine snow particles. The training set contains 4800 pairs of images, while the test set and validation set contain 600 image pairs separately.

Benchmark methods: To the best of authors' knowledge, no open-source methods are available for marine snow removal. Therefore, we select classic and state-of-the-art methods from the related fields of CNN-based (UGAN [28], Water-Net [15], FUnIE-GAN [16], NU2Net [29], U-shape [30]), and diffusion model-based (DM_underwater [31], PA-Diff [32]) underwater image enhancement methods, image denoising (DnCNN [33], VDN [34], DANet [35], MPRNet [36], and Restormer [37]), and image desnowing (DesnowNet [38], HDCW-Net [39], DDM-SNet [40], TKL [41], and TransWeather [42]) as benchmark methods. The two main reasons for our selection are as follows: 1) marine snow images share similar degradation characteristics with normal underwater images, thus underwater image enhancement methods are reasonable to be employed; 2) the marine snow effects, appearing as light spots resembling noise and veil-like blurs similar to snowflakes, may enable the use of denoising and desnowing methods. All benchmark methods are trained from scratch on the *MM-MSD* training set and tested on the *MM-MSD* testing set.

TABLE IV
QUANTITATIVE EVALUATION RESULTS OF UNDERWATER IMAGE ENHANCEMENT, DENOISING, AND DESNOWING METHODS ON THE *MM-MSD* TESTING SET, UTILIZING PSNR, SSIM, AND UIQM METRICS

Index	Method	PSNR (↑)	SSIM (↑)	UIQM (↑)
	Underwater Image Enhancement Methods			
(a)	UGAN [28]	16.0579	0.3734	2.8037
(b)	Water-Net [15]	16.6403	0.3801	2.8453
(c)	FUnIE-GAN [16]	15.9476	0.3643	2.8994
(d)	NU2Net [29]	16.3847	0.3687	2.7915
(e)	U-shape [30]	16.2729	0.3794	2.8328
(f)	DM_underwater [31]	16.1499	0.3850	2.7278
(g)	PA-Diff [32]	15.9074	0.3694	2.8814
	Denosing Methods			
(h)	DnCNN [33]	16.4263	0.3579	2.5385
(i)	VDN [34]	16.5668	0.3815	2.7843
(j)	DANet [35]	16.4153	0.3770	2.6913
(k)	MPRNet [36]	16.4392	0.3622	2.5169
(l)	Restormer [37]	16.4900	0.3629	2.5528
	Desnowing Methods			
(m)	DesnowNet [38]	16.4638	0.3612	2.5484
(n)	HDCW-Net [39]	16.5798	0.3831	2.7145
(o)	DDMSNet [40]	16.6209	0.3847	2.5303
(p)	TKL [41]	16.3744	0.3699	2.5974
(q)	TransWeather [42]	16.5253	0.3721	2.7834
	MSADiff (Ours)	**16.8126**	**0.3942**	**2.9262**

The symbol ↑ indicates that higher scores correspond to better performance. The best-performing value is highlighted in **bold**.

B. Overall Performance

A comprehensive and fair method validation has not been developed for evaluating marine snow removal. Therefore, in this part, we evaluate the state-of-the-art underwater image enhancement, denoising, and desnowing methods both qualitatively and quantitatively on our *MM-MSD*. Our qualitative evaluation results are shown in Fig. 6 and quantitative evaluation results are shown in Table IV. Fig. 6(a)–(g) denotes underwater image enhancement methods, Fig. 6(h)–(l) denotes denoising methods, and Fig. 6(m)–(q) denotes desnowing methods.

Qualitative evaluation: Upon reviewing all the results of the removal of marine snow in Fig. 6, our method proves to be the best of eliminating marine snow particles and their effects and retaining the detailed information of the original scene. The underwater image enhancement methods Fig. 6(a)–(g) also give good results. To some extent, this can be attributed to the fact that underwater images and marine snow images have similar image degradation characteristics. For example, blurring caused by light scattering, veil-like haze caused by suspended particles in water, and various noise points caused by equipment and

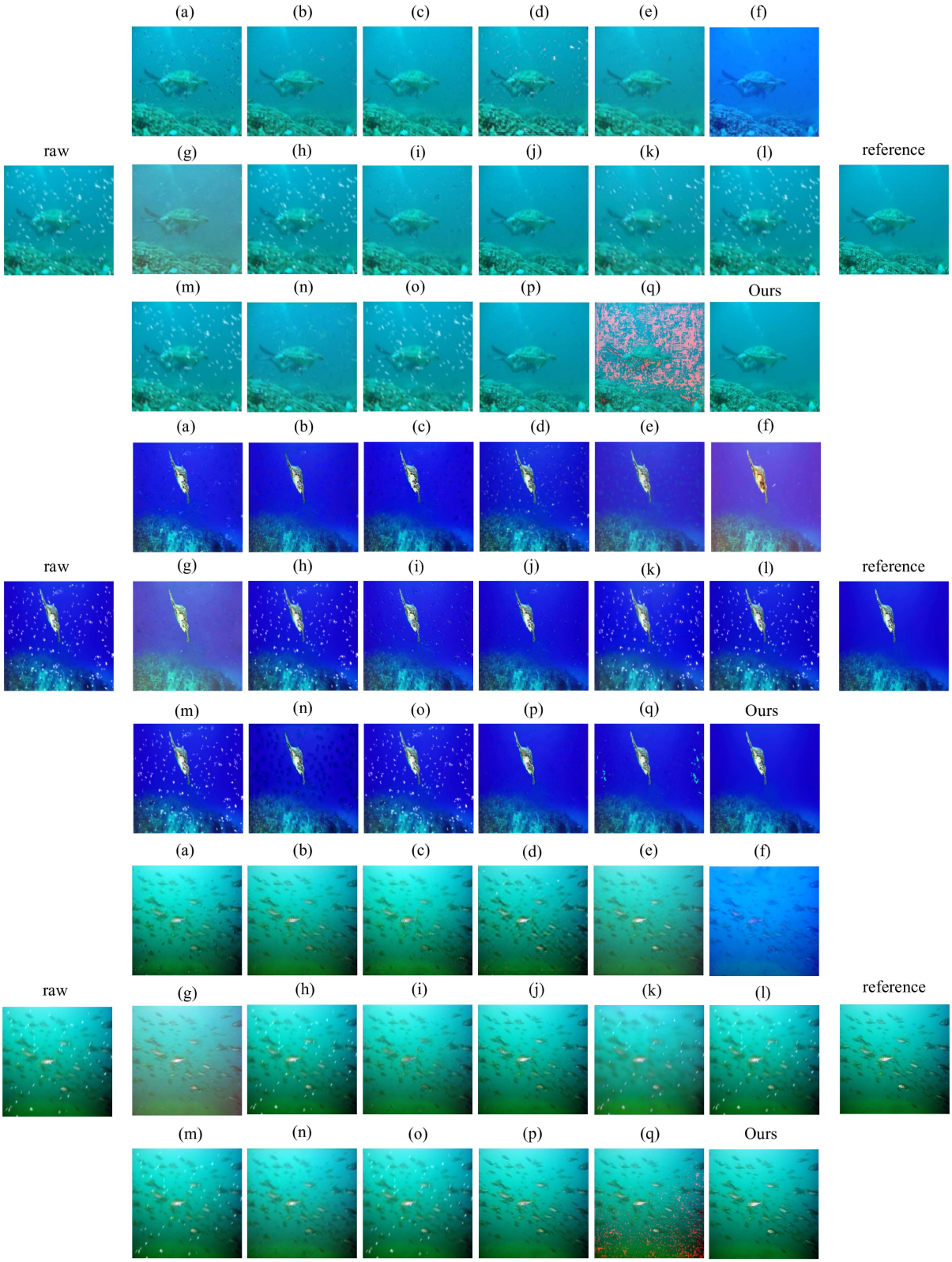

Fig. 6. Sample qualitative evaluation outcomes of underwater image enhancement, denoising, and desnowing methods on the *MM-MSD* testing set. Following the indexes: Raw marine snow images, results of underwater image enhancement methods—(a) UGAN [28], (b) water-net [15], (c) FUnIE-GAN [16], (d) NU2Net [29], (e) U-shape [30], (f) DM_underwater [31], and (g) PA-Diff [32], image denoising methods—(h) DnCNN [33], (i) VDN [34], (j) DANet [35], (k) MPRNet [36], and (l) restormer [37], image desnowing methods—(m) DesnowNet [38], (n) HDCW-Net [39], (o) DDMSNet [40], (p) TKL [41], and (q) TransWeather [42], and uurs *MSADiff*, and reference marine snow-free images.

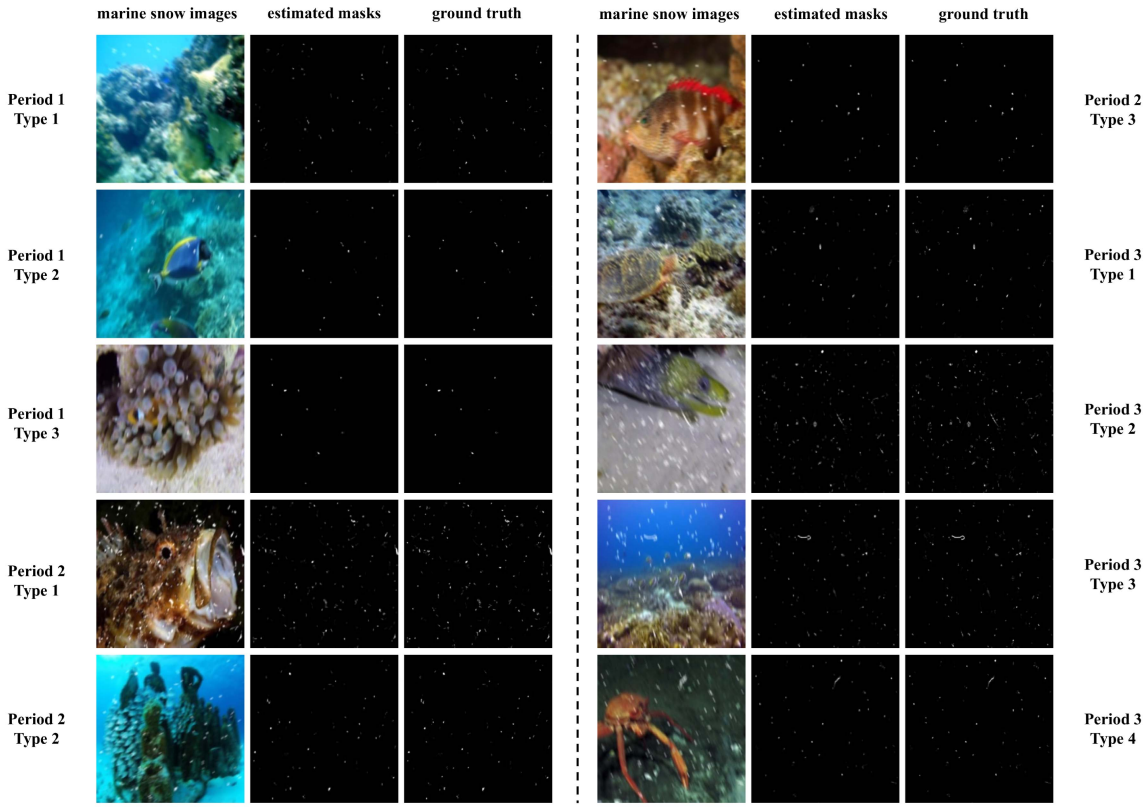

Fig. 7. Samples of estimated masks generated from the MSE.

bubbles. Therefore, those methods designed for degradation characteristics perform well.

As for the denoising Fig. 6(h)–(l) and desnowing Fig. 6(m)–(q) methods, although the desnowing methods generate images with higher clarity, both types of methods have their limitations to some degree. On the one hand, denoising methods are effective only for small marine snow particles but perform poorly for larger ones. This is mainly because denoising algorithms focus more on eliminating the random noise introduced by the acquisition equipment. On the other hand, desnowing methods do not perform well when handling large marine snow particles with motion blur. This is mainly because, while these methods are designed to remove snowflakes of various shapes, they all neglect the snow streaks caused by the motion of snowflakes.

Focusing on each specific method, Fig. 6(h) DDMSNet [40], Fig. 6(k) MPRNet [36], Fig. 6(m) DesnowNet [38], and Fig. 6(o) DnCNN [33] can hardly remove the marine snow particles because they focus more on smaller particles, exactly noise points. Fig. 6(a) UGAN [28], Fig. 6(e) U-shape [30], Fig. 6(i) VDN [34], Fig. 6(j) DANet [35], and Fig. 6(p) TKL [41] introduce artifacts at the positions of marine snow, that is, the pixels covered by marine snow particles are not well restored due to their inaccurate restoring algorithms. Fig. 6(b) Water-Net [15], Fig. 6(c) FUnIE-GAN [16], Fig. 6(f) DM_underwater [31], Fig. 6(g) PA-Diff [32]), Fig. 6(n) HDCW-Net [39], and Fig. 6(q) TransWeather [42] can remove marine snow particles better. In particular, within the underwater image enhancement methods, the results of Fig. 6(f) DM_underwater [31] look black as it focuses on correcting the green tone distortion to the color cast issues. Fig. 6(g) PA-Diff [32] generates a frog-like artifact to achieve the marine snow removal effects. The desnowing method Fig. 6(n) HDCW-Net [39] has good performance in removing particles with motion blur attributed to targeted processing of snow streak artifacts.

Quantitative evaluation: For quantitative evaluation, we conduct two full-reference evaluation metrics, peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) [43], and a nonreference evaluation metric, underwater image quality measure (UIQM) [44]. A higher PSNR score indicates a closer resemblance between the result and the reference concerning the content. A higher SSIM score indicates a greater degree of structural and textural consistency. A higher UIQM score indicates a closer match with human visual perception.

Table IV presents the quantitative results of different methods in terms of PSNR, SSIM, and UIQM on the testing set. All of these scores are computed based on the marine snow-removal results with the reference marine snow-free underwater images. Among these methods, our *MSADiff* achieves the best performance on all of the evaluation metrics. In addition, the underwater image enhancement methods perform the second best in terms of the nonreference evaluation metric UIQM, with moderate PSNR and SSIM scores, especially DM_underwater [31] and PA-Diff [32] (g). Both denoising and desnowing methods achieve comparable PSNR and UIQM scores, while desnowing methods show a slight advantage in SSIM.

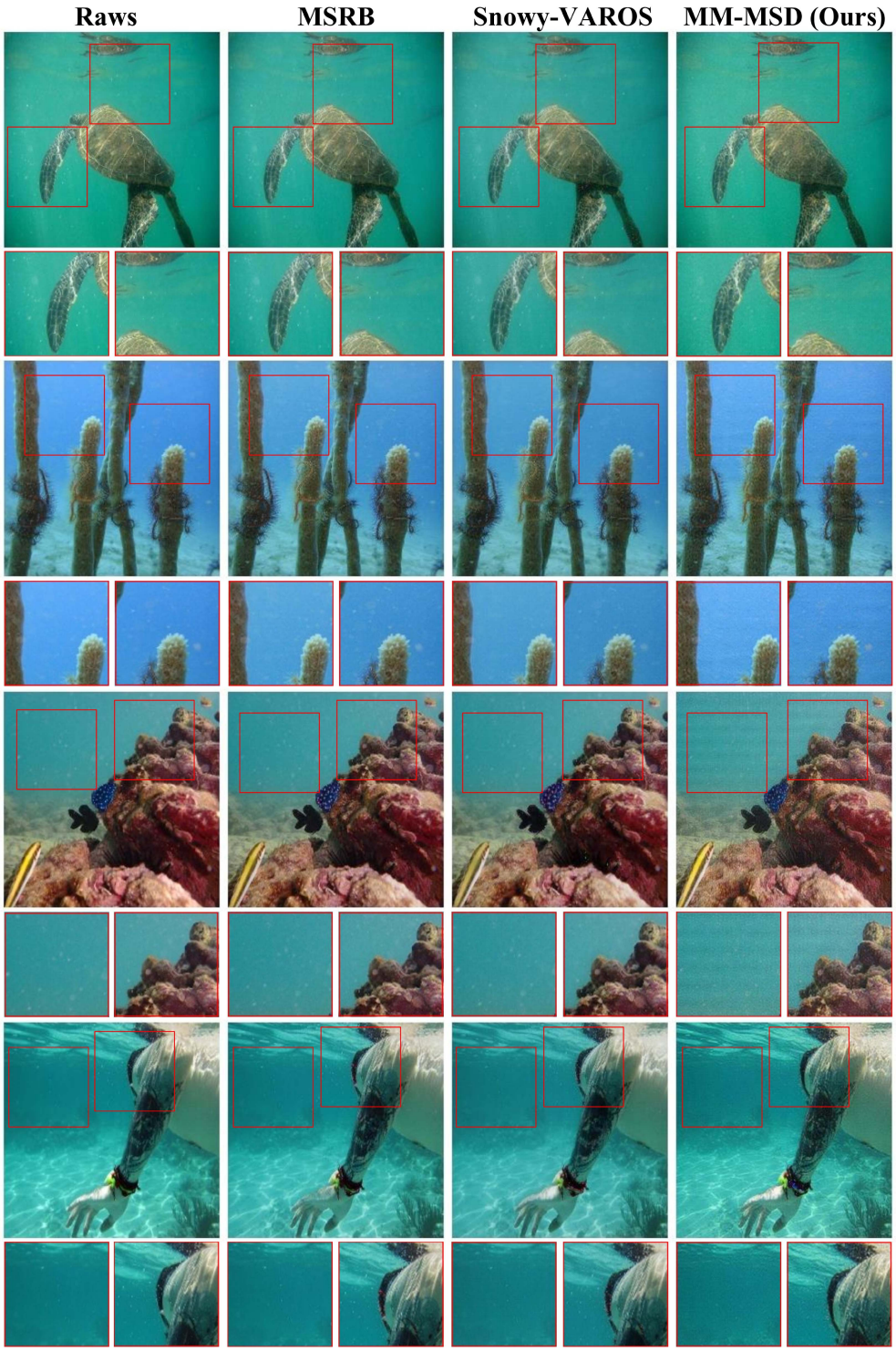

Fig. 8. Sample qualitative evaluation results of *MSADiff* trained on different marine snow image data sets. The test images are real-world marine snow images. Red boxes correspond to the zoomed-in patches for better comparison.

The evaluation metric scores also help explain the obvious visual performance differences. For example, DM_underwater (f) attains a high SSIM score, but its PSNR and UIQM values are low. The reason may be that although it manages to remove marine snow effects with good image structure and texture, the unexpected black tone deteriorates the image quality. Similarly, PA-Diff (g) introduces blurs to the results, thus leading to a decrease in both PSNR and SSIM scores. The reason for the lower SSIM scores of denoising methods aligns perfectly with the previous qualitative analysis. That is, denoising methods can hardly handle large marine snow particles, so they contribute little to the improvement of image structure and texture. Among the desnowing methods, when comparing the performance of DesnowNet (m) and HDCW-Net (n), HDCW-Net performs well

TABLE V
IMPACT OF DIFFERENT TRAINING APPROACHES

Training Strategy	PSNR (↑)	SSIM (↑)	UIQM (↑)
Joint Training	16.6678	0.3807	2.7843
Two-Stage Training	**16.8126**	**0.3942**	**2.9262**

The best performance is highlighted in **bold**.

TABLE VI
EFFECT OF THE MSAE MODULE IN MSE

MSAE	PSNR (↑)	SSIM (↑)	UIQM (↑)
✗	16.2518	0.3019	2.1813
✓	**16.8126**	**0.3942**	**2.9262**

in both SSIM and PSNR because it can handle snow streaks that are similar to motion blur.

Ablation study: Based on the above analyses, we have identified the advantages of our method. In addition, we will demonstrate the superiority of our method through ablation studies of some of its modules.

1) *Effect of Two-Stage Training Strategy:* As mentioned above, the training process of our *MASDiff* has two stages: The first stage is to train the MSE to recognize marine snow masks from marine snow images; the second stage is to train the marine snow remover to remove marine snow particles according to the mask-annotated positions. These two stages have distinct training objectives, with the performance of the second-stage removal dependent on the first-stage estimation results. We conduct quantitative comparisons of these two training strategies on the *MM-MSD* data set. As shown in Table V, the two-stage training strategy demonstrates marginally superior performance. This discrepancy stems from the challenges in multitask optimization: joint training forces simultaneous optimization of two different objectives, creating gradient interference that may degrade the task. Beyond algorithmic challenges, joint training increases GPU memory usage and training time, so we insist on the two-stage training strategy.
2) *Effect of MSE:* To display the accuracy of our MSE, we compare some samples of the estimated masks with the ground truth masks in Fig. 7. It is obvious that the estimation results are precise with the ground truth, underscoring the effectiveness of our estimator in accurately detecting and positioning marine snow effects.
3) *Effect of MSAE Module:* To confirm and investigate the contribution of the MSAE module in the MSE, we conduct an ablation study of it on the final experimental performance, with the results presented in Table VI. As shown in Table VI, the training of *MSADiff* incorporating the MSAE module within the MSE leads to enhanced performance. This slight difference serves to validate the rationale and motivation behind this module, highlighting its effectiveness in contributing to the overall performance improvement of *MSADiff* in the marine snow estimation stage.

C. Data Set Quality Assessment

As previously emphasized, the quality of existing marine snow image data sets significantly impacts the generalization ability of algorithms. To prove this point and validate the superiority of our data set, we comprehensively evaluate the quality of currently available marine snow image data sets, including MSRBD [9], Snowy-VAROS [11], and our *MM-MSD*.

We adopt a controlled evaluation framework using the *MSADiff* model with uniform training parameters, ensuring a fair comparison across data sets. The model is trained on each data set individually and tested on open-source real-world marine snow images spanning different lighting conditions. As visualized in Fig. 8, the *MM-MSD*-trained model demonstrates distinct advantages: It effectively removes both fine granular snow and large flocculent aggregates while preserving critical details, such as coral textures.

Visual analysis reveals that our data set-trained model maintains consistent color fidelity across varying illumination scenarios and mitigates motion blur artifacts caused by dynamic snow particle movements. In contrast, models trained on Snowy-VAROS often produce over-smoothed backgrounds with residual snow clusters in shaded regions, while MSRBD-trained models introduce slight color shifts under low-light conditions. These qualitative disparities highlight the *MM-MSD* data set's unique strength in balancing synthetic and real-world samples, which equips it to address complex underwater imaging challenges. Collectively, these results confirm that our *MM-MSD* data set establishes a new benchmark for marine snow removal tasks, enabling the development of robust algorithms capable of adapting to real-world underwater environments.

VI. CONCLUSION, LIMITATIONS, AND FUTURE WORK

In this study, we propose *OmniMSI*, a comprehensive pipeline for simulating marine snow images through four key steps: 1) preparing raw marine snow particle data from EcoTaxa and real-world underwater scenes, 2) classifying morphotypes and analyzing spatio-temporal distribution patterns, 3) generating diverse masks guided by ecological realism, and 4) synthesizing authentic marine snow images with postprocessing techniques. Researchers can flexibly apply *OmniMSI* to generate tailored data sets by selecting parameters, such as ocean regions, depths, or seasons, enabling the creation of marine snow images that align with specific research needs. Leveraging this pipeline, we construct *MM-MSD*, the first data set featuring multimorphology marine snow particles and multidistribution patterns across ocean conditions. To advance marine snow removal, we design *MSADiff*, a diffusion-based method that integrates a MSE and remover, achieving state-of-the-art performance in both qualitative and quantitative evaluations. Furthermore, we establish a benchmark by systematically evaluating existing algorithms,

providing critical insights into their strengths and limitations for marine snow removal tasks.

While our work demonstrates significant progress, limitations remain. First, *OmniMSI* balances realism and practicality, but discrepancies persist between synthetic and real-world marine snow scenarios, particularly in dynamic underwater environments. Second, the computational demands of *MSADiff* hinder real-time deployment on resource-constrained underwater platforms. These challenges highlight opportunities for future research. Expanding *OmniMSI* to incorporate advanced fluid dynamics or light propagation models could enhance ecological fidelity. In addition, lightweight architectures or hardware-aware optimizations may bridge the gap between algorithmic performance and practical applicability. Beyond technical refinements, this work lays a foundation for interdisciplinary applications, such as improving underwater robotic vision systems or advancing marine ecological studies through enhanced image analysis. By addressing these directions, we aim to foster robust solutions that empower both computational research and marine exploration in complex underwater environments.

REFERENCES

[1] E. Trudnowska et al., "Marine snow morphology illuminates the evolution of phytoplankton blooms and determines their subsequent vertical export," *Nat. Commun.*, vol. 12, no. 1, 2021, Art. no. 2816.

[2] U. Passow, K. Ziervogel, V. Asper, and A. Diercks, "Marine snow formation in the aftermath of the deepwater horizon oil spill in the gulf of Mexico," *Environ. Res. Lett.*, vol. 7, no. 3, 2012, Art. no. 035301.

[3] S. Banerjee, G. Sanyal, S. Ghosh, R. Ray, and S. N. Shome, "Elimination of marine snow effect from underwater image-an adaptive probabilistic approach," in *Proc. IEEE Students' Conf. Elect. Electron. Comput. Sci.*, Bhopal, India, 2014, pp. 1–4.

[4] Y. Wang, X. Yu, D. An, and Y. Wei, "Underwater image enhancement and marine snow removal for fishery based on integrated dual-channel neural network," *Comput. Electron. Agriculture*, vol. 186, 2021, Art. no. 106182.

[5] X. Mei et al., "UIR-Net: A simple and effective baseline for underwater image restoration and enhancement," *Remote Sens.*, vol. 15, no. 1, 2022, Art. no. 39.

[6] D. Lee, G. Kim, D. Kim, H. Myung, and H.-T. Choi, "Vision-based object detection and tracking for autonomous navigation of underwater robots," *Ocean Eng.*, vol. 48, pp. 59–68, 2012.

[7] L. Cai, N. E. McGuire, R. Hanlon, T. A. Mooney, and Y. Girdhar, "Semi-supervised visual tracking of marine animals using autonomous underwater vehicles," *Int. J. Comput. Vis.*, vol. 131, no. 6, pp. 1406–1427, 2023.

[8] Q. Jiang, Y. Chen, G. Wang, and T. Ji, "A novel deep neural network for noise removal from underwater image," *Signal Process. Image Commun.*, vol. 87, 2020, Art. no. 115921.

[9] R. Kaneko, Y. Sato, T. Ueda, H. Higashi, and Y. Tanaka, "Marine snow removal benchmarking dataset," in *Proc. Asia Pacific Signal Inf. Process. Assoc. Annu. Summit Conf., Taipei Int. Conventional Center*, Taipei, Taiwan, 2023, pp. 771–778.

[10] F. Galetto and G. Deng, "A deep learning approach for marine snow synthesis and removal," *Signal Image Video P.*, vol. 19, no. 1, p. 1, 2025.

[11] L. M. Hodne, E. Leikvoll, M. Yip, A. L. Teigen, A. Stahl, and R. Mester, "Detecting and suppressing marine snow for underwater visual SLAM," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New Orleans, LA, USA, 2022, pp. 5097–5105.

[12] D. Guo, Y. Huang, T. Han, H. Zheng, Z. Gu, and B. Zheng, "Marine snow removal," in *Proc. OCEANS 2022-Chennai*. Chennai, India, 2022, pp. 1–7.

[13] Z. Zhao and X. Li, "Image blending-based noise synthesis and attention-guided network for single image marine snow denoising," *Int. J. Mach. Learn. Cybern.*, vol. 1, pp. 2205–2219, 2023.

[14] J. Xiao, K. A. Ehinger, J. Hays, A. Torralba, and A. Oliva, "Sun database: Exploring a large collection of scene categories," *Int. J. Comput. Vis.*, vol. 119, no. 1, pp. 3–22, 2016.

[15] C. Li et al., "An underwater image enhancement benchmark dataset and beyond," *IEEE Trans. Image Process.*, vol. 29, pp. 4376–4389, 2020.

[16] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robot. Automat. Lett.*, vol. 5, no. 2, pp. 3227–3234, Apr. 2020.

[17] R. Li, L.-F. Cheong, and R. T. Tan, "Heavy rain image restoration: Integrating physics model and conditional adversarial learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Long Beach, CA, USA, 2019, pp. 1633–1642.

[18] P. G. O. Zwilgmeyer, M. Yip, A. L. Teigen, R. Mester, and A. Stahl, "The VAROS synthetic underwater data set: Towards realistic multi-sensor underwater data with ground truth," in *Proc. IEEE Int. Conf. Comput. Vis.*, Montreal, BC, Canada, 2021, pp. 3715–3723.

[19] F. Farhadifard, M. Radolko, and U. F. v. Lukas, "Single image marine snow removal based on a supervised median filtering scheme," in *Proc. Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, Lisbon, Portugal, 2017, pp. 280–287.

[20] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *J. Roy. Stat. Soc.: Ser. B*, vol. 63, no. 2, pp. 411–423, 2001.

[21] L. Vilgrain et al., "Trait-based approach using in situ copepod images reveals contrasting ecological patterns across an arctic ice melt zone," *Limnol. Oceanogr.*, vol. 66, no. 4, pp. 1155–1167, 2021.

[22] W. Wang et al., "PVT V2: Improved baselines with pyramid vision transformer," *Comput. Vis. Media*, vol. 8, no. 3, pp. 415–424, 2022.

[23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA: IEEE, 2018, pp. 4510–4520.

[24] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New Orleans, LA, USA, 2022, pp. 10674–10685.

[25] P. Esser, R. Rombach, and B. Ommer, "Taming transformers for high-resolution image synthesis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Nashville, TN, USA, 2021, pp. 12868–12878.

[26] D. P. Kingma and J. Ba, "ADAM: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[27] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," in *Proc. Int. Conf. Mach. Learn. Virtual Event*, Austria: OpenReview.net, 2021, pp. 8162–8171.

[28] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *Proc. IEEE Int. Conf. Robot. Automat.*, Brisbane, QLD, Australia, 2018, pp. 7159–7165.

[29] C. Guo et al., "Underwater ranker: Learn which is better and how to be better," in *Proc. AAAI Conf. Artif. Intell.*. Washington, DC, USA, 2023, pp. 702–709.

[30] L. Peng, C. Zhu, and L. Bian, "U-shape transformer for underwater image enhancement," *IEEE Trans. Image Process.*, vol. 32, pp. 3066–3079, 2023.

[31] Y. Tang, H. Kawasaki, and T. Iwaguchi, "Underwater image enhancement by transformer-based diffusion model with non-uniform sampling for skip strategy," in *Proc. ACM Int. Conf. Multimedia*, 2023, pp. 5419–5427.

[32] C. Zhao, C. Dong, and W. Cai, "Learning a physical-aware diffusion model based on transformer for underwater image enhancement," 2024, *arXiv:2403.01497*.

[33] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[34] Z. Yue, H. Yong, Q. Zhao, D. Meng, and L. Zhang, "Variational denoising network: Toward blind noise modeling and removal," in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, 2019.

[35] Z. Yue, Q. Zhao, L. Zhang, and D. Meng, "Dual adversarial network: Toward real-world noise removal and noise generation," in *Proc. Eur. Conf. Comput. Vis.*, Cham, Switzerland, 2020, pp. 41–58.

[36] S. W. Zamir et al., "Multi-stage progressive image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Nashville, TN, USA, 2021, pp. 14816–14826.

[37] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New Orleans, LA, USA, 2022, pp. 5718–5729.

[38] Y.-F. Liu, D.-W. Jaw, S.-C. Huang, and J.-N. Hwang, "DesnowNet: Context-aware deep network for snow removal," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 3064–3073, Jun. 2018.

[39] W.-T. Chen et al., "All snow removed: Single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss," in *Proc. IEEE Int. Conf. Comput. Vis.*, Montreal, QC, Canada, 2021, pp. 4176–4185.
[40] K. Zhang, R. Li, Y. Yu, W. Luo, and C. Li, "Deep dense multi-scale network for snow removal using semantic and depth priors," *IEEE Trans. Image Process.*, vol. 30, pp. 7419–7431, 2021.
[41] W.-T. Chen, Z.-K. Huang, C.-C. Tsai, H.-H. Yang, J.-J. Ding, and S.-Y. Kuo, "Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: Toward a unified model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New Orleans, LA, USA, 2022, pp. 17632–17641.
[42] J. M. J. Valanarasu, R. Yasarla, and V. M. Patel, "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New Orleans, LA, USA, 2022, pp. 2343–2353.
[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[44] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE J. Ocean. Eng.*, vol. 41, no. 3, pp. 541–551, Jul. 2016.

Yiqing Huang received the B.E degree in telecommunications engineering from Shandong Normal University, Jinan, China, in 2020, and the M.E. degree in electronic and communication engineering from the Ocean University of China, Qingdao, China, in 2023. She is currently working toward the Ph.D. degree in artificial intelligence with the Macau University of Science and Technology, Macau, China.

Her research interests include deep learning, underwater vision, and sign language processing.

Tianshun Han received the B.E degree in Internet of Things from the Xi'an University of Posts and Telecommunications, Xi'an, China, in 2020, and the M.E. degree in electronic and communication engineering from the Ocean University of China, Qingdao, China, in 2023. He is currently working toward the Ph.D. degree in artificial intelligence with the Macau University of Science and Technology, Macau, China.

His research interests include deep learning, underwater vision, and metaverse.

Haoru Zhao received the B.E degree in computer science and technology from Qingdao Agricultural University, Qingdao, China, in 2016, the M.E. degree in software engineering from the Qingdao University of Science and Technology, Qingdao, in 2019, and the Ph.D. degree in intelligence information and communication systems from the Ocean University of China, Qingdao, in 2024.

She is currently working as the Faculty of Shandong Police College, Jinan, China. Her research interests include image processing, underwater vision, and computer vision.

Yanyan Liang (Member, IEEE) received the B.S. degree in communication engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2004, and the M.S. and Ph.D. degrees in information technology and computer technology and its applications from the Macau University of Science and Technology (MUST), Taipa, Macau, in 2006 and 2009, respectively.

He is currently an Associate Professor with MUST. His research interests include computer vision, image processing, and machine learning.

Jun Wan (Senior Member, IEEE) received the B.E. degree from the China University of Geosciences, Beijing, China, in 2008, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University, Beijing, in 2015.

Since January 2015, he has been a Faculty Member with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science, Beijing, where he currently serves as an Associate Professor. His main research interests include computer vision and machine learning, especially for facial analysis, gesture, and action recognition.

Sergio Escalera (Senior Member, IEEE) received the B.S. and M.S. degrees from the Universitat Autònoma de Barcelona (UAB), Barcelona, Spain, in 2003 and 2005, respectively. He received the Ph.D. degree in multiclass visual categorization systems from the Computer Vision Center, UAB, in 2008. He is Full Professor with the Universitat de Barcelona, Barcelona, Spain, distinguished Professor with Aalborg University, Aalborg, Denmark, and Member of the Computer Vision Center, Barcelona, Spain. He is vice President of ChaLearn Challenges in machine learning. He is the Fellow of the European Laboratory for Learning and Intelligent Systems. His research interests include inclusive, transparent, and fair analysis of humans from visual and multimodal data.

Haiyong Zheng (Senior Member, IEEE) received the B.E. degree in electronic information engineering and the Ph.D. degree in ocean information sensing and processing from the Ocean University of China, Qingdao, China, in 2004 and 2009, respectively.

In 2009, he joined the College of Electronic Engineering, Ocean University of China, where he is currently a Professor. His research interests include computer vision, underwater vision, and deep learning.