

Marine Snow Remove

선행 연구 리뷰

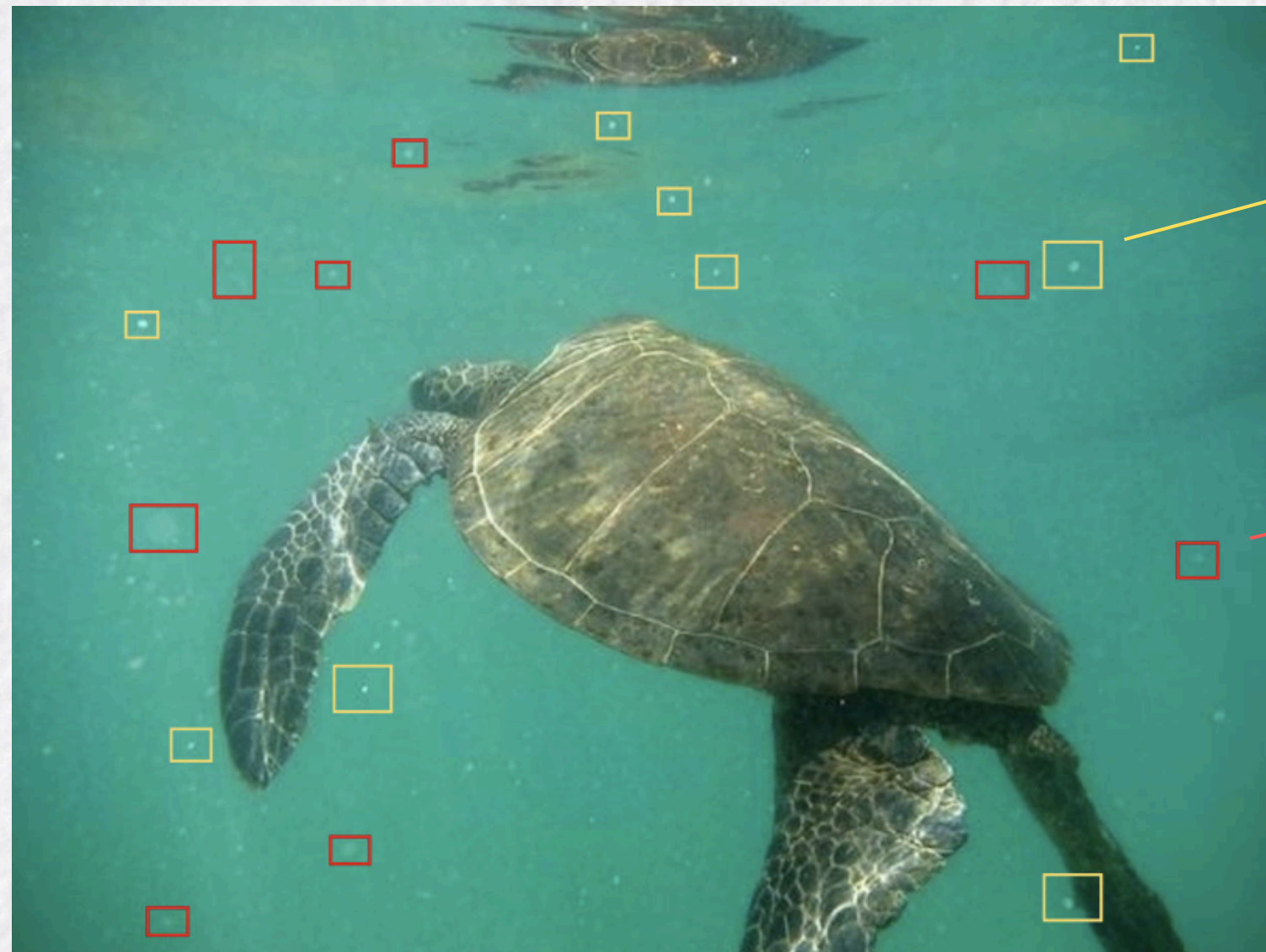
Deep Shark Lab 김태빈
2026.08.04

목차

- 01** 배경
- 02** OmniMSI 파이프라인
- 03** MM-MSD 데이터셋
- 04** MSADiff 모델
- 05** 실험 및 결과
- 06** 결론

Marine Snow 란 | 배경

수중 환경의 유기물 및 무기물 불순물들이 응집되어 형성된 다양한 형태의 입자



Light spots:

주로 불투명하거나 투명도가 낮은 입자들에 의해 발생

Veil-like blurs:

투명도가 높은 입자들이 이미지 전반에 퍼져 있을 때 나타남

문제점 | 배경

- **가시성 저하:** 빛의 산란을 유발하여 이미지의 선명도를 낮춤
- **객체 가림:** 입자들이 실제 객체를 가리거나 이미지상의 정보와 섞이는 현상

수중 시각 작업(수중 객체 탐지 및 수중 타겟 추적)의 발전 방해

기존 한계

| 배경

지나친 단순화

단순히 이미지에 작은 점들을 추가하는 방식

→ 실제 marine snow가 가지는 외형적 특징을 제대로 표현하지 못함

형태 과소평가

다양한 형태를 고려하여 수학적 모델이나 생성 모델을 사용

→ 관찰 데이터의 부족으로 인해 실제와 동떨어진 형태를 재현

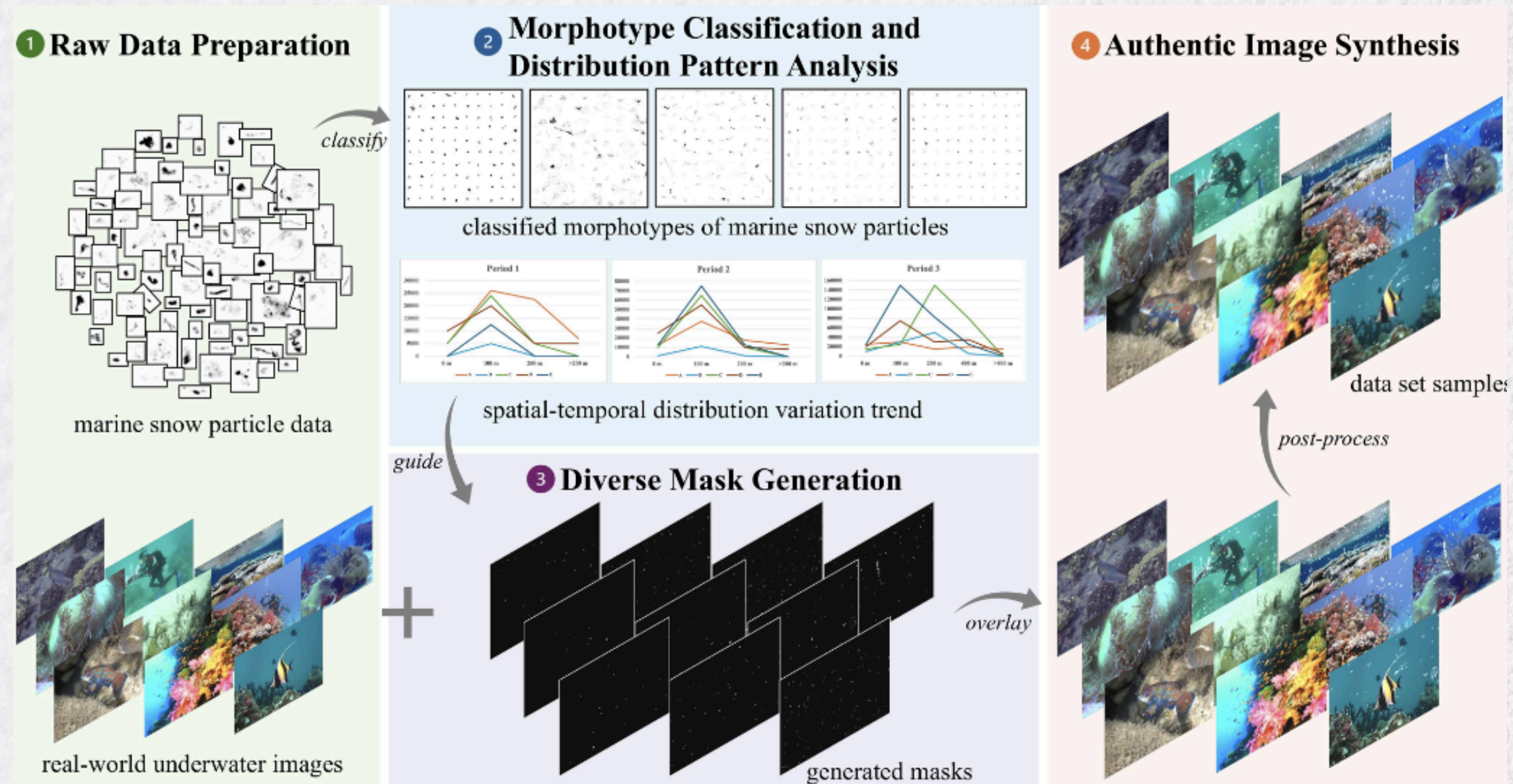
분포 패턴의 무시

최근 노력들은 실제 marine snow 입자 이미지를 사용

→ 해양 지역, 수심, 계절에 따라 달라지는 생태학적 분포 패턴을 반영하지 못함

OmniMSI 파이프라인

OmniMSI 파이프라인



원시 데이터 준비

분포 및 형태 분석

다양한 마스크 생성

실제 이미지 합성

실제 데이터의 확보

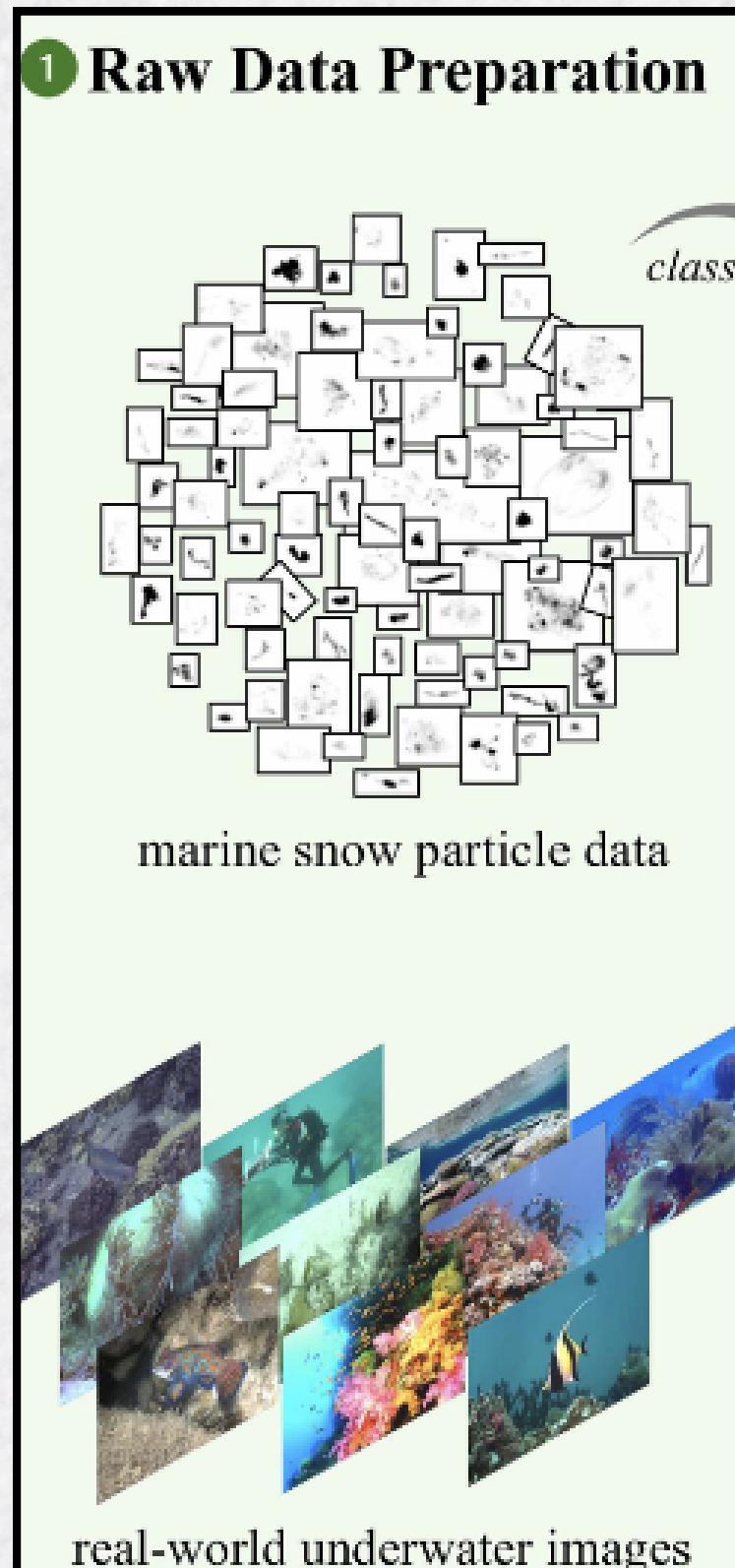
'환경에 따른 분포'와 입자의 '모양' 파악

입자들을 규칙에 맞게 배치

배경과 마린스노우의 자연스러운 혼합
6페이지

원시 데이터 준비

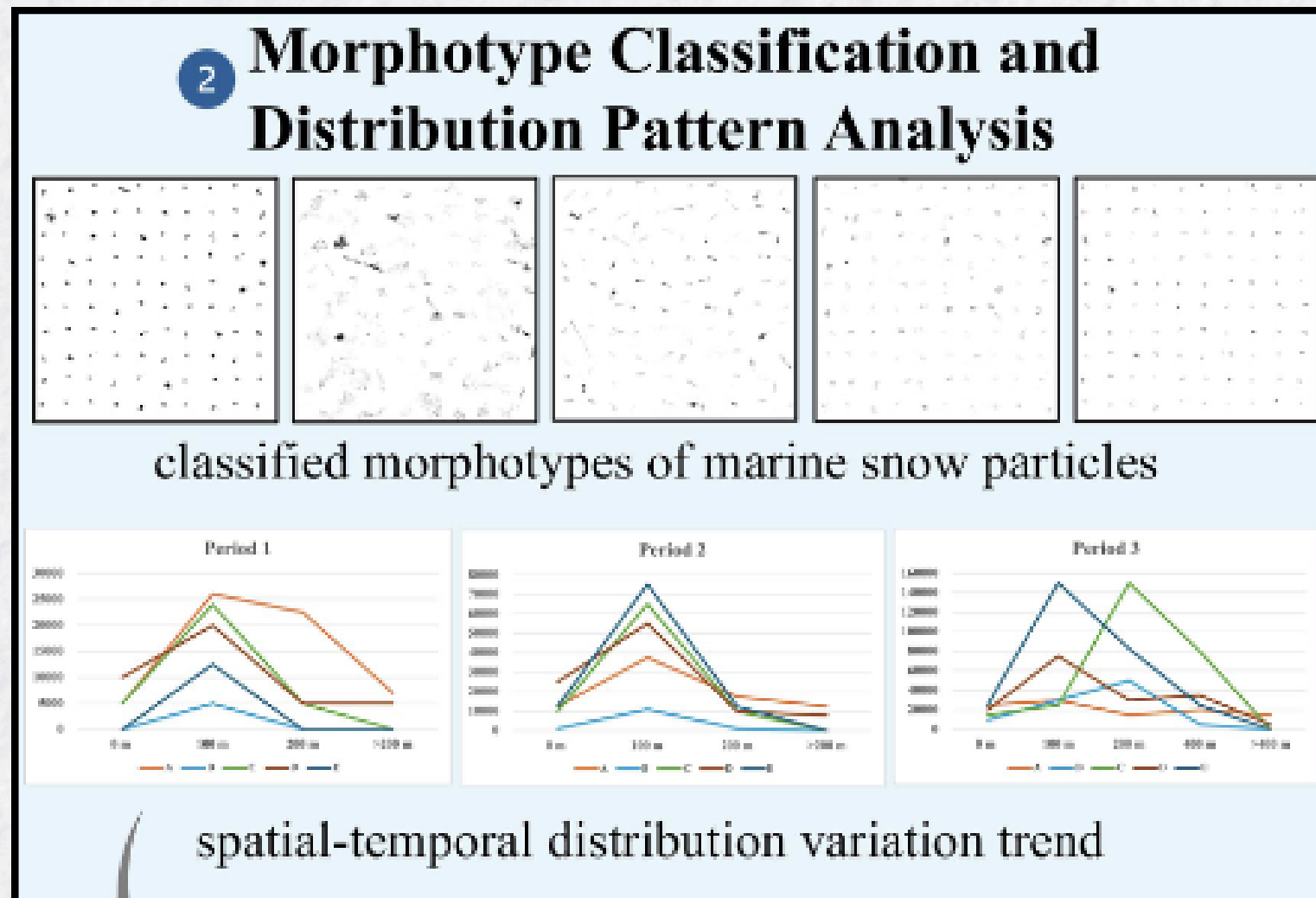
| OmniMSI 파이프라인



- **핵심:** 두 종류의 데이터 확보
- 'ECOTAXA'라는 실제 해양 입자 데이터 플랫폼에서 수집한 약 100만개 마린스노우 입자 이미지를 사용
- 마린스노우가 없는 깨끗한 수중 이미지를 데이터셋 (UIEB, EUVP 등)에서 가져와 배경으로 활용

분포 및 형태 분석

| OmniMSI 파이프라인



- **핵심:** 입자의 모양과 환경에 따른 분포 파악
- **분포 분석:** 해양 환경(깊이, 계절 등)에 따라 입자의 밀도가 어떻게 변하는지 통계적으로 분석하여, 실제 환경과 유사한 '분포 패턴'을 정의
- **클러스터링:** 수집된 입자들을 크기, 모양, 질감 등 24가지 특징을 기반으로 PCA와 K-MEANS 알고리즘을 사용하여 5가지 형태로 분류

크기6
모양5
질감13

입자 하나마다 24가지 특징을 PCA로 변환했을 때 만
들어진 상위 10개의 주성분, 즉 PC1~PC10이 전체 정보
를 얼마나 설명하는지 보여줌

TABLE III TOP-10 PRINCIPAL COMPONENTS (# PC) OF THE 24 MORPHOLOGICAL CHARACTERISTICS, INCLUDING EACH COMPONENT'S EIGENVALUE (# EV), VARIATION CONTRIBUTION (# CONTRL.) 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
# Contrb. (%)	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

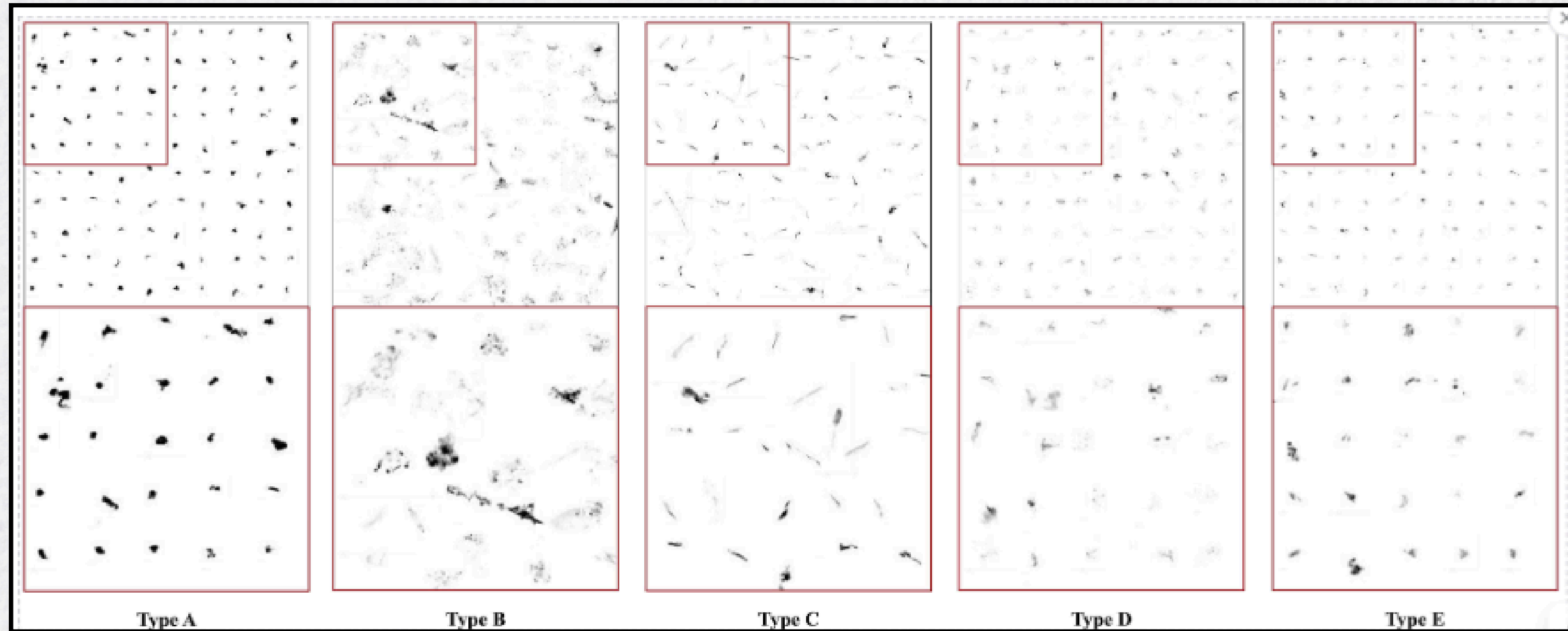
연구진은 상위 4개 주성분만으로도 원래 24가지 특징 정보
의 약 86.64%를 유지할 수 있다고 판단

PC1~PC4만 K-means에 입력했고, Gap Statistic으
로 적절한 군집 수를 계산한 결과 K=5가 나와 최종적으
로 입자를 5가지 형태로 분류

# 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	$trim(x)$
symetrie h	bilateral horizontal symmetry index	shape	$\log_{10} trim(x)$
symetrie v	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	$trim(x)$

5가지 형태 분류

| OmniMSI 파이프라인



Type A: 작고 규칙적·조밀

Type B: 크고 불규칙·느슨

Type C: 선형·타원형

Type D: 투명·안개형

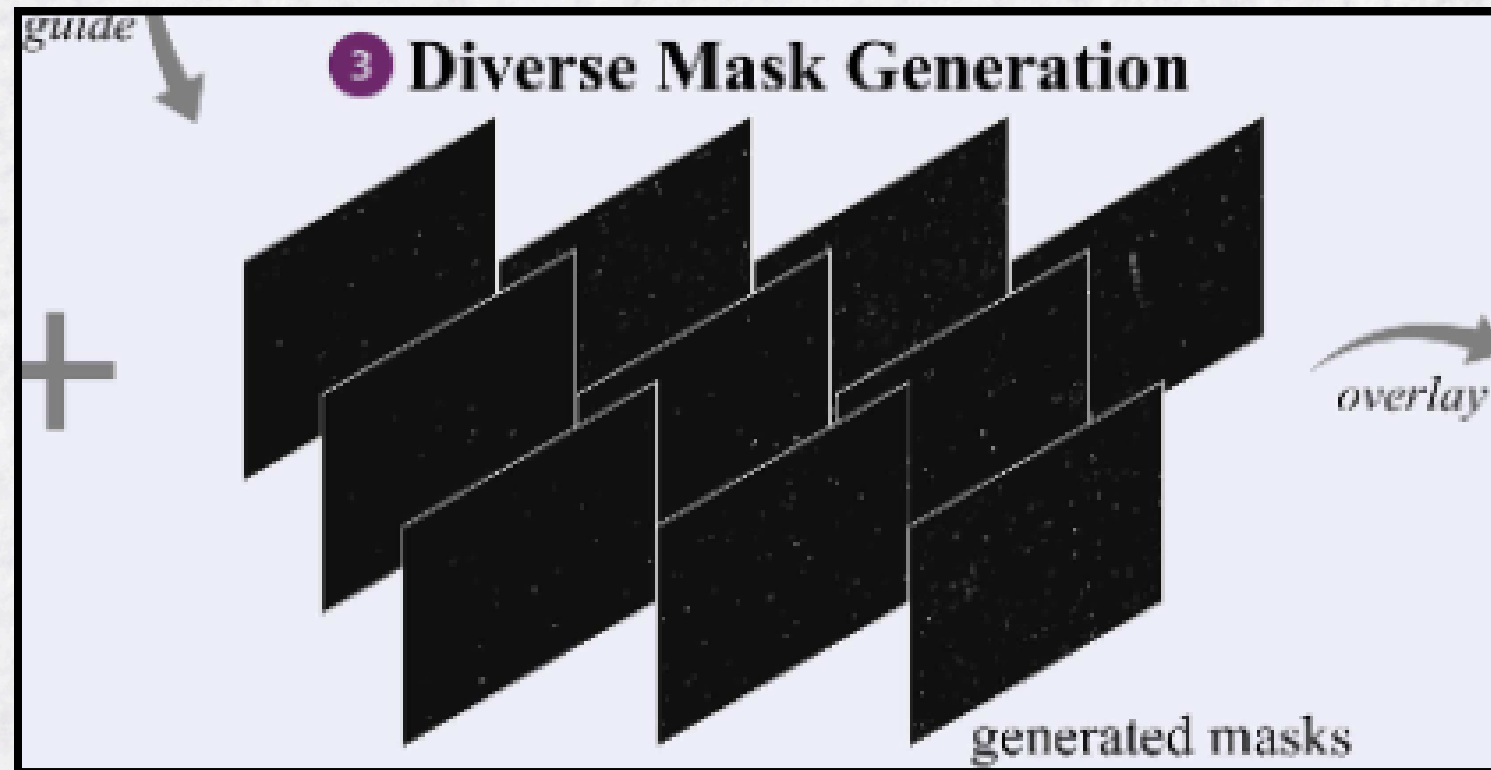
Type E: 매우 작고 규칙적·조밀

Type A·E → 밝은 점 형태

Type B·C·D → 베일처럼 흐릿한 형태

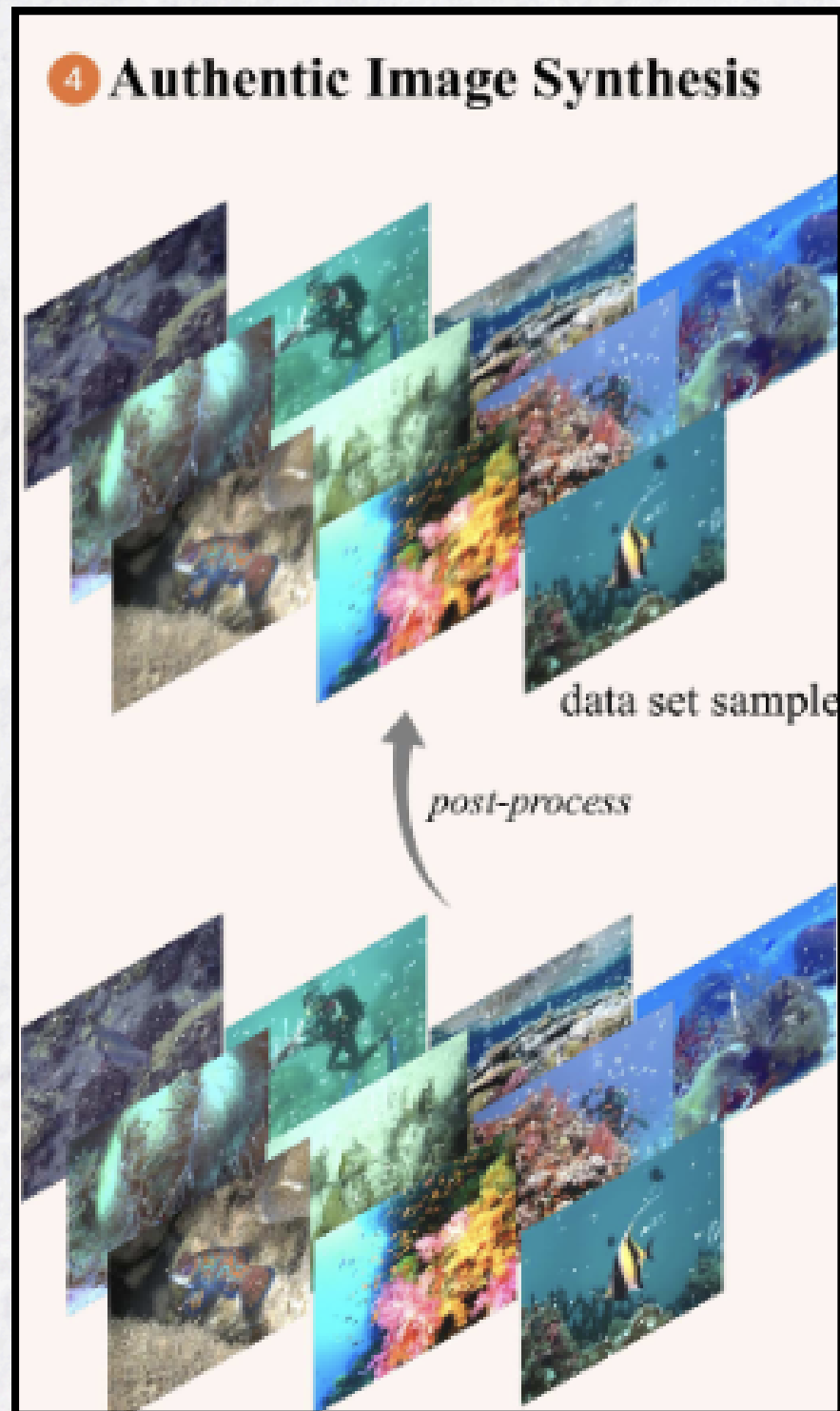
다양한 마스크 생성

| OmniMSI 파이프라인



- **핵심:** 입자들을 규칙에 맞게 배치
- 전 단계에서 분석한 분포 패턴을 가이드 삼아, 분류된 입자들을 마스크 위에 배치
- 입자의 크기나 회전 각도를 랜덤하게 조정하여 실제 자연스러운 느낌을 줌

실제 이미지 합성 | OmniMSI 파이프라인



- **핵심:** 배경과 마린스노우의 자연스러운 혼합
- **합성:** 생성된 마스크를 실제 수중 이미지 위에 중첩 시킴
- **후처리:**
 - 가우시안 블러: 빛의 산란 효과 재현
 - 동적 블러: 카메라와 물의 움직임에 의한 잔상 효과 재현

MM-MSD 데이터셋

| Multi-morphology and Multidistribution Marine Snow image Dataset

데이터 구성

- 마린 스노우 이미지 + 깨끗한 참조 이미지
참조 이미지 = Ground Truth

형태 및 분포 다양성

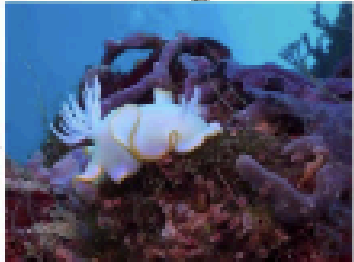
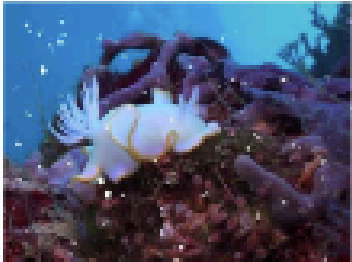
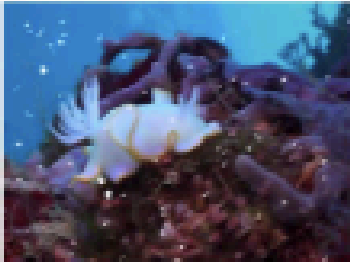
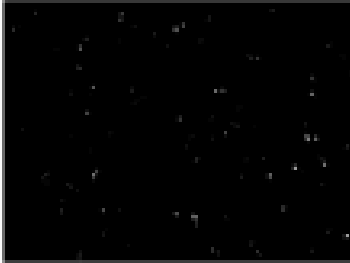
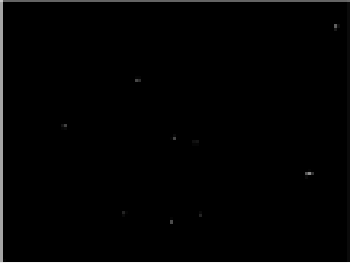
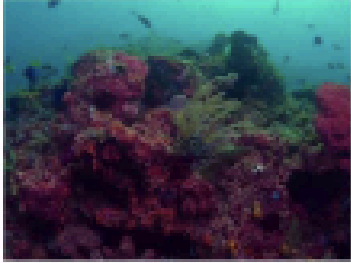
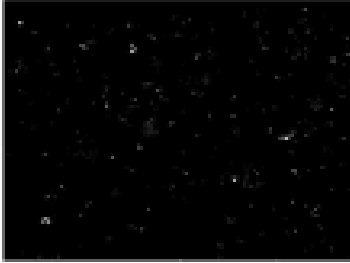
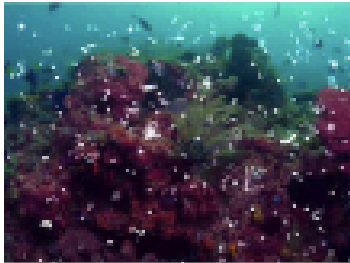
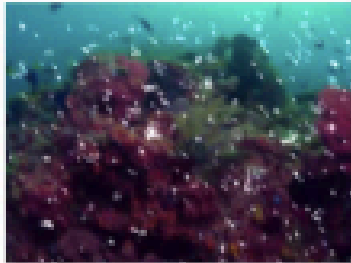
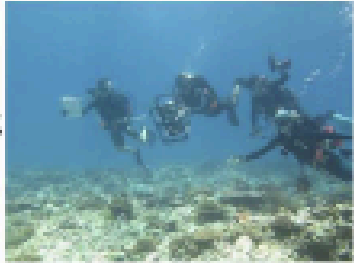
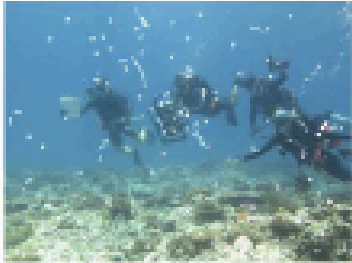
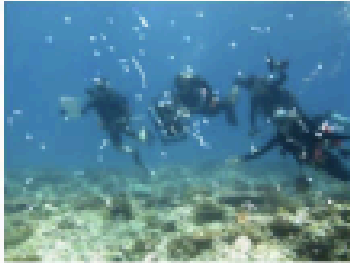
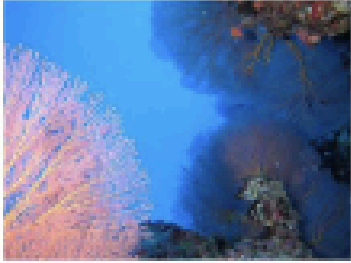
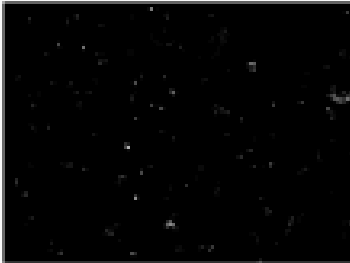
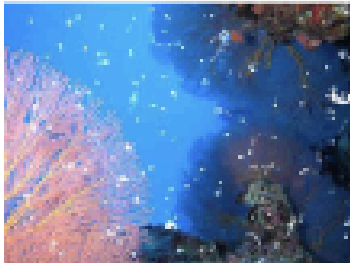
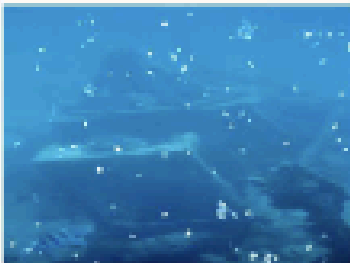
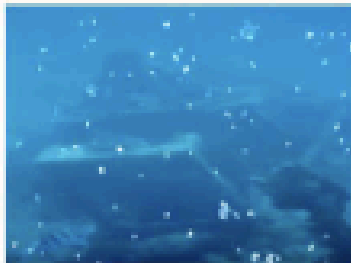
- 5가지 입자 형태
- 10가지 분포 패턴

데이터 규모

- 패턴당 600개 × 10가지 분포 = 총 6,000개 이미지 쌍
학습 4,800 · 검증 600 · 테스트 600

MM-MSD 예시

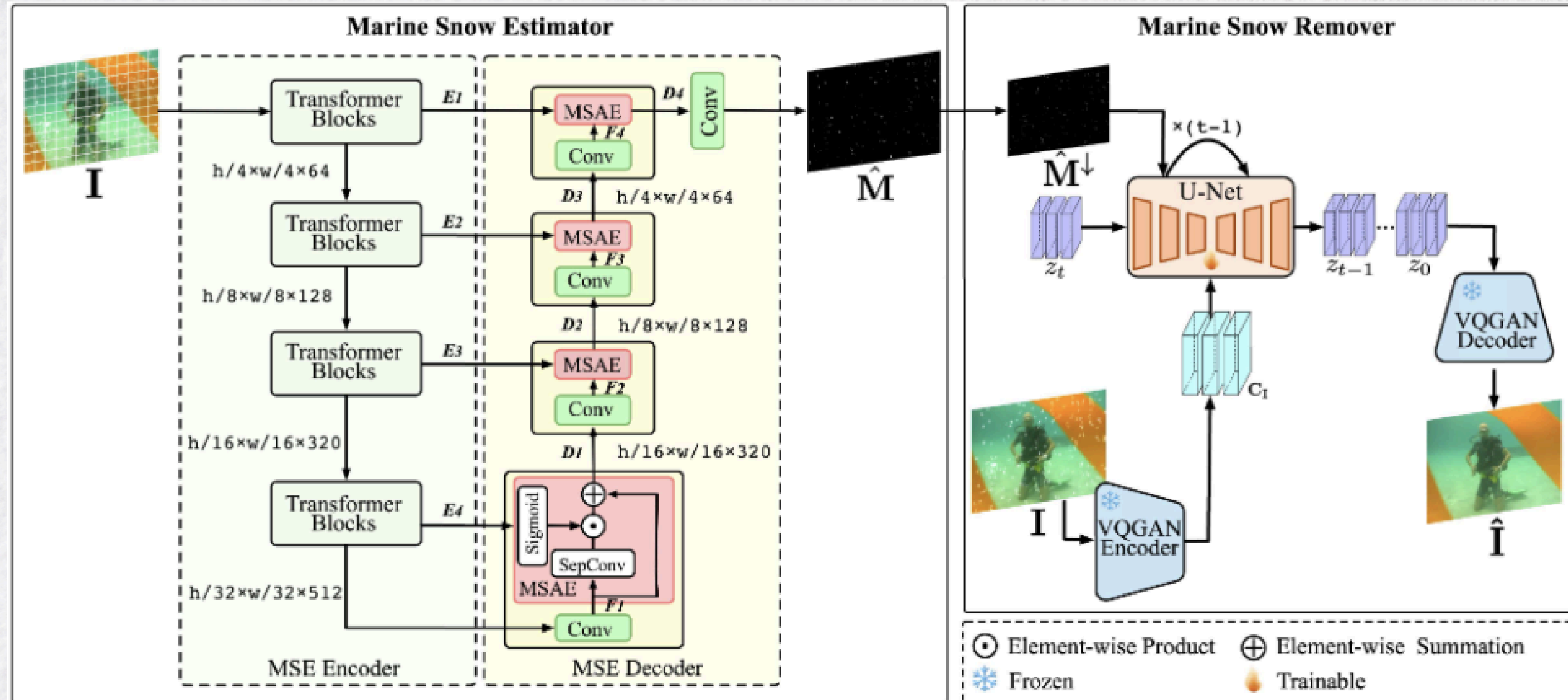
| MM-MSD 데이터셋

	Real-World Underwater Images	Marine Snow Masks	Direct Overlay Results	Post-Process Results		Real-World Underwater Images	Marine Snow Masks	Direct Overlay Results	Post-Process Results	
Period 1 Type 1										Period 2 Type 3
Period 1 Type 2										Period 3 Type 1
Period 1 Type 3										Period 3 Type 2
Period 2 Type 1										Period 3 Type 3
Period 2 Type 2										Period 3 Type 4

MSADiff 모델 | Marine Snow-Aware Diffusion

마린 스노우 추정기

마린 스노우 제거기



모델 핵심 기술

| MSADiff 모델

PVTv2

Pyramid Vision Transformer v2

기존 Transformer를 이미지 처리에 맞게 발전시킨 모델

역할: 이미지 전체의 문맥과 구조적 정보를 파악

- **피라미드 구조:** 단계가 깊어질수록 해상도는 줄고 특징 채널은 증가
- 세부적인 위치 정보부터 이미지 전체의 문맥까지 단계적으로 추출

MSAE 모듈

Marine Snow-Aware Enhancement

저자들이 직접 설계하고 제안한 모듈

역할: 마린 스노우로 보이는 특징을 강조하여 배경과 구분하고, 입자의 위치를 정확하게 찾아냄

- 여러 단계의 특징을 결합
- 마린 스노우와 관련된 특징을 강조

VQGAN

Vector Quantized GAN

기존 이미지 압축·복원 모델

- 이미지를 압축하고 다시 복원
- 중요한 구조와 질감 정보는 유지
- 복원된 특징을 다시 이미지로 변환

인코더: 원본 이미지를 중요한 특징만 담은 작은 크기의 이미지로 반환(이미지 압축)

디코더: 인코더에서 처리된 특징을 다시 사람이 볼 수 있는 이미지로 변환(이미지 복원)

Marine Snow Estimator(MSE) | MSADiff 모델

- **역할:** 입력된 이미지에서 마린스노우가 "어디에, 얼마나" 있는지를 찾아냄

입력 이미지 $I \longrightarrow$ MSE 인코더

MSE 인코더에 PVTv2를 적용하여 입력 이미지를 네 단계로 분석

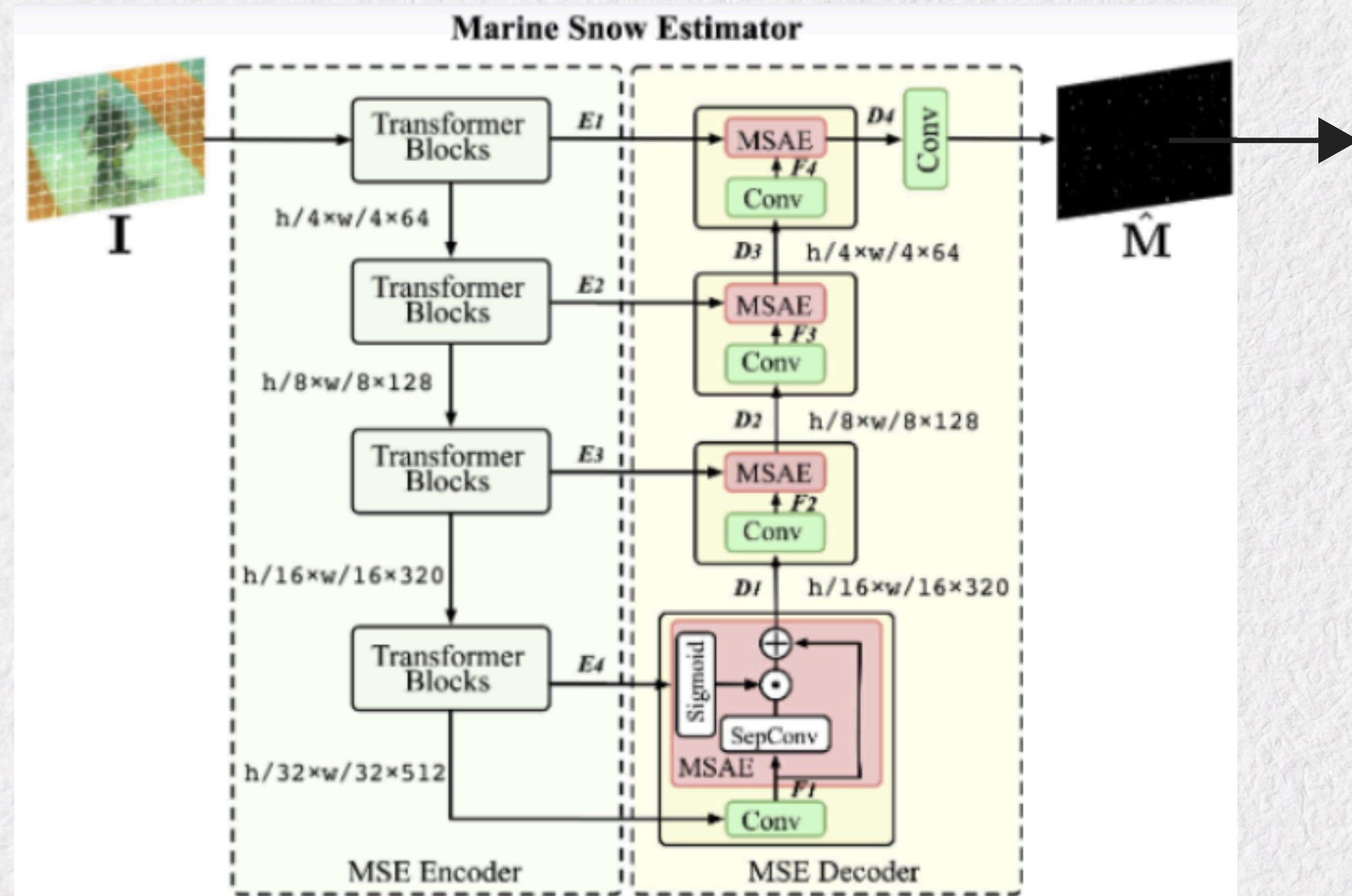
E_1	$(h/4 \times w/4 \times 64)$
E_2	$(h/8 \times w/8 \times 128)$
E_3	$(h/16 \times w/16 \times 320)$
E_4	$(h/32 \times w/32 \times 512)$

MSAE Decoder: D1부터 D4

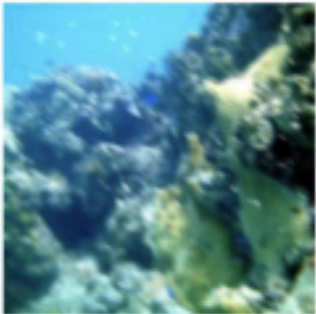
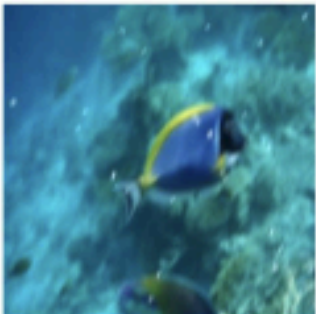
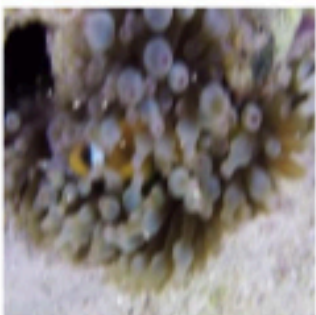
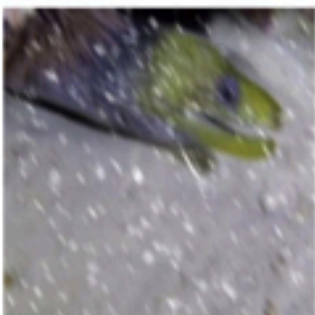
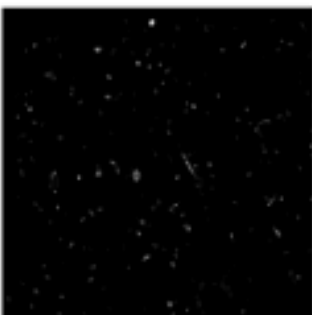
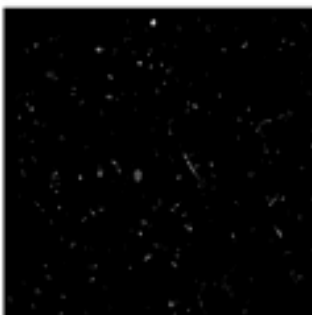
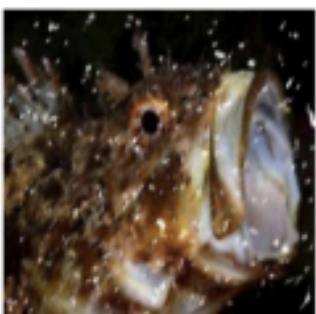
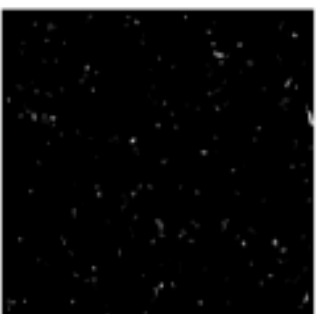
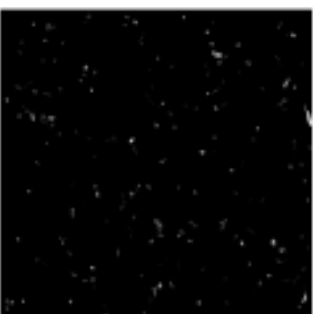
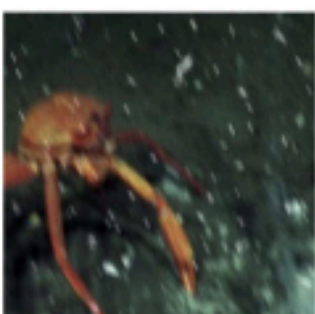
깊은 특징 E_4 부터 이미지 크기를 다시 키움

이과정에서 추출한 $E1 \sim E4$ 정보, 복원 중인 $D1 \sim D4$ 정보 **결합**
 각 단계의 Conv는 **결합**한 특징을 정리하고 다음 단계로 전달

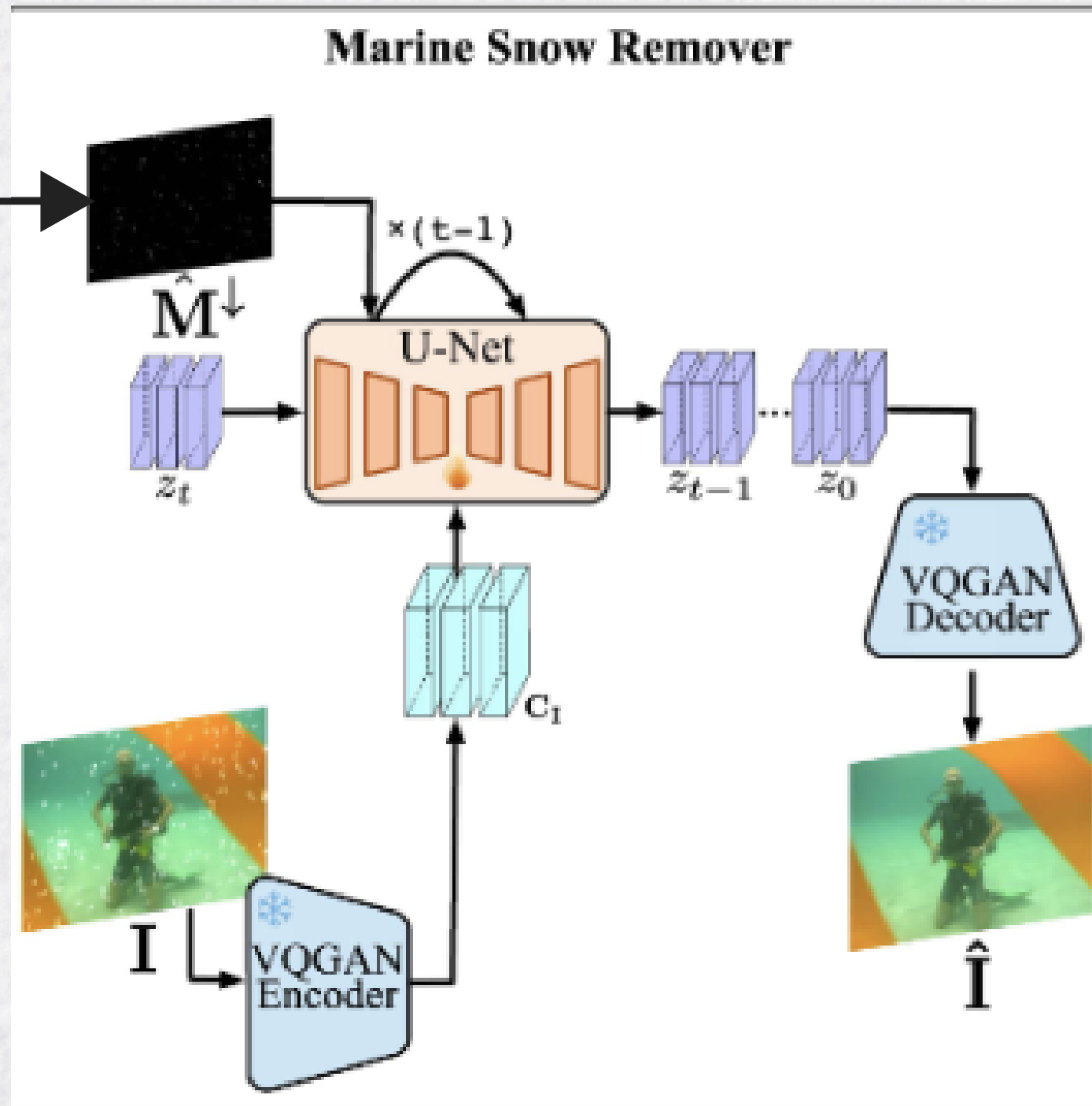
마지막 Conv가 특징을 한 채널로 변환하면
 마린 스노우 위치를 나타내는 이진 마스크 $\hat{M}$ 이 만들어짐



추정기 마스크 결과 (MSE) | MSADiff 모델

	marine snow images	estimated masks	ground truth		marine snow images	estimated masks	ground truth	
Period 1 Type 1								Period 2 Type 3
Period 1 Type 2								Period 3 Type 1
Period 1 Type 3								Period 3 Type 2
Period 2 Type 1								Period 3 Type 3
Period 2 Type 2								Period 3 Type 4

Marine Snow Remover | MSADiff 모델



역할: 추정된 마스크를 참고하여 마린 스노우를 실제로 제거하고 이미지를 복원

이미지 $I \rightarrow$ VQGAN 인코더

이미지의 배경과 물체, 색상 및 구조 정보를 압축한 특징 C_I

U-Net에는 총 세 가지 정보가 전달 $z_t, C_I, \hat{M}$

U-Net은

$z_t \rightarrow z_{t-1} \rightarrow z_0$

$z_0 \rightarrow$ VQGAN 디코더 $\rightarrow \hat{I}$

입력 이미지가 복원할 장면을 알려주고, 마스크가 복원할 위치를 알려주며, U-Net이 노이즈에서 깨끗한 이미지를 단계적으로 생성

TABLE V
IMPACT OF DIFFERENT TRAINING APPROACHES

Training Strategy	PSNR ($\uparrow$)	SSIM ($\uparrow$)	UIQM ($\uparrow$)
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 ($\uparrow$)	SSIM ($\uparrow$)	UIQM ($\uparrow$)
$\times$	16.2518	0.3019	2.1813
$\checkmark$	16.8126	0.3942	2.9262

벤치마크 평가 지표

실험 및 결과

기존 수중 이미지 개선

PSNR:복원 이미지가 깨끗한 정답 이미지와 픽셀 단위로 얼마나 비슷한지

SSIM:이미지의 구조·윤곽·질감이 정답과 얼마나 비슷한지

UIQM:수중 이미지의 색상·선명도·대비를 종합적으로 평가

잡음 제거

마린 스노우 제거 모델

MSADiff (Ours)

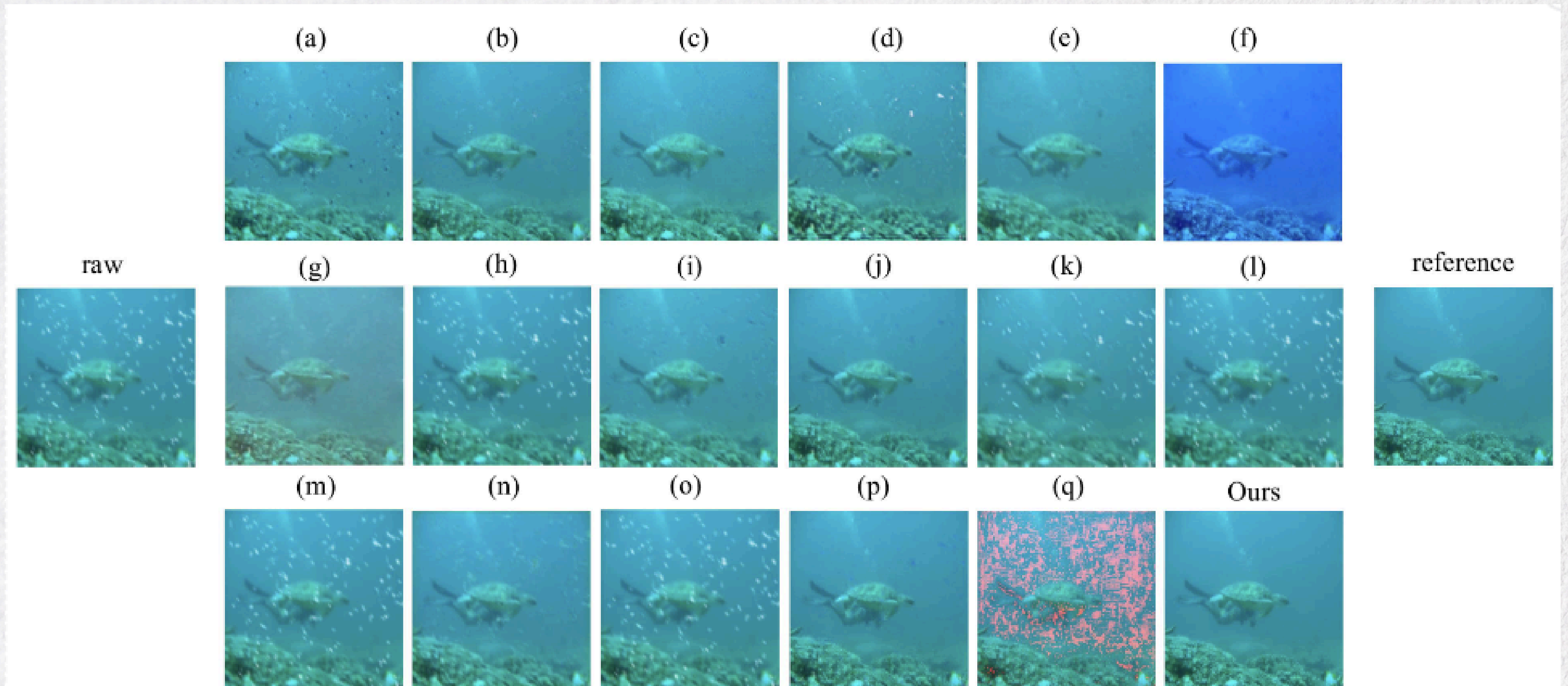
PSNR (↑)	SSIM (↑)	UIQM (↑)
16.8126	0.3942	2.9262

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
	<i>MSADiff</i> (Ours)	16.8126	0.3942	2.9262

성능 비교

| 실험 및 결과

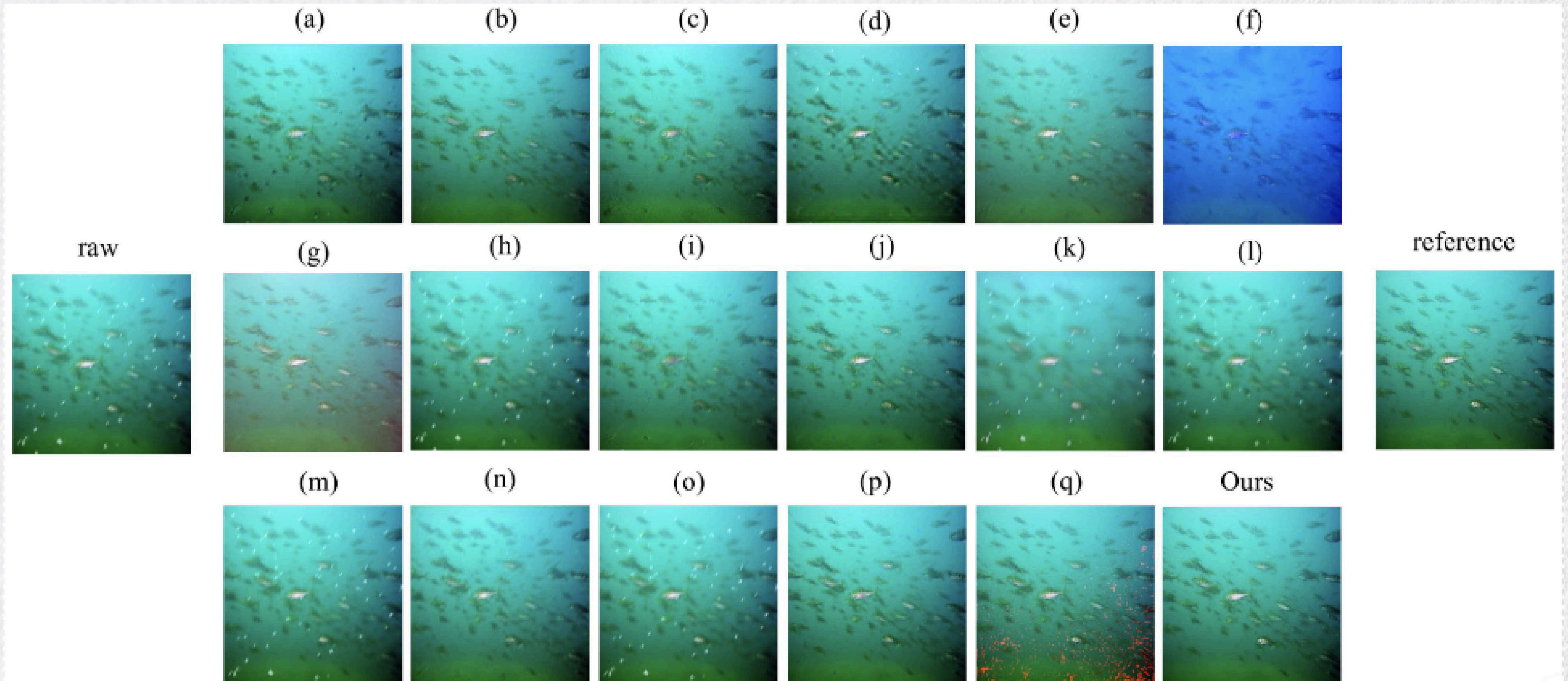
(a)~(g) 수중 이미지 개선 (h)~(l) 일반 잡음 제거 (m)~(q) snow 제거



성능 비교

| 실험 및 결과

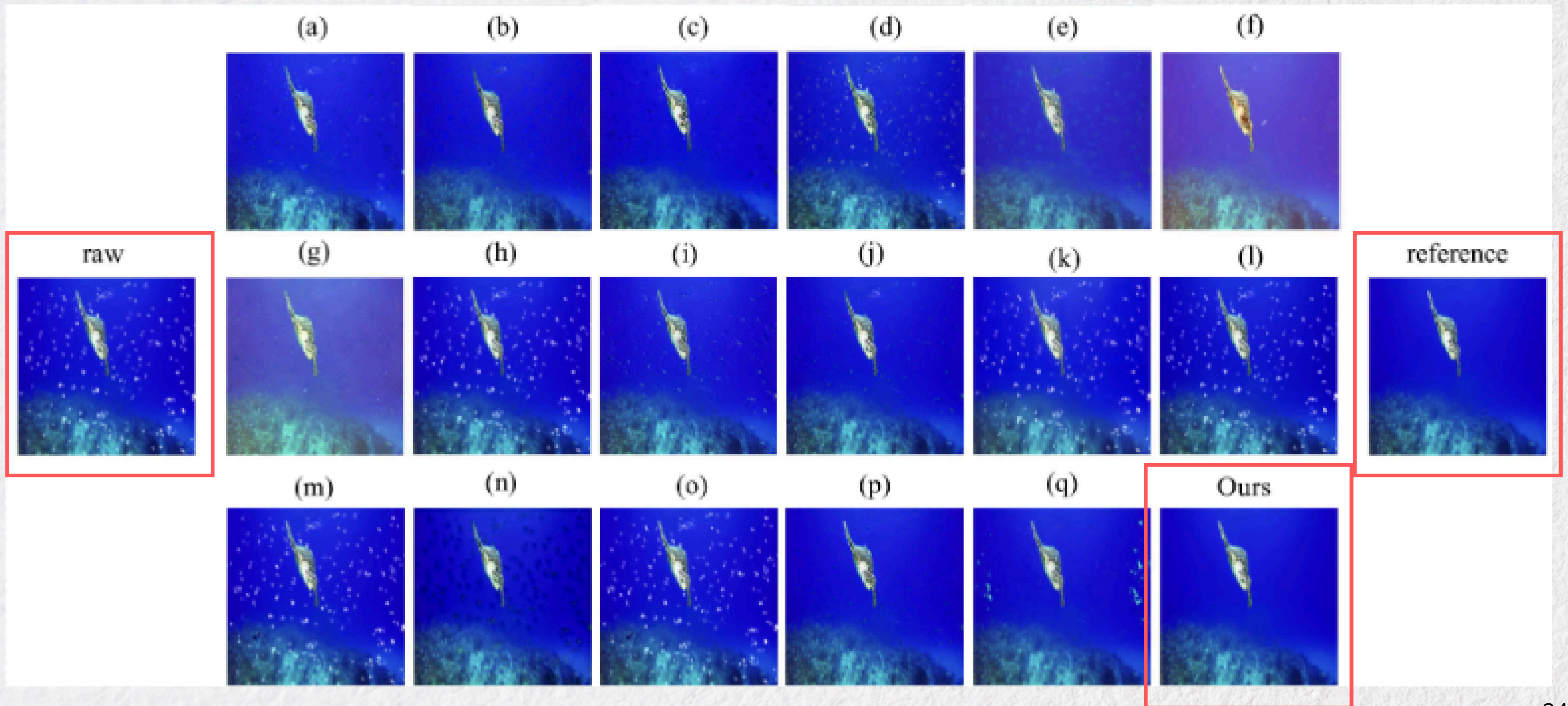
(a)~(g) 수중 이미지 개선 (h)~(l) 일반 잡음 제거 (m)~(q) snow 제거



성능 비교

| 실험 및 결과

(a)~(g) 수중 이미지 개선 (h)~(l) 일반 잡음 제거 (m)~(q) snow 제거



핵심 기여 | 결론

1. OmniMSI 파이프라인

실제 입자의 형태와 환경별 분포를 반영한 범용 마린 스노우 이미지 생성 파이프라인

2. MM-MSD 데이터셋

5가지 형태와 10가지 분포를 포함한 총 6,000쌍의 데이터셋

3. MSADiff 모델

마린 스노우 위치를 먼저 추정하고 확산 모델로 단계적으로 복원

4. 새로운 벤치마크

기존 수중 이미지 개선·잡음 제거·눈 제거 모델을 정량·정성적으로 비교

한계 및 향후 과제 | 결론

합성 데이터와 실제 환경 간의 격차

- **한계점:** OmniMSI 파이프라인이 리얼리티를 극대화했지만, 여전히 합성된 마린스노우와 실제 바닷속에서 발생하는 역동적인 환경 사이에는 차이가 존재
- **향후 과제:** 단순히 입자를 배치하는 방식을 넘어, 유체 역학이나 빛의 전파 모델을 결합하여 더욱 물리적으로 완벽한 환경을 시뮬레이션해야 함

계산 비용과 실시간성

- **한계점:** MSADiff는 확산 모델을 기반으로 하므로 성능은 압도적이지만, 연산량이 매우 많아 자원이 제한된 실제 수중 로봇 플랫폼에 즉시 탑재하여 실시간으로 구동하기는 어려움
- **향후 과제:** 성능을 유지하면서도 모델의 크기를 줄이는 경량화 아키텍처나, 하드웨어 특성에 맞춰 속도를 높이는 하드웨어 가속 최적화 연구가 필요

발표를 마치겠습니다