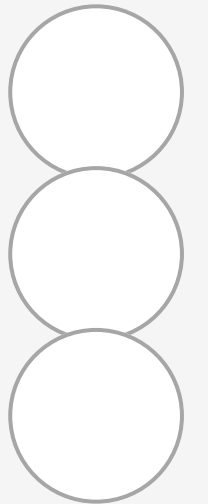


수중 드론 영상처리를 위한 선행연구 논문 리뷰



박정규

DeepShark Lab 학부연구생
Hongik University

MARINE SNOW SIMULATION AND ELIMINATION IN VIDEO

IEEE 2023

Contents

1

Introduction

문제 및 기존 방법의 한계

2

Methods

제안 방법

3

Training

모델 학습

4

Experiment

실험 결과

5

Conclusion

결과

INTRODUCTION

Introduction

What is Marine snow?

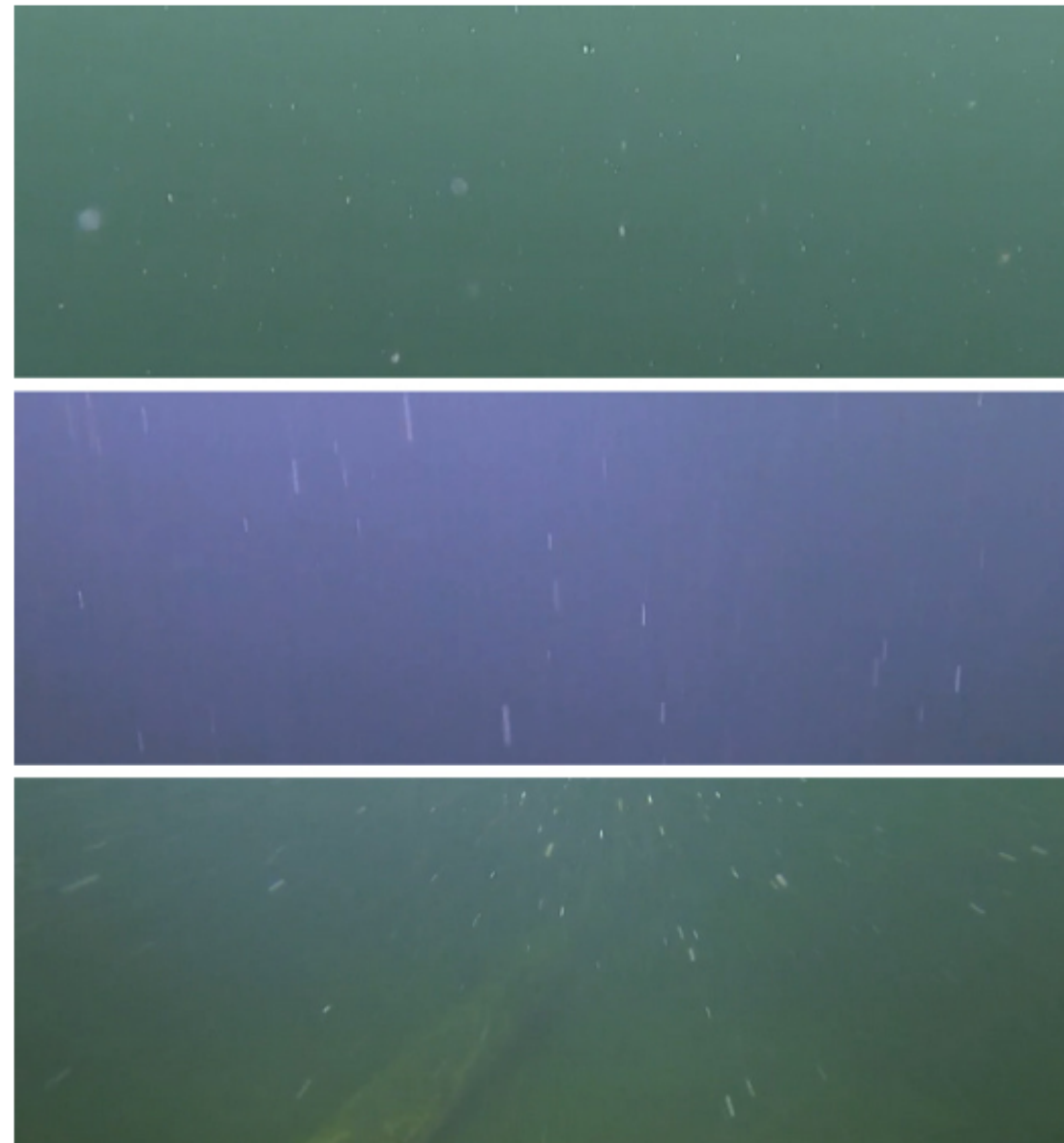


Fig. 2. Variation in marine snow appearance during a single mission. top: camera stationary near surface (aggregates appear circular), middle: diving (snow is vertically streaked), and bottom: moving forward near bottom (snow appears radially blurred).

Marine snow

- 유기물 잔해 와 무기 퇴적물의 작은 입자들이 서로 응집하여 형성된 입자로, 수중을 천천히 가라앉으며 하늘에서 눈이 내리는 것처럼 보이는 현상
- 흰색, 회색, 노란색, 갈색으로 나타날 수 있고 화학적 조성, 수질, 조명에 따라 달라진다
- 카메라의 이동에 따라 원, 타원, 선분등의 다양한 형태로 보여질 수 있다

Introduction

Problem Definition

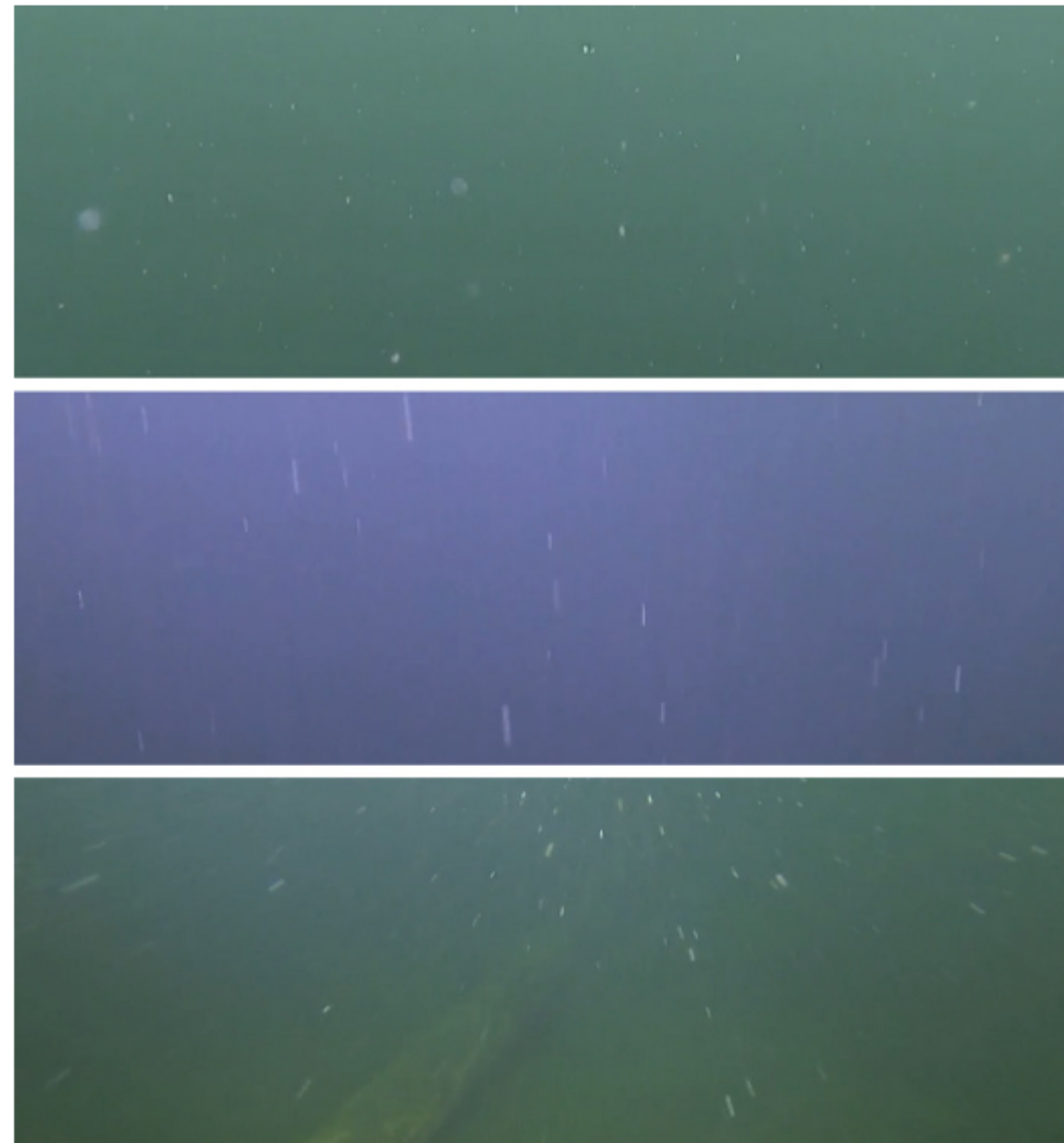


Fig. 2. Variation in marine snow appearance during a single mission. top: camera stationary near surface (aggregates appear circular), middle: diving (snow is vertically streaked), and bottom: moving forward near bottom (snow appears radially blurred).

Problem Definition

- 부유물은 빛을 산란시키는 특성을 가지며, 영상의 품질을 저하시킨다
- Computer Vision Algorithm의 성능이 감소한다
- Underwater Robot의 운용에 영향을 미친다

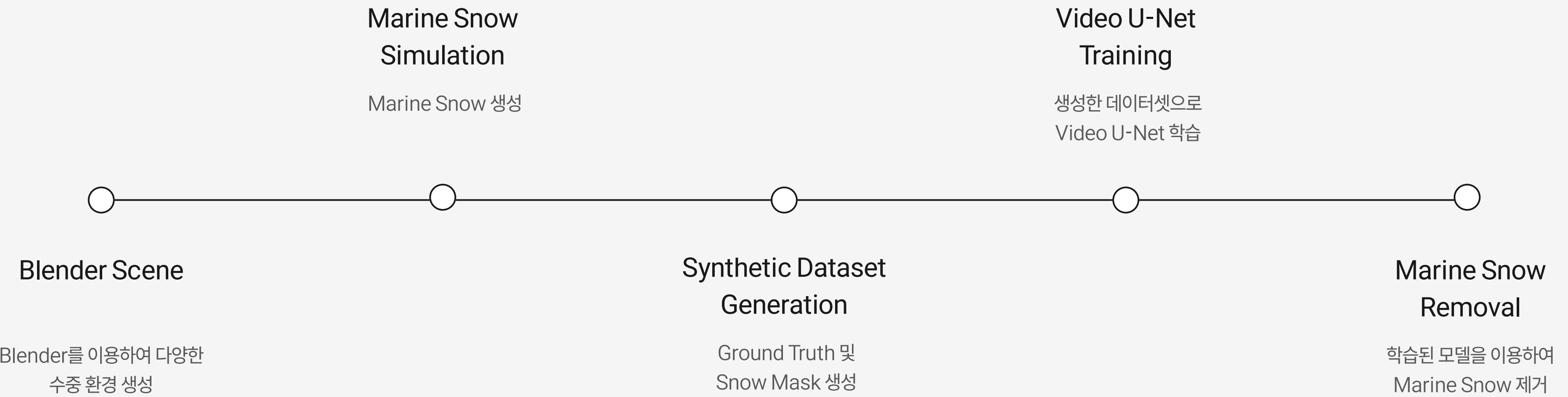
Limitations of Existing Methods

- 대부분의 부유물 제거 연구는 단일 이미지를 대상으로 한다
- 기존 Video 기반 방법은 대부분 고정된 카메라를 가정한다
- 딥러닝 기반 방법은 대량의 학습 데이터가 필요하지만, 실제 부유물 데이터의 수집과 정답(Ground Truth)생성이 매우 어렵다

METHODS

Methods

Overall Framework



Methods

Marine Snow Simulation in Blender

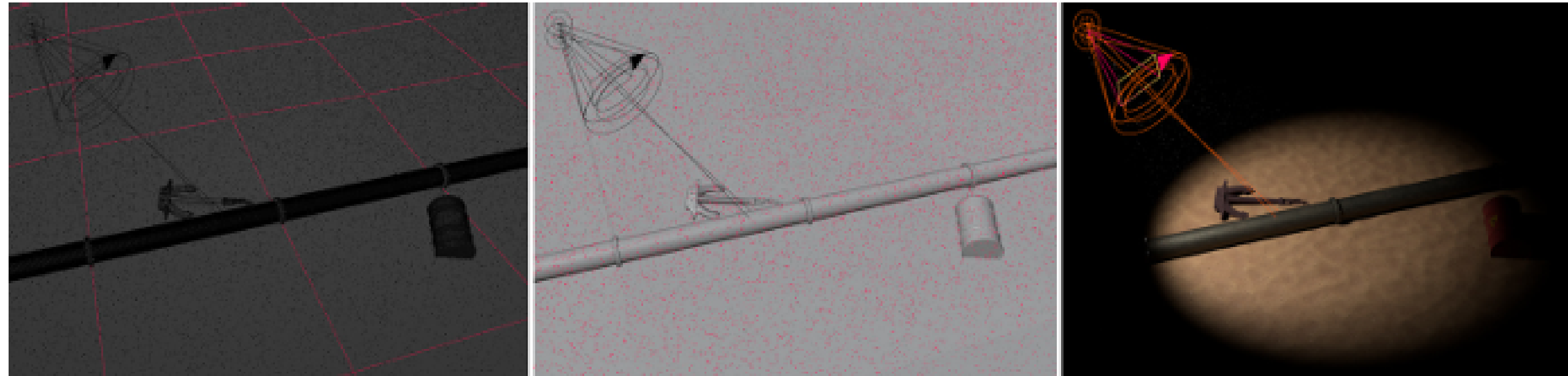


Fig. 3. Blender setup with artificial scene, camera, lighting, and marine snow: (left) mesh view with seafloor selected, (middle) solid view with snow system selected, and (right) material view with camera and lights selected.

Blender Scene

- 3D 환경을 모델링하고 렌더링할 수 있는 소프트웨어로, 논문에서는 실제와 유사한 수중 환경과 Marine Snow 영상을 생성하기 위한 시뮬레이션 도구로 사용

Methods

Marine Snow Simulation in Blender

Seafloor Generation

복잡한 해저 지형과 단순하거나 자유로운 지형 두 가지 방식으로 생성

Object Placement

해저면을 생성한뒤 직접 모델링한 수중 물체와 사전 제작된 3D 모델을 장면에 추가

Object Materials

물체를 장면에 추가한 후, 각 물체의 외형과 재질을 구체적으로 설정한다

Water Column

전체 객체를 덮는 Cube를 만들어 그 안의 색상, 밀도, 굴절률을 조절하여 실제 물처럼 보이게 하여 원하는 물속 환경 생성

Camera

수중 로봇의 촬영 환경을 재현하기 위해 카메라를 배치하고, 카메라의 이동 방향, 회전, 흔들림등을 적용하여 실제처럼 구성

Lighting Config

수면 위에 태양 광원을 배치하고, 카메라에도 조명을 붙혀 카메라와 조명이 함께 움직이는 환경을 재현

Methods

Marine Snow Generation



1. Geometry Nodes

기본적인 Marine Snow 생성 설정값으로, 갯수, 형태, 해상도등을 정의하는 Node 이다.



2. Randomization

각 부유물의 형태, 크기, 회전, 색상등을 서로 랜덤하게 생성한다.



3. Marine Snow Appearance

Shader를 이용하여 입자의 광학적 특성을 표현한다. Alpha, Emission 값을 조정하여 입자의 반투명 구조, 빛의 산란등을 표현한다.



4. Marine Snow Motion

중력과 부력을 설정하여 부유물의 물리적인 움직임 표현한다.
물의 흐름을 순류와 난류로 나누어 수중속 입자의 움직임 설정한다.

Marine Snow Generation

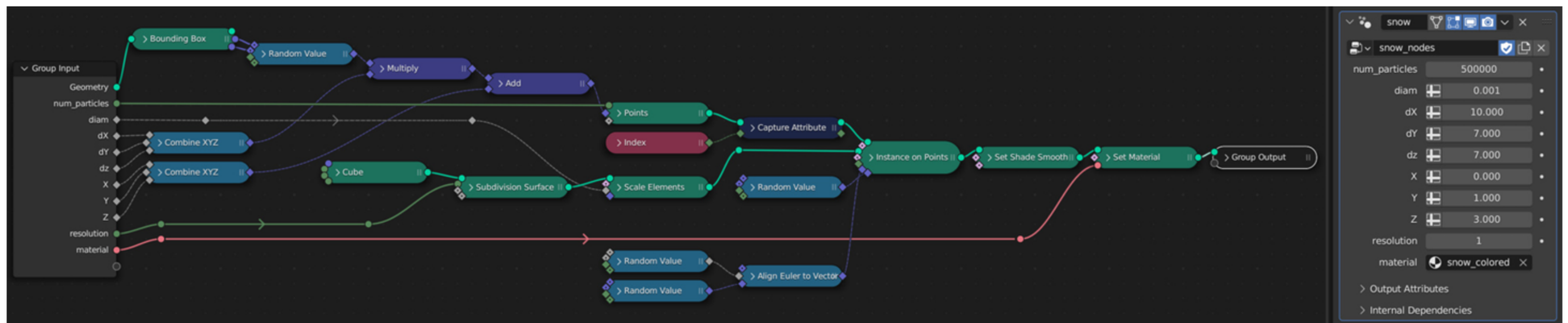


Fig. 4. Geometry node setup for marine snow system. Various nodes control aggregate count and randomized size, shape, and color.

Methods

Synthetic Dataset Generation

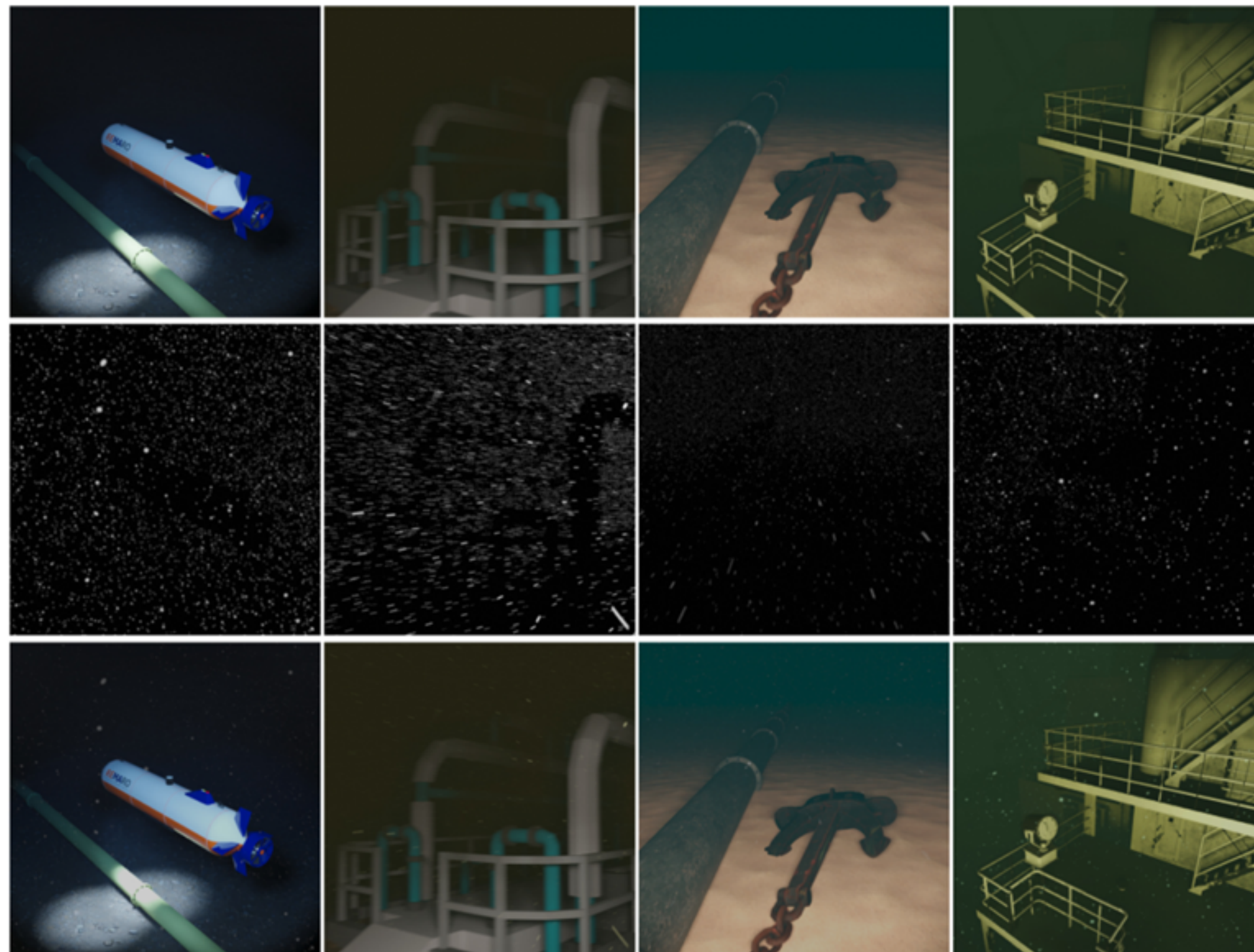
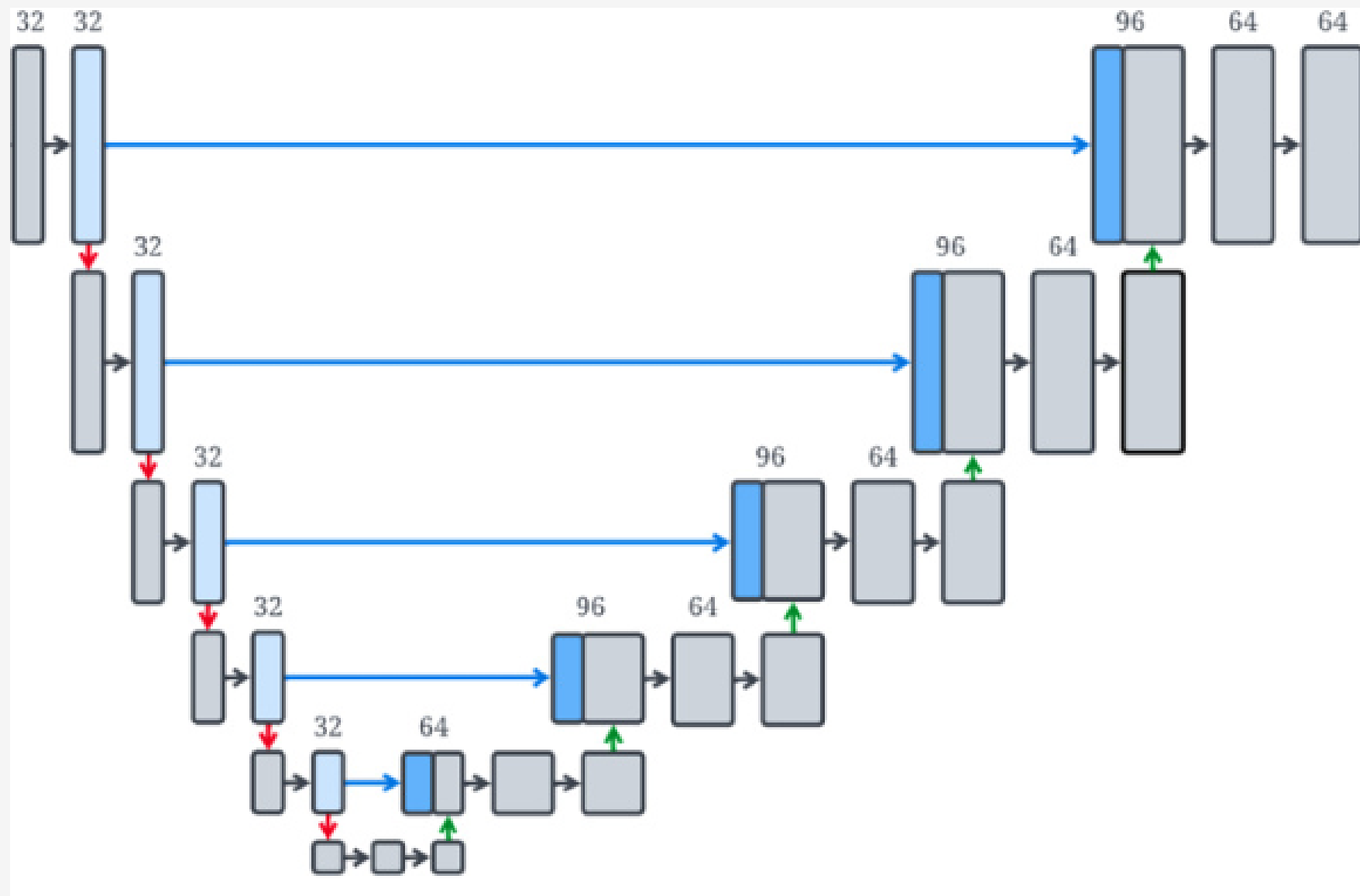


Fig. 5. Example frames from four scenes. (top) no snow, (middle) snow mask, and (bottom) snow added.

Video Generate

- 환경, 카메라 움직임, 광원, 수질등을 다양하게 구성하여 여러 Scene을 만들어 영상을 제작한다.
- 생성한 수중 환경과 부유물을 이용하여 총 3가지 영상을 제작한다.
- 부유물이 없는 일반 영상 (Ground Truth)
- 부유물 Mask 영상 (이번 학습에서는 사용하지 않음)
- 부유물이 추가된 영상 (input)

Video U-Net



U-Net

Encoder

RGB 영상으로부터 주변 픽셀의 공간을 추출하면서 영상을 점차 작은 Feature Map으로 압축

Bottleneck

Encoder에서 추출한 Feature를 가장 압축된 형태로 표현하는 부분

입력 영상의 중요한 특징만 유지하고, Marine Snow와 같은 불필요한 정보를 제거하도록 학습

Decoder

압축된 Feature를 이용하여 다시 원래 영상 크기로 복원

Skip Connection

Encoder의 Feature를 Decoder에 직접 전달하여 원본과 복원 이미지간 차이 즉, 잔차를 학습하여 복원 성능을 향상

Methods

Video U-Net

Limitation

기존 Image U-Net을 MSRB 데이터셋으로 학습한 후 실제 현장에서 촬영한 영상에 적용하였지만, 성능이 떨어지는 것을 확인하였다.

Cause Analysis

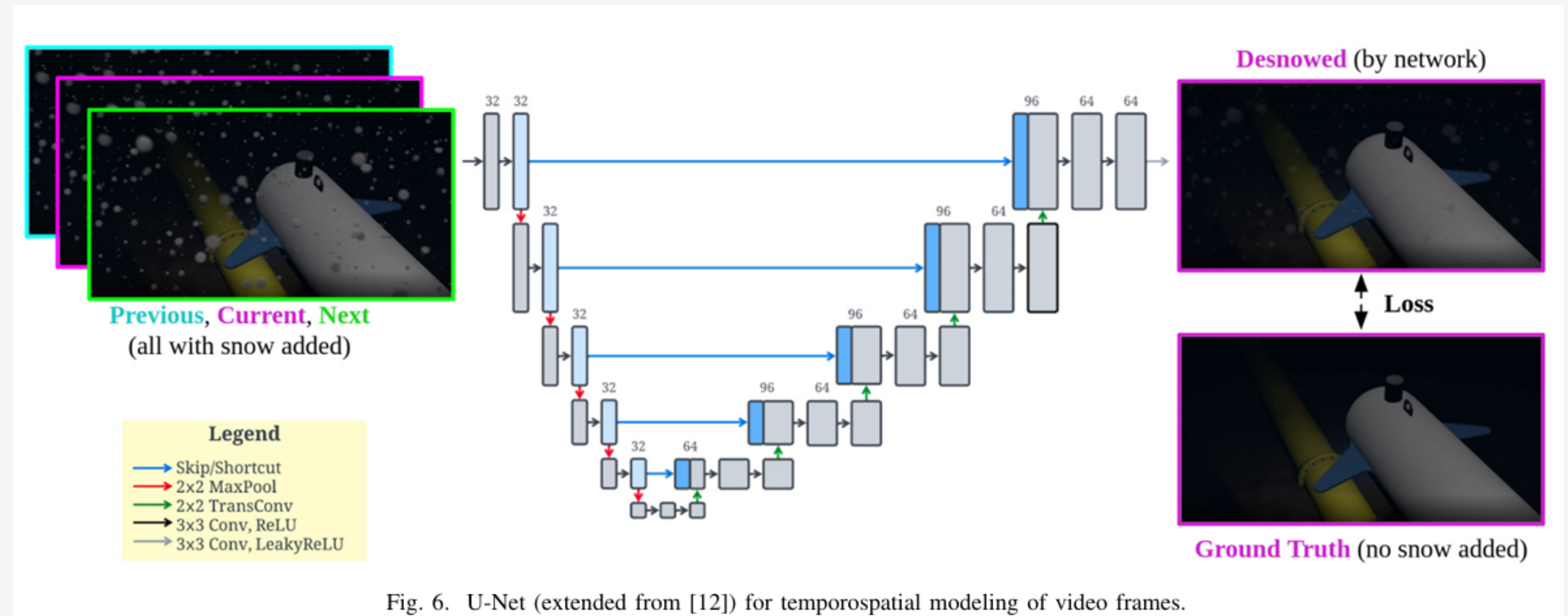
원인 분석을 위해 Blender로 제작한 데이터셋으로 동일한 네트워크를 재학습 시켰다.
기존 데이터셋은 실제 환경과 차이가 있어 성능이 저하되고, 기존 방식은 연속적 프레임 사이의 움직임 정보는 활용하지 못한다는 것을 확인하였다.

Proposed Solution

본 논문에서는 Image 기반의 U-Net 학습이 아닌, 연속된 세 프레임을 input으로 하는 Video 기반 U-Net을 구성하여 공간적 정보와 시간적 정보를 동시에 학습하도록 설계한다.

Methods

Video U-Net



TRAINING

Methods

Model Definition

Optimizer

1e-3의 Learning rate를 사용하는 Adam Optimizer 사용

Loss

픽셀 단위의 MSE LOSS 사용
100 Epoch 진행

Input

3개의 연속적인 프레임으로 구성된 4개의 Video 클립으로 구성
프레임을 수직 축을 기준으로 뒤집거나 재생 방향을 반전시키는 데이터 증강 과정 포함

Training

image U-Net Training

Image U-Net Classification by Dataset

MSRB Dataset

Flickr에서 수집한 2700장의 이미지를 small / Small + large Marine Snow 두 버전으로 5,400장으로 만들고 4,320장을 Training으로 사용

Proposed Synthetic Dataset

Blender로 제작한 실제 수중환경과 비슷한 이미지 데이터셋 7200장중 5,760장을 Training으로 사용

MSRB + Proposed Synthetic Dataset

두 데이터셋을 합쳐 10,080장을 Training으로 사용

- Marine Snow Removal Benchmarking Dataset (IEEE 2023)

Training

image U-Net Training

Video U-Net Classification by Dataset

Proposed Synthetic Dataset

Blender로 제작한 실제 수중환경과
비슷한 이미지 데이터셋 7200장중
5,760장을 Training으로 사용

MSRB + Proposed Synthetic Dataset

두 데이터셋을 합쳐 10,080장을
Training으로 사용

EXPERIMENT

Experiment

Evaluation Metrics

PSNR

복원된 영상이 Ground Truth와
얼마나 유사한지를 측정하는 영상
품질 평가 지표

부유물 제거 영상이 원본 영상과
얼마나 비슷한지 측정

값이 높을수록 오차가 적다는 의미

SSIM

두 영상의 구조적 유사도를 평가하
는 지표

단순히 픽셀 차이가 아니라, 밝기
대비 구조를 함께 비교하여 전체적
인 유사도를 평가

0 ~ 1사이의 값을 가지며 높을수
록 원본과 유사하다는 의미

Experiment

Result

Approach	Variant	Performance on ...					
		... MSRB Dataset		... Our Dataset		... Merged Dataset	
		PSNR (↑)	SSIM (↑)	PSNR (↑)	SSIM (↑)	PSNR (↑)	SSIM (↑)
Median Filter	3×3 kernel	25.81	0.810	37.93	0.966	32.74	0.899
	5×5 kernel	24.09	0.682	36.51	0.955	31.19	0.838
Image U-Net (from [12])	Trained Only on MSRB Dataset	36.44	0.970	39.09	0.967	37.95	0.968
	Trained Only on Our Dataset	25.28	0.840	44.92	0.991	36.50	0.926
	Trained on Both Datasets	35.86	0.967	43.65	0.991	40.31	0.981
Video U-Net	Trained Only on Our Dataset	25.49	0.836	45.78	0.991	37.08	0.925
	Trained on Both Datasets	35.64	0.966	43.87	0.988	40.34	0.978
Original Snowy Images	—	28.04	0.909	38.79	0.963	34.18	0.940

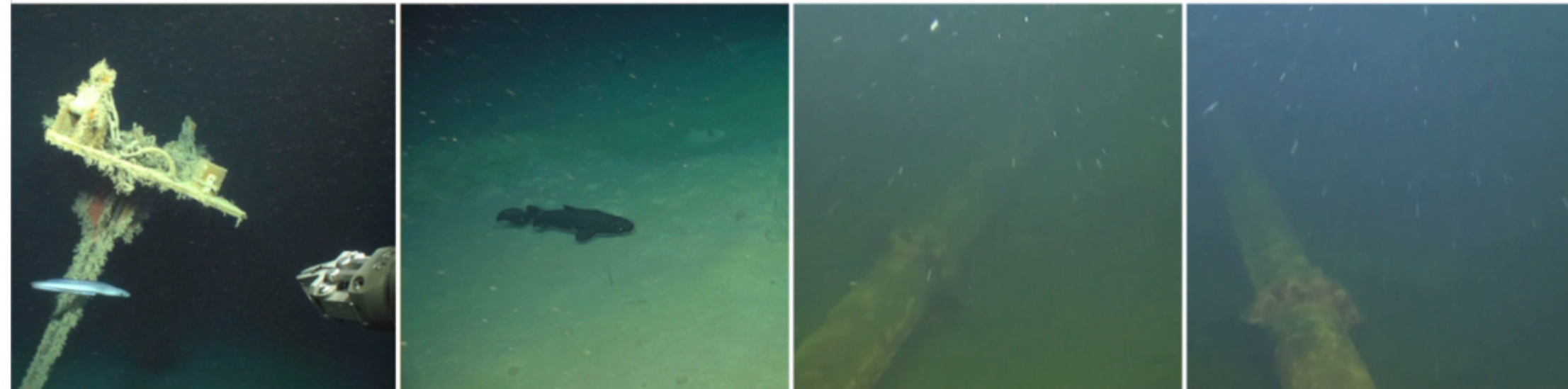
Result

- 논문에서 해당 결과를 학습 데이터와 테스트 데이터의 유사성이 가장 중요하다고 설명 한다.
- 실제 바다에서 촬영한 영상으로 테스트 했을 때, 해당 논문에서 제작한 데이터셋으로 학습한 모델이 더 좋은 결과를 보였다.

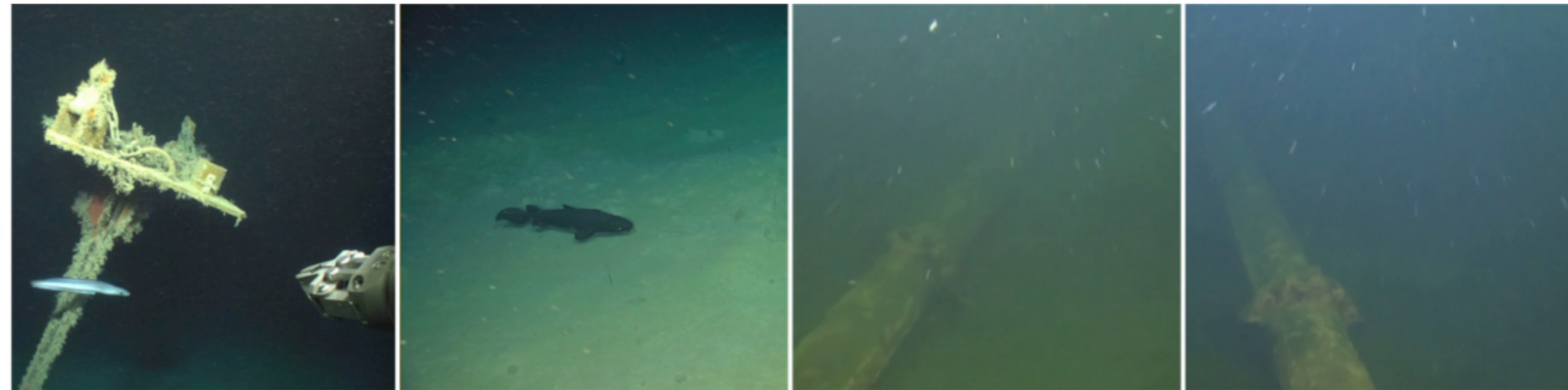
Experiment

Result

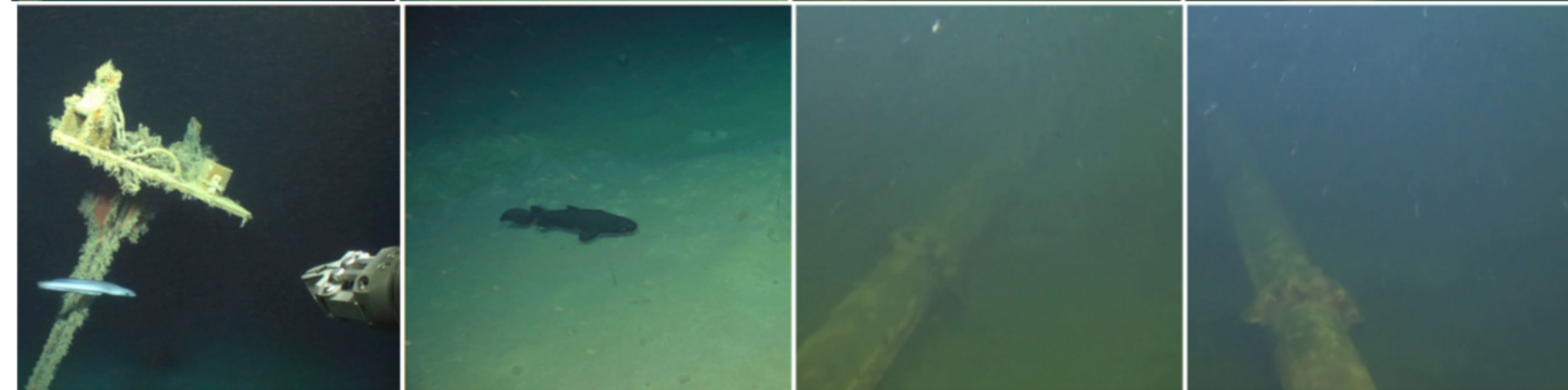
Original, Snowy Input



Output of
Image U-Net (from [12])
Trained Only on MSRB Dataset



Output of
Video U-Net
Trained on Both Datasets



CONCLUSION

Conclusion

Conclusion 1

학습 데이터와 테스트 데이터의 유사성이 가장 중요하다

Conclusion 2

해당 논문의 Dataset이 실제환경에 더 적합하다

Conclusion 3

Video U-Net은 시간적 정보를 활용할 수 있다

Final Conclusion

가장 유망한 개선 방향은 모델을 복잡하게 하는것이 아니라, 실제 환경과 유사한 데이터를 더 많이 생성하여 재학습하는 것이 가장 중요하며, Video U-Net은 이러한 데이터에서 시간 정보를 활용해 추가적인 성능 향상을 제공한다.

Q&A

A DEEP LEARNING APPROACH FOR MARINE SNOW SYNTHESIS AND REMOVAL.

Springer 2025

Contents

1

Introduction

문제 및 기존 방법의 한계

2

Methods

제안 방법

3

Training

모델 학습

4

Experiment

실험 결과

5

Conclusion

결과

INTRODUCTION

Introduction

What is Marine snow?

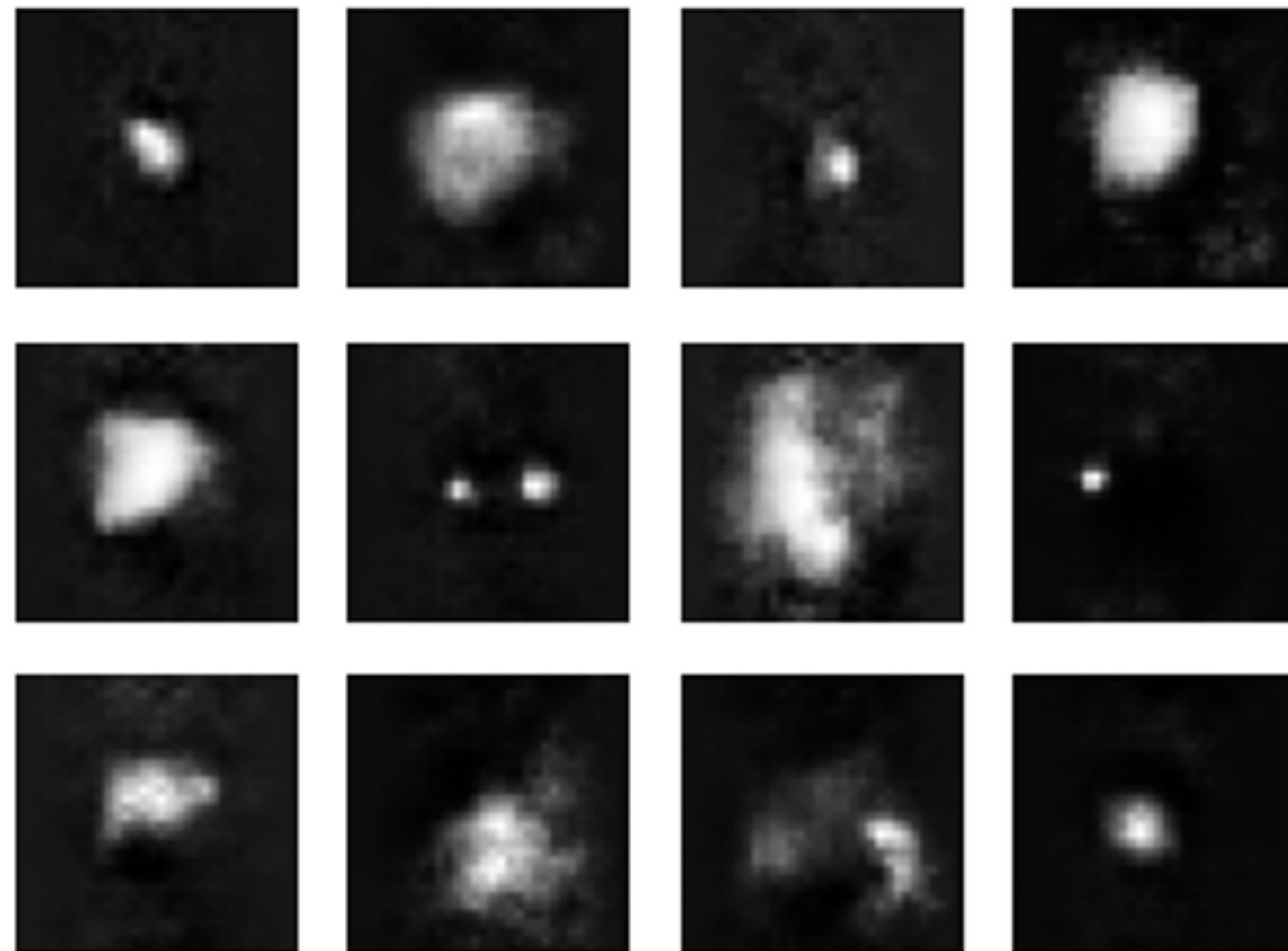


Figure 1: Synthetic marine snow samples.

Marine snow

- 수중에 떠다니는 유기물, 무기물, 침전물등의 부유 입자
- 카메라 조명에 의해 강한 빛을 반사하는 입자
- 영상에서 밝은 흰 점 또는 얼룩 형태로 나타남

Introduction

Problem Definition

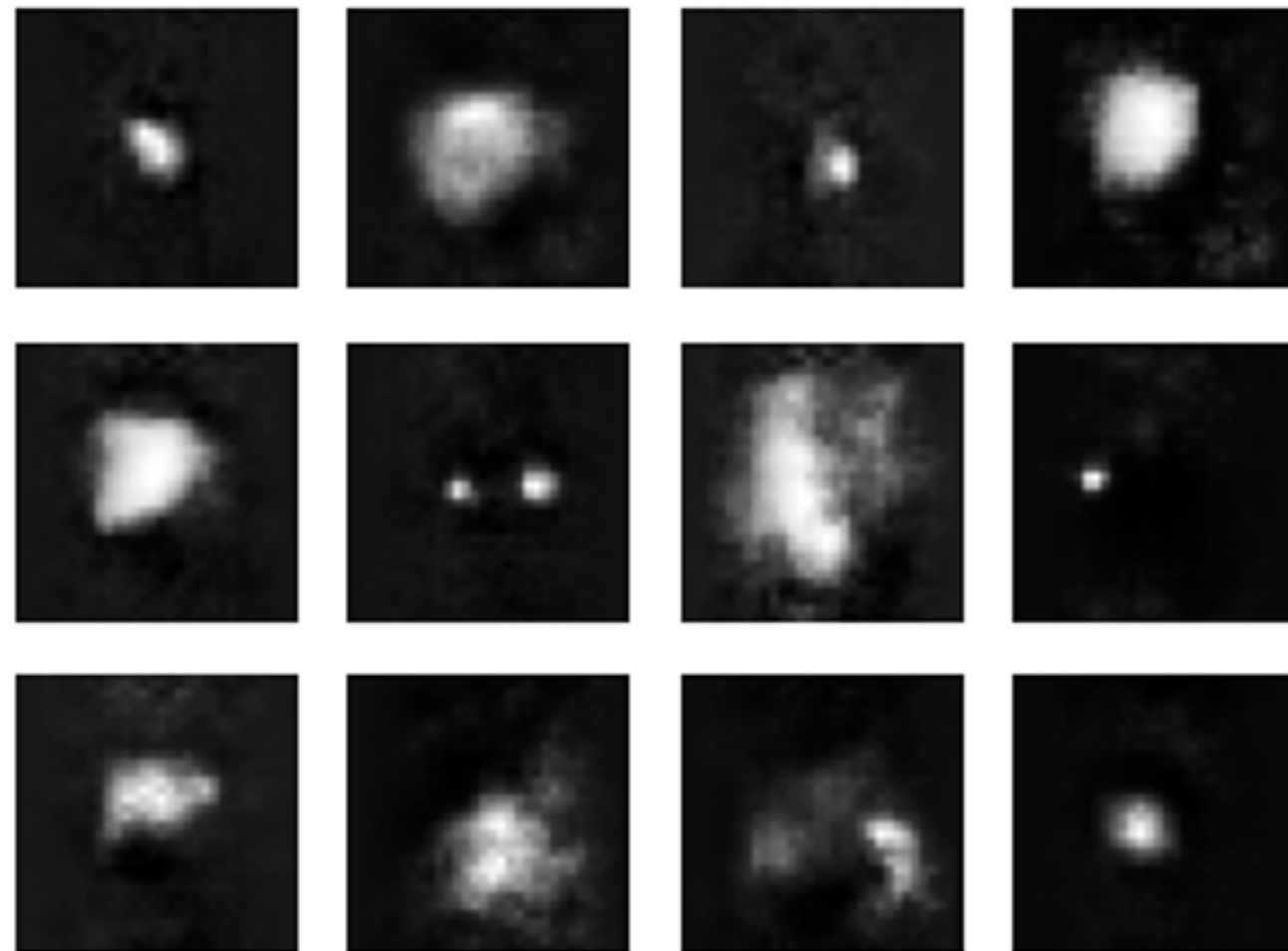


Figure 1: Synthetic marine snow samples.

Problem Definition

- 부유물은 빛을 산란시키는 특성을 가지며, 영상의 품질을 저하시킨다
- Computer Vision Algorithm의 성능이 감소한다
- 선체 청소 및 무인 설비 검사와 같은 수중 작업에서 중요한 문제로 작용한다

Limitations of Existing Methods

- 기존 수중 영상 복원 연구는 주로 색 왜곡, 낮은 대비의 문제를 해결하는데 초점을 맞추었다
- Marine Snow 제거를 다룬 연구는 제한적이며, 주로 Median Filter 기반 기법이 주로 사용되었지만, 큰 부유물 제거에는 한계가 있다
- 신경망 기반 방법도 제안되었지만, 다양한 Marine Snow 데이터의 부족으로 인해 성능에 한계가 있다

Introduction

Objective

GAN 구조 이용

GAN 기반 Marine Snow 샘플을
생성

Paired Dataset 구축

자연 수중 영상과 합성 Marine
Snow를 결합하여 Paired
Dataset 구축

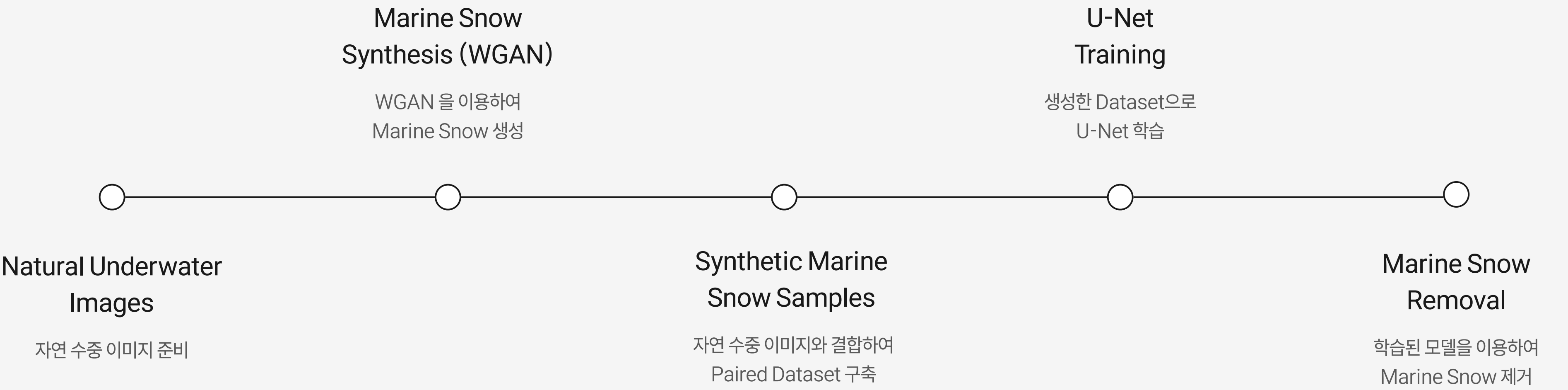
U-Net 학습

CNN 기반의 U-Net을 직접 구성
한 Dataset으로 학습시켜 효과적
으로 수중 이미지를 개선하는
Architecture 제안

METHODS

Methods

Overall Framework



Methods

GAN

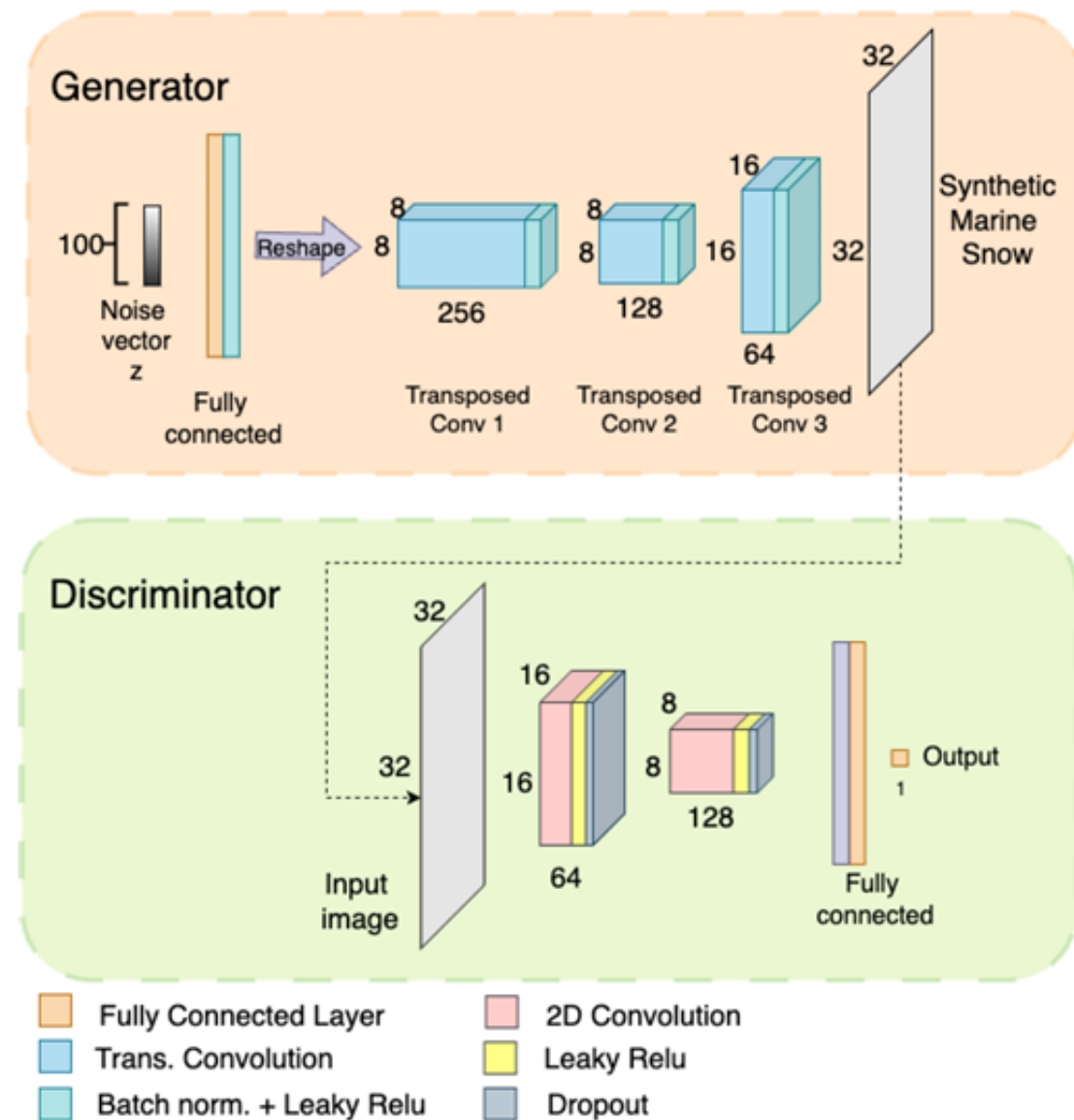


Figure 2: GAN architecture to synthesize marine snow artifacts.

GAN

- 새로운 데이터를 생성하기 위한 딥러닝 모델
- Generator와 Discriminator 두 개의 신경망이 서로 경쟁하면서 학습
- Generator는 Discriminator를 속이기 위해 실제 Marine Snow와 비슷한 이미지를 생성하고, Discriminator는 가능한 정확하게 구별하도록 학습
- 기존에는 Gaussian과 같은 단순 수학 모델로 Marine Snow를 만들었지만, 실제로는 크기, 모양, 밝기가 매우 다양하기 때문에 더 현실적인 DataSet을 만들기 위해 GAN계열 모델 사용

Methods

WGAN Loss

$$\mathcal{L}_d = \frac{1}{N} \sum_{i=1}^N (D(G(z_i)) - D(x_i))$$

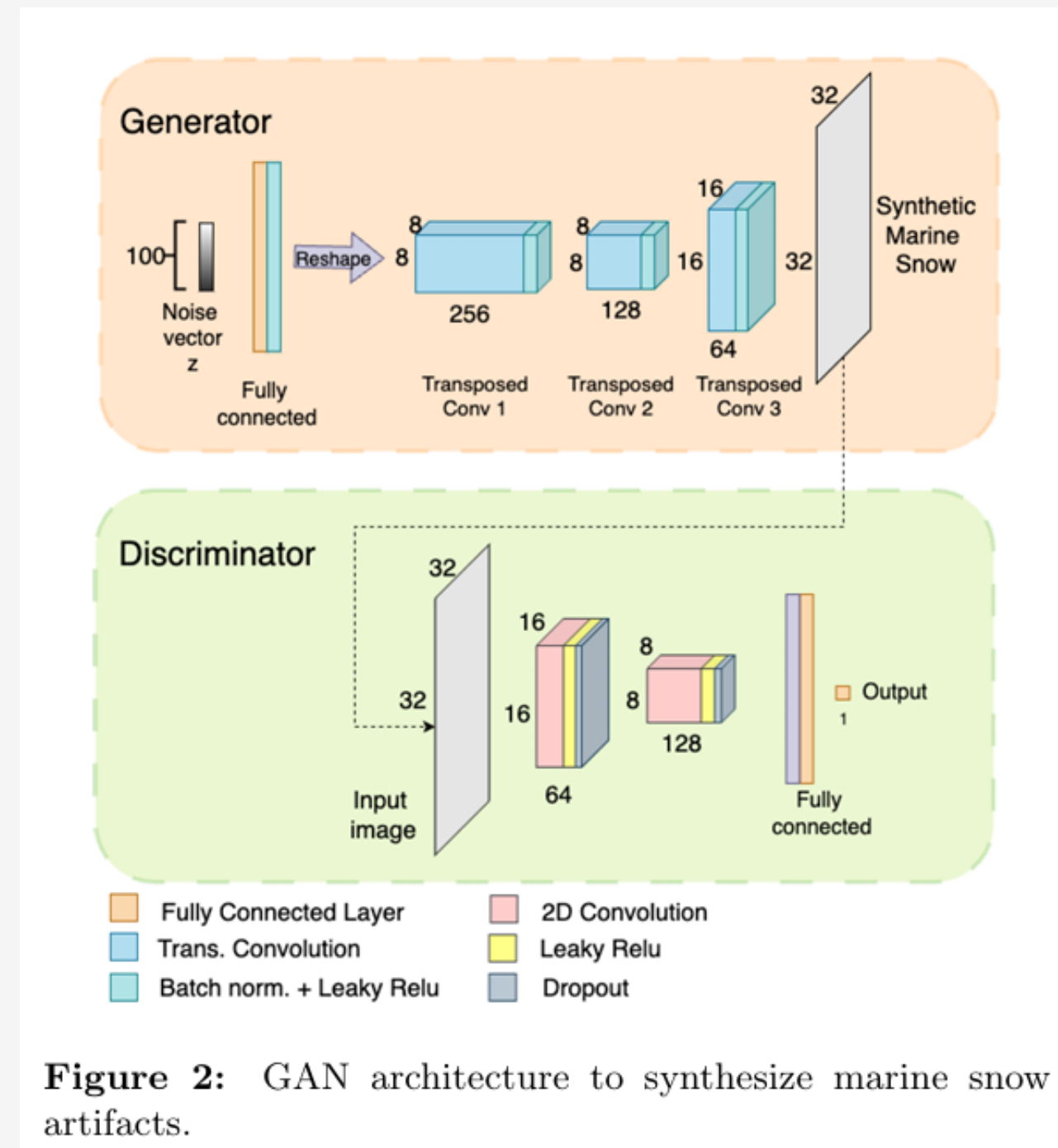
$$\mathcal{L}_g = -\frac{1}{N} \sum_{i=1}^N D(G(z_i))$$

Wasserstein GAN

- 해당 논문에서는 GAN의 단점인 Mode Collapse와 기울기 소실 문제를 완화하기 위해 WGAN을 사용한다.
- 기존 GAN은 생성 이미지가 실제 이미지일 확률을 출력하는 반면, WGAN은 생성 이미지와 실제 이미지 사이의 Wasserstein Distance를 근사하기 위한 Critic Score를 출력한다. 이를 통해 학습 안정성과 수렴성을 향상시킨다.
- Discriminator Loss는 랜덤노이즈 z 를 입력으로 한 Generator의 Critic Score와 실제 이미지의 Critic Score 사이의 차이이다.
- Generator Loss는 랜덤노이즈 z 를 입력했을때 생성된 이미지의 Critic score 이다.

Methods

WGAN Training



WGAN Training

- 2600개의 Marine Snow Sample을 추출 후 32 X 32 패치 크기로 조정하고, 픽셀값을 -1 에서 1 사이로 스케일링하여 사용

Generator

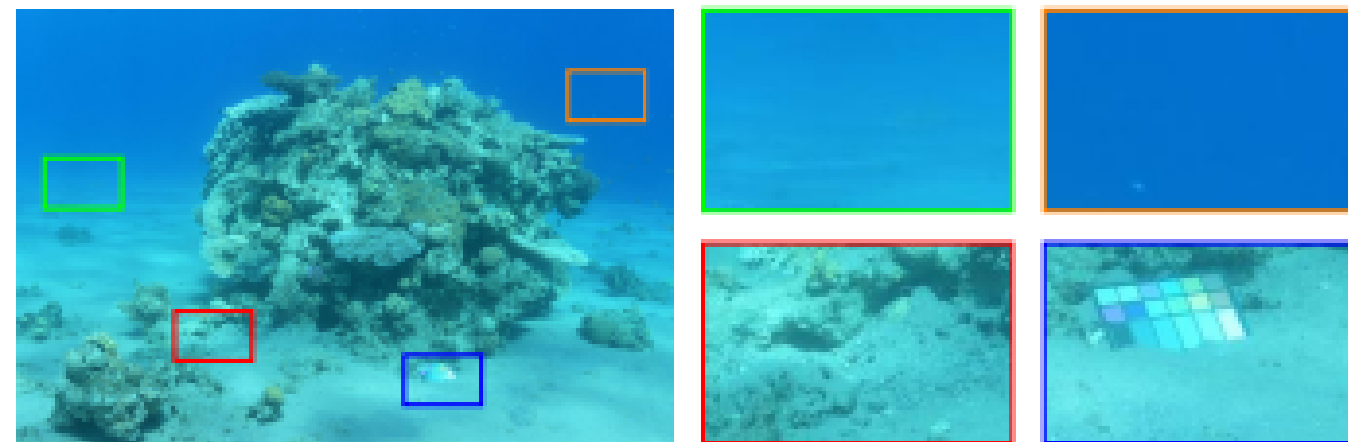
- 100차원의 랜덤 노이즈 벡터를 입력으로 받아 완전연결 계층을 통과해 8 X 8 X 256 형태로 Reshape
- 점점 채널수를 줄이고, 해상도를 높이며 32 X 32의 흑백 부유물 이미지를 생성
- 학습 안정성을 위해 각 전치 합성곱층 이후 배치 정규화와 Leaky ReLu 적용, 마지막 층에서 Tanh 함수를 사용해 픽셀값 스케일링

Discriminator

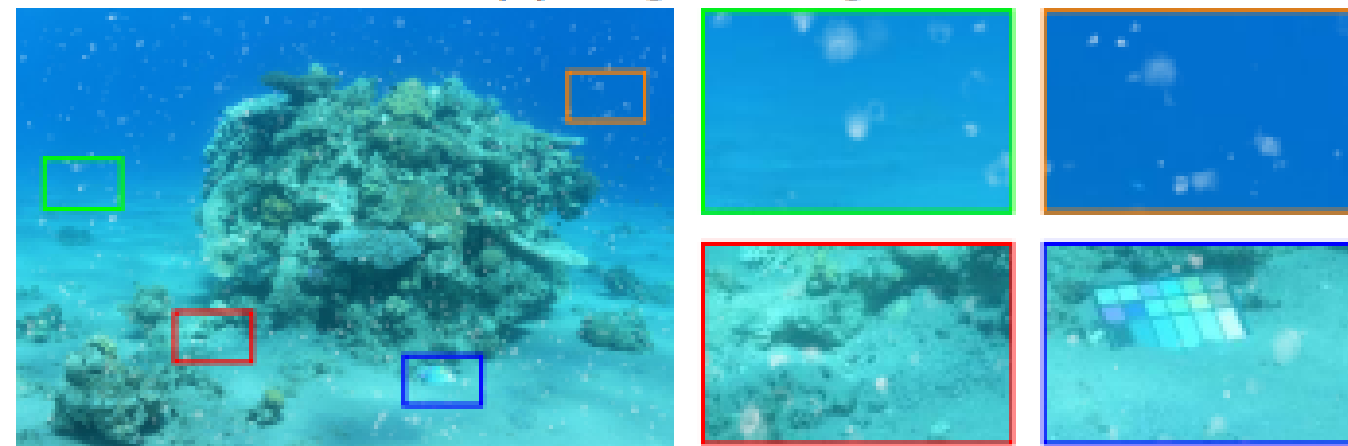
- Generator가 생성한 이미지를 입력받아 합성곱 층을 통과해 채널수를 늘려 완전연결 계층을 통과하고, 이미지가 실제일 확률을 출력
- WGAN에서는 확률이 아닌, 실제 데이터 분포에 얼마나 가까운지를 나타내는 실수 점수를 출력해 Generator에게 Critic 역할을 수행

Methods

Synthetic Marine Snow Samples



(a) Original Image



(b) Image with synthetic marine snow

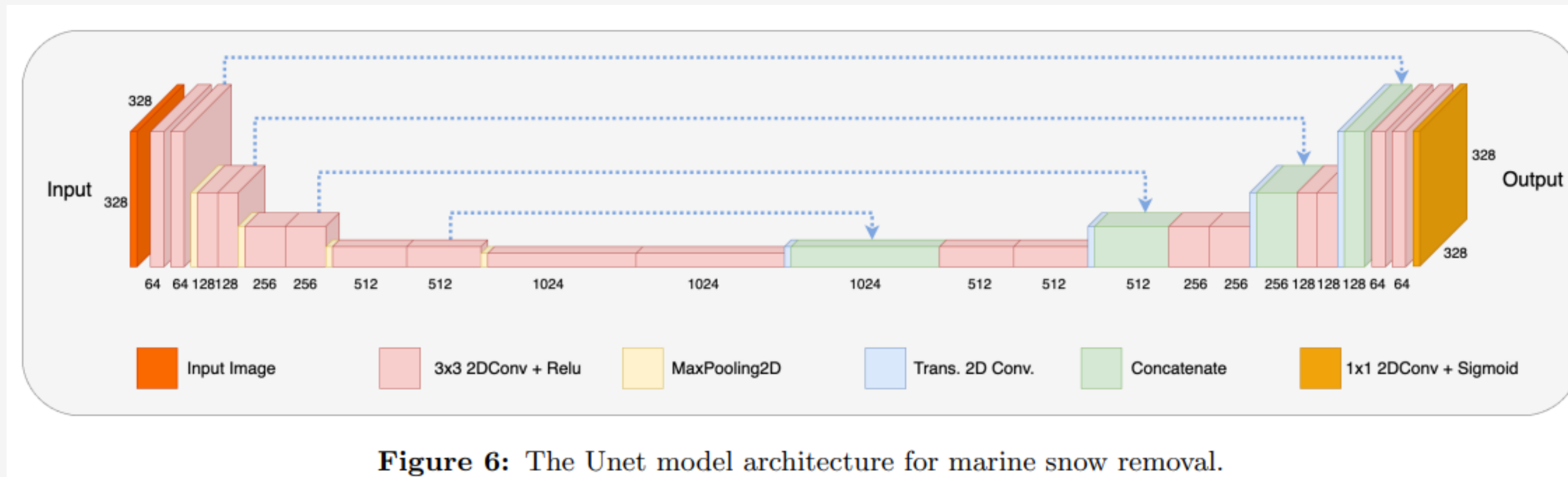
Figure 4: Example of adding synthetic marine snow to a natural underwater image.

Dataset Creation

- 학습된 WGAN을 이용하여 원본 이미지에 Marine Snow를 생성하고, 실제 환경과 비슷하게 만들기 위해 Impulse Noise, Gaussian Noise, Poisson Noise를 추가하여 U-Net 학습을 위한 Dataset을 구성한다.
- Impulse Noise : 단일 픽셀 크기의 Marine Snow 생성
- Gaussian Noise : 랜덤한 밝기변화를 모사
- Poisson Noise : 저조도 환경에서 발생하는 노이즈 모사

Methods

U-Net



U-Net Definition

- Skip connection을 포함한 인코더-디코더 구조를 한 형태
- 인코더는 다중 합성곱층, max-pooling, 공간차원의 다운샘플링을 통해 고수준 특징을 추출
- 디코더는 전치 합성곱을 사용하여 Feature Map을 업샘플링하고, 세부 정보를 포함한 이미지를 재구성

Methods

U-Net Loss

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$\mathcal{L}_{\text{p}} = \frac{1}{N} \sum_{i=1}^N (\text{VGG}(y)_i - \text{VGG}(\hat{y})_i)^2$$

$$\mathcal{L}_{\text{U-Net}} = \mathcal{L}_{\text{MSE}} + \gamma \mathcal{L}_{\text{p}}$$

U-Net Loss

- Ground Truth의 i 번째 픽셀과 생성된 image의 i 번째 픽셀간의 차이 평균 사용
- 전체 평균은 비슷해도, 경계나 질감이 흐릿한 이미지가 나올 수 있기 때문에 사전 학습된 VGG 19를 사용한 Loss를 추가함으로써 형태, 구조, 패턴등의 구체적인 특징까지 잡도록 유도
- Ground Truth 이미지와 생성된 이미지의 Feature Map의 차이의 평균을 Loss로 사용
- 해당 논문에서 $\gamma = 1$ 로 설정하여 MSE Loss 와 VGG Loss의 중요도를 동일하게 반영

TRAINING

Methods

Model Definition

Optimizer

1e-3의 Learning rate를 사용하는 Adam Optimizer 사용

Loss

픽셀 단위의 MSE LOSS 와
Feature Map VGG19 Loss 사용
20 Epoch 학습

Dataset

384 X 384 pixels RGB 이미지로 구성
각 이미지를 수평으로 뒤집는
data augmentation 진행

Train : 12,869

Valid : 3,217

Test : 2,760

EXPERIMENT

Experiment

Evaluation Metrics

PSNR

복원된 영상이 Ground Truth와
얼마나 유사한지를 측정하는 영상
품질 평가 지표

부유물 제거 영상이 원본 영상과
얼마나 비슷한지 측정

값이 높을수록 오차가 적다는 의미

SSIM

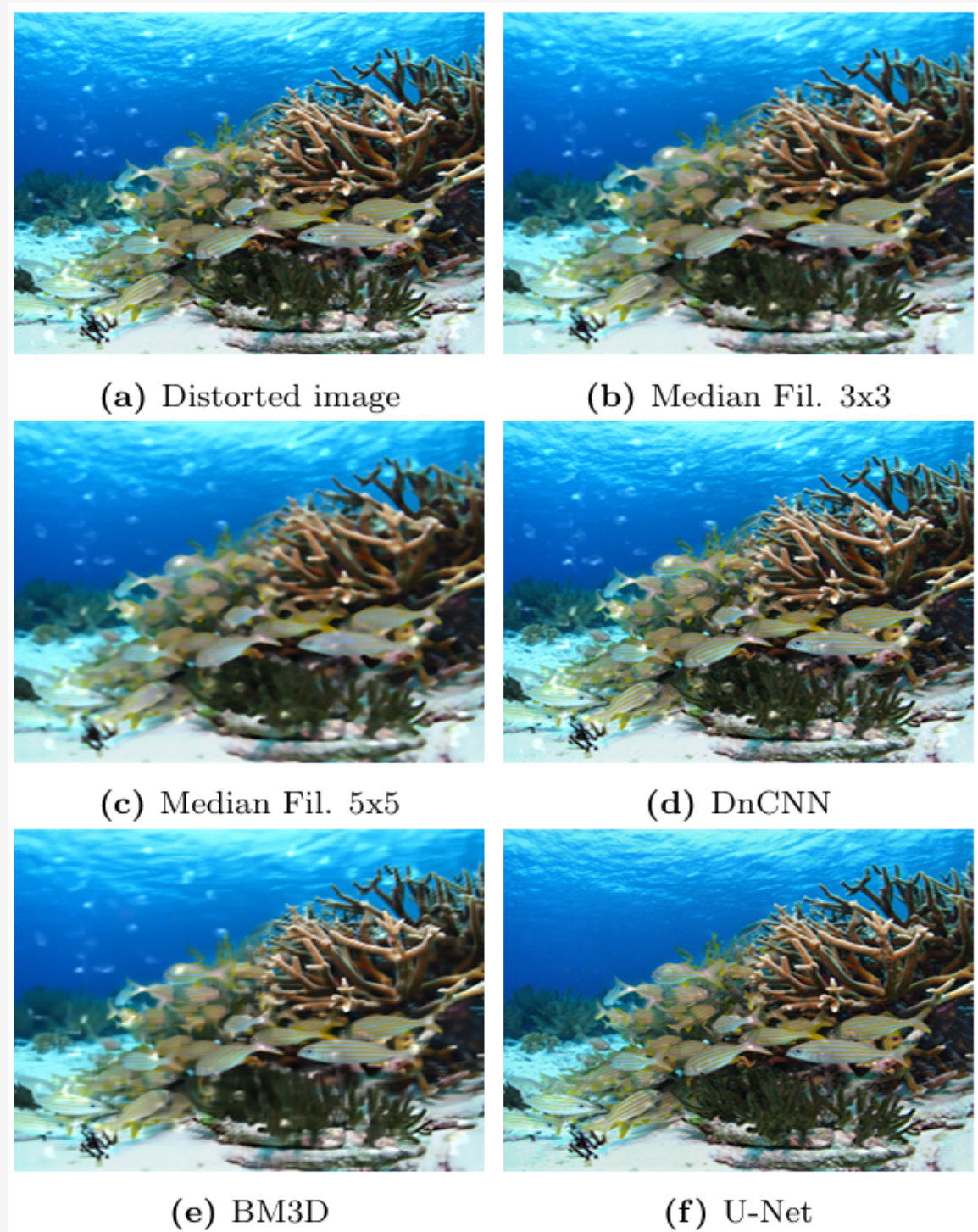
두 영상의 구조적 유사도를 평가하
는 지표

단순히 픽셀 차이가 아니라, 밝기
대비 구조를 함께 비교하여 전체적
인 유사도를 평가

0 ~ 1사이의 값을 가지며 높을수
록 원본과 유사하다는 의미

Experiment

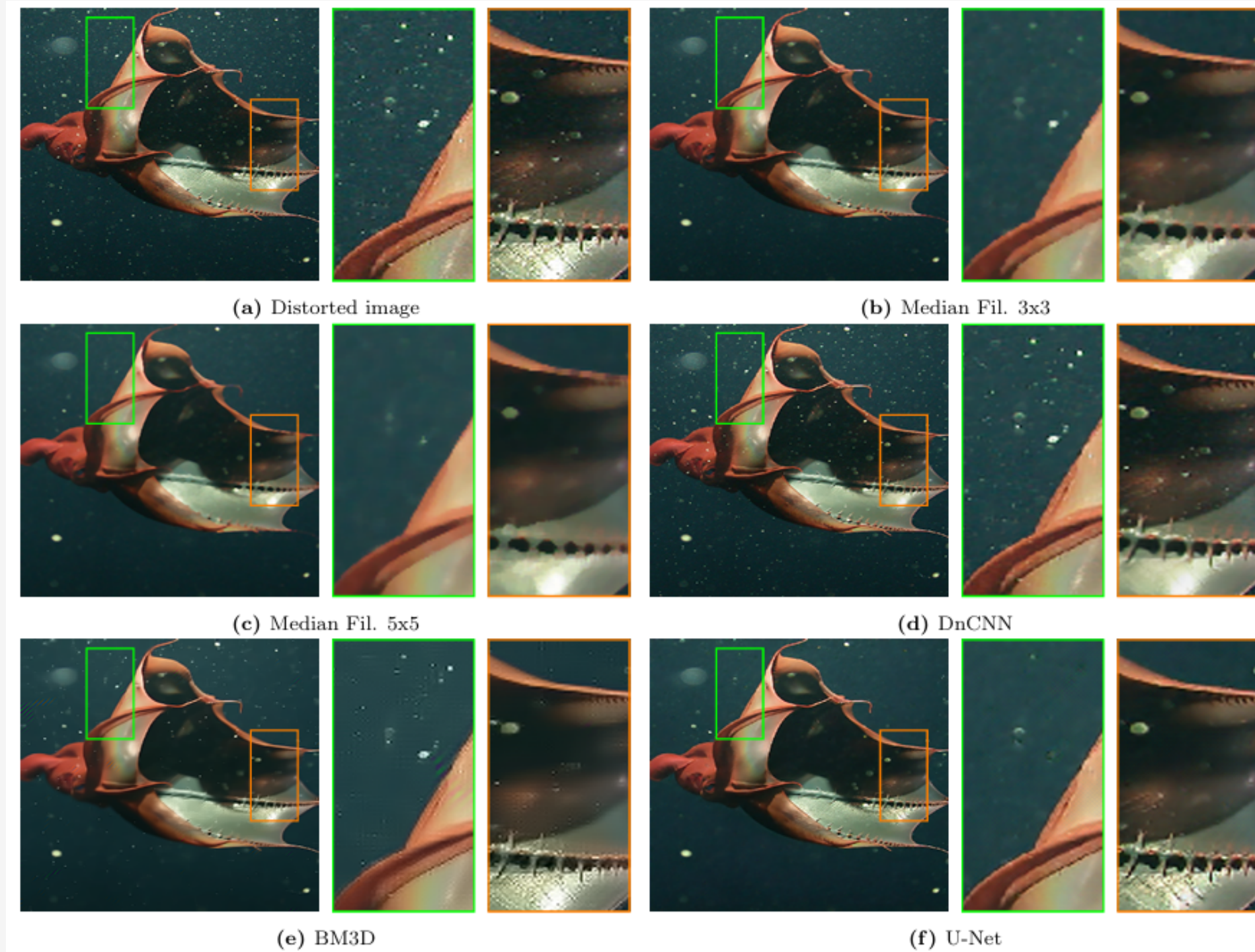
Synthetic Image



	MSE	PSNR [dB]	SSIM
Test images	59×10^{-4}	22.26	0.71
Med. Filt. 3×3	47×10^{-4}	23.90	0.83
Med. Filt. 5×5	56×10^{-4}	23.22	0.77
BM3D	46×10^{-4}	23.95	0.88
DnCNN	42×10^{-4}	24.50	0.91
U-Net	13×10^{-4}	29.18	0.91

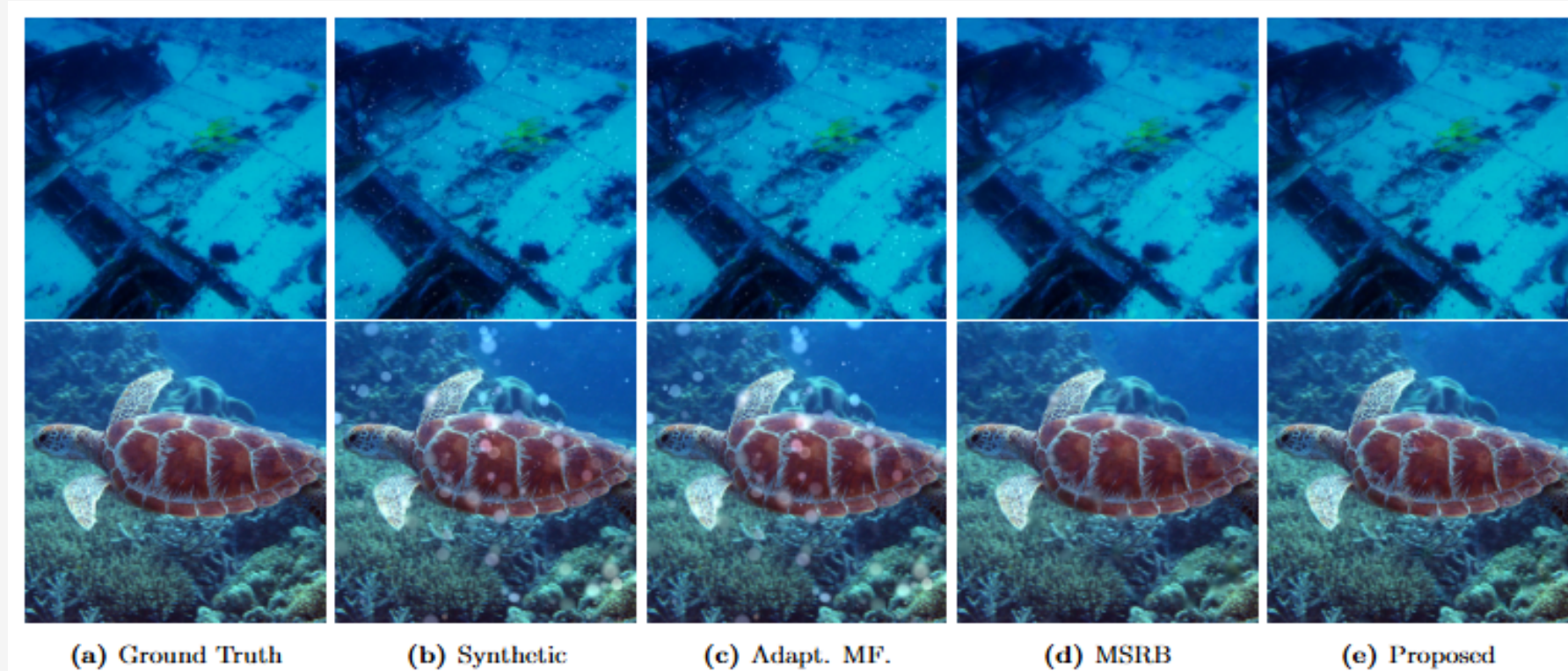
Experiment

Natural Image



Experiment

MSRB Image



	Task 1		Task 2	
	PSNR	SSIM	PSNR	SSIM
Test images	32.20	0.94	23.83	0.88
Med.Filt. 3×3	28.55	0.85	22.81	0.77
Med.Filt. 5×5	25.98	0.71	21.93	0.64
Adapt.Med. 3×3	29.88	0.91	23.35	0.84
Adapt.Med. 5×5	28.08	0.86	22.83	0.79
MSRB model	36.82	0.98	30.95	0.93
Proposed	31.11	0.94	27.27	0.90

Additional Experiment

- MSRB Image로 본 논문의 U-Net을 테스트 하였고, MSBR Dataset을 학습하지 않아도 MSRB 벤치마킹 모델과 비슷한 성능을 보였다.
- 이 실험을 통해 본 논문의 GAN 모델은 다양하고 사실적인 Marine Snow를 생성할 수 있고, U-Net도 효과적으로 식별하고 제거할 수 있는 능력을 입증하였다.

Experiment

Additional Experiment



(a) Input



(b) Ancutti [4]



(c) Ancutti+U-Net

Figure 9: Effect of removing marine snow on image enhancement.

Ancutti Algorithm

- Fusion-based Underwater image Enhancement 알고리즘
 - 하나의 영상을 여러가지 방식으로 보정한 후 좋은 부분만 합쳐 최종 영상 제작
 - 색 보정, 명암 대비 향상, 선명도 향상 목적의 알고리즘
- Enhancing underwater images and videos by fusion (IEEE 2012)

Experiment

Additional Experiment Evaluation Metrics

UIQM

$$\text{UIQM} = c_1 \times \text{UICM} + c_2 \times \text{UISM} + c_3 \times \text{UIConM}$$

색상, 선명도, 명암대비를 각각 계산 후 가중합으로 계산

$$C_1 = 0.0282$$

$$C_2 = 0.2953$$

$$C_3 = 3.5753$$

UCIQE

$$\text{UCIQE} = c_1 \times \sigma_c + c_2 \times \text{conl} + c_3 \times \mu_s$$

색 다양성, 명암대비, 평균 채도를 각각 계산 후 가중합으로 계산

$$C_1 = 0.4680$$

$$C_2 = 0.2745$$

$$C_3 = 0.2576$$

- An Underwater Color Image Quality Evaluation Metric (IEEE 2015)
- Human-Visual-System-Inspired Underwater Image Quality Measures (IEEE 2016)

Experiment

Additional Experiment

Table 2: Underwater image enhancement evaluation.

	UIQM [54]↑	UCIQE [55]↑
Original	2.0677	0.5023
Ancutti [4]	3.3875	0.5472
Ancutti [4]+Unet	4.3477	0.5487

Additional Experiment

- Ancutti를 이용하여 색을 보정하고, 본 논문에서 제시한 U-Net을 이용하여 부유물을 제거하였다.
- 부유물을 먼저 제거하면 기존 영상 향상 알고리즘의 성능도 개선되는지를 확인하기 위해 추가 실험 수행, 그 결과 U-Net으로 Marine Snow를 제거한 후 Ancutti를 적용했을 때 성능 향상을 실험적으로 확인
- 본 논문의 알고리즘이 기존 다른 알고리즘에 도움이 된다는 것을 나타낸다.

CONCLUSION

Conclusion

Conclusion 1

MSE Loss와 VGG19 Loss를 이용하여 학습한 U-Net이 합성 부유물 이미지를 높은 정확도로 제거할 수 있다.

Conclusion 2

MSRB Dataset 실험을 통해 본 논문의 GAN 모델이 사실적인 합성 부유물을 생성하는데 있어 효과적이다.

Final Conclusion

향후 연구에서 더 많은 실제 부유물 샘플로 WGAN 모델을 보강하여 더욱 사실적인 데이터셋을 만들 수 있고, 향상된 데이터셋을 통해 U-Net을 재학습시켜 부유물 제거의 성능과 적응성을 향상시킬 수 있다.

Q&A

수중 드론 영상처리를 위한 선행연구 논문 리뷰

발표를 마칩니다.

박정규

DeepShark Lab 학부연구생
Hongik University

