



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

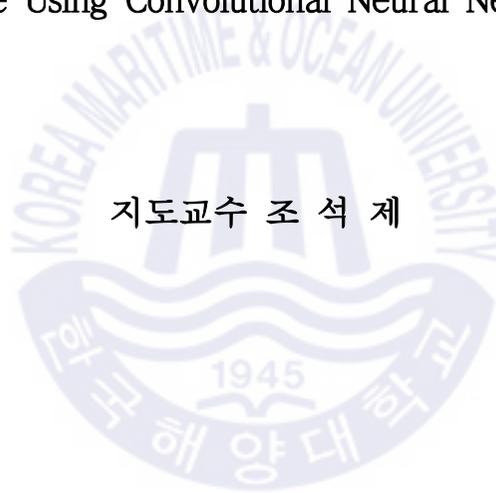
[Disclaimer](#)

공학석사 학위논문

Convolutional Neural Network를 이용한 수중 영상의 가시성 개선 및 색상 보정

Visibility Enhancement and Color Correction of Underwater
Image Using Convolutional Neural Network

지도교수 조 석 제



2020년 8월

한국해양대학교 대학원

제어계측공학과

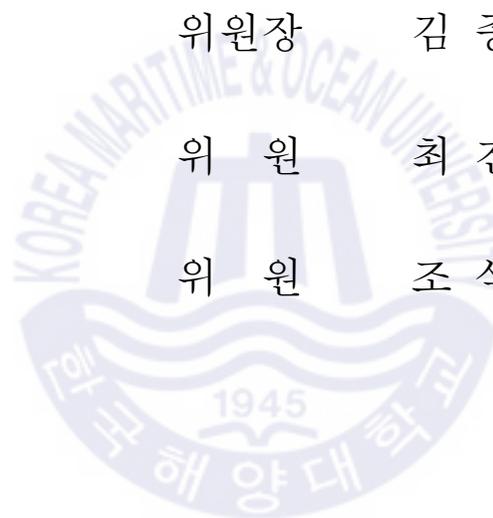
김도균

본 논문을 김도균의 공학석사 학위논문으로 인준함

위원장 김 종 화 (인)

위 원 최 진 규 (인)

위 원 조 석 제 (인)



2020 년 7 월 31 일

한국해양대학교 대학원

목 차

List of Figures	ii
Abstract	iii
제 1 장 서 론	1
제 2 장 Auto-encoder를 기반으로 한 수중 영상 개선	
2.1 수중 영상의 특징	4
2.2 AE를 기반으로 한 수중 영상 개선 방법	5
제 3 장 수중 영상 개선을 위한 Underwater Auto-Encoder	
3.1 수중 영상 형성 모델을 통한 학습 데이터 셋 합성	11
3.2 제안하는 UAE의 구조	12
3.3 UAE의 학습 방법	17
제 4 장 실험 및 고찰	18
제 5 장 결 론	30
참고문헌	31

List of Figures

Fig. 2.1	Image acquisition processing in underwater.....	4
Fig. 2.2	Light penetration in ocean water according to water depth.....	5
Fig. 2.3	Sample images taken in various environment	6
Fig. 2.4	Convolution with stride 2	8
Fig. 2.5	Transposed-convolution using nearest neighborhood interpolation...	8
Fig. 2.6	Comparison of actual underwater images and underwater images generated by cycleGAN and Jerlov scattering coefficient.....	10
Fig. 3.1	Training dataset synthesized by underwater image formation model equation, NYU dataset, and Jerlov light coefficients.....	13
Fig. 3.2	Training dataset synthesized by underwater image formation model equation, NYU dataset, and Jerlov light coefficients.....	14
Fig. 3.3	Training dataset synthesized by underwater image formation model equation, NYU dataset, and Jerlov light coefficients.....	15
Fig. 3.4	Architecture of proposed UAE with skip-connection.....	16
Fig. 4.1	Test dataset of type I and the result images of each methods...	20
Fig. 4.2	Test dataset of type IA and the result images of each methods..	21
Fig. 4.3	Test dataset of type IB and the result images of each methods..	22
Fig. 4.4	Test dataset of type II and the result images of each methods...	23
Fig. 4.5	Test dataset of type III and the result images of each methods..	24
Fig. 4.6	Result image of each methods about actual underwater image.....	25
Fig. 4.7	Result image of each methods about actual underwater image.....	26
Fig. 4.8	Result image of each methods about actual underwater image.....	27
Fig. 4.9	Result image of each methods about actual underwater image.....	28
Fig. 4.10	Result image of each methods about actual underwater image...	29

Visibility Enhancement and Color Correction of Underwater Image using Convolutional Neural Network

Do Gyun Kim

Department of Control and Instrumentation Engineering
Graduate School of Korea Maritime and Ocean University

Abstract

The underwater images obtained by optical sensor such as camera have low visibility and color distortion by underwater environments. In order to enhance visibility and color of underwater images, AE(Auto-Encoder) model based CNN(Convolutional Neural Network) is mainly used recently. In this paper, we propose the AE model which is effective for visibility enhancement and color correction of underwater images. The proposed UAE(Underwater Auto-Encoder) enhances underwater images by connecting skip-connection between encoder and decoder. For training UAE, we use underwater dataset consisted of no-distorted images and distorted underwater images generated by underwater image formation modeling equation. In order to verify the performance of UAE, we compare with traditional methods for underwater image enhancement. As a result of comparison using test dataset, UAE quantitatively outperforms than traditional methods on full reference metrics consisted of PSNR, SSIM, and color difference. Also UAE is qualitatively effective on actual underwater images.

제 1 장 서 론

최근 해양 자원에 대한 관심이 높아짐에 따라 수중 환경을 연구하고 탐사하기 위한 수중 로봇의 개발이 한창이다. 수중 로봇은 수중 환경을 확인하기 위한 다양한 장비가 부착된다. 그중 카메라와 같은 광학 장비는 로봇의 주변 환경을 직관적으로 확인하기 위한 기본적인 장비이다. 그러나 수중 영상은 물속의 여러 부유물에 의해 혼탁하고 물 분자에 의한 광 산란, 흡수, 그리고 배경광의 영향으로 가시성이 저하되고 색상이 왜곡된다[1-2]. 따라서 수중 영상의 가시성 개선과 색상을 보정하기 위해 기본적인 영상처리 방법을 활용한 방법, 영상 형성 모델링 기반 방법, 그리고 CNN(Convolutional Neural Network) 등이 연구되고 있다.

수중 영상에 사용되는 기본적인 영상처리 방법은 영상의 색상 보정을 위해 Gamma Correction[3], 영상의 대비를 향상시키기 위해 HE(Histogram Equalization)[4] 등을 응용하여 수중 영상을 개선한다. HE는 화소값의 분포를 균일하게 하여 대비가 높은 영상을 만들지만 본래 이미지가 가진 대비를 잃어 자연스럽지 못한 영상을 만든다. HE의 단점을 보완한 CLAHE(Contrast-Limited Adaptive Histogram Equalization)는 영상을 여러 구역으로 나누고 각 구역의 대비를 고려한 적응형 HE 방법이다[5]. 그러나 기본적인 영상처리 방법을 단독으로 사용할 경우 수중 영상의 가시성과 색상을 개선하는데 한계가 있다. 그래서 영상을 전처리한 후 기본적인 영상처리 방법을 병렬로 구성해 결과를 합하는 방법들이 제안되었다. Ancuti[6]등이 제안한 Image fusion 방법은 수중 영상을 White Balancing한 후 Gamma Correction과 Unsharp Masking을 병렬로 구성하고 결과를 합해 수중 영상을 개선한다. 그러나 White Balancing으로 본래 색을 찾기가 어렵고 병렬처리에서 영상에 적합한 기본적인 영상처리 방법을 선택하기가 어렵다.

영상 형성 모델링 기반 방법은 영상 형성을 수식으로 모델링하고 영상처리 방법을 적용하여 왜곡이 없는 영상을 구하는 방법이다. DCP(Dark Channel

Prior)는 영상 형성 모델 수식을 사용하여 안개가 제거된 영상을 얻었다[7]. 안개 영상은 안개 입자로 인한 빛의 산란과 배경광 때문에 자욱한 특징이 있다. 수중 영상 역시 물 분자 인한 자욱함 때문에 수중 영상을 개선하기 위해 안개 제거 방법들이 활용된다. Drews[8]등은 수중 환경에서 적색, 녹색, 청색광의 상대적인 전달량 차이를 고려한 DCP인 UDCP(Underwater DCP)를 제안하였다. 그러나 안개 제거 방법의 특성상 전반적인 색조를 개선하는데 한계를 보인다.

딥러닝(Deep Learning)의 방법 중 CNN은 영상 형성 모델링 수식에 필요한 정보를 얻기 위해 사용되거나[9] 개선된 영상을 직접 만드는데 사용된다. CNN의 모델 중 하나인 AE(Auto-Encoder)는 입력 영상의 내부 특징을 추출하고 한정된 특징을 바탕으로 입력 영상으로 복원시킨다. AE는 왜곡된 영상에서 본래 영상의 특징을 추출하고 본래 영상을 복원하는데 응용되어 안개 영상[10]이나 저조도 영상[11]을 개선했다. AE의 변형된 구조인 U-net[12]을 활용한 UGAN(Underwater Generative Adversarial Network)은 cycleGAN[13]을 통해 만든 수중 영상을 입력으로 하여 수중 영상을 개선하고자 하였다[14]. 그러나 cycleGAN으로 만든 수중 데이터 셋(Dataset)은 거리에 따른 왜곡 정도를 제대로 반영하지 못해 부자연스럽다. 이를 사용하여 학습한 UGAN은 인위적인 잡음을 생성하고 실제 수중 영상의 색조를 개선하는데 한계가 있다.

본 논문에서는 수중 영상의 가시성 개선과 색상 보정을 위해 AE에 Skip-connection을 적용한 UAE(Underwater AE)를 제안한다. 기존의 AE는 영상을 여러 차례 압축하면서 원래 영상의 정보가 손실되는 단점이 있다. UAE는 단점을 보완하기 위해 encoder와 decoder 전체에 skip-connection을 연결하여 수중 영상을 개선한다. UAE에 사용된 학습 데이터 셋은 수중 영상 형성 모델링 수식과 Jerlov 광 계수를 사용해 만들었다[15]. UAE의 성능을 입증하기 위해 검증 데이터 셋과 실제 수중 영상을 사용하여 UGAN을 비롯한 기존의 수중 영상 개선 방법들과 비교한다. 검증 데이터 셋을 통한 비교에서 UAE는 full reference metrics인 PSRN, SSIM[16], 그리고 색차에서 정량적으로 뛰어났고 실제 수중 영상에서도 정성적으로 가시성 개선과 색상 보정에 우수했다.

본 논문은 다음과 같이 구성된다. 2장에서는 AE를 기반으로 한 수중 영상 개선 방법에 대하여 설명한다. 3장에서는 제안하는 UAE에 관한 내용으로 수중

영상 형성 모델을 통한 학습 데이터 셋 합성, UAE의 구조, 그리고 UAE의 학습 방법에 대하여 설명한다. 4장에서는 검증 데이터 셋과 실제 수중 영상을 사용하여 UAE와 기존의 수중 영상 개선 방법들과 정량적, 정성적인 비교를 한다. 5장의 결론을 끝으로 본 논문을 마무리한다.



제 2 장 Auto-encoder를 기반으로 한 수중 영상 개선

2.1 수중 영상의 특징

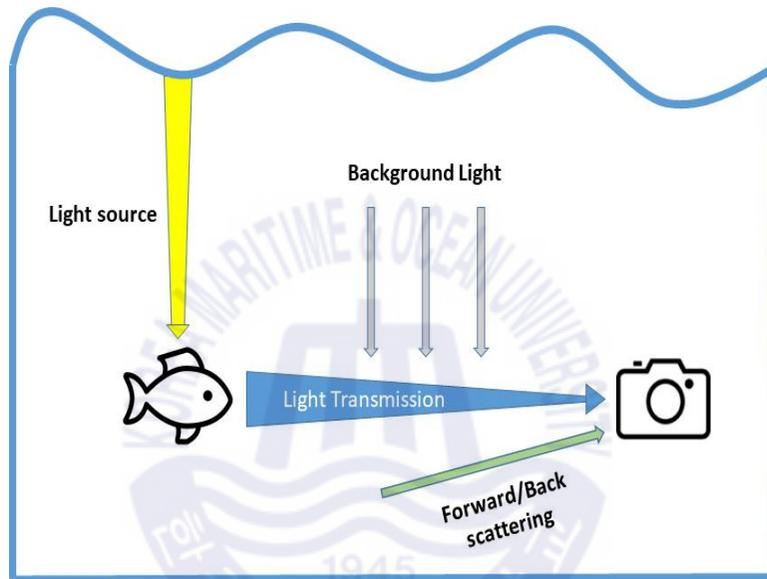


그림 2.1 물속에서의 영상 획득 과정

Figure. 2.1 Image acquisition processing in underwater

물속으로 투과된 광원은 물체에 반사되어 광학 장비에 도달한다. 이때 물속으로 투과된 광량은 수심에 따라 감소하고 물체와 광학 장비 사이의 거리에 따라 재차 감소한다. 또한 물 분자와 여러 부유물에 의한 빛의 흡수와 산란, 배경광 등의 원인으로 수중 영상은 가시성이 떨어지고 색상이 왜곡된다. 수중에서 광학 장비를 통해 영상이 획득되는 과정을 간략히 나타내면 그림 2.1과 같다.

수중 영상의 색상이 왜곡되는 주요 원인은 빛의 파장에 따른 왜곡 정도가 다르기 때문이다. 대양 바다에서 수심에 따른 빛의 파장별 전파는 긴 파장인 빨

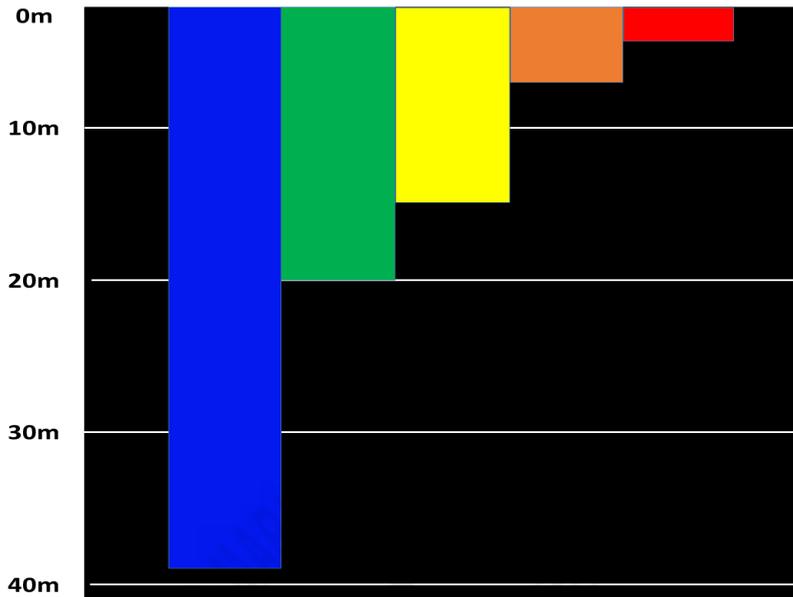


그림 2.2 대양에서 수심에 따른 빛의 전파

Figure. 2.2 Light penetration in ocean water according to water depth

간색에서 상대적으로 짧은 파장인 초록색과 파란색 보다 감소가 크다. 이것은 수중 영상이 전반적으로 초록빛 혹은 푸른빛의 색조를 띠는 원인이다. 그림 2.2는 대양에서 수심에 따른 빛의 전파 양상을 나타낸 것이다. 따라서 수중 영상은 수중이 아닌 안개 영상, 야간 영상 등의 저품질 영상과는 차이가 있고 더욱 복잡한 문제를 가진다. 그림 2.3은 쾌적한 환경에서 찍은 야외 영상, 안개 영상, 저조도 영상, 그리고 수중 영상의 차이를 보여준다.

2.2 AE를 기반으로 한 수중 영상 개선 방법

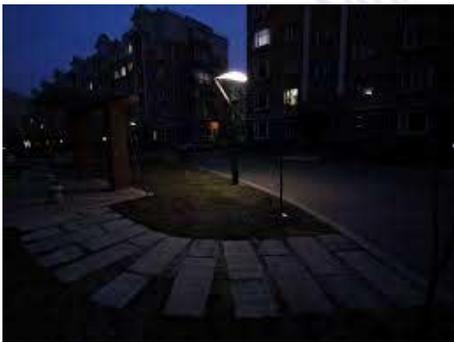
수중 영상은 안개 영상이 지닌 자욱함, 낮은 조도, 그리고 색상 왜곡 등의 특징이 있다. 그러나 수중 영상을 개선하기 위한 기본적인 영상처리 방법과 이를 조합한 Image fusion 방법, 형성 모델링 기반 방법 등은 수중 영상의 일



(a)



(b)



(c)



(d)

그림 2.3 다양한 환경에서 찍은 영상의 예

(a) 쾌적한 야외 영상, (b) 안개 영상

(b) 저조도 영상. (c) 수중 영상

Figure. 2.3 Sample images taken in various environment

(a) Clean image, (b) Haze image

(c) Low light image, (d) Underwater image

부 문제만을 해결하는데 초점을 두어 단편적인 성능을 보인다. 최근 CNN 기반의 딥러닝 방법은 진보한 컴퓨터 성능과 복잡한 네트워크 구조를 통해 기존의 영상처리 방법들을 대체하여 수중 영상이 가진 복합적인 문제를 해결하는데 활용되고 있다.

영상처리에서 convolution은 커널을 이용해 영상을 순회하여 영상을 필터링하는 방법이며 커널의 계수에 따라 스무딩, 샤프닝, 경계 검출 등 다양하게 활용된다. 그러나 단일 convolution은 객체 인식, 영상 분할 등의 고수준의 영상처리 문제를 해결하는데 한계가 있다. 따라서 convolution을 신경망처럼 연결하여 복잡한 영상처리 문제를 해결하는 CNN이 기존의 방법들을 대체하고 있다. CNN은 임의의 커널 계수를 가진 convolution을 여러 층으로 구성하고 학습을 통해 계수를 갱신한다. CNN은 사용 목적에 알맞은 학습 데이터 셋과 손실 함수를 구성하여 다양하게 활용된다. 여기서 손실함수란 CNN을 통해 얻고자 하는 목표값과 실제 CNN의 출력값 사이의 오차를 나타낸다. 많은 CNN 모델 중 AE는 데이터를 압축해 메모리 사용에 장점이 있어 많은 데이터를 한 번에 처리해야 하는 수중 영상 개선에 주로 활용된다.

AE는 Encoder와 Decoder로 이루어져있고 각각은 여러 개의 convolution 층으로 구성되어 convolution을 통한 특징맵을 출력한다. 이때 Encoder는 convolution의 커널이 순회하는 간격인 stride를 통해 출력하는 특징맵의 크기를 점차 줄여나가 입력 영상이 지닌 특징을 압축한다. **그림 2.4**은 convolution이 stride를 이용해 데이터를 압축하는 과정이다. 반면 decoder는 특징맵이 가진 압축된 정보를 점차 확대하여 원래 입력 영상으로 복원한다. 이때 먼저 up-sampling을 통해 특징맵의 크기를 키운 다음 convolution을 수행하는데 이를 transposed convolution이라 한다. **그림 2.5**는 transposed convolution을 통해 압축된 특징맵을 확대하는 과정이다.

또한 AE는 비지도 학습(Unsupervised Learning)의 대표적인 방법 중 하나이다. 여기서 비지도 학습이란 목표값에 대한 사전 학습 없이 입력 데이터의 특징 및 구조 등을 자율적으로 찾아내는 방식을 말한다. 따라서 AE는 원래 입력 영상을 특징을 추출하고 이를 다시 영상으로 재복원하는데 주로 사용된다. 그러나 Vincent[17]등은 AE에 지도 학습 개념을 도입하여 입력 데이터의 잡

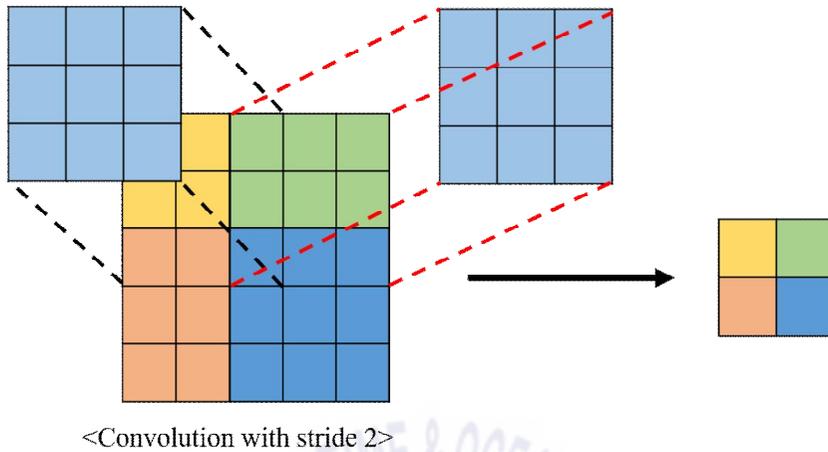


그림 2.4 Stride 2를 사용한 Convolution
 Figure. 2.4 Convolution with stride 2

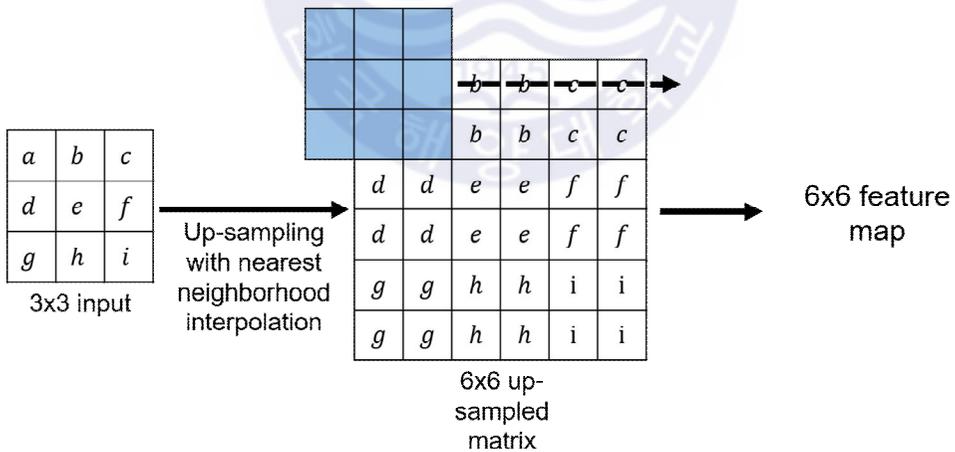


그림 2.5 최근접 이웃 보간법을 이용한 transposed-convolution
 Figure 2.5 Transposed-convolution using nearest neighborhood interpolation

음을 없애는 Denoising AE를 제안하였다. Denoising AE는 본래 데이터를 입력으로 하는 것이 아닌 잡음을 더한 새로운 데이터를 AE에 입력하여 본래 데이터에 근사하도록 학습을 한다. 이러한 지도 학습(Supervised Learning)을 바탕으로 AE는 본래 영상에 안개 추가, 조도 변형 등의 왜곡된 영상을 만들어 데이터 셋을 구성하고 이를 학습하여 영상 개선에 활용된다.

Fabrizi[14]등은 GAN에 cycleGAN[13]을 통해 만들어진 수중 데이터 셋을 학습하여 수중 영상을 개선하였다. 여기서 GAN은 AE를 기반으로 한 generator와 discriminator를 가진 CNN 모델 중 하나이다. Generator는 GAN을 통해 얻고자 하는 목표 영상에 유사한 위조 영상을 만들고 discriminator는 목표 영상과 generator가 만든 위조 영상을 비교해 위조 여부를 판단한다. GAN의 generator와 discriminator는 서로의 경쟁적인 관계를 손실함수로 사용해 서로를 학습시킨다. 이를 응용한 cycleGAN은 수중 영상이 아닌 영상을 수중 영상화하여 UGAN의 학습 데이터 셋을 만들었다. **그림 2.6**은 실제 수중 영상과 cycleGAN이 만든 수중 영상을 비교한 것이다. cycleGAN이 만든 수중 영상은 실제 수중 영상과 달리 광학 장비와 물체 사이의 거리에 따른 왜곡 정도를 표현하는데 한계가 있고 **그림 2.6**의 **(b-3)**, **(b-4)**의 영상처럼 인위적이다. cycleGAN이 만든 데이터 셋을 학습한 UGAN은 비교적 먼 거리에 있는 물체의 가시성을 개선하는데 한계가 있고 인위적인 잡음을 포함한 결과를 만든다.

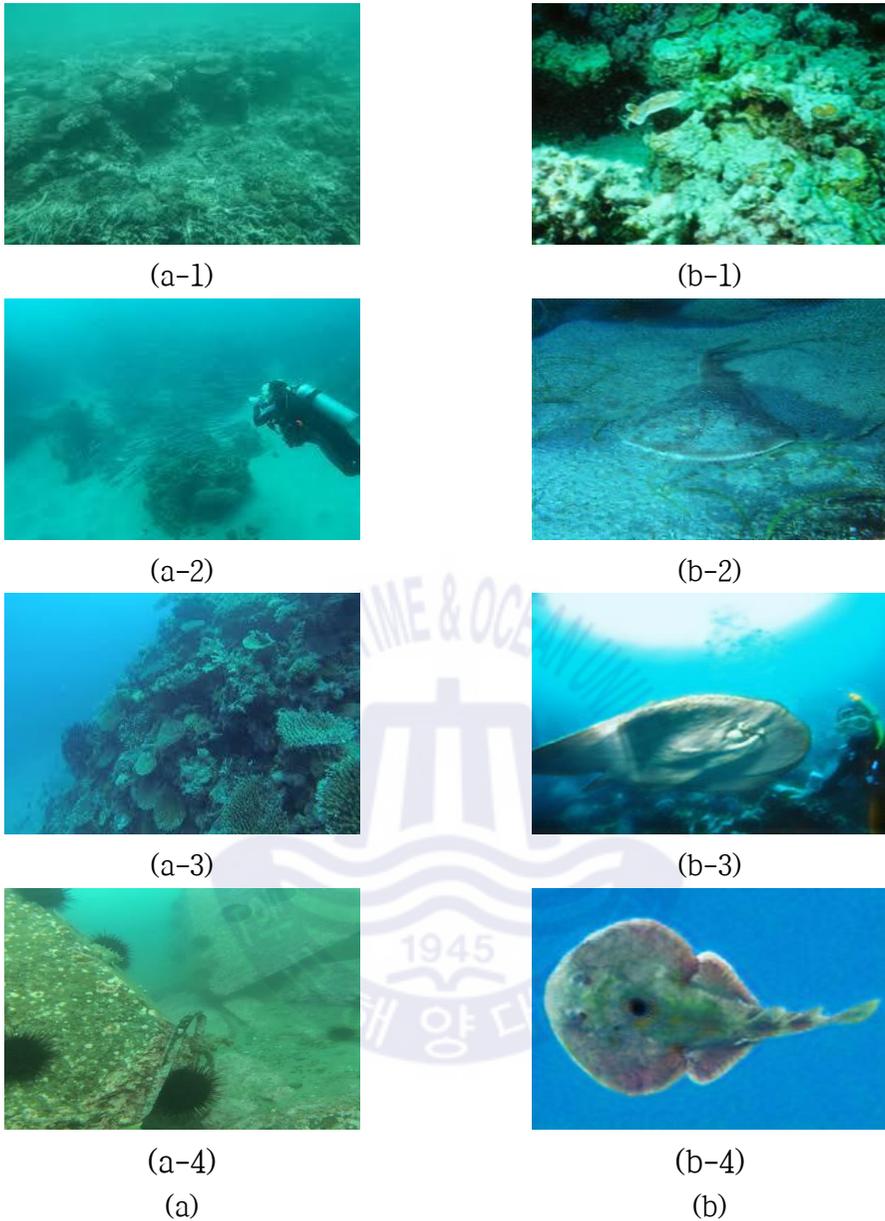


그림 2.6 실제 수중 영상과 cycleGAN이 만든 수중 영상의 비교

(a) 실제 수중 영상

(b) cycleGAN이 만든 수중 영상

Figure. 2.6 Comparison of actual underwater images and underwater images generated by cycleGAN

(a) Actual underwater images

(b) Underwater images generated by cycleGAN

제 3 장 수중 영상 개선을 위한 Underwater Auto-Encoder

3.1 수중 영상 형성 모델을 통한 학습 데이터 셋 합성

그림 2.1에서와 같이 물속에서 영상이 획득되는 과정을 수심, 카메라와 물체 간의 수평거리, 빛의 감쇄 정도, 그리고 배경광 등을 고려하여 식 (3.1)과 같이 모델링하였다[15].

$$\begin{aligned} I_{\lambda}^{under}(i,j) &= C_{\lambda}(i,j)I_{\lambda}^{clean}(i,j)T_{\lambda}(i,j) + (1 - T_{\lambda}(i,j))B_{\lambda} \\ &= e^{-\alpha(\lambda)D(i,j)}I_{\lambda}^{clean}(i,j)e^{\beta(\lambda)d(i,j)} + (1 - e^{\beta(\lambda)d(i,j)})B_{\lambda} \end{aligned} \quad (3.1)$$

i, j 는 영상에서 현 좌표를 의미한다. $I_{\lambda}^{under}(i, j)$ 는 모델링을 통해 얻어진 수중 영상이다. $I_{\lambda}^{clean}(i, j)$ 는 왜곡이 없는 영상을 나타낸다. $C_{\lambda}(i, j)$ 는 색상 왜곡 항으로 수심, $D(i, j)$ 와 광 흡수 계수 $\alpha(\lambda)$ 을 고려한 감소함수이다. $T_{\lambda}(i, j)$ 는 물체에서 카메라로 전달되는 빛의 양으로 물체와 카메라 사이의 거리, $d(i, j)$ 와 산란 계수, $\beta(\lambda)$ 를 고려한 감소함수이다. B_{λ} 는 배경광이다. 따라서 식 (3.1)에 $I_{\lambda}^{clean}(i, j)$, $D(i, j)$, $d(i, j)$, $\alpha(\lambda)$, 그리고 $\beta(\lambda)$ 를 대입하여 수중 영상을 합성한다.

본 논문에서는 학습 데이터 셋을 만들기 위해 왜곡이 없는 영상과 해당 영상의 거리정보를 제공하는 NYU dataset을 사용한다. NYU dataset은 마이크로소프트사의 depth camera인 Kinect로 찍은 실내 영상 데이터로 총 1449장의 R-G-B 값과 거리정보를 제공한다. 또한 실제 수중의 광 산란과 흡수 현상이 재현된 수중 영상을 합성하기 위해 Jerlov 광 계수를 적용한다. Jerlov는 실험에서 빛의 전달량 측정을 바탕으로 5가지 유형의 연안 바다와 5가지 유형의 대양 바다에 대한 광 산란 계수와 흡수 계수를 제시하였다. 본 논문에서는 수중 영상 합성을 위해 표 3.1과 같이 5가지 대양 바다의 적색, 녹색, 청색의 광 계수를 사용하였다.

표 3.1 5가지 대양의 유형별 빛의 파장에 따른 광 계수

Table 3.1 Light coefficient of 5 types ocean water according to the wavelength of light

Wave-length	I		IA		IB		II		III	
	$\alpha(\lambda)$	$\beta(\lambda)$								
650nm (Red)	0.334	0.0009	0.334	0.0023	0.334	0.393	0.334	0.27	0.336	0.74
525nm (Green)	0.046	0.0021	0.047	0.0040	0.047	0.078	0.047	0.387	0.051	1.06
450nm (Blue)	0.018	0.0038	0.022	0.0063	0.024	0.062	0.024	0.504	0.039	1.38

수중 영상 형상 모델링 수식에 필요한 수심 정보와 배경광의 크기는 $D(i, j) \in \{0.5m, 15m\}$, $B_{Red} \in \{0.2, 0.45\}$, $B_{Green} \in \{0.5, 0.75\}$, $B_{Blue} \in \{0.5, 0.8\}$ 의 랜덤한 값을 적용한다. 학습을 위해 1449개의 NYU dataset 중 1000개의 데이터 셋을 사용해 5가지 대양 바다 유형이 혼합된 10000장의 데이터 셋을 만든다. 나머지 449개의 데이터 셋은 검증 데이터 셋을 만들어 네트워크 성능 검증에 활용한다. 그림 3.1 ~ 3.3은 위 내용을 바탕으로 구성한 학습 데이터 예를 보여 준다.

3.2 제안하는 UAE의 구조

제안한 UAE는 4개의 convolution 층을 가진 encoder, 4개의 residual block, 그리고 4개의 transposed convolution 층을 가지며 그림 3.4와 같이 대칭적인 구조를 갖는다. Residual block은 그림 3.4의 아래와 같이 2 개의 convolution 층과 skip-connection으로 구성되어있다. Residual block의 skip-connection은 CNN을 구성할 때 생기는 기울기 소실문제를 극복하여 convolution 층을 늘릴 수 있도록 한다.

기존의 AE는 encoder가 깊은 convolution 층을 가질수록 입력 영상의 특징을



(a)



(b)



(c)



(d)



(e)



(f)

그림 3.1 수중 영상 형성 모델 , NYU dataset,

Jerlov 광 계수를 적용해 만든 학습 데이터 셋

(a) 원본 NYU dataset 영상

(b) ~ (f) Jerlov의 5가지 대양 바다의 광 계수를 적용한 영상

(b) 유형 I, (c) 유형 IA, (c) 유형 IB, (c) 유형 II, (c) 유형 III

Figure. 3.1 Training dataset synthesized by underwater image formation model equation, NYU dataset, and Jerlov light coefficients

(b) ~ (f) Synthesized underwater images using 5 ocean water types of Jerlov

(b) Type I, (b) Type IA, (b) Type IB, (b) Type II, (b) Type III



(a)



(b)



(c)



(d)



(e)



(f)

그림 3.2 수중 영상 형성 모델 , NYU dataset,
Jerlov 광 계수를 적용해 만든 학습 데이터 셋
(a) 원본 NYU dataset 영상

(b) ~ (f) Jerlov의 5가지 대양 바다의 광 계수를 적용한 영상
(b) 유형 I, (c) 유형 IA, (c) 유형 IB, (c) 유형 II, (c) 유형 III

Figure. 3.2 Training dataset synthesized by underwater image formation model equation, NYU dataset, and Jerlov light coefficients

(b) ~ (f) Synthesized underwater images using 5 ocean water types of Jerlov
(b) Type I, (b) Type IA, (b) Type IB, (b) Type II, (b) Type III



(a)



(b)



(c)



(d)



(e)



(f)

그림 3.3 수중 영상 형성 모델 , NYU dataset,

Jerlov 광 계수를 적용해 만든 학습 데이터 셋

(a) 원본 NYU dataset 영상

(b) ~ (f) Jerlov의 5가지 대양 바다의 광 계수를 적용한 영상

(b) 유형 I, (c) 유형 IA, (c) 유형 IB, (c) 유형 II, (c) 유형 III

Figure. 3.3 Training dataset synthesized by underwater image formation model equation, NYU dataset, and Jerlov light coefficients

(b) ~ (f) Synthesized underwater images using 5 ocean water types of Jerlov

(b) Type I, (b) Type IA, (b) Type IB, (b) Type II, (b) Type III

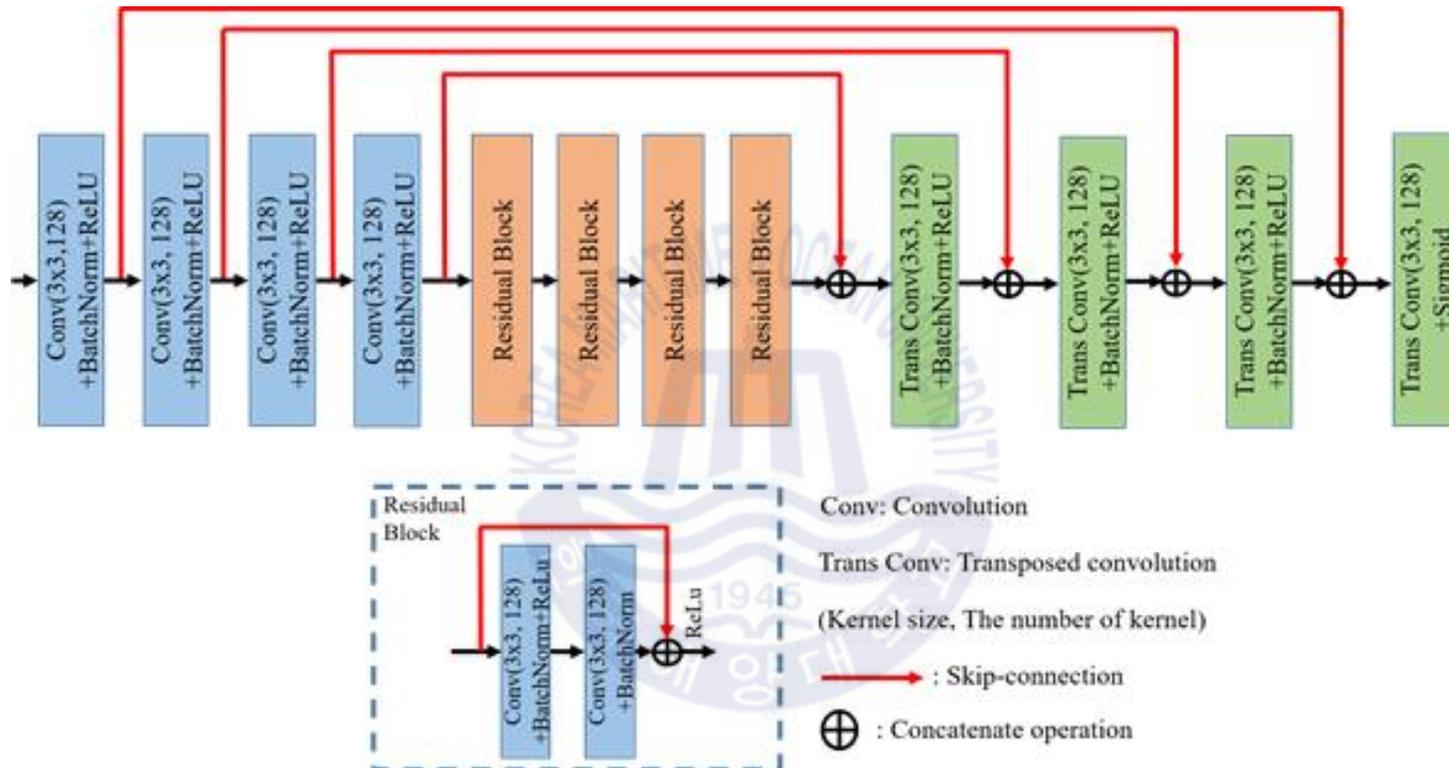


그림 3.4 Skip-connection을 가진 제안하는 UAE의 구조
 Figure 3.4 Architecture of proposed UAE with skip-connection

더 작은 크기의 특징맵으로 압축하게 되어 원본 영상이 갖는 형태와 세부 정보가 손실된다[12]. 그 결과 색상 개선 능력이 떨어지고 잡음이 추가된 영상을 만든다. 이 문제를 극복하기 위해 encoder와 decoder를 skip-connection으로 연결하였다. Skip-connection은 encoder의 출력 특징맵을 decoder의 입력 특징맵과 합친다(Concatenate Operation). 이는 손실되기 전 입력 영상의 특징을 decoder에서 재차 고려하여 decoder의 영상 복원 성능을 개선한다.

커널 크기는 AE를 사용할 때 중요한 파라미터 중 하나이다. 커널 크기가 작을수록 원본 이미지의 세부 정보, 특히 에지 정보를 보존하는 데 효과적이고 [18] 학습에 사용되는 메모리 및 가중치 계산 비용을 줄인다. 그러나 가장 작은 커널 크기인 1x1 커널은 주변 화소의 크기 및 형태를 고려하지 않은 지역적 정보를 추출한다. 따라서 주변 화소를 고려하고 네트워크 학습 비용을 줄이기 위해 네트워크 전체에 3x3 커널을 사용하고 커널 개수는 128개로 설정하였다. 또한 학습 효율을 위해 BN(Batch Normalization)과 ReLU(Rectified Logical Unit) 활성화 함수를 사용한다. 영상을 복원하는 마지막 transposed convolution 층에서는 결과 화소값이 0에서 1이 되도록 sigmoid 함수를 활성화 함수로 사용했다.

3.3 UAE의 학습 방법

UAE는 출력 영상과 수중 데이터 셋의 왜곡 없는 영상 간의 화소값 차이를 표현하는 MSE (Mean Squared Error)를 손실함수로 적용하였다.

$$MSE = \frac{\sum_{\lambda \in \{Red, Green, Blue\}} (I_{\lambda}^{result}(i,j) - I_{\lambda}^{clean}(i,j))^2}{3N} \quad (3.2)$$

$I_{\lambda}^{result}(i,j)$ 는 UAE의 결과 영상이고 $I_{\lambda}^{clean}(i,j)$ 은 수중 데이터 셋의 왜곡이 없는 영상이다. i, j 는 영상에서 현재 화소의 위치이다. N 은 영상의 총 화소 개수이다. UAE의 학습은 MSE를 줄이기 위한 최적의 convolution의 계수를 찾는 과정이다. 이때 optimizer는 현재 MSE를 반영하여 learning rate 만큼 계수의 크기를 변화시킨다. 본 논문에서는 손실함수를 줄이는데 효과적인 Adam optimizer를 사용하여 학습을 진행한다.

제 4 장 실험 및 고찰

본 논문에서는 python과 딥러닝 라이브러리인 pytorch를 사용해 네트워크를 구현하였고 CPU: AMD Ryzen 7 2700X, GPU: Nvidia Geforce RTX 2070 super, RAM: 16G 환경에서 UAE의 학습을 진행하였다. UAE는 10000장의 학습 데이터 셋을 사용하여 8 batch size와 0.0005의 학습률을 적용하여 50 epochs 동안 학습을 진행하였다.

제안하는 UAE의 성능을 입증하기 위해 검증 데이터 셋을 사용하여 UDCP, Image fusion, 그리고 UGAN과 성능을 비교하였다. 검증 시 UGAN은 cycleGAN 데이터 셋을 통해 미리 학습된 것을 사용하였다. 검증 데이터 셋은 학습 데이터에서 사용하지 않은 449개의 NYU dataset 사용해 학습 데이터 합성과 동일한 방식으로 구성하였다. 각 방법의 성능을 정량적으로 비교하기 위하여 full-reference metrics인 PSNR, SSIM, 그리고 R-G-B 채널에 대한 색차를 사용하였다. Full-reference metrics는 각 방법의 결과 영상을 왜곡되지 않은 참조 영상과 비교하여 영상의 품질 차이를 나타내는 비교 지표이다. PSNR은 최대 신호에 대한 잡음 비로 압축된 영상이나 재생성된 영상의 원본 영상 대비 손실율을 나타내며 결과 영상의 PSNR값이 높을수록 손실율이 적은 영상이다. SSIM은 인간이 시각적으로 영상의 품질을 판단할 때 고려되는 영상의 휘도, 대비, 그리고 구조를 고려한 품질 지표이며 결과 영상의 SSIM값이 높을수록 품질이 우수한 영상이다. 색차는 서로 다른 R-G-B 화소값의 절대적 크기 차이를 유클리드 거리로 표현한 것이다. 본 논문에서는 영상의 전체 화소에 대해 색차를 구하고 그것의 평균치를 결과 영상의 색차로 나타냈다. 결과 영상은 색차값이 작을수록 참조 영상과 유사한 영상이다.

그림 4.1 ~ 4.5는 대양 바다의 5가지 유형별 검증 데이터 셋의 예시와 각 방법의 결과를 보여준다. UDCP는 모든 유형에서 가시성 과 색상 개선에 효과적이지 못하며 오히려 색조를 더욱 어둡게 만들거나 적색을 과하게 보정 하는 경향을 보였다. Image fusion은 왜곡이 적은 I 유형에서 가시성 개선과 색상 보정

표 4.1 전체 검증 데이터 셋에 대한 PSNR, SSIM, 그리고 색차의 평균값

	UDCP	Image fusion	UGAN	UAE
PSNR	10.5885	14.0734	14.7615	22.0598
SSIM	0.4726	0.7199	0.7220	0.8797
Color difference	0.4760	0.3525	0.3031	0.1404

에 효과적이지만 그 외 모든 유형에서 원본 영상보다 밝은 영상을 만들었고 적색을 과하게 보정하였다. 또한 초록빛 색조가 강한 II 유형에서 색상을 보정하는데 한계를 보였다. UGAN은 모든 유형에서 가시성 개선과 색상 보정에 미흡한 성능을 보였다. 반면 UAE는 모든 유형에서 색상을 개선하는데 우수하며 왜곡이 심한 III의 경우를 제외한 나머지 유형에서 가시성 개선 효과적인 성능을 보였다.. 표 4.1은 전체 검증 데이터 셋에 대한 PSNR, SSIM, 그리고 색차의 평균값을 보여준다.

정성적인 결과와 마찬가지로 UDCP가 모든 수치에서 좋지 못한 결과를 보였다. Image fusion과 UGAN은 PSNR과 SSIM값이 비슷하나 빨간색을 과하게 보정하는 Image fusion의 색차값이 UGAN에 비해 높게 나왔다. UAE는 정성적으로 우수함과 동시에 정량적인 수치에서도 좋은 결과를 보였다.

그림 4.6 ~ 4.10은 실제 수중 영상을 각 방법에 적용한 결과를 나타낸다. UDCP은 검증 데이터 셋의 경우와 마찬가지로 영상의 색조를 더욱더 어둡고 진하게 하는 경향을 보였다. Image fusion은 가시성 개선에 효과가 있으나 보편적인 색 보정 성능을 기대할 수 없고 과한 보정으로 잡음이 생겼다. UGAN은 보편적인 색 보정 성능을 기대할 수 없고 영상이 불분명하며 인위적인 잡음을 만드는 경향을 보였다. 반면 UAE는 검증 데이터 셋의 경우와 마찬가지로 모든 실제 수중 영상에서 가시성 개선과 색상 보정에 보편적인 성능을 보였다.



(a)



(b)



(c)



(d)



(e)



(f)

그림 4.1 유형 I에 대한 검증 데이터 셋과 각 방법들의 결과 영상

(a) 원본 영상, (b) 수중 영상, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE

Figure 4.1 Test dataset sample of type I and result images of each methods

(a) Clean image, (b) Underwater image, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE



(a)



(b)



(c)



(d)



(e)



(f)

그림 4.2 유형 IA에 대한 검증 데이터 셋과 각 방법들의 결과 영상

(a) 원본 영상, (b) 수중 영상, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE

Figure 4.2 Test dataset sample of type IA and result images of each methods

(a) Clean image, (b) Underwater image, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE



(a)



(b)



(c)



(d)



(e)



(f)

그림 4.3 유형 IB에 대한 검증 데이터 셋과 각 방법들의 결과 영상

(a) 원본 영상, (b) 수중 영상, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE

Figure 4.3 Test dataset sample of type IB and result images of each methods

(a) Clean image, (b) Underwater image, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE



(a)



(b)



(c)



(d)



(e)



(f)

그림 4.4 유형 II에 대한 검증 데이터 셋과 각 방법들의 결과 영상

(a) 원본 영상, (b) 수중 영상, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE

Figure 4.4 Test dataset sample of type II and result images of each methods

(a) Clean image, (b) Underwater image, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE



(a)



(b)



(c)



(d)



(e)



(f)

그림 4.5 유형 III에 대한 검증 데이터 셋과 각 방법들의 결과 영상

(a) 원본 영상, (b) 수중 영상, (c) UDCP

(d) Image fusion, (e) UGAN, (f) UAE

Figure 4.5 Test dataset sample of type III and result images of each methods

(a) Clean image, (b) Underwater image, (c) UDCP

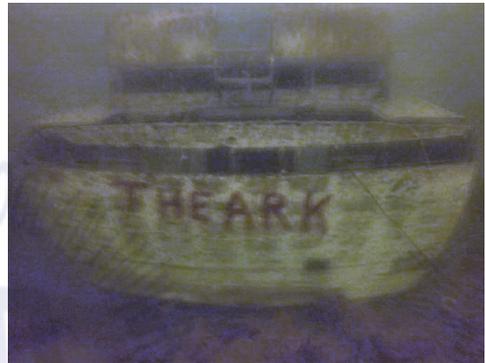
(d) Image fusion, (e) UGAN, (f) UAE



(a)



(c)



(d)



(e)



(f)

그림 4.6 실제 수중 영상 대한 결과 영상

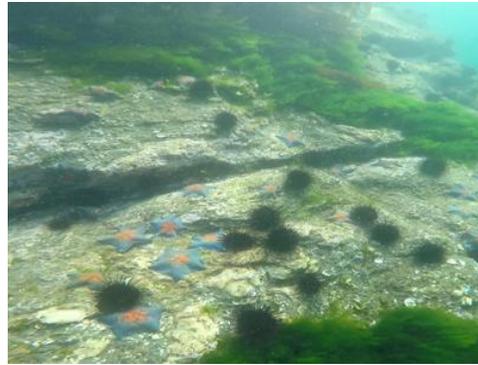
(a) 수중 영상

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE

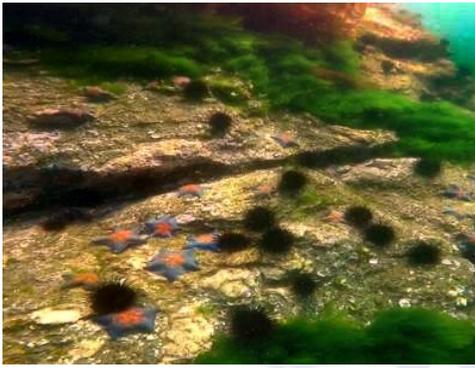
Figure 4.6 Result image of each methods about actual underwater image

(a) Underwater image

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE



(a)



(c)



(d)



(e)



(f)

그림 4.7 실제 수중 영상 대한 결과 영상

(a) 수중 영상

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE

Figure 4.7 Result image of each methods about actual underwater image

(a) Underwater image

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE



(a)



(c)



(d)



(e)



(f)

그림 4.8 실제 수중 영상 대한 결과 영상

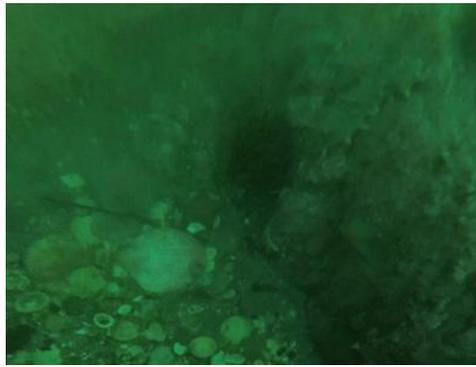
(a) 수중 영상

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE

Figure 4.8 Result image of each methods about actual underwater image

(a) Underwater image

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE



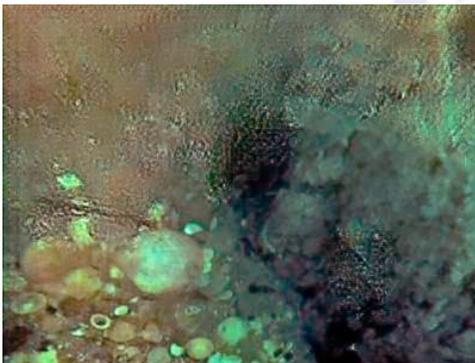
(a)



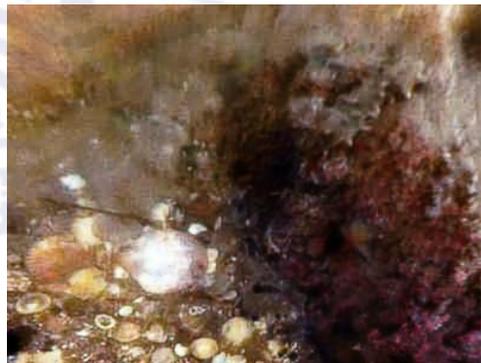
(c)



(d)



(e)



(f)

그림 4.9 실제 수중 영상 대한 결과 영상

(a) 수중 영상

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE

Figure 4.9 Result image of each methods about actual underwater image

(a) Underwater image

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE



(a)



(c)



(d)



(e)



(f)

그림 4.10 실제 수중 영상 대한 결과 영상

(a) 수중 영상

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE

Figure 4.10 Result image of each methods about actual underwater image

(a) Clean image

(b) UDCP, (c) Image fusion, (d) UGAN, (e) UAE

제 5 장 결론

본 논문에서는 자욱함, 낮은 조도, 그리고 색상 왜곡 등의 문제를 가지는 수중 영상을 개선하기 위해 CNN 기반의 딥러닝 방법을 활용하였다. 다양한 CNN 모델 중 영상이 가진 데이터를 압축해 메모리 사용에 효율적이고 영상 복원에 효과적인 AE 모델을 사용하였다. 그러나 기존의 AE는 영상이 가진 정보를 압축하는 과정에서 원본 영상이 갖는 형태와 세부 정보가 손실되는 단점이 있다. 이를 해결하기 위해 제안하는 UAE는 encoder와 decoder 전체에 skip-connect를 연결하여 모델의 성능을 개선했다. 또한 실제 수중 영상과 유사한 학습 데이터 셋을 만들기 위해 수중 영상 형성 모델에 NYU dataset 과 Jerlov 광 계수를 적용하였다.

제안한 UAE 모델의 우수성을 검증하기 위해 학습 데이터 셋을 합성하는 것과 동일한 방법으로 검증 데이터 셋을 구성하고 UDCP, Image fusion, UGAN에 각각 적용하여 결과를 정성적 정량적으로 비교 분석하였다. 제안하는 UAE는 검증 데이터 셋에서 왜곡 정도가 적은 I, IA, 그리고 IB 유형에서 다른 방법들 보다 가시성 개성과 색상 보정에 모두 뛰어난 결과를 보였다. 하지만 왜곡 정도가 심한 II와 III 유형의 경우 전반적인 색조를 개선하였지만 영상의 세부 특징과 형태를 개선하는데 미흡한 결과를 보였다. 그러나 다른 방법들에 비해 정성적으로 우수한 성능을 보였다. 검증 데이터 전체에 대한 정량적 평가에서도 제안하는 UAE가 가장 뛰어난 결과를 보였다. 또한 몇 가지 실제 수중 영상을 통한 검증에서도 제안하는 UAE가 정성적으로 가장 보편적인 성능을 보였다.

제안한 UAE는 학습 시 수중 데이터 셋을 사용하여 수중 영상 개선에 특화된 모델이다. 하지만 UAE는 내부구조, 학습 데이터 셋 등에 따라 목적과 성능이 달라지는 만큼 안개 영상, 저조도 영상 등 영상 품질 향상을 위한 다양한 분야에 적용할 수 있을 것이다. 또한 추후 연구를 통해 왜곡이 심한 영상에서도 세부 특징과 형태를 개선할 수 있도록 모델을 수정 보완할 예정이다.

참 고 문 헌

- [1] K. Lebart, C. Smith, E. Trucco and D. M. Lane, “Automatic Indexing of Underwater Survey Video: Algorithm and Benchmarking Method,” in *IEEE Journal of Oceanic Engineering*, vol. 28, no. 4, pp. 673–686, doi: 10.1109/JOE.2003.819314, 2003.
- [2] I. Leonard, A. Alfalou, and C. Brosseau, “Sensitive Test for Sea Mine Identification Based on Polarization-Aided Image Processing,” *Optics express*, vol. 21, pp. 29283–97, 10.1364/OE.21.029283, 2013.
- [3] G. Bianco, M. Muzzupappa, F. Bruno, R. Garcia, and L. Neumann, “A New Color Correction Method for Underwater Imaging,” *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol.XL-5/W5, pp.25–32, 2015.
- [4] O. Deperlioglu, U. kose, and G. Guraksin, “Underwater Image Enhancement with HSV and Histogram Equalization,” *Proceedings of International Conference on Advanced Technologies*, 2018.
- [5] S. Anilkumar, P. R. Dhanya, A. A. Balakrishnan, and M. H. Supriya, “Algorithm for Underwater Cable Tracking Using CLAHE based Enhancement,” *International Symposium on Ocean Technology (SYMPOL)*, 2019, pp. 129–137, 2019
- [6] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, “Color Balance and Fusion for Underwater Image Enhancement,” in *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 379–393, 2018.
- [7] K. He, J. Sun and X. Tang, “Single Image Haze Removal Using Dark Channel Prior,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.

- [8] P. Drews-Jr, E. Nascimento, F. Moraes, S. Botelho, and M. Campos, "Transmission Estimation in Underwater Single Images," Proceedings of IEEE International Conference on Computer Vision Workshops, pp. 825-830, 2013.
- [9] M. Hou, R. Liu, X. Fan, and Z. Luo, "Joint Residual Learning for Underwater Image Enhancement," Proceedings of IEEE International Conference on Image Processing, pp. 4043-4047, 2018.
- [10] S. Bianco, L. Celona, F. Piccoli, and R. Schettini, "High-Resolution Single Image Dehazing Using Encoder-Decoder Architecture," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1927-1935, doi: 10.1109/CVPRW.2019.00244, 2019.
- [11] K. Lore, A. Akintayo, and S. Sarkar, "LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement," Pattern Recognition, vol. 61, 10.1016/j.patcog.2016.06.008, 2015.
- [12] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional Networks for Biomedical Image Segmentation," <http://arxiv.org/abs/1505.04597>, 2015.
- [13] J. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," arXiv preprint arXiv:1703.10593, 2017.
- [14] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing Underwater Imagery Using Generative Adversarial Networks," Proceedings of IEEE International Conference on Robotics and Automation, pp. 7159-7165, 2018.
- [15] X. Ding, Y. Wang, Y. Yan, Z. Liang, Z. Mi, and X. Fu, "Jointly Adversarial Network to Wavelength Compensation and Dehazing of Underwater Images," arXiv:1907.05595, 2019
- [16] W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE

Transactions on Image Processing. Vol. 13, Issue 4, pp. 600–612, 2004.

- [17] P. Vincent, H. Larochelle, Y. Bengio, and P. Manzagol, “Extracting and Composing Robust Features with Denoising Autoencoders,” Proceedings of the 25th International Conference on Machine Learning. pp. 1096–1103, 10.1145/1390156.1390294, 2008.
- [18] J. Yan, C. Li, Y. Zheng, S. Xu, and X. Yan, “MMP-Net: A Multi-Scale Feature Multiple Parallel Fusion Network for Single Image Haze Removal,” in IEEE Access, vol. 8, pp. 25431–25441, 2020.

