

순환 손실 함수를 이용한 딥러닝 기반 위상 홀로그램 초해상도

차준영, 반현민, 최승미, 김휘용
경희대학교

{cjy8922, hmban1996, 2018102242, hykim.v}@khu.ac.kr

Deep Learning-based Phase-Only Hologram Super Resolution using Circular Loss

Junyeong Cha, Hyunmin Ban, Seungmi Choi, Hui Yong Kim
Kyunghee University

요약

홀로그램(Hologram)은 3차원 물체에서 나오는 빛의 정보를 제어하는 기술이다. 현재는 컴퓨터 생성 홀로그램(CGH)으로 생성한 디지털 홀로그램에 관한 연구, 특히 물체에서 나오는 빛의 정보를 최대한 기록하고 재현하여 디지털 홀로그램의 해상도를 향상시키려는 연구가 활발히 진행되고 있다. 이에 본 논문에서는 고해상도 홀로그램 영상을 얻기 위해 딥러닝 기반 초해상도(Super Resolution) 네트워크를 훈련 및 최적화하여, 저해상도 위상 홀로그램 영상으로부터 높은 화질의 홀로그램 영상을 재현하는 고해상도 위상 홀로그램 영상을 생성하는 것을 목표로 한다. 이때 위상 홀로그램 영상의 특성을 이용한 순환 손실 함수(Circular loss function)를 새롭게 제안하며, 기존의 이미지 초해상도 신경망 모델을 학습시킬 때 자주 사용하는 L1 손실 함수와 비교했을 때 약 0.13dB 정도의 성능 향상이 있었다.

1. 서론

홀로그래피는 3차원상에 존재하는 물체에서 나오는 빛의 진폭과 위상을 특수한 재료로 만들어진 2차원의 홀로그램 필름에 기록하는 기술이다^[1]. 이후, 실존하는 물체가 아닌 가상의 물체에서 나오는 빛의 정보를 컴퓨터로 계산하여 기록하는 컴퓨터 생성 홀로그램(CGH)^[2]과 이를 디스플레이 하기 위한 디지털 공간광변조기(SLM) 관련 기술이 등장하였다. 컴퓨터로 생성한 홀로그램을 완벽히 복원하여 디스플레이 하기 위해서는 빛의 진폭과 위상을 모두 표현할 수 있는 공간광변조기가 요구되지만, 현재 기술적으로 빛의 진폭 혹은 위상만 기록할 수 있으며, 현재는 빛을 더 효율적으로 표현할 수 있는 위상 공간광변조기가 주로 사용된다. 이에 맞추어 위상 공간광변조기에서 홀로그램 영상을 복원할 수 있게 빛의 위상만 기록한 홀로그램을 위상 홀로그램(Phase-only Hologram)이라 한다.

기술이 발전함에 따라 고해상도 이미지를 디스플레이 할 수 있게 되면서, 최근에는 저해상도 이미지를 고해상도 이미지로 변환하는 초해상도(Super Resolution) 기술에 대한 연구가 이루어지고 있다. 초해상도 기술은 Bilinear, Bicubic 등의 보간법(interpolation)을 이용하는 방식에서 SRCNN (Super Resolution Convolution Neural Network)^[3]을 시작으로 CAR (Content Adaptive Resampler)^[4], SRGAN (Super Resolution Generative Adversarial Network)^[5] 등 딥러닝 기반 초해상도 신경망 모델을 이용하는 방식으로 발전해왔다. 위상 홀로그램 영상은 일반적인 2차원 이미지에 비해 공간적인 상관도가 낮아 한 픽셀 단위의 정보가 중요해지게 되는데, 고해상도 홀로그램 영상을 얻기 위해 보간법을 사용하게 되면 위상 홀로그램의 중요한 정보가 변질되어 잡음과 잔상이 나타나는 결과를 초래한다^[6]. 이러한 이유로 딥러닝 기반으로

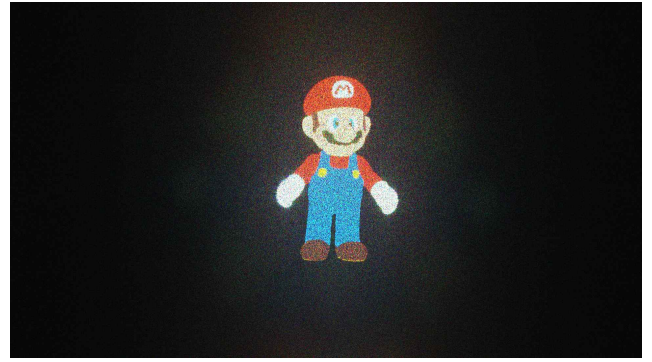
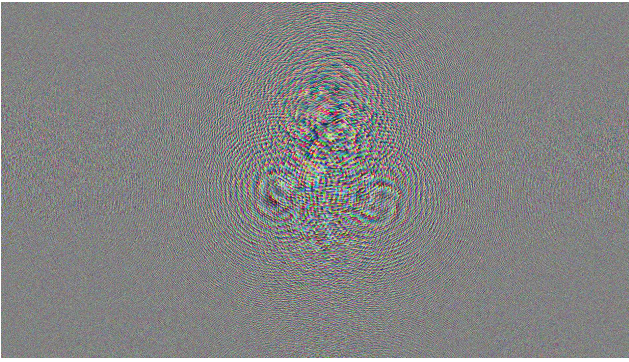
고해상도 위상 홀로그램 영상을 얻기 위한 연구가 활발하게 진행되고 있으며, 2차원 이미지에서 고해상도 영상을 얻기 위한 딥러닝 기반 초고해상도 신경망 모델을 고해상도 홀로그램 영상을 얻기 위한 목적으로 사용했을 때 좋은 성능을 보이기도 했다.

본 논문에서는 고해상도 홀로그램 영상을 얻기 위한 딥러닝 기반 초해상도 신경망 모델의 성능을 높이는 데 주력했다. 특히 2차원 이미지의 단일 픽셀 유사도를 계산하는 L1 손실 함수가 아닌, 2π 주기로 같은 정보를 담고 있는 위상 홀로그램의 특성을 이용하여 L1 기반의 순환 손실 함수, L1-Circular 손실 함수를 제안한다. 논문의 구성은 다음과 같다. 먼저, 위상 홀로그램 데이터와 2차원 이미지에서 좋은 성능을 보인 초고해상도 신경망 모델을 소개한다. 둘째, 위상 홀로그램 영상의 특성을 이용한 L1 기반의 순환 손실 함수를 제시한다. 셋째, 앞서 소개한 신경망 모델에 L1 loss function을 사용하여 학습한 결과와 L1-Circular loss function을 사용하여 학습한 결과를 비교한다. 마지막으로 결론을 지으면서 향후 연구 과제를 도출한다.

2. 위상 홀로그램 데이터 및 초해상도 기술

1) 위상 홀로그램 데이터

디지털 홀로그램 영상을 생성할 때, Point Cloud 기반으로 생성한 데이터를 RGBD (RGB + depth)로 변환한 후, 랜덤으로 생성한 위상 데이터를 추가한다. 그리고 layer 기반의 Fresnel CGH 방식을 이용해 복소 홀로그램을 계산하고 위상 데이터만을 뽑아 위상 홀로그램으로 만든다. 이때, 홀로그램 생성에서 공간광변조기의 화소 크기는 $8\mu\text{m}$ 이고, RGB 파장은 각각 660nm, 532nm, 473nm이다. 수치 복원 시에도 동일하게 Fresnel 방식을 이용한다. 이렇게 생성된 위상 홀로그램 영상은



<Figure 1> Phase-only hologram (Left), Numerical reconstruction of phase-only hologram (Right)

1920×1080의 해상도를 가진다. 총 7가지 가상의 물체를 이용해 위상 홀로그램 영상을 생성했으며, 하나의 물체를 가지고 만든 위상 홀로그램 영상을 하나의 시퀀스라고하면, 각 시퀀스는 시점을 다르게 하며 총 33개의 프레임을 생성한다. Figure 1은 위 방법으로 생성한 위상 홀로그램 영상과 그것을 복원한 결과이다. 여기서 위상 이미지는 $0\sim 2\pi$ 의 값을 이미지의 bit level에 맞춰 양자화하여 표현한 것으로 8bit level에서 위상 홀로그램 영상은 $0\sim 255$ 사이의 값으로 균일 양자화된다.

2) 초해상도 신경망 모델 구조

본 논문에서 사용된 초해상도 신경망 모델은 해상도를 높이기 위해 사용된 신경망 구조를 경량화하면서도 화질을 유지하기 위한 목적으로 개발된 모델이다. AWSRN (Adaptive Weighted Super Resolution Network)은 크게 비선형적 특징을 잘 잡아내기 위해 LFB (Local Fusion Block)을 쌓은 부분과 이미지 재구성을 위한 AWMS (Adaptive Weighted Multi-Scale), 추가적인 합성곱 연산과 Pixel Shuffle 연산을 통해 LSC (Long Skip Connection) 연산을 이루어낸 부분으로 나눌 수 있다.

먼저, LFB는 다시 효율적으로 low-level의 feature들을 잘 뽑기 위한 AWRU (Adaptive Weighted Residual Unit)과 다양한 level에서 추출한 feature를 모으는 LRFU (Local Residual Fusion Unit)으로 나눌 수 있다. AWRU의 구조는 파라미터 수를 늘리지 않으면서 비선형적 특징을 잘 추출하기 위해 Inverted Residual Block 구조를 사용했다. AWRU의 입력 데이터와 Inverted Residual Block을 통해 나온 결과에 각각의 가중치를 곱해 더한다. LRFU에서는 AWRU를 순차적으로 연결해 각각의 AWRU에서 나온 결과를 채널 단위로 연결한 후, bottleneck 구조에 통과시키면서 LFB의 입력 데이터와 동일한 채널 수로 변경하여 마찬가지로 가중합한다. AWMS에서는 이렇게 이루어진

LFB 구조를 여러 개 지나면서 얻어진 비선형적 특징을 다양한 크기의 window와 합성곱한다. 합성곱하여 나온 결과에 각각 가중치를 곱한 후, Pixel Shuffle Layer를 통해 해상도를 키운 결과를 모두 합한다. 마지막으로 처음 입력 데이터를 합성곱하고 해상도를 키운 결과와 합한다.

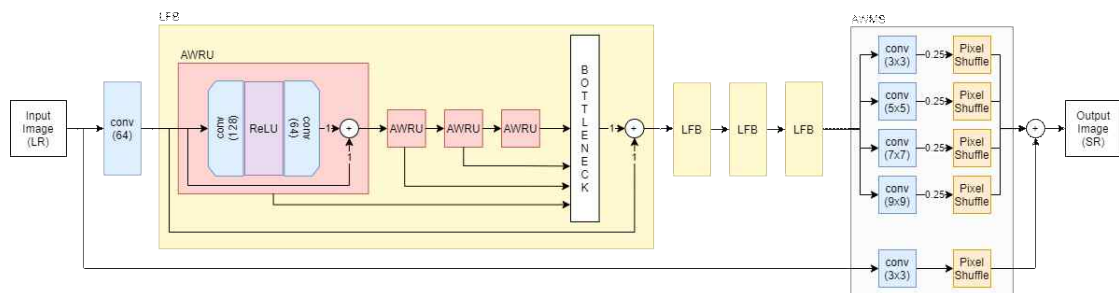
해당 신경망 구조에서 위상 홀로그램의 복잡한 특징을 더 잘 잡아내기 위해 AWRU의 Inverted Residual Block의 연산 결과 채널을 64개로 증가시켰다. AWRU에서 필요한 가중치는 모두 1로 설정했으며, 총 4개의 Unit을 사용했다. 마찬가지로 LFB에서 필요한 가중치는 모두 1로 설정했으며, 총 4개의 Block를 사용했다. AWMS에서는 3×3, 5×5, 7×7, 9×9 크기의 window를 사용해 합성곱했으며, 가중치는 모두 0.25를 사용했다.

3. 순환 손실 함수

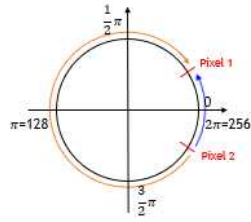
기존 딥러닝 기반 초해상도 신경망 모델에 대한 연구가 많아지면서 그 성능이 좋아지고 있다. 초해상도 신경망 모델의 결과로 예측한 고해상도 이미지와 원본 이미지 간의 차이를 최소화하기 위해 주로 MSE 손실 함수를 사용했으며, 최근에는 예측 이미지와 원본 이미지의 픽셀 간 유사도를 계산하는 L1 손실 함수를 자주 사용한다. 하지만 이런 손실 함수는 위상 홀로그램의 특징인 주기성을 고려하지 않아 위상 홀로그램을 위한 초해상도 신경망 모델의 손실 함수로 적절하지 않다.

어떤 위상 홀로그램의 예측 픽셀값을 Pixel 1, 원본 고해상도 위상 홀로그램의 픽셀값을 Pixel 2라고 하면, 두 픽셀의 오차는 Figure 3과 같이 주황색처럼 계산할 수 있지만, 파란색처럼 계산할 수도 있다. 이를 바탕으로 신경망에서 위상 홀로그램의 주기성을 고려해 2가지 오차에 대해 계산할 수 있게 하여 더 작은 쪽을 선택하게 하면 더 효율적인 계산을 할 수 있을 것으로 예상된다.

bit level에 의한 픽셀값의 차이를 없애기 위해 0과 1 사이로 정규화



<Figure 2> AWSRN architecture



<Figure 3> Errors caused by periodicity of phase

하고 L1 기반 순환 손실 함수(L1-circular)를 수식으로 표현하면 다음과 같다.

$$l_n = \begin{cases} |y_{pred} - y_{true}| & (0 \leq |y_{pred} - y_{true}| < 0.5) \\ 1 - |y_{pred} - y_{true}| & (0.5 \leq |y_{pred} - y_{true}| \leq 1) \end{cases}$$

$$Loss = \frac{\sum_{n=1}^{W*H*B*C} l_n}{W*H*B*C}$$

y_{pred} 는 초해상도 신경망 모델을 통해 고해상도 이미지의 픽셀값을 예측한 결과이고, y_{true} 는 원본 고해상도 이미지의 픽셀값을 의미한다.

W 와 H 는 고해상도 이미지의 너비와 높이를 의미하고, B 와 C 는 배치 사이즈의 크기와 RGB 채널의 개수를 의미한다. 둘 사이의 오차의 절댓값이 0.5보다 작은 경우, 일반적인 이미지에서 L1 손실 함수를 이용해 값을 구하는 것과 같은 방식으로 값을 구한다. 하지만 오차의 절댓값이 0.5보다 크거나 같고 1보다 작거나 같은 경우, 위상의 주기성을 고려하여 더 작은 오차값을 선택하기 위해 1에서 둘 사이의 오차의 절댓값을 빼 0.5보다 작은 값을 갖도록 한다. 각 픽셀 단위로 구한 l_n 의 평균값을 최종적으로 구한다.

4. 실험 결과

총 7개의 위상 홀로그램 시퀀스 중 마리오 시퀀스는 테스트 데이터셋으로 사용하였고, 나머지 6개의 위상 홀로그램 시퀀스를 이용하여 초해상도 신경망 모델을 학습시켰다. 각각의 시퀀스에서 3장의 프레임은 임의로 선택하여 검증 데이터 셋으로 사용하였고, 나머지 프레임들은 모두 훈련 데이터 셋으로 사용하였다. 실제 신경망 모델을 학습시키기 위해서 원본 이미지는 훈련 데이터 셋과 검증 데이터 셋 모두 각각의 프레임마다 120×120의 크기로 임의의 위치에서 100장씩 잘라내어 사용하였

고, 신경망 모델에 입력할 이미지는 원본 이미지에서 bicubic 보간법을 이용해 scale을 2배로 줄인 이미지를 사용하였다. 또한, 각 이미지 간 픽셀값의 편차를 줄이기 위해 0과 1 사이로 정규화하였으며, 복잡한 위상 홀로그램의 특징을 더 잘 잡아내기 위해, 잘라낸 100장 중 절반에 해당하는 50장을 수평 대칭 이동하여 데이터를 증강하였다.

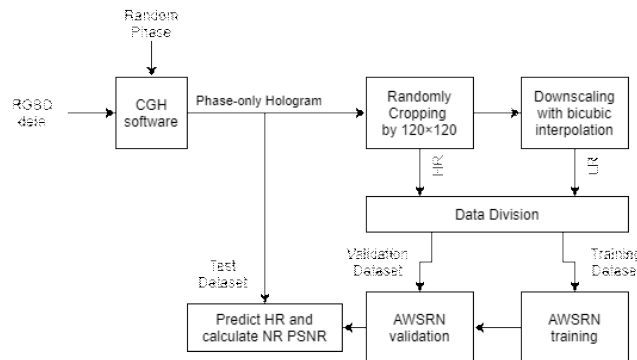
120×120 크기의 원본 이미지의 Scale을 2배로 줄여 신경망 모델에 입력으로 넣어 고해상도 홀로그램 영상을 얻기 위한 가중치를 최적화한다. 이렇게 훈련된 모델에 검증 데이터셋을 입력으로 넣어 올바른 방향으로 학습되고 있는지 확인한다. 마지막으로 학습에 사용하지 않은 테스트 데이터 셋을 입력으로 넣어 만들어낸 고해상도 위상 홀로그램 영상을 수치적으로 복원한 결과의 PSNR을 계산하여 비교한다.

학습에 사용된 데이터는 총 18,000개, 검증 데이터는 총 1,800개이며, 학습률(Learning Rate)는 0.001로 매 200 epoch마다 절반씩 감소하도록 하였다. 최적화 함수로는 Adam 알고리즘을 사용했으며, RAM 크기를 고려해 배치 사이즈를 16으로 사용하였다.

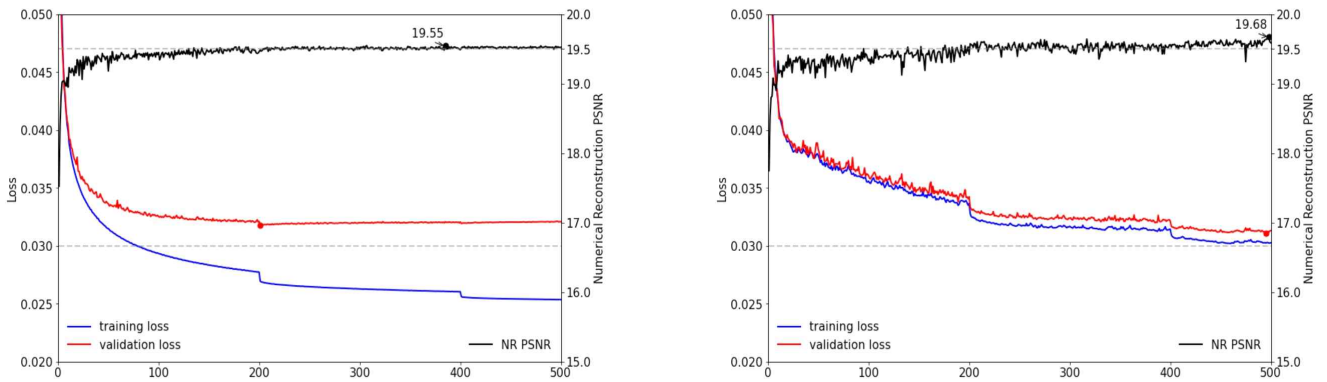
일반적으로 딥러닝을 이용한 초해상도 신경망 모델에서 주로 사용하는 L1 손실 함수와 새롭게 제안한 L1-circular 손실 함수를 비교해보았다. 둘 모두 같은 데이터 셋과 같은 학습 방법을 사용해 500 에폭까지 학습한 결과는 Figure 5와 같다. 왼쪽은 L1 손실 함수를 사용한 결과이고, 오른쪽은 L1-circular 손실 함수를 사용한 결과이다. 파란색 꺾은선 그래프는 training loss, 빨간색 꺾은선 그래프는 validation loss를 의미하며 왼쪽 y축의 scale을 따른다. 검은색 꺾은선 그래프는 각각의 모델에서 테스트 데이터셋을 수치 복원했을 때의 PSNR을 의미하며, 오른쪽 y축의 scale을 따른다. L1 손실 함수를 사용한 경우, 초반에는 빠르게 학습이 되지만 200 epoch 부근에서 과적합이 발생하는 모습을 확인할 수 있다. L1-circular 손실 함수를 사용하는 경우, L1 손실 함수에 비해 다소 느리게 학습되지만 계속 학습되는 경향을 보인다. 서로 다른 손실 함수를 사용해 학습한 신경망 모델로 고해상도 홀로그램 영상을 각각 만들어 수치 복원했을 때 PSNR 결과는 L1-circular를 사용했을 때 약 0.13db만큼의 차이를 보이며 더 좋은 성능을 보였다.

Reconstruction method	PSNR	SSIM
bicubic	16.1710	0.1225
ASWRN (L1 loss)	19.5502	0.2632
ASWRN (L1-circular loss)	19.6764	0.2733

<Table 1> PSNR, SSIM of numerical reconstruction image



<Figure 4> Block diagram of the experiment



<Figure 5> Result of training with L1 loss (Left) and L1-circular loss (Right)

5. 결론

본 논문에서는 위상 홀로그램의 주기성을 이용한 L1 기반의 순환 손실 함수, L1-circular Loss Function을 제안한다. 일반적인 2차원 이미지에서 예측 픽셀값과 원본 픽셀값 사이의 유사도를 이용한 L1 손실 함수와 같은 초해상도 신경망 모델에서 비교해봤을 때, L1-circular 손실 함수의 성능이 우수함을 보였다. 또한, L1 손실 함수는 빠르게 학습된다는 점과 L1-circular는 계속해서 학습된다는 장점을 이용하여 초반에는 L1 손실 함수로 초해상도 신경망 모델을 빠르게 학습시키고, 이후 L1-circular 손실 함수로 바꾸어 학습시킨 결과 성능 향상이 가능함을 보였다. 이에 앞으로 딥러닝 기반 위상 홀로그램 초해상도 신경망 모델 개발에 앞서 더 나은 성능의 결과를 얻을 수 있을 것이라 기대된다.

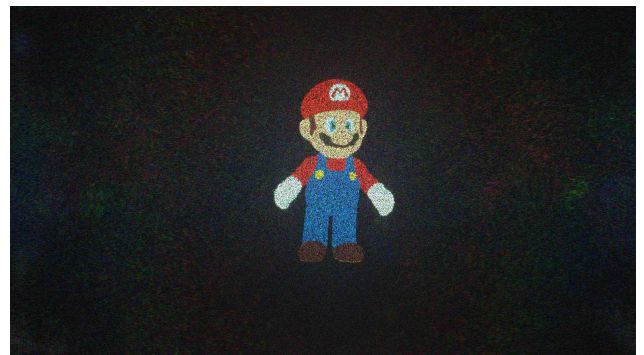
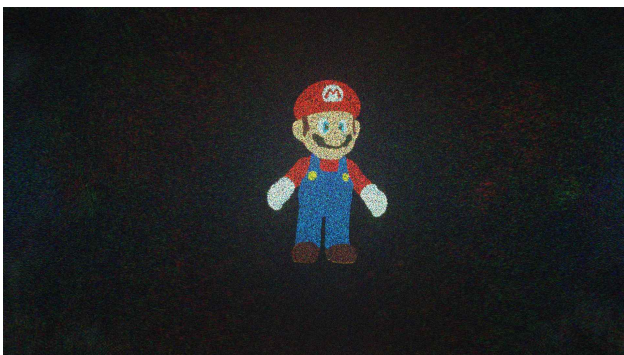
감사의 글

본 연구는 2020학년도 경희대학교 연구비 지원(KHU-20201116) 및 과학기술정보통신부 Giga Korea 사업의 디지털 홀로그래픽 테이بل 타입 단말 기술 개발 사업[GK20D0100]의 지원에 의한 결과입니다.

6. 참고문헌

[1] P. Hariharan, "Basics of Holography", Cambridge University Press, 2002.
 [2] U. Schnars, C.Falldorf, J. Watson, W.Juptner, "Digital Holography and Wavefront Sensing", Springer, 2005.

[3] C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, Jun 2015.
 [4] W. Sun, Z. Chen, "Learned Image Downsampling for Upscaling using Content Adaptive Resampler", IEEE Transactions on Image Processing, Feb 2020.
 [5] C. Ledig, et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4681-4690, 2017.
 [6] W. Kim, B. S. Park, J. K. Kim, K. J. Oh, J. W. Kim, D. W. Kim and Y. H. Seo, "Deep Learning-based Super Resolution for Phase-only Holograms", Journal of Broadcast Engineering, Volume 25 Issue 6, pp. 935-943, 2020.
 [7] K. J. Oh, J. Park, "Research and Standardization Trends of Digital Hologram Compression", Electronics and Telecommunications Trends, Volume 34 Issue 6, pp. 145-155, 2019.
 [8] Y. Zhang, Kungpeng Li, Kai Li, L. Wang, B. Zhong, Y. Fu, "Image Super-Resolution Using Very Deep Residual Channel Attention Networks", Proceedings of the European Conference on Computer Vision (ECCV), pp. 286-301, 2018.



<Figure 6> Numerical reconstruction image of training with L1 loss (Left) and L1-circular loss (Right)