

비지도 학습 기반 영상 노이즈 제거 기술을 위한 정규화 기법의 최적화

이강근[°] 정원기*
울산과학기술원 고려대학교
leekanggeun@unist.ac.kr wkjeong@korea.ac.kr

Optimized Normalization for Unsupervised Learning-based Image Denoising

Kanggeun Lee[°] Won-Ki Jeong*
UNIST Korea University

요약

최근 노이즈 제거를 위한 심층 학습 모델에 대한 연구가 활발하게 진행되고 있다. 특히 블라인드 노이즈 제거 (blind denoising) 기술이 발전하면서 깨끗한 영상을 얻기가 불가능한 영상의 영역에서 노이즈 영상만으로 심층 학습 기반 노이즈 제거 모델의 학습이 가능해졌다. 우리는 관찰된 노이즈 영상으로부터 깨끗한 영상을 얻기 위해 더는 깨끗한 영상과 노이즈 영상의 짝을 이루는 데이터를 필요하지 않는다. 하지만 노이즈 영상과 깨끗한 영상 간의 차이가 큰 데이터라면 노이즈 영상만으로 학습된 노이즈 제거 모델은 우리가 원하는 품질의 깨끗한 영상을 복원하기 어려울 것이다. 이 문제를 해결하기 위해서 짝지어지지 않는 깨끗한 영상과 노이즈 영상으로 학습한 모델 기반 노이즈 제거 기술은 최근 연구되고 있다. 가장 최신 기술인 ISCL은 깨끗한 영상과 노이즈 영상의 짝을 기반으로 한 지도학습 기반 모델의 성능과 거의 근접한 성능을 보여 주었다. 우리는 제안된 방법이 ISCL을 포함한 다른 최신 짝을 이루지 않는 영상 기반 노이즈 제거 기술보다 성능이 우수함을 보여준다.

Abstract

Recently, deep learning-based denoising approaches have been actively studied. In particular, with the advances of blind denoising techniques, it become possible to train a deep learning-based denoising model only with noisy images in an image domain where it is impossible to obtain a clean image. We no longer require pairs of a clean image and a noisy image to obtain a restored clean image from the observation. However, it is difficult to recover the target using a deep learning-based denoising model trained by only noisy images if the distribution of the noisy image is far from the distribution of the clean image. To address this limitation, unpaired image denoising approaches have recently been studied that can learn the denoising model from unpaired data of the noisy image and the clean image. ISCL showed comparable performance close to that of supervised learning-based models based on pairs of clean and noisy images. In this study, we propose suitable normalization techniques for each purpose of architectures (e.g., generator, discriminator, and extractor) of ISCL. We demonstrate that the proposed method outperforms state-of-the-art unpaired image denoising approaches including ISCL.

키워드: 정규화, 노이즈 제거, 비지도 학습, 심층 학습

Keywords: Normalization, Denoising, Unsupervised Learning, Deep Learning

1. 서론

노이즈는 다양한 환경적 요인들로 인해 디지털 영상에서 흔하게 관찰된다. 이러한 노이즈들은 우리가 디지털 영상을 통해 얻고자 하는 정보들을 분석하고 관찰하는데 방해요소가 될 수 있다.

예를 들어 의료진단 영역에서는 이러한 노이즈들이 정확한 진단을 방해하고 더 나아가 오진율을 높이는 요소가 된다. 이러한 문제들을 해결하기 위해 오래전부터 디지털 이미지에서 노이즈를 제거하는 기술은 계속해서 발전해왔다. 디지털 센서와 장비의 발

*corresponding author: Won-Ki Jeong/Korea University(wkjeong@korea.ac.kr)

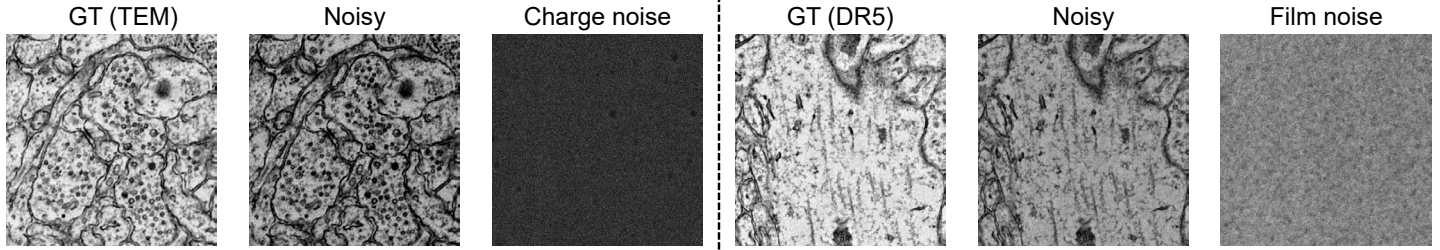


Figure 1: Examples of heterogeneous charge noise and film noise reported by Lee and Jeong [1]

전으로 노이즈가 감소했지만 특정 영역에서는 아직도 노이즈가 필연적인 요소로 남아있다. 뇌의 구조를 분석하고 원리를 이해하기 위해 연구가 되는 커넥톰릭스 (connectomics) 연구에서 가장 중요한 과정은 뇌의 구조를 디지털화하는 것이다. 뇌의 구조를 디지털화하기 위해서는 많은 연구가 진행되고 있는데, 최근 초고속 자동 이미징 [2, 3, 4] 장비와 기술이 발전하면서 병목현상을 유발하는 이미징 과정이 대폭 단축됐다. 하지만 Quan *et al.* [5]에서 소개된 것처럼 이러한 이미징 방법은 많은 노이즈가 수반된다. 초고속 자동 이미징 장비에서 발생하는 노이즈들은 여러 다른 종류들이 혼합된 노이즈이기 때문에 전통적인 노이즈 제거 방식으로는 제거가 어려운 문제점이 있다.

이러한 영상의 노이즈를 제거하기 위하여 우리가 관찰한 노이즈 영상 이외에 참조할 수 있는 깨끗한 영상이 존재한다면 그 영상을 활용할 수 있는 방법이 활발하게 연구되어 왔다. 최근 몇 년 동안 심층학습의 발전으로 지도학습의 노이즈 제거 [6, 7] 성능은 비약적으로 발전했다. 이러한 성능을 얻기 위해서는 지도학습은 깨끗한 영상과 노이즈가 섞인 두 쌍의 영상들에 의존한다. 초고속 자동 이미징 장비의 예시에서와같이 관찰된 노이즈가 섞인 영상과 짝을 이루는 깨끗한 영상은 얻기가 불가능하다. Figure 1에서 보이는 것처럼 우리는 기존에 알고 있는 가우시안 노이즈, 소금-후추 노이즈와 같은 잘 알려진 것과 다른 형태의 노이즈를 볼 수 있다. 이런 경우에는 노이즈의 수학적 통계를 알 수 없기 때문에 주어진 노이즈 이미지와 깨끗한 이미지의 분포만을 참조해야만 한다. 짝을 이루지 않는 (unpaired) 깨끗한 영상의 참조는 성능을 높이기 위한 필수적인 요소이다. 하지만 짝을 이루지 않는 노이즈 영상과 깨끗한 영상을 사용해 심층학습 모델을 학습하는 것은 어렵다. 두 영상이 짝을 이루지 않는 경우라면 심층학습모델을 학습하기 더욱더 까다로워진다. Quan *et al.* [5]에서는 관찰된 노이즈 영상에 대한 다른 짝이 존재하는 경우 사용할 수 있는 심층학습 기반 노이즈 제거 기술을 개발했다. 하지만 깨끗한 영상과 노이즈 영상의 쌍으로 학습시킨 지도학습 기반 심층학습 모델의 성능과 큰 격차가 존재했다.

최근 Lee and Jeong [1]은 자가협동 학습을 통해 짝을 이루지 않는 데이터를 이용한 노이즈 제거 모델의 성능을 크게 개선했고 지도학습 기반의 방법의 성능에 거의 근접한 성능을 보여줬

다. 자가협동 학습은 자가 적대 공격 (self-adversarial attack) [8]과 같은 문제점을 보여주는 CycleGAN [9]의 한계를 극복하고 더 나아가 커넥톰릭스에서 사용되는 초고속 자동 이미징 장비에서 발생하는 노이즈를 획기적으로 복원했다. 이 연구를 통해 우리는 좋은 품질의 노이즈 이미지의 복원 영상을 얻기 위해 더는 심층학습 모델을 깨끗한 이미지와 노이즈 이미지의 쌍의 데이터를 통해 학습시킬 필요가 없어졌다. 하지만 Lee and Jeong [1]이 제안한 상호의존적 자가 협동 학습 (ISCL)은 복잡한 다섯 개의 네트워크로 이루어져 있는데 이 네트워크의 정규화 방법에 대한 최적화가 이루어지지 않았다. 정규화 방법은 위 논문에서 언급된 것과 같이 성능에 큰 영향을 미치기 때문에, 이번 연구에서는 Lee and Jeong [1]이 제안한 상호의존적 자가 협동 학습 (ISCL) 방법에서 정규화 방법에 대한 논의와 실험을 통해 최적화된 정규화 방법을 찾으려고 한다.

2. 관련 연구

2.1 지도 학습 기반 노이즈 제거

영상의 사전 지식 (prior)을 사용한 패치 (patch) 기반 방법들은 심층 학습 기법의 발전 이전에 영상 노이즈 제거에서 뛰어난 성능을 보여줬다. 특히 BM3D [10]의 경우 가우시안 노이즈에 대해서 심층 학습 기법에 비견될 수 있는 수준을 보여준다. 하지만 최근에는 심층 학습의 계속된 발전으로 전통적인 영상의 사전 지식 기반 방법들의 성능을 뛰어넘었다. DnCNN [6]에서는 기존의 지도 학습 방법과는 다르게 수학적 노이즈 모델을 통해 만들어진 노이즈 영상을 깨끗한 영상에 더해 합성된 데이터를 만든다. 그렇게 되면 나머지 (residual) 영상과 노이즈 영상 쌍으로 지도 학습을 적용했을 때에 더 안정적인 수렴과 높은 성능을 보여줄 수 있음을 확인했다. 거기에 더해 배치 정규화 (batch normalization) [11]이 주어진 모델의 성능을 더욱 신장시킬 수 있음을 확인했다. CaGAN [7]에서는 생성적 적대 신경망 (generative adversarial network (GAN))을 사용해 DnCNN의 과도한 평활화 (smoothing)을 줄여주고 더욱 실제적인 질감을 살릴 수 있는 노이즈 제거 방법을 제시했다. 생성적 적대 신경망에 더해 집중 (attention) 신경망을

더해 성능을 증진했다.

2.2 비지도 학습 기반 노이즈 제거

Noise2Noise [12]는 비지도 학습 기반 노이즈 제거 방법으로 처음 소개가 됐다. 같은 영상에 대한 서로 다른 노이즈를 더해 쌍을 만들어 학습시키는 방법인데 이 방법은 노이즈 통계에 대한 사전 지식이 필요하다. 실제로 우리가 접하는 노이즈 영상들은 노이즈 통계에 대한 사전 지식이 없는 경우가 많아 실제로 적용이 어렵다. 이러한 문제를 해결하기 위해 자가감독 학습 기반 노이즈 제거 기술 [13, 14, 15, 16]들이 발전했다. 자가감독 학습 기반 방법들은 노이즈 영상만을 가지고 있을 때 노이즈 영상만으로 심층 인공 신경망을 학습시키는 방법이다. 노이즈 통계와 깨끗한 영상에 대한 조건 없이 노이즈 제거를 하는 방식을 블라인드 노이즈 제거 (blind denoising) 이라 부른다. 이 블라인드 노이즈 제거 방식의 문제점은 노이즈 영상에 의존하는 방식이기 때문에 얻고자 하는 깨끗한 영상이 노이즈가 있는 영상의 특성과 너무 다르게 되면 복원하기 어렵다는 단점이 ISCL [1]에서 보고되었다. 이를 해결하기 위해서 짝지어지지 않는 깨끗한 영상과 노이즈 영상을 사용해 심층 신경망을 학습시키는 기법들이 개발되었다. ADN [17]과 DRGAN [18]은 풀어진 신경망(disentanglement network)과 생성적 적대 신경망(GAN)을 사용해 짝지어 지지 않는 깨끗한 영상과 노이즈 영상을 사용한 노이즈 제거(unpaired image denoising) 성능을 높였다. 하지만 풀어진 신경망의 경우 많은 오토인코더(auto-encoder) 손실 함수의 계산으로 인해 계산 비용이 증가하게 되고 이로 인해 학습에 드는 시간이 길어진다. ISCL은 새로운 접근 방식을 제시했는데, 만약 관찰된 노이즈 영상과 얻고자 하는 깨끗한 영상이 일대일 대응(one-to-one correspondence)라면 노이즈를 의미하는 나머지 영상 또한 한 개 이상 존재한다라는 가정을 제시했다. 이를 통해 노이즈 제거 모델의 치역을 제약할 수 있는 노이즈 추출 인공 신경망을 모델에 추가했고 이를 자가나머지 학습(self-residual learning)이라고 부른다. ISCL은 기존의 방법들보다 뛰어난 노이즈 제거 성능을 보여줬고 더 나아가 모델의 크기는 다른 최신 방법들보다 훨씬 적었다. 이 방법에서는 배치-인스턴스 정규화 [19] 기법을 모든 모델에 적용했다.

2.3 인공신경망을 위한 정규화 기법

심층 인공신경망의 경우 신경망의 깊이가 복잡하고 길어질수록 그래디언트가 사라지거나(gradient vanishing) 발산하는(gradient exploding) 문제가 있었다. 이를 해결하기 위해서 배치 정규화 [11] 방법이 제시됐다. 배치 정규화의 경우 배치 단위로 특성의 통계를 정규화하므로 인스턴스 사이의 변화를 보존할 수 있지만, 인스턴스 정규화 [20] 방법은 인스턴스 사이의 변화를 정규화를 통해 제거하게 된다. 즉, 인스턴스 정규화 기법은 주어진 입력 특성의 질감이나 빛의 세기를 정규화를 통해 제거하게 된다. 이로 인해 정규화 기법마다 이미지의 모양 또는 질감의 특성을 강조하거나 무시하게 된다. 이 두 문제를 해결하기 위해서 배치-인스턴스

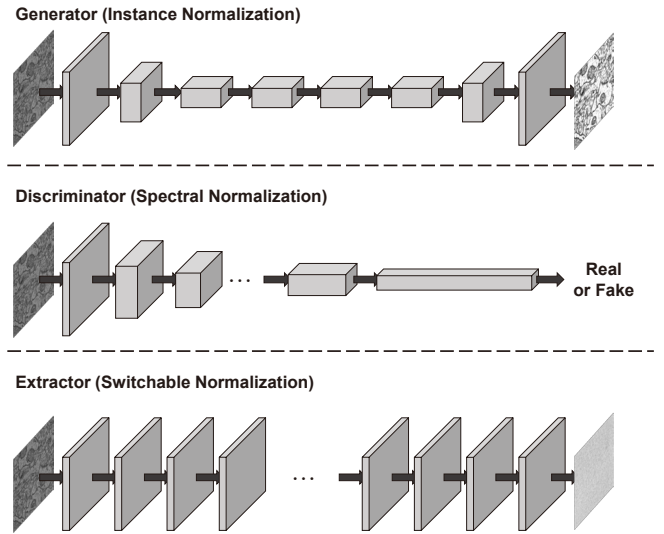


Figure 2: A modified architecture of ISCL [1] in which a suitable regularization technique is applied for the purpose of each model.

스 정규화 [19] 기법이 제안됐다. 배치-인스턴스 정규화 기법은 훈련 가능한 변수를 적용해 배치 정규화로 얻어진 특성과 인스턴스 정규화로 얻어진 특성을 결합해 손실함수를 최소화 할 수 있는 적절한 정규화 기법이 선택될 수 있도록 하는 정규화 기법을 제안했다. 해당 방법은 분류모델에서 기존 배치 정규화 기법보다 더 좋은 성능을 보여줬다. 전환 가능한 정규화 기법 (switchable normalization) [21]은 배치, 인스턴스 정규화 방법 그리고 더 나아가 층 (layer) 정규화 방법 [22]을 서로 전환 가능한 정규화 기법을 제시했다. 이 정규화 방법은 학습 과정을 통해 적절한 정규화 방법을 선택하기 때문에 적절한 정규화 기법을 찾지 않아도 된다. 하지만 추가적인 계산 비용과 훈련 가능한 파라미터의 증가는 불필요한 경우가 있을 수 있다. 예를 들어 층 정규화를 적용하면 오히려 학습에 방해가 되는 경우에는 전환 가능한 정규화 기법 대신에 배치-인스턴스 정규화 기법을 적용하는 것이 학습에 더 도움이 될 것이다. 스펙트럼 정규화(spectral normalization) [23]는 생성적 적대 신경망에서 판별자(discriminator)의 불안정한 학습을 보이는 문제점을 해결 할 수 있음을 보여줬다. 다른 정규화 기법과는 다르게 각 층의 출력에 대한 정규화가 아닌 각 층의 가중치에 대한 수렴 가능 공간을 제한하는 방법이다. 이 논문에서는 상호의존적 자가협동 학습에서 사용된 생산자, 구분자, 노이즈 추출을 위한 세 종류의 함수에 적절한 정규화 기법을 적용해 성능을 높이는 것이 목표이다.

3. 방법

3.1 상호의존적 자가협동 학습

상호의존적 자가협동 학습 [1]은 심층 학습 기반 노이즈 제거 기술들의 데이터 활용의 한계를 해결하기 위해 개발된 기술이다.

관찰된 노이즈 영상과 짝지어지지 않는 깨끗한 이미지의 통계에 대한 학습을 통해서 비지도 학습 기반인 최신 방법들의 성능을 뛰어넘었다. 이 방법은 크게 생산자, 판별자, 추출자 (extractor)로 이루어져 있다. 생산자가 학습할 때에 판별자와 적대적 학습을 진행하고 추출자에 의해 학습 가능한 잠재 공간이 제약된다. 판별자가 학습 할때에 생산자에 의해 만들어진 가짜(fake) 데이터뿐 아니라 추출자로 만들어진 가짜 데이터를 추가로 학습에 적용해 판별자의 판별성능을 높인다. 추출자는 생성자에 의해 만들어진 깨끗한 모조 영상을 통해 학습된다. 추출자는 생성자에 의해 만들어진 모조 영상에 의존하는 학습을 하므로 자가협동 학습이라 불린다. 이렇게 세 가지 모델은 상호의존적으로 반복 학습을 진행하게 되는데 각 모델은 정규화 층들을 가지고 있다. 모든 모델은 공통으로 배치-인스턴스 정규화를 사용해 학습의 안정성과 성능을 높인다. 본 연구에서는 Figure 2에서와 같이 생성자와 판별자, 그리고 추출자에 적합한 정규화 방법을 적용하고 성능을 높이는 방법을 제시한다.

3.2 생성자의 정규화 최적화

최근 정규화 기법은 특정 문제에 따라 적합하게 발전해왔다. 가장 먼저 개발된 정규화 기법 중에서 인스턴스 정규화는 다음 수식과 같이 배치를 정규화 한다. $x \in \mathbb{R}^{N \times H \times W \times C}$ 는 입력 특성이라고 가정한다. B, W, H, C 들은 각각 배치, 넓이, 높이, 채널의 크기를 나타낸다. 인스턴스 정규화의 경우 다음과 같다.

$$u_{nk} = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W x_{nijk} \quad (1)$$

$$\sigma_{nk}^2 = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W (x_{nijk} - u_{nk})^2 \quad (2)$$

$$x_{nijk}^{(I)} = \frac{x_{nijk} - u_{nk}}{\sqrt{\sigma_{nk}^2 + \epsilon}} \quad (3)$$

, 여기에서 n, i, j, k 는 각 차원의 원소에 대한 색인이다. 여기서 $\gamma, \beta \in \mathbb{R}^C$ 는 학습 가능한 어파인 변환(affine transformation) 변수를 적용해 다음 방정식과 같은 인스턴스 정규화를 사용한다.

$$x_{nijk}^{(I)} = \frac{x_{nijk} - u_{nk}}{\sqrt{\sigma_{nk}^2 + \epsilon}} \cdot \gamma + \beta \quad (4)$$

인스턴스 정규화의 장점은 배치 크기에 상관없이 동일한 결과를 도출한다. 합성곱 신경망에서 장점은 훈련 과정과 검증 단계에서 사용되는 영상의 크기가 동적이다. 따라서 테스트 노이즈 영상의 크기와 상관없이 훈련과정에서는 영상을 쪼개서 쌓는 방식으로 배치의 크기를 키울 수 있다. 미니배치(mini-batch)의 크기가 커질수록 성능이 향상되는 것에 대한 연구 [24]는 많이 보고되어 왔다. 이렇게 노이즈 제거 모델의 성능을 높게 되더라도 테스트 단계에서 노이즈가 있는 영상의 크기가 크다면 GPU 메모리의 한계 때문에 배치 크기를 과도하게 줄일 수밖에 없다. 이렇게

된다면 미니배치 안에 데이터가 적어 표본 간의 평균값과 변화량의 의미 있는 값을 얻을 수 없어 배치 정규화를 사용해 훈련한 모델의 경우 낮은 성능을 야기할 수 있다. 테스트 과정에서 배치 크기를 늘리기 위해 노이즈 영상을 쪼개 쌓게 된다면 나중에 다시 합치는 과정에서 패치의 경계면에서 노이즈가 발생할 수 있고 이를 해결하기 위해 여러 후처리가 진행되어야 한다. 후처리하게 된다면 실제 심층 학습 모델로 얻어진 결과에서 성능이 달라질 수 있다. 이런 부분을 해결하기 위해서 훈련과정 동안 표본 간의 평균값과 변화량을 저장해두는 방식을 사용하지만, 테스트 영상에 의존하는 값이 아닌 훈련데이터에 의존하는 값이기 때문에 성능 저하를 야기할 수 있다. 또한, 기존의 배치-인스턴스 정규화의 경우 두 종류의 정규화된 특성에 대해서 융합을 하기 위한 학습을 동시에 해야 하기 때문에 특성을 거의 보존해야 하는 노이즈 제거 방법과 다르게 모델의 복잡도가 올라가기 때문에 적합하지 않다. 더 나아가 모델의 복잡성으로 인해 생성자의 초기 학습속도가 느려지게 된다면 이후 판별자와 추출자의 성능에 영향을 미치기 때문에 학습의 복잡도가 낮고 배치 크기에 민감하지 않은 인스턴스 정규화를 적용했다.

3.3 판별자의 정규화 최적화

스펙트럼 정규화의 경우 판별자에 적용했을 때에 하이퍼변수(hyperparameter)의 의존도가 줄어든다고 보고됐다. 그에 더해 학습을 안정화 하는 WGAN [25] 또는 WGAN-GP [26]과 비교해서도 적대적 생성자 신경망의 성능이 좋았다. 판별자의 각 합성곱 층에 대해서 우리는 보고된 안정성을 부여하기 위해 다음 방정식과 같은 스펙트럼 정규화를 적용한다.

$$W_{SN}(W) := W/\sigma(W) \quad (5)$$

$$\text{where } \sigma(W) := \max_{h: h \neq 0} \frac{\|Ah\|_2}{\|h\|_2} = \max_{\|h\|_2 \leq 1} \|Ah\|_2 \quad (6)$$

여기서 $\sigma(W)$ 의 경우 각 층의 가중치 W 에 대한 스펙트럼 노름(spectral norm)이다. 방정식 (5)의 경우 각 층의 가중치를 Lipschitz norm으로 나눠줌으로써 Lipschitz 상수의 상계(upper bound)가 1이 되도록 유도한다. 스펙트럼 정규화의 경우 입력 특성에 대한 정규화가 아닌 각 합성곱 신경망의 층을 구성하는 가중치에 대한 제약조건이다.

3.4 추출자의 정규화 최적화

우리는 전환 가능한 정규화를 추출자에 대해서 적합한 정규화로 제안한다. 전환 가능한 정규화란 인스턴스, 배치, 층 정규화에 대한 융합을 최종 정규화로 제안하고 훈련 가능한 상수를 통해 융합의 정도를 학습한다. 가장 먼저 배치 정규화와 층 정규화는

Table 1: Specifications for our EM experiment cases.

Case	Noise-Free Images	Noise Types	Noisy Images (Scenario)
1	TEM_{ZB}	Charge	$TEM_{ZB} + \text{Charge}$ (Synthetic)
2	TEM_{DR5}	Film	$TEM_{DR5} + \text{Film}$ (Synthetic)
3	TEM_{ZB}	Charge	SEM_{ZB} (Real)
4	TEM_{DR5}	Film	TEM_{PPC} (Real)

Table 2: P.S.: Paired Supervision and U.S.: Unpaired Supervision. Quantitative PSNR and SSIM results on case 1 and 2. The first and second best results of unpaired supervision methods are highlighted in red and blue, respectively.

Type	Method	Charge noise		Film noise	
		PSNR	SSIM	PSNR	SSIM
P.S.	DnCNN [6]	28.27	0.9172	27.55	0.8964
	RED-CNN [27]	28.61	0.9230	28.02	0.9049
	CaGAN [7]	28.60	0.9186	28.03	0.9020
U.S.	Quan <i>et al.</i> [5]	22.32	0.8785	23.44	0.8288
	UIDNet [28]	23.11	0.8592	21.34	0.7826
	ADN [17]	25.67	0.8686	24.37	0.8535
	ISCL [1] + Batch Normalization [11]	24.62	0.8969	20.63	0.7568
	ISCL [1] + Layer Normalization [22]	24.61	0.8945	23.99	0.8829
	ISCL [1] + Group Normalization [29]	24.86	0.9016	24.74	0.8879
	ISCL [1] + Switchable Normalization [21]	25.12	0.9044	26.80	0.8948
	ISCL [1] + Batch-Instance Norm [19]	27.12	0.9054	27.06	0.8915
	Proposed	27.11	0.9076	27.86	0.8992

다음 방정식과 같다.

$$u_k = \frac{1}{NHW} \sum_{n=1}^N \sum_{i=1}^H \sum_{j=1}^W x_{nijk} \quad (7)$$

$$\sigma_k^2 = \frac{1}{NHW} \sum_{n=1}^N \sum_{i=1}^H \sum_{j=1}^W (x_{nijk} - u_k)^2 \quad (8)$$

$$x_{nijk}^{(B)} = \frac{x_{nijk} - u_k}{\sqrt{\sigma_k^2 + \epsilon}}$$

$$u_n = \frac{1}{HWC} \sum_{i=1}^H \sum_{j=1}^W \sum_{k=1}^C x_{nijk} \quad (10)$$

$$\sigma_n^2 = \frac{1}{HWC} \sum_{i=1}^H \sum_{j=1}^W \sum_{k=1}^C (x_{nijk} - u_n)^2 \quad (11)$$

$$x_{nijk}^{(L)} = \frac{x_{nijk} - u_n}{\sqrt{\sigma_n^2 + \epsilon}} \quad (12)$$

전환 가능한 정규화는 이전에 언급한 세가지 정규화 방법을 다음과 같이 결합한다

$$x_{nijk}^{(S)} = \frac{x_{nijk} - (w_b u_k + w_l u_n + w_i u_{nk})}{\sqrt{w_b' \sigma_k^2 + w_l' \sigma_n^2 + w_i' \sigma_{nk}^2 + \epsilon}} \cdot \gamma + \beta, \quad (13)$$

$$\text{where } w_b + w_l + w_i = 1 \text{ and } w_b' + w_l' + w_i' = 1. \quad (14)$$

(9) 여기서 $\gamma, \beta \in \mathbb{R}^C$ 는 학습 가능한 어파인 변환(affine transformation) 변수이다. 다음 방정식과 같이 훈련 가능한 변수 $w_b, w_b', w_l, w_l', w_i, w_i'$ 에 의해 배치 정규화된 특성, 층 정규화된 특성, 그리고 인스턴스 정규화된 특성이 결합한다. 전환 가능한 정규화는 모델의 복잡도가 높아지게 되지만 추출자의 경우 DnCNN [6]에서 보고된 것과 같이 학습의 복잡도가 낮기 때문에 모델의 복잡도가 높아지더라도 학습에 미치는 영향이 적다. 더 나아가 모델의 복잡도는 학습 이후 간단한 모델보다 더 좋은 성능을 보여주기 때문에 추출자에 적합한 정규화로 전환 가능한 정규화를 제안한다.

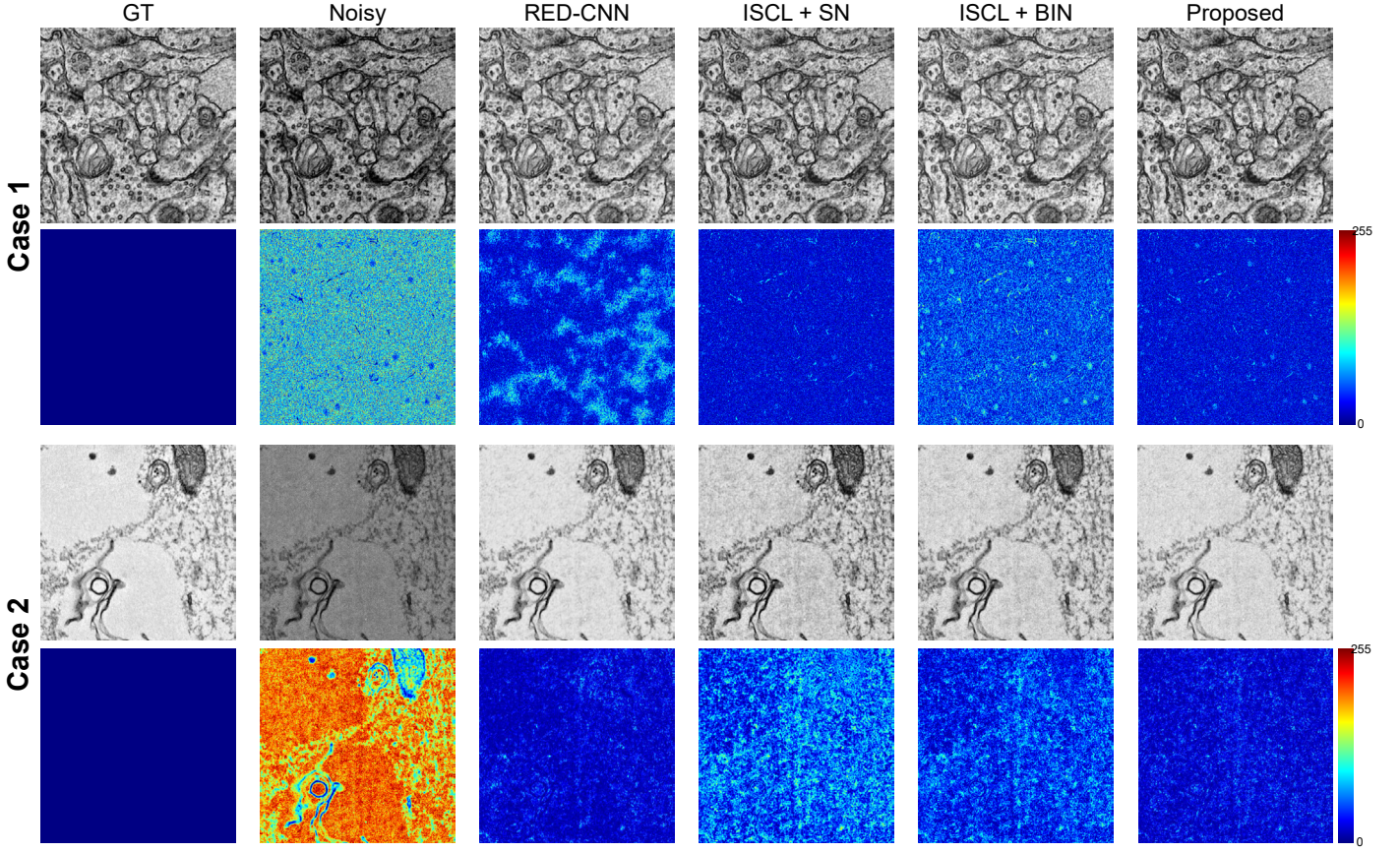


Figure 3: Qualitative results of RED-CNN [27], ISCL [1]+SN (switchable normalization) [21], ISCL [1]+BIN (batch-instance normalization) [19], and proposed method. First and third row indicate the result of case 1 (charge noise) and case 2 (film noise), respectively. Second and fourth rows are the heat maps of error compared to ground truth.

4. 결과

이번 실험을 통해 우리가 제안한 방법의 성능을 보여주기 위해서 우리는 기준인 ISCL의 정규화 방법을 최신 정규화 방법으로 대체하면서 실험을 진행했다. 기준점인 ISCL은 배치-인스턴스 정규화를 사용했고, 더 나아가 우리는 배치-인스턴스 정규화를 대신해 배치 정규화 [11], 층 정규화 [22], 그룹 정규화 [29], 전환 가능한 정규화 [21]를 사용해 비교했다. 우리는 다음 방정식과 같은 그룹 정규화 방법을 사용했다.

$$u_{ng} = \frac{1}{HW|\mathcal{G}_g|} \sum_{i=1}^H \sum_{j=1}^W \sum_{k \in \mathcal{G}_g} x_{nijsk} \quad (15)$$

$$\sigma_{ng}^2 = \frac{1}{HW|\mathcal{G}_g|} \sum_{i=1}^H \sum_{j=1}^W \sum_{k \in \mathcal{G}_g} (x_{nijsk} - u_{ng})^2 \quad (16)$$

$$x_{nijsk}^{(G)} = \frac{x_{nijsk} - u_{ng}}{\sqrt{\sigma_{ng}^2 + \epsilon}} \text{ and } k \in \mathcal{G}_g \quad (17)$$

g 는 그룹의 색인이고 \mathcal{G}_g 는 g 번째 그룹의 집합이다 (e.g., $\mathcal{G}_1 = \{1, 2, \dots, 32\}$, $\mathcal{G}_2 = \{33, 33, \dots, 64\}$). 본 연구에서 제안된 방법은

각 모델마다 생성자에 대해서는 인스턴스 정규화 [20], 판별자에 대해서는 스펙트럼 정규화 [23], 그리고 추출자에 대해서는 전환 가능한 정규화 [21]를 사용했다. 공정한 비교를 위해서 ISCL에서 사용한 모든 하이퍼변수들에 대해서 똑같은 환경과 조건으로 실험을 진행했다. 또한 Table 1에서와 같이 합성 데이터 실험인 case 1과 2의 경우 4번의 교차검증 역시 같은 표본들로 검증했다. 비교 방법은 지도 학습 기반인 DnCNN [6], RED-CNN [27], CaGAN [7]을 선택했다. 비지도 학습으로서는 본 연구의 기초가 되는 Quan *et al.* [5] ISCL [1]과 최신 짝지어지지 않는 영상 기반 노이즈 제거 방법인 UIDNet [28], ADN [17]을 선택했다.

4.1 실험 데이터

본 연구에서는 공평한 실험을 위해서 Quan *et al.* [5]과 Lee and Jeong [1]에서 사용한 데이터와 같은 데이터를 사용했다. Table 1에서와 같이 관찰 대상이 없는 영역에서 촬영한 노이즈 영상을 실제 깨끗한 관찰 영상에 합성해 케이스 1번과 2번 사례를 만들었다. 케이스 3번과 4번의 경우 실제 사례에서 노이즈만 관찰 가능한 영상을 사용해 실제 사례에서의 성능을 평가했다. 실제 사

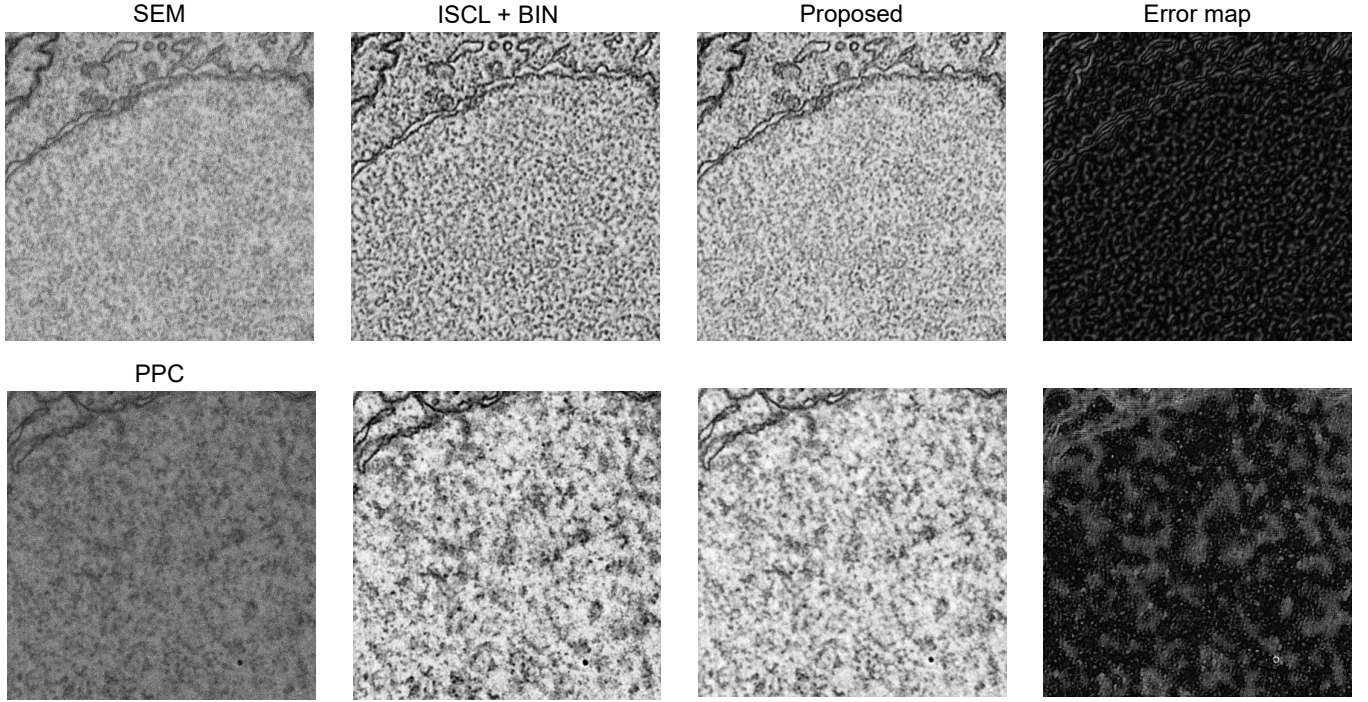


Figure 4: Qualitative assessment on real EM images corrupted by charge noise and film noise which are specified in Table 1. We compare the proposed method with ISCL+BIN which is the state-of-the-art unpaired image denoising method as a baseline. Error map represents the proportion of error between the baseline and the proposed method.

레에서는 관찰된 노이즈 영상에 해당되는 깨끗한 영상이 없기 때문에 시각적인 질적 평가만 가능하다.

4.2 결과

ISCL의 구현은 공개된 공식 코드를 사용했다. 모든 하이퍼변수는 기존 ISCL의 실험과 동일하게 세팅했다. 그룹 정규화의 경우 그룹의 크기 $|G_g|$ 는 32를 사용했고 채널의 수가 32개일 경우 층 정규화와 같아진다. Table 1에서 명시된 것과 같이 case 1과 2에 대한 교차검증의 결과는 Table 2에 요약되어 있다. Table 2에서 보이는 것과 같이 정규화 방법에 따라서 성능 차이가 크다. 3.2에서 설명한 것 처럼 배치 정규화의 경우 훈련 과정과 테스트 과정에서 크기가 다르기 때문에 훈련과정 동안 각 배치 정규화 층마다 평균과 변화량 값을 저장했다가 테스트 과정에서 사용한다. 하지만 이 값들은 테스트 영상의 값에 따라 변화하는 값이 아닌 훈련 데이터에 의존하는 값이기 때문에 성능에 저하가 발생한다. Table 2에서 처럼 배치 정규화로 대체했을 때에 성능이 크게 저하되는 것을 볼 수 있다. 더 나아가 배치 정규화는 인스턴스 정규화의 장점인 질감의 변화량을 제거하지 못해 판별자의 성능을 저하할 수 있다. 짝을 이루지 않는 영상 노이즈 제거에서 배치 정규화보다 배치-인스턴스 정규화가 더 좋은 성능을 보여줬다. 또한 층 정규화와 그룹 정규화의 경우 배치 정규화보다 더 좋은 성능을 보여줬다. 배치-인스턴스 정규화는 다른 정규화에 비해 좋은 성능을 보여줬고 필름 노이즈 (film noise)의 경우에는 전환 가능한 정규화와 비슷하거나 더 좋은 성능을 보여줬다. 생성자,

판별자, 추출자에 대해서 모두 같은 정규화를 사용할 때에는 배치-인스턴스 정규화가 가장 좋은 성능을 보여줬지만 제안된 방법 처럼 각 모델의 특성에 적합하게 적용했을 때에는 가장 좋은 영상 품질을 보여줬다. PSNR의 경우 필름 노이즈 실험에서 0.8dB만큼 성능이 향상됐다. 특히 주목할 점은 지도학습인 DnCNN과의 비교에서 필름 노이즈에서 더 좋은 품질의 영상을 얻을 수 있었다. 더 나아가 지도학습 기반의 최신 방법인 CaGAN과 비교해서도 PSNR 0.17dB 정도 차이가 난다. SSIM의 경우 0.0028정도 차이가 나는 것을 봤을 때 필름 노이즈에서는 지도학습과 거의 근접한 성능을 얻을 수 있음을 확인했다. Figure 3에서와 같이 제안된 방법은 생산자, 판별자, 추출자 모두 같은 정규화 방법을 사용하는 방법인 두 방법 (i.e., ISCL+BIN, ISCL+SN)과 비교해서 더 좋은 품질의 결과를 보여줬다. Figure 3의 두 번째와 네 번째 행에서 제안된 방법이 실측(ground truth)과 비교해 더 적은 오차를 보여준다. 특히 RED-CNN의 경우 전하 노이즈에 대해서 많은 노이즈가 존재한다. Figure 3에서 제안된 방법이 시각적 평가에서도 기존의 방법보다 충분한 개선점을 확인했다. Figure 4는 실측이 존재하는 노이즈 영상이 아닌 실제 사례에서의 전하 노이즈 (i.e., case 3)와 필름노이즈 (i.e., case 4)의 결과이다. 기준이 되는 방법 (ISCL+BIN)과 제안된 방법 모두 노이즈 영상에서 경계면과 입자들이 노이즈 영상보다 명확해졌다. 하지만 기존 ISCL은 노이즈 영상에서 보이는 입자들이 과도하게 강조되는 경향이 있다. 또한 노이즈 영상에서 관측되는 얼룩진 부분들이 강조되어 나타난다. 오차 지도 (error map)에서 얼룩진 부분에서 차이가 나는 것을 볼

수 있다.

5. 결론

본 연구에서는 최신 심층 학습 기반 노이즈 제거 방법들의 문제점인 짝을 이루는 깨끗한 영상과 노이즈 영상의 의존성을 해결하는 ISCL 방법을 개선했다. ISCL에서 보여준 성능을 개선하기 위해서 우리는 생성자, 판별자, 추출자에 따른 각각의 모델의 특성에 적합한 정규화 기법을 통해 최적의 모델을 찾아냈다. 공정한 실험을 통해서 우리가 제시한 방법이 기존의 배치-인스턴스 정규화를 사용한 것보다 더 좋은 성능을 보임을 확인했다. 더 나아가 본 연구에서는 정규화 방법이 모델 성능에 미치는 영향을 실험 결과를 통해 간접적으로 확인할 수 있었다. 또한 실제로 깨끗한 데이터가 존재하지 않는 사례에서 노이즈 영상을 복원함으로써 실제 사례에서도 적용 가능함을 보여줬다. 하지만 기존 방법과 비교해서 시각적 평가의 개선점에 대한 논의가 더 필요하다. 이후 연구에서는 복원된 영상의 차이와 활용 면에서의 차이에 대해 연구를 할 계획이다.

감사의 글

본 연구는 과학기술정보통신부 재원의 정보통신기획평가원의 ICT명품인재양성 사업 (IITP-2021-2020-0-01819), 교육부 재원의 한국연구재단의 기초연구사업 (No. NRF-2021R1A6A1A13044830), 그리고 과학기술정보통신부 재원의 한국연구재단의 초융합AI원천기술개발사업 (NRF-2019M3E5D2A01063819)의 지원을 받아 수행되었음.

References

- [1] K. Lee and W.-K. Jeong, "ISCL: Interdependent self-cooperative learning for unpaired image denoising," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3238–3248, 2021. [Online]. Available: <https://doi.org/10.1109/TMI.2021.3096142>
- [2] D. G. C. Hildebrand, B. J. Graham, and W.-C. A. Lee, "Grid-tape for fast nanoscale imaging," 2017.
- [3] B. J. Graham, D. G. C. Hildebrand, and W.-C. A. Lee, "Grid-tape imaging stage," 2018.
- [4] B. J. Graham, D. G. C. Hildebrand, A. T. Kuan, J. T. Maniates-Selvin, L. A. Thomas, B. L. Shanny, and W.-C. A. Lee, "High-throughput transmission electron microscopy with automated serial sectioning," *bioRxiv*, 2019. [Online]. Available: <https://www.biorxiv.org/content/early/2019/06/02/657346>
- [5] T. Minh Quan, D. Grant Colburn Hildebrand, K. Lee, L. A. Thomas, A. T. Kuan, W.-C. Allen Lee, and W.-K. Jeong, "Removing imaging artifacts in electron microscopy using an asymmetrically cyclic adversarial network without paired training data," in *Proc. IEEE Int. Conf. Comput. Vis. Workshop*, Oct. 2019.
- [6] 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.
- [7] Z. Huang, Z. Chen, Q. Zhang, G. Quan, M. Ji, C. Zhang, Y. Yang, X. Liu, D. Liang, H. Zheng, *et al.*, "Cagan: a cycle-consistent generative adversarial network with attention for low-dose ct imaging," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1203–1218, 2020.
- [8] D. Bashkurova, B. Usman, and K. Saenko, "Adversarial self-defense for cycle-consistent gans," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019.
- [9] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2223–2232.
- [10] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [11] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," ser. *Proc. Int. Conf. Mach. Learn.*, F. Bach and D. Blei, Eds., vol. 37, Lille, France, Jul. 2015, pp. 448–456.
- [12] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2Noise: Learning image restoration without clean data," in *Proc. Int. Conf. Mach. Learn.*, J. Dy and A. Krause, Eds., vol. 80, Jul. 2018, pp. 2965–2974.
- [13] K. Lee and W.-K. Jeong, "Noise2kernel: Adaptive self-supervised blind denoising using a dilated convolutional kernel architecture," 2020. [Online]. Available: <https://arxiv.org/abs/2012.03623>
- [14] A. Krull, T.-O. Buchholz, and F. Jug, "Noise2void-learning denoising from single noisy images," in *Proc. IEEE conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 2129–2137.

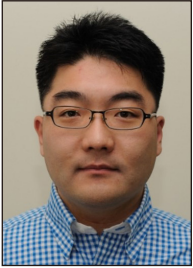
- [15] J. Batson and L. Royer, “Noise2Self: Blind denoising by self-supervision,” in *Proc. Int. Conf. Mach. Learn.*, Mar. 2019, pp. 524–533.
- [16] Y. Quan, M. Chen, T. Pang, and H. Ji, “Self2Self with dropout: Learning self-supervised denoising from single image,” in *Proc. IEEE conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 1890–1898.
- [17] H. Liao, W.-A. Lin, S. K. Zhou, and J. Luo, “Adn: Artifact disentanglement network for unsupervised metal artifact reduction,” *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 634–643, 2019.
- [18] Y. Huang, W. Xia, Z. Lu, Y. Liu, H. Chen, J. Zhou, L. Fang, and Y. Zhang, “Noise-powered disentangled representation for unsupervised speckle reduction of optical coherence tomography images,” *IEEE Trans. Med. Imag.*, 2020.
- [19] H. Nam and H.-E. Kim, “Batch-instance normalization for adaptively style-invariant neural networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018, pp. 2558–2567.
- [20] D. Ulyanov, A. Vedaldi, and V. Lempitsky, “Instance normalization: The missing ingredient for fast stylization,” 2016. [Online]. Available: <https://arxiv.org/abs/1607.08022>
- [21] P. Luo, R. Zhang, J. Ren, Z. Peng, and J. Li, “Switchable normalization for learning-to-normalize deep representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 2, pp. 712–728, 2019.
- [22] J. L. Ba, J. R. Kiros, and G. E. Hinton, “Layer normalization,” 2016. [Online]. Available: <https://arxiv.org/abs/1607.06450>
- [23] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, “Spectral normalization for generative adversarial networks,” in *Proc. Int. Conf. Learn. Represent.*, 2018.
- [24] A. Brock, J. Donahue, and K. Simonyan, “Large scale gan training for high fidelity natural image synthesis,” in *Proc. Int. Conf. Learn. Represent.*, 2018.
- [25] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein generative adversarial networks,” in *Proc. Int. Conf. Mach. Learn.*, vol. 70, Aug. 2017, pp. 214–223.
- [26] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, “Improved training of wasserstein gans,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5767–5777.
- [27] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, “Low-dose ct with a residual encoder-decoder convolutional neural network,” *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, 2017.
- [28] Z. Hong, X. Fan, T. Jiang, and J. Feng, “End-to-end unpaired image denoising with conditional adversarial networks,” in *Proc. Assoc. Adv. Artific. Intell.*, Apr. 2020, pp. 4140–4149.
- [29] Y. Wu and K. He, “Group normalization,” in *Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.

〈 저 자 소 개 〉



이 강 근

- 2017 : UNIST 컴퓨터공학과 학사
- 2017~현재 : UNIST 컴퓨터공학부 석박
통합과정
- 관심분야 : 영상 처리, 기계 학습, 컴퓨터 비전
- <https://orcid.org/0000-0002-0800-2129>



정 원 기

- ~2008 : University of Utah 전산학 박사
- 2008~2011 : Center for Brain Science at
Harvard University 연구원
- 2011~2020 : UNIST 컴퓨터공학과 조교수,
부교수
- 2020~현재 : 고려대학교 컴퓨터공학과 정교수
- 관심 분야 : 시각화, 영상 처리, 병렬 컴퓨팅
- <https://orcid.org/0000-0002-9393-6451>