

이미지 복원을 위한 네트워크 파라미터의 동적 업데이트를 위한 기법

□ 김태현 / 한양대학교

요약

최근 많은 연구 결과물에서 빅데이터를 이용하여 학습된 뉴럴 네트워크가 영상 내 노이즈를 제거하는데 매우 효과적임이 입증되었다. 여기에서 한 걸음 더 나아가, 입력으로 주어진 노이즈가 있는 영상의 특징을 분석하여, 사전에 학습된 네트워크의 파라미터를 테스트 타임에 동적으로 업데이트함으로써 주어진 입력 영상을 더욱 잘 처리할 수 있도록 하는 연구들이 시도되고 있다. 본 원고에서는 이와 같이 테스트 타임에 주어지는 입력 영상을 네트워크 학습에 사용하는 (self-supervision) 이미지 복원 기법들을 소개한다. 다음으로, 기존의 self-supervision을 이용하는 기법들 대비 학습 효율성과 정확도를 더욱 향상시킬 수 있는 새로운 형태의 네트워크 파라미터 업데이트 기법을 설명하고, 제안하는 기법의 우수성을 다양한 실험 결과를 통해 분석 및 입증한다.

I. 서론

디지털 카메라를 이용하여 촬영한 사진에는 여러

가지 이유들로 인해 화질 열화가 발생하고, 이런 대표적인 화질 열화 요소 중의 하나가 바로 노이즈이다. 특히, 광량이 부족한 저조도 환경에서는 카메라의 ISO를 높게 설정하게 되고, 이로 인해 노이즈가 큰 폭으로 증폭이 된다. 일반적으로 이와 같은 노이즈를 분석하기 위해서 많은 경우에 Additive White Gaussian Noise(AWGN) 모델을 이용하여 가정하고, 노이즈가 없는 깨끗한 원본 신호 X 와 노이즈 n 을 이용하여 노이즈가 있는 관측 신호 Y 를 다음과 같이 표현한다.

$$Y = X + n. \quad (1)$$

특히, 식 (1)은 Y 만 주어진 상태에서 노이즈 n 을 추정하고, 이와 동시에 원본 신호 X 를 찾아야 하는 전형적인 ill-posed 형태의 문제로 다루기가 쉽지 않다. 따라서 지난 수십 년간 영상 내 노이즈를 제

거하기 위한 많은 연구가 수행되었다[1-3].

근래에는 뉴럴 네트워크를 이용한 연구가 많이 수행되고 있다. 노이즈가 없는 선명한 영상과, 여기에 노이즈를 더한 흐搠된 영상을 쌍으로 한 학습 데이터셋을 이용하여 뉴럴 네트워크를 학습하고, 학습이 완료된 네트워크는 노이즈가 있는 영상을 입력으로 받아 노이즈가 제거된 영상을 출력하게 된다. 이와 같이 지도 학습(supervised learning)에 기반한 기법들이 모델링 및 최적화에 기반한 전통적인 방식들에 비해 훨씬 더 우수한 성능을 보여주는 것으로 많은 연구에서 입증이 되고 있다[8,9].

그러나 이런 지도 방식으로 학습된 뉴럴 네트워크들은 학습에 사용한 노이즈의 분포와 실제 테스트 영상의 노이즈 분포가 다른 경우에는 잘 동작하지 않는다. 따라서 최근에는 학습된 네트워크 파라미터를 이용하여 입력 영상의 노이즈를 곧장 제거하는 기존 방식에서 벗어나, 입력 영상의 특성을 분석하고, 이를 이용하여 네트워크의 파라미터를 업데이트 한 후, 업데이트 된 네트워크를 이용하여 비로서 노이즈를 제거하는 자기 지도 학습(self-supervised learning)에 기반한 기법들이 소개되고 있다[4-7].

구체적으로 Lehtinen 등이 제안한 noise-to-noise[6] 학습 기법은 입력 영상 내 존재하는 랜덤 노이즈의 기대값이 $0(E[n]=0)$ 이면서, 독립항등분포(independent and identically distributed)를 따르는 경우, 깨끗한 영상 없이 노이즈가 있는 영상 쌍들을 이용하여 네트워크를 학습할 수 있다는 사실을 입증하였고, 이를 바탕으로 Krull 등의 noise-to-self[5]와 Batson 등의 noise-to-void[7]와 같은 후속 연구들이 수행되었다. 특히, noise-to-self[5]와 noise-to-void[7] 학습 기법은 noise-

to-noise[6] 기법과는 다르게 테스트 타임에 주어진 입력 영상을 이용하여 네트워크 파라미터를 업데이트 하고, 업데이트 된 네트워크 파라미터를 이용하여 입력 영상의 노이즈를 제거할 수 있다. 특히, 입력 영상의 노이즈 분포가 학습 과정 동안 사용했던 노이즈의 분포와 다른 경우에도 입력 영상의 노이즈를 제거할 수 있다는 큰 장점이 있다. 그러나 noise-to-self와 noise-to-void에서 제안된 네트워크 파라미터를 동적으로 업데이트 하는 메커니즘이 매우 비효율적이고, 또한, 가우시안 노이즈와 같이 알려진 분포를 따르는 노이즈를 제거하는 경우 사전에 학습된 네트워크 파라미터보다 업데이트 된 파라미터의 디노이징 성능이 더 나쁘게 되는 문제점도 있다. <표 1>에는 가우시안 노이즈를 제거하도록 학습된 DnCNN[8] 네트워크의 파라미터를 noise-to-void[7] 기법을 이용하여 업데이트 하고, 업데이트 전후의 디노이징 성능을 비교하였고, 파라미터 업데이트 후에 디노이징 성능이 더 저하가 되는 문제를 확인할 수 있다.

<표 1> DnCNN 8 의 파라미터 업데이트 전후 디노이징 성능의 비교. 네트워크 파라미터 업데이트를 위해 Noise-to-void[7] 기법을 사용함. Urban100 데이터셋 영상들이 입력 영상으로 사용되었고 다양한 노이즈 레벨($\sigma=10$, $\sigma=40$)의 가우시안 노이즈를 더하여 노이즈가 있는 영상을 생성함.

	DnCNN[8]	DnCNN[8]+noise-to-void[7]
PSNR (dB) (가우시안 노이즈 레벨: $\sigma=10$)	35.85	35.72
PSNR (dB) (가우시안 노이즈 레벨: $\sigma=40$)	29.07	29.03

따라서 본 원고에서는, 입력 영상에 잘 알려진 분포를 따르는 노이즈가 첨가되어 있는 경우에도 자

기지도 학습 기법을 활용한 네트워크 파라미터의 동적 업데이트를 통해 디노이징 성능을 더욱 향상시킬 수 있는 효과적이고 새로운 기법을 설명한다.

II. 이미지 복원을 위한 네트워크 파라미터의 동적 업데이트 기법의 개요

일상에서 접할 수 있는 일반적인 영상들은 유사한 패치들이 반복해서 나타나는 패치 리커런스(patch-recurrence, self-similarity)[10-12]라는 특성을 내포하고 있다. 특히, 이러한 패치 리커런스 정보를 활용하면, 입력 영상에 첨가된 랜덤 노이즈 n 을 제거할 수 있다. 예를 들면, 노이즈는 서로 다르지만 모양이 같은 패치들의 평균값을 취하게 되면, L2 관점에서 최적으로 디노이징 된 패치를 얻을 수 있게 된다.

그러나 기존의 방법들은 서로 유사한 패치들의 정보를 이용하기 위해, 단일 영상 내에서 모양이 유사한 패치들을 찾아야 하고, 이를 위해 논로컬 모듈(non-local module) 같이 계산량이 많은 네트워크 모듈을 활용해야 한다[13,14]. 즉, 기존의 지도 학습 기반 영상 복원 기법들은 영상 내 패치 리커런스 정보를 활용하는데 아키텍처 설계상 제약이 있고, 같은 모양의 두 패치가 멀리 떨어져 위치하는 경우에는 두 패치 내 중복되는 정보를 활용하기 어려운 문제점이 있다. 그러나 자기지도 방식에 기반한 학습 기법들은 패치 리커런스 정보를 매우 손쉽게 이미지 복원에 활용을 할 수 있다.

1. Noise-to-void 및 noise-to-self 학습 기법

Noise-to-void[7] 및 noise-to-self[5] 학습 기법들은 자기지도 방식으로 테스트 타임에 네트워크 파라미터를 업데이트 할 수 있다. 구체적으로 테스트 타임에 학습을 수행하기 위한 손실 함수(loss function)는 다음과 같이 정의된다.

$$\text{Loss}(\theta) = \sum_i (\mathbf{f}_\theta(\mathbf{B}(\mathbf{y}_i)) - \mathbf{y}_i)^2. \quad (2)$$

식 (2)에서 θ 는 네트워크의 파라미터를 의미하고, \mathbf{f}_θ 는 파라미터가 θ 인 이미지 복원 네트워크, \mathbf{y}_i 는 픽셀 좌표 i 를 중심으로 위치한 패치를 의미한다. 특히, noise-to-void 및 noise-to-self 학습 기법에서는 블라인드 함수 \mathbf{B} 의 역할이 매우 중요한데, 블라인드 함수 \mathbf{B} 는 이름과 같이 입력 패치 \mathbf{y}_i 의 가운데 픽셀값을 네트워크에 보여주지 않는 역할을 한다. 이를 위해 실제로는 픽셀 i 주변의 인접 픽셀값들 중 하나를 랜덤으로 선택하여 채우는 방식을 사용하고, 이렇게 함으로써 네트워크가 입력 영상을 단순히 출력으로 바이패스하지 않고, 디노이징 능력을 갖출 수 있도록 하는 역할을 한다.

이 때, 픽셀 i 의 참값(ground-truth)이 주변 픽셀값들의 기대값과 같고, 또한, 주어진 입력 영상으로부터 정확하게 얼라인이 된 M 개의 패치 $\{\mathbf{y}_1, \dots, \mathbf{y}_M\}$ 가 주어지는 경우, 손실 함수 (2)를 최적화 함으로써 최적 파라미터 $\theta^\#$ 을 계산할 수 있다. 최적 파라미터 $\theta^\#$ 을 이용하면 입력 패치들의 노이즈를 제거할 수 있고 그 때 노이즈가 제거된 패치는 입력 패치들의 평균값 $\frac{1}{M} \sum_{i=1}^M \mathbf{y}_m$ 이 된다. 또한, 원본 영상 내 노이즈의 분산이 σ^2 인 경우 노이즈가 제거된 패치 내 노이즈 분산은 $\frac{1}{M} \sigma^2$ 로 줄어든다. 따

라서, 식 (2)를 최소화함으로써 얻은 파라미터를 이용하여 입력 패치의 노이즈를 제거할 수 있게 된다.

특히, 주어진 영상 복원 네트워크가 전부 컨벌루션 레이어로 구성이 된 경우(fully convolutional)에는 네트워크의 translation-invariant한 특성으로 인해서, 논로컬 모듈을 이용하여 유사한 패치들을 직접적으로 찾는 과정이 없어도, 학습 과정 동안 유사한 패치들의 통계치가 파라미터 업데이트에 자동으로 반영이 된다. 그러나 기존의 자기 지도 방식에 기반한 방법들은 블라인드 함수 B로 인해 학습 과정이 매우 비효율적이다. 실제적으로, 식 (2)를 최적화 하기 위해서는 네트워크의 용적률(receptive field)보다 더 큰 입력 패치가 필요하지만, 실제로는 가리어진 가운데 픽셀에 대해서만 미분이 수행되기 때문에 학습 속도가 매우 느리고 비효율적이게 된다.

2. Restore-from-restored 학습 기법

기존의 자기 지도 학습법의 비효율성을 극복하기 위해 다음과 같은 새로운 손실 함수를 정의할 수 있다.

$$\text{Loss}(\theta) = \sum_{j=1}^N \left(\mathbf{f}_\theta(\mathbf{f}_{\theta_0}(\mathbf{Y}) + \mathbf{n}_j) - \mathbf{f}_{\theta_0}(\mathbf{Y}) \right)^2. \quad (3)$$

식 (3)에서 θ_0 는 특정 노이즈 분포(예: 가우시안)에 대해 사전에 충분히 학습된(fully pre-trained) 네트워크 파라미터를 의미하고, \mathbf{Y} 는 노이즈가 있는 입력 영상, \mathbf{n}_j 는 입력 영상의 노이즈와 같은 분포를 따르는 새로운 노이즈를 의미한다. 즉, 입력 영상의 노이즈가 분산이 σ^2 인 가우시안 랜덤 노이즈인 경우, \mathbf{n}_j 도 같은 분포를 따르는 랜덤 노이즈가 된다.

일반적으로 자기 지도 학습에 기반하는 기법들은 테스트 타임에 학습을 위한 참값(ground-truth)이

주어지지 않는다. 마찬가지로, 식 (3)에서도 참값을 사용할 수 없기 때문에, 사전에 학습된 네트워크의 출력값 $\mathbf{f}_{\theta_0}(\mathbf{Y})$ 을 수도 타겟으로 설정하고 학습을 수행한다. 따라서 테스트 타임 학습에 사용되는 트레이닝 데이터셋은 입력 영상을 한번 디노이징한 결과 $\mathbf{f}_{\theta_0}(\mathbf{Y})$ 와 여기에 노이즈 \mathbf{n}_j 를 다시 더한 영상들의 쌍으로 구성이 된다.

손실 함수 (3)을 최적화 하게 되면, 최적 파라미터 θ^* 를 계산할 수 있고 이를 이용하면 입력 영상 \mathbf{Y} 로부터 노이즈가 제거된 버전 $\frac{1}{N} \sum_{j=1}^N (\mathbf{f}_{\theta_0}(\mathbf{Y}) + \mathbf{n}_j)$ 을 출력으로 얻을 수 있고, 이 때 출력 영상 내 노이즈의 분산 값은 $\sigma_{\theta_0}^2 + \frac{1}{N} \sigma^2$ 이 된다. $\sigma_{\theta_0}^2$ 은 사전에 학습된 네트워크의 출력값 $\mathbf{f}_{\theta_0}(\mathbf{Y})$ 에 남아 있는 노이즈의 분산을 의미한다. 특히, 노이즈의 분포가 알려진 경우에는 \mathbf{n}_j 를 매우 쉽게 생성할 수 있으므로 $N = \infty$ 로 가정할 수 있다. 이 때, 업데이트된 네트워크 파라미터 θ^* 를 이용하여 출력한 영상의 기대값은 당연하지만 $\mathbf{f}_{\theta_0}(\mathbf{Y})$ 가 되고, 그 때 출력 영상 내 노이즈의 분산 역시 $\sigma_{\theta_0}^2$ 이 된다. 결국, 초기 네트워크 \mathbf{f}_{θ_0} 의 출력 값이 자연스레 출력되게 된다. 그러나 이 때 영상 내 유사한 패치가 존재하는 경우, 즉 패치 리커런스가 있는 경우, 새로운 결과를 기대할 수 있다.

가령, 영상 내 M 개의 동일한 패치 $\{y_1, \dots, y_M\}$ 가 존재하는 경우, 이 패치들을 θ^* 를 이용하여 디노이징 하게 되면 출력으로 패치들의 평균 값 $\frac{1}{M} \sum_{j=1}^M \mathbf{f}_{\theta_0}(y_j)$ 을 얻을 수 있다. 이 때, 출력된 패치 내 노이즈의 분산은 $\frac{1}{M} \sigma^2$ 으로 더욱 줄어들게 된다. 일반적으로 σ^2 에 비해 $\sigma_{\theta_0}^2$ 이 훨씬 작기 때문에, 패치 리커런스가 있는 경우, 사전에 학습된 파라미터 θ_0 를 이용할 때보다 θ^* 를 이용하여 출력 영상의 노이즈를 더욱 줄일 수 있게 된다. 재미있는 부분은 영상 내 패치 리커런스의 정도에 따라 다른 디노이징 성능을 보

<표 2> 제안하는 기법과 기존의 자기 지도 학습 방법을 이용하여 M 개의 동일한 패치에 대한 디노이징을 수행했을 때 디노이징 결과 비교

	$E[y] (N \rightarrow \infty)$	$Var[y] (N \rightarrow \infty)$	학습 난이도	임의의 노이즈
Noise-to-void [5] 및 Noise-to-self [7]	$\frac{1}{M} \sum_{j=1}^M y_j$	$\frac{1}{M} \sigma^2$	어려움 (패치기반 학습)	다룰 수 있음
Restore-from-restored	$\frac{1}{M} \sum_{j=1}^M f_{\theta_0}(y_j)$	$\frac{1}{M} \sigma_0^2$	쉬움 (이미지기반 학습)	다룰 수 없음

일 수 있다는 점이다. 즉, 영상 내 동일한 패치가 많은 패치는 그렇지 않은 패치들에 비해 디노이징이 더욱 잘 수행될 수 있다.

추가적으로 noise-to-void나 noise-to-self 기법처럼 이미지 복원 네트워크가 모두 컨벌루션 연산으로 구성이 된 경우(fully convolutional neural network) 동일한 모양의 패치를 찾는 명시적인 과정 없이도 자연스레 패치 리커런스 정보를 학습 중에 활용할 수 있는 장점이 있다. 또한, 제안하는 식 (3)은 패치기반이 아니라 이미지 기반으로 학습이 가능하고, 블라인드 함수 같은 학습 효율을 떨어뜨리는 모듈이 필요 없기 때문에 매우 빠르고 효과적으로 학습을 수행할 수 있다. <표 2>는 기존의 자기 지도 학습 방법과 제안하는 손실 함수 (3)을 최적화하여 학습을 했을 때, 패치 y 에 대한 디노이징 결과를 비교하여 정리한 것이다.

III. 이미지 복원 네트워크의 동적 업데이트 알고리즘

본 장에서는 손실 함수 (3)을 주어진 노이즈가 있는 입력 영상을 이용하여 최적화하여 네트워크 파라미터를 동적으로 업데이트하고, 입력 영상으로부

터 노이즈를 제거한 결과를 얻는 알고리즘을 소개 한다. 먼저 미리 학습된 파라미터 θ_0 를 이용하여 입력 영상의 노이즈를 제거한 버전 $f_{\theta_0}(Y)$ 을 획득한다. 다음으로 $f_{\theta_0}(Y)$ 에 추가적으로 노이즈 n_j 을 더한 후 새로운 학습 데이터셋 쌍 $(f_{\theta_0}(Y) + n_j, f_{\theta_0}(Y))$ 를 구성하고 이를 이용하여 식 (3)을 최적화 한다. 제안하는 “restore-from-restored” 알고리즘의 수도 코드는 <알고리즘 1>과 같다.

IV. 실험 결과

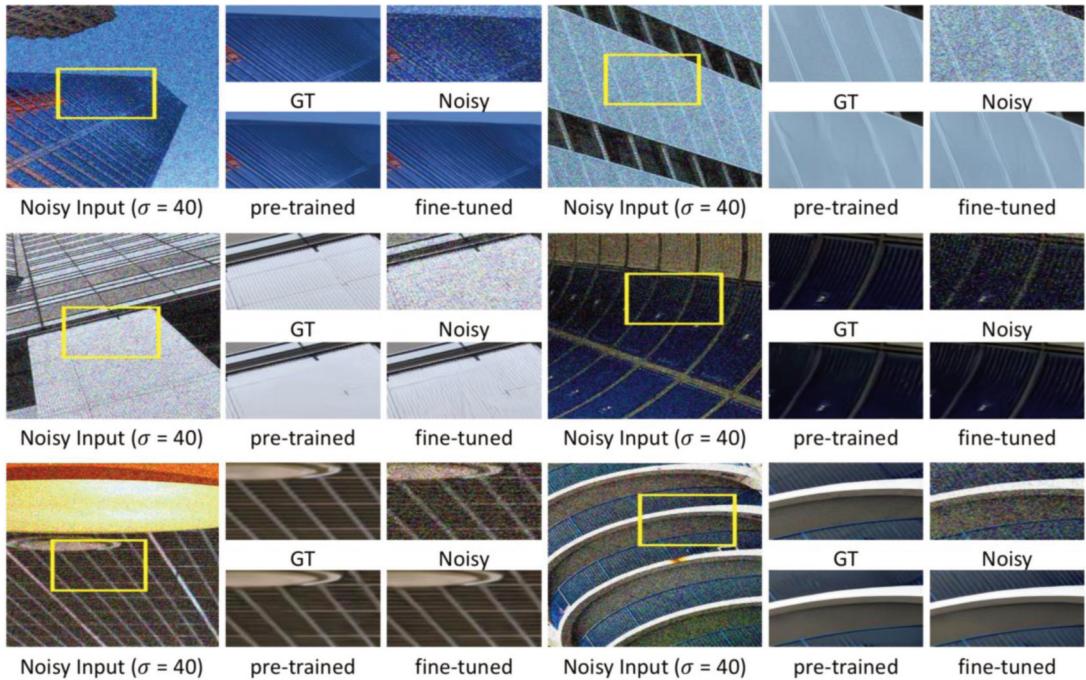
제안하는 restore-from-restored 알고리즘을 이용하면 기존의 최신의 디노이징 네트워크의 파라미터를 입력 영상에 특화되도록 만들 수 있고, 따라서 디노이징 성능을 더욱 향상시킬 수 있다. 본 비교 실험을 위해서 가우시안 디노이징 네트워크로 DnCNN[8], RIDNet[15], 그리고 RDN[16]의 네트워크 아키텍처가 사용되었고, DIV2K train 데이터셋을 이용하여 사전에 충분히 학습을 진행하였다. 사전 학습을 위해서 입력 영상에는 가우시안 노이즈가 첨가되었고, 가우시안 노이즈의 분산은 0에서 40 사이의 실수값으로 랜덤하게 결정되도록 하여 학습을 수행하였다.

<알고리즘 1> 제안하는 *restore-from-restored* 알고리즘의 수도 코드**Algorithm 1: Restore-from-restored algorithm****Input:** noisy input image \mathbf{Y} **Output:** fine-tuned parameter θ^* , restored image $\mathbf{f}_{\theta^*}(\mathbf{Y})$ **Require:** pre-trained network \mathbf{f}_{θ_0} , number of training set N , learning rate α , noise level σ 1 $i \leftarrow 0$ while $i < N$ do2 $\mathbf{n} \sim \mathcal{N}(0, \sigma^2)$ // generate random noise3 $Loss(\theta_i) = (\mathbf{f}_{\theta_i}(\mathbf{f}_{\theta_0}(\mathbf{Y}) + \mathbf{n}_i) - \mathbf{f}_{\theta_0}(\mathbf{Y}))^2$ // calculate the loss4 $\theta_{i+1} \leftarrow \theta_i - \alpha \nabla_{\theta_i} Loss(\theta_i)$ // update the network parameter5 $i \leftarrow i+1$

end

6 $\theta^* \leftarrow \theta_N$ **Return:** θ^* , $\mathbf{f}_{\theta^*}(\mathbf{Y})$ <표 3> 다양한 데이터셋, 다양한 노이즈 레벨에 대한 최신 디노이징 기법들의 디노이징 성능 비교. 특히, DnCNN 8, RIDNet [15], RDN [16] 네트워크의 파라미터 업데이트 전후로 디노이징 결과가 표기되어 있고, 자기지도 학습에 사용하는 데이터의 수인 N 이 커질수록 디노이징 성능이 더욱 개선되는 것을 확인가능 함.

Method	Adaptation	Dataset	Urban100				DIV2K testset				BSD68			
			Noise	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$	$\sigma = 40$	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$	$\sigma = 40$	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$
BM3D [1]	-		35.77	31.92	29.38	27.06	36.15	32.13	29.58	27.49	35.75	31.52	20.07	27.19
FFDNet [17]	-		35.43	31.87	29.51	27.60	36.36	32.55	30.08	28.11	35.82	31.87	29.55	27.82
RIDNet [15]	Fully pre-trained N = 5 N = 10 N = 15 N = 20		36.18	32.98	31.09	29.71	36.99	33.41	31.39	29.99	36.28	32.51	30.47	29.11
			36.26	33.07	31.20	29.84	37.05	33.47	31.46	30.07	36.33	32.56	30.53	29.18
			36.28	33.10	31.24	29.88	37.07	33.49	31.48	30.09	36.35	32.57	30.54	29.19
			36.30	33.12	31.26	29.91	37.09	33.50	31.49	30.10	36.35	32.58	30.55	29.20
			36.31	33.14	31.28	29.93	37.10	33.51	31.50	30.11	36.36	32.59	30.56	29.21
DnCNN [8]	Fully pre-trained N = 5 N = 10 N = 15 N = 20		35.85	32.49	30.51	29.07	36.73	33.12	31.08	29.66	36.16	32.35	30.29	28.91
			35.93	32.60	30.63	29.20	36.78	33.17	31.14	29.72	36.21	32.40	30.34	28.96
			35.95	32.64	30.68	29.25	36.81	33.20	31.16	29.75	36.22	32.41	30.36	28.99
			35.97	32.66	30.71	29.29	36.82	33.21	31.18	29.76	36.23	32.43	30.37	29.00
			35.99	32.68	30.74	29.31	36.83	33.22	31.19	29.78	36.24	32.44	30.38	29.01
RDN [16]	Fully pre-trained N = 5 N = 10 N = 15 N = 20		36.27	33.01	31.08	29.65	37.07	33.48	31.44	30.02	36.22	32.43	30.39	29.02
			36.33	33.11	31.18	29.76	37.12	33.53	31.49	30.08	36.25	32.46	30.43	29.06
			36.37	33.15	31.23	29.81	37.14	33.55	31.52	30.11	36.27	32.48	30.44	29.08
			36.39	33.18	31.27	29.85	37.16	33.57	31.53	30.12	36.27	32.49	30.45	29.09
			36.41	33.20	31.30	29.89	37.17	33.58	31.54	30.14	36.28	32.49	30.46	29.10



<그림 1> 주관적인 디노이징 성능 비교. RIDNet[15]의 네트워크 파라미터 업데이트 전후로 디노이징 결과물이 비교 제시됨

〈표 3〉에는 여러가지 데이터셋, 그리고 다양한 노이즈 레벨에 대한 최신 디노이징 기법들의 디노이징 성능이 비교되어 있다. 특히, DnCNN[8], RIDNet[15], RDN[16] 기법들은 제안하는 “restore-from-restored” 방법을 이용하여 네트워크 파라미터를 업데이트 한 후의 성능도 함께 표기를 하였고, 네트워크 파라미터 업데이트를 통해 매우 큰 폭으로 디노이징 성능을 향상시킬 수 있는 것을 확인할 수 있다.

다음으로 〈그림 1〉에는 사전에 학습된 RIDNet[15]의 네트워크 파라미터를 테스트 타임에 동적으로 업데이트 함으로써 점진적으로 개선되는 디노이징 결과물을 시각적으로 보여주는 결과물이다. 특히 패치 리커런스가 풍부한 영상에 대해 매우 높은 성능 개선효과를 확인할 수 있다.

제안하는 방법에 대한 더욱 자세한 설명과 다양

한 실험결과를 확인하기 위해서는 필자의 테크니컬 리포트[18]를 참조하기 바란다.

V. 결 론

본 원고에서는 테스트 타임에 네트워크의 파라미터를 동적으로 업데이트 함으로써 훼손된 입력 영상에 대한 복원 성능을 더욱 향상시킬 수 있는 새로운 학습 방법인 “restore-from-restored” 알고리즘을 제시하였다. 제안한 알고리즘을 이용하면 영상 내 패치 리커런스 정보를 활용함으로써 기존의 자기 지도 학습 기법들에 비해 훨씬 빠르고 효과적으로 네트워크 파라미터를 업데이트 할 수 있게 되고, 더 나아가 더 높은 수준의 복원 성능을 달성할 수 있음을 다양한 실험을 통해 확인하였다.

참고문헌

- [1] Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing* 16(8), 2080-2095 (2007)
- [2] Portilla, J., Strela, V., Wainwright, M.J., Simoncelli, E.P.: Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Transactions on Image Processing* 12(11), 1338-1351 (2003)
- [3] Buades, A., Coll, B., Morel, J.: A non-local algorithm for image denoising. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2005)
- [4] Ehret, T., Davy, A., Morel, J.M., Facciolo, G., Arias, P.: Model-blind video denoising via frame-to-frame training. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2019)
- [5] Krull, A., Buchholz, T.O., Jug, F.: Noise2void-learning denoising from single noisy images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2019)
- [6] Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M., Aila, T.: Noise2Noise: Learning image restoration without clean data. In: *International Conference on Machine Learning (ICML)*. vol. 80, pp. 2965-2974 (2018)
- [7] Batson, J., Royer, L.: Noise2self: Blind denoising by self-supervision. In: *International Conference on Machine Learning (ICML)* (2019)
- [8] Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing* 26, 3142-3155 (2017)
- [9] Anwar, S., Barnes, N.: Real image denoising with feature attention. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (2019)
- [10] Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (2009)
- [11] Shocher, A., Cohen, N., Irani, M.: "zero-shot" super-resolution using deep internal learning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
- [12] Huang, J.B., Singh, A., Ahuja, N.: Single image super-resolution from transformed self-exemplars. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015)
- [13] Lefkimiatis, S.: Non-local color image denoising with convolutional neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016)
- [14] Liu, D., Wen, B., Fan, Y., Loy, C.C., Huang, T.S.: Non-local recurrent network for image restoration. In: *Advances in Neural Information Processing Systems (NIPS)* (2018)
- [15] Anwar, S., Barnes, N.: Real image denoising with feature attention. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (2019)
- [16] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image restoration. *CoRR* abs/1812.10477 (2018)
- [17] Zhang, K., Zuo, W., Zhang, L.: Ffdnet: Toward a fast and flexible solution for cnn based image denoising. *IEEE Transactions on Image Processing* 27, 4608-4622 (2018)
- [18] Lee, S., Cho, D., Kim, J., Kim, T.H., Restore from Restored: Single Image Denoising with Pseudo Clean Image. arXiv:2003.04721

필자소개

김태현

- 2011년 ~ 2016년 : 서울대학교 전기정보공학부 박사
- 2016년 ~ 2018년 : 맥스플링크 연구소 연구원
- 2018년 ~ 현재 : 한양대학교 컴퓨터소프트웨어학부 조교수
- 주관심분야 : 컴퓨터비전, 머신러닝

