

일반논문 (Regular Paper)

방송공학회논문지 제26권 제1호, 2021년 1월 (JBE Vol. 26, No. 1, January 2021)

<https://doi.org/10.5909/JBE.2021.26.1.88>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 저조도 환경 기반 색상 잡음 검출 및 영상 복원

오 교 혁<sup>a)</sup>, 이 재 린<sup>a)</sup>, 전 병 우<sup>a)†</sup>

### Color Noise Detection and Image Restoration for Low Illumination Environment

Gyoheak Oh<sup>a)</sup>, Jaelin Lee<sup>a)</sup>, and Byeungwoo Jeon<sup>a)†</sup>

#### 요 약

CCTV를 사용하여 저조도와 같은 열악한 환경에서도 범죄 예방 및 특정 대상을 정확히 확인하는 것이 최근 더욱 중요해지고 있다. 저조도 환경하의 CCTV 응용에서는 눈에 거슬리지 않는 근적외선 조명을 이용하여 영상을 획득하는데, 이 경우, 비록 사람 눈에는 어두운 저조도 환경이지만 근적외선 조명을 사용하기 때문에 영상의 상세 텍스처 정보를 얻을 수 있는 장점은 있지만, CCTV 영상내의 물체 판별이나 인물 확인을 위하여 매우 요긴한 정보인 색상 정보는 얻기 힘들다는 단점이 있다. 본 논문에서는 저조도 환경에서 근적외선 조명을 사용하여 얻은 CCTV 영상으로부터 DCGAN을 사용하여 색상정보를 획득하는 방법과 이때 재구성된 색상 영상에 생기는 색상 잡음을 제거하는 방법을 제시한다.

#### Abstract

Recently, the crime prevention and culprit identification even in a low illuminated environment by CCTV is becoming ever more important. In a low lighting situation, CCTV applications capture images under infrared lighting since it is unobtrusive to human eye. Although the infrared lighting leads to advantage of capturing an image with abundant fine texture information, it is hard to capture the color information which is very essential in identifying certain objects or persons in CCTV images. In this paper, we propose a method to acquire color information through DCGAN from an image captured by CCTV in a low lighting environment with infrared lighting and a method to remove color noise in the acquired color image.

**Keyword:** Machine learning, DCGAN, color, filter, Near-Infrared

---

a) 성균관대학교 전자전기컴퓨터공학과(Department of Electrical and Computer Engineering)

† Corresponding Author : 전병우 (Byeungwoo Jeon)

E-mail: bjeon@skku.edu

Tel: +82-31-290-7144

ORCID: <https://orcid.org/0000-0002-5650-2881>

※ 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (NO. 2018-0-00348, CCTV 제약점 개선을 통해 범인 검거율 저하 문제 해결을 지원하는 지능형 영상 보안 시스템 기술 개발). 연구 데이터 수집에 도움을 준 이준형 연구원에게 감사를 드립니다.

※ This work was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2018-0-00348, Development of Intelligent Video Surveillance Technology to Solve Problem of Deteriorating Arrest Rate by Improving CCTV Constraint). Appreciation to Researcher Joon-Hyung Lee for helping us to collect research data.

· Manuscript received October 21, 2020; Revised December 22, 2020; Accepted January 6, 2021.

Copyright © 2021 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

## I. 서론

대부분의 CCTV는 주변 조도가 일정 이하가 될 경우, 근적외선 (infrared) 조명을 사용하여 회색조 (graylevel) 영상을 획득하는 NIGHT 모드로 동작한다. 근적외선을 이용하여 얻어진 회색조 영상은 일반적인 RGB 영상에서 획득하기 어려운 세부적인 텍스처 정보를 제공할 수 있는 장점은 있지만, CCTV 상의 물체 판별이나 인물 확인을 위해 매우 요긴한 정보인 색 정보를 제공할 수는 없다는 단점이 있다. 이러한 문제점을 해결하기 위해서 [1]에서는 Illumination map을 이용한 저조도 이미지 향상 방법이 진행되고 있으며, [2]에서는 Global Illumination-Aware and Detail Preserving Network를 이용한 저조도 이미지 향상에 관한 연구가 진행되고 있다. 최근에는 Y. Jiang이 제시한 GAN을 통해 저조도 영상의 attention map을 이용해 화질 향상하는 방법이 제시되었다<sup>[3]</sup>.

저조도 영상의 화질 향상에 관한 연구 중에서, Illumination map을 이용한 화질 향상 방법은 빠른 실행시간을 통해 쉽게 영상을 생성할 수 있다는 점 때문에 실용성이 뛰어나 최근 연구에서 각광을 받고 있다. Illumination map을 이용한 화질 향상 연구는 빛이 존재하는 환경에서는 유용하지만, CCTV 영상과 같이 극히 낮은 조도 환경에서는 제대로 된 Illumination map을 생성하지 못해 향상이 잘 이루어 지지 않는다.

이에 따라 본 연구에서는, 저조도 환경에서도 통상적인 CCTV의 NIGHT 모드 대신, 주간 때와 마찬가지로 RGB 센서를 통하여 컬러영상을 획득하되, 근적외선 차단 필터를 제거하여 가시영역 및 근적외선 혼합 영상을 획득하도록 하였다. 또한, CCTV의 특성인 한 곳만 바라본다는 특징을 반영하여, DCGAN<sup>[4]</sup>을 사용하여 근적외선 영상과 가시영역 영상을 분리한 후, 가시영역 영상 값을 이용해 원래의 색과 유사한 색으로 복원하는 중단 간 네트워크를 설계하여 영상을 복원하였다. 또한, 본 논문에서는 CCTV의 저조도 환경에서 취득한 근적외선 데이터를 통해 복원한 색상 영상의 배경에서 생성된 잡음을 제거하기 위해, 근적외선 영상을 이용해 물체와 배경을 분리하고, 생성된 영상의 색

상 잡음을 제거하는 방법을 제시한다. 또한 원본 영상과 수정된 DCGAN을 이용해 복원한 방법과 기존 DCGAN을 이용해 복원한 방법과 Illumination map을 이용한 색 복원 방법의 비교를 통해 성능을 비교한다.

## II. Deep Convolutional Generative Adversarial Network

### 1. DCGAN을 통한 영상 생성 방법

DCGAN<sup>[4]</sup>은 비지도 학습으로 적대 신경망 간의 경쟁을 통해 영상을 생성하거나, 영상 변환과 같은 특정 작업에 활용한다. 비지도 학습은 정답 라벨이 없는 데이터를 비슷한 특징끼리 군집화 하여 새로운 데이터에 대한 결과를 예측하는 방법이다. DCGAN은 훈련 데이터를 통해 데이터의 분포를 생성하고 이를 기반으로 영상을 생성한다. DCGAN은 기본적으로 생성자와 판별자로 분류된다. 생성자는 훈련 데이터를 가지고 원본과 유사하게 영상을 생성하려 하고, 판별자는 생성자가 원본 영상과 비교하여 생성한 영상의 진실 유무를 판별하는 것을 학습한다. 적대학습은 분류 모델을 먼저 학습한 후, 생성 모델을 학습하는 과정을 반복한다. 분류 모델 학습은 진짜 데이터를 입력한 후 진짜로 분류하도록 학습하는 과정과 생성자에 의해 생성된 가짜 데이터를 입력한 후 가짜로 분류하도록 학습하는 과정으로 이루어진다. 여기서, 생성 모델은 분류 모델을 오인시킬 수 있는 정도의 데이터를 생성하도록 학습하고, 분류 모델은 생성 모델이 생성한 데이터와 진짜 데이터를 정확하게 구분하도록 학습한다. DCGAN의 손실 함수는 다음과 같다.

식 (1)의 손실함수에서 G는 생성자, D는 판별자를 의미한다. 여기서  $x \sim p_{data}$ 는 실제 데이터의 확률 분포에서 샘플링한 데이터를 의미하고, y는 class label 등과 같은 보조 정보를 의미하며,  $z \sim p_z(z)$ 는 임의의 잡음에서 샘플링한 데이터를 의미한다.

$G(z|y)$ 는 잡음변수 z를 원하는 실제 데이터 공간 x에 매핑하는 것을 의미한다. 이와 달리,  $D(x|y)$ 는 판별모델을

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}} [\log D(x|y)] + E_{z \sim p_z(z)} [\log (1 - D(G(z|y)))] \quad (1)$$

모델링하는데, 주어진 데이터  $x$ 가 실제 데이터 세트  $y$ 에서 나올 확률로 계산한다. 이 판별 모델은 실제 입력 영상에 대해서는 진짜로 분류될 확률을 최대화하고 생성된 데이터가 진짜로 분류될 확률은 최소화되도록 학습된다. 이는  $D(x|y)$ 를 최대화하고  $D(G(z|y))$ 를 최소화하는 것으로 실현될 수 있다. 생성모델은 생성된 가짜 영상데이터를 판별 모델이 진짜라고 분류하도록  $G(z|y)$ 를 최대화한다. 이때, 생성모델은 실제 데이터와 유사한 영상을 생성하여 판별 모델의 손실함수를 최적화하도록 하고, 판별 모델은 데이터 분류를 통해 생성 모델에서 생성된 영상이 실제 데이터로 분류될 확률을 전달하여 생성 모델의 손실함수를 최적화한다.

DCGAN 모델은 기존 GAN<sup>[5]</sup> 모델을 변형한 것이다. 기존의 GAN은 U-Net<sup>[6]</sup> 구조를 기반으로 생성된 모델로, pooling layer를 사용하지만 DCGAN<sup>[4]</sup> 모델은 전치된 컨볼루션을 사용해 pooling layer를 대체한다. 모든 층이 컨볼루션을 통해 진행되기 때문에 생성 모델과 판별 모델이 각각 up sampling 될 수 있게 되어 구조적 자유도가 증가하고 해상도가 기존 GAN에 비해 상승하였다. 또한 마지막에 있는 모두 연결된 신경망을 제거하고, 생성 모델의 마지막 출력 데이터와 판별 모델의 처음 입력 데이터를 제외한 신경망들에 배치 정규화를 실시하였다. 배치 정규화 층은 모드 붕괴 현상을 완화하기 위해서 추가하였지만, 완전히 해결되지 않을 경우 잡음이 발생하게 된다. 모드 붕괴 현상이란 판별자와 생성자중 하나의 모델이 다른 모델에 비해 월등

히 학습되어 학습의 불균형으로 인해 모델이 제대로 학습되지 못하는 상태를 의미한다. 마지막으로 생성 모델에서 마지막 출력 데이터를 제외하고 RELU layer를 사용하였고 마지막 출력에서는 Tanh layer를, 판별 모델에서는 leaky RELU layer 사용하여 모델에서 발생하는 그라디언트 손실을 방지하였다.

## 2. GAN을 이용한 저조도 영상 향상

Illumination map을 이용한 저조도 영상의 조도 및 색 향상에 관한 연구는 지속적으로 진행되고 있다. [3]의 논문은 기존의 Illumination map을 이용하였으나, 기존에 알고리즘과 다르게 영상에 대해 입력 영상과 attention map을 결합하여 GAN 네트워크를 통해 저조도 영상의 조도 및 색 향상을 진행하였다. 기존의 Illumination map을 이용한 연구와 달리 attention map을 저조도 영상에 결합하여 영상을 복원한 뒤 국부 판별자와 지역 판별자 2개를 이용해서 생성 모델의 학습을 진행한다.

위 논문에서는 밝은 영역에 비해 어두운 영역에 대해서 향상을 진행하고자, RGB 영상에서 조도 채널인 I 채널을 이용해 attention map을 생성한다. 사용한 방법으로는 입력 영상의 I 채널을  $[0, 1]$  값으로 정규화 시키고, 각각의 좌표에 대해  $1 - I_{x,y}$  연산을 통해 attention map을 생성한다.

Convolution을 시행한 각각의 layer에 대해 attention map이 생성되고 이를 디코더 측에 전송하여 up sampling 한 영

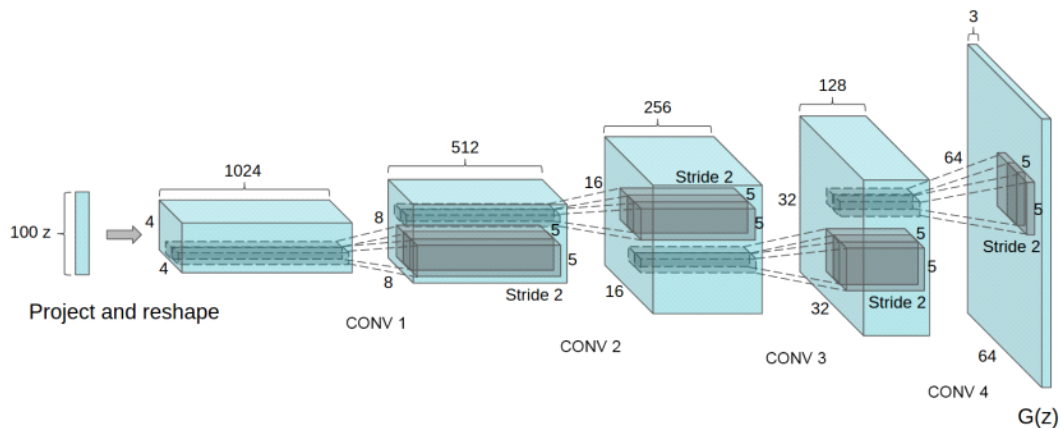


그림 1. DCGAN의 생성 모델 구조<sup>[4]</sup>

Fig. 1. DCGAN generator model architecture<sup>[4]</sup>

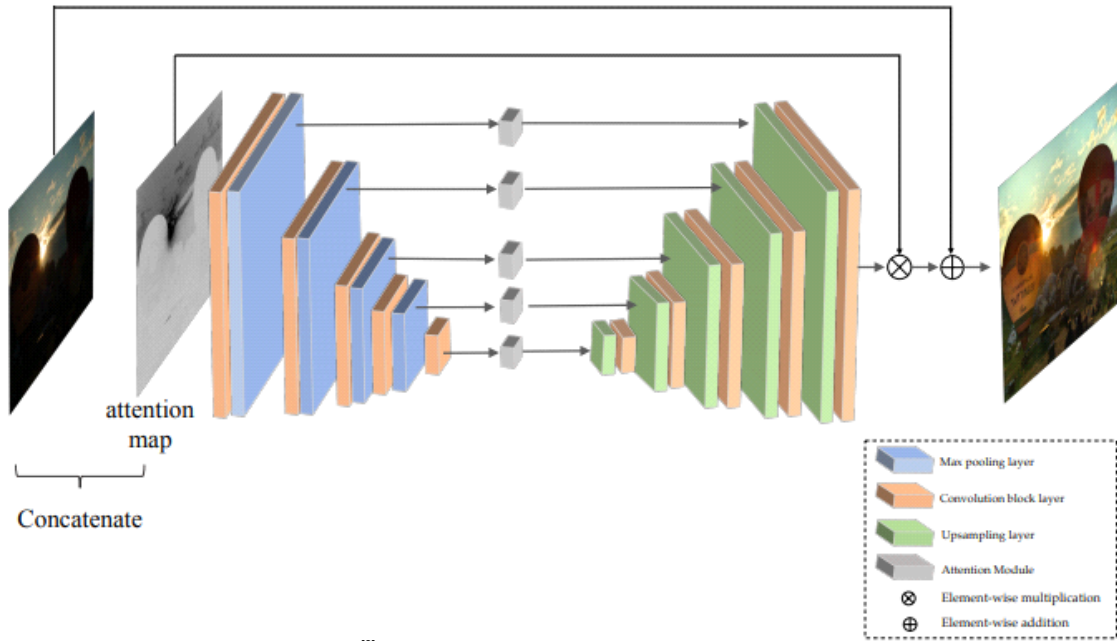


그림 2. Enlighten GAN 생성 모델 구조도<sup>[3]</sup>

Fig. 2. Enlighten GAN generator model architecture<sup>[3]</sup>

상에 attention map을 연결해준다. 최종적으로 생성된 영상과 기존의 attention map을 요소별 곱셈을 통해 보강해주고, 입력 영상을 더해주는 것으로 밝은 부분을 보완해준다. 판별 모델의 경우 국부 판별자와 지역 판별자로 나뉘었다고 서술되어 있다. 이는 전체적인 영상에 대해 판별하는 국부 판별자와 전체적인 영상 중 일부를 잘라내어 판별하는 지역 판별자로 둘의 네트워크 구조는 동일하게 구성되어 있다. 이 네트워크의 유리한 점은 국부 판별자와 지역 판별자의 특성을 동시에 활용할 수 있는 것에 있다.

### III. 사용한 DCGAN 생성 및 판별 모델

#### 1. 수정된 DCGAN 생성 모델

본 논문은 CCTV 응용에 사용할 수 있는 색상 복원 모델을 개발하는 것을 목표로 한다. 기존의 DCGAN<sup>[4]</sup>을 통한 근적외선 색상화 연구와 달리 본 논문은 저조도 환경에서의 가시영역 및 근적외선 신호가 혼합된 영상을 획득한 후, 이 데이터로부터, 가시영역 및 근적외선 영상 신호를 분리

한 후, 정상조도에서 촬영한 것과 최대한 유사한 색상과 텍스처 정보를 갖는, 개선된 가시영역 영상으로 복원하는 것을 목표로 한다. 본 연구에서 특히 문제가 되었던 것은 저조도 영상을 사용하여 색상을 복원할 경우, 잡음이 매우 많이 발생하는 것이었다. 본 논문은 이 색상잡음 문제를 해결하기 위하여, 혼합영상신호에서 분리한 근적외선 영상과 단일 네트워크가 아닌 중단 간 네트워크를 사용하여 잡음이 현저하게 감소된 색상 영상을 복원한다.

그림 3은 본 논문에서 사용한 생성 모델이다. 기존의 DCGAN 네트워크는 1x1x256까지 줄어든다. 그렇게 되면 공간 해상도가 낮아지게 되어 흐릿한 영상이 출력된다. 따라서, 채널의 개수는 동일하게 두고 선, 가장자리 등과 같은 상세정보를 보존하기 위해 신경망 층을 줄여 32x32x256까지만 컨볼루션 연산을 진행하여 공간 해상도를 향상시켰다. 또한 기존 DCGAN과 달리 잔차 신경망 층을 중간에 추가하였다. 잔차 신경망 층에는 shortcut 연결이 존재한다.

Shortcut 연결은 입력과 출력의 연결을 통해 영상의 모든 값에 대해서 연산하지 않고 잔차 신호에 대해서만 연산을 하여, 연산속도를 빠르게 하고 깊은 네트워크의 오류 값을 줄여 성능을 개선하기 위해 추가한 것이다. 또한 잔차 블록

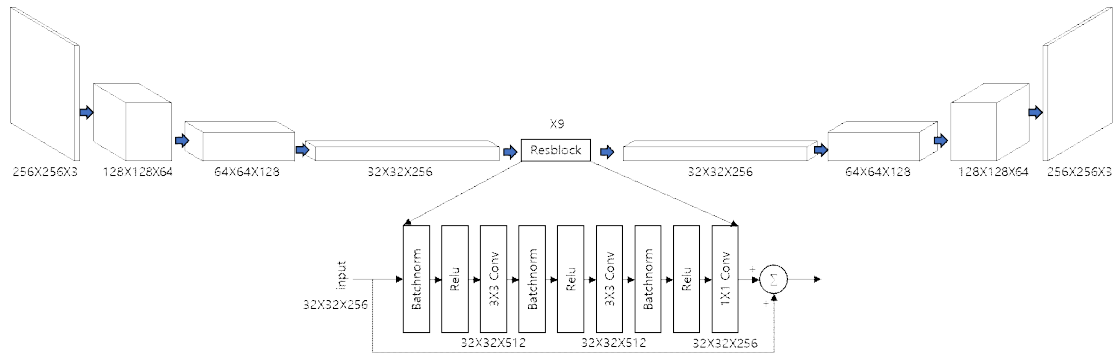


그림 3. 본 논문에서 사용한 수정된 DCGAN 생성 모델

Fig. 3. Modified DCGAN generation model used in this paper

에 1x1 컨볼루션 레이어를 추가하여 차원을 줄여 연산량을 줄여 네트워크 학습 속도를 빠르게 만드는 것으로 생성 모델을 설계하였다. 동일한 학습 횟수로 학습시켰을 경우, 기존 DCGAN을 이용한 학습에는 26시간이 소요되었으나, 본 논문이 설계한 생성 모델을 사용해 학습하였을 경우 20시간이 소요되는 것을 확인하였다.

## 2. 수정된 DCGAN 판별 모델

기존의 판별 모델은 전체적인 배경 및 전역 콘텐츠에 대해서 학습하였다. 하지만 본 논문의 목적은 근적외선 및 가시영역 분리 및 복원이다. 이 때, 신호세기가 매우 작은 RGB 값을 이용해 정상조도 가시영역 영상을 복원해야 된다. 이러한 경우 전역 콘텐츠에 대해서 학습하는 것 보다 국부 콘텐츠에 대해서 학습하여 신호세기가 작은 신호들을 복원해 주는 것이 유리하다. 기존의 U-NET<sup>[6]</sup> 구조의 판별 모델은 컨볼루션 연산과 배치 정규화 방법만 사용하여 학습을 진행한다. 이는 전역 콘텐츠에 대해서 학습하는 방법

이다. 이를 해결하기 위해 논문에서는 기존의 U-NET<sup>[6]</sup> 구조의 판별 모델에서 국부 콘텐츠 학습에 유리한 FPN<sup>[7]</sup> 네트워크를 구성해 판별 모델을 생성했다. 그림 4를 보면 FPN 네트워크는 위로 올라가는 forward 단계를 통해 각 레이어의 특징 정보를 생성한다. 이때 생성된 각 레이어의 특징 정보는 top-down의 skip connection 단계를 통해 손실된 지역적 정보를 보충하며 판별 모델을 생성한다. Skip connection은 encoder단에서 소실된 낮은 단계부터 높은 단계의 공간 해상도 특징을 decoder단에 보충해주는 역할을 한다.

그림 5는 본 논문의 전체적인 구조도이다. 위의 생성 모델과 판별 모델은 마지막 부분에 나오는 Denoising을 수행하기 전에 사용되는 네트워크 모델이다. 또한, 출력되는 최종 결과물은 분리된 저조도 가시영역 영상, 복원된 적정조도 가시영역 영상, 분리된 근적외선 영상이다. 본 논문은 분리된 근적외선 영상을 이용해 적정조도 가시영역 영상에 생겨난 색 잡음을 Denoising 블록을 통하여 제거하여 원본 영상과 유사한 영상을 생성한다.

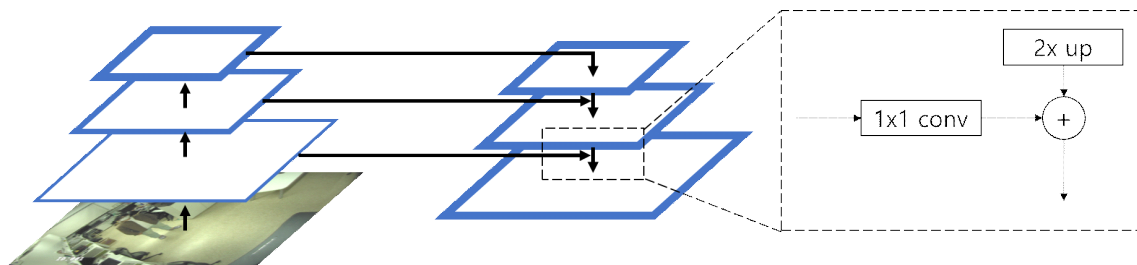


그림 4. 본 논문에서 사용한 수정된 DCGAN 판별 모델

Fig. 4. Modified DCGAN discriminator model used in this paper

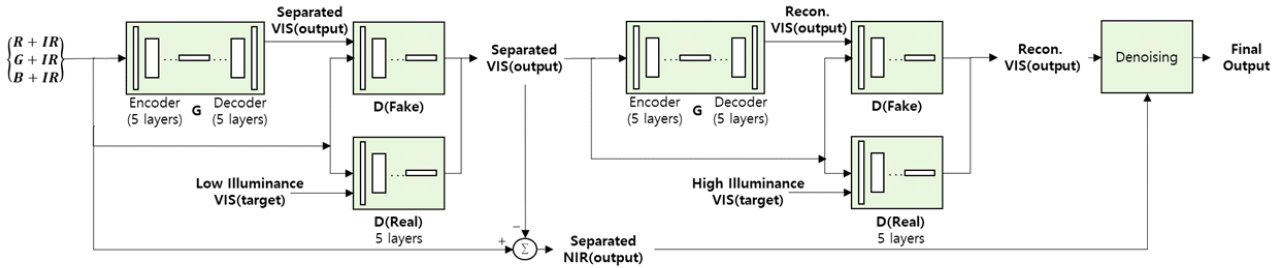


그림 5. 전체 프로세스 구조도  
Fig 5. Overall process structure diagram

#### IV. DCGAN을 통한 신호 분리, 색 복원 및 물체 영역 별 색상 잡음 제거

##### 1. DCGAN을 이용한 근적외선 및 가시영역 혼합영상 신호분리 및 복원

본 논문이 제안하는 방법의 DCGAN<sup>[4]</sup>은 중단 간 네트워크이다. DCGAN 네트워크는 저조도에 촬영된 근적외선 및 가시영역 혼합 영상에서 가시영역 및 근적외선 신호분리를 먼저 진행한다. 이를 위하여, 근적외선 및 가시영역 혼합 영상을 입력영상으로 하고 별도로 수집된 저조도 가시영역 영상을 참조영상으로 하여 생성 모델이 저조도 가시영역 영상을 생성하도록 학습한다. 가시영역 및 근적외선 신호분리와 색상 복원에 사용한 네트워크는 III장에서 설명한 네트워크를 사용하였다. 가시영역 및 근적외선 혼합 영상의 경우 저조도 환경에서 포화되지 않으므로, 분리해낸 저조도 가시영역 영상을 입력인 혼합영상에서 감산함으로써 근적외선 영상을 획득한다. 그 후, 분리된 저조도 가시영역 영상을 이용해 정상조도 가시영역 영상을 복원한다.

실제 CCTV환경에서는 하나의 카메라만을 사용하기 때문에, 가시영역과 근적외선 영역 모두를 하나의 카메라로 획득하도록 하였다. 이를 위하여 근적외선 영역을 차단하는 핫미러를 제거하여 RGB 센서의 센싱 영역에 들어오는 근적외선 신호도 센싱 되도록 하여 작은 RGB 값에 근적외선 값을 합하여 영상을 생성한다. 근적외선 정보와 가시영역 정보가 혼합된 혼합영상으로부터 근적외선 분리 및 가시영역 복원영상이 생성되는 신호분리 네트워크를 설계하였다. 또한, 분리된 가시영역 영상에는 색 잡음이 존재한다. 이 경우, 잡음에 대한 영향을 최소화하기 위하여, 근적외선 영상을 활용하여 색상 잡음을 제거한 후, 물체가 위치하지 않는 배경 영역은 원본 배경영상으로 대체해 줌으로써 영상의 화질을 향상시키도록 하였다. 일반적인 응용에서는 배경이 수시로 바뀔 수 있으므로 이렇게 하는 것이 가능하지 않지만, 본 연구의 응용은 고정지역 감시용 CCTV 이어서 안정적인 배경영상을 항상 획득할 수 있기 때문에 가능하다. 즉, 새로운 물체가 출현하지 않는 일상 감시상태에서 CCTV는 항상 배경 영상만을 수집하고 있기 때문이다.

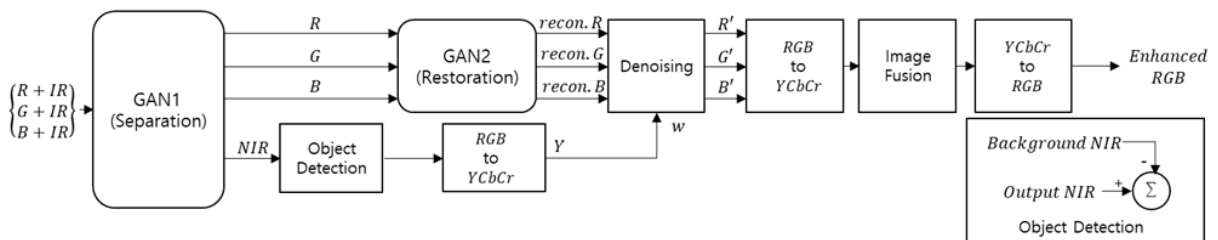


그림 6. 잡음 제거 알고리즘 순서도  
Fig. 6. Algorithm flow chart for noise removal

## 2. 색상 잡음 제거

### 2.1 물체 검출 방법

본 논문은 위에 언급한 고정지역 감시용 CCTV라는 특성을 이용해 새로운 광원이 들어오지 않은 경우, CCTV의 배경영상의 화소 값은 변화하지 않는다는 점을 이용해 배경과 물체의 영역을 분리한다. 본 논문은 감시영역내에 새로 등장한 물체에 대한 근적외선 및 가시영역 영상과 배경에 대한 근적외선 및 가시영역 영상 간의 화소 값의 차이를 이용해 새로 들어온 물체에 대한 영역을 지정하는 매우 간단한 방식으로 물체를 검출한다. 물체 검출에 대한 매우 정교하고 더욱 효과적인 방법<sup>[8,9]</sup>이 많이 있지만, 본 연구의 주안점이 아니기 때문에 물체 검출을 위한 매우 간단한 방법을 사용하였다.

### 2.2 색상 잡음 검출 및 복원

학습된 모델을 통해 생성된 영상의 색상 잡음은 근적외선의 밝기 값(Intensity)이 높은 영역에서 주로 발생한다. 따라서 근적외선 영상을 이용해 복원된 컬러영상의 잡음을 제거하기 위해 근적외선의 밝기 값에 기반하여 중요도를 나타내는 행렬인 가중치 맵을 사용한다. 가중치 맵은 후에 색상 잡음을 검출하고 제거하기 위해 사용된다. 가중치 맵을 구성하는 각 화소에 대한 가중치  $w$ 는 다음과 같다.

$$w = 1 - \frac{Y_{nir} - \min(Y_{nir})}{\max(Y_{nir}) - \min(Y_{nir})} \quad (2)$$

본 논문의 실험을 위해 획득한 영상은 동일한 RGB 카메라를 통해 획득한 저조도 근적외선 및 가시영역 혼합 영상, 저조도 가시영역 영상, 적정조도 가시영역 영상이다. 식 (2)에서  $Y_{nir}$ 은 획득한 저조도 혼합영상에서 가시영역영상을 제거하여 획득한 근적외선 영상을  $YCbCr$ 로 변환한 영상의  $Y$ 채널의 화소 값이다. 가중치 맵은 전체 영상의 화소 값을  $[0,1]$  사이 값으로 정규화 시키는 방법으로 생성한다. 근적외선 영상의  $Y$ 채널값이 큰 부분에서 색상 잡음이 자주 발생하기 때문에 근적외선  $Y$  화소 값이 큰 영역에 대해 작은 가중치를 두어 작은 변화에도 민감하게 반응하게 설계하였다. 잡음 검출 필터는 감시영역 내로 물체가 들어온 영

상과 배경 영상의 혼합 영상간 화소 값의 차이를 이용해 물체와 배경 분리를 시행한 후, 물체 각각의 영역에 대한 평균과 표준편차를 계산하여 각 영역의 잡음을 검출한다. 색상 잡음을 검출하는 필터는 영상 라인별로 식 (3)과 같이 수행된다.

$$i_{x,y} < m - w\sigma \quad \text{or} \quad i_{x,y} > m + w\sigma \quad (3)$$

$m$ 과  $\sigma$ 는 채널별로 각행의 평균과 표준편차를 의미하고,  $i_{x,y}$ 의  $x,y$ 는 화소의 위치 그리고  $i_{x,y}$ 는  $x,y$  위치에서의 화소 값을 의미한다. 식 (3)의 필터는 가중치 맵을 사용하여 각 행의 평균값에서 가중치에 표준편차를 곱한 값 이상으로 차이나는 경우, 색상 잡음이 발생한 것으로 판단한다. 즉, 각 라인 별로, 화소 값이  $[m - w\sigma, m + w\sigma]$ 범위를 넘어가는 경우 잡음이 존재하는 것으로 판단한다. 식 (3)의 필터를 통해 검출한 색상 잡음을 주변과 유사한 화소 값으로 대체해 주기 위해, 식 (4)의 매우 간단한 색상 복원 필터를 적용한다.

$$i'_{x,y} = m - w\sigma + 2 \left( \frac{I(x,y) - \min(I(x,y))}{\max(I(x,y)) - \min(I(x,y))} \times w\sigma \right) \quad (4)$$

$i'_{x,y}$ 는 색상 복원 필터에 의해 복원된  $x,y$  위치의 화소 값이다. 색상 복원 필터는 R, G, B 각각의 채널 별로, 잡음이라고 판별된 화소 값에 대해  $[m - w\sigma, m + w\sigma]$  사이 값으로 변환한다. 위의 간단한 필터를 적용하여 검출된 색상 잡음 화소 값을 주변 화소와 유사하게 변환하도록 하여, 색상 잡음을 복원하고 최종적으로 색상이 복원된 RGB 영상을 획득한다. 획득한 영상을  $YCbCr$  영상으로 변환한 후, 근적외선 영상의 휘도 성분인  $Y$  채널로 대체한다. 최종적으로 합성한  $YCbCr$  영상을 RGB 영상으로 변환하여 결과를 획득한다.

## V. 실험 획득 환경, 결과 및 분석

본 연구의 학습을 위한 영상데이터 획득을 위하여, 저조



도 환경인 2 Lux에서 촬영을 진행하였고, 노출 시간은 1/30s 로 고정시켜 사용하였다. 여기서 2 Lux로 정한 것은, 가시영역 영상의 화소 값이 0이 되지 않는 저조도 환경으로 설정하였기 때문이다. 근적외선 영상을 획득하기 위해 사용한 근적외선 조명은 CCTV에 부착되어 있는 근적외선 LED 12개를 사용하였으며, 실험에 사용된 CCTV는 통상적인 감시환경에 사용되는 일반적인 카메라로 선정하였다.

주변 조명에 따라 화이트 밸런스가 자동 조절되는 것을 방지하고자 화이트 밸런스는 RED 7 GREEN 7 BLUE 7로 제한시킨 후, 영상을 획득하였다. 화이트 밸런스의 경우 자동으로 하면 영상의 화질 향상을 기대할 수 있다. 하지만 근적외선 및 가시영역 혼합영상의 경우 사람이 들어와 전

체적인 색이 달라질 경우, 입력 영상을 다르게 인식하여 색이 달라지는 경우가 존재하는 것을 확인하여, 일정한 영상 획득을 위해 제한한 후 촬영을 진행하였다.

근적외선 LED를 켜 놓은 상태로 핫 미러 필터 사용 또는 미사용의 상태를 반복하며 저조도 가시영역 영상 및 근적외선 혼합 영상을 획득하였다. 이때, 저조도 영상을 획득하기 위해 기존에 특정 조도 이하가 되면 바뀌는 Day/Night 모드를 day 모드로 고정한 후 촬영하였으며, 정상조도 영상의 경우 핫 미러 필터를 장착한 상태로 120 lux 환경에서 영상이 포화되지 않도록 노출 시간을 1/120s 로 조정한 후, 동일 구도에서 촬영하였다.

근적외선 및 가시영역 혼합 영상의 경우 일정 조도 이하

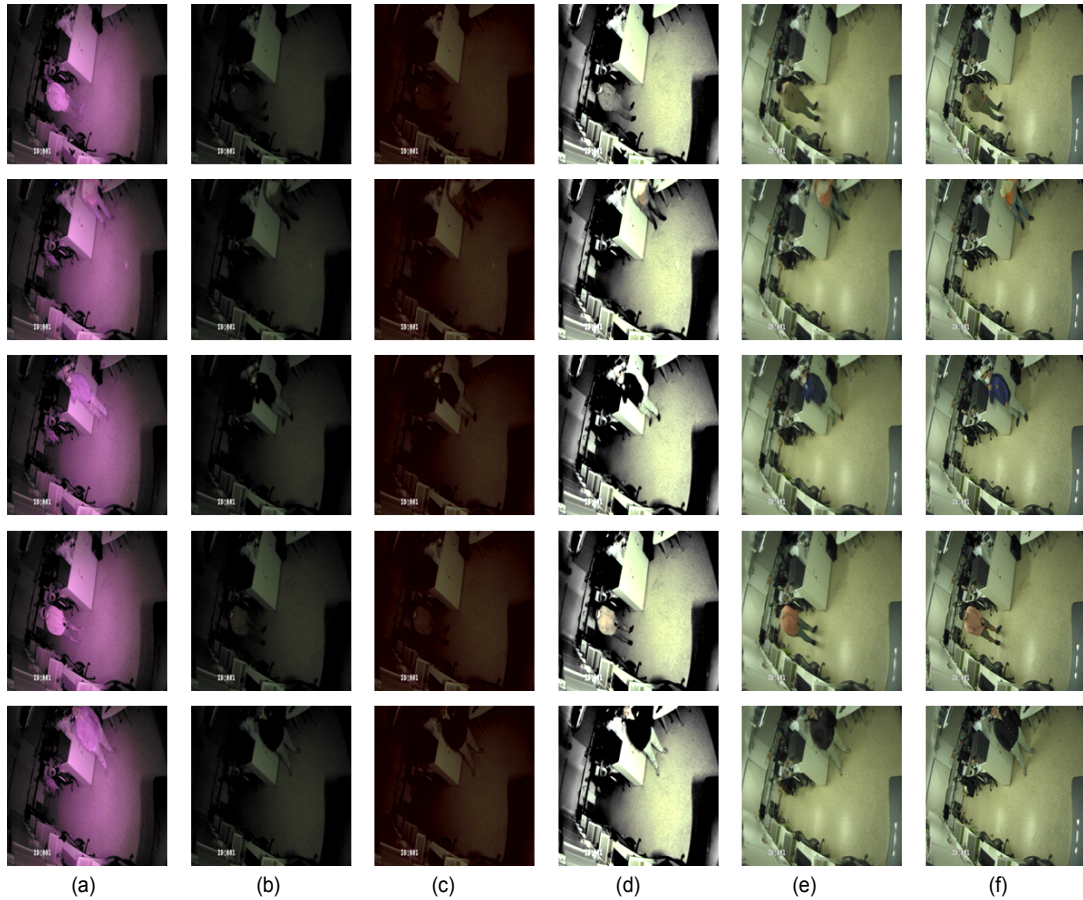


그림 7. 실험 결과 영상. (a) 실험 입력 영상, (b) 수정된 DCGAN을 통해 분리해낸 가시영역 영상, (c) Enlighten GAN[3]을 통해 복원한 영상 (d) Lime[1]을 통해 복원한 영상 (e) 수정된 DCGAN을 통해 복원한 영상 (f) 원본 영상

Fig. 7. Experiment result images. (a) Test input image (b) Separated visible image by our modified DCGAN (c) Reconstructed visible image by Enlighten GAN (d) Reconstructed visible image by Lime (e) Reconstructed visible image by our modified DCGAN (f) Ground Truth



의 환경에서 day 모드를 통해 RGB 영상을 획득하게 조정 한 후, 근적외선 조명을 켜 놓은 상태로 핫 미러 필터를 제거하여 근적외선 및 혼합 영상을 획득하였다. 위 조건으로 촬영한 가시영역 및 근적외선 혼합 영상은 그림 7(a)와 같다. 모든 데이터 영상은 1920x1080의 해상도로 촬영되었고, 총 3300장의 데이터가 학습에 사용되었다. 학습한 epoch는 총 1000회 진행하였으며,  $l1\_weight$ 는 100으로  $gan\_weight$ 는 1로 고정시켜 학습을 진행하였다. CCTV의 근적외선 조명의 특성을 학습시키기 위해 획득한 영상은 crop 및 augmentation을 하지 않고 이미지를 down sampling 하여 영상을 학습시켰다.

표 1. CCTV 설치 및 촬영 환경

Table 1. CCTV installation and capturing environment

Environment mode	Low illuminance	Normal illuminance
Lux	2.0 lux	120 lux
Exposure time	1/30s	1/120s
NIR light	CCTV NIR LED	X
Hot mirror	On/off	On

그림 7(a)는 저조도 환경 가시영역 및 근적외선 혼합 입력 영상이다. 입력 영상을 받아 본 논문이 설계한 네트워크를 통해 저조도 가시영역 영상이 분리된다. 그림 7(b)는 입력영상으로부터 분리된 저조도 가시영역 영상이다. 영상의 잡음을 제거하기 위해 생성한 근적외선 영상은 입력영상인 가시영역 및 근적외선 혼합 영상으로부터 그림 7(b)의 분리된 가시영역 영상을 차분하는 것을 통해 생성한다. 그림 7(c)는 분리한 가시영역 영상을 Enlighten GAN을 이용해 복원한 적정조도 가시영역 영상이다. 그림 7(d)는 분리한 가시영역 영상을 Lime을 이용해 복원한 적정조도 가시영역 영상이다. 그림 7(c)와 (d)의 경우에는 Illumination map 기반의 복원 영상이다. Illumination 기반 영상은 영상 조도의 차이를 이용해 저조도 영상을 복원한다. 하지만 CCTV의 저조도 영상의 특성상 영상 안에서의 조도 차이가 적기 때문에 CCTV 영상에는 적합하지 않다. 그림 7(e)는 분리한 가시영역 영상을 수정된 DCGAN을 이용해 복원한 적정조도 가시영역 영상이다. 복원된 적정조도 가시영역 영상을 확인하면 배경 부분에 대해서는 반복 학습을 통해 유사하

게 색이 복원된 것을 확인할 수 있다. 하지만 물체의 일부분에 대해서 색 잡음이 발생하는 것을 확인할 수 있다. 그림 7(f)는 복원한 영상과의 비교를 위해 넣은 원본 영상이다.

그림 8은 기존의 pix2pix<sup>[4]</sup>와 차이를 확인하고 잡음을 제거한 결과 영상이다. 그림 8(a)는 잡음을 판별하고, 제거하기 위해 사용한 근적외선 이미지다. 근적외선 이미지의 생성 방법은 그림 7에서 서술하였다. 그림 8(b)는 pix2pix를 사용하여 복원한 영상이다. Pix2pix를 사용하여 복원한 8(b)의 경우, 수정된 DCGAN을 통해 복원한 영상인 8(c)와 비교했을 때 전체적으로 색의 복원 성능이 떨어지는 것을 확인할 수 있다. 하지만 수정된 DCGAN의 경우에도 잘못된 복원으로 인한 color noise가 발생하고, 동시에 영상을 압축하여 진행한 결과 detail 성분이 많이 제거된 것을 확인할 수 있다. 그림 8(d) 영상은 잡음을 제거하고, detail 성분을 추가해준 영상이다. 잡음 부분에 대해서는 근적외선 영상을 이용해 잡음으로 인식하고, 물체 영역의 성분과 유사한 색으로 색이 복원된 것을 확인할 수 있다. 그림 8(e)는 결과 비교를 위해 넣은 원본 영상이다. 기존 알고리즘<sup>[10]</sup>과 비교했을 때 기존에는 근적외선 이미지로 가시영역 영상의 Y채널을 대체해 줌으로써 ROI 영역 이외의 영역에 대해 contrast가 저하되었으나, 근적외선 영상을 detail과 base부분을 분리한 뒤 image fusion을 통해 ROI 영역의 잡음 제거 및 detail 성분을 강화하였다.

$$PSNR = 10 \log \left( \frac{2^{bitdepth} - 1}{\frac{1}{MN} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|x(i, j) - \tilde{x}(i, j)\|_2^2} \right)^2 \text{ dB} \quad (5)$$

실험 결과는 PSNR를 통해 성능을 비교하였다.  $x(i, j)$ 는 원본 영상의 픽셀 값이며  $\tilde{x}(i, j)$ 는 산출 영상의 픽셀 값이다. 원본 저조도 가시영역 영상과 분리된 가시영역 영상은 평균 PSNR이 38dB가 나왔다. Enlighten GAN을 이용해 복원한 영상은 원본 가시영역 영상과 비교하였 때 평균 PSNR이 10.2dB가 나왔으며, LIME은 11.44dB의 복원 성능이 나왔다. 본 논문이 제안한 방법으로 복원된 가시영역 영상의 성능은 평균 24.23dB의 복원 성능을 갖는다. 기존 방법들의 경우 영상의 조도의 차이를 이용해 Illumination map을 생성하여 영상을 복원하는 방식을 사용하였으나, 이

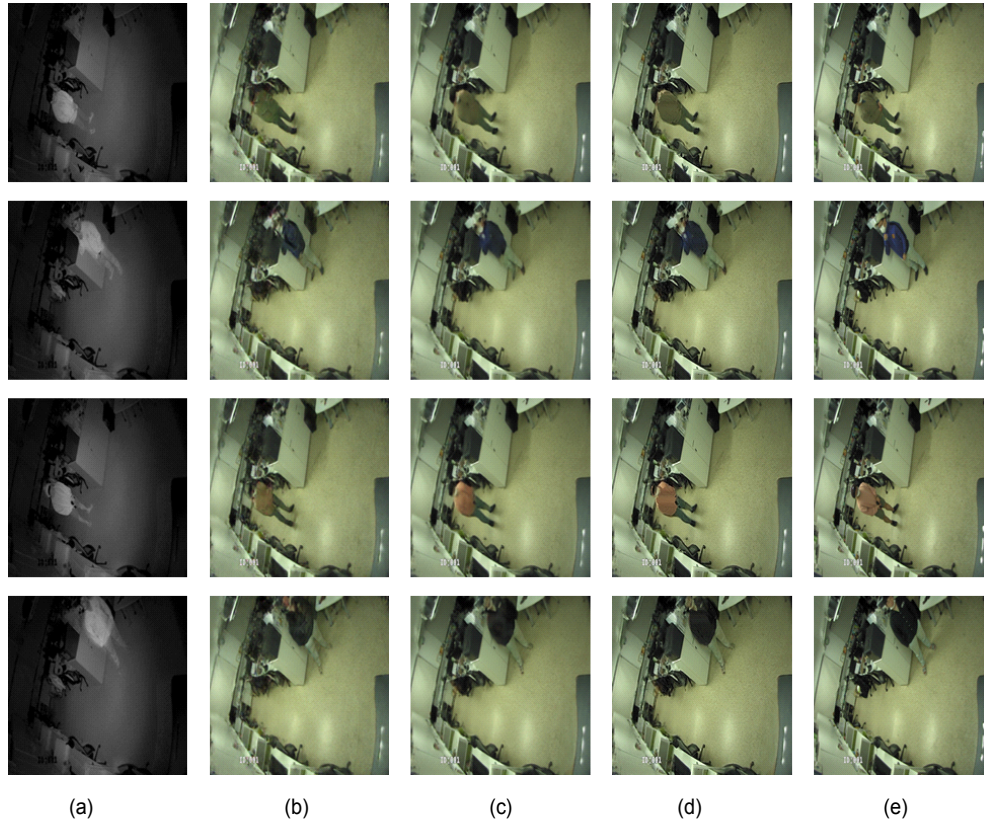


그림 8. DCGAN 알고리즘과 비교 및 잡음 제거 결과 영상 (a) 근적외선 영상 (b) Pix2pix를 통해 생성된 영상 (c) 수정된 DCGAN을 통해 생성된 영상 (d) 잡음 제거 영상 (e) 원본 영상

Fig. 8. Comparison with DCGAN algorithm and image of noise removal results (a) Near-Infrared image (b) Image created via Pix2pix (c) Image created via modified DCGAN (d) Noise-removed image (e) Original image

는 영상의 조도 차이가 적을 경우, 전반적인 복원이 제대로 이루어지지 않는 것을 확인하였다. 이와 달리 본 논문에서 제안한 방법을 통해 복원한 경우, 전체적인 조도가 기존 방법에 비해 향상되고, 그림 8을 확인하면 기존 pix2pix와 달리 전체적인 형상 및 색 정보에 대해서 기존 구조보다 복원이 잘 이루어진 것을 확인할 수 있다.

얻은 근적외선 영상을 이용해 얻어진 색상 영상을 원본과 비슷한 색으로 복원시킬 수 있었다. 하지만 근적외선 영상을 이용해 들어온 물체에 대한 구분이 정확하게 이루어지지 않아 신발의 색과 같은 부분이 정확하게 복원되지 않는 문제도 있었다. 물체를 영역별로 더 정확히 분할하여 잡음을 판별하는 추가 연구가 필요하다.

## VI. 결 론

본 논문에서는 수정된 DCGAN을 이용해 시간을 단축함과 동시에 공간 해상도를 향상시켜 복원한 적정조도 가시 영역 영상에서 색상 잡음을 제거하는 방법에 대한 연구를 진행하였다. 본 논문의 제안방법을 통해 기존에 CCTV에서

## 참 고 문 헌 (References)

- [1] X. Guo, Y. Li, and H. Ling, "LIME: Low-Light Image Enhancement via Illumination Map Estimation," IEEE Transactions on Image Processing, Vol. 26, No. 2, pp. 982-993, Feb, 2017.
- [2] W. Wang, C. Wei, W. Yang, J. Liu, "GLADNet: Low-Light Enhancement Network with Global Awareness," IEEE international

- Conference on Automatic Face & Gesture Recognition, Xi'an, China, pp. 751-755, May, 2018.
- [3] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "EnlightenGAN: Deep Light Enhancement without Paired Supervision," arXiv preprint arXiv:1906.06972, 2019.
- [4] A. Radford, and L. Metz, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," International Conference on Learning Representations, Caribe Hilton, Puerto Rico, pp. 1-16, 2016.
- [5] I. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville, Y. Bengio, "Generative Adversarial Networks," Proceedings of the International Conference on Neural Information Processing Systems, Montreal, Canada, pp. 2672 - 2680, 2014.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," International Conference on Medical Computing and Computer-Assisted Intervention, Munich, Germany, pp. 234-241, 2015.
- [7] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, Hawaii, pp. 2117-2125, 2017.
- [8] G. Huang, Z. Liu, and L. Maaten, "Densely Connected Convolutional Networks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, Hawaii, pp. 4700-4708, 2017.
- [9] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, Hawaii, pp. 2961-2969, 2017.
- [10] G. Oh, J. Lee, and B. Jeon, "Noise Removal in Reconstructed Color Image by GAN for Low Light CCTV Application," International conference on Electronics, information, and Communication 2020, Barcelona, Spain, 2020.

## 저 자 소 개



### 오 교 혁

- 2013년 ~ 2019년 : 성균관대학교 전자전기컴퓨터공학과 학사
- 2019년 ~ 현재 : 성균관대학교 전자전기컴퓨터공학과 석사과정
- ORCID : <https://orcid.org/0000-0002-9048-6002>
- 주관심분야 : 머신러닝, 멀티미디어, 영상신호처리



### 이 재 린

- 2008년 ~ 2015년 : 연세대학교 원주캠퍼스 의공학과
- 2015년 ~ 2017년 : 아이엠 헬스케어 연구원
- 2017년 ~ 현재 : 성균관대학교 전자전기컴퓨터공학과 대학원 석박통합과정
- ORCID : <https://orcid.org/0000-0001-9172-2882>
- 주관심분야 : 라이트필드 영상처리, 계산사진학 기반 영상처리, 머신러닝, 영상신호처리



### 전 병 우

- 1985년 : 서울대학교 전자공학과 졸업 (학사)
- 1987년 : 서울대학교 전자공학과 졸업 (석사)
- 1992년 : Purdue Univ. School of Elec. 졸업 (공학박사)
- 1993년 ~ 1997년 : 삼성전자 신호처리연구소 선임/수석연구원
- 1997년 ~ 현재 : 성균관대학교 전자전기공학부 교수
- ORCID : <https://orcid.org/0000-0002-5650-2881>
- 주관심분야 : 멀티미디어, 영상압축, 영상인식, 신호처리