

# 가상 데이터 증식을 이용한 Pix2pix 기반의 물방울 제거 학습

한유근\*<sup>1</sup>, 정성운\*<sup>2</sup>, 권혁주\*<sup>3</sup>, 이성학\*<sup>4</sup>

## Pix2pix-based Water Drop Removal Learning using Virtual Data Augmentation

Yu-Keun Han\*<sup>1</sup>, Sung-Woon Jung\*<sup>2</sup>, Hyuk-Ju Kwon\*<sup>3</sup>, and Sung-Hak Lee\*<sup>4</sup>

이 논문은 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임  
(NRF-2021R1I1A3049604)

### 요약

본 논문은 딥러닝 기법을 활용하여 영상에서 카메라 렌즈 상에 물방울이 맺힌 부분을 효율적으로 제거하는 영상 변환 방법을 제안한다. 제안 방법에서 다른 도메인의 영상을 새롭게 형성한다는 점에 있어 두 영상의 특징을 잘 파악할 수 있는 GAN(Pix2pix Generative Adversarial Network) 모델을 통해 효과적으로 물방울이 제거된 영상을 얻는 방법을 제안한다. 또한, 촬영 영상을 기반으로 하는 학습 방법에 있어 많은 학습용 데이터가 필요하다는 점과 모델 학습 특성상 불필요한 노이즈가 발생하거나 물방울 사진을 정확하게 얻지 못하는 경우가 발생할 수 있는데, 가상의 물방울 영상 데이터를 생성하고 CNN(Convolution Neural Network)을 이용하여 효과적으로 식별함으로써 학습 데이터를 효율적으로 확보하는 방법을 제안한다.

### Abstract

In this paper, we propose a system that efficiently removes pictures with water droplets on the camera lens using deep learning techniques. We propose a method to effectively obtain a water droplet removal image through the Pix2pix Generative Adversarial Networks(GAN) model, which can understand the characteristics of the two images in terms of newly forming images of different domains. In addition, by applying the fake water droplet data formed by applying the fake water droplet data to the Convolution Neural Network(CNN), paying attention to the fact that a lot of training data is required for image learning, as well as unnecessary noise or water droplet pictures cannot be obtained accurately due to the nature of model learning. We propose a method to efficiently secure data by effectively classifying the formed photos.

### Keywords

deep learning, GAN, CNN, Pix2pix, augmentation learning, rainwater removal

\* 경북대학교 IT대학 전자공학부(\*<sup>4</sup> 교신저자)  
- ORCID<sup>1</sup>: <https://orcid.org/0000-0002-1262-7789>  
- ORCID<sup>2</sup>: <https://orcid.org/0000-0002-2258-3396>  
- ORCID<sup>3</sup>: <https://orcid.org/0000-0002-7665-4049>  
- ORCID<sup>4</sup>: <https://orcid.org/0000-0002-1030-381X>

Received: Mar. 02, 2022, Revised: Mar. 22, 2022, Accepted: Mar. 25, 2022  
Corresponding Author: Sung-Hak Lee  
School of Electronics Engineering, Kyungpook National University, 80  
Daehak-ro, Buk-gu, Daegu, 41566, Korea  
Tel.: +82-53-950-7216, Email: shak2@ee.knu.ac.kr

## I. 서 론

인공지능기반 사물 인식에 대한 관심이 증가함에 따라 이를 이용하여 자율주행차 개발이 활발해지고 있다. 자율주행차의 핵심 기술 중 하나는 운전자의 시각이 아닌 카메라 센서로부터 주행 시야를 잘 확보하는 것이다. 그러나 악천후 등에 의한 기후 변화로 인해 주행경로에 대한 시야확보가 어려워지는 문제가 자주 발생한다. 특히 우천 상황으로 렌즈에 물방울이 맺혀서 전방에 있는 장애물을 인식하지 못해서 생기는 자율주행 사고가 발생할 수 있다. 이는 인명피해도도 이어질 수 있기 때문에 자율주행 기술에 있어 카메라 렌즈에 묻은 이물질은 영상에서 제거하는 기술이 중요하다 할 수 있다.

물방울 제거를 위한 영상 학습을 위해서는 동일 시야에 대해 렌즈에 이물질이 없는 깨끗한 영상과 물방울이 맺힌 영상 쌍의 데이터 셋이 필요하다. 이 두 가지의 영상을 이용하여 물방울 검출 학습을 진행하고 이를 제거하는 작업을 거쳐야 한다. 기존의 GAN(Generative Adversarial Networks)을 이용한 학습은 영상 쌍이 아닌 단일 영상 셋을 요구하기 때문에 원하는 조건을 지정할 수 없어 물방울 제거에 있어 한계가 있다[1]-[3]. 페어(Pair)가 아닌 데이터로 학습을 진행하는 cycleGAN은 Pix2pix에 비해 데이터 셋의 확보가 용이하다는 장점이 있으나 그림 1에서처럼 cycleGAN은 변환 결과 영상에서 영상의 색이나 질감의 변화를 보이고 원하는 물방울 제거 효과가 나타나지 않는 단점을 보인다.

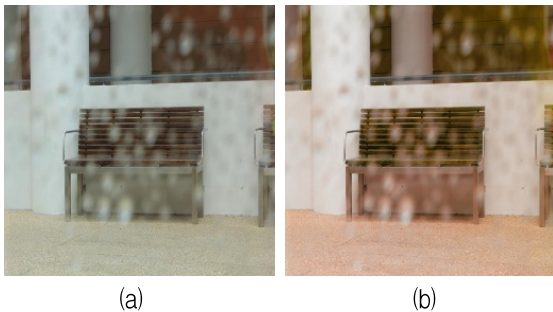


그림 1. cycleGAN을 이용한 물방울 입력 및 제거 결과 영상, (a) 물방울 영상, (b) cycleGAN 물방울 제거 결과  
Fig. 1. Image of water droplet input and removal result using cycleGAN, (a) Water droplet image, (b) cycleGAN water droplet removal result

본 논문에서는 Pix2pix 모델을 기반으로 하여 물방울 제거를 3단계 보완 학습을 통해 개선하는 방법을 제안한다. Pix2pix는 페어 영상으로 학습을 진행하는 모델로서 영상들을 비교함으로써 영상 간 차이점을 잘 파악하는 특징을 가지고 있다[3]. 이는 물방울 제거에 더 유리함을 의미한다. 그러나 충분한 페어 데이터 셋 확보가 어렵다는 단점이 있기 때문에 본 연구에서는 먼저 1차적으로 기존에 확보한 페어 영상을 바탕으로 가상의 물방울 영상을 생성하는 학습을 시행한다. 1차 학습 결과 생성되는 영상들은 물방울이 선명하게 잘 생성된 영상이 있는 반면 그렇지 않은 영상들이 공존한다. 따라서 물방울이 잘 형성된 영상을 효과적으로 분류하기 위해 CNN 기반 학습을 도입하여 물방울이 잘 생성된 사진과 그렇지 않은 사진을 분류한다. 마지막으로 분류된 사진을 다시 기존 사진과 함께 Pix2pix에 학습 영상으로 입력하여 물방울을 제거하는 학습을 보완하여 진행한다.

결과적으로 Pix2pix를 이용한 첫 번째 단계인 물방울 데이터 증식 학습, 두 번째 단계인 CNN을 이용한 분류학습, 마지막 단계인 물방울 제거학습의 총 3단계에 걸쳐 물방울 제거 학습을 진행한 모델은 단순한 물방울 제거 변환 학습만을 거친 결과 영상과 비교했을 때, 영상의 물방울 부분을 더 잘 제거함을 확인할 수 있다.

## II. 관련 연구

### 2.1 물방울 데이터셋 확보

Qian, R 등은 물방울 페어영상을 확보하기 위해 같은 배경을 두고 두 개의 유리를 이용해 데이터 셋을 확보하였다[4]. 유리가 공기와 다른 굴절률을 가지기 때문에 유리를 렌즈에 덧대어서 촬영하여 광선의 오정렬을 피할 수 있도록 하였다. 이때 영상 1은 물방울이 묻어 있는 유리를 사용하였고, 그리고 영상 2는 물방울이 묻어 있지 않은 깨끗한 유리를 사용하여 페어 데이터 셋을 확보하였다. 이 방법의 경우 데이터 셋을 높은 해상도로 확보할 수 있다는 장점이 있지만 일일이 각기 다른 배경에 대해 촬영

을 진행해야 하므로 많은 시간이 소요되는 단점이 있다.

구수한은 알파혼합을 이용해 학습 데이터 셋을 확보하였다[5]. 알파혼합이란 반투명 영상과 배경영상을 결합하여 새로운 혼합영상을 만드는 과정이다. 즉, 물방울 왜곡영상과 일반 배경영상을 혼합하여 배경영상에 물방울이 나타나는 원리를 이용한 방법이다. 그러나 알파 혼합을 통해 생성한 영상에서 학습 시 사용한 물방울 왜곡영상의 잔상이 생성된 영상에 남아있게 되는 단점이 발견된다.

## 2.2 Pix2pix 변환 학습

Pix2pix는 Image to image translation의 한 방법으로서 어떤 영역의 영상을 다른 영역의 영상으로 변환시켜주는 모델이다[3]. Pix2pix는 새로운 영역의 영상을 형성하는 생성자(Generator)와 생성자에 의해 생성된 영상이 진짜 영상인지 가짜 영상인지를 판별하는 판별기(Discriminator)로 구성되어 있다. 생성자의 경우, 일반적인 GAN의 생성자와 동작 방식이 유사하며 입력데이터를 제공함으로써 새로운 영역의 영상을 생성한다. 그러나 GAN이 임의의 노이즈 벡터를 입력으로 요구하는 반면 Pix2pix는 입력 영상을 추가적으로 요구한다.

또한 판별기는 두 가지의 입력 영상을 요구한다. 하나는 생성자에 제공한 영상이고 나머지 하나는 목표 영상 혹은 생성자에 의해 생성된 영상이다. 이 두 개의 페어 영상을 판별기에 제공함으로써 판별기는 제공된 페어 영상을 비교하고 영상이 생성된 영상인지 진짜 영상인지를 구분한다. 다음 수식에서 Pix2pix의 손실 함수를 확인할 수 있다.

$$L_{cGAN}(G,D) = E_{x,y}[\log D(x,y)] + E_{x,z}[\log(1 - D_Y(x,G(x,z)))] \quad (1)$$

$$L_{L1}(G) = E_{x,y}\|y - G(x)\|_1 \quad (2)$$

$$G = \arg \min_G \max_D L_{GAN}(G,D) + \lambda L_{L1}(G) \quad (3)$$

여기서 G와 D는 각각 생성자와 판별기를 의미하며  $G(x,z)$ 는 생성자에 의해 생성된 영상이다.

수식 (1) ~ (3)에서처럼 Pix2pix는 기본적으로 적대적 학습신경망으로 구성된다. 즉, 입력 영상을 넣었을 때 생성된 영상을 판별기가 제대로 판별하지 못하도록 생성기는 학습되며 반대로 판별기는 두 영역의 영상을 제대로 판별할 수 있도록 학습된다. 이때, 식 (1)과 같이 적대적 손실(Adversarial loss)만을 사용하게 되면 영상이 다소 흐리게 보임으로 식 (2)와 같이 L1 loss를 사용한다. 즉, 목표 영상과 생성 영상의 픽셀 거리를 구함으로써 이를 최소화하여 일반적인 영상처럼 생성하도록 하는 역할을 한다. 따라서 식 (1)과 식 (2)를 결합해 최종적인 Pix2pix의 손실함수는 식 (3)과 같이 계산된다.

## 2.3 CNN 분류 학습

CNN 학습모델은 영상을 분석하기 위해 패턴을 찾는데 유용한 알고리즘이다[6]. CNN은 특징을 추출하는 부분인 합성곱 계층(Convolutional layer) 및 이미지를 분류하는 Fully connected layer로 구성되어 있다. 특징 추출 영역인 합성곱 계층은 픽셀별로 특징을 추출하는 필터와 필터의 값을 비선형 값으로 바꾸는 활성화 함수(Activation function)로 구성되어 있다.

CNN에 입력되는 데이터는 높이, 너비, 채널의 3차원 텐서(Tensor)로 구성되어 있고 행렬(Matrix)형태로 나타난다. 하나의 합성곱 계층에는 영상의 채널 개수만큼 필터가 존재하며 각 채널에 필터를 적용함으로써 합성곱 계층의 출력이 결정된다. 일반적으로 필터는  $3 \times 3$ ,  $5 \times 5$  등 정사각행렬의 형태가 사용된다. 각 픽셀에 대해 필터를 적용할 때 그림 2에서 확인되듯이 계산 결과로 피쳐맵(Feature map)을 얻을 수 있다. 피쳐맵의 필터가 지정된 간격으로 순회하며 각 픽셀 값은 Element-wise product 방식으로 계산되어 출력된 결과이다. 이때 필터의 간격은 스트라이드(Stride)를 지정함으로써 변경 가능하다. 그림 2는 스트라이드가 1로 지정되어 계산된 피쳐맵을 나타낸 것으로 필터가 1칸씩 이동하여 픽셀 연산을 실행한다.

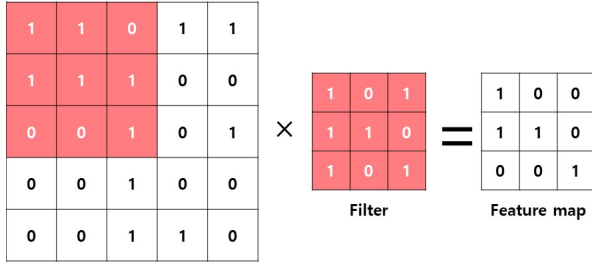


그림 2. 합성곱 계층 필터 적용 동작 방식  
 Fig. 2. Operation method for applying the convolutional layer filter

그러나 위와 같이 합성곱 계층을 거치며 출력되는 영상의 크기는 점점 작아지며 이는 픽셀에 대한 정보가 사라짐을 의미한다. 이러한 문제는 패딩(Padding)을 이용해 해결할 수 있다. 패딩이란 영상의 가장자리에 특정 값을 가진 픽셀을 추가함으로써 입력 영상과 출력 영상의 크기를 비슷하게 조정할 수 있다. 즉, 특정 값을 추가함으로써 합성곱 계층의 출력 영상의 크기가 줄어드는 것을 방지할 수 있다.

추가적으로 CNN은 풀링(Pooling) 계층이 존재한다. 풀링계층은 합성곱 계층의 출력값을 입력으로 받아 데이터의 차원 크기를 줄이거나 특정 데이터를 강조하는 용도로 사용된다. 풀링 레이어는 크게 최대 풀링(Max pooling)과 평균 풀링(Average pooling)으로 구성된다. 최대 풀링의 경우 데이터의 값들 중 최댓값을 모아 표현하는 방식이며 평균 풀링은 데이터의 영역 내 값들을 평균 내어 표현하는 방식이다. 식 (4)와 식 (5)는 합성곱 계층의 출력영상 크기를 나타낸다.

$$O_{height} = \frac{(I_{height} + 2P - KH)}{S} + 1 \quad (4)$$

$$O_{width} = \frac{(I_{width} + 2P - KH)}{S} + 1 \quad (5)$$

여기서  $O_{height}$ 와  $O_{width}$ 는 출력 데이터의 높이와 너비를 의미한다. 또한  $I_{height}$ 와  $I_{width}$ 는 입력데이터의 높이와 너비를 의미하며  $P$ 는 패딩크기를,  $S$ 는 스트라이드 크기를 의미한다. 마지막으로  $KH$ 와  $KW$ 는 각각 필터의 너비와 높이를 의미한다.

컨볼루션 레이어와 풀링 레이어의 출력 값에서 적용되는 활성화 함수는 Sigmoid 함수, Tanh 함수, ReLU 함수 등이 존재한다. CNN 동작시 학습 진행 도중 기울기(Gradient)값이 줄어들어 학습률이 떨어지는 문제가 발생할 수 있다. 따라서 CNN에서는 ReLU 함수를 주로 활성화 함수로 사용한다[7].

또한 CNN에는 Fully connected layer가 존재하는데 이는 기존의 뉴럴 네트워크를 의미한다. 즉, 합성곱 계층에서의 일련의 계산과정이 입력 영상에 대한 정보 및 특징 추출을 위한 것이었다면 Fully connected layer에서는 분류를 위한 계산과정이 진행된다.

마지막으로 데이터를 분류하기 위해선  $n$ 개의 채널로 구성된 합성곱 계층의 출력 결과 값을 1차원의 데이터로 만드는 Flatten 과정을 거치고 이로써 특징 값들을 1차원 데이터로 만든 후 분류를 진행하게 된다.

### III. 제안 방법

본 논문에서는 물방울 제거를 위한 영상 쌍 변환 학습을 위해서 기본적으로 Pix2pix구조를 사용한다. 또한 생성기의 경우 U-net을 사용한다. U-net구조는 기존 인코더-디코더 구조에서 Skip connection을 사용하여 저수준(Low-level)정보들의 소실을 방지하는 기능을 가지고 있다[3]. 판별기의 경우 16x16의 patchGANs을 사용하여 패치 단위로 입력영상이 진짜인지 가짜인지를 판별한다. 패치단위의 판별기는 전체 영상에 대해 판별을 진행하는 판별기에 비해 더 적은 매개변수를 가진다는 특징을 가지며 패치 단위로 판별을 진행하므로 좀 더 자세한 결과를 출력할 수 있다는 장점이 있다.

그림 3은 제안 방법의 흐름도를 나타낸다. 첫번째 학습블록은 1차 학습을 의미하며 초기 학습 영상 쌍을 이용하여 가상 물방울 영상 생성 학습을 실행한다. 두번째 학습블록은 CNN을 통한 영상분류 학습을 나타낸다. 마지막으로 세번째 학습블록은 증식된 학습셋을 이용한 3차 보완된 학습 과정을 나타낸 것이며 분류된 증식 영상 및 기존 데이터셋이 학습에 이용된다.

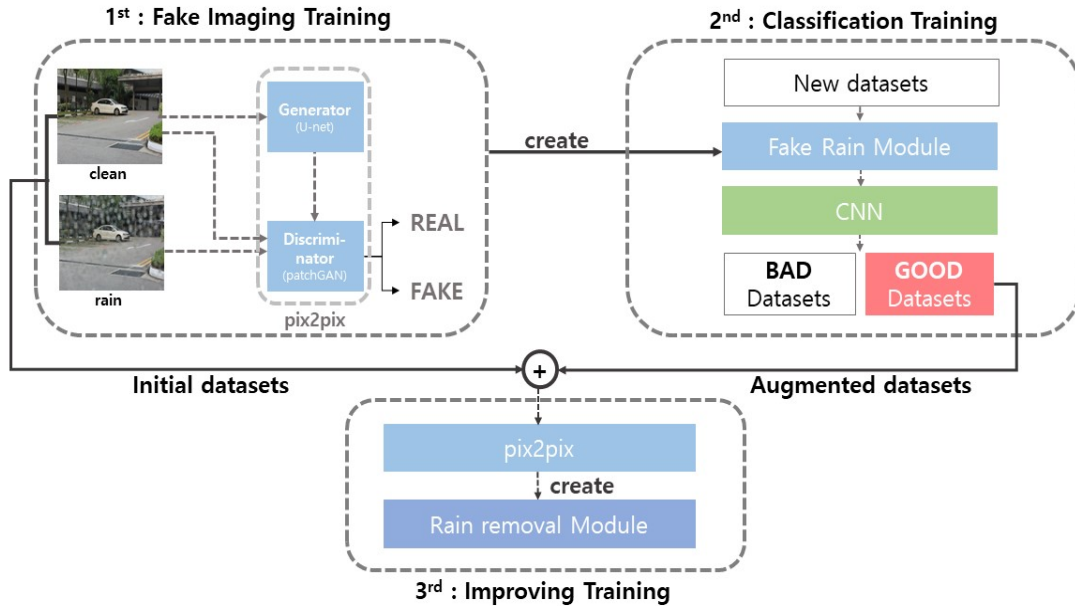


그림 3. 제안 방법 흐름도

Fig. 3. Flowchart of the proposed method

### 3.1 데이터 셋 확보를 위한 1차 증식 학습

딥러닝 네트워크의 경우 적절한 개수의 학습 데이터 셋을 확보해야 한다. 학습 데이터 셋이 부족하면 학습이 제대로 이루어지지 않아 물방울이 잘 제거되지 않고, 부족한 학습 데이터 셋을 이용하여 과도하게 학습할 경우 과적합(Overfitting)이 발생하게 된다. 따라서 다양한 환경이 고려된 충분한 수의 학습 데이터 셋을 확보하는 게 중요하다. 그러나 Pix2pix의 경우 학습 데이터 셋으로 페어 영상을 요구하기에 이를 확보하기가 매우 어렵다. 특히 물방울 제거와 같이 다양한 환경에서 어려 물방울 패턴의 영상 쌍을 직접 촬영하기는 매우 번거롭다. 따라서 이를 보완하기 위해 기존에 확보한 데이터 셋을 활용하여 Pix2pix 학습 모델을 이용하여 가상 물방울을 생성하는 학습을 진행한다. 본 연구에서는 기존 데이터 영상 쌍 200장 및 추가적으로 확보한 일반 도심, 도로 시각 영상 약 6600장에 물방울을 형성시켜 물방울 데이터 셋을 확보하였다.

그러나 앞서 언급했듯, 부족한 데이터 셋으로 인해 물방울 생성결과는 매우 상이한 결과들을 보여 준다. 그림 4에서처럼 특정 영상들은 물방울이 비교적 선명하게 생성된 반면, 특정 영상들은 물방울 생성 효과가 미미하였다.



그림 4. 1차 물방울 생성학습 결과, (a) 물방울이 형성 양호한 경우 (b) 물방울이 형성 안된 경우

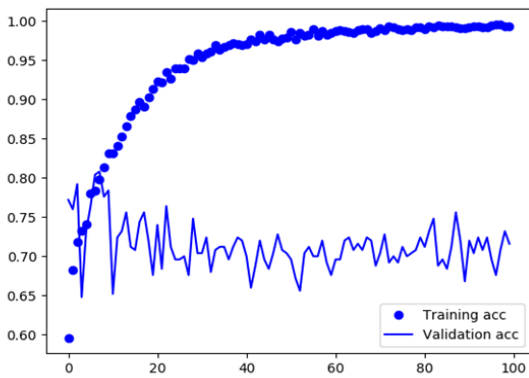
Fig. 4. Results of the first droplet generation training, (a) Image with good water droplet formation, (b) Image with poor water droplet formation

### 3.2 CNN을 이용한 가상 데이터 셋 분류

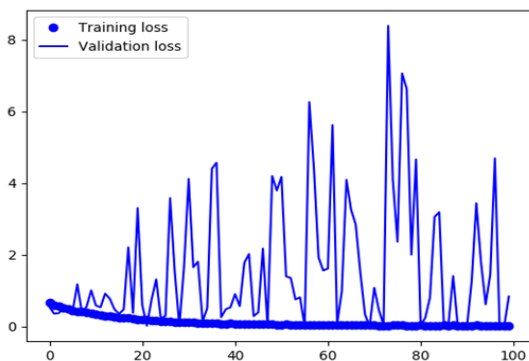
학습 성능 향상을 위해 물방울이 비교적 온전하게 생성된 영상들을 활용해야 한다. 그러나 생성된 모든 영상을 육안으로 판별하는 것은 매우 소모적이다. 따라서 영상 판별에 있어 유용한 딥러닝 네트워크인 CNN을 활용하여 양호한 물방울 형성 사진을 분류한다. 해당 영상 결과를 분류하는 CNN모델을 학습하기 위해서는 1차 증식학습에 사용된 기존 200개 영상 쌍을 사용하였으며, 분류학습에 따라 6600장의 물방울 형성 데이터 중 양호한 영상과 그렇지 않은 영상을 분류하였다. 물방울 영상으로 분

류된 영상 약 1100장에 기존 물방울 영상 200장을 더해 2차 학습에 활용하였고 나머지 5500장을 물방울 미생성 비교군으로 사용하였다. 또한 학습 시 epoch 100으로 설정하였으며 활성화 함수는 ReLU를 사용하였다. 그 결과 그림 5와 같은 학습 수치를 획득하였다.

그림 5(a)는 학습 횟수에 따른 정확도이고 Training accuracy는 학습 데이터 셋의 정확도이며 Validation accuracy는 중간 테스트 데이터 셋의 정확도이다. 두 수치 모두 1에 가까워야 성능이 좋다. 그러나 그림에서 학습 데이터 셋의 정확도는 높으나 이에 반해 테스트 데이터 셋의 정확도는 학습 정확도와 다소 차이가 나는 것을 확인 할 수 있다. 이는 학습데이터 셋의 부족으로 인한 과적합이 발생했음을 알 수 있다. 또한 그림 5(b)는 영상에 대한 손실도를 나타낸 것이다. Training loss는 학습 손실도를, Validation loss는 중간 테스트 손실도를 뜻하며, 0에 가까울수록 성능이 좋다.



(a) 정확도  
(a) Accuracy



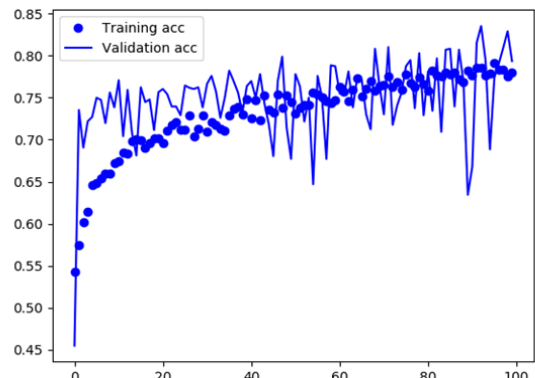
(b) 손실도  
(b) Loss

그림 5. 초기 CNN모델 학습결과

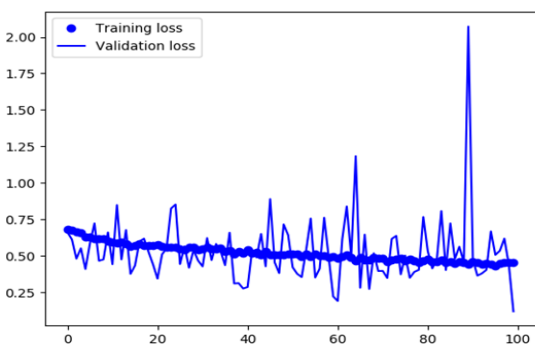
Fig. 5. Training results of initial CNN model

그러나 그림 5(b)에서도 그림 5(a)에서와 마찬가지로 학습 손실도는 0에 수렴하지만 테스트 손실도는 등락이 매우 심한 것을 확인할 수 있다. 이를 개선하기 위해 비교군으로 분류한 영상들을 좌우, 상하 반전 및 일정 각도 회전을 적용해 데이터 영상들을 증식한 후 학습을 진행한다. 그 결과 그림 6과 같은 수치를 얻을 수 있다.

그림 5와 마찬가지로 그림 6(a)와 6(b)는 각각 수정된 정확도와 손실도를 나타낸 지표이다. 비록 학습 정확도 및 학습 손실도는 초기 모델 수치들과 비교했을 때 성능이 약간 감소했지만 테스트 영상에 대한 정확도 및 손실도는 학습 수치 값에 수렴하여 나타나는 것을 확인할 수 있으며, 테스트 정확도의 경우 이전에 비해 약 10% 증가한 약 80%의 정확도를 보였고, 테스트 손실도도 이전에 비해 비교적 안정된 등락을 가지는 것을 확인할 수 있다. 따라서 1차 데이터 증식이 사용된 CNN 모델을 활용하여 물방울 생성 분류를 진행하였다.



(a) 정확도  
(a) Accuracy



(b) 손실도  
(b) Loss

그림 6. 데이터 증식 CNN모델 학습결과

Fig. 6. Training results of data augmentation CNN model

### 3.3 물방울 제거모델 형성을 위한 2차 학습

데이터 증식을 이용한 CNN모델을 생성한 후 해당 모델에 기존에 확보한 영상 쌍 데이터셋 200장 및 CNN모델로 구분된 가상 물방울 생성 영상을 Pix2pix에 입력으로 제공한다. 최종 학습 시에는 물방울 제거가 목표이므로 1차 학습과 달리 영상 쌍의 순서를 바꾸어 입력으로 제공해야 한다. 학습환경은 1차학습과 동일하게 배치사이즈를 3으로 지정하여 학습하고, 활성화 함수는 Hyperbolic tangent를 사용한다. 또한 최종적인 Epoch는 55번으로 지정하여 학습을 실행하였다.

본 논문에서는 200장의 물방울이 없는 영상과 물방울이 있는 영상 쌍을 기본 데이터로 활용하며 새로운 데이터 셋으로 약 6600장의 로드뷰(Roadview) 데이터 셋을 사용하였다[4][8]. 1차 학습의 경우 Pix2pix의 배치 크기를 3으로 지정하며, 400번의 반복횟수를 설정하였다. 2차 학습의 경우 배치크기를 3으로 1차와 동일하게 지정하며, 55번의 반복횟수를 설정하였다. 이 때 학습횟수는 학습 정확도가 포화되는 지점을 사용하였다. 학습에 사용된 CNN 모델은 총 4개의 합성곱 계층을 사용하였으며 기울기 손실 최적화 방법으로 RMSprop을 사용하였다[9].

GAN의 경우 생성 결과에 대한 정량적인 평가 기준이 존재하지 않는다. 따라서 육안으로 기존 200장의 데이터 셋으로 학습한 모델의 물방울 제거 결과를 제안 방법의 결과와 비교 진행하였다.

## IV. 실험 결과



그림 7. 기존 방법 및 제안 방법을 이용하여 물방울을 제거한 영상: (a) 물방울 영상, (b) 기본 데이터 셋을 활용한 학습한 결과, (c) 제안한 학습결과, (d) 목표 영상

Fig. 7. Image of water droplets removed using the existing method and the proposed method: (a) Original images with water droplets, (b) Results of removing water droplets using only the existing dataset, (c) Results of removing water droplets using the proposed method, (d) Target images

그림 7은 확보한 물방울 영상 쌍을 기존 방법 및 제안 방법을 이용하여 물방울을 제거한 결과를 나타내었다. 그림 7(a)는 물방울이 포함된 원본 영상을 나타내었고 그림 7(b), 7(c)는 그림 7(a)를 각각 기존 데이터셋 및 제안 방법을 이용하여 물방울을 제거한 결과를 나타내었다. 마지막으로 그림 7(d)는 물방울이 전혀 묻어있지 않은 깨끗한 원본 영상을 나타낸 것이다. 이때 그림 7(b)와 그림 7(c)의 경우 동일한 학습 횟수로 비교를 진행하였다.

기존 방법과 제안된 방법과의 동일한 55회 학습 횟수로 학습한 결과 비교를 그림 7(b)와 그림 7(c)에서 각 물방울 영상에서 물방울 패턴 제거 정도를 통해 비교할 수 있다. 제안 방법은 기존 방법에 비해 물방울을 더 효과적으로 제거한 것이 확인되며 특히 블러된 백색 점 및 흑색 점 현상이 완화된 것을 확인할 수 있다. 또한 제안 모델의 경우 추가적으로 물방울에 의해 왜곡된 부분을 선명하게 보정된 것을 확인할 수 있었다.

## V. 결 론

본 논문에서는 자율주행 자동차 등에서 활용되는 카메라 및 렌즈 상에 물방울이 묻었을 때 이를 효과적으로 제거하기 위한 영상 변환 처리 방법을 제안하였다. 제안 방법은 Pix2pix 딥러닝 모델을 기반으로 부족한 데이터 셋을 확보하기 위해 기존 데이터 셋을 바탕으로 1차 생성모델을 학습한 후 이를 통해 형성된 물방울을 이용해 기존 데이터 셋과 추가적인 학습을 진행함으로써 해결할 수 있음을 확인하였다. 또한 CNN을 추가하여 생성된 가상 데이터를 효과적으로 분류함을 확인할 수 있고 제안 방법을 이용하여 기존 데이터 셋을 활용한 모델보다 물방울 제거능력이 더 뛰어난 것을 확인하였다. 이는 기존 방법보다 물방울 제거에 더욱 효과적인 방법으로 고려된다.

## References

[1] I. J. Goodfellow, J. Pouget-Abadie, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y.

Bengio, "Generative adversarial nets", *Advances in neural information processing systems* 27, Montreal, Canada, pp. 1-9, Dec. 2014. <https://doi.org/10.48550/arXiv.1406.2661>.

[2] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks", *Proceedings of the IEEE international conference on computer vision*, Venice, Italy, pp. 2223-2232, Oct. 2017. <https://doi.org/10.48550/arXiv.1703.10593>.

[3] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks", *Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, HI, USA, pp. 5967-5976, Jul. 2017. <https://doi.org/10.1109/CVPR.2017.632>.

[4] R. Qian, R. T. Tan, W. Yang, J. Su, and J. Liu, "Attentive generative adversarial network for raindrop removal from a single image", *Proceedings of the IEEE conference on computer vision and pattern recognition*, Salt lake city, Uta, pp. 2482-2491, Jun. 2018. <https://doi.ieeecomputer society.org/10.1109/CVPR.2018.00263>.

[5] S. H. Koo, "Restoring water drop on window using Conditional Generative Adversarial Network", *Doctoral dissertation*, Department of Computer Engineering Graduate School of Chosun University, 2018.

[6] Y. Lecun, L. Bottou, Y. Bengio, and P. Hafner, "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278-2324, 1998. <https://doi.org/10.1109/5.726791>.

[7] C. H. Jang and T. Y. Kim, "Hand Feature Enhancement and User Decision Making for CNN Hand Gesture Recognition Algorithm", *The Institute of Electronics and Information Engineer*, Vol. 57, NO. 2, pp. 60-70, Feb. 2020. <https://doi.org/10.5573/ieie.2020.57.2.60>.

[8] A. R. Zamir and Mubarak Shah, "Image geo-



localization based on multiplenearest neighbor feature matching using generalized graphs", IEEE transactions on pattern analysis and machine intelligence, Vol. 36, No. 8, pp. 1546-1558, 2014. <https://doi.org/10.1109/TPAMI.2014.2299799>.

- [9] S. Ruder, "An overview of gradient descent optimization algorithms", arXiv preprint arXiv:1609.04747, Sep. 2016. <https://doi.org/10.48550/arXiv.1609.04747>.

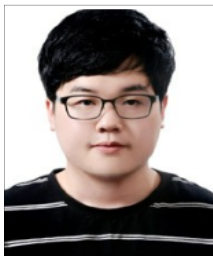
저자소개

한 유 근 (Yu-Keun Han)



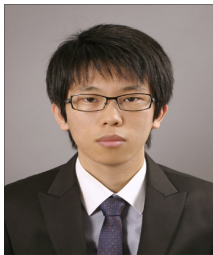
2017년 3월 ~ 현재 : 경북대학교  
IT대학 전자공학부 학부생  
관심분야 : Deep Learning, 컴퓨터  
비전, 신호 처리, 이동 통신

정 성 운 (Sung-Woon Jung)



2019년 2월 : 안동대학교  
전자공학과(공학사)  
2021년 2월 : 경북대학교  
전자공학부(공학석사)  
2021년 3월 ~ 현재 : 경북대학교  
전자전기공학부(박사과정)  
관심분야 : 영상처리, 인공지능

권 혁 주 (Hyuk-Ju Kwon)



2010년 2월 : 경북대학교 전자전기  
컴퓨터학부(공학사)  
2012년 2월 : 경북대학교  
전자전기컴퓨터학부(공학석사)  
2017년 9월 : 경북대학교  
전자공학부(박사후 연구원)  
2019년 8월 : 아진산업 선행연구팀

연구원

2019년 8월 ~ 현재 : 경북대학교 전자공학부 계약교수  
관심분야 : Color Image Processing, Color Management  
HDR 영상처리, 적외선 영상처리, 인공지능

이 성 학 (Sung-Hak Lee)



1997년 2월 : 경북대학교  
전자공학과(공학사)  
1999년 2월 : 경북대학교  
전자공학과(공학석사)  
1999년 2월 ~ 2004년 6월 : LG  
전자 영상제품연구소 선임연구원  
2008년 2월 : 경북대학교

전자공학과(공학박사)

2009년 8월 ~ 2017년 7월 : 경북대학교 IT대학

전자공학부 연구초빙교수

2018년 3월 ~ 현재 : 경북대학교 IT대학 전자공학부 교수

관심분야 : Color Image Processing, Color Appearance  
Model, Color Management, HDR 영상처리, 영상융합,  
인공지능영상처리