

Generative Adversarial Networks를 이용한 Face Morphing 기법 연구

한운 · 김형중*

고려대학교 빅데이터응용및보안학과

Face Morphing Using Generative Adversarial Networks

Yoon Han · Hyoung Joong Kim*

Department of Big Data Application and Security, Korea University

[요 약]

최근 컴퓨팅 파워의 폭발적인 발전으로 컴퓨팅의 한계 라는 장벽이 사라지면서 딥러닝 이라는 이름 하에 순환 신경망(RNN), 합성곱 신경망(CNN) 등 다양한 모델들이 제안되어 컴퓨터 비전(Computer Vision)의 수많은 난제들을 풀어나가고 있다. 2014년 발표된 대립쌍 모델(Generative Adversarial Network)은 비지도 학습에서도 컴퓨터 비전의 문제들을 충분히 풀어나갈 수 있음을 보였고, 학습된 생성기를 활용하여 생성의 영역까지도 연구가 가능하게 하였다. GAN은 여러 가지 모델들과 결합하여 다양한 형태로 발전되고 있다. 기계학습에는 데이터 수집의 어려움이 있다. 너무 방대하면 노이즈를 제거를 통한 효과적인 데이터셋의 정제가 어렵고, 너무 작으면 작은 차이도 큰 노이즈가 되어 학습이 쉽지 않다. 본 논문에서는 GAN 모델에 영상 프레임 내의 얼굴 영역 추출을 위한 deep CNN 모델을 전처리 필터로 적용하여 두 사람의 제한된 수집데이터로 안정적으로 학습하여 다양한 표정의 합성 이미지를 만들어 낼 수 있는 방법을 제시하였다.

[Abstract]

Recently, with the explosive development of computing power, various methods such as RNN and CNN have been proposed under the name of Deep Learning, which solve many problems of Computer Vision have. The Generative Adversarial Network, released in 2014, showed that the problem of computer vision can be sufficiently solved in unsupervised learning, and the generation domain can also be studied using learned generators. GAN is being developed in various forms in combination with various models. Machine learning has difficulty in collecting data. If it is too large, it is difficult to refine the effective data set by removing the noise. If it is too small, the small difference becomes too big noise, and learning is not easy. In this paper, we apply a deep CNN model for extracting facial region in image frame to GAN model as a preprocessing filter, and propose a method to produce composite images of various facial expressions by stably learning with limited collection data of two persons.

색인어 : 대립쌍 기계학습, 얼굴합성, 합성곱 신경망, 비지도 학습

Key word : Generative adversarial network, face morphing, DCGAN, dCNN, unsupervised learning

<http://dx.doi.org/10.9728/dcs.2018.19.3.435>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 26 November 2017; Revised 20 December 2017

Accepted 25 March 2018

*Corresponding Author; Hyoung-Joong Kim

Tel: +82-2-3290-4895

E-mail: khj-@korea.ac.kr

1. 서론

기존의 얼굴 합성(face morphing)은 특징점을 기반으로 이루어졌다[1]. 각 이미지에서 특징점들을 추출하고, 두 이미지에서 얼굴로 인식 되는 부분의 평균 벡터를 통해서 얼굴 이미지를 합성하는 방법이 일반적이다. 얼굴 합성은 포토샵 같은 이미지 편집툴을 이용한 수작업 생성도 가능하다. 하지만 이러한 방법들은 각각의 원본이 존재하며, 결과물이 단일 이미지라는 한계가 있다.

2014년에 Ian Goodfellow는 원본을 복제하는 생성기와 그것을 분류하는 분류기가 서로 적대적으로 대립하여 학습하며 성능을 끌어올리는 GAN(Generative adversarial network)[2]을 제안하였는데, 이는 기계학습 연구의 경향을 지도학습기반(CNN, RNN 등)에서 비지도 학습 및 반지도학습(semi-supervised)[3]으로 크게 바꾸어 놓았다.

기계 학습의 연구에는 데이터 셋 확보의 어려움이 존재한다. 지도학습에서 데이터가 너무 방대하면 데이터 셋의 분류도 어렵고, 노이즈로 인해 유의미한 학습이 어려운 경우가 많다. 확보된 데이터를 일일이 분류하고, 노이즈를 제거하는 일은 시간과 비용이 많이 든다. 반대로 데이터가 작으면 과적합의 문제를 피하기가 어렵다. GAN은 랜덤 노이즈를 통해서 이미지를 변형하여 생성하고, 분류기가 원본과 비교하며 분류기의 분류 확률, 생성기의 복제확률을 미니맥스(minimax) 알고리즘을 통해 학습하므로 이러한 문제점에 좋은 성능을 보인다.

기존의 기계 학습 방법들이 분류, 인식, 예측의 문제를 풀었다면, 분류기와 생성기를 동시에 학습하는 모델인 GAN은 분류뿐만 아니라 “생성”의 문제도 풀 수 있게 되었다. 이러한 특성은 수목화를 재현해 내거나[4], 음악을 만들어내는[5] 등 다양한 분야에서 활용 또는 연구되고 있다.

본 연구에서는 2명으로 제한된 데이터 셋으로 GAN 모델을 통해 유의미한 특성을 찾아내고, 이를 통해서 다양한 형태의 합성 이미지를 생성해 내는 것을 목적으로 한다. 데이터 셋은 정면의 얼굴이 촬영된 10분가량의 영상을 활용한다. 각각 2명의 정면 얼굴을 포함하되 표정을 변화를 주어 촬영한다. 자연스러운 표정의 변화를 학습함에 있어서 과적합되는 것을 방지하기 위하여 연속 프레임에 수직하지 않고, 매 초당 랜덤하게 4개의 연속되지 않은 이미지를 추출한다. 기존의 연구에서는 방대한 데이터 셋을 기반으로 하였고, 과적합을 방지하기 위하여 유사한 이미지들을 제거 하였으나[6], 본 연구에서는 유사성을 표정의 변화, 입꼬리, 눈모양 변화 등에 따른 얼굴 특성 학습을 목표로 하여, 연속 프레임 제거로 선형적인 표정 변화를 막는 것으로 과적합을 방지한다. 각각의 인물별로 학습하여, 부드러운 표정의 변화를 그려내는지 여부를 통해 학습 여부를 판단한다. 마지막으로 2명의 데이터를 동시에 학습하여, 한사람의 이미지를 기준으로 다른 사람의 이미지로 부드러운 변화를 그려내는지 확인하고, 이를 통해서 다양한 형태의 합성 이미지를 획득한다.

본 논문은 다음의 순서로 구성된다. 2장에서는 대립쌍 모델

과 관련된 선행 연구들을 확인하고, 3장에서는 본 연구에 앞선 사전실험 및 결과를 살펴보고, 4장에서는 사전 실험을 통해 확인된 방법으로 데이터 셋을 재조정하여 학습을 진행하고, 결과를 분석하고, 마지막으로 5장에서는 연구 결과를 한계점, 향후 연구 방안을 논의한다.

II. 선행 연구

2-1 Convolutional Neural Network

CNN(convolutional neural network)은 오래전부터 연구된 모델이었으나, 현재의 표준 형태를 처음 갖추게 된 것은 1998년이다[7]. CNN은 컨볼루션 레이어(convolution layer)와 풀링 레이어(pooling layer) 두 가지의 핵심 구조로 이루어진다. 컨볼루션 레이어는 일련의 학습 가능한 필터들로 이루어져 있다. 각 필터는 입력 볼륨의 가로/세로 차원으로 슬라이딩하며 필터와 입력 요소들 사이의 내적연산을 실행한다. 이를 통해서 입력데이터의 특징을 추출한다. 풀링 레이어는 입력 공간을 추상화하는 역할을 담당하고, sub-sampling을 통해서 차원을 축소한다. 일반적인 신경망 구조가 이미지를 다루기에 적절하지 않은 것에 비해 CNN은 이미지에 적합하게 3차원의 레이어로 구성되어 있다. 이로 인해 사물을 인식하거나, 얼굴을 인식 하는 등 이미지에서 특징을 찾아내고 인식하는데 많이 쓰이며[8], 인식된 형태에 대한 유사성을 비교하는 방법으로 쓰이기도 한다[16].

본 연구에서는 GAN 학습을 위한 이미지 전처리 과정에서 눈과 입 등의 얼굴 기준 특성을 인식하기 위해 사용되었다.

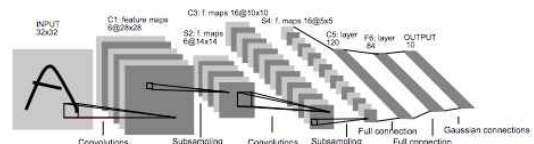


그림 1. LeNet-5의 구조(convolution neural net)
Fig. 1. Architecture of LeNet-5, convolution neural net

2-2 Deep Neural Network

인공 신경망 구조에서 입력 레이어와 출력 레이어 사이에 존재하는 은닉층(hidden layer)이 다중으로 이루어진 심층 신경망 구조를 말한다. 흔히 딥러닝[15] 이라고 이야기하는 상당수의 모델이 이 심층 신경망 구조이다. 컴퓨팅 성능의 한계와, 과적합의 문제를 해결 할 수 없어서 주목 받지 못하였으나, 2000년대 들어서 드롭아웃, ReLU(Rectified Linear Unit)등의 과적합을 줄이는 방법론이 등장하면서 다시 발전되기 시작하였다. 과적합을 방지하는 방법들은 현재도 활발하게 연구 되고 있다[9].

CNN의 구조에서 은닉층을 다중으로 구성하여 DNN구조를 가지게 되면 deep CNN이라 통칭한다. 뒤에서 살펴볼 DCGAN은 GAN에 이 deep CNN을 적용한 모델이다.

본 연구에서는 GAN 학습을 위한 이미지 전처리 필터에서 미리학습된 deep CNN 모델이 사용되었다.

2-3 Generative Adversarial Networks

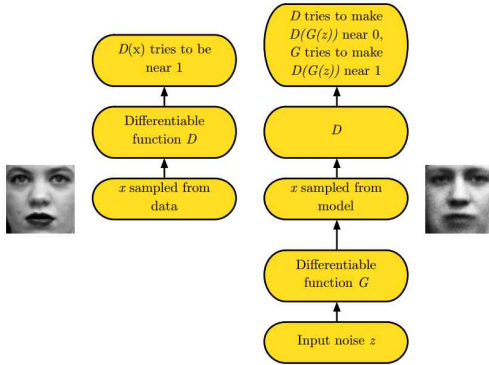


그림 2. GAN 모델의 framework
Fig. 2. Generative adversarial network framework

GAN은 2014년 Ian Goodfellow에 의해 처음으로 제안된 모델이다[2]. GAN은 생성자(generator)와 판별자(discriminator) 두 가지로 구성되어 있고 두 가지 모델이 서로 적대적으로 경쟁하며 학습하여 학습 성능이 높이는 비지도학습 기반의 학습 모델이다.

$D(x)$ 는 판별자가 입력된 이미지가 원본인지 가짜인지를 판별해낼 확률이다. $D(x)$ 가 1인 것은 진짜는 진짜로, 가짜는 가짜로 완벽히 분류 해낸다는 의미가 된다. $G(z)$ 는 생성자가 입력잡음(uniform random) z 를 입력 받아 원본을 z 샘플링한 결과이고, $D(G(z))$ 는 생성자가 생성해낸 이미지를 판별자가 분류해낼 확률이다. 생성자는 $D(G(z))$ 가 1에 가까워지도록, 다시 말해 생성해낸 가짜를 판별자가 진짜라고 분류하는 것을 목표로 학습된다. 판별자는 그림에도 불구하고 원본과 가짜를 분류해 낼 수 있도록 학습된다. 모델을 수식으로 표현하면 아래와 같다.

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

생성자와 판별자의 value function $V(D, G)$ 에 대한 미니맥스 문제를 푸는 것으로 모델의 성능이 향상된다.

학습이 반복됨에 따라 원본과 생성자의 분포는 점점 비슷해지다가 완벽히 학습이 된 후에는 일치하게 된다. 그때 판별자가 원본인지 가짜인지 가려낼 확률은 0.5로 유지된다[그림 3].

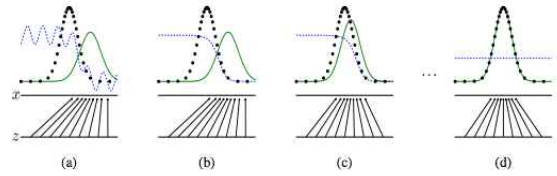


그림 3. GAN의 학습 시간에 따른 변화
Fig. 3. Change according to learning epoch of GAN

GAN은 지도학습 중심의 기계학습 흐름을 비지도 학습으로 가지고 오는데 큰 역할을 하였고, 분류기와 생성기를 동시에 학습하는 구조 덕분에 분류기의 학습 결과로는 분류의 문제를, 생성기의 학습 결과로는 생성의 문제까지도 풀 수 있게 되었다. 하지만 여전히 학습과정과 구현이 까다롭고, 미니맥스를 푸는 과정이 안정성이 떨어진다는 단점이 존재한다.

2-4 DCGAN

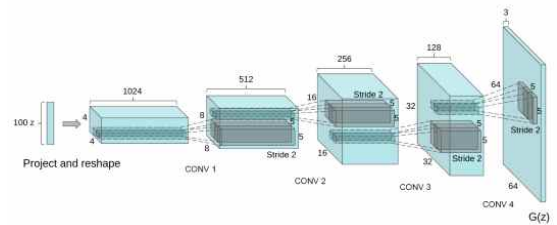


그림 4. DCGAN의 generator 구조
Fig. 4. Deep generator architecture

DCGAN(Deep convolutional GAN)은 비지도 학습인 CNN을 적용한 모델이다. 기존의 GAN에서 단점으로 지적 받던 안정성을 크게 개선하여 대부분의 환경에서도 안정적인 훈련을 제공한다라고 밝히고 있다[10].

DCGAN은 GAN의 기본 형태에 다층 은닉층을 둔 deep convolutional network를 적용한 모델이다. 또한 기존의 CNN 구조에 대해서도 3가지의 변화를 주었다. 즉, all convolutional net 사용[11], fully connected layer 제거, batch normalization의 사용이다. 덧붙여 DCGAN은 rectified linear unit을 사용하였다.

실험 데이터 셋으로 수십만 장의 방안의 풍경 사진과, 35만 장의 인물 데이터를 사용하였고, 실험을 통해서 DCGAN 모델이 학습을 통해서 이미지의 형태 속에서 유의미한 특성을 찾아냈다. Gaussian 잡음 벡터를 잡음 입력 Z 로 이용함으로써 학습된 결과에 z 산술이라는 새로운 방법도 제시하였다[그림 5]. “안경 쓴 남자” 학습결과에 “남자” 학습결과를 빼고, “여자” 학습결과를 더하면 “안경 쓴 여자”를 생성해낸다. 학습모델이 안경, 남자, 여자를 각각 학습하였음을 알 수 있다.

최근 연구되는 대부분의 GAN 연구모델들은 DCGAN 을 기반으로 연구되고 있다, 본 연구도 DCGAN을 기반으로 한다.

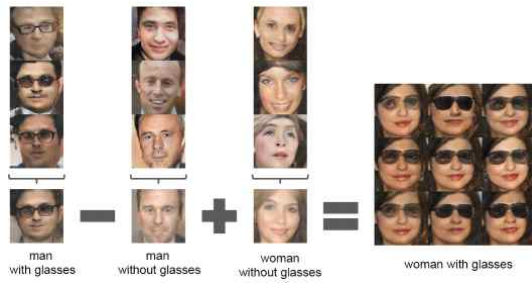


그림 5. DCGAN의 z 연산을 이용한 이미지 생성
 Fig. 5. Image generation using z arithmetic of DCGAN

III. 사전 실험

본 연구에 앞서, 임의의 데이터 셋 및 조건으로 실험을 진행하였다. 실험의 목적은 본 연구에 앞서 다음의 사항들을 확인하기 위함이다.

- 1) 잡음이 심한 데이터 셋의 학습은 어떤 결과를 보이는가?
- 2) 실패한 학습의 학습 진행은 어떤 양상을 보이는가?
- 3) 영상 중 얼굴이 포함된 프레임이 그대로 넣었을 때, 얼굴의 위상차에도 불구하고 학습이 되는가?

1과 2는 본 연구에서의 학습이 노이즈에 의한 학습 실패인지 여부를 비교하기 위함이다. 3의 결과를 살펴보고 본 연구의 실험 데이터의 설정 및 연구 모델을 설정하고자 하였다.

1의 실험에서 실험 데이터 셋은 드론이미지로 설정하였다. 날개형 드론, 쿼드형 드론, 원형 드론 등 형태가 굉장히 다양하기에 노이즈가 큰 데이터라고 판단하였다. 데이터의 수집은 Image-net과 Amazon을 크롤링하여 약 3,000개의 샘플을 수집하였다. 실험은 64개의 배치(batch) 형태로 원본을 정렬하고, 생성자도 입력 잡음에 맞추어 64개의 배치 형태로 생성된 샘플을 배치하여 일정한 주기로 출력 이미지로 생성하여 학습 상태를 이미지로 확인할 수 있도록 하였다.

실험의 결과 학습모델은 3,000개의 노이즈가 심한 데이터 셋으로는 어느 정도의 학습 성과는 보였으나, 선행 연구가 보여준 성능을 보이지 못했다. 생성자가 만들어낸 이미지는 어느 정도 드론의 형태를 닮아 있기는 하지만, 똑같은 패턴의 반복, 혹은 흐릿해진 영역 등이 많다, DCGAN 페이지에 따르면 이는 언더피팅이 일어난 모습이다[10].



그림 6. 원본과 학습결과로 생성된 드론이미지의 비교
 Fig. 6. Comparison of original and generated image

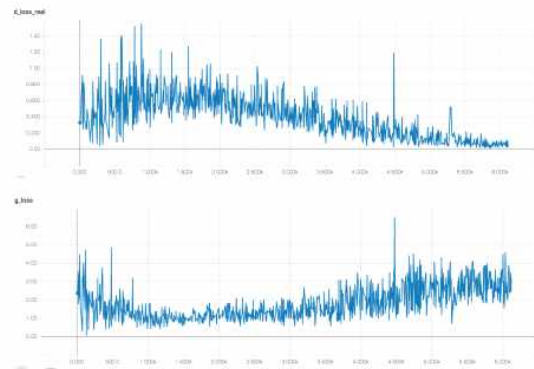


그림 7. 드론 학습의 반복에 따른 g-loss, d-loss 그래프
 Fig. 7. g-loss, d-loss graph of dataset 'drone'

GAN에서 학습 상태를 나타내는 지표로 loss function을 사용하는데, 이 loss function 각 모델의 성취도, 성능을 나타낸다. 분류기가 실제 이미지를 잘 구분 하였을 때를 의미 하는 1과 실제 이미지의 구분 결과와의 차이를 real-loss, 가짜 이미지와 0과의 차이를 fake-loss로 표현한다. 판별자의 loss 값은 d-loss 라고 표현하고 fake-loss와 real-loss의 합으로 표현한다. 생성자의 loss 값은 g-loss라고 표현하고 fake-loss를 쓰되, 판별자와 반대로 가짜 이미지와 1과의 차이로 표현한다.

앞선 실험의 d-loss와 g-loss 변화의 양상을 살펴보면 y축의 진동 폭이 굉장히 크고, d-loss는 학습이 반복될수록 0에 가까워지며, g-loss 는 학습이 반복될수록 증가한다. 정상적인 학습 상태에서는 d-loss와 g-loss가 일정 학습기간 동안 낮아지고 둘 다 안정된 그래프를 보여준다. 한 값이 0으로 수렴하거나, 급격히 꾸준히 상승하는 것은 학습 실패를 나타낸다.

다음으로는 영상에서 추출한 프레임을 대상으로 얼굴이 있는 영역을 임의로 잡고 잘라 사용하였다. 남성, 여성 각각 1인씩의 영상으로 데이터를 수집하였다[그림 8].

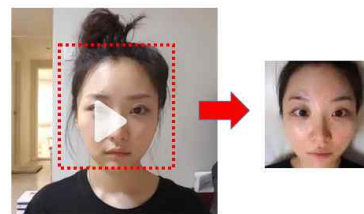


그림 8. 임의로 설정한 영역의 이미지 추출 샘플
 Fig. 8. Image extraction sample of arbitrary rectangular area

남자와 여자 샘플의 학습 결과가 큰 차이를 보였다[그림 9]. 원본 데이터 셋을 검증해본 결과, 남성의 데이터 셋은 얼굴 위치에 대한 위상차이가 크지 않았던 것에 비해, 여성의 데이터 셋은 얼굴 위치와 각도 등의 위상차이가 크게 나타났다.



그림 9. 남, 여 데이터 셋의 학습 결과 비교
 Fig. 9. Comparison of learning results for each dataset

학습에 실패한 여성 dataset을 1000번 이상 반복학습하여 d-loss, g-loss 그래프를 추출 하여 그래프의 경향을 살펴보았다. drone dataset 과 비교하여 y축의 진폭은 크지 않았지만, 학습이 반복 될수록 d-loss 는 0으로 수렴하고, g-loss 는 상승하는 경향은 비슷함을 확인 할 수 있었다[그림 10].

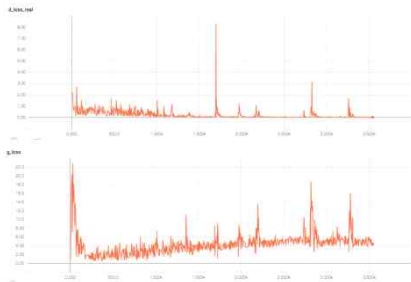


그림 10. 여성 데이터셋 학습의 g-loss, d-loss 그래프
 Fig. 10. g-loss, d-loss graph of dataset 'female'

사전 실험의 결과로, 제한적인 학습데이터를 대상으로 학습을 할 때 형태의 변화가 너무 다양하거나, 학습하고자하는 목표 형태의 평균적인 위치가 비슷하지 않으면, 이러한 특성들이 데이터 셋의 노이즈가 되어 학습이 잘 되지 않음을 알 수 있었다. 효과적인 학습을 위하여 데이터 셋을 어느 정도의 기준에 맞추어 정렬할 필요가 있음을 확인하였다.

IV. 본 연구

4-1 연구모델

사전 연구에서 보았던 데이터 셋의 노이즈 제거를 위하여 dCNN을 이용한 전처리 필터가 추가된 GAN의 학습 모델을 설정하였다[그림 11].

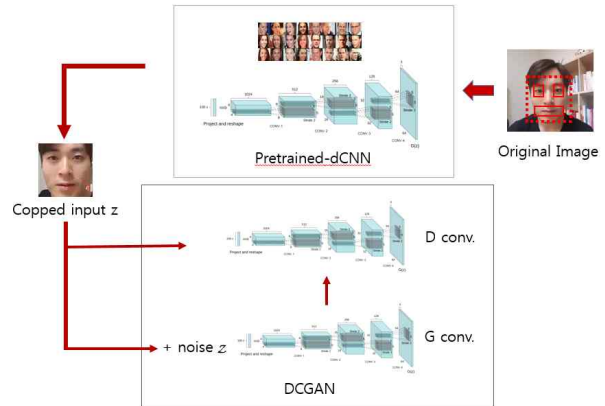


그림 11. 얼굴 인식 전처리과정이 추가된 학습모델
 Fig. 11. Learning model with a face recognition preprocess

dCNN으로 얼굴의 각 부분(part)을 인식[13]하고 중심점을 기준으로 일정한 크기만큼의 이미지를 잘라내는 과정을 전처리 필터로 설정하였다. dCNN이 얼굴의 각 부분을 인식하기 위해서는 사전 학습의 과정이 필요하다. 사전 학습을 위해서 데이터 셋은 Labeled Faces in the Wild [14]의 데이터 셋을 이용하였다. 얼굴 인식을 통하여 정렬(aligned)된 이미지는 DCGAN모델로 학습한다.

4-2 학습 데이터의 수집

데이터의 사전 실험과 같은 방법으로 수집하였다. 특정한 사람의 이미지를 효과적으로 추출하는 방법으로 동영상 촬영을 선택하였다. 24프레임 동영상 기준으로 초당 24장의 사진을 수집할 수 있다. 하지만 연속된 사진 수집은 표정 변화 등 모든 과정이 수집되므로 과적합을 피하기 위하여, 연속되지 않은 사진을 초당 4개씩만 수집한다. 동영상은 남성과 여성 각각 1명으로 한정하여 수집하였다. 10분가량의 영상 촬영을 통해 각 2,500장씩 총 5천장의 이미지를 확보하고 실험을 진행하였다. 20~100만장 정도의 데이터 셋으로 학습하는 일반적인 학습 방법과 비교하면 5,000장의 이미지도 작은 데이터 셋이다. 하지만 5천장 실험 이후 조금씩 데이터 셋을 더 줄여 나가면서 실험을 진행하였다.

4-3 실험 목적 및 방법

정렬(aligned)된 이미지 수집을 위하여 dCNN으로 얼굴 인식을 하고, 인식점을 기준으로 노이즈가 적은 이미지 수집이 가능하다. 노이즈나 변화가 심한 이미지는 더욱 많은 학습데이터를 필요로 하지만, 본 연구에서는 “얼굴 합성”으로 목적이 뚜렷하기 때문에, 얼굴 인식을 통한 데이터 수집은 학습데이터를 통해서 GAN 모델이 얼굴의 표정의 변화, 두 인물간의 차이점에 집중할 수 있게 해준다.

학습이 완료되면 학습된 데이터 z 에 대하여 생성기를 이용하여 학습된 결과 이미지를 생성해낸다. 이미지가 잘 생성되었는지 판단하기 위하여 3가지 요소를 확인한다.

- 1) 각 인물의 얼굴에 대하여 표정의 변화가 학습된 데이터 z 의 변화에 따라 자연스럽게 변화하는가?
- 2) 한 인물에서 다른 인물로 변화 하는 학습된 z 가 존재하며, z 의 변화에 따라 자연스럽게 변화하는가?
- 3) 여러 차례 반복된 학습을 통해서 여러 가지의 표정의 인물 합성 이미지가 생성 되는가?

4-3 실험결과

첫 번째로, 남성 여성 각각에 대하여 따로 DCGAN 학습을 수행하였다. 각각 수행한 학습의 결과로 원본 이미지에 대하여 다양한 표정의 변화를 학습하였음을 확인하였다. 학습된 데이터는 입력 z 를 0에서 1까지 일정하게 증가시킨 값에 따라 생성기를 이용하여 이미지로 생성하였다. 입력 z 의 변화에 따라 부드러운 표정의 변화를 보여주고 있다[그림 12]. 선행연구에 따르면 이러한 변화의 형태를 walking in the latent space라고 표현한다[10].



그림 12. z 의 변화에 따른 부드러운 변화로 생성된 이미지
Fig. 12. Image generated with smooth change by z

다음으로는 남성과 여성을 혼합하여 DCGAN 학습을 수행하였다. 학습데이터는 앞에서 밝힌 바와 같이 각각 2,500장씩 5,000장의 이미지로 학습하였다. 실험의 결과로 입력 z 의 변화에 따라 남성에서 여성으로 혹은 여성에서 남성으로 부드러운 변화를 보여주었다[그림 13]. 이 과정을 통해 학습이 성공적으로 이루어졌음을 확인할 수 있었다. 또한 중간 정도의 z 값으로, 두 인물이 혼합된 합성 이미지를 얻을 수 있었다. 잘 정렬된 학습데이터로 동시에 학습하기만 하는 것으로 face morphing의 결과물을 얻을 수 있다. 간단히 z 를 조작하는 것만으로 두 데이터 분류 중 어느 쪽에 가까운 결과물을 얻을 것인지도 선택할 수 있다. 비슷한 표정의 유사성을 학습함에 따라 다양한 표정의 face morphing 결과물도 얻을 수 있었다[그림 14].

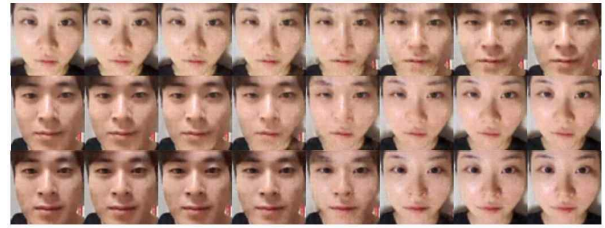


그림 13. 두 학습데이터가 부드럽게 혼합된 이미지
Fig. 13. Smoothly mixed images of both training data

사전 조사에서 d-loss가 0으로 수렴하고, g-loss가 지속적으로 상승하던 데이터 셋을 dCNN 을 통한 정렬(align) 이후의 변화를 살펴보면, d-loss는 다소 높은 진동 폭을 보이지만, 0.3-1 사이에 분포가 높은 것을 확인할 수 있다. g-loss 값은 학습 횟수가 늘어나도 크게 변하지 않고 안정적인 양상을 보여주었다[그림 15]. 이것은, 학습이 안정적으로 이루어지고 있음을 보여준다.

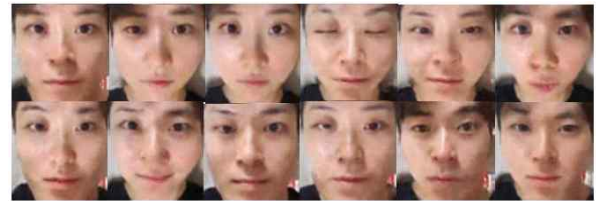


그림 14. 학습을 통해 생성된 다양한 표정의 합성이미지들
Fig. 14. Generated images of various facial expressions

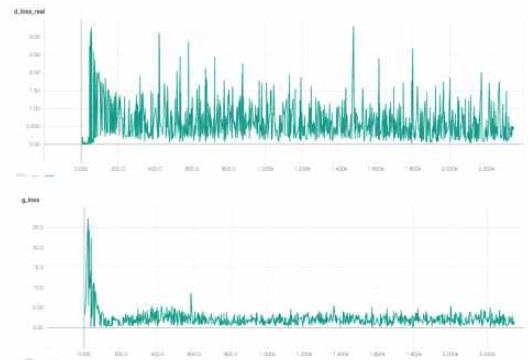


그림 15. 혼합 학습데이터의 g-loss, d-loss 그래프
Fig. 15. g-loss and d-loss graph of 'mixed' dataset

마지막으로 데이터 셋의 수를 줄이며 실험하였다. 데이터 셋의 수를 500개로 줄였을 때도 안정적인 그래프를 보여주었다. 극단적으로 100개로 줄였을 때도 안정적인 그래프를 보여주었다[그림 17]. 생성기로 생성된 이미지는, 비록 원본을 충분히 학습하지 못하여 품질은 떨어지지만[그림 16], z 의 변화에 따른 자연스러운 이미지의 변형은 학습한 것을 볼 수 있다.



그림 16. 100개의 데이터로 학습된 여성 이미지
Fig. 16. Generated Images of women with 100 sample

IV. 결론

본 연구에서 GAN 학습모델로 동영상 촬영이라는 간단한 방법으로 수집된 두 사람의 데이터 셋으로 학습하여 합성이미지를 생성해낼 수 있음을 확인하였다. GAN의 학습데이터를 통한 이미지 합성은 이미지 학습에 대한 전문적인 지식이나 기술 없이도 가능하며, 학습데이터의 다양성에 따라 다양한 합성 이미지를 얻을 수 있음을 보여주었다.

또한, 실험을 통해서 GAN 학습 시 형태의 심한 변화가 있거나, 위상의 큰 차이가 있는 경우 제한된 데이터에서는 학습이 되지 않음을 확인하였고, 이를 보완하기 위해서 dCNN 기계학습으로 사전 학습된 데이터를 이용하여 원본 사진의 학습 대상 영역을 추출하는 방법을 적용하였다. 즉, 학습의 앞 단계에서 선행학습된 데이터를 이용하여 기준이 될 영역을 잡고, 그에 맞추어 이미지를 정렬함으로써, GAN의 학습 모델이 학습하기 위한 기준이 존재하고, 데이터가 그 기준에 맞게 잘 정렬되어 있으면 제한된 데이터 안에서도 충분히 좋은 성능을 낼 수 있음을 보였다.

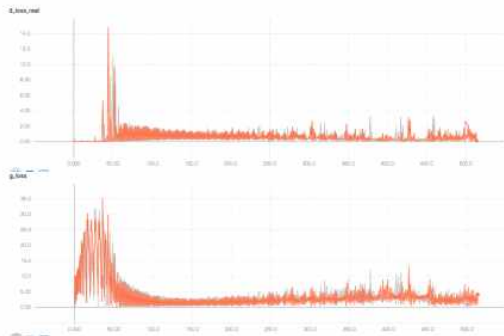


그림 17. 100개 데이터의 g-loss, d-loss 그래프
Fig. 17. g-loss and d-loss graph of 100 sample

V. 연구의 한계 및 향후 연구 방향

본 연구의 한계점은 다음과 같다. 첫째, 합성된 이미지를 예측할 수 없다. GAN의 비지도학습 특성으로 인하여, 학습모델이 데이터 셋을 통하여 자유도 높은 샘플들을 생성해내지만, 원하는 형태로 생성해내기는 힘들다. 그로 인하여, 결과물에 대한 목표를 세우기가 힘들다. 두 번째, 데이터 셋 정렬을 위하여 또 다른 기계학습이 선행되어야 한다는 것이다. GAN학습모델의

학습 피드백으로 원본의 기준을 세우고 또 그에 따라 정렬해내며 학습할 수 있다면 더 좋은 결과를 이끌어 낼 수 있을 것으로 기대된다.

참고문헌

- [1] J. Areeyapinan, and P. Kanongchaiyos, "Face morphing using critical point filters," in *Proceedings of International Joint Conference on Computer Science and Software Engineering*, pp. 283-288, 2012.
- [2] I. J. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," in *Proceedings of the Neural Information Processing Systems*, pp. 2672-2680, 2014.
- [3] R. Tachibana, T. Matsubara, and K. Uehara, "Semi-supervised learning using adversarial networks," in *Proceeding of the International Conference on Computer and Information Science*, pp. 1-6, 2016.
- [4] M. S. Ko, H. K. Roh, and K. H. Lee "GANMOOK: Generative adversarial network to stylize images like ink wash painting," in *Proceedings of the Korea Computer Congress*, pp. 793-795, 2017.
- [5] L. C. Yang, S. Y. Chou, and Y. H. Yang "MidiNet: A convolutional generative adversarial network for symbolic-domain music generation," in *Proceedings of the 18th International Society of Music Information Retrieval Conference*, pp. 324-331, 2017.
- [6] A. Radford, L. Metz, and S. Chintala "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proceedings of the International Conference on Learning Representations*, pp. 1-15, 2016.
- [7] Y. Le Cun, L. Bottou, Y. Bengio, and P. Haffner "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
- [8] S. Lawrence, C.L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Transactions on Neural Networks* vol. 8, no. 1, pp. 98-113, 1997.
- [9] X. Sun, X Ren, S Ma, and H. Wang, "meProp: Sparsified back propagation for accelerated deep learning with reduced overfitting," in *Proceedings of the International Conference on Machine Learning*, pp. 3299-3308, 2017.
- [10] A. Radford, L. Metz, and S. Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proceedings of the International Conference on Learning Representations*, pp. 1-15, 2016.

arXiv preprint arXiv:1511.06434

- [11] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," in *Proceedings of the International Conference on Learning Representations*, pp. 1-14, 2015. arXiv preprint arXiv:1412.6806
- [12] X. Chen, Y. Duan, R. Houthoof, J. Schulman, I. Sutskever, and P. Abbe, "InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets," in *Proceedings of Advances in Neural Information Processing Systems*, pp. 2180-2188, 2016.
- [13] D. Triantafyllidou and A. Tefas, "Face detection based on deep convolutional neural networks exploiting incremental facial part learning," in *Proceeding of the International Conference on Pattern Recognition*, pp. 3560-3565, 2016.
- [14] E. L. Miller, G. Huang, A. RoyChowdhury, H. Li, and G. Hua, "Labeled faces in the wild: A survey," in *Advances in Face Detection and Facial Image Analysis*, pp. 189-248, 2016.
- [15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [16] S. H. Choi, and S. H. Jung, "Performance Improvement of Fake Discrimination using Time Information in CNN-based Signature Recognition", *Journal of Digital Contents Society*, Vol. 19, No. 1, pp. 206-212, 2017



한윤(Yoon Han)

2003년~2012년 : 고려대학교 전자통신공학과 학사
2016년~현재 : 고려대학교 정보보호대학원
빅데이터응용및보안학과(석사과정)

2007년~2008년: (주)그린시스 연구원
2008년~2010년: 인포뱅크 솔루션&웹팀 연구원
2011년~2012년: (주)타임스프레드 공동대표
2013년~현재: 고려대학교 산학관 소재 미디어오션 개발팀 팀장
※관심분야 : 머신러닝, 빅데이터분석, Security Assurance, 머신러닝해킹



김형중(Hyoung-Joong Kim)

1978년 : 서울대학교 전기공학과 학사
1986년 : 서울대학교 제어계측공학과(공학석사)
1989년 : 서울대학교 제어계측공학과(공학박사)

1989년~2006년: 강원대학교 교수
2006년~현재: 고려대학교 정보보호대학원 교수
※관심분야 : 컴퓨터보안, 패턴인식, 가역정보은닉, 머신러닝, 빅데이터분석 등