

심층 GAN을 이용한 이미지 완성 어플리케이션

조상현*, 김종득**

*부산대학교 정보컴퓨터공학부

**LG 유플러스 해외사업팀

e-mail:delosycho@gmail.com

A application for Image completion with Deep GAN

Sang-Hyun Cho*

Jong-Deug Kim**

*Dept of Computer Science and Engineering, Pusan National University

**LG Uplus Oversea Business Team

요 약

사진에는 의도하지 않은 노이즈나 찍는 과정 중에 발생한 실수나 문제로 원치 않게 가려진 부분이 있을 수 있는데, 이미지 완성 어플리케이션은 사용자가 전문적인 프로그램이나 전문가의 도움 없이 노이즈나 가려진 부분을 제거할 수 있도록 하였다. 본 논문에서는 GAN(Generative Adversial Network) 모델에 노이즈가 있는 사진을 입력으로 넣어 노이즈가 제거 된 사진을 생성하도록 하였고, 생성 된 사진과 기존 이미지가 자연스럽게 합성 될 수 있도록 보정을 하여 완성 된 이미지를 출력하는 어플리케이션을 제안한다. GAN 분류 모델의 시그모이드 교차-엔트로피와 생성이미지와 원본이미지간의 평균 제곱 오차를 함께 최소화 하도록 생성 모델을 학습시켰고, 낮은 평균 제곱 오차를 가지는 완성 이미지를 생성 할 수 있었다. 이미지 보정을 통해 생성 된 이미지와 입력 이미지와의 밝기 차이를 해소시켜 좀 더 자연스러운 완성 이미지 결과를 얻을 수 있었다.

1. 서론

사진을 찍기 간편해짐에 따라서 사람들은 일상에서 많은 사진들을 찍고 있다. 찍은 사진에는 의도하지 않은 노이즈나 찍는 과정 중에 발생한 실수나 문제로 인해서 원치 않게 가려진 부분이 있을 수 있다. 하지만 이러한 부분을 제거하기 위해서는 전문프로그램이나 전문가의 도움이 필요할 수 있고 그로 인해서 많은 비용이나 시간이 든다.

본 연구에서는 사진의 일부가 검은색 픽셀로 가려진 사진을 입력으로 넣어서 검은 픽셀의 부분을 GAN을 통해 자연스러운 사진이 되도록 채워주는 이미지 완성 어플리케이션을 제안한다.

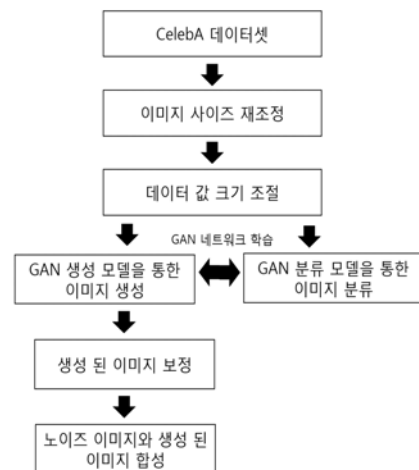
2. 이미지 완성 어플리케이션 구성

이미지 완성 어플리케이션은 이미지 파일들을 읽어 와서 처리하는 데이터 입력 부분, 그리고 일부분이 가려진 노이즈가 포함된 데이터를 받아서 자연스러운 사진이 되도록 픽셀을 채워주는 이미지 생성 부분, 생성 되어진 이미지를 보정하여 좀 더 자연스러운 이미지가 되도록 하는 이미지 보정 부분으로 이루어져 있다.

본 연구에서는 약 20만개의 유명인사 얼굴 이미지들을 포함하고 있는 CelebA 데이터셋을 사용하였다. CelebA 데이터셋은 다양한 연령층과 다양한 복장 그리고 다양한 헤어스타일을 가진 이미지들을 포함하고 있다. 어플리케이션

의 데이터 입력 부분에서는 데이터셋의 이미지를 읽어와 이미지의 크기를 64 × 64로 조정하고 이미지의 픽셀 값을 255로 나누어 데이터의 범위가 0 ~ 1이 되도록 조정하고 무작위로 위치와 범위를 선택하여 선택 된 범위의 값을 모두 0으로 설정하여 이미지의 일부분에 검은색 노이즈를 추가하도록 하였다.

이미지 생성 부분에서는 노이즈가 추가 된 이미지를 GAN의 생성 모델에 입력으로 넣어 노이즈가 제거 된 이미지를 얻고 분류 모델을 통해 네트워크를 학습하여 자연스러운 이미지를 출력하도록 하였다.



(그림1) 이미지 완성 어플리케이션 구성도

3. 방법

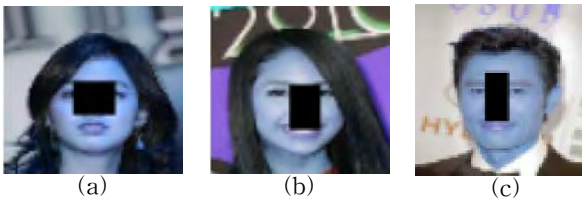
3.1 Denoising Auto-Encoder

일반적인 Auto-Encoder는 입력 값과 목표 값이 모두 주어진 감독학습과 다르게 입력 값만이 주어진 상태에서 데이터의 특징을 찾아내는 비감독학습 방법이다. 입력 값 x 를 인코더 s 를 통해서 압축하고, 압축 된 데이터 y 를 디코더를 통해 입력 값과 같은 크기의 데이터로 복원되도록 한다.

$$\begin{aligned} y &= s(Wx + b) \\ z &= s(W'y + b') \end{aligned} \quad (1)$$

Denoising Auto-Encoder[2]는 이러한 모델에서 입력 데이터에 노이즈가 추가되었을 때, 추가 된 노이즈를 제거하고 원래의 데이터를 출력하는 모델이다.

본 논문에서는 입력 데이터에 임의로 노이즈를 주기 위해 23과 33 사이의 무작위 숫자를 선택하여 X와 Y 좌표로 설정하고, 5와 20 사이에서 무작위 숫자를 선택하여 직사각형의 너비와 높이로 설정하고, 64의 높이와 너비를 가지는 입력 이미지에서 만들어진 직사각형 영역의 RGB 픽셀을 모두 0으로 설정하고 임의로 노이즈를 생성시켜 실험을 진행하였다.



(그림2) 임의로 노이즈를 생성한 이미지 예시

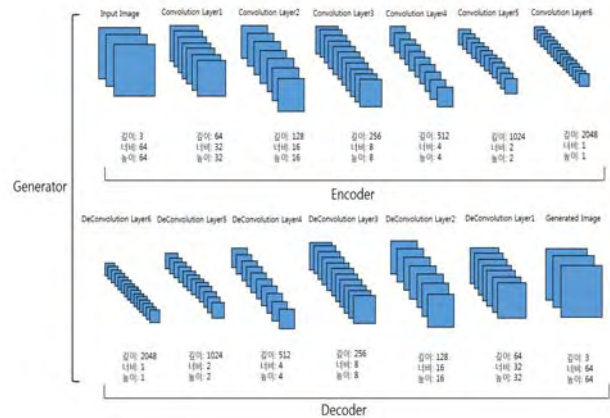
3.2 GAN

GAN은 실제 데이터에 가까운 데이터를 생성하는 생성 모델과 데이터가 실제인지 아닌지 구분해내는 분류 모델이 있다. 생성 모델은 분류 모델에서 실제 데이터로 분류되게끔, 분류 모델은 생성 모델에서 생성 된 데이터가 실제 데이터로 분류되지 않게끔 신경망을 학습시킨다.

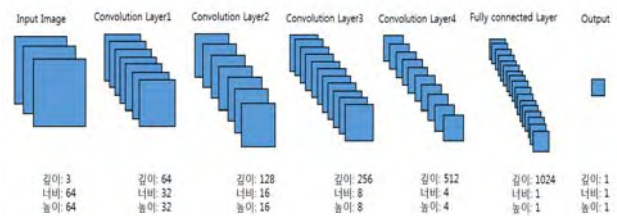
$$\begin{aligned} \min_G \max_D V(D, G) &= E_{x \sim P_{data}(x)} [\log D(x)] + \\ &E_{z \sim P_{data}(z)} [\log(1 - D(G(z)))] \end{aligned} \quad (2)$$

생성 모델에서는 데이터의 높이와 너비가 1이 될 때까지 합성곱 네트워크를 통해서 연산을 했고, 압축 된 데이터를 다시 역합성곱 네트워크를 통해서 데이터의 처음 높이와 너비가 될 때 까지 연산을 했다. 각 합성곱과 역합성곱의 층의 활성화 함수로는 기존의 ReLu의 변형으로 음수에 대해 1/10 값으로 줄여서 사용하는 Leaky ReLu를 사용했다. 과적합 문제를 피하기 위해서

각 층의 연산이 끝난 후 배치 정규화와 드롭아웃을 적용시켰다.



(그림3) 제안하는 어플리케이션에 적용 된 GAN의 생성 모델



(그림4) 제안하는 어플리케이션에 적용 된 GAN의 분류 모델

3.3 이미지 보정

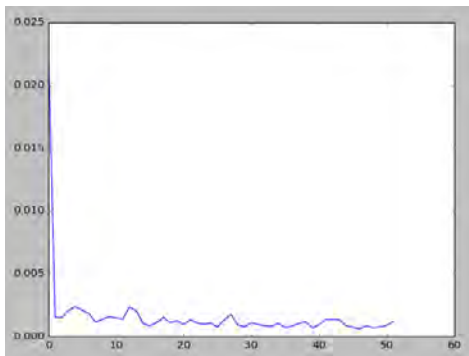
GAN의 생성 모델을 통해 생성 된 이미지는 원본 이미지의 모양과 비슷하지만 밝기의 차이 때문에 합성을 했을 때 부자연스러운 현상을 줄이기 위해 생성 된 이미지에 보정 기법을 사용해서 더 자연스러운 이미지를 생성 할 수 있도록 하였다. 생성 된 이미지에서 노이즈가 있는 부분의 RGB 픽셀의 평균을 각각 구하고 노이즈 영역 주변 픽셀의 RGB 픽셀의 평균을 각각 구하여 노이즈 영역 주변의 평균에 노이즈 영역 평균값을 나눈 값을 이미지를 합성할 때 곱해주었다. 다음 수식은 이미지 보정을 위한 값이 계산되는 방식을 나타내며 n width와 n height는 노이즈 영역의 너비와 높이, n_x 와 n_y 는 노이즈 영역의 X, Y 좌표를 나타내며, 다음 수식은 R, G, B 픽셀에 대해서 각각 적용되도록 하였다.

$$\begin{aligned} m_g &= \sum_{x=n_x}^{nwidth} \sum_{y=n_y}^{nheight} G_{x,y} / ((nwidth - n_x) * (nheight - n_y)) \quad (3) \\ m_i &= \left(\sum_{x=n_x}^{nwidth} (I_{x, n_y-1} + I_{x, nheight+1}) \right. \\ &\quad \left. + \sum_{y=n_y}^{nheight} (I_{n_x-1, y} + I_{nwidth+1, y}) \right) / ((nwidth - n_x) \\ &\quad \quad \quad + (nheight - n_y)) \\ G &= G * (m_i / m_g) \end{aligned}$$

4. 실험 결과 및 고찰

제안하는 어플리케이션의 GAN 모델에서는 생성 모델에서 생성된 이미지가 분류 모델에서 실제 이미지로 분류되도록 하기 위한 시그모이드 교차-엔트로피와 함께 완성된 이미지와 생성 모델을 통해 생성된 이미지와의 평균 제곱 오차를 함께 최소화하도록 학습시켰다. 평균 제곱 오차 손실 값은 일정 수준이상 수렴한 후에는 일정 값 이하로 떨어지지 않았다.

노이즈가 얼굴의 대부분을 가리거나, 모자나 선글라스 같은 부분이 노이즈 영역으로 가려지거나, 얼굴이 정면이 아닌 부분이 찍힌 이미지의 경우 부정확한 생성 결과를 내는 것을 확인 할 수 있었다.



(그림5) 학습에 따른 평균 제곱오차 손실 값 변화



(그림6) 원본 이미지



(그림7) 일정 부분이 가려진 얼굴 이미지



(그림8) GAN 모델을 통해 생성된 얼굴 이미지



(그림9) 생성 된 이미지를 합성해 만든 완성 이미지



(그림10) 보정 처리가 된 완성 이미지

평균 제곱오차가 최소가 되도록 학습 된 값을 보정 방법에 따라 변형시킨 것이기 때문에 보정 된 이미지의 원본 이미지와의 평균 제곱오차는 보정이 되지 않은 이미지보다 높게 나왔지만, 합성 될 이미지의 밝기를 주변의 픽셀에 맞게 보정함으로써 좀 더 자연스러운 완성 이미지가 나오는 것을 볼 수 있었다.

(표1) 보정이 되기 전과 후의 평균 제곱오차 손실 값 비교

보정 되기 전 평균 제곱오차	보정 된 후 평균 제곱오차
0.000902673	0.000963396

5. 결론 및 향후 연구

본 논문에서는 심층 GAN을 이용하여 일부분이 검은색 픽셀로 가려진 이미지를 입력으로 받아 원본 이미지와 비슷한 이미지를 생성하고, 생성 된 이미지를 통해 가려진 부분이 원본 이미지와 같이 자연스러운 픽셀로 채우는 방법을 제안하였다.

GAN의 생성 모델을 통해 입력 이미지의 원본 이미지와 비슷한 이미지를 생성하여 낮은 평균 제곱 오차를 가지는 이미지를 완성할 수 있었지만 채워진 부분과 이미지와 원본이미지의 미묘한 밝기차이로 인해서 자연스러운 이미지를 생성하지는 못했다.

향후에 좀 더 보완된 GAN 학습방법과 이미지 보정 방법을 통해 좀 더 자연스러운 완성 된 이미지를 생성할 수 있는 연구를 진행하려 한다.

참고문헌

[1] 김진영, 부석준, 조성배, “심층 전이-GAN을 이용한 악성 소프트웨어 탐지”, 한국정보과학회 학술발표논문집, pp.1893-1895, 2017.

[2] Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P. A., “Extracting and composing robust features with denoising autoencoders”, In Proceedings of the 25th international conference on Machine learning, pp.1096-1103, 2008.

[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., “Generative Adversarial Nets”, Advances in neural information processing systems, pp.2672-2680, 2014.

[4] Radford, Alec, Luke Metz, Soumith Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks.” arXiv preprint arXiv:1511.06434, 2015.

[5] Isola, P., Zhu, J. Y., Zhou, T., Efros, A. A., “Image-to-image translation with conditional adversarial networks”, arXiv preprint arXiv:1611.07004, 2017.