

GAN을 이용한 보행자 학습 데이터 생성

김다운, 송현철, 최광남*
중앙대학교 컴퓨터공학과
*e-mail : knchoi@cau.ac.kr

Generating Pedestrian Training Dataset using GAN

Daeun Dana Kim, Hyun Chul Song, Kwang Nam Choi*
Computer Science and Engineering, Chung-Ang University

요 약

GAN(Generative Adversarial Networks)가 발전함에 따라 네트워크를 이용한 학습을 하여 컴퓨터가 데이터를 생성하는 것이 가능해졌으며 이를 이용하여 보행자 학습 데이터를 생성하는 연구를 진행하였다.

1. 서론

머신러닝이 보편화 되어 감에 따라 사람이 하는 여러 가지 일을 컴퓨터가 대체하게 되었다. GAN(Generative Adversarial Networks)이 연구되고 발전하면서 인간만의 영역이라고 여겨졌던 ‘창조’의 영역에서 컴퓨터가 활용되기 시작했다[1]. 학문적인 분야뿐만 아니라 음악, 미술 등의 예술적 분야에서도 GAN을 이용한 데이터 생성이 이루어지고 있다.

2. 관련연구

인공지능을 소프트웨어적으로 구현한 결과인 머신러닝(Machine Learning)은 컴퓨터가 데이터를 학습하여 데이터의 패턴을 스스로 찾아내 주어진 문제에 대한 적절한 작업을 수행하도록 하는 알고리즘이다[2]. 머신러닝은 그림 1과 같이 지도학습(Supervised Learning), 비지도학습(Unsupervised Learning), 강화학습(Reinforcement Learning)의 분야로 연구되고 있다. 지도학습은 정답이 주어진 상태에서 알고리즘을 학습하는 것을 의미한다. 예를 들어, 수많은 개와 고양이의 사진이 주어지고, 해당 사진이 개인지 고양이인지 정답을 알려준 후 학습하도록 한다. 비지도학습은 정답이 주어지지 않은 상태에서 학습하도록 하는 알고리즘이며 가장 대표적인 비지도학습에는 clustering이 있다. 강화학습은 주어진 환경 안에서 컴퓨터가 행동을 했을 때 보상을 주어 최적의 행동을 하도록 하는 학습방법이다[3].

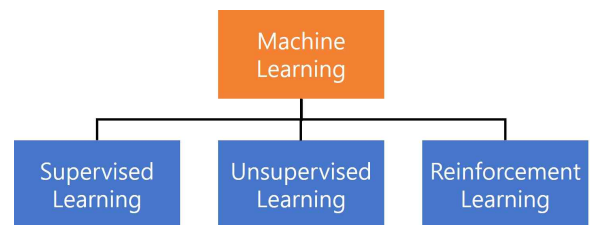


그림 1. Machine Learning의 종류

3. Generative Adversarial Networks

GAN이란 Generative Adversarial Networks로 네트워크를 이용한 학습을 하여 컴퓨터가 데이터를 생성할 수 있도록 한다[4]. GAN은 이미지 생성 능력이 뛰어나기 때문에 영상처리 등 여러 분야에서 활용된다[5]. GAN은 Generator와 Discriminator가 경쟁적으로 서로 발전하고 개선하는 방법을 이용한다. Generator는 존재하는 데이터를 이용하여 사실과 같은 새로운 이미지를 생성하는 네트워크이고 Discriminator는 사물을 인식하는 네트워크이다. Generator가 이미지를 생성하게 되면 Discriminator는 이 이미지가 real 이미지인지 fake 이미지인지를 판별하게 된다. Generator의 목적은 Discriminator가 fake라고 분별할 수 없는 데이터를 생성하는 것이고 Discriminator의 목적은 존재하는 이미지 데이터는 real로, Generator가 생성한 이미지는 fake로 판별하는 것이다. 위와 같은 경쟁이 반복된다면 상호 발전하여 Generator가 실제와 같은 이미지를 생성할 수 있게 된다. 두 목적을 수식화한 함수는 다음과 같다.

4. Deep Convolutional GAN

Deep Convolutional GAN(DCGAN)은 기존 GAN의 문제점을 보완한다[6]. 기존 GAN의 단점은 고해상도의 이미지를 생성하기 힘들다는 것과 Fully-Connected한 네트워크의 구조 때문에 불안정하다는 점이다. 이를 개선하기 위해 DCGAN은 Discriminator와 Generator의 중간에 네트워크 구조를 CNN 구조로 바꾼 형태며 이를 시각화하면 그림 2와 같다. Discriminator는 convolution을 거치며 입력된 이미지의 특징을 뽑아내고 특징 맵의 크기를 점점 줄여 한 개의 값으로 축소시킨 후 그 값을 activation function에 넣어 예측 값을 구한다. Generator는 입력 노이즈 z 를 받아서 z 의 분포 위치에 대응하는 이미지를 생성하도록 학습된다.

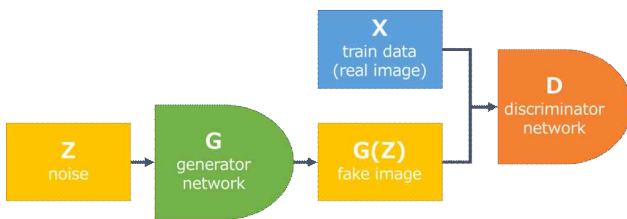


그림 2. DCGAN의 구조

5. 구현

본 논문에서는 DCGAN을 활용하여 보행자 dataset을 생성하는 프로그램을 구현하였다.

학습에 사용된 보행자 이미지 데이터는 24,770개이며 전체 데이터 셋에 대해 학습을 한 횟수를 나타내는 수치인 epoch은 10으로 설정하여 전체 데이터 셋에 대해 10번 학습을 반복하도록 하였다.

$$\min_D \max_G V(D, G) = E_{x \sim p_{\text{data}}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log \{1 - D(G(z))\}]. \quad (1)$$

수식 1은 GAN의 목적함수이다. Generator인 G 는 $\log(D(G(z)))$ 를 최대화하도록 하여 real 데이터와 같이 확률이 높게 나오도록 학습이 진행된다. Discriminator인 D 는 $\log(D(x)) + \log(1 - D(G(z)))$ 를 최대화하도록 하여 real 데이터를 입력하면 높은 확률이 나오도록 하고 fake 데이터를 입력하면 확률이 낮아지도록 학습이 진행된다. 수식 1을 이용하여 기존의 이미지는 real로, 생성된 이미지는 fake로 판별할 수 있도록 하였다. 그림 3은 학습의 과정과 결과를 나타낸다.



그림 3. 학습 과정과 결과(epoch 1, epoch 4, final result)

6. 결론

본 논문에서는 DCGAN을 이용하여 보행자 dataset을 생성하는 방법에 대하여 논하였다. 본 논문에서는 약 25,000개의 dataset을 학습에 활용하였지만 더 많은 dataset을 이용한다면 좀 더 실제 이미지에 가까운 결과를 얻어낼 수 있을 것이다. 딥러닝의 분야에서 문제가 되는 부분이 많은 dataset을 필요로 한다는 점이다. 이 점을 개선하기에 적합한 방법이 GAN을 이용하여 부족한 dataset을 생성하는 방법이다. 따라서 dataset이 많이 필요한 딥러닝 분야에서 유용하게 사용될 것으로 보인다. 학습 전에 데이터를 가공하는 기법을 추가하면 더욱 발전시킬 수 있을 것으로 보이며 Conditional GAN 등의 기법을 이용하면 조건에 맞는 dataset 생성도 가능하기 때문에 추후 이러한 방법으로 발전할 가능성이 있다.

사 사

본 논문은 중소기업청에서 지원하는 2018년도 산학협력 기술개발사업(20180746)의 연구수행으로 인한 결과물임을 밝힙니다.

참고문헌

- [1] 이호진, et al. "데이터 불균형 문제를 해결하기 위한 Generative Adversarial Networks (GAN) 기반의 가상 데이터 합성." *대한기계학회 춘추학술대회* (2017): 1142-1143.
- [2] 김지원, et al. "다양한 딥러닝 알고리즘과 활용." *정보과학회지* 33.8 (2015): 25-31.
- [3] Mnih, Volodymyr, et al. "Human-level control through deep reinforcement learning." *Nature* 518.7540 (2015): 529.
- [4] Gauthier, Jon. "Conditional generative adversarial nets for convolutional face generation." *Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition, Winter semester 2014.5* (2014): 2.
- [5] Kong, Xiangjie, et al. "Mobility dataset generation for vehicular social networks based on floating car data." *IEEE Transactions on Vehicular Technology* 67.5 (2018): 3874-3886.
- [6] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." *arXiv pre-print arXiv:1511.06434* (2015).