AI 기술을 이용한 탑 생성에 대한 방법에 대한 연구

고병국[O], 강은수[*], 이조순[*], 최하진[*], 김준오[*], 이병권[**]
[O]동국대학교 멀티미디어공학과,
[*]동국대학교 멀티미디어공학과,
[**]서원대학교 멀티미디어과
e-mail: sonic747@seowon.ac.kr

A Study on the Method of Tower Generation Using AI Technology

Ko Byeongguk[O], Eunsu Kang[*], JoSun Lee[*], Choi Hajin[*], Jun O Kim[*], Byongkwon Lee[**]
[O]Dept. of Multimedia Engineering, Dongguk University,
[*]Dept. of Multimedia Engineering, Dongguk University,
[**]Dept. of Multimedia Seowon University

● 요 약 ●

본 논문에서는 AI 기술을 이용한 객체 생성을 위해 윈도우 환경에서 DCGAN 알고리즘을 사용한 객체 생성에 대한 방법을 제안한다. 이 방법은 아나콘다, 리눅스 등의 가상환경을 요구하지 않기 때문에 사전 환경구성 시간을 최소화한다. 또한 이 방법은 Python을 기반으로 구현됐기 때문에 이해 및 추가개발 및 수정이 편리하다. 따라서 빠른 시간 내에 자신이 원하는 객체를 생성해낼 수 있는 시스템을 구축하는 방법을 제안한다.

키워드: DCGAN,윈도우 환경, 작업시간 최소화, 편리한 수정

I. Introduction

최근 원본과 다른 문화재 복원 작업에 대하여 논란이 존재한다. 원본에 대한 데이터가 존재할 경우 문화재 복원 작업이 원활하게 이루어질 수 있지만, 그렇지 않은 경우 원본과 유사한 복원 작업에 어려움이 생긴다.

따라서 손실된 문화재 영상을 GAN을 통하여 원본의 영상으로 복원시키는 연구를 진행하고자 한다. 특히 문화재 중 탑에 대한 연구를 진행하려 하며, 본 연구는 사전 연구로써, GAN 알고리즘을 통해 탑 이미지를 기반으로 새로운 탑 이미지를 생성하고자 한다.

II. Preliminaries

1. Related works

1) GAN(Generative Adversarial Network)

첫 글자인 'G'는 GAN이 생성(Generation) 모델이라는 것을 뜻한다. 생성 모델이란 '그럴듯한 가짜'를 만들어내는 모델이다. 언뜻 보면 진짜 같은 가짜 사람 얼굴 사진을 만들어내거나 실제로 있을 법한 고양이 사진을 만들어내는 것이 생성 모델의 예다.

두번째 글자인 'A'는 GAN이 두 개의 모델을 적대적(Adversarial)으로 경쟁시키며 발전시킨다는 것을 뜻한다. 새로운 데이터를 생성해내는 생성자(Generator)와 이를 구분하는 구분자(Discriminator)를 경쟁적으로 학습시킨다[1].

마지막 글자 'N'은 이 모델이 인공신경망(Artificial Neural Network) 또는 딥러닝(Deep Learning)으로 만들어졌기 때문에 붙었다.

요약하자면 GAN은 생성이라는 문제를 풀기 위해 딥러닝으로 만들어진 모델을 적대적 학습이라는 독특한 방식으로 학습시키는 알고리즘이다[2].

GAN은 학습이 불안정하기로 악명이 높다. 다음은 GAN에 대한 한계 3가지이다.

① GAN은 결과가 불안정하다

② Black-box method : Neural Network 자체의 한계라고 볼 수 있다. 결정 변수나 주요 변수를 알 수 있는 다수의 머신러닝 기법들과 달리 Neural Network는 처음부터 끝까지 어떤 형태로 그러한 결과가 나오게 되었는지 그 과정을 알 수 없다.

③ Generative Model 평가 : GAN은 결과물 자체가 새롭게 만들어진 Sample 이다. 이를 기존 sample과 비교하여 얼마나 비슷한 지 확인할 수 있는 정량적 척도가 없고, 사람이 판단하더라도 이는 주관적 기준이기 때문에 얼마나 정확한지, 혹은 뛰어난지 판단하기 힘들다.

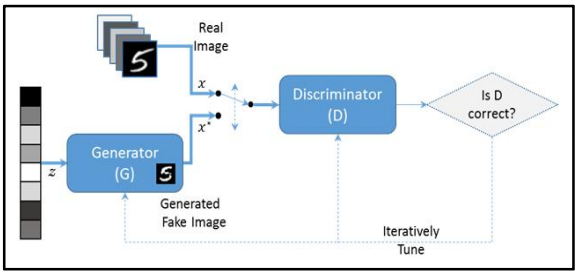

Fig. 1. GAN Structure

2) DCGAN : GAN은 학습이 불안정하기로 악명이 높다. 학습이 어렵다는 점은 GAN 모델이 다양한 곳에 응용되는 것을 가로막는 큰 장애물이었다. 이런 상황에서 수많은 실험 끝에 안정적인 학습이 가능한 GAN 모델의 구조를 찾아낸 것이 DCGAN이다.

DCGAN의 특징은 몇 가지로 요약할 수 있다. 먼저, 선형 레이어와 풀링 레이어(Pooling Layer)를 최대한 배제하고 합성곱(Convolution)과 'Transposed Convolution(Fractional-Strided Convolution)'으로 네트워크 구조를 만들었다. 풀링 레이어는 여러 딥러닝 모델에서 불필요한 매개변수의 수를 줄이고 중요한 특징만을 골라내는 역할을 하는 레이어지만 이미지의 위치 정보를 잃어버린다는 단점이 있다.

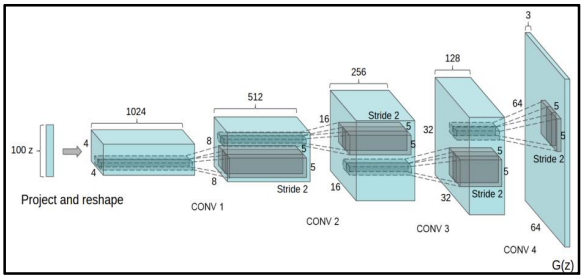

Fig. 2. DCGAN Structure

이미지를 생성하기 위해서는 위치 정보가 중요하기 때문에 DCGAN은 풀링 레이어를 배제했다. 선형 레이어 역시 마찬가지로 위치 정보를 잃어버리므로 모델의 깊은 레이어에서는 선형 레이어를 사용하지 않았다[3].

III. The Proposed Scheme

DCGAN을 윈도우에서 사용하기에 앞서 다음의 환경이 준비되어야 한다.

① Python 3.7(with php) 및 Python edit 프로그램

② CUDA 10.1 및 cuDNN

1. 환경 구성

cmd환경에서 php 업데이트 및 파이썬 라이브러리를 다운로드 받는다.

① pip 업그레이드

pip install —upgrade pip

② numpy 라이브러리 설치

python -m pip install numpy

③ matplotlib 라이브러리 설치

python -m pip install matplotlib

④ Pytorch 네이티브 설치

링크 : https://pytorch.org/

Pytorch 홈페이지에서 맞는 환경을 선택 후 나오는 Run this Command 내용을 cmd에서 실행

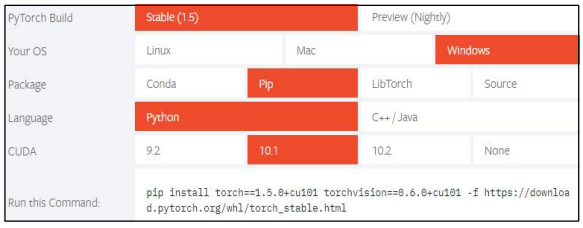

Fig. 3. Pytorch installation setup

⑤ 아래 그림의 명령어를 통해 GPU 종류, CUDA 사용가능 여부 및 설치된 Pytorch 버전을 알 수 있다. 결과가 제대로 나온다면 환경설정이 완료된다.

Fig. 4. Configuration confirmation command

2. DCGAN 실행

github 링크 :

https://github.com/Natsu6767/DCGAN-PyTorch

프로젝트의 py파일들의 내용을 목적에 맞게 수정한다.

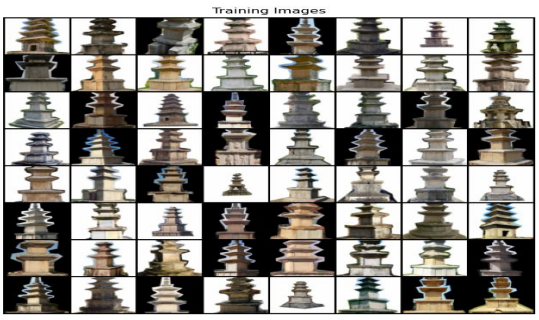

Fig. 5. Dataset for Training

① utils.py에서 root 값 수정 : 데이터셋이 있는 디렉토리의 상위 디렉토리 경로

② train.py에서 학습 파라미터를 설정한다.

nepochs : 학습 반복회수

save_epoch : save step

③ 학습을 시작하기 위해 cmd창에서 train.py 파일을 실행시킨다. 아래 이미지는 utils.py의 root 경로를 통해 수집한 데이터셋 이미지들이며, 문제가 없을시 창을 닫는다. 창을 닫게 되면 학습이 시작되고 학습이 종료되면 학습에 대한 Loss 그래프가 보여진다. 학습된 가중치 파일은 model 디렉토리에 저장된다.

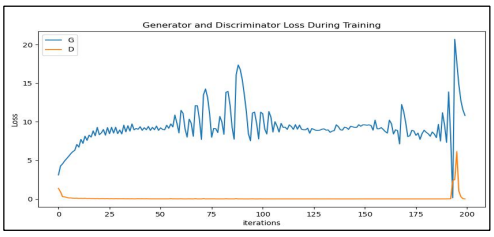

Fig. 6. Training Loss Graph

④ 학습된 가중치 파일로 이미지를 생성해내기 위해서 cmd창에서 generate.py 파일을 실행시킨다. 가장 마지막으로 학습된 가중치 파일인 model_final.pth 가중치 파일로 generate가 진행된다. 아래 이미지는 가중치 파일을 통해 generate된 탑 이미지이다.

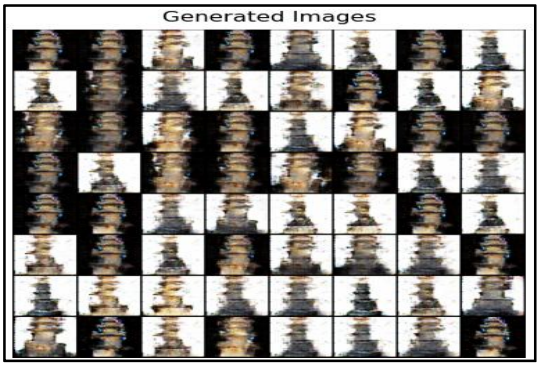

Fig. 7. Generated Image Based on Dataset

IV. Conclusions

최근 GAN 알고리즘을 개선한 많은 다른 알고리즘 연구들이 나오고 있기에 GAN 알고리즘을 문화재 복원에 활용하고자 했다. 본 연구는 DCGAN 알고리즘을 이용하여 새로운 탑 이미지를 생성하는 방법을 연구했다. 하지만 아직 고해상도 이미지를 생성해낼 수 없고 학습이 불안정한 부분이 존재한다. 차후 모델 개선을 통하여 손실된 탑 이미지를 기반으로 원본 탑 이미지를 생성해내는 GAN 모델을 연구하고자 한다.

본 연구에서 제안하는 방법은 아나콘다, 우분투 등 특별한 가상환경을 요구하지 않으며, Python을 사용하여 구현 또한 간편하다. 따라서 딥러닝을 통해 객체 생성 모델을 개발 또는 테스트하고자 하는 사람들에게 도움을 줄 수 있다.

ACKNOWLEDGEMENT

이 (성과)는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2020R1A2C100766811).

REFERENCES

[1] Ian J. Goodfellow, "Generative Adversarial Networks" NIPS 2014, Jun 2014.

[2] Alec Radford, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks" ICLR 2016, Jan. 2016.

[3] Joonho Lee and Donghwan Seo, "Recovery of Mosaic Facial Images by Separate Training for Key Facial Attributes" Journal of the Institute of Electronics and Information Engineers, Vol. 57, No. 5, pp. 87-95, Jan. 2016.