

## CycleGAN을 이용한 인터랙티브 웹페이지

\*김지원, 정해정 \*\*김동호  
서울과학기술대학교

wonjw3638@naver.com, 17101532@seoultech.ac.kr, dongho.kim@seoultech.ac.kr

### Interactive Web using CycleGAN

\*Jiwon Kim, Haejung Jung \*\*Dongho Kim  
Seoul National University of Science and Technology

#### 요약

최근에 딥러닝 기술인 GAN (Generative Adversarial Networks) 연구는 Image-to-Image translation 분야에서 활발하게 이뤄지고 있다. 이러한 기술을 바탕으로 사용자에게 편의와 재미를 제공하는 서비스가 애플리케이션 및 웹사이트의 형태로 개발되고 있다.

이에 본 논문은 CycleGAN 모델을 사용하여 이미지를 변환하고, 이를 인터랙티브 웹페이지를 통해 사용자와 실시간으로 상호 작용하며 결과 이미지를 제공할 수 있는 방법을 연구하였다. 모델을 구현하기 위해 Tensorflow 및 Keras를 사용하였고, Django 와 HTML5, CSS, JavaScript를 사용하여 웹사이트를 제작하였다.

#### 1. 작품의 제작 동기

우리는 사진의 전체적인 분위기를 보정하거나 필터를 적용하기 위해 다양한 편집 프로그램과 애플리케이션을 사용한다. 가끔 이러한 방식은 틀에 대해 사용 방법을 익혀야 할 때가 있고 시간이 많이 소요될뿐더러 얻은 결과물이 만족스럽지 않은 경우가 있다. 이에 우리는 위와 같은 고충을 해결하고 기술의 발전을 도모하기 위해, Image-to-Image translation 네트워크 중 CycleGAN 모델을 사용하여 이미지를 특정 타겟 도메인의 이미지로 변환하고 화질을 개선하여 인터랙티브 웹페이지를 통해 제공하는 구조의 서비스를 제안한다.

사용자는 웹페이지를 통해 변환하고자 하는 이미지를 입력한다. CycleGAN은 입력된 이미지를 해당 도메인과 매핑되는 타겟 도메인의 이미지로 변환한다. 이후 특정 도메인에서 발생하는 화질 저하의 문제를 해결하고 전체적인 결과 이미지의 품질을 높이고자 RDN (Residual Dense Network)을 한 번 더 통과한다. 모든 과정을 거친 이미지는 웹페이지에서 인터랙티브로 동작할 수 있도록 구현된 기능들에 의해 사용자의 움직임에 따라 표시된다.

우리는 CycleGAN을 이용한 인터랙티브 웹페이지를 통해 사진 색감 보정 및 필터 적용을 위한 방식 중 편리함을 극대화한 방법을 제공하고자 하였다. 또한, 쉽게 학습시켜 사용할 수 없는 딥러닝 모델을 실시간으로 서비스하여 인공지능 기술에 대한 사용자의 접근성을 높였다.[1] 덧붙여 사용자와 간단하게 상호작용하며 동작할 수 있는 방식으로 흥미를 유발하였다.

수많은 이미지 변환 모델 중 CycleGAN을 선택한 이유는 학습 데이터를 수집하는 과정에서 생길 수 있는 어려움을 최소화하기 위함이다.[2] CycleGAN은 서로 짝을 이루지 않는 데이터셋으로 학습시킬 수 있는 이미지 변환인 Unpaired image-to-image translation에 해당하는 네트워크이고, Patch 단위로 훈련을 진행한다. 따라서 현실적으로 수집하기 어려운 Paired dataset을 구하는 문제점을 해결할 수 있고, 비교적 적은 양의 데이터로 안정적인 결과 이미지를 얻는 것이 가능하다. CycleGAN은 객체의 모양은 바꿀 수 없다는 단점이 존재하나, 이미지의 전체적인 분위기 및 스타일을 변환하고자 하는 우리의 목표에는 큰 문제가 되지 않는다고 판단하였다.

이미지의 도메인은 낮과 밤, 봄과 가을, 여름과 겨울 도메인에서 상호 변환 할 수 있도록 설정하여 총 3개의 모델을 생성하였다. 각각에 해당하는 데이터셋을 수집하기 위해 Google에서 데이터를 크롤링하여 저장하였고, Tensorflow에서 제공하는 데이터셋 중 CycleGAN 논문에서 사용한 summer2winter\_yosemite 데이터를 사용하였다. 그리고 공공 데이터포털에서 지역관광공사와 국립공원공단이 제공하는 데이터를 활용하여, 실제 서비스 제공 시 입력될 이미지를 고려해 우리나라 풍경을 학습할 수 있게 하였다. 이후 데이터 증강을 위해 좌우변환, 확대, 랜덤 크롭 및 채도 조정을 수행하였다. 데이터 중 80%는 훈련에 사용하고 20%를 시험에 사용하였다 [표 1].

표 1. 데이터셋

	낮 ↔ 밤		봄 ↔ 가을		여름 ↔ 겨울	
Train	623	301	321	416	1231	971
Test	108	60	82	100	309	247
합계	731	361	403	513	1540	1218

#### 2. 작품의 설계 및 구현

##### 2.1 시스템제작

사용자에게서 input image를 받은 후, CycleGAN으로 변환된 이미지를 인터랙티브 웹페이지를 통해 보여주는 것이 본 시스템의 최종 목표이다. CycleGAN code의 사용을 용이하게 하기 위해서 python code를 사용하는 Django를 선택하여 Backend를 구성하였다.[3]

사용자에게서 받은 이미지를 저장할 수 있는 Database를 생성하였다. CycleGAN의 입력 이미지는 256x256 pixel이어야 한다는 조건이 있으므로, 사용자가 원하는 영역을 1:1 비율로 crop 할 수 있도록 하였다. 1:1로 크롭된 이미지는 256x256으로 resize 변환과정을 거친 후, 학습된 모델을 통해 이미지 변환을 거치게 된다.

사용자가 입력한 이미지가 어떠한 도메인인지를 선택하면 그에 따른 모델로 변환시켜준다. 후에 변환된 이미지는 인터랙티브 Web을 통해 확인할 수 있다. 낮↔밤 변환은 Drag에 따른 Scratch 반응을 보인다. 봄↔가을 변환은 마우스 hover에 따른 blur 반응을 보인다. 여름↔겨울 변환은 Scroll에 따라 이미지가 나타나게 된다. 인터랙티브 Web을 경험한 후에는 변환된 이미지를 다운로드 받을 수 있도록 하였다.

## 2.2 한계 및 개선

학습 과정을 checkpoint로 확인한 결과 작은 epoch에서는 checkerboard artifact나 타겟이 아닌 곳에서의 왜곡이 발생했으나, 전반적으로 epoch이 증가하면서 개선되었다. 그러나 낮 ↔ 밤 변환 모델에서 이미지 도메인의 특징으로 인해 화질의 저하가 눈에 띄게 나타나는 것을 확인하였다. 이 문제를 해결하고 특정 모델뿐만 아니라 전체 시스템의 품질을 높이고자 RDN (Residual Dense Network) 모델을 사용하여 화질을 개선하였다.[4] RDN의 학습된 가중치를 사용할 수 있는 네트워크 중 우리의 이미지에 적용했을 때 성능이 가장 뛰어난 noise-cancel 모델의 가중치를 사용해 화질 개선을 수행하였다. 비교를 위해 RDN으로 화질을 개선한 이미지, RDN을 통과한 후 Edge sharpening을 수행한 이미지, ESRGAN 모델을 적용한 이미지 총 3개의 PSNR을 측정해 비교하고, 추가적으로 주관적인 선호도를 평가한 결과 RDN만 적용한 버전을 최종적으로 선택하였다.

## 2.3 실험환경

본 연구에서 제안하는 시스템을 구현하기 위해서는 GPU를 탑재한 PC가 필요하다. Intel(R) Core(TM) i7-5960X CPU @ 3.00GHz 및 RTX 3080 을 탑재한 Window10 64비트 운영체제를 기반으로 하는 PC를 사용 하였다. 그리고 Google이 제공하는 Google Colaboratory의 nVIDIA tesla K80 GPU 환경에서 학습을 진행하였다. 각 모델당 200 epoch 동안 학습하였고, 환경과 데이터셋의 크기에 따라 5시간에서 이주일 정도의 시간이 소요되었다.

인터랙티브 Web을 구현하기 위해서 Django v3.1.7, Python v3.8.5을 사용하였다.

## 3. 작품의 구현 결과

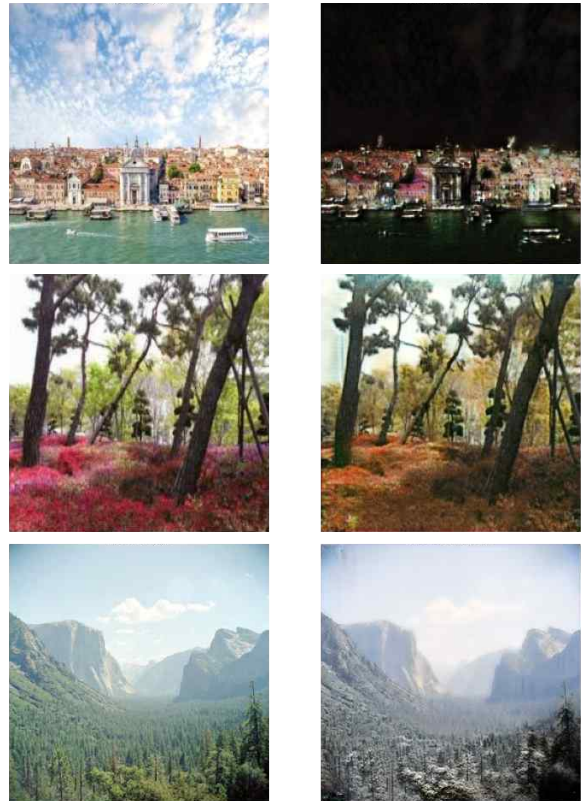


그림 1. 구현 결과 (차례로 낮→밤, 봄→가을, 여름→겨울)



그림 2. 구현 결과 (차례로 밤→낮, 가을→봄, 겨울→여름)

[그림 1]과 [그림 2]는 CycleGAN의 상호변환 결과 이미지이다. 전반적으로 해당하는 계절 및 시간대의 특징을 적절히 학습하여 목표하는 성능만큼 수행하는 것으로 보인다. 그러나 [그림 2]의 첫 번째 이미지인 밤→낮 변환 모델에서 건물의 색상이 다양하게 변환되지 않아 완성도가 떨어지고, 다른 모델에 비해 왜곡이 심하여 모델의 안정성이 부족함을 보인다.

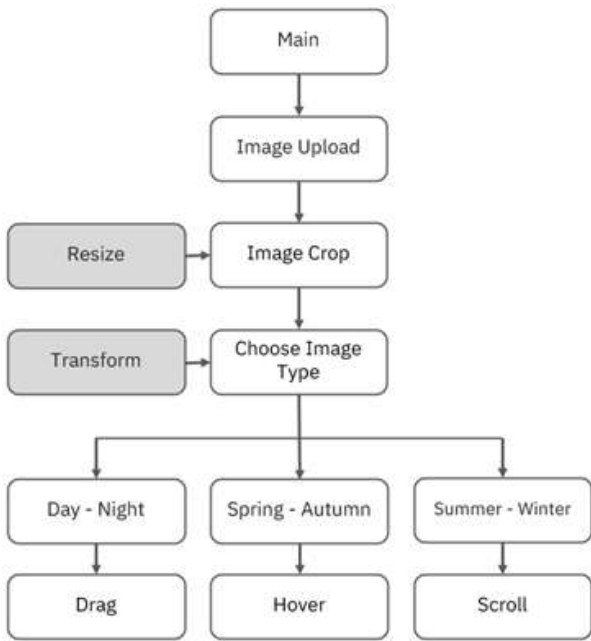


그림 3. 웹 흐름도

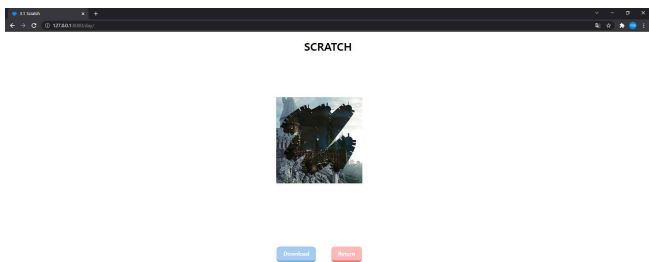


그림 4. Darg page

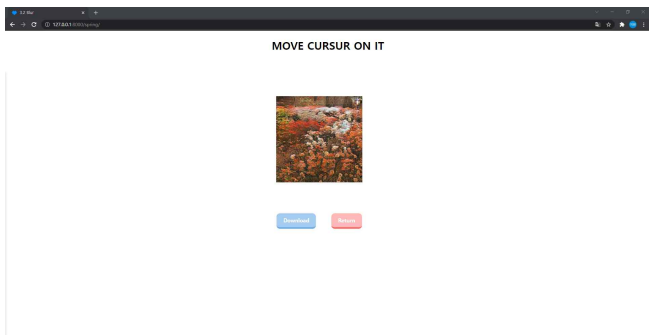


그림 5. Hover page

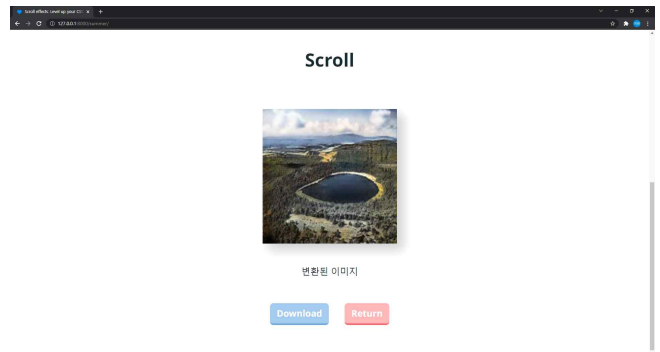


그림 6. Scroll page

CycleGAN을 학습한 모델을 사용하여, [그림 3]의 웹 흐름도를 바탕으로 웹 페이지를 구현하였다. 그 결과 사용자가 손쉽게 원하는 이미지를 crop하여 upload할 수 있었고, 학습한 모델을 바탕으로 이미지를 변환할 수 있었다. 또한, 인터랙티브 웹 페이지[그림 4 ~ 6]에서 이를 확인할 수 있었고, 변환된 이미지를 다운로드 할 수 있었다.

#### 4. 작품의 기대효과

결론적으로 본 연구에서는 Cycle GAN을 이용한 인터랙티브 웹페이지를 통해 사용자가 보다 쉽게 인공지능 기술을 사용하여 이미지를 변환하는 시스템을 제안하였다. 시스템에 더 다양한 모델을 학습시킨다면 사용자는 이미지 보정 및 편집과 같은 별도의 과정 없이도 자신이 원하는 이미지를 손쉽게 얻을 수 있다. 본 연구의 시스템을 이미지뿐만 아니라 video, 360영상, AR/VR등의 미디어에 적용 시킨다면 한정된 이미지 자원을 사용하여 더 다양한 서비스를 사용자에게 제공할 수 있을 것이다.

#### Acknowledgement

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00994,이용환경을 반영하는 자율적 VR·AR 콘텐츠 생성 기술개발)

#### 5. 참고문헌

[1] 김동원, "'포토샵 하나도 모르는데 전문가라고?'...사실 AI가 다했다", AI타임스, 2021.07.20.

[2] ZHU, Jun-Yan, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. 2017. p. 2223-2232.

[3] Vincent, William S. Django for Beginners: Build websites with Python and Django. WelcomeToCode, 2020.

[4] ZHANG, Yulun, et al. Residual dense network for image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. p. 2472-2481.