스케치의 이미지 변환 모델을 이용한 웹 이미지 검색 설계

박연우[1], 지혜정[2], 최채린[3], 김윤희[3]
[1]숙명여자대학교 통계학과
[2]숙명여자대학교 소프트웨어학부 소프트웨어융합전공
[3]숙명여자대학교 소프트웨어학부 컴퓨터과학전공
{balljieun, hyejung0906, lin0902, yulan}@sookmyung.ac.kr

A Design of Sketch Image Transformation and Its Web-Search Results

Yeonwoo Park[1], Hyejung Ji[2], Chae-lin Choi[3], Yoonhee Kim[3]
[1]Dept. of Statistics, Sookmyung Women's University
[2]Dept. of Software Convergence, Sookmyung Women's University
[3]Dept. of Computer Science, Sookmyung Women's University

요 약

직접 그린 스케치 그림을 실제 이미지처럼 변형하는 모델에 연구가 있다. 본 논문은 스케치를 실제 사진으로 변환할 수 있도록 모델을 학습시켜 웹검색이 가능한 웹어플리케이션의 개발을 연구한다. 이를 위하여 관련 데이터를 수집/선별해 학습시킨다. 웹서버와 모델을 연동하여 사진을 입력하면 학습된 이미지 그림 결과를 생성하고 웹에 검색 API를 연결해 해당 이미지 파일의 검색 결과를 바로 제공한다. 이를 통하여 손그림 이미지에 대한 상품 구매 등이 가능하다.

1. 서론

본 논문은 직접 그린 스케치를 통해 원하는 의류를 찾을 수 있도록 하는 서비스를 제안한다. 사용자의 옷에 대한 스케치를 실제 사진으로 변환할 수 있도록 모델을 학습시켰으며, 모델을 통해 얻은 결과 이미지(실제 사진처럼 변환된 이미지를 의미한다.)를 검색에 활용해 검색 결과를 반환한다. 특히 본 논문에서는 image-to-image 변환 모델 중 pix2pix 모델을 사용했으며, 검색 결과를 위해 Bing Visual Search API를 사용했음을 밝힌다. 이 논문은 서론에 이어 관련 연구, 설계, 결론으로 구성된다.

2. 관련 연구

CycleGAN[1]은 도메인을 통해 이미지 변환을 수행하는 모델이다. CycleGAN은 쌍으로 이루어진 데이터 없이도 이미지 간의 변환을 수행할 수 있는 모델로, paired data를 확보하기 어려운 경우에 유용하게 사용할 수 있다.

Image-to-Image Translation[2]은 조건부 네트워크를 사용한 이미지 간 변환에 대한 연구다. 이 연구에서 입력 이미지와 출력 이미지 간의 조건부 관계를 학습하여 이미지 간의 변환을 수행한다. 위의 기술은 그림을 실제 사진으로 변환하는 작업에 활용할 수 있다.

위와 같은 연구들은 이미지 간 변환 기술을 발전시키고 다양한 응용 분야에서 활용할 수 있는 가능성을 제시한다.

3. 설계

3.1 데이터 전처리

후보 중 하나였던 Pix2Pix 모델은 입력과 출력 이미지 간의 변환 과정을 학습하기 위해, 각 입력 이미지가 명확한 출력 이미지와 직접 매핑되는 paired dataset의 구성이 필수적이다. 본 논문에서의 목표는 사용자가 그린 스케치를 실제 옷 사진으로 변환하는 것이었기 때문에, 이미 준비된 출력 이미지 데이터를 바탕으로 이에 상응하는 입력 이미지를 생성해 데이터셋을 구축하고자 했다.

3.1.1 색상 입력

사용자가 원하는 옷을 찾기 위해서는 옷의 형태와 더불어, 색상도 큰 역할을 할 것이다. 따라서, 사용자가 지정한 색상의 이미지를 생성하고자 하였다.

3.1.1.1 Color Clustering

Color Clustering이란 K-Means 기반의 알고리즘으로, 이미지 내 색상의 개수를 k개로 감소시킨다. 이는 이미지를 그림처럼 변환시키고자 할 때 사용될 수 있다.

[그림1] 실제 이미지 [그림2] 결과 이미지

하지만, Color Clustering의 경우 옷에 패턴이 있을 경우 패턴이 매우 정교하게 잡히고, 모두 같은 색상으로 지정되기 때문에 이미지를 구성하는 색상의 개수가 적을 경우 명암이 뚜렷하게 표현되어 사용자의 그림과는 괴리가 생긴다

는 단점이 존재했다. 따라서, 한 단계를 추가해 사용했다.

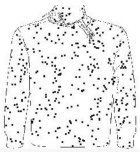
3.1.1.2 구간별 평균 색상을 활용

이미지를 특정 구간별로 구분하고 구간 내 색상의 평균으로 중심점에 색상을 지정하고자 했다. 사용했던 실제 옷 사진은 모두 배경이 없는 이미지였기 때문에, 평균으로 색상을 지정했을 때 배경과 이미지가 구분되지 않는 등의 문제가 발생하지 않았다.

3.1.2 Edge Detection

사용자가 옷의 형태를 표현하기 위해서는 라인 스케치를 활용할 것이다. 따라서, 출력 이미지 데이터에 Canny 알고리즘을 활용하여 Edge detection을 수행해 라인 스케치를 구현했다.

3.1.3 결합



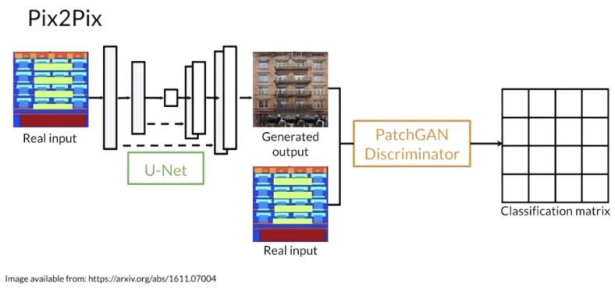
[그림3] 실제 이미지 [그림4] 결과 이미지

라인 스케치와 색상을 지정한 이미지를 결합해 최종적으로 입력 이미지를 생성했다. 색상을 지정한 이미지 위로 라인 스케치에서 색상이 흰색이 아닌 부분을 매핑해 라인이 더욱 뚜렷하게 표현되도록 이미지를 결합해 사용했다.[2]

3.2 Image-to-Image translation

대표적인 image-to-image translation 알고리즘인 pix2pix와 CycleGAN을 사용해 모델링하였다. 그림을 사진으로 변환하는 과정에는 매우 정확한 매핑이 요구되기 때문에, 도메인을 통해 이미지를 변환하는 CycleGAN보다 pix2pix의 성능이 우수해 pix2pix를 최종 모델로 선택하였다.

3.2.1 pix2pix 아키텍처



[그림7] 아키텍처 이미지[3]

pix2pix는 쌍으로 이루어진 데이터를 사용하는 조건부 GAN으로, U-Net 기반의 생성자와 PatchGAN 기반의 판별자가 특징이다. 생성자는 입력 이미지를 목표 이미지로 변환하고, 판별자는 실제 이미지와 생성된 이미지를 구별하는데 이때 L1 손실을 추가하여 세밀한 디테일을 재현하기 때문에 이미지 변환에 효과적인 모델이다.[4]

3.2.2 모델링 과정

생성한 paired data를 입력해 모델을 학습시켰고, pix2pix 논문 내에서 수행한대로 batch_size는 1로 설정했다. 이후 상, 하의 옷을 따로 학습시켜 각각의 모델을 개발했다.

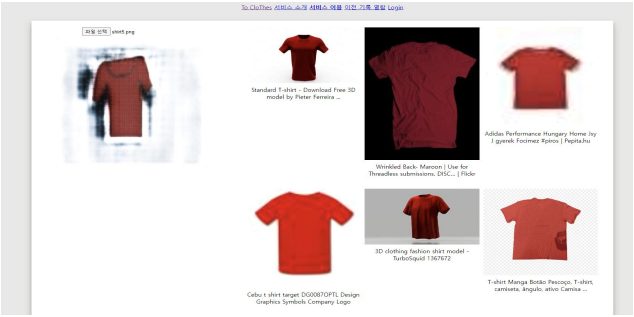
3.3 웹 설계

HTML, CSS5를 이용해 디자인하고, JavaScript로 웹 상에서 기능하는 간단한 동작들을 구현했다. Canvas API의 내장함수를 이용해 그림판 기능을 구현했다. 사용자는 웹에서 제공된 그림판을 사용하여 옷을 직접 그리거나 이미지를 업로드한다. 그려진 이미지나 업로드된 이미지는 학습시킨 모델에 입력되어 사진과 유사한 이미지로 변환된다.

[그림5] 그림판 예시 이미지

3.3.1 상품 검색

생성된 이미지를 기반으로 실제 상품을 검색한다. 사용자는 검색 결과로 나온 이미지를 클릭하면 해당 상품이 있는 링크로 이동할 수 있다. 이미지 검색 API로는 Bing Visual Search API[5]를 선택하였다.



[그림10] 검색 결과 이미지

4. 결론

스케치를 실제 사진처럼 변환하는 모델을 활용해 웹검색까지 가능한 웹 어플리케이션을 개발하였다. 이를 통해 사용자의 스케치가 상품 구매까지 이어지는 과정또한 검증하였다. 향후 연구로는 모델의 성능을 높이면 보다 나은 이미지 생성이 가능하므로 모델 성능 향상을 위한 연구를 진행한다.

참고문헌

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei AEfros(2018). “Image-to-Image Translation with Conditional Adversarial Networks”. Cornell University
[2]https://github.com/michaelnation26/pix2pix-edges-with-color.git
[3]https://medium.com/@vkhalokhi/how-to-implement-pix2pix-gan-in-pytorch-and-keras-e85d82220cf2
[4] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros(2020). “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks”. Cornell University
[5] Bing Visual Search API 공식 문서