

데이터 쌍이 없는 이미지의 형태 변화 개선에 관한 연구

조성무, 김동환, 김석찬*
부산대학교, 부산대학교, *부산대학교

kakawo98@gmail.com, dongh@pusan.ac.kr, *sckin@pusan.ac.kr

A Study on the Unpaired image-to-image Translation with improvement of Shape Transformation

Seongmu Jo, Donghwan Kim, Suk Chan Kim*
Pusan Univ., Pusan Univ., *Pusan Univ.

요약

본 논문에서는 데이터 쌍(Paired Data Set) 없이 학습하는 이미지 변환 모델의 취약점, 형태 변환을 개선하기 위한 방법들을 도입한다. 레이어를 수정하고 새로운 손실함수를 도입하여 모델을 개선하고, 도출된 결과를 기존 모델의 결과물과 비교하여 정성적으로 평가한다.

I. 서론

원본 이미지 내의 물체가 가지고 있는 크기, 각도 등 기하학적 정보와, 배경 이미지와의 맥락을 유지한 채로 이미지 내의 물체 A를 물체 B로 바꿀 수 있다면 다양한 컴퓨터 비전 분야에서 응용될 수 있을 것이다. 이러한 변환모델을 학습시키기 위해서는 변환시키고자 하는 두 도메인 간에 상관관계가 명확한 데이터 쌍이 필요하다. 원본 이미지의 특징적인 정보를 유지하면서 도메인이 다른 데이터 쌍은 그 확보와 전처리에 많은 비용이 든다. Cycle GAN[1]은 사이클-일관성 손실(Cycle Consistency Loss)를 도입하여 데이터 쌍이 존재하지 않는 이미지 간의 변환에 성공한다. 하지만 색, 질감 등은 성공적으로 변환시키지만 물체의 형태 변환에는 그 효과가 미미하다.

본 논문에서는 쌍이 없는 이미지 간의 변환에서 유의미한 형태 변환이 일어나도록 Cycle GAN을 기반으로 판별자와 생성자 레이어 구조를 개선하고, 추가적인 손실함수를 도입하여 형태변환에 대한 개선점을 확인한다. 본론에서는 모델 학습에 사용할 데이터 셋, 판별자와 생성자의 개선된 합성곱 신경망 구조, 추가적인 커스텀 손실 함수에 대해 설명하고, 실험 결과에서는 동일한 학습 데이터 셋에 대한 기존 모델과 개선 모델의 결과를 비교하여 개선점을 평가한다. 결론에서는 개선 모델의 결과물에 대한 분석과 추후 개선사항에 대해 논한다.

II. 본론

1. 데이터 셋

본 논문에서는 학습을 위한 데이터 셋으로 Apple2Banana, Cat2Dog 데이터 셋을 사용한다. 분류 모델 학습을 위해 사용되는 캐글(Kaggle) 오픈 데이터 셋을 이용한 것으로 변환 모델을 위한 데이터 쌍이 존재하지 않는다. 전자는 흰 바탕의 사과와 바나나를 다양한 측면으로 보여주는 데이터 셋으로, 두 도메인

간의 형태 차이가 분명하고, 흰 바탕으로 인해 변환된 형태의 경계가 뚜렷해야 하기 때문에 전체적인 형태 변환을 확인하기 좋은 데이터 셋이다. 후자의 경우 다양한 측면에서 개와 고양이를 보여주는 데이터 셋으로 본 논문에서는 기존 이미지의 특징정보를 유지하면서 타 도메인의 특성들을 제대로 학습하는지 확인하기 위한 데이터 셋으로 사용한다.

2. 모델 개선 및 손실 함수

패치(Patch)를 기반으로 하는 기존 판별자는 구역으로 나뉜 세부 특성은 잘 판별하지만 전체적인 형태 특성을 판단하지 못한다. 이 취약점을 개선하기 위해 각 패치의 크기를 키워 전체 패치의 개수는 줄인다. 또한 팽창 레이어를 추가하여 전체적인 연산량은 유지한 채 합성곱 신경망의 필터가 더 넓은 영역의 특성(형태)에 기반해 판별할 수 있도록 한다.

생성자 레이어는 잔차 블록을 사이사이에 배치하여 잔차를 학습하도록 해 학습의 난이도를 낮춘다.

일관성 손실은 원본 이미지 A에서 도메인 B로 변환된 이미지가 A의 특징정보를 어느정도 유지할 수 있게 해준다. 그러나 적대 손실(Adversarial Loss)이 도메인 간의 변화를 야기한다면 일관성 손실은 위와 같은 특징 때문에 도메인 간의 변화를 크지 않게 만든다. 즉, 일관성 손실을 비교적 크게 만드는 형태변환은 지양하는 쪽으로 학습하게 된다.

$$\frac{1}{n-1} \sum_i \|E_{b \sim p_{data}} f_i(b) - E_{a \sim p_{data}} f_i(G_{ab}(a))\| \quad (1)$$

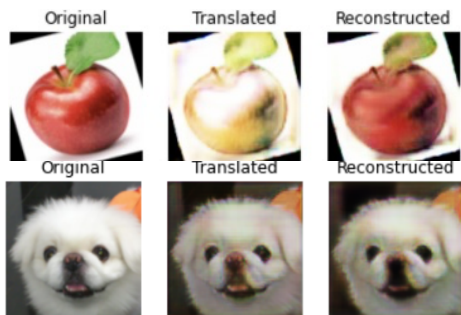
이를 보완하기 위해 판별자의 활성화 맵(Activation Map) 레벨에서 접하는 새로운 손실 함수를 도입한다. 그 손실 함수는 (1)과 같다. A를 생성자에 넣어 변환시킨 가짜 B와 도메인 B의 진짜 B 이미지를 판별자에 넣었을 때, 판별자 각 레이어의 활성화 맵 출력을 비교하여 그 차의 평균을 손실함수로 구성한다. 판별자의 각 합성곱 레이어의 출력인 특성 맵은 활성화 함수를 통해 활성화 맵으로 맵핑 된다. 그리고 활성화 맵은 판별자가 특정 도메인을 그 도메인이라고 판단하는

비선형적 특성 정보를 담고 있다. 따라서 도메인 A 에서 변환된 가짜 B 와 도메인 B 의 진짜 B 를 판별자의 입력으로 주었을 때, 활성화 맵 출력의 차이가 줄어들도록 생성자를 학습시키면 A 에서 B 로 변환시키는 생성자가 B 도메인의 특성들을 가져오는 방향으로 학습한다. 이 손실함수를 활성화 맵 손실 함수(λ_{act})라 하면 본 모델의 최종적인 손실함수는 (2) 와 같다.

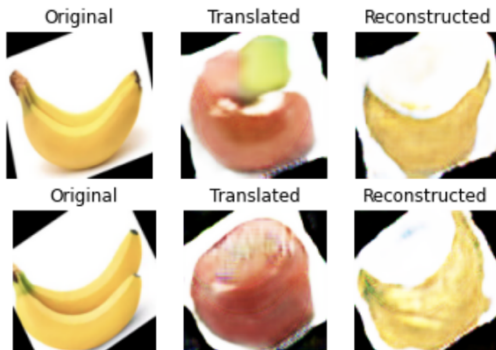
$$L_{total} = \lambda_{ad}L_{ad} + \lambda_{cyc}L_{cyc} + \lambda_{act}L_{act} \quad (2)$$

적대손실과 일관성 손실은 기존 Cycle GAN 과 동일하게 L2(mse) 손실 함수를 사용한다. 형 변환에 대한 개선점 확인이 주 목적이기 때문에 동일성 손실(Identity Loss)는 생략한다.

III. 실험 결과



[그림 1] 기존 Cycle GAN 의 결과



[그림 2] 개선 모델의 결과 (Apple2Banana)



[그림 3] 개선 모델의 결과 (Cat2Dog)

[그림 1]을 보면 기존 Cycle GAN 의 경우 두 데이터 셋 모두 색감은 잘 변환됐지만 형태 변화는 전혀 발생하지 않는다.

[그림 2]를 보면 개선 모델의 형태 변환이 매우 잘 나타난 것을 확인할 수 있다. 또한 생성된 사과의 구도가 원본 바나나 이미지의 구도 정보를 유지하여 적절한 측면 이미지로 생성되었다. 생성된 사과를 다시 바나나로 변환시키는 생성자에 넣어 재구성된 바나나 이미지가 원본 바나나 이미지의 형태, 구도를 따라간다. 이를 통해 생성자가 타 도메인의 무작위한 랜덤 이미지를 모방해 생성한 것이 아닌, 원본 이미지와의 상관관계를 유지하며 변환시킨 것을 확인할 수 있다.

[그림 3]를 보면 개에서 고양이 변환의 결과가 기존과 다르게 확실한 형태변환을 보인다. 특히 고양이 도메인의 특징이라고 할 수 있는 뽀족한 귀, 코, 날카로운 털 등이 원본 개 이미지에 추가되는 형태로 변환이 일어나는 것을 보아 활성화 맵 손실 함수가 고양이 도메인의 특성들을 학습하도록 만들어 형태 변환을 유도하는 것으로 보인다.

여러 개의 손실들 사이의 비율을 조정하여 학습하는데 어려움이 있고 그로 인해 이미지가 다소 흐릿한 모습을 보인다. 손실들 간의 비율이 학습 도중에 동적으로 조절되도록 모델을 개선한다면 더 깨끗한 결과를 도출할 수 있을 것이라 예상된다.

IV. 결론

본 논문에서는 데이터 쌍이 없이 학습되는 변환 모델의 취약점, 형태 변환을 개선시키기 위해 새로운 모델을 설계했다. 그 결과, 상관관계를 유지하며 변환되면서 취약한 부분이었었던 형태 변환이 나타나는 모습을 확인할 수 있었다. 또한 새로운 손실함수의 도입 목적대로 형태 변환이 타 도메인의 특성들을 학습하여 야기되는 것을 확인할 수 있었다. 결과물이 흐릿한 문제는 추후 개선이 필요하나 초 해상도 모델(Image Super Resolution)등과 결합하면 훌륭한 범용 이미지 변환기로 활용될 수 있을 것이다.

ACKNOWLEDGMENT

“이 논문은 국토교통부의 스마트시티 혁신인재육성사업으로 지원되었습니다.”

이 연구는 2022 년도 산업통상자원부 및 산업기술평가관리원 (KEIT) 연구비 지원에 의한 연구임('00144500')

참 고 문 헌

- [1] Unpaired image-to-image translation using Cycle-consistent Adversarial network, JY Zhu, 2017