

이중 생성자를 사용한 저용량 선화 자동채색 모델

이영섭^o, 이성진^{*}

^o경상대학교 기계항공정보융합공학부 항공우주및소프트웨어공학전공,

^{*}경상대학교 기계항공정보융합공학부 항공우주및소프트웨어공학전공

e-mail: ys2lee.ac@gmail.com^o, insight@gnu.ac.kr^{*}

A Lightweight Deep Learning Model for Line-Art Colorization Using Two Stage Generator Model

Yeongseop Lee^o, Seongjin Lee^{*}

^oDepartment of Aerospace and Software Engineering, Gyeongsang National University,

^{*}Department of Aerospace and Software Engineering, Gyeongsang National University

● 요약 ●

미디어 산업의 발전으로 스토리보드와 같은 선화 이미지의 자동채색 연구가 국내외에서 진행되고 있다. 하지만 자동채색 모델 용량에 초점을 두는 연구는 아직 진행되고 있지 않다. 기존 자동채색 연구는 모델 용량이 최소 567MB 이상으로 모델 용량이 큰 단점을 가지고 있다. 본 논문에서는 채색을 2단계로 나누는 이중 생성자 구조와 기존 U-Net을 개선한 생성자를 사용해 기존 U-Net에 비해 30%, VGG16/19를 사용한 기법과 비교해 최대 85% 작은 106MB 모델을 생성했고 FID(Fréchet Inception Distance)를 통한 이미지 평가결과 512x512px에서 153.69의 채색성능을 얻었다.

키워드: Machine Learning, GAN(Generative Adversarial Networks), Image Generation

I. Introduction

미디어 산업에서 펜 터치만 사용한 스토리보드는 감독이 의도하는 분위기를 표현하기 힘들기 때문에 채색된 스토리보드를 만든다. 이 과정에서 전문인력의 시간을 소비해야 하기 때문에 국내외에서 다양한 선화 자동채색 연구가 진행되고 있다[1, 2].

선화의 자동채색은 채색할 선화와 참고 이미지 혹은 컬러 스킴과 같은 힌트를 입력으로 이미지분할 및 컬러화에 대한 복합적인 해결방법이 필요해 어려운 문제로 분류된다. 기존 선화 자동채색은 입력에 따라 스타일 변환, 힌트 입력으로 나누어진다. 두 기법 모두 GAN (Generative Adversarial Networks)[3]에 기반을 두고 있으며, 이들 중 어느 것도 네트워크 용량에 초점을 두지 않았기 때문에 모델 용량이 최소 567MB 이상의 크기를 갖는다는 단점이 있다. 저용량 모델을 사용하는 것은 머신러닝을 사용한 온라인 서비스를 운영할 때 발생하는 모델의 추론속도, 서버 부담과 모델의 모바일 이식을 위해서 매우 중요한 일이다.

본 논문에서 모델 용량을 줄이기 위해 U-Net의 구조를 개선하고 이중 생성자 구조를 사용했다. 제안하는 기법을 사용해 생성자 모델 용량을 기존 U-Net을 사용한 모델과 비교해 약 30%, VGG16/19를 사용한 기법과 비교해 최대 85% 작은 100MB급으로 줄였다. 512x512px 해상도 이미지에서 FID 기준 기존 U-Net을 사용한 모델과 비교해 대략 7% 낮은 153.69의 채색성능을 얻었다.

II. Preliminaries

1. Related works

사용자 의도를 반영하기 위해서 많이 사용되는 채색방법으로 스타일 변환 자동채색 연구가 있다. Zhang et al.[1]은 생성자에 VGG16/19 네트워크를 사용해 압축된 참고 이미지 정보를 입력해 사용자 의도를 반영한다. 제한적으로 사용자 의도를 반영할 수 있지만 참고 이미지의 특징을 추출하기 때문에 네트워크 구성이 복잡하고 네트워크 용량이 크다는 문제가 있다.

사용자 의도를 반영하는 또 다른 채색방법으로 선화와 힌트를 입력하는 연구도 있다. Ci et al.[2]은 U-Net 구조 생성자의 디코더 레이어에 ResNextBlock을 사용하고 선화의 특징을 추출하는 네트워크(LFN)를 사용해 채색성능 향상을 하였다. 하지만 ResNextBlock과 LFN으로 인해 모델 용량이 커지는 문제가 있다.

III. The Proposed Scheme

1. 이중 생성자 구조

효율적인 채색을 위해 초안생성과 채색이라는 2단계로 나뉜 모델을 구성하였다. 각각의 작업을 전담하는 U-Net 구조의 생성자(Draft

Model, Colorization Model)는 1 단계에서 선화와 사용자 힌트를 입력받아 128x128px 해상도의 초안 이미지를 생성한다. 2 단계에서 선화와 초안 이미지의 크기를 입력받은 선화의 크기인 512x512px 해상도로 키워서 채색 작업을 진행한다. 첫 번째 생성자의 적대적 학습을 위한 분류자는 DCGAN[4]구조를 사용하였다.

2. Additive U-Net

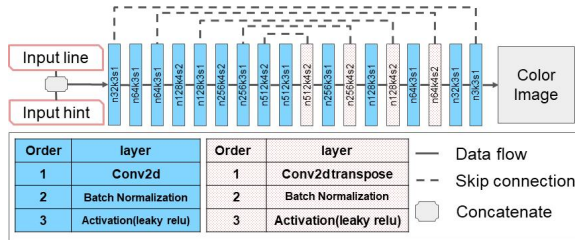


Fig. 1. AU-Net 구조

$$Parameter\# = (IF * K + 1) * OF \quad (25)$$

기존 U-Net의 경우 인코더(좌측)에서 받은 feature map을 디코더(우측)에 연결하는 구조를 가진다. U-Net을 구성하는 CNN(Convolutional Neural Network)에서 파라미터는 식 (1)과 같이 입력층 필터에 비례하기 때문에 연결하는 과정에서 디코더(우측)에 입력되는 필터 수가 늘어나 파라미터 수가 늘어나게 된다.

본 연구에서 제안하는 Additive U-Net (AU-Net, 그림 3)의 경우 추출된 특징을 연결하지 않고 더하는 방법을 사용한다. 이 방법을 사용하면 기존 U-Net과 비교해 필터 수는 반으로 줄어든다. 결과적으로 파라미터 수(모델 용량)를 줄일 수 있다.

3. 실험 및 결과

모델 용량은 생성자의 파라미터 수 그리고 직렬화시켜 저장한 생성자 모델 용량을 비교한다. 실험결과 그림 4와 같이 기존 U-Net과 비교해 약 30% 작은 13,071,072개의 파라미터가 나오는 것을 확인했다. 모델 용량의 경우 생성자 두 개를 합쳐 약 기존 U-Net 모델 용량 148MB에서 제안한 기법을 사용하면 106MB로 용량이 줄었다.

채색성능을 비교하기 위해 FID 을 사용하여 원본과 생성 이미지를 평가하였다. FID 정량법은 값이 낮을수록 높은 품질을 나타낸다. 원본에서 추출한 힌트 수(픽셀 개수)를 140개로 고정하였을 때 기존 U-net 구조와 비교해 6.89% 높은 153.69의 값을 얻었다. 하지만 약 7%의 정량적인 수치 차이로는 시각적인 결과물에서는 큰 차이를 보이지 않는다.

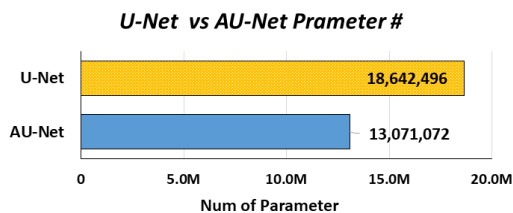


Fig. 2. 파라미터 수 비교

IV. Conclusions

본 논문에서는 선화 자동채색 모델의 경량화를 위해 기존 U-Net의 구조를 개선한 이중 생성자 채색기법을 제안했다. 실험결과 두 개의 생성자를 합쳐 106MB의 모델 용량을 사용해 기존 U-Net 구조와 비교해 FID 기준 대략 7% 높은 153.69의 채색성능을 가진다. 기존 U-Net 구조와 비교해 30% 그리고 VGG16/19를 사용한 기법[2, 3]에 비해 최대 85% 작은 모델 용량을 가지는 장점이 있다.

ACKNOWLEDGEMENT

이 논문은 2018년도 정부(교육부)의 재원으로 한국과학창의재단 (2019년도 학부생 연구프로그램)의 지원을 받아 수행된 연구이며 또한, 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구입(No. 2019R1G1A1100455).

REFERENCES

- [1] Zhang, Lvmin, Yi Ji, Xin Lin, and Chungping Liu. "Style transfer for anime sketches with enhanced residual u-net and auxiliary classifier gan." In 2017 4th IAPR Asian Conference on Pattern Recognition, pp. 506-511. IEEE, 2017.
- [2] Ci, Yuanzheng, Xinzhu Ma, Zhihui Wang, Haojie Li, and Zhongxuan Luo. "User-guided deep anime line art colorization with conditional adversarial networks." In 2018 ACM Multimedia Conference on Multimedia Conference, pp. 1536-1544. ACM, 2018.
- [3] Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." In Advances in neural information processing systems, pp. 2672-2680. 2014.
- [4] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv preprint arXiv:1511.06434, 2015.