

색의 채널을 보상하는 L1 손실 알고리즘

김범식^o, 이성진^{*}

^o경상대학교 기계항공우주공학부 항공우주및소프트웨어공학전공,

^{*}경상대학교 기계항공우주공학부 항공우주및소프트웨어공학전공

e-mail: {bumdori2, insight}@gnu.ac.kr^{o*}

Color Channel Compensating L1 Loss Algorithm

Bumsik Kim^o, Seongjin Lee^{*}

^oDept. of Aerospace and Software Engineering, Gyeongsang National University,

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

● 요약 ●

본 논문에서는 머신러닝을 이용한 이미지 생성을 위한 새로운 오차 함수모델을 제안한다. 제안된 함수모델은 기존 오차함수가 반영하지 못하던 채널 간 오차비용정보를 반영하여 기존 오차함수에 비해 빠른 초기 수렴속도와 더 좋은 FID값을 보인다. 본 논문에서는 하나의 네트워크 모델을 통해 기존의 오차함수모델에 비해서 우수함을 보인다.

키워드: AI, 기계 학습 (Machine Learning), 손실 모델 (Loss Model)

I. Introduction

최근 GPU의 발전과 딥 러닝을 활용한 기술들이 각광을 받으면서 사용자의 입력을 통해 생성자의 출력을 원하는 방향으로 조절하려는 시도가 있다[1, 2]. 이러한 접근 방향은 크게 두 가지로 나뉜다. 원본 이미지와 생성된 이미지간의 차이를 norm을 사용하여 정량화하는 방법[1]과 원본과 생성된 이미지를 미리 훈련된 네트워크에 통과시킨 뒤 norm을 사용하여 정량화하는 방법이다[2]. 이를 통해 기존의 GAN 기법보다 현실적인 이미지를 생성할 수 있었다[1, 2]. 하지만 norm만을 사용한 정량화는 색 채널이 가지는 정보를 충분히 반영하지 못하는 단점이 있다. 본 논문에서는 색 정보를 반영할 수 있는 새로운 정량화 기법을 제시한다.

II. Preliminaries

1. Related works

Image to image 변환에서 원본과 생성자의 출력 간의 차이를 줄이기 위해 오차함수를 사용하여 생성자를 조절하는 방법들이 제안되었다[1, 2, 3]. 먼저, 이미지 간의 차이를 RGB 공간에서 L1/ L2 오차 함수로 계산하는 방식이 보편적이다[1, 3]. 생성자의 오차함수는 식 (1)과 같이 단순화된 형태로 나타낼 수 있다.

$$L_{total} = \lambda_{img} L_{img} + \lambda_{adv} L_{adv} \quad (1)$$

여기서 L_{img} 는 feature space에서의 이미지간 차이, L_{adv} 는

adversarial loss를 나타낸다. L_{img} 를 통해서 학습을 안정화[3] 했고 실제에 가까운 이미지[1]를 생성할 수 있었다. Ci et al[2]는, 색 공간 외의 특징을 반영하기 위해 식 (2)와 같은 오차함수를 사용했다[2].

$$L_{feat} = \sum_i^n |C(G(x_i)) - C(y_i)| \quad (2)$$

여기서 C 는 feature space로 매핑하는 함수, G 는 생성자, x 는 noise vector, y 는 ground truth이다.

III. The Proposed Scheme

1. 제안 기법

관련 연구들이 제안하는 오차 함수 기법들은 RGB 색 공간 전체를 feature space로 본다. 이 접근은 어떤 오차함수를 사용해도 오차는 RGB 채널 간 오차의 평균이기 때문에 오차 비율 정보는 버려진다는 문제가 있다. 이런 문제를 해결하기 위해 본 논문에서는 각 채널을 독립된 feature space로 보는 새로운 오차함수를 제안한다. 먼저, 각 채널의 L1 loss 식 (3)과 같이 구한다. 여기서 RGB는 채널, t 는 ground truth, f 는 생성자의 출력을 의미한다.

$$l_{R,G,B} = \frac{1}{n} \sum_{i=0}^n |t_{R,G,B_i} - f_{R,G,B_i}| \quad (3)$$

식 (3)에서 구한 각 채널의 L1 loss를 모두 더한 뒤 이를 채널 오차를 나눠 비율을 구하면 다음과 같다.

$$l_{RGB} = l_R + l_G + l_B, w_{R,G,B} = \frac{l_{R,G,B}}{l_{RGB}} \quad (4)$$

식(4)을 통해 비율이 보정된 L1 loss를 구하면 다음과 같다.

$$l_{comp} = w_R l_R + w_G l_G + w_B l_B \quad (5)$$

제안하는 기법을 실현하기 위해 다음과 같은 목적함수를 사용했다[4].

$$V_{GAN}(G, D) = E_x [-\log(D(x))] + E_{y,z} [-\log(D(G(y,z)))] \quad (6)$$

(6)에 (5)를 적용하여 다음과 같은 목적 함수를 추가하면 생성자의 학습 방향을 원하는 방향으로 유도할 수 있다[1].

$$V_{L_1}(G) = E_{l,y,z} [|l_{comp}|_1] \quad (7)$$

식 (6)과 (7)을 고려하면 식 (8)과 같은 최종 목적함수를 얻는다.

$$G^* = \arg \min_G \max_D L_{GAN}(G, D) + \lambda V_{L_1}(G) \quad (8)$$

식 (6), (7), (8)에서 x는 원본 이미지이고, y는 원본이미지의 특징을 추출한 선화 이미지, z는 원본이미지에서 추출한 컬러 픽셀이다. 식(8)에서 λ는 L1 loss 가중치이다.

2. 실험환경 및 모델

제안된 기법의 실험은 Intel i7, 64GB RAM, RTX2080Ti, Tensorflow로 진행했고, 사용된 CNN 모델[4]은 그림 1과 같이, U-Net 구조를 사용했고 원본에서 약간의 변형을 주었다. 데이터 셋은 자체 데이터 셋 150만장을 사용했다. ADAM을 사용해 최적화 했고 배치는 4이며 하이퍼파라미터는 다음 식 (9)과 같다.

$$\beta_1 = 0.9, \beta_2 = 0.99, lr = 1e - 4, \lambda = 100 \quad (9)$$

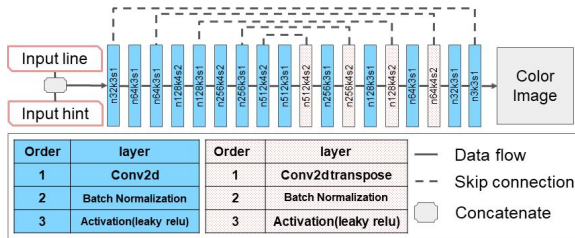


Fig. 1. Network model 구조

3. 결과

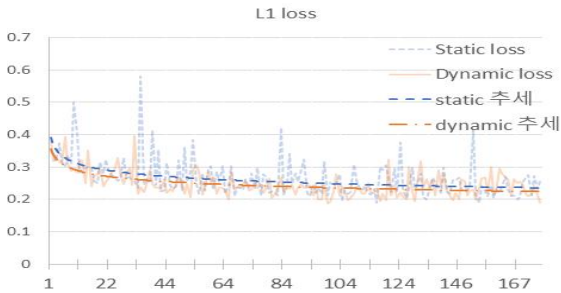


Fig. 2. 검증 오차

Table 1. 실험 결과

	제안된 기법	기존 기법
L1 loss 값	0.17352	0.17346
FID 값	307.87	351.02
평균 기울기 값	-1.3	-1(기준값)

동일한 함수로 오차를 정량화 한 결과는 그림 2와 같다. 여기서 오차의 차이를 평균으로 계산한 결과 약 8%가량 낮은 값을 가졌다.

IV. Conclusions

이후 진행된 학습에서 유사한 L1 loss값을 기준으로 이미지를 샘플링해 FID(frechet inception distance)를 계산한 결과, 다음 표 1과 같은 결과를 얻었다.

제안된 기법을 이용하여 빠른 수렴 속도와 좋은 FID값을 얻었다. 이후 다른 네트워크 모델을 통한 추가적인 검증은 할 것이다.

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 연구임(No. 2019R1G1A1100455).

REFERENCES

- [1] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. "Image-to-image translation with conditional adversarial networks," In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 1125-1134, 2017.
- [2] Ci, Y., Ma, X., Wang, Z., Li, H., & Luo, Z. (2018, October). User-guided deep anime line art colorization with conditional adversarial networks. In ACM Multimedia Conference on Multimedia Conference. p. 1536-1544, 2018.
- [3] Dosovitskiy, A., & Brox, T. (2016). Generating images with perceptual similarity metrics based on deep networks. In Advances in neural information processing systems (pp. 658-666).
- [4] Yeongseop Lee, Seongjin Lee. "Automatic Colorization of Anime Style Sketch using User-Hint and Deep Neural Network." KCC2019: 736-738. 2019