

판별자를 활용한 적대적 생성 신경망 프루닝

이동준, *이승현, **송병철

인하대학교, *인하대학교, **인하대학교

dj31665@gmail.com, *lsh910703@gmail.com, **bcsong@inha.ac.kr

Generative Adversarial Network Pruning using Discriminator

Dongjun Lee, *Seunghyun Lee, **Byungcheol Song

Inha University, *Inha University, **Inha University

요 약

본 논문에서는 판별자를 활용하여 Image to Image translation(I2I) 분야에서 사용되는 적대적 생성 신경망(GAN)을 압축하는 방법을 제시한다. 우선, 잘 학습된 판별자와 생성자 사이의 adversarial loss 를 활용하여 생성자 내 필터들의 중요도 점수를 매겨준다. 그리고 생성자 내의 필터들을 중요도 점수를 기준으로 나열한 후 점수가 낮은 필터들을 제거하는 필터 프루닝을 한번 수행하여 적은 시간 비용으로 생성자를 압축한다. 마지막으로 지식 증류를 활용해 압축된 생성자를 학습시켜 기존의 생성자와 유사한 성능을 보이도록 하였다. 이 과정들을 통해 효과적이고 빠르게 GAN 모델을 압축할 수 있음을 확인하였다.

1. 서론

최근 적대적 학습(adversarial training)을 통해 영상을 생성, 합성하는 GAN^[1] 모델은 엄청난 발전을 이루어 왔고 영상의 도메인을 변환하는 I2I 분야에서 많이 활용되고 있다. 하지만 이러한 발전과 함께 GAN 모델의 연산량, 파라미터 수 또한 엄청나게 증가하였고 그로 인해서 스마트폰과 같은 자원이 제한된 환경에서는 GAN model 이 적용되기 힘들게 되었다.

위의 문제들을 해결하기 위해 GAN 모델을 압축하려는 여러 연구들이 진행되어 왔다. 그 중 GAN compression^[2] 기법은 Neural Architecture Search(NAS)를 통해 성능의 손실을 최소화하면서 모델을 상당히 압축하였다. 하지만 그 과정이 매우 느리고 비용이 상당히 크다는 단점이 있다.

본 논문에서는 GAN model 을 학습시킬 때 같이 학습되는 판별자를 활용하여 효과적이고 빠르게 GAN model 을 프루닝하는 방안을 제안한다. 본 논문의 구성은 다음과 같다. 2 절에서는 GAN

기반 I2I 모델인 CycleGAN^[3]과 GAN 경량화 기법인 GAN compression 에 대해 설명한 후 3 절에서는 본 논문의 기법에 대해 설명한다. 그리고 4 절에서는 제안된 기법의 성능을 확인하고 5 절에서는 결론을 맺는다.

2. CycleGAN, GAN compression

CycleGAN 은 적대적 학습을 통해 생성자 A 가 도메인 X 인 이미지가 입력되었을 때 타겟 이미지의 도메인 Y 로 변환하는 법을 학습한다. 이때 도메인 Y 의 이미지를 X 로 바꾸는 생성자 B 를 동시에 학습하여 도메인 X 에서 Y 로 변환된 이미지가 다시 도메인 X 로 돌아올 수 있도록 한다. 그리고 두 생성자에 의해 도메인 X 로 되돌아온 이미지는 입력된 이미지와 동일해야 한다는 cycle consistency loss 를 통해 쌍을 이루지 않는 데이터셋에서도 학습이 가능하도록 한다.



그림 1. 말을 얼룩말로 변환하는 작업에 대한 CycleGAN, MobileNet기반 CycleGAN, GAN compression, 본 논문 기법의 결과

기법	MACs	#Params	FID(↓)
CycleGAN	56.8G	11.30M	61.53
CycleGAN*	18.3G	1.98M	59.29
GANcomp*	2.67G	0.34M	64.95
Ours*	2.39G	0.32M	71.60

표 1. CycleGAN과 MobileNet기반 CycleGAN, GAN compression, 본 논문 기법의 정량적 평가 비교
*표시는 MobileNet기반의 모델

GAN compression 은 GAN 을 활용한 I2I 모델을 압축하는 기법으로 우선 기존 ResNet 기반의 CycleGAN 모델을 MobileNet^[4] 기반으로 변환하여 성능을 유지한 채 연산량을 효과적으로 줄였다. 그리고 적대적 학습과 channel distillation, reconstruction loss 를 활용해 MobileNet 기반의 supernet 을 학습한 후 NAS 를 통해 효과적이고 압축된 모델 아키텍처를 찾는다. 하지만 NAS 과정은 상당한 시간과 비용을 요구하게 된다. 본 논문에서는 GAN compression 과 마찬가지로 ResNet 기반의 CycleGAN 모델을 MobileNet 기반의 CycleGAN 모델로 변환한 후 프루닝했다.

3. 본론

적대적 학습은 생성자가 판별자를 속이고 판별자는 이를 구별하려고 하면서 학습이 진행된다. 본 논문에서는 CycleGAN 모델 내의 필터들을 제거했을 때 학습된 판별자를 속이는 능력이 덜 감소하도록, 즉 생성자 관점에서의 adversarial loss 가 덜 증가하도록 중요도 점수를 매겼다. 이를 위해 [5]와 같이 Taylor expansion 을 활용한 결과 필터가 출력하는 feature map 의 값과 gradient 의 곱으로 그 필터의 중요도를 매길 수 있다는 것을 확인하였다. 이는 합성곱 레이어 뒤에 마스크를 붙이고 마스크의 gradient 를 구하면 쉽게 얻을 수 있다. 이렇게 필터들의 점수를 매기기 위해 파라미터의 업데이트 없이 1 에폭 정도 데이터를 입력하고 역전파를 통해 마스크들의 gradient 를 구했다.

위와 같은 방법으로 필터들의 점수를 한번 만 매긴 후

나열한 다음, binary search 를 통해 타겟 연산량을 만족할 수 있는 threshold 를 찾았다. 그리고 이 threshold 보다 낮은 점수를 갖는 필터들을 제거하였다. 이렇게 하여 한번의 점수 매기기와 binary search 를 통해 NAS 과정보다 훨씬 빠르고 적은 비용으로 프루닝을 진행하였다.

한 스텝의 프루닝을 거친 후 압축된 모델의 망가진 성능을 기존 모델 수준으로 복구하기 위해 기존의 적대적 학습에 더해 channel distillation, perceptual loss, SSIM loss 를 활용하였다. 그리고 압축된 생성자를 적대적 학습할 때 미리 학습된 판별자를 활용하였다.

4. 실험 및 결과

본 실험에서는 horse2zebra 데이터 셋을 활용하여 말 영상을 얼룩말 영상으로 변환하는 작업을 수행하였다.

표 1 에서와 같이 ResNet 기반의 CycleGAN 을 MobileNet 으로 변환하면서 약 3 배 정도의 연산량을 줄이는 동시에 정량적으로 더 좋은 성능을 보일 수 있었다. 그리고 GAN compression 기법과 본 논문의 기법을 정량적으로 비교했을 때 본 논문의 기법이 GAN compression 기법보다 압축률은 높았지만 좀 더 낮은 성능을 보였다. 하지만 정성적으로 비교했을 때 큰 차이를 보이지 않았다.

5. 결론

본 논문에서는 판별자를 활용한 adversarial loss 를 통해 필터 프루닝을 진행하면 효과적이고 빠르게 생성자를 압축할 수 있음을 보였다. 정량적인 평가는 기존의 GAN compression 보다 조금 뒤쳐지지만 압축률은 조금 증가했으며 정성적으로 평가시에도 큰 차이가 없었다. 하지만 압축하는 시간과 그 비용의 경우 NAS 를 활용하는 GAN compression 과 비교했을 때, 거의

없다시피 할 정도로 줄일 수 있었다.

6. Acknowledgement

The authors greatly thank the three anonymous reviewers for their kind comments which have really helped improve the quality of our paper. This work was supported by IITP grants funded by the Korea government (MSIT) (No. 2021-0-02068, AI Innovation Hub and RS-2022-00155915, Artificial Intelligence Convergence Research Center (Inha University)), and was supported by the NRF grant funded by the Korea government (MSIT) (No. 2022R1A2C2010095 and No. 2022R1A4A1033549).

7. 참고문헌

- [1] I.Goodfellow, J.Pouget-Abadie, M.Miraza, B.Xu, D.Warde-Farley, S.Ozair, A. Courville, Y.Bengio, "Generative Adversarial Nets", in Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS), 2014, pp.2672-2680.
- [2] M. Li, J.Lin, Y.Ding, Z.Liu, J.Zhu, S.Han, "GAN Compression: Efficient Architectures for Interactive Conditional GANs", in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp.5284-5294.
- [3] J.Zhu, T.Park, P.Isola, A.Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks" in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2223-2232.
- [4] A.Howard, M.Zhu, B.Chen, D.Kalenichenko, W.Wang, T.Weyand, M.Andreetto, H.Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications", arXiv preprint arXiv:1704.04861, 2017.
- [5] P.Molchanov, S.Tyree, T.Karras, T.Aila, J.Kautz, "Pruning Convolutional Neural Networks for Resource Efficient Inference", in Proceedings of the International Conference on Learning Representations (ICLR), 2017.