

폐기물 분류 개선을 위한 이미지 생성 모델 비교 분석

고준혁¹, 박정현¹, 김시웅¹, 문남미²

¹호서대학교 컴퓨터공학과 석사과정

²호서대학교 컴퓨터공학부 교수

junhyeok970306@gmail.com, jh.park970609@gmail.com,

kimsiung990811@gmail.com, nammee.moon@gmail.com

Comparative Analysis of Image Generation Models for Waste Recognition Improvement

Jun Hyeok Go, Jeong Hyeon Park, Siung Kim, Nammee Moon
Dept. of Computer Science and Engineering, Hoseo University

요 약

이미지 기반 폐기물 처리시스템에서 품목별 상이한 수집 난이도로 인해 발생하는 데이터 불균형으로 분류 모델 학습에 어려움이 따른다. 따라서 본 논문에서는 폐기물 분류 모델의 성능 비교를 통해 적합한 이미지 생성 모델을 탐색한다. 데이터의 불균형을 해결할 수 있도록 VAE(Variational Auto-Encoder), GAN(Generative Adversarial Networks) 및 Diffusion Model을 이용하여 이미지를 생성한다. 이후 각각의 생성 방법에 따라 학습데이터와 병합하여 객체 분류를 진행하였다. 정확도는 VAE가 84.41%로 3.3%의 성능 향상을, F1-점수는 Diffusion Model이 91.94%로 6.14%의 성능 향상을 이루었다. 이를 통해, 데이터 수집에서 나타나는 데이터 불균형을 해결하여 실 사용환경에 알맞은 시스템을 구축이 가능함을 확인하였다.

1. 서론

최근 폐기물 처리 문제는 전 세계적으로 증가하고 있는 문제 중 하나이다[1]. 이에 대한 대안으로 딥러닝을 이용한 이미지 기반 폐기물 분류 시스템을 구축하는 것이 제안되었다[2].

자동화 시스템을 구현하기 위해서는 객체 분류가 핵심적인 요소이다. 하지만 폐기물은 종류가 다양하고, 수집 난이도가 품목마다 달라 데이터 불균형이 발생하여 모델 성능 저하 요소가 된다.

이러한 문제를 해결하기 위해, 본 논문에서는 VAE(Variational Auto-Encoder), GAN(Generative Adversarial), Diffusion Model 기반의 증강방식을 실험함으로써 이미지 생성 모델 성능을 비교한다.

직접 수집한 이미지 데이터를 이용하여 학습 데이터 셋을 구성하고, 다양한 형태를 가진 이미지를 생성한다. 이후에 각 증강 방법에 대해 객체 분류 성능 검증 및 평가를 진행하여 데이터의 불균형을 가장 잘 해결한 생성 모델을 탐색한다. 이를 통해 데이터 증강 과정에서 적절한 생성 모델을 이용한 데이터 생성으로 다양한 객체와 모양을 가진 폐기물에 대한 인식 정확도 향상을 기대할 수 있다.

2. 관련 연구

2.1 객체 분류

객체 분류는 컴퓨터 비전 기술로, 데이터를 인공 신경망에 학습시켜 객체를 분류하는 방법이다. 최근 객체 분류를 사용하여 폐기물을 분류하려는 다양한 시도가 있었다[3].

ResNet은 깊은 신경망을 효율적으로 학습시키기 위한 분류 모델 중 하나로, Skip Connection을 이용하여 기울기 소실 문제를 해결하였다[4]. 본 연구에서는 폐기물 생성 이미지 성능 평가로 보편적으로 많이 사용되는 ResNet을 사용한다.

2.2 데이터 생성

데이터 생성은 기존의 데이터를 이용해 이미지를 다양한 방식으로 증가시켜 데이터 불균형을 해소하는 방법이다. 최근 연구로 생성 네트워크 모델을 사용해 데이터를 생성하여 모델의 일반화 성능을 향상시키는 방법이 대두 되었다[5].

본 논문에서는 각 생성 네트워크 모델 중 성능이 뛰어난 VQ-VAE[6], BIGGAN[7], DDPM[8] 중 어느 모델이 객체 분류 모델의 성능 향상에 가장 큰 기여를 하는지 실험하고자 한다.

3. 수집 데이터셋 및 전처리

데이터 생성 및 학습에 사용할 데이터는 직접 수집한 폐기물 이미지를 사용하였다. 수집 데이터의 품목과 개수는 <표 1>과 같다.

<표 1> 수집한 데이터의 품목과 개수

자전거	소화기	키보드	빈화분	의자	헤어 드라이기	합계
1000	923	657	788	1000	823	
도마	아령	가방	다리미	전자레인지	선풍기	9,059
474	325	1000	951	493	625	

아령, 전자레인지, 도마 품목은 다른 데이터에 비해 적게 수집되어 데이터 불균형이 발생하였다. 데이터 불균형 해결하기 위해, 데이터가 500개 이하 품목인 아령, 전자레인지, 도마에 생성 모델을 적용하였다. 생성 모델이 이미지의 특징점을 원활하게 학습하기 위해, 학습 이미지를 128*128 pixel 크기로 통일하고, 정규화하는 전처리를 진행하여, 이미지 배치의 일반화 성능을 향상시켰다.

4. 실험

4.1 데이터 증강

본 연구에서는 생성 모델을 이용하여 이미지 데이터를 생성한다. BIGGAN, VQ-VAE, DDPM과 <표 2>의 파라미터를 사용하여 생성 모델을 학습하였다.

<표 2> 생성모델 학습 파라미터

Model	VQ-VAE	BIGGAN	DDPM
ImageSize	128 pixel	128 pixel	128 pixel
Batch	32	32	64
Loss	MSE	hinge	L1
learning rate	0.0004	D 0.0002 G 0.00005	0.0004

$$q(x_{1:T}|x_0) := \prod_{t=1}^T q(x_t|x_{t-1}) \quad q(x_T|x_{t-1}) := N(\sqrt{1-\beta_t}x_{t-1}, B_t I) \quad (식1)$$

$$p_\theta(x_{0:T}) := p(x_T) \prod_{t=1}^T p_\theta(x_{t-1}|x_t) \quad p_\theta(x_{t-1}|x_t) := N(\mu_\theta(x_t, t), \sum_\sigma(x_t, t)) \quad (식2)$$

(식1)은 DDPM의 확산 프로세스, (식2)는 역프로세스로 (식1)을 이용하여 이미지에 가우시안 노이즈를 추가하고, (식2)를 이용하여 초기 이미지의 확률 분포를 학습하였다[8].

$$V_D(\hat{G}, D) = E_{x \sim p_{data}(x)} [\min(0, -1 + D(x))] + E_{z \sim p_z(z)} [\min(0, -1 - D(\hat{G}(z)))] \quad (식3)$$

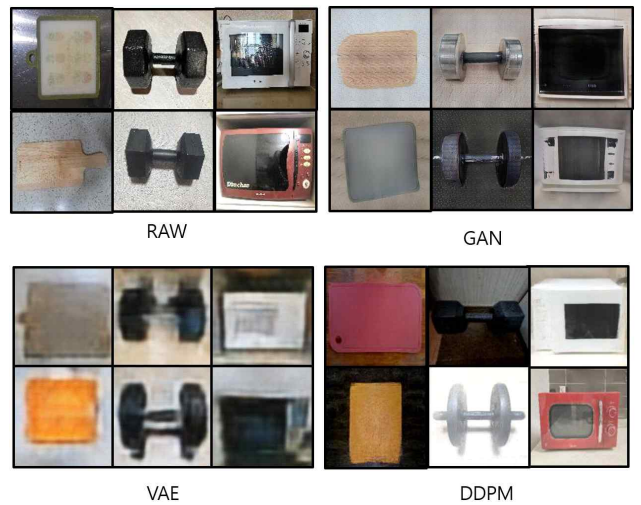
$$V_D(G, \hat{D}) = -E_{z \sim p_z(z)} [\hat{D}(G(z))]$$

(식3)은 BIGGAN의 판별자의 손실 함수로 hinge Loss를 사용한다[9]. 생성자가 생성한 가짜 이미지와 판별자의 진짜 이미지에 대한 출력값에 따라 학습 유무를 결정하여, GAN모델의 학습 불균형 문제를 해결하였다.

$$L = \log p(x|z_q(x)) + \left| |sg[z_e(x)] - e| \right|_2^2 + \beta \left| |z_e(x) - sg[e]| \right|_2^2, \quad (식4)$$

(식4)는 VQ-VAE의 손실함수로, 식을 이용하여 VQ-VAE의 인코더와 디코더를 학습하였다[6].

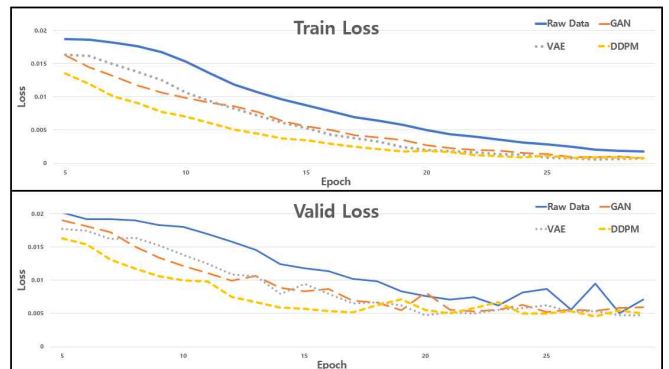
위의 파라미터와 수식으로 학습한 생성모델 이미지는 (그림 1)과 같다.



(그림 1) 생성된 폐기물 이미지 데이터 결과

4.2 분류 모델 학습

수집 데이터에 대해 8:2로 학습데이터와 검증데이터로 분류하였으며, ResNet50 분류 모델을 이용하여 학습을 진행하였다. 학습 횟수 30, 배치 크기는 128, 학습률은 0.1로 학습하였다. 학습 중의 손실 값은 (그림 2)와 같다.



(그림 2) ResNet 학습 손실 그래프

검증 데이터셋에서 손실값은 원본데이터 0.005, GAN 0.00479, DDPM 0.00453, VAE 0.0047로 DDPM이 가장 낮은 손실 값을 기록하였다. DDPM, GAN, VAE의 증강 데이터셋은 학습시에 안정적으로 손실 값이 수렴하는 모습을 확인할 수 있으나, 비증강 데이터셋은 간헐적으로 손실 값이 상승하는 현상을 발견할 수 있다.

4.3 학습 결과

학습 결과 검증 데이터의 정확도, F1-점수는 <표 4>와 같다.

<표 4> 학습의 정확도, F1-점수 결과

Model	정확도	F1-점수
Raw Data	81.16	85.8
VQ-VAE	84.41	86.97
BIGGAN	81.67	90.50
DDPM	84.16	91.94

정확도는 VQ-VAE를 활용한 증강이, F1-점수는 BIGGAN이 제일 높았다.

원본 데이터 학습 모델과 증강 데이터 모델의 차이는 정확도가 최대 3.3%, F1-점수는 최대 6.14%로 정확도보다 F1-점수가 더 많이 차이 나는 것을 확인할 수 있다. 생성 모델을 활용한 증강 방법이 모델의 손실 값을 안정적으로 학습시킬 뿐 아니라, 데이터 불균형 문제에 민감한 F1-점수에서도 높은 점수 상승률을 보였다. 폐기물 데이터 불균형 문제에 대해서도 효과적인 학습 방법이 될 것으로 판단된다.

5.3 결론

본 논문은 수집된 폐기물 데이터를 VQ-VAE, BIGGAN, DDPM 모델을 활용하여 불균형 문제를 해결하고자 하였다. 실험에서 DDPM 모델이 분류 모델 성능 비교에서 가장 뛰어난 성능을 보여주었다. 그 다음 BIGGAN, VAE 순으로 불균형 문제를 잘 해결하는 모델로 나타났다.

본 연구의 결과는 데이터 수집을 진행할수록 커지는 객체 간의 데이터 불균형을 효과적으로 해결하는데 의의가 있다. 하지만, 이미지 기반 폐기물 처리 시스템에 활용하기 위해서는 더 많은 종류의 객체를 안정적으로 분류할 필요가 있다. 추후, 후속 연구로 다양한 객체에 대한 학습 데이터를 추가 수집함으로써 안정적으로 데이터를 분류하는 연구가 진행하고자 한다.

ACKNOWLEDGEMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(No. 2019-0-01834).

참고문헌

- [1] Minelgaité, A., Liobikienė, G., “Waste problem in European Union and its influence on waste management behaviours”, Science of the Total Environment, Vol. 667, pp. 86-93, 2019.
- [2] 김형주, 박찬, 박정현, 김진아, 문남미, “GAN 기반 데이터 증강을 통한 폐기물 객체 인식 모델 설계” 한국방송미디어공학회 하계학술대회, 한국방송미디어공학회, 2022, pp. 571-573
- [3] Thokrairak S., Thibuy K., Jitngernmadan P., “ValuableWaste Classification Modeling based on SSD-MobileNet”, In2020-5th International Conference on InformationTechnology (InCIT), pp. 228-232, 2020.
- [4] He K., Zhang X., Ren S., Sun J., “Deep residual learning for image recognition.”, Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770-778, 2016.
- [5] Shorten, Connor, Taghi M., Khoshgoftaar., “A survey on image data augmentation for deep learning.”, Journal of big data, Vol. 6, 1, pp. 1-48, 2019.
- [6] Van Den Oord A., Vinyals O., Kavukcuoglu K., “Neural discrete representation learning.”, Advances in neural information processing systems, Vol. 30, 2017.
- [7] Brock A., Donahue J., Simonyan K., “Large scale GAN training for high fidelity natural image synthesis.”, arXiv preprint arXiv:1809.11096, 2018.
- [8] Ho ., Jain A., Abbeel P., “Denoising diffusion probabilistic models.”, Advances in Neural Information Processing Systems, Vol. 33, pp. 6840-6851, 2020.
- [9] Jae Hyun L., Jong Chul Y., “Geometric gan.”, arXiv preprint arXiv:1705.02894, 2017.