

데이터 균형을 위한 Chat-GPT와 Diffusion Model 기반 폐기물 생성모델 설계

김시웅¹, 고준혁¹, 박정현¹, 문남미²

¹호서대학교 컴퓨터공학과 석사과정

²호서대학교 컴퓨터공학부 교수

kimsiung990811@gmail.com, junhyeok970306@gmail.com,

jh.park970609@gmail.com, nammee.moon@gmail.com

Design of a Waste Generation Model based on the Chat-GPT and Diffusion Model for data balance

Siung Kim, Junhyeok Go, Jeonghyeon Park, Nammee Moon
Dept. of Computer Science and Engineering, Hoseo University

요 약

데이터의 균형은 객체 인식 분야에서 영향을 미치는 요인 중 하나이다. 본 논문에서는 폐기물 데이터 균형을 위해 Chat-GPT와 Diffusion model 기반 데이터 생성 모델을 제안한다. Chat-GPT를 사용하여 폐기물의 속성에 해당하는 단어를 생성하도록 질문하고, 생성된 단어는 인코더를 통해 벡터화시킨다. 이 중 폐기물과 관련 없는 단어를 삭제 후, 남은 단어들을 결합하는 전처리 과정을 거친다. 결합한 벡터는 디코더를 통해 텍스트 데이터로 변환 후, Stable Diffusion model에 입력되어 텍스트와 상응하는 폐기물 데이터를 생성한다. 이 데이터는 AI Hub의 공공 데이터를 활용하며, 객체 인식 모델인 YOLOv5로 학습해 F1-score와 mAP로 평가한다.

1. 서론

데이터의 균형은 객체 인식 분야에서 영향을 미치는 요인 중 하나이다. 폐기물을 인식하는 연구는 컴퓨터 비전분야에서의 다양한 객체 인식 알고리즘을 사용되고 있다. 객체인식 알고리즘 중에서 YOLO(You Only Look Once)v5의 mAP(mean Average Precision)가 92%로 다양한 알고리즘 중 가장 높은 성능을 기록하였다[1].

또한, 불균형한 폐기물 데이터를 조정하는 연구에는 GAN(Generative Adversarial Networks) 기반 데이터 증강을 진행하여 폐기물을 인식하는 방법을 제안하였다[2]. 하지만 GAN은 이미지를 생성할 때 정보손실에 대한 문제점을 가지고 있어 원하는 이미지를 정확하게 생성하지 못하는 경우가 발생할 수 있다[3].

본 논문에서는 정보손실로 이미지 정확도가 하락하는 것을 방지하기 위해 GAN보다 손실이 적은 Diffusion model을 사용해 이미지의 정확도를 향상하고자 한다. 이렇게 생성한 데이터로 불균형 폐기물 데이터 균형 유지 모델을 제안한다.

2. 관련 연구

2.1 Chat-GPT

Chat-GPT는 GPT라는 Transformer 기반의 대규모 텍스트 데이터를 학습하여 만들어진 언어모델을 사용하여 채팅을 진행하는 챗봇이다. 이 챗봇은 입력된 문장을 토큰화(Tokenization)하여 각 단어를 벡터 형태로 변환 후, 여러 층의 신경망에 통과시켜 문장의 의미를 추출한다. 또한, 추출된 의미를 바탕으로 새로운 문장을 생성하는데 GPT는 언어 모델(Language Model)이라는 개념을 사용하여 다음 단어가 무엇인지 예측한다. 그 후, 예측한 단어를 다시 입력으로 사용하는 것을 반복하여 최종적으로 완성된 응답을 출력한다[4].

본 논문은 폐기물과 접목할 속성에 대한 단어를 생성하도록 질문을 입력하여 답변을 생성하는 Chat-GPT를 사용하여 Stable Diffusion model의 입력으로 사용할 텍스트 데이터 생성을 제안한다.

2.2 Diffusioun Model

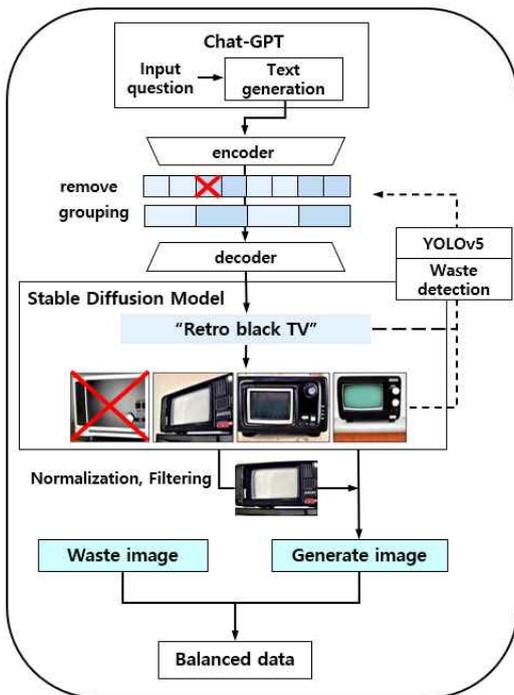
Diffusion model은 이미지에 임의의 가우시안 노이즈(Gaussian noise)를 추가하여 이미지를 생성한

다. 이렇게 생성된 노이즈 데이터를 다시 노이즈가 없는 데이터로 돌아오는 과정에서, 이전 단계에서 생성된 이미지의 정보를 추가하여 새로운 이미지를 생성하는 모델이다. 반면, Stable Diffusion model은 이미지에 일정한 노이즈를 추가한 후, 제거하는 과정에서 이전 단계에서 생성된 이미지와 현재 생성된 이미지를 비교하여 이미지를 생성하는 모델이다 [5][6].

Stable Diffusion model은 이미지 생성 시 안정화 과정을 거쳐 효과적으로 노이즈를 제거하여 신뢰성과 다양성을 보장한다. 따라서 본 논문에서는 Stable Diffusion model을 사용하여 데이터 불균형 문제 해결을 제안한다.

3. 폐기물 데이터 생성 모델

(그림 1)은 본 논문에서 제안하는 모델의 과정이다. 먼저, Chat-GPT를 사용해 대형 폐기물의 속성을 표현할 텍스트 데이터를 생성한다. 생성된 텍스트 데이터를 인코딩, 삭제, 그룹화 순으로 전처리 후, 디코더를 통해 텍스트로 출력하여 Stable Diffusion model의 입력에 사용한다. 입력된 텍스트와 상응하는 이미지를 생성 후, 기존 이미지와 생성된 이미지를 YOLOv5를 사용하여 학습하여 점수를 도출한다. 최종적으로 생성된 데이터는 불균형한 데이터셋의 입력으로 사용한다.



(그림 1) 모델 과정

3.1 Chat-GPT 기반 텍스트 생성

폐기물과 접목할 텍스트를 생성하기 위해서는 명사를 꾸며주는 Adjective나 Color같은 속성이 필요하다. Chat-GPT를 사용하여 폐기물에 해당하는 명사를 꾸며주는 Adjective와 Color를 생성한 후, <표 1>과 같이 표현하였다.

<표 1> Chat-GPT로 생성한 텍스트

| | Adjective | Color |
|----|-------------|--------|
| 1 | Vintage | Red |
| 2 | Classic | Blue |
| 3 | Retro | Yellow |
| 4 | Antique | Green |
| 5 | Elegant | Brown |
| 6 | Nostalgic | Black |
| 7 | Stylish | White |
| 8 | Discarded | Silver |
| 9 | Distinctive | Gold |
| 10 | Recyclable | Gray |

3.2 텍스트 전처리

각각의 Adjective, Color 텍스트를 벡터로 변환하기 위해 워드임베딩(word embedding)을 사용한다. 벡터화된 단어 중 폐기물 텍스트와 접목되지 않는 불필요한 단어를 삭제하고, Adjective와 Color를 그룹화하는 전처리 작업을 진행한다. 전처리된 데이터를 디코더를 통해 텍스트로 변환함으로써 Stable Diffusion model의 입력으로 한다.

3.3 이미지 생성

텍스트 데이터를 Stable Diffusion model의 입력으로 이미지를 생성한다. 이 때, 노이즈나 왜곡이 발생한 이미지에 대해 정규화, 필터링 등을 통해 복원을 진행한다. 이후, AI Hub의 공공 데이터와 함께 YOLOv5로 학습하여 F1-score와 mAP를 통해 생성 데이터가 실제 데이터와 얼마나 유사한지 확인한다. 이 중에서 실제 이미지와 유사하지 않은 생성 이미지 즉, 텍스트 데이터와 매치 되지 않는 이미지를 생성한 Adjective와 Color의 그룹은 삭제하거나 재그룹화 한다. 이러한 텍스트 재구성 과정을 통해 Stable Diffusion model이 텍스트 데이터와 매치되는 이미지를 생성하도록 개선한다.

4. 결론

본 논문에서는 Chat-GPT와 Stable Diffusion model을 사용하여 이미지를 생성하는 시스템을 제안하였다. Chat-GPT를 통해 생성된 텍스트는 전처

리 과정을 거쳐 Stable Diffusion model의 입력으로 사용하였다. 이 모델은 입력된 텍스트에 상응하는 이미지를 생성하며, 텍스트 재구성 과정으로 모델을 개선하는 방법을 제안하였다.

논문에서 제안한 방법을 통해 생성된 데이터를 활용한다면 폐기물 데이터의 균형을 유지 할 수 있을 것으로 생각되며, 이를 통해 폐기물 인식 성능향상을 기대할 수 있다.

ACKNOWLEDGEMENT

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

참고문헌

- [1] N. Singh, P. Sastry, N. B S, A. Sinha and U. V, "Performance Analysis of Object Detection Algorithms for Waste Segregation." 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), India, 2023, pp. 940-945.
- [2] 김형주, 박찬, 박정현, 김진아, 문남미. "GAN 기반 데이터 증강을 통한 폐기물 객체 인식 모델 설계." 한국방송미디어공학회 학술발표대회 논문집, 2022, pp. 571-573.
- [3] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. "Improved techniques for training gans." Advances in neural information processing systems, 2016.
- [4] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. "Language models are few-shot learners." Advances in neural information processing systems, 2020, pp. 1877-1901.
- [5] Ho, J., Jain, A., & Abbeel, P. "Denoising diffusion probabilistic models." Advances in Neural Information Processing Systems, 2020, pp. 6840-6851.
- [6] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. "High-resolution image synthesis with latent diffusion models." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10684-10695.