

심층 신경망 기반의 생활폐기물 자동 분류

남준영, 이혜민, Asif Ashraf Patankar, Hanxiang Wang, Yanfen Li, 문현준
 세종대학교 컴퓨터공학과
 hmoon@sejong.edu

Object classification for domestic waste based on Convolutional neural networks

Junyoung Nam, Christine Lee, Asif Ashraf Patankar, Hanxiang Wang, Yanfen Li,
 Hyeonjoon Moon
 Department of Computer Science and Engineering, Sejong University

요 약

도시화 과정에서 도시의 생활폐기물 문제가 빠르게 증가되고 있고, 효과적이지 못한 생활폐기물 관리는 도시의 오염을 악화시키고 물리적인 환경오염과 경제적인 부분에서 극심한 문제들을 야기시킬 수 있다. 게다가 부피가 커서 관리하기 힘든 대형 생활폐기물들이 증가하여 도시 발전에도 방해가 된다. 생활폐기물을 처리하는데 있어 대형 생활폐기물 품목에 대해서는 요금을 청구하여 처리한다. 다양한 유형의 대형 생활폐기물을 수동으로 분류하는 것은 시간과 비용이 많이 든다. 그 결과 대형 생활폐기물을 자동으로 분류하는 시스템을 도입하는 것이 중요하다. 본 논문에서는 대형 생활폐기물 분류를 위한 시스템을 제안하며, 이 논문의 4 가지로 분류된다. 1) 높은 정확도와 강 분류(robust classification) 수행에 적합한 Convolution Neural Network(CNN) 모델 중 VGG-19, Inception-V3, ResNet50의 정확도와 속도를 비교한다. 제안된 20 개의 클래스의 대형 생활폐기물의 데이터 셋(data set)에 대해 가장 높은 분류의 정확도는 86.19%이다. 2) 불균형 데이터 문제를 처리하기 Class Weight VGG-19(CW-VGG-19)와 Extreme Gradient Boosting VGG-19 두 가지 방법을 사용하였다. 3) 20 개의 클래스를 포함하는 데이터 셋을 수동으로 수집 및 검증하였으며 각 클래스의 컬러 이미지 수는 500 개 이상이다. 4) 딥 러닝(Deep Learning) 기반 모바일 애플리케이션을 개발하였다.

1. 서론

UN(United Nations) 통계에 따르면 전 세계 국가의 약 60%가 2015 지구 정상 회의에서 생활폐기물 및 기타 환경 문제 관리에 대해 높은 우려를 강조했다[1]. 이러한 문제에서 합리적인 생활폐기물 처리와 관리를 통해 생활폐기물을 줄이는 것은 세계적으로 중요한 문제이다. 끊임없이 증가하는 생활폐기물을 적절하게 취급하지 않으면 환경을 오염 시킬 뿐만 아니라 사람들의 삶과 질이 떨어지게 된다. 18 세기, 사람들은 고철, 종이 및 목재와 같은 재사용 가능한 재료를 재활용하기 시작했다. 1900 년대 말, 환경 인식이 개선되면서 재활용은 전세계 모든 국가에서 생활폐기물을 처리하는 부분은 중요한 부분이 되었다. 한국 정부는 도시 생활폐기물 문제에 대해서 생활폐기물 분류 정책을 추진하고 생활폐기물을 처리하는데 있어 크기에 따라 돈을 지불하는 시스템을 제안했다. 생활폐기물의 분류와 대규모 생활폐기물의 지불은 생활폐기물 발생을 효과적으로 줄였다. 하지만 전통적인 방법은 수동 분류 방식으로, 시간과 비용이 많이 발생하고 생활폐기물이 분류되는 범주가 너무 작다는 단점을 갖고 있다. 따라서 재활용 생활폐기물의 자동분류는 경제적 및 환경적 측면에서 가장 중요한 부분이다. 효율적이고 정확한 분류 방법을 제시하기 위해 많은 연구자들은 기계학습과 딥 러닝 방법을 사용하여 생활폐기물을 분류하고 계산 시간과 인식 정확도 모두에서 우수한 실험 결과를 도출해냈다.

본 논문에서는 대형 생활폐기물에 대해 생활폐기물 분류의 부족을 해결하고 높은 분류를 하기 위해 딥러닝 기반으로 가장 일반적인 20 가지 유형의 생활폐기물 분류 시스템을 개발했다. 또한 이 논문에 사용된 구조는 클라이언트-서버 구조를 갖고 있다. 클라이언트 측은 서버에 직접 연결되어 중간 링크를 제거하고 데이터 전송이 다른 방법 보다 빠르다. 또한 불균형 데이터 세트는 실 세계에 존재하므로 분류 작업에 큰 어려움이 있다. 수집된 데이터 셋에서 불균형 데이터 문제도 발생한다. 이 문제를 해결하기 위해 많은 연구자들은 resampling 기술 [2]과 cost-sensitive learning [3]과 같은 효과적인 방법들을 제안하였다. 본 논문은 1 장 서론에 이어 2 장 관련연구, 3 장 본문, 4 장 실험, 5 장 결론 순으로 기술한다.

2. 관련연구

환경 오염의 심각성으로 생활폐기물 분류는 인기 있는 연구형태가 되었고 많은 연구자들은 생활폐기물을 식별하기 위해 딥 러닝 아키텍처를 사용하여 생활폐기물을 분류했다. [4]는 자동 생활폐기물 분류를 위해 딥러닝과 Support Vector Machine(SVM)을 비교했다. 세 가지 주요 범주를 기반으로 SVM 은 94.80%의 높은 분류 정확도를 달성 한 반면 CNN 은 83%를 달성했다. SVM 알고리즘은 이 실험에서 더 나은 분류 결과를 보여주지만 큰 표본 크기 데이터 셋에는 적합하지

않으며 훈련 샘플 데이터 수가 증가함에 따라 훈련 속도가 급격히 떨어진다. [5]는 컴퓨터 비전과 딥 러닝을 사용하여 생활폐기물 수거 통을 인식한다. 특징을 검출하기 위해 Vector of Locally Aggregated Descriptors(VLAD), You Only Look Once(YOLO)를 사용했으며 90% 정확도를 보였다. [6]은 생활폐기물을 유리, 종이, 금속, 플라스틱, 판지 및 일반 쓰레기를 포함하여 6 가지로 분류한다. Scale-invariant Feature Transform(SIFT) 기능과 CNN 을 갖춘 SVM 이다. 실험 결과 SVM 은 63%, CNN 이 22%의 정확도를 보여주었다. [7]은 스마트 폰을 사용하여 8 가지 유형 또는 도로 손상을 분류했고 실험 결과 MobileNet 모델과 Inception V2 모델을 사용하여 시간이 1.5 초인 75% 및 77%의 정확성을 나타냈다. 불균형 데이터 문제를 해결하기 위해 여러 방법이 제시되었다.

본 논문은 이전 연구를 기반으로 딥 러닝 모델을 사용하여 생활폐기물을 분류하는 시스템을 설계했다. 그 후에 20 개의 클래스로 분류하고 이를 애플리케이션에 적용한다. 애플리케이션은 서버·클라이언트 모델이고 각 클라이언트 측은 서버에 요청을 보내고 빠른 속도로 응답을 얻을 수 있다.

3. 본론

3.1 데이터셋 수집 및 데이터 분류

데이터 셋에서 모든 이미지는 224x224 이미지로 저장되며 시스템에는 이미지 크기 조정을 위한 python 스크립트도 포함되어 있다. 본 논문은 Yahoo, Bing, Google, Baidu 및 Naver 와 같은 다양한 웹 사이트에서 20 가지 종류의 생활 폐기물을 수집한다. 오래된 웹 사이트나 주요 웹 사이트에서 관련 데이터를 찾기 어려운 카테고리를 포함한다. 반대로 데이터 셋에는 식기 세척기, 에어컨 등과 같은 일반적인 클래스가 있다. 따라서 각 범주의 데이터 양은 매우 다르며 데이터의 불균형 문제로 이어진다. 최소 이미지의 수는 각 클래스에 대해 500 개 이미지이고 최대 3,470 이며 데이터 집합의 양은 21,047 이다. 모든 클래스에는 각 생활폐기물 유형의 해당 이름이 표시되어 있다. 데이터를 훈련 세트와 테스트 세트로 나누고 트레이닝 세트는 전체 데이터 세트의 90%를 구성하며 심층 신경망에 해당하는 가중치를 얻는데 사용한다. 테스트 셋은 이미지의 10%로 구성되며 모델에 사용된다. [표 1]은 각 클래스의 데이터베이스를 구성하고 있다. Name 은 클래스 분류의 이름을 나타내고 No.는 클래스에 대한 이미지의 개수를 나타낸다.

표 1. 클래스의 데이터베이스

Index	Name	No.
1	Armoire	855
2	Mungab	506
3	Game machine	500
4	Drawer	3225
5	Bad frame	500
6	Wheelchair	2197
7	Window blind	1938

8	Bicycle	625
9	Mirror	3470
10	Gas stove	726
11	Ice box	500
12	Pad	500
13	Bed	713
14	Mattress	565
15	Computer	987
16	Water tank	1182
17	Bookshelf	531
18	Dining table	518
19	Blender	509
20	Printer	500

대형 생활폐기물 분류는 딥 러닝을 기반으로 생활폐기물 분류하고 분류 범주의 부족한 문제를 해결한다. 제안된 방식의 애플리케이션 시스템은 [그림 1]과 같다. 안드로이드 시스템 기반이며 인터페이스는 세 가지로 구성하고 있다. 첫 번째 단계는 사용자가 찍은 생활폐기물 이미지나 이미 갖고 있는 이미지를 선택할 수 있다. 두 번째 단계는 상위 4(Top-4)의 클래스를 이용하여 분류하고 분류가 완료되면 분류 클래스와 생활폐기물 가격이 표시된다. 3 번째 단계는 2 번째 단계에서 분류된 항목에 해당 품목이 없을 경우에 서브 클래스를 두어 한번 더 분석을 하게 되고 추가적인 분석이 이뤄지게 된다.

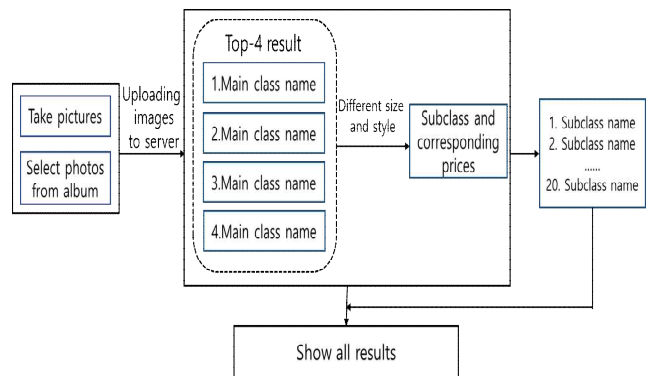


그림 1. 모바일 애플리케이션의 구조

3.2 Convolution Neural Networks model

Convolution Neural Networks (CNN)는 대부분의 분류 작업에서 기계학습 방법인 K-NEAREST Neighbor(KNN), Support Vector Machine (SVM) 보다 더 정확한 분류 결과와 빠른 계산 속도를 갖고 있다. CNN 의 핵심적인 아이디어는 로컬 수용필드 및 가중치 공유로 네트워크 매개 변수를 단순화하고 네트워크가 어느 정보의 변위, 스케일링 및 비선형 변형 안정성을 갖고 있다. 풀링 레이어(pooling layer)는 일반적으로 연속적인 컨벌루션 레이어(convolutional layer)

사이에 배치되어 점차적으로 데이터의 공간을 줄이고 계산 리소스 소비를 줄일 수 있다. 완전 연결된 (fully connected) 레이어는 학습된 모든 기능을 통합하고 네트워크 전체에서 분류기 역할을 한다.

이 연구에서는 VGG-19, ResNet50 및 Inception-V3 를 포함한 3 가지 CNN 모델을 사용하여 수집된 데이터셋을 학습한다. 이미지넷 (ImageNet) 데이터 셋 기반으로 널리 사용되는 훈련된 모델을 사용하였다. 데이터 셋의 크기와 모양이 이미지넷 데이터 셋과 약간 다르고 처음 몇 개의 컨볼루션 레이어가 좀 더 일반적인 기능을 학습하기 위해 사용되었으므로 사전 훈련 모델의 모든 컨볼루션 레이어를 초기화 하고 첫 번째의 파라미터를 고정한다. 배치 정규화 (Batch Normalization)는 신경망을 최적화하는 방법이다. 흩어진 데이터를 통합하고 그라디언트 소실 (gradient disappearance) 또는 그라디언트 폭발 (gradient explosion) 문제를 피하고 네트워크의 견고성을 높이는 것이다. 이 기술은 학습 성능을 크게 향상시키는 것으로 입증되었다. 따라서 배치 정규화 계층은 두 개의 전체 연결 계층 사이에 추가되었다. [그림 2]는 전이 학습 후 VGG-19 모델의 구조이며, 붉은색 점선 상자 구조는 고정 레이어 구조이고 보라색 점선 상자 내의 레이어는 학습이 가능하다. 배치 정규화 층은 2 와 3 의 풀 커넥션 사이 (pull connection)에 추가된다.

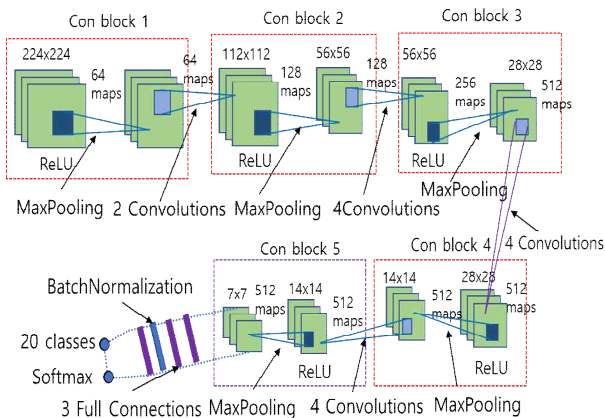


그림 2. 전이 학습 후 VGG-19 의 모델 구조

3.3 불균형 데이터 문제

대부분의 이미지가 다중 분류 작업에서 매우 많은 유형의 데이터가 있다. 데이터 내 각각의 클래스들이 갖고 있는 데이터를 균일하게 수집하기 어렵기 때문에 데이터들의 수들은 균일하지 않고 한쪽으로 치우치게 된다.

실제 훈련과정에서 훈련 셋의 클래스가 다른 클래스 보다 데이터가 적으면 분류 결과는 편향되어진다. 우리는 이러한 문제를 해결하기 위해 [그림 3]과 같이 두 가지 방법을 사용한다.

두 가지 방법 중 하나는 트레이닝 모델에서 다른 카테고리들의 수에 대해 훈련 가중치 값들을 다르게 설정하는 것이다. 다른 하나는 Extreme Gradient Boosting (XGBoosting) 알고리즘을 사용하여 불균형 데이터 문제를 해결한다. 몇몇 분류 작업에서 분류 정확도가 모델 평가의 유일한 기준이라고 말할 수는 없다. 예를 들어 카테고리의

균형이 맞지 않으면 정확도가 정확하지 않고 의미가 없다. 이 연구에서는 혼동 행렬 (confusion matrix), ROC 곡선 (Receiver Operating Characteristic Curve), Area Under Curve (AUC) 값을 사용하여 분류기의 성능을 측정한다.

훈련 과정에서 특정 숫자에 따라 카테고리별로 다른 가중치를 설정하여 분류의 정확도를 향상시킬 수 있다. 가중치는 총 데이터 셋에서 특정 클래스의 발생 빈도를 계산하여 얻을 수 있다. 특정 클래스의 가중치가 크면 다른 클래스의 경우 가중치는 손실된다. 이에 따라 많은 클래스들에 대해서 시스템 학습을 줄이기 위해 가중치를 낮게 설정하였다.

XGBoost 알고리즘의 기본 원리는 학습 성능에 따라 샘플 분포를 조정하여 성능이 좋지 않은 샘플이 더 많은 집중을 할 수 있게 만든 다음 기본 학습의 수가 지정된 수에 도달할 때까지 조정된 샘플로 다음 기본 학습자를 반복적으로 학습한다. 그리고 병렬 컴퓨팅 및 희소 훈련 데이터 처리를 지원하므로 다른 Gradient boosting 방식보다 효율적이다.

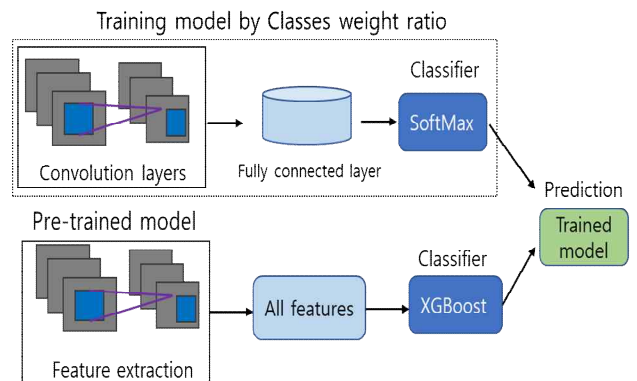


그림 3. 불균형 데이터 문제를 해결하기 위한 방법

4. 실험결과

실험을 하기 위한 시스템 환경은 서버와 클라이언트로 나뉘고 서버의 사양은 GPU: NVIDIA PACAL TITIAN Xp GDDR5 12GB PCI-express, CPU: Intel Core i7-5280k (6Core/3GHz/15MB), RAM: DDR4 16GB PC4-17000, Library: Tensorflow/Keras Library, OS: Ubuntu 14.04 이다. 클라이언트는 디바이스: HUAWEI nova lite 2, CUP Version: Android 8.0.0 이다. CNN 모델에서 각 매개 변수에 대한 최적의 값을 얻었으며, 분류 정확도와 실행속도를 비교하여 최상의 모델을 얻었다. 그리고 최종 분류 정확도를 향상시키기 위해 선택한 모델의 구조가 수정된다. 그 후, 두 가지 다른 방법을 사용하여 불균형 데이터 문제를 해결하였다.

평가 과정에서 매개 변수에 대한 최적의 값을 구하고 더 정확한 값을 위해 추가적으로 다른 매개 변수들을 이용하였다. CNN 모델 중 VGG-19, Inception-V3, ResNet50 모델들을 비교하였고 비교를 통해 실행속도, 정확도, 반복횟수를 [표 2]에서 확인할 수 있다.

정확도를 살펴보면 VGG-19 는 86.19%, Inception-V3 는 81.15%, ResNet50 는 79.63%이다. 실행 속도가 빠른 순으로 ResNet50, VGG-19, Inception-V3 이다. 결과적으로 정확도, 계산 복잡도, 실행 속도를 고려했을 경우 가장 적합한

모델은 VGG-19 인 것을 확인할 수 있다.

표 2. 다른 종류의 CNN 모델 비교

Model	Learning rate	Iterations	Accuracy (%)
VGG-19	0.001	1,764	86.19
Inception-V3	0.0001	1,000	81.15
ResNet50	0.0001	441	79.63

불균형 데이터에 대한 실험은 두 가지로 분류된다. 가중치 값을 변경하여 설정하고 다른 하나는 XGBoost 알고리즘을 적용하는 것이다. CNN 모델은 데이터 셋에 적합한 VGG-19 모델을 사용하였다. 불균형 데이터를 이루고 있는 Mungab 과 Drawer 클래스를 선택하여 사용하였다. 두 클래스의 이미지에 대해서 비율을 변경하여 실험하였고, 비율은 1:1, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:100 와 같이 나눈다. 데이터의 확률은 가중치를 계산하는데 사용하고 Keras 의 클래스들 가중치 매개변수를 사용하여 훈련 가중치를 VGG-19 모델에 사용한다. 그런 다음 XGBoost 분류기에 학습하면 [그림 4]와 같은 결과를 확인할 수 있다.

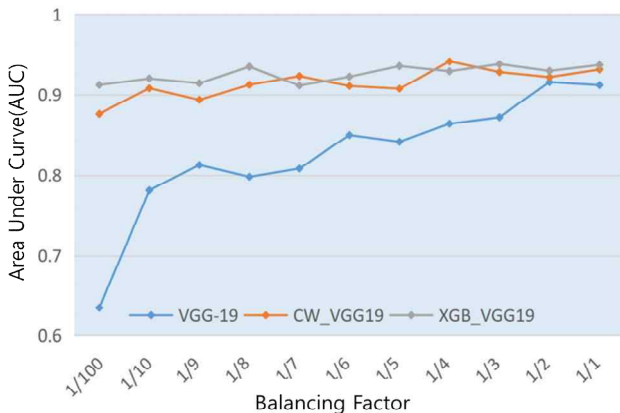


그림 4. 불균형 데이터 셋의 AUC 결과

결과를 보면 비율이 1:100 인 경우 VGG-19 는 0.63 이고 XGB_VGG-19 는 Area Under Curve(AUC)가 0.91, CW_VGG19 는 0.88 이다. 상대적으로 XGB_VGG19 모델은 CW_VGG19 모델보다 안정적인 것을 확인할 수 있다.

5. 결론

본 연구는 배경이 복잡한 이미지에서 생활폐기물을 인식하기 위해 딥러닝 방법을 사용했으며, 이 방법을 기반으로 실 생활에서 사용할 수 있는 애플리케이션을 개발하였다. 총 21,047 개의 이미지를 웹 사이트에서 다운로드 한 후 검증하였다. CNN 모델 세 가지 중 VGG-19 는 두 번째로 빠른 속도와 20 개의 클래스에 대해서 86.1%로 가장 높은 정확도를 갖고 있어 가장 적합한 모델이다.

이 연구에서 중요한 두 가지 문제가 있다. 그 중 하나는

클래스 간의 유사성이 높고 다른 하나는 불균형 데이터 문제이다. 클래스의 유사성 문제를 해결하기 위해 상위 4 개 클래스로 분류하고 정확도는 평균 86.19%에서 97.01% 사이의 높은 정확도를 얻었다. 불균형 데이터 문제를 해결하기 전과 해결한 후의 결과를 비교했다. 실험을 통해 제안된 방법으로 문제를 해결했을 때 좋은 성능을 확인할 수 있었다. 또한 안드로이드 시스템에서 VGG-19 모델을 기반으로 실 생활에서 사용할 수 있는 생활폐기물 분류 애플리케이션을 개발하였다.

이 연구에서 사용된 시스템은 다른 심층 CNN 모델보다 더 나은 정확도와 빠른 실행 속도를 달성할 수 있었지만 이 연구에는 몇 가지 제한 사항이 있다. 컴퓨터 메모리에 대한 수요가 높고 수집된 데이터 셋이 크기 때문에 배경이 복잡한 이미지에 대해서는 이미지 전처리 작업을 수행하지 않을 경우 인식 결과에 큰 영향을 줄 수 있다. 따라서 이미지 전처리 작업에 대한 연구가 추가적으로 이뤄져야 할 것이다.

Acknowledgement

이 논문은 2019 년도 과학기술 정보통신부의 재원으로 정보통신 기획평가원의 지원을 받아 수행된 연구임(2019-0-00136, 스마트시티 산업 생산성 혁신을 위한 AI 융합 기술 개발)

6. 참고문헌

- [1] Noorani, S. and M. Fernandes (2017). Evaluation of Convolutional neural networks for Waste Identification. 2017 International Conference on Computing Methodologies and Communication (ICCMC), IEEE.
- [2] L., R. Zhou, L. Tang and R. Chen (2018). "A DBN-based resampling SVM ensemble learning paradigm for credit classification with imbalanced data." Applied Soft Computing 69: 192-202.
- [3] Fernández, A., S. García, M. Galar, R. C. Prati, B. Krawczyk and F. Herrera (2018). Cost-sensitive learning. Learning from Imbalanced Data Sets, Springer: 63-78.
- [4] Sakr, G. E., M. Mokbel, A. Darwich, M. N. Khneisser and A. Hadi (2016). Comparing deep learning and support vector machines for autonomous waste sorting. 2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET), IEEE.
- [5] Valente, M., H. Silva, J. M. Caldeira, V. N. Soares and P. D. Gaspar (2019). "Detection of Waste Containers Using Computer Vision." Applied System Innovation 2(1): 11.
- [6] Yang, M. and G. Thung (2016). "Classification of trash for recyclability status." CS229 Project Report 2016.
- [7] Maeda, H., Y. Sekimoto, T. Seto, T. Kashiyama and H. Omata (2018). "Road damage detection and classification using deep neural networks with smartphone images." Computer-Aided Civil and Infrastructure Engineering 33(12): 1127-1141.