

<https://doi.org/10.7236/JIIBC.2023.23.3.199>

JIIBC 2023-3-27

딥러닝기반 건축폐기물 이미지 분류 시스템 비교

A Comparison of Image Classification System for Building Waste Data based on Deep Learning

성재경*, 양민철**, 문경남***, 김용국****

Jae-Kyung Sung*, Mincheol Yang**, Kyungnam Moon***, Yong-Guk Kim****

요약 본 연구는 건축시 발생하는 폐기물의 자동분류를 위해 딥러닝 알고리즘을 활용해 건축 폐기물 데이터를 각각 목재 폐기물, 플라스틱 폐기물, 콘크리트 폐기물로 분류하는 두 모델들을 통해서 성능 비교를 한다. 건축 폐기물의 분류를 위해 사용된 딥러닝 알고리즘은 합성곱 신경망 이미지 분류 알고리즘 VGG-16과 NLP를 기반으로 이미지를 시퀀스화 시킨ViT, Vision Transformer 모델을 사용했다. 건축 폐기물 데이터 수집을 위해 이미지 데이터를 전 세계 검색엔진에서 크롤링 하였고, 육안으로도 명확히 구분하기 어렵거나, 중복되는 등 실험에 방해되는 이미지는 전부 제외하여 각 분류당 1천장씩 총 3천장의 이미지를 확보했다. 또한, 데이터 학습시에 모델의 정확도 향상에 도움을 주기 위해 데이터 확대 작업을 진행해 총 3만장의 이미지로 실험을 진행 하였다. 수집된 이미 데이터가 정형화 되어있지 않은 데이터임에도 불구하고 실험 결과는 정확도가 VGG-16는 91.5%, ViT 는 92.7%의 결과가 나타났다. 이는 실제 건축폐기물 데이터 관리 작업에 실전 활용 가능성을 제시한 것으로 보인다. 본 연구를 바탕으로 추후에 객체 탐지 기법이나 의미론적 분할 기법까지 활용한다면, 하나의 이미지 안에서도 여러 세밀한 분류가 가능해 더욱 완벽한 분류가 가능할 것이다.

Abstract This study utilizes deep learning algorithms to automatically classify construction waste into three categories: wood waste, plastic waste, and concrete waste. Two models, VGG-16 and ViT (Vision Transformer), which are convolutional neural network image classification algorithms and NLP-based models that sequence images, respectively, were compared for their performance in classifying construction waste. Image data for construction waste was collected by crawling images from search engines worldwide, and 3,000 images, with 1,000 images for each category, were obtained by excluding images that were difficult to distinguish with the naked eye or that were duplicated and would interfere with the experiment. In addition, to improve the accuracy of the models, data augmentation was performed during training with a total of 30,000 images. Despite the unstructured nature of the collected image data, the experimental results showed that VGG-16 achieved an accuracy of 91.5%, and ViT achieved an accuracy of 92.7%. This seems to suggest the possibility of practical application in actual construction waste data management work. If object detection techniques or semantic segmentation techniques are utilized based on this study, more precise classification will be possible even within a single image, resulting in more accurate waste classification

Key Words : Automatic Classifier, Convolutional Neural Network(CNN), Data Augmentation, Deep Learning, VGG-16, Vision Transformer(ViT)

*정회원, 에이아이씨랩 연구이사

**준회원, 에이아이씨랩 대표

***준회원, 메타크레용 대표

****정회원, 세종대학교 컴퓨터공학과 (교신저자)

접수일자 2023년 5월 4일, 수정완료 2023년 5월 30일

게재확정일자 2023년 6월 9일

Received: 4 May, 2023 / Revised: 30 May, 2023 /

Accepted: 9 June, 2023

*Corresponding Author: ykim@sejong.ac.kr

Dept. of Computer Eng., Sejong University, Seoul, Korea

I. 서 론

건축시 필연적으로 발생하는 건축 폐기물은 단순히 폐기 되는 것 뿐만 아니라 퇴비, 합판, 아스팔트 등으로 재사용이 가능하다. 하지만 건축폐기물은 각각 목재폐기물, 콘크리트폐기물 등 자재마다 폐기, 재활용 방식이 다르고, 이에 따라 폐기를 담당하는 업체가 다르다. 이를 취급하는 업체가 다양함에 따라 효율적으로 건축폐기물을 재활용하기엔 여러 어려움이 있다. 현재 폐기물을 처리하는 방식인 단순한 업무 담당자가 일일이 폐기물을 파악하고 분류해 폐기물 업체를 부르는 것은 시간적으로 효율적이지 않다. 보다 빠르고 정확한 건축 폐기물 폐기 및 재활용을 하기 위해 우선적으로 해결해야 할 문제는 건축폐기물 분류의 자동화이다. 본 연구는 건축시 발생하는 폐기물을 쉽게 취득할 수 있는 이미지 데이터로 제공받아 이를 이미지처리 딥러닝 알고리즘을 활용해 각각 목재 폐기물, 플라스틱 폐기물, 콘크리트 폐기물로 분류하는 두 모델을 제시를 통해서 성능 비교를 한다. 건축 폐기물의 분류를 위해 사용된 딥러닝 알고리즘은 합성곱 신경망 이미지 분류 알고리즘 VGG-16^[1]를 기반으로 사용했다. 또 다른 모델은 자연어 처리 분야 Attention 메커니즘을 기반으로 하는 트랜스포머(Transformer)를 사용하였다. 건축 폐기물 데이터 수집을 위해 이미지 데이터를 전 세계 검색엔진에서 크롤링^[2] 하였고, 육안으로도 명확히 구분하기 어렵거나, 중복되는 등 실험에 방해되는 이미지는 전부 제외하여 각 분류당 1천장씩 총 3천장의 이미지를 확보하였다. 또한, 데이터 학습시에 모델의 정확도 향상에 도움을 주기 위해 데이터 증강^[3] 작업을 진행해 총 3만장의 이미지로 실험을 진행 하였다. 수집된 이미 데이터가 정형화 되어 있지 않은 데이터 임에도 불구하고 실험 결과는 VGG-16과 트랜스포머 모델에서 각각 91.5%와 99.7%의 정확도를 보였다. 이는 실제 건축폐기물 데이터 관리 작업에 용이 할 것으로 보인다. 본 연구를 바탕으로 추후에 객체 탐지 기법^[4]이나 의미론적 분할 기법^[5]까지 활용시킨다면, 하나의 이미지 안에서도 여러 세밀한 분류가 가능해 더욱 정확한 분류가 가능할 것이다. 이러한 접근 방식은 건축 폐기물 처리 과정의 효율성을 높이고, 재활용 가능한 자원의 활용도를 높여 환경에 긍정적인 영향을 미칠 것으로 기대된다.

II. 관련 연구

인공신경망^[6]은 인간의 두뇌에 있는 신경세포, 뉴런이

연결된 형태를 본 따 만든 모델이다. 현대 컴퓨터 공학기술은 인공지능을 효과적으로 학습하기 위해, 즉 머신러닝^[7] 기술을 적용하고 이를 최적화시키기 위해 발전된 형태의 인공신경망의 구조인 딥러닝^[8]을 차용한다. 본 장에서는 이러한 인공신경망의 개념 및 정의와 여러 논문에서 발췌한 현재까지 사용되는 인공신경망 구조 중 딥러닝의 대표적인 구조인 합성곱 신경망(CNN)과 트랜스포머를 소개한다. 또한 이를 활용한 이미지 데이터 분류 기술, 이미지 처리에 관련된 딥러닝 기술 대해서 살펴본다.

1. 인공신경망(Artificial Neural Network, ANN)

인간의 뇌는 1000억 개가 넘는 신경세포가 100조 개 이상의 시냅스를 통해 병렬적으로 연결되어 있다고 한다. 각각의 뉴런은 수상돌기를 통해 다른 뉴런에서 입력신호를 받아서 축색돌기를 통해 다른 뉴런으로 신호를 내보낸다. 시냅스는 뉴런과 뉴런을 연결하는 역할을 한다. 출력신호는 입력된 신호가 모여서 일정한 용량을 넘어설 때 일어난다.

인공신경망 뉴런 모델은 생물학적인 뉴런을 수학적으로 모델링한 것이다. 즉, 생물학적인 뉴런이 위의 그림과 같이 다른 여러개의 뉴런으로부터 입력값을 받아서 세포체에 저장하다가 자신의 용량을 넘어서면 외부로 출력값을 내보내는 것처럼, 인공신경망 뉴런은 여러 입력값을 받아서 일정 수준이 넘어서면 활성화되어 출력값을 내보낸다.

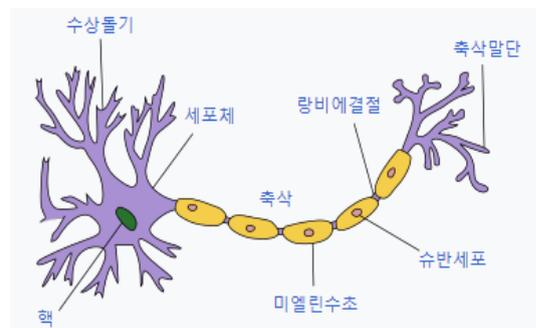


그림 1^[9]. 뉴런의 예시

Fig. 1. Example of a neuron

아래 그림과 같이 신경망은 입력값이 들어가는 입력층, 그리고 입력값들을 잘 섞어서 변환하는 중간층(은닉층) 그리고 출력이 나오는 출력층이 있다. 각각 층간을 연결하는 것들은 가중치(Weight)로 구성되어 있다. 또한 신경망의 특징으로는 조금 더 복잡한 데이터도 학습할

수 있도록 입력값을 잘 섞는 중간층의 노드가 하나가 아니라 여러 개가 될 수 있다는 점이다. 중간층의 노드를 여러 개 만들어 입력값의 특성을 다양한 방법으로 학습할 수 있도록 한다. 또한 중간층은 신경망 알고리즘 내부에 존재해 사람이 가중치가 어떻게 되어가고 있는지 명확히 알기 어려운 것이 특징이다. 신경망과 회귀 분석 모델^[11]의 가장 큰 차이는 회귀모델은 중간층 없이 입력값과 출력값의 관계가 직접적으로 연결되어 있다는 것이다.

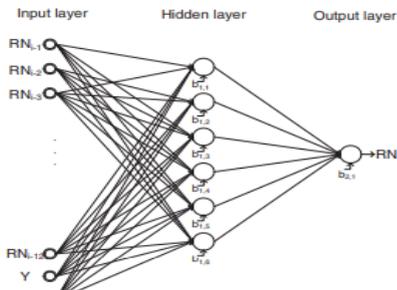


그림 2^[10]. 인공신경망의 노드 예시
 Fig. 2. An example of a node in an artificial neural network

인공신경망의 입력층에서 입력들을 중간층으로 전달하며 임의의 가중치와 편향(Bias)^[12] 해 출력층에서 출력값을 결정하는 과정을 함수화 한 것을 활성화 함수(Activation Function)^[13] 고 한다. 대표적인 활성화 함수로는 시그모이드 함수(Sigmoid Function)^[14]가 있다.

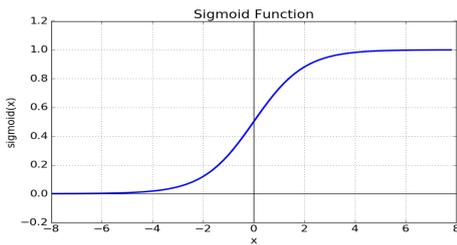


그림 3^[15]. 시그모이드 함수의 예시
 Fig. 3. Example of sigmoid function

시그모이드 함수를 사용한 인공신경망의 학습과정은 다음과 같다. 인공신경망은 입력에 대하여 순전파(Forward Propagation) 연산을 하고, 그리고 순전파 연산을 통해 나온 예측값과 실제값의 오차를 손실 함수를 통해 계산하고, 이 손실을 미분을 통해서 기울기(Gradient)를 구하고 이를 통해 역전파(Back Propagation)를 수행한다.

그리고 시그모이드 함수의 문제점은 미분을 해서 기울기를 구할 때 발생한다. 위 그림3 에서와 같이 시그모이드 함수의 출력값이 0 또는 1에 가까워지면 기울기가 완만해지게 된다. 역전파 과정에서 0에 가까운 아주 작은 기울기가 곱해지게 되면, 앞단에는 기울기가 잘 전달되지 않게 된다. 이러한 현상을 기울기 소실 문제(Gradient vanishing problem)^[16] 라고 한다.

위와 같은 문제는 초기 인공신경망에 어려움을 겪게 만든 문제였지만, 현대 컴퓨터 공학에서는 이러한 문제를 최대한 해결하기 위해 ReLU(Rectified Linear Unit) 함수^[17]와 같은 발전된 형태의 활성화 함수를 사용한다.

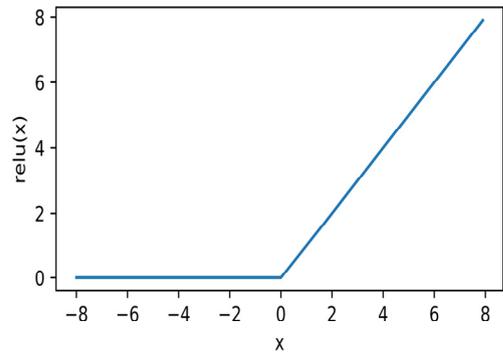


그림 4^[18]. ReLU 함수의 예시
 Fig. 4. Example of ReLU function

2. 합성곱 신경망(convolutional neural network)

합성곱 신경망(convolutional neural network)^[19]은 이미지를 더욱 효과적으로 처리하기 위해 필터링 기법을 인공신경망에 적용함으로써 제시되었다. CNN은 합성곱(Convolution) 연산을 사용하는 ANN의 한 종류이다. 합성곱 연산이란 두함수 f, g 가운데 하나의 함수를 반전, 전이 시킨 다음 다른 하나의 함수와 곱한 결과를 적분하는 것을 의미한다.

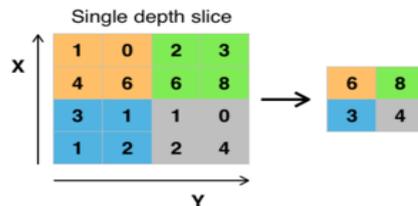


그림 5^[19]. convolution max pooling의 예시
 Fig. 5. Example of convolution max pooling

CNN은 앞 이미지와 같이 이미지의 특징을 추출하는 부분과 클래스를 분류하는 부분으로 나눌 수 있다. 특징 추출영역은 필수요소인 Convolution Layer와 선택요소인 Polling Layer를 여러겹 쌓는 형태로 구성된다. 그리고 마지막 부분에 이미지 분류를 위한 Fully Connected Layer가 추가된다. 이미지의 특징을 추출하는 부분과 이미지를 분류하는 부분 사이에 이미지 형태의 데이터를 배열 형태로 만드는 Flatten Layer가 위치한다.

3. VGGnet

VGGnet은 Karen Simonyan과 Andrew Zisserman이 만든 합성곱 신경망 모델로 네트워크의 깊이가 모델이 좋은 성능을 보이는 데 중요한 역할을 한다는 것을 보여준 모델이다. VGGnet은 모든 필터 커널의 크기를 3 x 3으로 구성해 네트워크의 깊이를 깊게 만들었다. 네트워크의 깊이가 깊어질수록 더 좋은 성능을 내지만, 파라미터의 갯수가 비약적으로 늘어난다는 단점이 있어, 연산하는데 소요되는 시간이 많다. 본 모델은 A부터 E까지 총 6가지의 서로 다른 구조로 실험을 하였고 이중 VGG-16, VGG-19(각각 D, E이다)와 같이 더 많은 층을 사용한 모델의 성능이 더 좋다. 본 연구에서는 VGG-16을 활용해 건축폐기물 데이터를 분류한다.

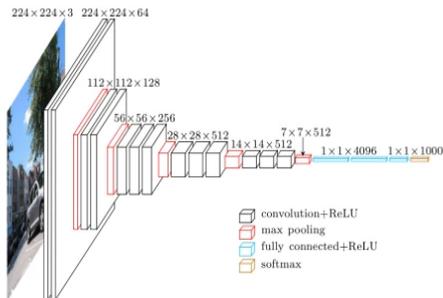


그림 6^[22]. VGG-16 모델의 구조
Fig. 6. The VGG-16 model architecture

4. 트랜스포머(transformer) 모델^[22]

트랜스포머는 자연어 처리 분야에서 일반적으로 사용되는 순환 신경망(RNN)^[23] 보다 높은 성능을 보이는 딥러닝 모델^[24]로 어텐션(attention) 메커니즘을 사용한다. 어텐션 메커니즘은 자연어 처리 분야에서 입력 시퀀스 내에서 각 단어들 간의 상호작용을 파악하여 중요도를 계산한다. 그리고 단어의 중요도에 따라 가중치를 부여하는 기술로 현재 단어인 쿼리(Query)와 다른 단어인

키(Key), 밸류(Value)의 개념으로 이루어져 있다. 트랜스포머 모델의 대표적인 예로 Multi-Head Attention이 있다. Multi-Head Attention은 여러 개의 셀프 어텐션으로 병렬로 수행후 이를 합치는 방식의 동작으로 입력 시퀀스의 여러 위치에서 필요 정보를 동시에 수집한다. 그리고 트랜스포머 모델에서는 인코더와 디코더에서 모두 사용한다. 인코더에서는 입력문장을 embedding(임베딩) 과정을 통해서 고차원의 벡터로 변환 후에 각 단어들 간의 관계를 파악하고 의미를 추출하는 데 사용한다. 디코더에서는 이전에 생성된 단어들을 여러 개의 디코더층에 입력하여 입력 문장과 이전에 생성된 단어들을 이용하여 새로운 단어를 생성하는 데 사용된다. 이를 통해 트랜스포머 모델은 자연스러운 문장 생성을 위해 입력 시퀀스의 각 위치에서 필요한 정보를 집중적으로 수집할 수 있다.

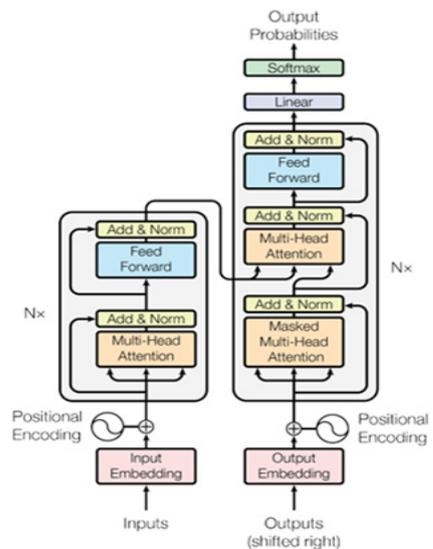


그림 7^[22]. Transformer 모델의 구조
Fig. 7. The Transformer model architecture

III. 3장 건축폐기물 분류 시스템

본 장에서는 연구에 사용된 건축 폐기물 데이터 수집 방법과 이미지 데이터 분류 시스템의 설계 및 구현에 대해 설명한다.

1. 이미지 데이터 분류 시스템 구조 설계 및 구현
본 연구에서는 구글과 같은 전세계 검색엔진에서 수집

한 이미지 데이터와 CNN과 ViT의 딥러닝 알고리즘을 활용해 분류 시스템을 구현하였다. 본 시스템은 3천장의 이미지데이터를 Augmentation하여 3만장으로 증폭시킨 뒤 학습데이터와 테스트데이터로 분리하여 사용하였다. 구현 기술은 VGG-16 모델은 Tensorflow API를 활용해 구현했다. Tensorflow는 딥러닝을 하기위한 엔드 투엔드 서비스를 제공하는 오픈소스 API로서 파이썬 환경에서 데이터 전처리부터 모델학습에 필요한 모델의 기본적인 Architecture를 로드하는데 사용된다.

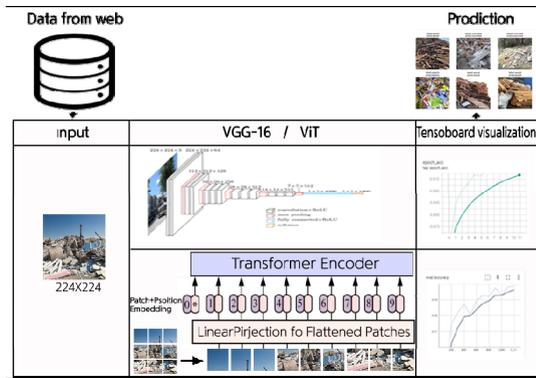


그림 8. VGG-16 및 ViT기반 건축폐기물 분류 시스템의 구조도
 Fig. 8. Structure of the construction waste classification system based on VGG-16 and ViT

수집된 데이터를 VGG-16 모델의 Input으로 사용하기 위해 224 x 224 사이즈로 변환하여 학습시킨 뒤, Tensorboard로 결과를 시각화 하였다.

ViT를 사용한 구현 설명 본 연구에서 사용된 ViT 모델은 이미지를 시퀀스화하여 입력으로 사용한다. 이미지를 고정 크기의 패치로 분할하고, 각 패치를 1차원 벡터로 변환한다. 그런 다음 각 벡터에 위치 임베딩을 추가하고, 이러한 임베딩된 벡터를 트랜스포머 인코더에 입력한다. 트랜스포머 인코더는 전역적인 정보를 포착하고, 분류 토큰을 사용하여 최종 예측을 생성한다. 또한 트랜스포머 모델은 이미지는 동일한 조건으로 구현기술은 PyTorch API를 사용하였다. PyTorch는 Torch 라이브러리를 기반으로 한 오픈 소스 머신러닝 라이브러리로 그래픽 처리 장치(GPU)를 통해 강력한 가속을 지원하는 텐서 연산(NumPy와 유사)과 타입 기반 자동 미분 시스템을 기반으로 하는 딥 뉴럴 네트워크를 제공한다. PyTorch와 함께 여기에서 사용된 이미지 분류 기술은 Vision Transformer (ViT) 모델을 사용하여 특성 추출 및 이미지 분류를 하였다.

2. 실험데이터

실험에 사용한 데이터는 전세계 검색엔진에서 이미지를 크롤링해 수집했다. 각국의 언어로 번역한 목재 폐기물, 콘크리트 폐기물, 플라스틱 폐기물을 키워드로 크롤링한 약 8,000개의 이미지 데이터중 중복되는 데이터, 잘못 수집된 데이터, 육안으로도 구분이 어려운 부정확한 데이터등 실험에 방해되는 데이터를 제외하고 각 클래스당 1,000개씩 총 3,000개의 데이터를 확보했다. 모델 학습을 위해 데이터를 학습 데이터셋(Train dataset)과 테스트 데이터셋(Test dataset)으로 각각 8:2의 비율로 나눠 실험하였다.



그림 9. 실험에 사용된 Data와 argumentation의 예시
 Fig. 9. Example of data and argumentation used in the experiment

IV. 실험 및 분석

본 장에서는 건축 폐기물 분류 시스템과 전세계 검색엔진에서 크롤링해 수집한 이미지 데이터로 학습한 자동 분류 시스템을 실험하였다. 먼저 실험환경과 방법에 대한 설명과 각 분류 시스템의 실험결과를 분석을 한다.

1. 실험환경 및 방법

본 연구의 분류 모델 학습을 위한 실험환경은 AMD 5600x CPU, 16.0GB RAM과 NVIDIA GeForce 3070 GPU를 사용하였다. 실험을 위한 환경은 Python3.9, Tensorflow 2.5 버전을 사용했고, 이를 Anaconda 가상환경에서 구성했다. 실험 데이터는 각각 목재 폐기물, 콘크리트 폐기물, 플라스틱 폐기물로 분류해 클래스를 지정했다. 또한 분류된 총 3000개의 데이터를 Data augmentation 작업을 통해 30000개로 확대하였으며, Tensorflow API의 기능중 하나인

Tensorboard를 통해 실시간으로 모델의 학습 결과를 획득하였다. 이후 결과값이 좋은 모델을 사용해 Validation data에서 각 클래스 별로 200장씩 랜덤하게 추출한 Predict dataset을 추가로 구성해 실제로 학습된 모델을 적용시켜 시각화 하였다. 총 600개의 이미지 데이터중 4행 4열로 16개의 이미지를 랜덤하게 지정해 클래스의 이름을 표기하고, 모델이 예측한 예측값을 pred으로 표기하여 실험하였다.

2 실험 데이터 분석 및 결과

아래 표는 실험을 위해 수집된 건축 폐기물 데이터를 Data augmentation을 통해 수량을 늘린 총 데이터의 분류별 개수이다. 목재폐기물, 플라스틱 폐기물, 콘크리트 폐기물을 각각 10,000장씩 사용하였고 이중 24000장은 Train data, 6000장은 Test data로 나누어 사용했다. 모델 학습을 하는데에 생성된 parameter의 수는 40,480,899개이다.

표 1. 3모델 실험데이터(feature / data)
Table 1. 3model experimental data (feature / data)

category	wood	plastic	concrete	parameters
number	10,000	10,000	10,000	40,480,899

본 연구에서는 VGG-16모델의 batch size와 epoch 및 learning rate를 다르게 설정하여 수차례 정확도를 실험하였고 각각의 학습된 모델의 성능중 실제 검증 과정에서 정확도가 가장 높게 나온것은 VGG-16은 91.5%이고, 트랜스포머는 92.7%이다.

그림10, 11은 epoch(x축) 진행에 따라 학습결과(y축) 정확도(왼쪽) 및 loss(오른쪽)를 나타내고 있다.

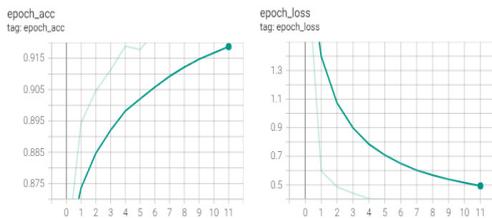


그림 10. VGG16 모델의 학습 결과와 Loss 그래프
Fig. 10. VGG16 Model learning result and loss graph

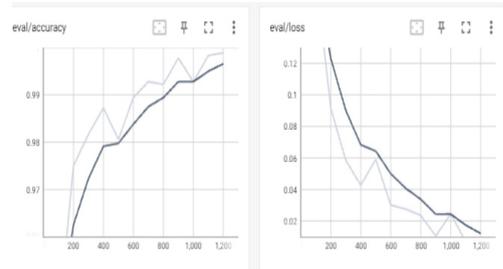


그림 11. ViT 모델의 학습 결과와 Loss 그래프
Fig. 11. ViT Model learning result and loss graph

다음 그림은 학습된 해당 모델의 예측결과를 시각화 한 것이다. label은 미리 지정된 클래스에 따라 validation data에 저장되어있는 값이고 pred는 학습된 모델이 예측한 결과값이다. 아래와같이 학습이 잘 된 것을 확인할 수 있다. 여러 모델을 비교해본 결과 test data에서 정확도가 더 높았던 모델보다 모델의 경우 validation data에서 예측을 할 시 더 부정확한 모습을 보인다. 또한 VGG-16모델 이외에 다른 모델도 실험해보아 수치상 비슷한 결과를 얻었으나 ViT 모델이 실제 validation시 더 높은 정확도를 보였다.



그림 12. 분류 시스템의 예측 결과
Fig. 12. Predicted results by the classification system

앞 그림은 600개의 Predict dataset에서 랜덤하게 추출한 16개의 이미지 중 일부분이다. 각각의 이미지에 상단에 위치한 텍스트 중 Label은 미리 지정해놓은 클래스를 의미하고 아래에 pred는 모델이 예측한 결과값이다.

V. 결 론

본 연구에서는 VGG-16과 ViT 모델을 활용하여 건축 폐기물 이미지 데이터를 분류하는 실험을 진행하였다. 이를 위해 육안으로 구분하기 어려운 건축 폐기물 이미지를 각각 1,000장씩 수집하였다. 실험 결과, 본 분류 시스템이 성공적으로 작동함을 확인하였다. 앞으로 더 많은 데이터와 높은 해상도의 이미지를 수집하여 모델의 과적합을 방지함으로써, 분류 정확도를 더욱 향상시킬 수 있을 것이다. 그리고 ViT 모델을 건축 폐기물 이미지 분류 문제에 적용해 보았다. 실험 결과, ViT 모델은 VGG-16 모델과 비교하여 높은 분류 정확도를 달성하였다. 이를 통해 건축 폐기물 이미지 분류에 ViT 모델이 효과적임을 확인할 수 있었다. 또한, 본 연구에서는 ViT 모델의 성능을 더욱 향상시키기 위해 데이터 증강 기법을 활용하였다. 데이터 증강은 원본 이미지에 여러 가지 변형을 적용하여 새로운 이미지를 생성하는 기법으로, 모델의 일반화 성능을 향상시키는데 도움이 된다. 본 연구에서는 이미지 회전, 이동, 확대/축소, 노이즈 추가 등의 방법을 사용하여 데이터를 증강하였으며, 이를 통해 모델의 성능이 향상된 것으로 확인되었다.

이상의 실험 결과를 바탕으로, 본 연구에서 제안하는 건축 폐기물 분류 시스템이 실제 건축 폐기물 처리 업무에서 효율적으로 활용될 수 있을 것으로 기대된다. 향후 연구에서는 이미지에 YOLOv5와 같은 객체 탐지 알고리즘과 의미론적 분할 기법을 적용하여, 한 이미지 안에 여러 종류의 건축 폐기물이 존재하는 경우에도 해당 폐기물 영역을 보다 간편하게 분할할 수 있는 방법을 탐구할 계획이다.

References

- [1] Karen Simonyan, Andrew Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition" arXiv 1409.1556, 2014
DOI:https://doi.org/10.48550/arXiv.1409.1556
- [2] https://en.wikipedia.org/wiki/Web_crawler
- [3] D Han, Q Liu, W Fan."A new image classification method using CNN transfer learning and web data augmentation." Expert Systems with Applications, 2018 - Elsevier
DOI:https://doi.org/10.1016/j.eswa.2017.11.028
- [4] A Bochkovskiy, CY Wang, HYM Liao. "Yolov4: Optimal speed and accuracy of object detection."

arXiv preprint arXiv:2004.10934, 2020
DOI:https://doi.org/10.48550/arXiv.2004.10934

- [5] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, Jose Garcia-Rodriguez. "A Review on Deep Learning Techniques Applied to Semantic Segmentation." arXiv:1704.06857, 2017
DOI:https://doi.org/10.48550/arXiv.1704.06857
- [6] S Agatonovic-Kustrin, R Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research" Journal of Pharmaceutical and Biomedical Analysis. Volume 22, Issue 5, pp. 717-727. 2000.
DOI:https://doi.org/10.1016/S0731-7085(99)00272-1
- [7] TM Mitchell. "Does Machine Learning Really Work?". 1997
DOI:https://doi.org/10.1609/aimag.v18i3.1303
- [8] LeCun, Y., Bengio, Y. & Hinton, G. "Deep learning." Nature 521, pp.436-444 2015
DOI:https://doi.org/10.1038/nature14539
- [9] https://ko.wikipedia.org/wiki/%EC%8B%A0%EA%B2%BD_%EC%84%B8%ED%8F%AC
- [10] P.G. Benardos, G.-C. Vosniakos, "Optimizing feedforward artificial neural network architecture," Engineering Applications of Artificial Intelligence, Volume 20, Issue 3, Pages 365-382, ISSN 0952-1976, 2007
DOI:https://doi.org/10.1016/j.engappai.2006.06.005
- [11] Gülden Kaya Uyanık, Neşe Güler, "A Study on Multiple Linear Regression Analysis." Procedia - Social and Behavioral Sciences, Volume 106, Pages 234-240, ISSN 1877-0428, 2013
DOI:https://doi.org/10.1016/j.sbspro.2013.12.027
- [12] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. "A Survey on Bias and Fairness in Machine Learning." ACM Comput. Surv. 54, 6, Article 115, 35 pages. 2022
DOI:https://doi.org/10.1145/3457607
- [13] Prajit Ramachandran, Barret Zoph, Quoc V. Le, "Searching for Activation Functions", arXiv:1710.05941, 2017
DOI:https://doi.org/10.48550/arXiv.1710.05941
- [14] Jun Han, Claudio Moraga. "The influence of the sigmoid function parameters on the speed of backpropagation learning" From Natural to Artificial Neural Computation, Volume 930, ISBN : 978-3-540-59497-0, 1995
DOI:https://doi.org/10.1007/3-540-59497-3_175
- [15] https://en.wikipedia.org/wiki/Sigmoid_function
- [16] Y. Bengio, P. Simard and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult" in IEEE Transactions on Neural Networks, vol. 5, no. 2, pp. 157-166, March 1994
DOI:https://doi.org/10.1109/72.279181
- [17] Abien Fred Agarap, "Deep Learning using Rectified Linear Units (ReLU)", arXiv:1803.08375, 2019
DOI:https://doi.org/10.48550/arXiv.1803.08375

[18] [https://en.wikipedia.org/wiki/Rectifier_\(neural_networks\)](https://en.wikipedia.org/wiki/Rectifier_(neural_networks))

[19] Lee, D., Sun, Y.-G., Kim, S.-H., Sim, I., Lee, K.-S., Song, M.-N., & Kim, J.-Y. (2020). CNN-based Image Rotation Correction Algorithm to Improve Image Recognition Rate. The Journal of The Institute of Internet, Broadcasting and Communication, 20(1), 225-229. <https://doi.org/10.7236/JIIBC.2020.20.1.225>

[20] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), pp. 1-6, 2017
DOI:<https://doi.org/10.1109/ICEngTechnol.2017.8308186>

[21] <https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/>

[22] <https://arxiv.org/pdf/1706.03762.pdf>
Ashish Vaswani, "Attention Is All You Need", arXiv:1706.03762v5 [cs.CL] 6 Dec 2017

[23] Kim, Hyun-Su "Control Performance Evaluation of Smart Mid-story Isolation System with RNN Model" Journal of the Korea Academia-Industrial cooperation Society, Volume 21 Issue 1, Pages.774-779, 2020, 1975-4701(pISSN), 2288-4688(eISSN)
DOI:<https://doi.org/10.5762/KAIS.2020.21.1.774>

[24] Jeong-Jae Kim, Sang-Min Park, Byung-Won On."(2023). A Pooled RNN-based Deep Learning Model based on Data Augmentation for Clickbait Detection.", JKIIIT, 21(4), 45-56.
DOI:<https://doi.org/10.14801/jkiit.2023.21.4.45>

양민철(준회원)



- 1997 ~ 1999 : 중앙대학교 경영학 (MIS) 석사
- 에이아이씨랩 대표이사
- 주관심 분야 : 메타버스, HCI, 인공지능 등

문경남(준회원)



- 2004 ~ 2007 : 성균관대학교 행정대학원 석사
- 메타크레용 대표이사
- 주관심 분야 : 인공지능응용업 등

김용국(준회원)



- 1997 : Cambridge Univ. Computer Vision(박사)
- 1986 ~ 1989 : KT, 전임연구원
- 1998 ~ 2001 : Smith-Kettlewell Vision institute USA, Research Associate
- 2001 ~ 현재 : 세종대학교 컴퓨터공학과 정교수
- 주관심분야 : 컴퓨터 비전, HCI

저자 소개

성재경(준회원)



- 2003 ~ 2007 : 세종대학교 소프트웨어 (박사)
- 에이아이씨랩 연구이사
- 주관심분야 : 인공지능, UI/UX, HCI, 모바일게임 등

※ 본 연구는 중소기업기술정보진흥원의 2022년도 지역특화산업육성+(R&D)-지역주력산업육성 2차의 일환으로 수행하였음.
과제번호:S3317531