

Development of Location Image Analysis System design using Deep Learning

Jin-Wook Jang*

*Professor, Cooperative Industry Major, Agricultural Cooperative University, Gyeonggi-do, Korea

[Abstract]

The research study was conducted for development of the advanced image analysis service system based on deep learning. CNN(Convolutional Neural Network) is built in this system to extract learning data collected from Google and Instagram. The service gets a place image of Jeju as an input and provides relevant location information of it based on its own learning data. Accuracy improvement plans are applied throughout this study. In conclusion, the implemented system shows about 79.2 of prediction accuracy. When the system has plenty of learning data, it is expected to predict various places more accurately.

▶ **Key words:** Deep-Learning, Unsupervised Learning, Image Analysis, Convolutional Neural Network, Location Image

[요 약]

본 연구는 장소 이미지를 수집하고 학습하여 사용자가 관심이 있어 하는 이미지의 장소를 예측하여 알려주는 서비스 개발을 목적으로 한다. 이미지 학습을 위한 이미지 데이터들은 크롤링 부분을 통해 수집되도록 설계되었다. 이미지 수집 이후 수집된 이미지들은 장소별로 라벨링 되어 CNN의 다양한 층을 통하여 학습된다. 각 층을 거칠 때마다 입력받은 학습 데이터는 최적화하여 특징 맵과의 비교를 반복하며 특정 장소 이미지의 특징 정보를 뽑아낸다. 충분한 학습 데이터가 쌓이면 다양한 장소 이미지들에 대해 예측이 가능하다. 학습 결과 모델의 정확도는 79.2로 높은 학습 정확도를 보였다.

▶ **주제어:** 딥러닝, 비지도형 기계학습, 이미지 분석, 콘볼루션 신경망, 지역 정보

I. Introduction

본 연구는 특정 제주 장소 이미지를 자동으로 수집하고 딥러닝하는 과정을 시스템으로 개발 및 연구하였다.

모델을 통해 학습한 뒤 그 학습 모델을 기반으로 사용자가 관심이 있는 제주 장소 이미지를 입력하였을 때 장소 정보제공을 목적으로 연구되었다. 그리고 차후 장소정보 주변의 관광명소나 주변정보 제공 등을 연계하여 발전시킬 예정이다.

구현된 모델은 방대한 이미지 데이터를 가진 구글에서 장소 이미지 데이터를 크롤링(crawling) 코드를 통해 자동 수집한다. 그리고 CNN(convolutional neural network)을 통하여 다양한 장소 이미지 학습시켜 장소 이미지 정보가 없는 사진이 주어졌을 때 위치 정보를 제공할 수 있도록 설계되었다. 본 시스템은 위치 정보가 없는 이미지의 장소를 예측한다는 점에서 다양한 서비스와 연계될 수 있다.

본 논문의 구성은 2장에서는 관련 연구로 CNN, 맥스 풀링(max pooling), 손실함수(loss function) 기술하였으며 3장에서는 본론으로 장소 이미지의 수집, 분류를 통한 데이터 셋을 통한 예측처리 과정에 관하여 기술하였다. 4장에서는 제안한 시스템의 테스트를 통하여 검증하였다. 그리고 5장에서는 제안한 시스템의 효용성과 활용방안을 제안하였다.

II. Preliminaries

1. Related works

1.1 CNN

딥러닝 분야에서 심층 신경망의 한 종류인 CNN은 합성곱 신경망이라고도 불린다. 이 신경망은 학습을 통해 자동으로 필터를 생성하는 게 특징이다. 위 필터를 이용하여 행렬 형태로 된 이미지 파일에 행렬곱셈을 적용해 이미지의 특징을 뽑아 높은 정확도의 학습 결과를 제공한다. CNN은 필터를 통해 특징 지도(feature map)를 만들고 중요한 특징만 남긴다. 그리고 차원을 줄이는 풀링을 통해 크기를 최소화하여 학습 결과를 생성한다[9].

이러한 CNN은 2차원 데이터를 학습하는데 적합한 신경망 모델로, 제주 이미지를 학습시키는 목적의 위 연구에 가장 적합한 신경망 모델이다. 본 연구에서는 수집된 장소 이미지를 CNN을 이용하여 특징을 추출하였다. 그리고 각각의 장소 이미지를 학습시키고 식별할 수 있다. 개념도는 그림 1와 같다[1,10].

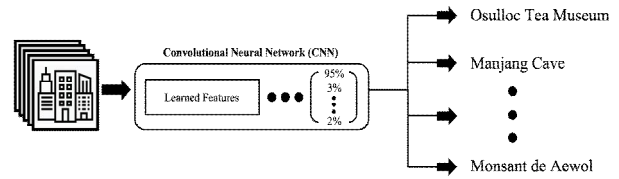


Fig. 1. Process of an Input Image Learning Process on CNN

텐서플로우(tensorflow)는 구글(google)에서 개발한 머신 러닝 라이브러리, 엔진으로 구글 검색, 음성 인식, 번역, AI 서비스 등 다양한 인공지능 분야에서 유용하게 쓰이고 있다[2].

1.2 Loss Function

CNN을 이용하여 예측을 원하는 사진 파일들의 저장장소를 입력하면 데이터 셋(dataset) 부분에서 저장된 이미지 데이터를 학습하고 학습 데이터를 가지고 예측을 진행한다.

풀링(pooling) 방식으로는 이미지를 최적화 시키는 맥스 풀링이 사용된다. 그림 2에서 볼 수 있듯 음수와 같은 불필요한 값을 받으면 무조건 0값을 리턴 하여 다른 수식을 이용하였을 때보다 현저히 빠른 처리 속도를 보이는 ReLU 함수가 손실함수가 이용되었다[7].

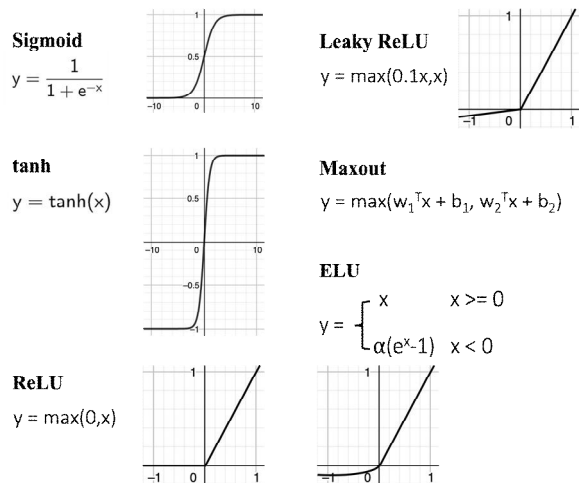


Fig. 2. loss function

유명한 인공지능 바둑 로봇인 알파고도 텐서플로우를 이용한 인공지능 머신이다. 텐서플로우를 이용하면 이미지 인식, 반복 신경망, 신경망 학습 등을 위한 알고리즘을 손쉽게 구현할 수 있다[8].

1.3 Max Pooling

텐서플로우 라이브러리는 다양한 연산 처리를 작성할 수 있도록 제작되어 이미지 파일, 즉 행렬 타입의 파일을

처리해야 하는 위 연구에 유용한 라이브러리가[3]. 본 연구에서 텐서플로우는 이미지 학습을 위한 맥스 풀링, 손실 함수 등 다양한 연산 처리에 사용된다[5].

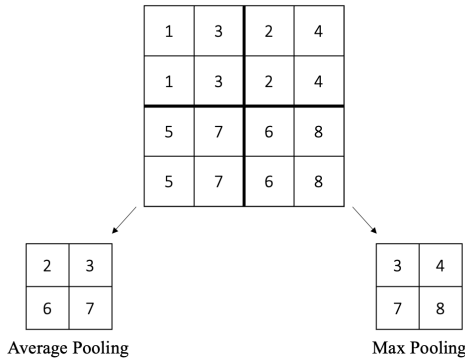


Fig. 3. Max pooling

위와 같은 정보들을 얻고 연구에 적용하기 위해 딥러닝의 전체적인 구조와 학습을 하는 원리, 알고리즘을 학습하였고 이후에는 CNN을 구성하는 풀링 계층(Pooling Layer), 활성화 함수(activation function), 손실함수(Loss Function), 합성곱 계층(convolutional layer) 등이 있다.

III. The Proposed Scheme

본 구현은 이미지 데이터를 수집하는 크롤링 부분, 수집된 이미지 파일들을 학습에 적합하게 변환시켜주는 데이터 셋 부분, 변환된 데이터 셋을 학습하는 CNN 부분, 그리고 학습된 데이터를 기반으로 주어진 장소 이미지를 예측하는 부분으로 구성된다.

개발된 알고리즘의 이미지 크롤링 부분은 그림 4와 같이 엑셀 파일에 수집을 원하는 장소 이미지를 작성하고 라벨링 하면 해당 라벨 명을 갖는 하위 폴더를 생성한다. 그리고 각 장소 이미지를 구글에서 자동 검색 및 수집한 후에 상응하는 폴더에 저장되는 형식으로 설계하였다.

	A	B	C	D	E	F	G	
1	0	1	2	3	4	5	6	
2	하이 엔드	원앤원리	월드	선릉상대	에펠	명진전복	종달리엔	수우동

Fig. 4. Excel file contents

위 과정을 통해 무작위로 저장된 이미지들은 전처리 과정을 통해 잘못 수집되었거나 중복수집 된 이미지는 제거되고 남은 이미지들은 240*240 변경되고 통일된다.

전처리 프로그램이 저장된 이미지를 확인하는 과정에서

이미지 파일을 인식할 수 없다는 메시지가 발생하면 try와 except 함수를 통해 예외를 발생시키고 파일은 제거된다. 이러한 방식으로 프로그램은 시스템 에러를 처리하고 유효하지 않은 이미지를 처리한다. 전처리 과정을 통과한 이미지 데이터들은 딥러닝 모델에게 학습 데이터로 주어지는데, 높은 학습률을 얻기 위해서는 주어진 장소와 관련 없는 데이터를 걸러내는 작업을 요구한다. 이것은 사용자가 직접 걸러주는 것이 가장 정확하다.

데이터 셋 부분은 이전 과정에서 통과된 이미지들과 각 라벨을 딥러닝 모델이 쉽게 받고, 읽을 수 있게 포맷을 변경해주는 단계이다. 학습 이미지가 저장된 경로를 입력해 주면 데이터 셋은 경로에 저장된 모든 파일을 이들의 라벨과 함께 .npy 형태로 만들어준다. 이 포맷의 파일은 텐서플로우에서 지원하는 이미지 데이터 셋인 MNIST와 같은 구조로 CNN 파트에서 데이터를 불러오는 데 유용하게 활용된다.

실질적인 학습이 진행되는 CNN 부분에서는 모델로부터 학습 데이터를 받을 수 있는 단계이다. 컨볼루션층의 필터, 스트라이드, 패딩의 사이즈를 입력해주면 이미지에 패딩 사이즈 만큼의 NULL 값이 생성되어 필터가 스트라이드 사이즈에 따라 주어진 이미지 위를 움직이며 이미지 특징값을 계산한다.

원본 이미지의 사이즈가 줄어드는 것을 방지하기 위해 패딩 사이즈 만큼의 빈 유닛을 돌려주는 것이다. 위 과정에서 필터가 원본 이미지를 지나가며 수식 2와 같이 데이터의 볼륨은 줄이고 선과 점이 주를 이루는 저수준(low-level) 특징들만 선택하게 된다[4], 컨볼루션 레이어의 결과값을 연산하기 위하여 컨볼루션 행과 열의 크기를 입력된 특성을 연산하여 출력을 산출하였다. 수식은 (1)와 같다[6].

$$n_{out} = \frac{n_{in} + 2p - k}{s} + 1 \tag{1}$$

n_{out} : number of output feature
 n_{in} : number of input features
 k : convolution kernel size
 p : convolution padding size
 s : convolution stride size

이후 무작위로 수집된 이미지들 전처리작업으로 28x28 픽셀로 변환시키고 유효하지 않은 파일은 자동으로 걸러내도록 하였다. 전처리작업은 이미지 파일을 오픈하는 과정에서 예외처리(exception)를 이용하여 “이미지 파일을 식별할 수 없다(cannot identify image file)”의 에러 문과 함께 운영체제 오류가 발생할 때 예외를 주어 파일을 삭제하였다.

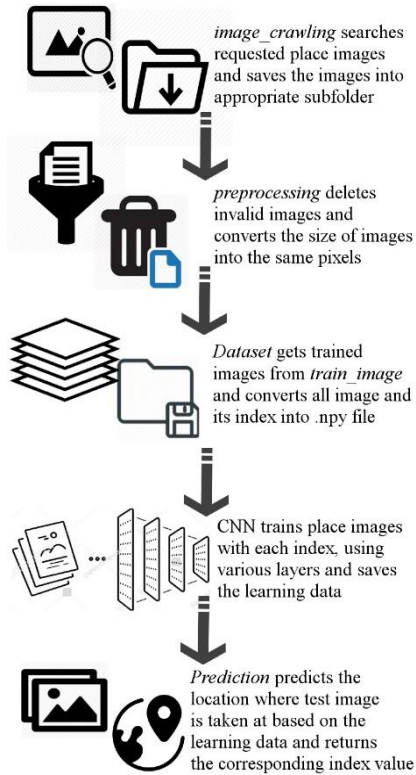


Fig. 5. System structure of Place Image Crawling, Learning, and Prediction Model

데이터 셋 부분은 학습된 이미지가 저장된 최상위 폴더의 경로를 입력하면 for 문과 주어진 경로의 파일을 읽는 OS 라이브러리를 통해 최초 입력한 경로 속 모든 하위 폴더를 읽어온다. 위 작업을 통해 컴퓨터는 장소별로 라벨링된 각 하위 폴더로부터 사용자가 높은 정확도를 위해 미리 분류해둔 학습 이미지를 전달한다.

그리고 CNN 부분에서 쉽게 데이터를 이용할 수 있도록 행렬을 다루는 numpy 라이브러리와 이미지를 다루는 CV 라이브러리를 이용해 이미지 데이터들을 라벨과 함께 .npy 형태의 파일로 만들고 저장한다. 텐서플로우에서 제공하는 MNIST 파일과 같이 저장된 .NPY 파일로부터 X_train, X_test, Y_train, Y_test 값들을 얻을 수 있어 CNN 코드에서 편리하게 학습하는 라인을 개발할 수 있다.

마지막으로 CNN 코드는 학습 결과를 바로 알 수 있는 학습과 예측하는 코드에 예측을 원하는 사진 파일들의 저장장소를 입력하면 데이터 셋 부분에서 저장된 이미지 데이터를 학습하고 학습 데이터를 가지고 예측을 진행한다. 풀링 방식으로는 이미지를 최적화 시키는 맥스 풀링이 사용되었대[4,5].

그림 5는 구현된 모델의 알고리즘을 시각화한 것으로 다음과 같이 크게 5가지 스텝으로 정리할 수 있다. 첫째,

사용자가 원하는 이미지를 수집하기 위해 16진수(hexadecimal)를 이용해 라벨링 해주고, 원하는 장소명들과 레벨을 엑셀 파일에 그림 5와 같이 작성한다. 둘째, 이미지 크롤링 코드가 전 단계에서 저장한 엑셀 파일을 읽고 자동으로 각 장소별 폴더를 제작한다[11].

이후 장소 이미지들이 자동 검색, 수집되고 알맞은 하위 폴더에 순차적으로 저장된다. 셋째, 정확한 학습을 위해 크롤링한 사진을 토대로 사용자가 관련 없는 이미지를 제거한다. 네 번째로 전처리 코드를 통해 유효하지 않은 파일을 삭제한 후 학습 이미지들을 28×28픽셀로 변환시키고 처리한다[12].

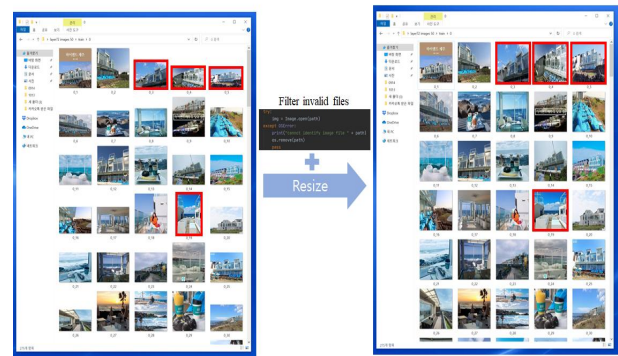


Fig. 6. Resizing

다섯 번째, 이후 데이터 셋 코드를 통해 이미지들을 .npy 형태의 데이터 셋으로 변경한다.



Fig. 7. Convert

여섯 번째 스텝에서 CNN 코드를 통해 딥러닝 머신의 합성곱 계층, 맥스 풀링 계층, ReLU 함수를 거쳐 학습이 진행된다. 학습 후에는 계층적 데이터가 담긴 h5 파일 형태로 학습 결과가 저장된다.

예측 결과 16진수의 형태로 반환되며 라벨링 표를 통해 예측된 이미지의 장소를 확인할 수 있다.

IV. Experimental results

테스트를 위한 장소로 제주 유명 카페인 ‘하이엔드’와 ‘원앤온리’의 정확한 장소 사진을 선별 후 수집하고 각각 0과 1로 라벨링 하였다. 그리고 그림 2에서 확인할 수 있듯이 각 3,368개 학습된 이미지를 cnn_sample 폴더에 라벨링 된 0과 1 하위 폴더를 만들어 수집한 이미지를 알맞게 저장하고 CNN 딥러닝 모델로 학습시켰다.

학습 결과 모델의 정확도(accuracy)는 그림 8와 같이 0.9994, 손실 값(loss)은 그림 9와 같이 0.0015로 높은 학습 정확도를 보였다.

```

1472/1775 [=====>.....] - ETA: 28s - loss: 0.0016 - accuracy: 0.9993
1504/1775 [=====>.....] - ETA: 25s - loss: 0.0016 - accuracy: 0.9993
1536/1775 [=====>.....] - ETA: 22s - loss: 0.0016 - accuracy: 0.9993
1568/1775 [=====>.....] - ETA: 19s - loss: 0.0016 - accuracy: 0.9993
1600/1775 [=====>.....] - ETA: 16s - loss: 0.0015 - accuracy: 0.9994
1632/1775 [=====>.....] - ETA: 13s - loss: 0.0015 - accuracy: 0.9994
1664/1775 [=====>.....] - ETA: 10s - loss: 0.0015 - accuracy: 0.9994
1696/1775 [=====>.....] - ETA: 7s - loss: 0.0015 - accuracy: 0.9994
1728/1775 [=====>.....] - ETA: 4s - loss: 0.0015 - accuracy: 0.9994
1760/1775 [=====>.....] - ETA: 1s - loss: 0.0015 - accuracy: 0.9994
1775/1775 [=====] - 168s 95ms/step - loss: 0.0015 - accuracy: 0.9994

```

Fig. 8. Accuracy of the learning process

이후 예측 결과를 테스트하기 위해 검색엔진을 통해 무작위로 수집된 카페 이미지별 각 500장의 이미지들이다. 총 500장의 이미지들을 예측해 본 결과 그림 8에서 확인할 수 있듯이 예측 결과와 테스트 이미지의 장소 라벨값이 정확히 일치하는 확률은 그림 9와 같이 79.2%의 정확도를 보였다.

```

-----
396 predictions success out of 500 tests
It has a success rate of 79.2%
Finished predictions on jeju_images_filtered_size50 with google_first1
-----

```

Fig. 9. Success rate of the prediction

각각의 이미지에 대한 예측값은 그림 10와 같다. 테스트 이미지 중 무작위로 포함된 인물 사진 등 관련성이 없는 이미지들을 제외한다면 더 높은 정확도를 보일 것으로 예상된다.

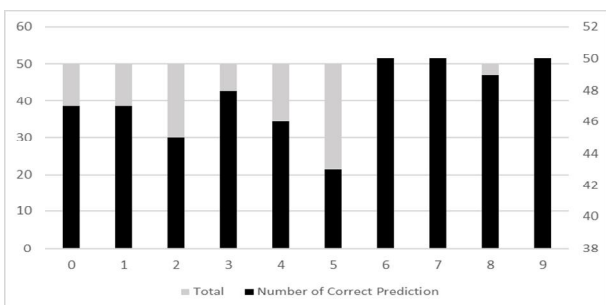


Fig. 10. The Result of The Predictions

위 모델은 엑셀 파일을 통해 크롤링 부분에 방대한 제주 장소명을 제공해주어 다양한 장소별로 많은 이미지를 수집했을 때 더 많은 장소 정보와 이미지 데이터를 딥러닝 모델에 학습시킬 수 있다. 학습 데이터양이 많고 다양한 수록 예측 정확도와 제공할 수 있는 장소 정보가 증가한다[3].

따라서 위 내용을 통해 모델을 발전시킨다면 다양한 제주 장소 이미지를 대상으로 높은 정확도를 제공할 수 있을 것으로 기대된다.

V. Conclusions

위 연구에는 다양한 제주 장소 이미지를 수집하고 딥러닝 머신을 통해 수집한 이미지를 학습시켜 위치 정보가 없는 제주 장소 이미지를 제공 받았을 때 위치 정보를 제공하기 위한 목적으로 알고리즘을 설계하고 구현하였다. 딥러닝 모델을 구현하기 위해 이미지 학습에 가장 적합한 신경망 모델인 CNN을 설계하였으며 높은 정확도와 빠른 학습을 위해 맥스 풀링과 ReLU함수를 이용하였다. 또한 위 모델을 구현하기 위하여 구글에서 제공하는 머신러닝 라이브러리인 텐서플로우를 이용하여 2차원 배열로 이미지 계산 및 학습시켰다.

테스트 결과 검색엔진을 통해 무작위로 수집된 카페 이미지별 각 500장의 이미지들을 대상으로 예측해 본 결과 79.2%의 정확도를 보였다. 향후에는 설정 및 모델 튜닝 등에 대한 다양한 테스트를 통하여 신뢰성을 확보할 필요가 있다.

ACKNOWLEDGEMENT

This research was supported by National Research Foundation of Korea in 2021(No. NRF-2020R1G1A1005872)

REFERENCES

- [1] Marina Chatterjee, “What is TensorFlow? The Machine Learning Library Explained”, 2020.
- [2] Tensorflow, “Customization basics: tensors and operations”, 2020.
- [3] R. Bevans, An introduction to the two-way ANOVA(2020), Retrieved Sept., 13, 2020.

- [4] N. Taback, "Design of Experiments and Observational Studies: Factorial Designs at Two Levels -2^k Designs", Retrieved Sept., 13, 2020.
- [5] J. Brownlee, A Gentle Introduction to the Rectified Linear Unit (ReLU)(2020), Retrieved Sept., 12, 2020.
- [6] Nathan Taback. "Design of Experiments and Observational Studies", 2020.
- [7] Jason Brownlee, "A Gentle Introduction to Pooling Layers for Convolutional Neural Networks", 2019,
- [8] Jason Brownlee, "A Gentle Introduction to the Rectified Linear Unit (ReLU)", 2019.
- [9] Prabhu, "Understanding of Convolutional Neural Network (CNN) Deep Learning", 2018.
- [10] M. J. Evan and J. S. Rosenthal, Probability and Statistics: The Science of Uncertainty, Second Edition, CA: University of Toronto, pp. 586-590, 2005.
- [11] Grace Karimi, "Introduction to YOLO Algorithm for Object Detection", 2021.
- [12] Bum-Suk Jang et al., " Sub-Frame Analysis-based Object Detection for Real-Time Video Surveillance", International Journal of Internet, Broadcasting and Communication, Vol.11, No.4, pp.76-85, 2019.

Authors



Jin-Wook Jang received in Ph.D. of Management Engineering from Konkuk University. The main interests are digital transformation, location Image deep learning Dr. Jang worked in Korea Defense

Intelligence Command of Ministry National Defense as a Computational officer and in SK communications as a PMO Manager. He is Professor at Agricultural Cooperative University