

Design of Deep Learning-based Location information technology for Place image collecting

Jin-wook Jang*

*Professor, Office of Industry-Academy Cooperation Foundation, Hanyang Women's University, Seoul, Korea

[Abstract]

This research study designed a location image collecting technology. It provides the exact location information of an image which is not given in the photo to the user. Deep learning technology analysis and collects the images. The purpose of this service system is to provide the exact place name, location and the various information of the place such as nearby recommended attractions when the user upload the image photo to the service system. Suggested system has a deep learning model that has a size of 25.3MB, and the model repeats the learning process 50 times with a total of 15,266 data, performing 93.75% of the final accuracy. This system can also be linked with various services potentially for further development.

▶ **Key words:** Deep Learning, Place Image, CNN(Convolutional Neural Network), Image Collecting, Web Service

[요 약]

본 연구에서는 딥러닝 처리기술을 이용한 이미지 분석을 통하여 위치정보가 없는 사진의 위치를 사용자에게 제공하는 장소이미지 수집기술을 설계하였다. 본 서비스는 사용자가 생활 중에 관심 있는 장소의 이미지 사진을 서비스에 업로드하면 해당 장소의 이름과 위치뿐만 아니라 관련 주변 정보를 확인 할 수 있는 서비스 개발을 목적으로 설계되었다. 본 연구는 이미지에 해당하는 정보를 제공하고 그 위치 정보를 기반으로 사용자가 관심 있는 주변정보를 제공할 수 있는 서비스의 기반기술이다. 이를 통하여 다양한 서비스에 활용이 가능하다.

▶ **주제어:** 딥러닝, 장소 이미지, 컨볼루션 뉴럴 네트워크, 이미지 수집, 웹 서비스

-
- First Author: Jin-wook Jang, Corresponding Author: Jin-wook Jang
 - *Jin-wook Jang (jinwchng@hywoman.ac.kr), Office of Industry-Academy Cooperation Foundation, Hanyang Women's University
 - Received: 2020. 06. 11, Revised: 2020. 08. 29, Accepted: 2020. 08. 29.

I. Introduction

일반적으로 사진의 위치를 찾기 위해서는 사진파일에 있는 위도, 경도의 정보가 담긴 위치태그를 이용한다. 혹은 검색엔진을 이용하여 해당 위치이미지를 검색하는 방법을 사용한다.

여행자들은 대부분 스마트 폰으로 사진을 촬영하며 이는 다른 사람이 촬영한 여행사진을 보고 사진 속 장소를 찾아가려는 여행객이 있는 경우 위치정보를 찾기 어렵다. 그리고 이미지파일을 웹사이트에 업로드 시 원본파일이 그대로 올라가지 않고 수정되어 업로드 되는 경우 파일의 위치태그가 사라진다. 그 외 스마트폰의 기본 카메라 애플리케이션이 아닌 무음카메라 애플리케이션의 경우 대부분 스마트 폰의 화면을 캡처하는 방식의 촬영이기 때문에 위치태그가 담기지 않는다.

본 연구에서는 위치태그가 없는 사진의 위치를 알 수 있도록 딥러닝을 이용하여 위치를 추정하여 사용자에게 제공하는 서비스를 설계하고 그 정확도를 측정하였다.

본 서비스는 사용자에게 장소 이미지 위치와 연계하여 주변 관광명소, 음식점, 숙박업소 등 주변 환경 정보를 제공하는 웹 서비스의 기반기술을 목표로 하고 있다.

단, 서비스의 정확도를 향상하기 위하여 충분한 장소 이미지 데이터가 필요하다. 그 양이 많을수록 정확도가 증가한다. 그래서 자동으로 장소 이미지 데이터를 수집할 수 있는 이미지 수집(collect image) 부분과 이미지를 이용한 딥러닝(deep learning)엔진 설계를 통하여 정확도를 개선하고 하나의 서비스로 발전시키기 위한 설계 및 측정을 하였다.

본 논문의 구성으로 2장에서는 본 연구에 활용되고 있는 구글의 이미지 검색 및 관련 기술들을 소개한다. 3장에서는 시스템의 구성 및 특징에 대하여 기술한다. 4장에서는 딥러닝 기반의 장소이미지 수집 알고리즘 및 처리방식을 제시한다. 5장에서는 서울 대표 명소 이미지를 대상으로 이미지를 학습하고 정확도를 측정하였다. 6장에서는 그 결과를 평가하고 7장에서 결론을 통하여 활용방안을 제시하고 마무리 하였다.

II. Relavant Service and Technology

구글의 이미지 검색 서비스는 이미지의 URL(uniform resource locator)을 입력하거나 이미지 파일을 업로드하면, 해당 이미지와 유사한 이미지들과 해당 이미지들에서 공통된 키워드를 알려준다. 형태적으로 유사한 이미지와

공통키워드에 해당하는 검색어와 연관된 키워드가 나오는 등 다용도의 목적으로 유용하게 사용이 가능하여 이용자가 많다.[1]

하지만 다목적에는 전문성의 부족이 따른다. 데이터의 양이 정확도와 비례하는 기계학습의 특성 상 데이터가 세계에서 가장 많다고 해도 과언이 아닌 구글의 정확도가 부족한 것은 아니지만 전체 검색결과에서 많이 나오는 키워드를 출력하여 위치만을 확인하기에는 제한적이다.

딥러닝 분야에서 심층 신경망의 한 종류인 CNN(Convolutional Neural Network)은 합성곱 신경망이라고도 불린다. 이 신경망은 학습을 통해 자동으로 필터를 생성하는 게 특징이다. 위 필터를 이용하여 행렬 형태로 된 이미지 파일에 행렬곱을 적용시켜 이미지의 특징을 뽑아 높은 정확도의 학습 결과를 제공한다. CNN은 필터를 통해 특징 맵(feature map)을 만들고 중요한 특징만 남기며 차원을 줄이는 풀링을 통해 크기를 줄여 최적화된 학습 결과를 남긴다.[1] 이러한 CNN은 2차원 데이터를 학습하는데 적합한 신경망 모델로, 제주 이미지를 학습시키는 목적의 위 연구에 가장 적합한 신경망 모델이다. 따라서 위 연구에서는 딥러닝 모델을 CNN을 이용해 제주의 장소 이미지들의 특징을 추출, 학습시켜 높은 정확도의 학습 데이터를 갖는다.[2,3]

텐서플로로는 구글에서 개발한 머신 러닝 라이브러리, 엔진으로 구글 검색, 음성 인식, 번역, AI서비스 등 다양한 인공지능 분야에서 유용하게 쓰이고 있다. 유명한 인공지능 바둑 로봇인 알파고도 텐서플로로를 이용한 인공지능이다. 텐서플로로를 이용하면 이미지 인식, 반복 신경망, 신경망 학습 등을 위한 알고리즘을 손쉽게 구현할 수 있다.[2] 텐서플로로 라이브러리는 다양한 연산처리를 작성할 수 있도록 제작되어 이미지 파일, 즉 행렬 타입의 파일을 처리해야하는 위 연구에 유용한 라이브러리이다.[3] 본 연구에서 텐서플로로는 이미지 학습을 위한 맥스풀링(Max Pooling), 손실 함수(Loss Function)등 다양한 연산처리에 사용되었다.[4]

위 와 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), 손실 함수(Loss Function), 합성곱 계층(Convolutional Layer) 등이 있다.[5,6]

III. System Structure

딥러닝 기반의 위치정보 제공 서비스의 구조는 Fig. 1과 같이 장소이미지를 수집하는 부분과 이미지를 학습하는 과정으로 이루어진다.

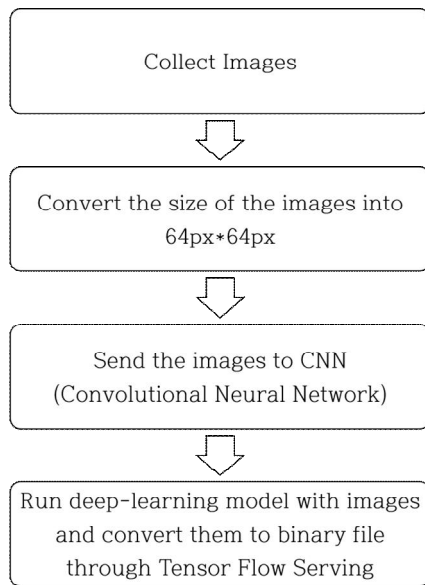


Fig. 1. The System Structure

장소 이미지가 수집되면 64x64 사이즈로 변경한다. 그리고 변경된 이미지는 CNN 알고리즘을 통하여 학습되는 과정을 거치고 장소정보와 매핑한다.

제한된 시스템 구조에서 정확도에 영향을 미칠 수 있는 부분은 최초 장소 이미지를 수집하는 과정에서 정확한 장소 이미지와 잘못된 장소 이미지를 구분하는 것이며 이러한 과정에서 발생할 수 있는 중복 이미지를 처리하는 것이다. 이러한 과정이 전처리 되어 학습대상 데이터 셋을 구성한다.

테스트는 서울 주요명소 정확한 장소사진을 선별 후 수집하고 각각 0과 1로 라벨링 하였다. 그리고 학습된 이미지를 CNN폴더에 라벨링된 0과 1 하위 폴더를 만들어 수집한 이미지를 알맞게 저장하고 CNN 딥러닝 모델로 학습한다. 마지막으로 학습결과 모델의 정확도를 측정한다.

IV. Design of Deep Learning-based Local Image Collecting Model

딥러닝 엔진의 프레임워크로는 구글의 텐서플로로(tensorflow)를 사용하였고 개발 언어는 파이썬(python)

을 사용하였다. Fig. 2 장소 이미지 딥러닝 모델은 빌드가 완료된 이후부터 사용자에게 새로운 모델이 배포되기까지의 과정을 나타내고 있다. score 변수와 classes 변수는 신경망에서 발생하는 출력이다.

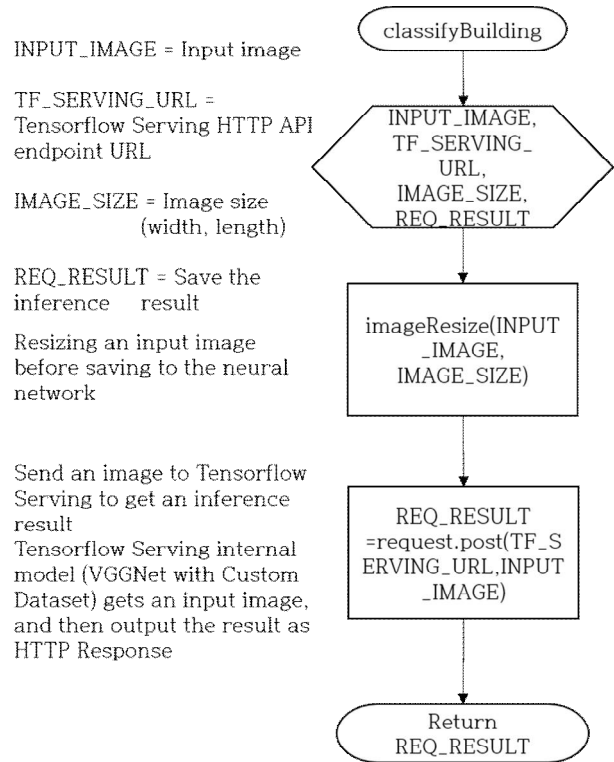


Fig. 2. Process of Tensorflow-Serving

output_signature 는 모델 추출 시 저장된 모델의 입력과 출력을 정의하는 객체이며 builder 는 텐서플로로에 내장된 모델 빌더 객체이며 세션(session)은 프로그램의 세션으로 즉, 데이터를 의미한다. 필요한 객체와 변수를 선언하고 저장될 모델의 구조를 정의한 다음 builder 객체에서 모델을 저장한다.

그리고 텐서플로로서빙(tensorflow-serving) 파트 모듈은 텐서플로로의 서브 프로젝트로서 어떻게 하면 “저장된 모델을 효과적으로 제공 할까?”하는 관점에서 개발되었다. 텐서플로로서빙은 Fig. 2와 같이 저장된 모델에 이미지를 입력받고 HTTP Response 형태로 보내주게 되는 것이다.

텐서플로로서빙은 저장된 모델에 대한 추론을 인공지능 전문가가 아니더라도 사용이 편리하다는 특징이 있다.[2]

장소 이미지 딥러닝 모델의 전체 흐름은 Fig. 3과 같다. 사용자로부터 이미지 입력이 들어오면 가장먼저 입력 받은 파일이 유효한지 확인하는 과정을 거친다.

이후 유효성 체크에 실패하면 바로 알고리즘을 종료시키고 성공할 경우 다음 과정인 물체 감지 신경망에 파일을

넣고 결과를 저장한다. 이후 위 결과를 조건으로 거짓일 경우를 분류한다.

2차 분류 시 활용하기 위하여 저장하고 알고리즘을 끝내며 참일 경우 다음 단계인 장소 분류 신경망으로 넘어가게 된다. 위 신경망의 추론 결과를 저장하고 그 추론 결과를 바탕으로 저장 폴더를 결정한다. 이후 2차 분류 작업을 통하여 더욱 개선된 정확도의 데이터 셋으로 신경망을 재학습 시킨다.[4,5]

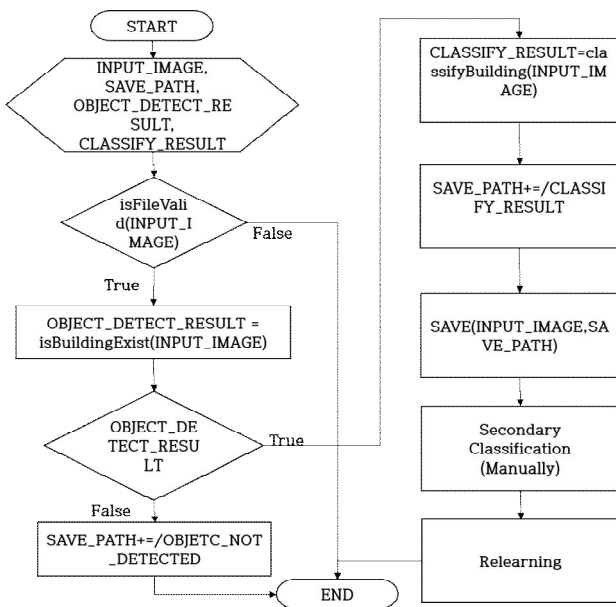


Fig. 3. Design of Local Image Deep Learning Model

이러한 과정을 통하여 중복 이미지에 대한 처리를 강화하였다. 각 레이어 들이 상호 연결된 텐서는 Fig. 4와 같이 텐서보드를 통하여 확인 할 수 있다.

V. Performance Analysis

시스템 정확도를 향상시키기 위하여 해당 시스템의 성능을 측정하였다. 측정은 컴퓨터 행렬(confusion matrix)을 이용하였다. 분류기준은 Table 1과 같다.[3]

Table 1. Confusion Matrix

Criteria	Expected Value		
	Division	True	False
	True	True Positive (TP)	False Positive (FP)
False	False Negative (FN)	True Negative (TN)	

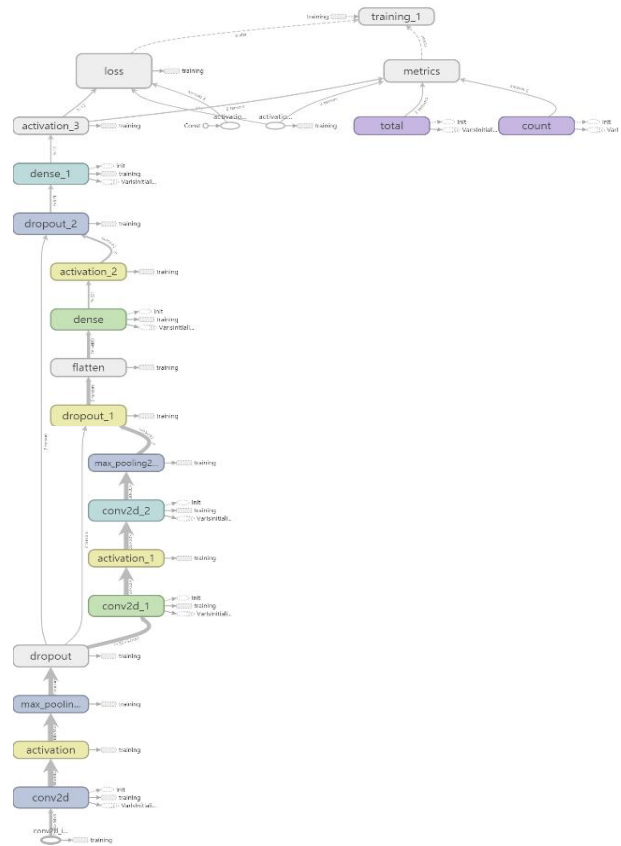


Fig. 4. The Interrelationship of each layer

평가 장소이미지는 서울의 주요 명소인 63 빌딩, 서울 시청, 이순신 장군 동상, 세종대왕 동상 등의 12곳을 대상으로 평가하였다.

지표는 딥러닝의 정확도를 측정하기 위하여 정확도 (accuracy), 정밀도(precision), 재현율(recall) F1 점수를 사용하였다. 정확도 측정 측정결과 정밀도 $\frac{TP}{TP+FP}$ 는

0.84, 재현율 $\frac{TP}{TP+FN}$ 은 0.33, 정확도

$\frac{TP+TN}{TP+FN+FP+TN}$ 는 0.75, 1 score

$$F1 - score = 2 \times \frac{1}{\frac{1}{\text{정밀도}} + \frac{1}{\text{재현율}}} = 2 \times \frac{\text{정밀도} \times \text{재현율}}{\text{정밀도} + \text{재현율}}$$

은 7.66 로 측정되었다.

학습모델의 크기는 25.3MB이며 전체 15,266 개의 데이터를 50 회 반복 학습하였으며 최종 정확도는 다음 Fig. 5 과 같이 0.94% 이다.

학습 데이터 셋인 장소이미지를 수집하는 과정에서 중복 이미지의 수집, 파일네임은 일치하나 관련성이 없는 이미지 파일인 경우, 장소이미지 수집 주기의 변화에 따른 중복이미지 등 다양한 학습이미지 오류가 발생하고 있다. 학습 데이

터의 정확도를 개선하기 위하여 장소이미지를 수집하는 과정에서 정확도를 향상하기 위한 노력이 필요하다.

```
600/600 [=====] - 1s 2ms/sample - loss: 0.9174 - acc: 0.9375
loss = 0.9174310119946798
accuracy = 0.9375
Process finished with exit code 0
```

Fig. 5. The Accuracy Calculation Result

정확도는 반복횟수가 증가 할수록 Fig. 6와 같이 향상되는 모습을 보인다. 10회까지는 증가정도가 높고 차츰 낮아지며 “1”에 수렴하는 모습을 보인다.

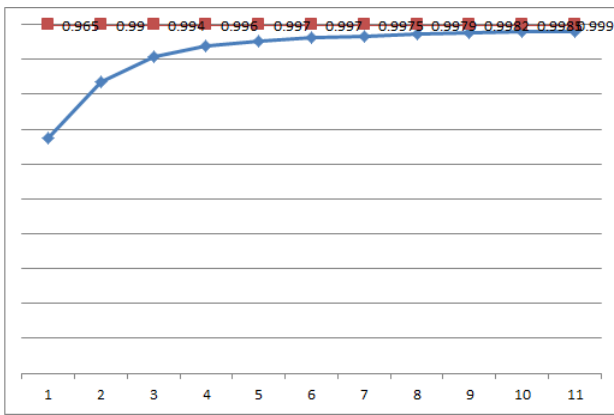


Fig. 6. Graph of The Deep Learning Model Accuracy

VI. Substantiation




시스템의 서비스 적용 가능성을 측정하기 위하여 Table 2와 같이 서울의 주요 장소이미지는 선정하고 인덱싱 하였다.

Table 2. Index of Place Image

Index	Place
0	63 Building
1	Seoul City Hall
2	A statue of Admiral Yi Sun-shin
3	A statue of King Sejong
4	Geunjeongjeon Hall in Gyeongbokgung Palace
5	Olympic Park
6	Lotte World Tower
7	Myeongdong Catholic Cathedral
8	The Namsan Tower
9	Seokchon Lake
10	Seodaemun Prison History Museum
11	Seoul Station Old Platform

Table 2의 측정값의 인덱싱 값이 “1.0”인 경우 측정결과가 성공을 의미한다. 대표 장소 이미지로 서울시, 롯데타워, 남산타워를 대상으로 측정하였으며 결과는 Table 3과 같다. 기타 장소의 경우 장소 이미지 수집과정에서 동일이름의 다른 이미지, 유사 중복 이미지의 등 수집 전처리과정에서 오류가 발생하였으며 차후 보완이 필요하다.

Table 3. Prediction Result Example of Place Image

Image	prediction	Index
	[0.0, 1.0 , 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]	Label Index 1 indicates that the probability that given image is 'Seoul City Hall' is 100%
	[0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 1.0 , 0.0, 0.0, 0.0, 0.0, 0.0]	Label Index 6 indicates that the probability that given image is 'Lotte World Tower' is 100%
	[0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 1.0 , 0.0, 0.0]	Label Index 8 indicates that the probability that given image is 'The Namsan Tower' is 100%

VII. Conclusion

본 서비스는 지역정보를 기반으로 제공하는 서비스로의 실용화가 용이하며 대표 장소이미지를 대상으로 적용하였을 때 그 효과가 크다.

그러나 제한된 장소 이미지를 수집하고 학습시켰으며 실용화를 위하여 수집과정의 중복 및 유사 이미지 제거가 필요하며 다양한 장소 이미지의 수집 및 학습 알고리즘 적용이 필요하다.

본 연구에서는 장소 이미지에 위치태그가 없는 사진을 수집하여 장소 이미지를 답러닝하는 모델을 설계 하였다. 학습모델을 통하여 반복 학습하였으며 최종 정확도는 93.75%로 측정되었다. 모델 설계를 통하여 인식된 장소를 기반으로 주변 관광명소, 음식점, 숙박업소 등 주변 환경 등의 정보를 제공하는 서비스로 활용가능하다.

ACKNOWLEDGEMENT

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

REFERENCES

- [1] Google Lens. <https://lens.google.com/>
- [2] Google Inc, "Tensorflow Serving Documentation. <https://www.tensorflow.org/tfx/guide/serving>, 2016.
- [3] Salma Ghoneim. Accuracy, Recall, Precision, F-Score & Specificity, which to optimize on?. Towards Data Science. <https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124>, 2019.
- [4] H. J'egou, M. Douze, C. Schmid, and P. P'erez, "Aggregating Local Descriptors into a Compact Image Representation," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 3304-3311, 2010.
- [5] Jiazheng Lu, Yu Ye, Xunjian Xu and Qinpu Li, "Application research of convolution neural network in image classification of icing monitoring in power grid", EURASIP Journal on Image and Video Processing, 2019.
- [6] Prabhu, "Understanding of Convolutional Neural Network (CNN) Deep Learning", 2018.

Authors



Jin-wook Jang received in Ph.D. of Management Engineering from Konkuk University. The main interests are location Image deep learning, Software Quality Assurance, Test Process.

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 Office of Industry-Academy Cooperation Foundation, Hanyang Women's University.