

ISSN: 2508-7894 © 2017 KAIA. <http://www.kjai.or.kr>

Doi: <http://dx.doi.org/10.24225/kjai.2017.5.1.10>

A Study on the Evaluation of Optimal Program Applicability for Face Recognition Using Machine Learning

기계학습을 이용한 얼굴 인식을 위한 최적 프로그램 적용성 평가에 대한 연구

¹ Min-Ho Kim (김민호), ² Ki-Yong Jo (조기용), ³ Hee-Won You (유희원), ⁴ Jung-Yeal Lee (이정렬),
⁵ Un-Bae Baek (백운배)

¹ First Author Department of Medical IT Marketing, Eulji University, Korea, E-mail:
minho5976@naver.com, jky6873@naver.com

^{2,3,4} Department of Medical IT Marketing, Eulji University, Korea, E-mail:
youheewon6@naver.com, jungyeal91@hanmail.net

⁵ Corresponding Author Professor, Daegu Mirae University, Korea. E-mail: woenb@hanmail.net

Received: May 23, 2017. Revised: June 01, 2017. Accepted: June 17, 2017.

Abstract

This study is the first attempt to raise face recognition ability through machine learning algorithm and apply to CRM's information gathering, analysis and application. In other words, through face recognition of VIP customer in distribution field, we can proceed more prompt and subdivided customized services. The interest in machine learning, which is used to implement artificial intelligence, has increased, and it has become an age to automate it by using machine learning beyond the way that a person directly models an object recognition process. Among them, Deep Learning is evaluated as an advanced technology that shows amazing performance in various fields, and is applied to various fields of image recognition. Face recognition, which is widely used in real life, has been developed to recognize criminals' faces and catch criminals. In this study, two image analysis models, TF-SLIM and Inception-V3, which are likely to be used for criminal face recognition, were selected, analyzed, and implemented. As an evaluation criterion, the image recognition model was evaluated based on the accuracy of the face recognition program which is already being commercialized. In this experiment, it

was evaluated that the recognition accuracy was good when the accuracy of the image classification was more than 90%. A limit of our study which is a way to raise face recognition is left as a further research subjects.

Keywords: : CNN, Machinen Learning, Inception-V3, TF-SLIM, Image Recognition. CRM

1. Introduction

본 연구는 기계학습 알고리즘을 통해 고객얼굴 인식률을 끌어올려 CRM의 정보수집, 분석, 활용에 적용하고자 하는 최초의 시도다. 즉, 유통분야의 VIP 고객 얼굴 인식을 통해 고객등급별로 보다 신속하고 세분화된 맞춤형 서비스를 진행할 수 있다. 수년 전부터 스스로 판단하고 행동하는 지능형 시스템에 대한 관심이 급격하게 증가하고 있다. 구글, 페이스북과 같은, IT 분야의 유명 기업들이 관련 기술 개발을 위해 노력하고 있으며 기계학습을 아이টে으로 하는 스타트업 기업들이 대거 등장하고 있다.

딥러닝은 깊은 신경망(DNN, Deep Neural Networks)을 이용해 학습 데이터로부터 높은 수준의 정보를 추출하고 활용하기 위한 기계학습 알고리즘의 한 종류이다. 딥러닝의 등장은 컴퓨터가 인간의 인식 능력을 능가하게 된 계기로 꼽히며 실제로 딥러닝의 등장을 기점으로 기계학습을 사용하는 시스템의 성능이 대폭 향상되었다. 최근 영상이나 음성신호 인식에서 인간의 능력을 상회하는 성능을 보이는 학습 알고리즘은 대부분 딥러닝을 기반으로한다. 딥러닝은 눈부신 성능 향상을 통해, 연구 패러다임을 바꾸고 있으며, 인공지능 기계학습 분야에서 가장 중요한 방법론이 되고 있다.

과거 방향성 소실 문제 등 이론적 문제와 학습 데이터 수집의 어려움, 계산 성능의 한계 등 현실적 문제로 인해 깊은 신경망을 사용하는데 어려움이 있었다. 그러나 수 년전부터 새로운 학습 알고리즘이 개발되고 빅데이터 및 하드웨어 기술이 발전함에 따라 이론적, 현실적 문제들이 많이 극복되면서 IT의 경계를 넘어 다양한 응용 분야에 까지 널리 적용되어 뛰어난 성능을 보이고 있다.

특히, 컨볼루션 네트워크(CNN, Convolutional Neural Networks)는 ILSVRC(ImageNet Large Scale Visual Recognition Challenge)를 비롯한 여러 영상 인식 컨테스트에서 타 방법론들을 압도하는 성능을 보이며 시각 인식 분야에서 가장 중요한 핵심 방법론이 되었다. 이미지 인식이 중요해짐에 따라 얼마나 정확히 이미지를 인식할 수 있는지 여러 모델이 개발되었는데 그중 CNN 이미지 인식 모델인 Inception-V3와 TF-SLIM을 사용하여 이미지 분류를 하는 연구를 진행하였다. 따라서 본 연구의 궁극적 목적은 오픈소스로 구하기 쉬운 모델을 이용하여 이미지 인식을 하였을 때, 얼마나 정확한 결과를 도출하는 지 알아보고 이 결과를 토대로 유통분야의 서비스 접점에서 VIP 고객 얼굴을 인식하여 신속한 서비스를 수행할 수 있는지 알아 보기 위함이다. 결론적으로 나비나 호랑이에 비해 사람얼굴의 인식률은 기대에 미치지 못했다.

본 연구의 한계인 사람 얼굴 인식률을 끌어올리는 과제는 향후 연구과제로 남겨두었다.

2. 이미지 분석 모델

2.1. CNN(Convolutional Neural Network)

CNN 은 CONV, ReLU, POOL 을 번갈아서 적용하고 마지막에는 Fully connected network 으로 구성된다. Start with an image(width × high × depth)으로 구성되고 마지막 depth 는 색의 수를 의미한다. 즉, 3 이라고 하면 red, green, blue 를 나타내고<Figure 1>의 경우 gray scale 를 나타낸다.

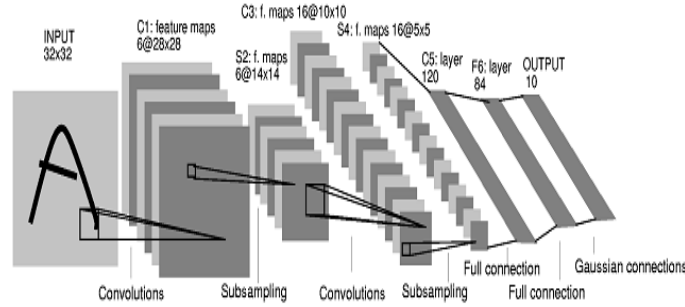


Figure 1. Convolutional neural network (LeCun et al., 1998)

우선 입력된 벡터값에 Convolution 으로 값들을 추출한다. 영상 처리 분야에서 convolution 은 주로 filter 연산에 사용되며, 영상으로부터 특정 feature 들을 추출하기 위한 필터를 구현할 때 convolution 을 사용한다.

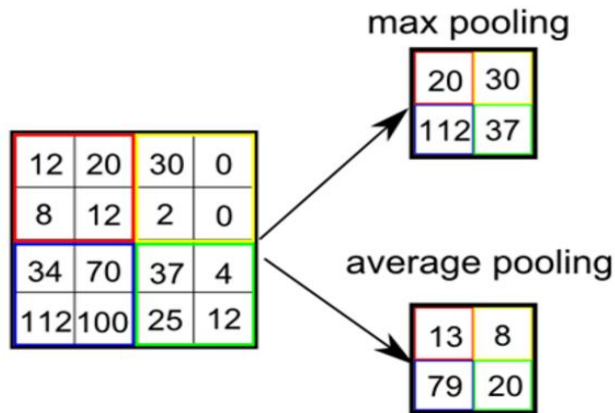


Figure 2. CNN stride

<Figure 2>처럼 전체 이미지에서 노란색 부분이 현재 convolution 이 일어나고 있는 영역이며 stride 값만큼 이동하며 mask 에 대해 연산을 수행하여 오른쪽 결과를 도출한다. 이 때 영역 내에서 가장 큰 값을 선택하는 max-pooling 과 각 영역의 평균을 취하는 average-polling 이 있다.

Convolution 과정 이후 Sub-sampling 과정을 통해 데이터의 크기를 작게 만든다. 이 과정을 여러번 반복해 얻은 특징을 추출하여 최종적으로 fully connected 를 통해 학습 시켜 인식능력을 갖춘다.

2.2. 텐서플로(Tensorflow)

TensorFlow 는 기계 학습과 딥러닝을 위해 구글에서 제작한 오픈소스 라이브러리로서, 데이터 플로우 그래프(Data Flow Graph) 방식을 사용하였다. 데이터 플로우 그래프는 수학 계산과 데이터의 흐름을 노드(Node)와 엣지(Edge)를 사용한 방향 그래프(Directed Graph)로 표현한다. 노드는 수학적 계산, 데이터 입/출력, 그리고 데이터의 읽기/저장 등의 작업을 수행합니다. 엣지는 노드들 간 데이터의 입출력 관계를 나타낸다. 엣지는 동적 사이즈의 다차원 데이터 배열(=텐서)을 실어나르는데, 여기에서 텐서플로우라는 이름이 지어졌다. 텐서(Tensor)는 과학과 공학 등 다양한 분야에서 이전부터 쓰이던 개념이지만, 본 논문에서는 학습 데이터가 저장되는 다차원 배열을 의미한다. TensorFlow 는 원래 기계학습(ML)과 깊은 신경망(deep neural networks,DNN) 연구수행 목적으로 구글의 기계지능 연구조직 내 구글브레인팀의 연구자와 엔지니어들이 개발 하였지만, 다양한 다른 분야에 적용할 수 있을 만큼 그 시스템이 일반적이다.

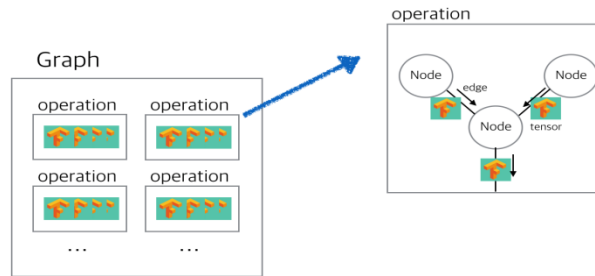


Figure 3. Tensorflow 구성 요소

TensorFlow 의 기본 기능은 값을 입력하여 결과값을 예측하는 것(지도학습)과 값을 입력하여 이 값을 분류하는 것(비지도학습)이다. 지도학습의 예로는 날씨, 요일, 온도, 휴일 여부,등을 입력하여 자전거 대여 숫자를 예상 수입, 주택보유 여부, 연체 여부 등을 판단하여 대출 적합성을 판단하는 것이다. 비지도학습의 예로 시장에서 소비자군을 분류할 때 목표를 주지 않고 컴퓨터에게 알아서 분류하게 하는 것과 알파벳을 무엇인지 알려주지 않고 알아서 분류하도록 해주는 것이다. 즉, 데이터만 많이 있다면 그 속에서 이전에 찾지 못했던 규칙을 찾을 수 있고 이 것을 이용해서 문제를 해결할 수 있는 기술이다.

2.3. TF-SLIM

TF-SILM 이란 이미지 분류, 분할을 위한 tensorflow 라이브러리이다. Google Research 에서 발표한 자료에 따르면 TF-Slim 은 AlexNet, VGG, ResNet, Inception-V3 와 같이 이미지 분류에 사용되는 Deep Learning CNN 모델을 제공하고 있다. 단순히 모델 네트워크 구성만 제공하는 것이 아니라, Datasets(ImageNet, CIFAR10, MNIST, FLOWER), 모델 Training 과 Evaluating 할 수 있는 코드까지 제공하고 있다. 심지어 ImageNet 을 이용한 Pre-trained 모델도 제공하며 새로운 이미지에 대한 Fine Tune 기능도 제공하고 있다.

2.4. Inception-V3

Inception 은 Tensorflow 공개 후 우리가 가장 많이 사용하는 Conv-Net (CNN)모델이다. Inception-V2 은 기존의 Inception-V1 의 기술에 인수분해 개념을 집어 넣어 망을 더욱 깊게 만들고 BN(batch normalization)기술을 반영한 모델을 말하며, 여기에 구조실험을 위해 convolution factorization, label smoothing 및 Auxiliary classifier 에 BN 까지 모든 개념이 적용된 모델이 Inception-V3 이다.

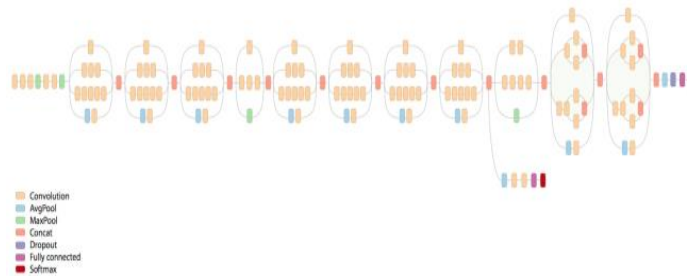


Figure 4. Tensorflow의 Inception-V3 모델

inception v3 은 GoogLeNet 모델에서 사용되었고 2014 년도에 최첨단 이미지 인식 네트워크로써 사용되었다.



3. 이미지 인식 프로그램 성능 비교



본 절에서는 이미지를 입력 받아 이미지의 특징에 대한 분류를 실험하였다. 동일한 PC 와 동일한 버전의 실험도구로 Python, Tensorflow 를 이용하였다. 동일 이미지를 이미지 분석 모델인 Inception-V3 와 TF-SLIM 을 통해 비교 분석하여 분류 정확도를 측정하였다.


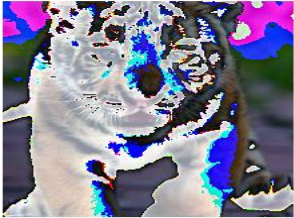
3.1. 이미지 분석 결과

아래 <Table 1> 은 각 그림 앞에 있는 숫자에 해당하는 인식결과를 보여준다.

Table 1. Comparison of TF-SLIM and Inception-V3

TF-Slim	Inception-V3
<p data-bbox="305 331 448 352">Downloaded image</p>  <p data-bbox="220 615 639 636">Resized, Cropped and Mean-Centered input to network</p>  <p data-bbox="212 890 789 978"> Probability 0.64 => [monarch, monarch butterfly, milkweed butterfly, Danaus plexippus] Probability 0.27 => [ringlet, ringlet butterfly] Probability 0.06 => [lycaenid, lycaenid butterfly] Probability 0.03 => [sulphur butterfly, sulfur butterfly] Probability 0.00 => [lacewing, lacewing fly] </p>	<p data-bbox="805 575 1446 720"> ringlet, ringlet butterfly (score = 0.17076) monarch, monarch butterfly, milkweed butterfly, Danaus plexippus (score = 0.15572) lycaenid, lycaenid butterfly (score = 0.12823) admiral (score = 0.04628) lacewing, lacewing fly (score = 0.04339) </p>

TF-Slim	Inception-V3
<p data-bbox="256 1178 461 1199">Downloaded image</p>  <p data-bbox="220 1430 808 1451">Resized, Cropped and Mean-Centered input to network</p>  <p data-bbox="212 1680 675 1757"> Probability 0.20 => [groom, bridegroom] Probability 0.08 => [suit, suit of clothes] Probability 0.05 => [seat belt, seatbelt] Probability 0.04 => [gown] Probability 0.03 => [notebook, notebook computer] </p>	<p data-bbox="846 1383 1338 1539"> gown (score = 0.18808) kimono (score = 0.05965) maillot (score = 0.05869) bikini, two-piece (score = 0.05299) torch (score = 0.03544) </p>

TF-Slim	Inception-V3
<p data-bbox="256 279 461 298">Downloaded image</p>  <p data-bbox="209 562 792 581">Resized, Cropped and Mean-Centered input to network</p>  <p data-bbox="201 842 797 926"> Probability 0.67 => [tiger, Panthera tigris] Probability 0.25 => [tiger cat] Probability 0.02 => [jaguar, panther, Panthera onca, Felis onca] Probability 0.01 => [leopard, Panthera pardus] Probability 0.01 => [tabby, tabby cat] </p>	<p data-bbox="829 543 1414 659"> tiger, Panthera tigris (score = 0.74040) tiger cat (score = 0.15136) jaguar, panther, Panthera onca, Felis onca (score = 0.00433) leopard, Panthera pardus (score = 0.00232) snow leopard, ounce, Panthera uncia (score = 0.00075) </p>

첫 번째 나비 이미지를 TF-SLIM 으로 분류하였을 때는 64 퍼센트 확률로 monarch butterfly 로 분류했고, Inception-V3 로 분류했을 때는 ringlet butterfly, monarch butterfly, lyceanid butterfly 를 각각 17, 15, 13 퍼센트의 비슷한 수치로 분류하였다. 이는 TF-SLIM 가 Inception-V3 에 비해 더욱 정확한 분류를 했다고 볼 수 있다. 실제 이미지로도 monarch butterfly 에 분류된 것이 정확하다는 것을 알 수 있다.

두 번째 유명 연예인 사진을 분류하였을 때 연예인 본인으로 분류될 것을 예상했지만 두 가지 방법 모두 사람으로 분류하지 못하고 주변환경 즉, 악세사리나 옷 또는 남자자체의 특성으로 분류를 하였다. TR-SLIM 의 경우 신랑으로 분류했고, Inception-V3 는 옷으로 분류하였다. 이 결과를 보아 사람의 외모는 특징이 너무 많고 특징 자체가 불분명하다고 추측하게 되었다. 따라서 두 모델은 인물 인식 기술에는 적합하지 않는 것으로 보인다.

세 번째 새끼 호랑이 사진은 두 가지 방식 모두 대략 70 퍼센트로 호랑이로 분류하였다. 이는 호랑이와 같이 특징이 분명한 이미지의 경우에는 두 모델 모두 정확한 판단을 한다는 것을 알 수 있다. 3 가지 분석을 통하여 TF-SLIM 과 Inception-V3 는 특징이 분명하고 표본이 많은 이미지를 추출하기 적합한 모델임을 확인할 수 있다.

4. Conclusion

딥러닝은 철저히 학습 데이터에 기반하여 최적의 판단이나 예측 값을 찾아낼 뿐이므로, 융통성과 유연성이 부족하다. 따라서 학습에서 접하지 못했던 전혀 새로운 입력을 접하는 경우 치명적인 오류를

유발할 수 있다. 따라서 딥러닝은 학습 데이터 확보에 많은 시간과 비용이 들고, 학습데이터를 확보하더라도 학습 데이터의 범위를 벗어나는 패턴에 대해서는 제대로 해석하지 못한다.

본 실험을 통하여 Inception-V3 보다 TF-Slim 의 성능이 뛰어남을 확인할 수 있었지만, 분류 알고리즘에 따라 이미지 분류 정확도에 차이가 있음을 확인할 수 있었다. 실험에 사용한 동물, 사물 등 표본데이터가 많고 특징이 분명한 사진 즉, 학습데이터가 확보된 사진은 인식이 비교적 잘되어 분류되었지만, 표본 데이터가 적은 사람의 경우 특징이 너무 많고 불분명하기 때문에 얼굴이 아닌 악세서리나 옷과 같은 주변 환경을 인식하게 되었다. 두 모델의 경우 가장 높은 정확도 값이 TF-SLIM 이 67% 그리고 Inception-V3 이 74%를 보여주어 저조한 정확도 값을 보이기 때문에 두 이미지 모델은 인물 인식 프로그램 제작에 적합하지 않지만 다른 학습데이터가 정확한 이미지를 분석하기에 적합하다고 보여진다. 본 연구의 한계인 사람 얼굴 인식률을 끌어올리는 과제는 미래 연구과제로 남겨둔다.

References

- Ahn, S. M. (2017). Deep Learning Model and Its Applications. Kookmin University, College of Business Administration
- Ahn, M. H., Lee, D. S., & Lee, J. T. (1995). Developed algorithm trading system for machine learning. Proceedings of the Korea Simulation Society.
- Kensorflowkorea (2017). https://tensorflowkorea.gitbooks.io/tensorflow-kr/content/g3doc/tutorials/image_recognition/
- Kim, I. J. (2015). Recent Development Trends of Deep Learning Technology for Visual Recognition. *Information Science Society*, 32, 15-20.
- Lyu, B. I., Hwang, Y. J., Han, S. J., Lee, S. M., Kim, J. B., & Han, J. J. (2015). Trend of Deep Learning Based Image Recognition Near Human Level. *Information Science Society*, 32, 32-41.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE 86 (11), 2278-2324.
- Moon, S. E., Park, J. K., Park, Y. K., On, H. Y., & Kang, D. J. (2015). Object Recognition Technique for Image Using Deep Learning. *Journal of Control, Robotics and Systems*, 13, 21-26.
- Pythonkim (2017). <http://pythonkim.tistory.com/53>