

Deep Residual Network와 Distributed Learning을 이용한 신발 상품 분류

Classification of Shoes Using Deep Residual Network with distributed learning

요 약

SNS와 인터넷에 게시되는 신발 이미지에 대해서 자세한 정보(상품명, 브랜드, 이와 유사한 신발 상품)를 얻고자 하는 요구가 있을 때 이에 대한 객관적인 솔루션을 Neural Network 소프트웨어로 개발하여 제시하고자 한다. 이미지 Classification 알고리즘 중에서 성능이 높다고 알려진 Deep Residual Network(이하 ResNet)와 분류해야 되는 Class가 많을 때 효과적인 Label Mapping을 이용하여 Neural Network Classifier를 구현하게 될 것이다.

1. 서론

SNS와 인터넷에 게시되는 신발 사진은 대부분 판매용이 아닌 이상 그 신발의 브랜드, 정확한 상품명 등 신발에 대한 자세한 정보를 기재하지 않은 채 게시된다. 특정 신발을 구매하고 싶거나 자세한 정보를 찾아보려는 요구가 있을 때 이를 충족시킬 수 있는 객관적인 솔루션이 존재하지 않는다.

이에 대해 본 연구에서는 신발 이미지에 대해서 신발 상품명, 브랜드, 유사한 신발 상품들을 추천해주는 Neural Network Software를 솔루션으로 제공하고자 한다.

신발 상품명을 예측하는 Neural Network Classifier의 정확도를 높이기 위해서, ResNet(Residual Neural Network)를 사용할 것이다.

분류 대상이 신발의 브랜드가 아닌 신발의 상품명이기 때문에, 분류해야 되는 Class의 개수가 많다. 이러한 Large-scale classification 문제를 단일의 ResNet 구조로는 해결하지 못한다. 따라서 본 연구에서는 Label Mapping concept에서 비롯된 distributed learning를 도입하여 수많은 신발 상품에 대해서 각각 분류하고자 한다.

2. 관련 연구

2-1. Deep Residual Network

AlexNet, VGG Model, Inception Modules 등의 CNN(Convolutional Neural Network)모델에서, Convolution Layer의 깊이가 깊어질 수록 Vanishing 혹은 Exploding Gradient 문제가 발생하여 학습을 시키기가 점점 어려워졌다. 즉, 파라미터 update 시에, gradient 값이 너무 큰 값이나 작은 값으로 포화되어 더이상 움직이지 않아 학습의 효과가 없어지거나 학습 속도가 느려지는 문제이다. 이 문제를 해결하기 위해서 Batch Normalization, 파라미터의 초기값 설정 방법 개선 등의 기법이 적용되었지만, layer 개수가 일정 수를 넘어서면 학습의 결과가 좋지 않았다.

따라서 ResNet에서는 Layer를 100개 이상 깊게 하면서, <그림 1>과 같은 Residual Learning Block을 도입하였다. 입력에서 바로 출력으로 연결되는 shortcut 연결이 생기게 되었으며, 이 shortcut은 파라미터가 없이 바로 연결되는 구조이기 때문에 연산량 관점에서는 덧셈이 추가되는 것외에는 차이가 없다. 학습을 진행하면서 $F(x)$ 는 0이 되는 방향으로 파라미터를 조금씩 수정하게 되고, 입력의 작은 움직임을 쉽게 검출할 수 있게된다. 이를 통해서 깊은 layer를 가진 CNN 구조에서도 쉽게 최적화가 가능하고, 정확도 또한 개선할 수 있게 된다.

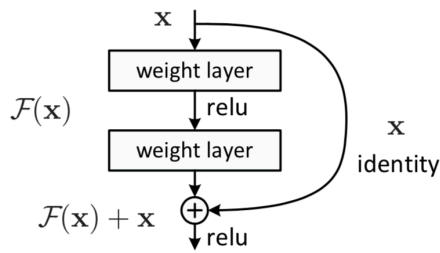


그림 1. Residual Block의 구조

2-2. Large Scale classification in Deep Neural Network with Label Mapping

Image Classification 문제에서 분류해야 하는 class의 개수가 massive할 때, 기존의 deep neural network의 구조로는 해결할 수 없다. 이러한 massive class들을 적절한 함수를 통해 분할하여 여러 개의 base learner network에게 할당한 다음, 각 network 별로 적은 수의 class들을 예측하는 것이 Label Mapping 방법이다. 즉, N 개의 Class를 분류 해야하는 큰 classification 문제를 여러개의 작은 classification 문제로 나눈 것이다.

<그림 3>에서 Label Mapping에서는 n개의 network에서 각 model은 input에 대해서 함수를 통해 나뉘어진 class의 확률을 softmax classifier를 통해 예측하게 된다.

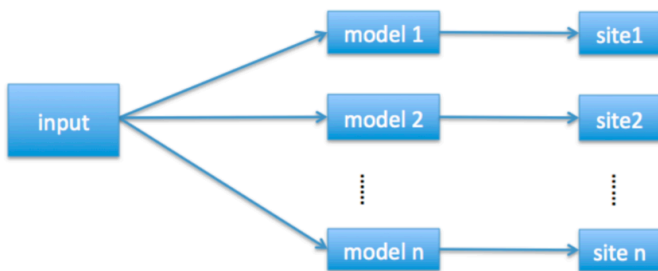


그림 2. Label Mapping Classification Neural Network 구조

3. 제안 Model

Label Mapping Classification Neural Network 구조에서는 Input data는 split하지 않고, 예측해야 하는 class를 함수를 통해 split하게 된다. 이렇게 split된 class들을 예측하는 network를 여러개 두어 학습을 진행하게 된다.

본 연구에서는 Large scale의 신발 이미지에 대해서 classification하는 학습 모델을 <그림 3>와 같이 구성한다. 앞서

말한 방법과 다르게 모든 class에 대해서 label mapping을 먼저 진행하지 않고, input data를 신발 브랜드에 따라 split하고 그에 따른 ResNet model을 여러개 두어 각 network model은 해당 브랜드의 신발 상품 class에 대해서만 classification하도록 distributed network를 구성할 것이다.

테스트를 진행할 때에는, 어떤 신발 이미지를 각각의 ResNet에 input으로 주게 되고, 각 ResNet이 가장 높은 확률로 예측된 class(상품명) 정보를 사용자는 최종적으로 받게 된다.

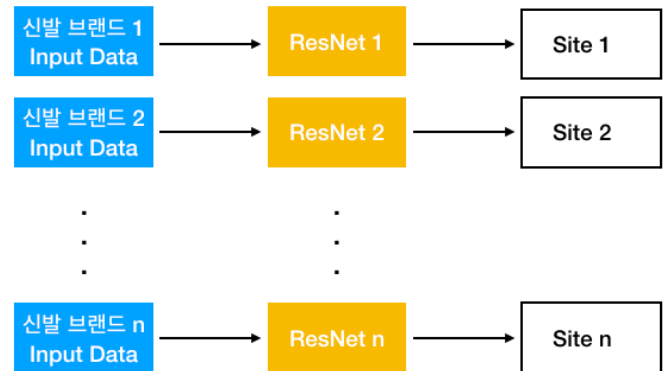


그림 3. Distributed Residual Network

4. 학습 데이터 셋 구성

브랜드 별로 500 X 600 크기의 다양한 각도에서 찍은 신발 이미지를 쇼핑몰 웹사이트로부터 크롤링 하여 총 8400개 정도의 이미지를 확보한다.



그림 4. 학습을 위한 신발 이미지의 예

5. 결론 및 향후 계획

본 연구에서는 distributed learning과 Deep Residual Network를 이용한 신발 상품 이미지 분류를 진행한다.

Distribute 된 Residual Network 각각이 split된 데이터를 가지고 학습하기 때문에, 각 network는 학습 할 수 있는 데이터가 적다. 따라서 classification하는 정확성을 높이기 위해서는 신발 상품 이미지를 최대한 많이 확보하고, 똑같은 신발 상품에 대해서 다양한 각도에서 찍은 많은 이미지를 확보하는 것이 본 연구의 중요한 지표가 될 것이다.

참 고 문 헌

- (1) Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", 2014.
- (2) Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jean Sun, "Deep Residual Learning for Image Recognition", 2015.
- (3) Qizhi Zhang, Kuang-Chih Lee, Hongying Bao, Yuan You, Wenjie Li, Dongbai Guo, "Large scale classification in deep neural network with Label Mapping", 2018.