

다중 클래스의 이미지 장면 분류

신성윤^{1*} · 이현창² · 신광성² · 김형진³ · 이재완¹

¹군산대학교 · ²원광대학교 · ³전북대학교

Image Scene Classification of Multiclass

Seong-Yoon Shin^{1*} · Hyun-Chang Lee² · Kwang-Seong Shin² · Hyung-Jin Kim³ · Jae-Wan Lee²

¹Kunsan Natl. University · ²Wonkwang University · ³Jeonbuk Natl. University

E-mail : s3397220@kunsan.ac.kr / hclglory@wku.ac.kr / waver0920@wku.ac.kr /

kim@jbnu.ac.kr / jwlee@kunsan.ac.kr

요 약

본 논문에서는 변환 학습에 기반을 둔 다중 클래스 영상 장면 분류 방법을 제시한다. ImageNet 대형 이미지 데이터 세트에서 사전 훈련된 네트워크 모델에 의존하여 다중 클래스의 자연 장면 이미지를 분류한다. 실험에서는 최적화된 ResNet 모델을 Kaggle의 Intel Image Classification 데이터 셋에 분류하여 우수한 결과를 얻었다.

ABSTRACT

In this paper, we present a multi-class image scene classification method based on transformation learning. ImageNet classifies multiple classes of natural scene images by relying on pre-trained network models on large image datasets. In the experiment, we obtained excellent results by classifying the optimized ResNet model on Kaggle's Intel Image Classification data set.

키워드

multi-class image scene classification method, transformation learning, image datasets, ResNet

I. 서 론

최근 몇 년 동안 그래픽 이미지 처리 알고리즘의 연구와 함께 컴퓨터 비전 기술의 급속한 발전이 촉진되었으며 그 중 이미지 인식 및 분류 기술이 가장 두드러졌다. 이미지 인식 및 분류에서는 이미지를 입력하면 이미지 내용의 분류에 대한 설명이 출력되는데, 이는 컴퓨터 비전 연구의 핵심 과제이다. 이미지 처리 알고리즘은 객체 감지, 의미론적 분할 등에 널리 사용된다. 이미지 분류의 전통적인 방법은 특징 설명 및 탐지[1, 2, 3]이다.

CNN 연구에서 이론상 CNN의 깊이가 깊을수록 분류 정확도는 높아지지만 시간이 더 많이 걸리고 컴퓨터 컴퓨팅 성능이 더 많이 필요하다. 따라서 분류 및 인식 작업이 유사하고 데이터 세트의 크

기가 유사한 경우 사전 훈련된 모델은 변환 학습 방법을 통해 이러한 문제를 더 잘 해결할 수 있다.

II. ResNet Model

2015 ImageNet Image Classification, Image Object Orientation 및 Image Object Detection Algorithm Competition에서 ResNet 모델이 우승을 차지했습니다. 경쟁에서 컨볼루션 신경망을 훈련할 때 ResNet 모델은 잔차 학습 방법을 사용하여 더 깊은 네트워크 차원 축소 및 향상된 분류 정확도에 성공했습니다. 구체적인 방법은 CNN을 기반으로 하는 잔여 모듈을 추가하는 것입니다. 잔차 모듈의 구조는 그림 1에 나와 있습니다. 잔차 모듈에서 x 는 입력이고 $F(x)$ 는 매핑 함수입니다. 각 잔차 모듈은 두 개의 경로를 포함하며, 그 중 하나는 입

* corresponding author

력 기능의 직접 통신 경로이고, 다른 경로는 기능의 잔차를 얻기 위해 기능에 대해 2-3 컨볼루션 연산을 수행하고, 마지막으로 해당 기능에 대한 기능을 수행합니다. 두 개의 경로가 함께 추가됩니다. 그림 2는 ResNet34 및 ResNet50/101/152 모델의 잔여 모듈을 보여줍니다.

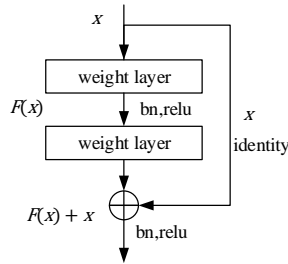


그림 1. 잔여 학습: 빌딩 블록

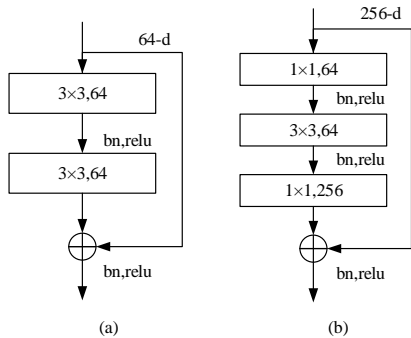


그림 2. 잔여 학습 모듈의 적용. (NS).

동일한 수의 출력 채널을 갖는 2개의 3x3 컨볼루션으로 구성된 기본 모듈 연결 방법. (NS). 병목 모듈 연결.

III. 실험

비교 실험을 사용하여 인텔 이미지 분류 데이터 세트에 대해 실험을 수행했습니다. 먼저 CNN 모델에 대한 훈련과 테스트를 진행한 후 ResNet50 모델에 대한 훈련과 테스트를 진행하였다. 실험에서 training sample은 batch size가 32, training step size가 batch당 439, epoch가 35였다. ResNet50 모델의 학습에서는 최적의 모델 검사 알고리즘을 사용하여 모델을 판단하였다. 테스트 손실이 더 이상 떨어지지 않으면 훈련이 종료됩니다.



그림 3. ResNet50 모델에 의해 잘못된 레이블이 지정된 이미지의 몇 가지 예

IV. 결론

본 논문은 변환 학습을 기반으로 하는 다중 클래스 이미지 분류 방법을 제시하였다. 이미지를 분류하기 위해 top layer를 제거한 pre-trained ResNet50 모델을 사용하여 Intel Image Classification 데이터 셋의 CNN 모델과 비교하는 방법을 제시하였다.

Acknowledgement

"This research is partially supported by Institute of Information and Telecommunication Technology of KNU."

References

- [1] Begüm Demir, Sarp Ertürk, "Improving SVM classification accuracy using a hierarchical approach for hyperspectral images," 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 2849-2852, 2009.
- [2] Xiangrong Zhang, Peng Weng, Jie Feng, Erlei Zhang, Biao Hou, "Spatial-spectral classification based on group sparse coding for hyperspectral image," 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS, pp. 1745-1748, 2013.
- [3] W. Wang, I. Pollak, T.-S. Wong, C.A. Bouman, M.P. Harper, J.M. Siskind, "Hierarchical Stochastic Image Grammars for Classification and Segmentation," IEEE Transactions on Image Processing, vol. 15, pp. 3033-3052, 2006.