

학습을 이용한 영상 분류 방법

신성윤^o, 이현창*, 신광성*

^o군산대학교 컴퓨터정보통신공학부,

*원광대학교 디지털콘텐츠공학과

e-mail: s3397220@kunsan.ac.kr^o, {hclglory, waver0920}@wku.ac.kr*

Image Classification Method Using Learning

Seong-Yoon Shin^o, Hyun-Chang Lee*, Kwang-Seong Shin*

^oSchool of Computer Inf. & Communication Eng., Kunsan National University,

*Dept. of Digital Contents Eng., Wonkwang University

● 요약 ●

본 논문에서는 변환 학습을 기반으로 한 다중 클래스 이미지 장면 분류 방법을 제안하도록 한다. ImageNet 대형 이미지 데이터 세트에서 사전 훈련 된 네트워크 모델을 사용하여 다중 클래스의 자연 장면 이미지를 분류하였다. 실험에서 최적화 된 ResNet 모델은 Kaggle의 Intel Image Classification 데이터 세트에 분류되어 우수한 결과를 얻었다.

키워드: 변환학습(transform learning), ImageNet, ResNet

I. Introduction

본 논문에서는 변환 학습을 기반으로 한 다중 클래스 이미지 장면 분류 방법을 제안한다. 이미지 분류를 위해 큰 이미지 데이터 세트 ImageNet에 대해 사전 학습 한 ResNet (ResNet) 모델[1,2]을 사용하는 방법이다. CNN 모델의 이미지 분류 방법에 비해 분류 정확도와 효율성을 크게 향상시킬 수 있다. 실험에서 이미지 분류를 위해 ResNet 모델을 사용하고 ResNet 모델의 사전 훈련 된 기준선을 채택했다. 모델 기중치는 imagenet에서 사용되며 완전히 연결된 모델 계층이 제거되어 모델 학습 속도가 빨라진다. 모델에서 컨벌루션 계층의 데이터를 고차원에서 저 차원으로 줄인다.

II. Suggested Methoded

이 섹션에서는 먼저 변환 학습의 기본 이론을 간략히 소개 한 다음 ResNet 모델의 기본 구조를 자세히 소개한다. 그런 다음 모델의 특성을 분석하고 이미지 분류의 작업 특성을 결합하여 모델 구조를 최적화한다. 데이터 전처리는 데이터 확장 및 정규화의 두 가지 측면에서 수행되어 모델 학습의 신속한 수렴을 달성하고 학습 결과의 과적합을 억제한다.

1.1 데이터 최적화

CNN 훈련에서 일반적인 접근 방식은 입력 데이터 전처리를 통해 훈련 결과의 과적합을 제거하는 것이다. 일반적으로 사용되는 데이터 전처리 방법은 데이터 증대 기술과 정규화이다. 데이터 증강 기술은 훈련 세트 샘플이 작은 경우 이미지 뒤집기, 자르기, 패닝 및 크기 조정과 같은 일련의 작업으로 훈련 세트 수를 곱한다. 실험에 사용 된 데이터 증가 매개 변수는 표 1에 나와 있다.

데이터 최적화의 또 다른 방법은 입력 데이터의 정규화이다. 다중 클래스 CNN 모델에서도 이미지의 최종 분류 결과는 이진 분류, 즉 포지티브 또는 네거티브 클래스로 처리 될 수 있다. 입력 데이터가 디지털화되면 샘플의 특성에 해당하는 행렬 요소가 크거나 작으며 간격이 매우 크다. 이러한 데이터가 학습 및 예측에 사용되면 과적합으로 이어지고 분류 결과에 영향을 미친다. 원본 데이터를 정규화 한 후 각 지표는 동일한 크기로 종합적인 비교 및 평가에 적합하다. 이 방법의 장점은 정규화로 인해 기술기가 빨라져 최적의 솔루션을 찾을 수 있다는 것이다. 반면에 정규화는 정확도를 향상시킬 수 있다. 실험에서는 Z-score 표준화 방법을 사용하고 처리 된 데이터는 표준 정규 분포를 따랐다. 즉, 평균값은 0이고 표준 편차는 1이었다.

Table 1. Data augmentation parameters

No.	Parameter	Augmentation
1	Rotation_range(°)	30
2	Width_shift_range(%)	20
3	Height_shift_range(%)	20
4	Shear_range(%)	20
5	Zoom_range(%)	20
6	Horizontal_flip	TRUE
7	Fill_mode	Nearest

III. Experiment

비교 실험을 사용하여 인텔 이미지 분류 데이터 세트에 대한 실험을 수행했다. 먼저 CNN 모델의 학습 및 테스트를 수행 한 다음 ResNet50 모델의 학습 및 테스트를 수행했다. 실험에서 학습 샘플의 배치 크기는 32, 배치 당 학습 단계 크기는 439, epoch는 35이다. ResNet50 모델의 학습에서는 모델을 판단하기 위해 최상의 모델 검사 알고리즘을 사용한다. 테스트 손실이 더 이상 떨어지지 않으면 교육이 종료된다. 섹션 III에 따르면 학습률은 1e-02로 설정되고 학습률 감소는 ResNet50 모델에서 1e-6으로 설정되며 운동량의 가중치는 0.9이다. 실험 출력 결과는 모델 학습 및 손실률 및 정확도 곡선 테스트, 예측 결과의 혼동 행렬이다. CNN 모델 테스트의 손실률은 0.4570, 정확도는 0.8487, 손실률과 정확도 곡선은 그림 1에 나와 있다. ResNet50 모델 테스트의 손실률은 0.2177, 정확도는 0.9233이다. 손실률과 정확도 곡선은 그림 2에 나와 있다. CNN 모델의 테스트 결과와 비교하여 ResNet50 모델의 손실률은 52.36 % 감소하고 정확도는 8.79 % 증가했다. 모델의 혼동 행렬은 그림 3에 나와 있다.

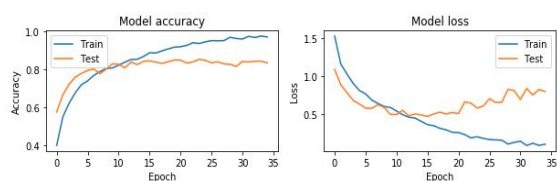


Fig. 1. Loss rate and accuracy curve of CNN model

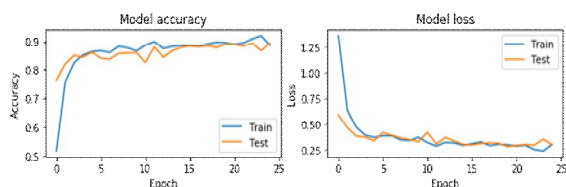


Fig. 2. Loss rate and accuracy curve of ResNet50 model

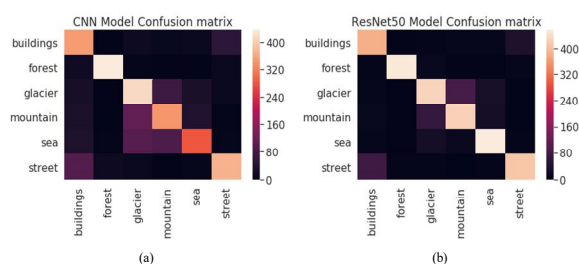


Fig. 3. Confusion matrix, (a), CNN model confusion, (b), ResNet50 model confusion matrix.

IV. Conclusions

이 논문은 변환 학습을 기반으로 한 다중 클래스 이미지 분류 방법을 제시했다. 이미지를 분류하기 위해 최상위 계층이 제거 된 사전 학습 된 ResNet50 모델을 사용하여 Intel Image Classification 데이터 세트의 CNN 모델과 비교하는 방법이다. 실험에서는 변환 학습 방법을 통해 ResNet50에 대한 매개 변수 튜닝 및 데이터 최적화를 수행했다. 실험을 통해 모델의 손실률과 정확도 곡선, 혼동 행렬, 예측 결과 그래프, 예측 오차 그래프를 얻었다. ResNet50의 손실률과 정확도는 각각 0.2177과 0.9233이었다.

REFERENCES

- [1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition," Computer Vision and Pattern Recognition, pp.1-12, 2015. arXiv: 1512.03385 [cs.CV]
- [2] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, Kaiming He, "Aggregated Residual Transformations for Deep Neural Networks," Computer Vision and Pattern Recognition, pp.1-10, 2016. arXiv: 1611.05431 [cs.CV].