

동물 이미지 패치 분류를 위한 향상된 딥 러닝

신성윤^o, 이현창^{*}, 신광성^{*}

^o군산대학교 컴퓨터정보통신공학부,

^{*}원광대학교 디지털콘텐츠공학과

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

Enhanced Deep Learning for Animal Image Patch Classification

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 동물 이미지 분류를 위한 작은 데이터 세트를 기반으로 하는 향상된 딥 러닝 방법을 제안한다. 먼저, CNN을 사용하여 작은 데이터 세트에 대한 훈련 모델을 구축한다. 데이터 증대를 사용하여 훈련 세트의 데이터 샘플을 확장한다. 다음으로, VGG16과 같은 대규모 데이터 세트에서 사전 훈련된 네트워크를 사용하여 작은 데이터 세트의 병목 현상 기능을 추출한다. 그리하여 두 개의 NumPy 파일에 새로운 훈련 데이터 세트 및 테스트 데이터 세트로 저장한다. 마지막으로 완전히 연결된 네트워크를 훈련시킨다.

키워드: deep learning, convolutional neural network, VGG16

I. Introduction

최근 몇 년 동안 CNN을 사용하여 두 범주 분류 문제를 해결하는 딥 러닝 CNN의 연구 개발은 더 나은 결과를 얻을 수 있었다. 본 논문에서는 Kaggle Dogs vs Cats 데이터셋[1]을 예로 들어 개선된 CNN 예측 방법을 제안한다. 먼저 CNN을 사용하여 작은 데이터 세트를 기반으로 훈련 모델을 구축하고 데이터 증대 기법을 사용하여 훈련 세트의 데이터 샘플 수를 개선하여 훈련 결과의 과적합을 억제한다. 둘째, VGG16과 같은 대규모 데이터 세트에서 사전 훈련된 네트워크를 사용하여 작은 데이터 세트의 병목 현상 기능을 추출하여 두 개의 NumPy 파일에 새로운 훈련 데이터 세트 및 테스트 데이터 세트로 저장한다. 마지막으로, 완전히 연결된 네트워크가 새 데이터 세트로 활성화한다. 실험 데이터 세트를 기반으로 실험 결과를 개선된 방법과 비교한다.

컨볼루션 계층으로 전송된다. 그런 다음 풀링 계층은 획득한 이미지를 분할하고 주요 특징 정보를 유지하려고 시도한다. 다음으로 추출된 특징 정보에 가중치를 부여하고 완전 연결 계층에서 연결한다. 마지막으로 분류에 해당하는 출력 뉴런 이미지가 출력 레이어에 출력된다. CNN의 다섯 부분에서 input layer, convolutional layer, pooling layer는 실제로 영상의 특징 정보를 추출하는 과정이다. 후속 완전 연결 계층과 출력 계층은 이미지 분류 작업을 수행하는 데 사용된다. CNN을 기반으로 VGG16 및 VGG19와 같은 신경망 모델이 파생된다. VGG16 모델은 그림 2와 같이 5개의 conv 블록과 1개의 완전 연결 계층으로 구성된 CNN 모델이다. 다섯 번째 컨볼루션 블록의 가장 큰 풀링 계층에서 입력 데이터의 병목 특성을 추출한다. 마지막 계층은 완전 연결 계층이다.

II. Convolutional Neural Networks

기본 컨볼루션 신경망 모델은 그림 1과 같이 입력 레이어, 컨볼루션 레이어, 풀링 레이어, 전체 연결 및 출력 레이어의 다섯 부분으로 구성된다. 입력 레이어는 입력 이미지를 균일한 크기 형식으로 처리하는 역할을 한다. 처리된 이미지는 이미지의 중요한 특징을 추출하기 위해 공유 가중치를 사용하여 피쳐 학습 및 컨볼루션 처리를 위해

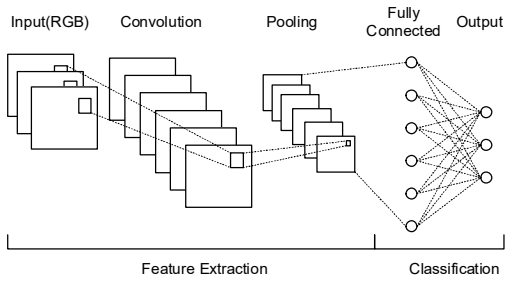


Fig. 1. Convolutional block structure of a CNN

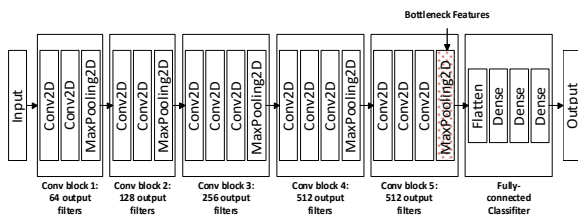


Fig. 2. VGG16 CNN structure diagram

III. Data Augmentation

데이터 증대는 과적합을 효과적으로 억제할 수 있는 방법이다. 이 방법은 원본 이미지 값의 특정 범위 내에서 회전, 크기 조정, 이동, 미러링 등의 작업을 수행하여 원본 이미지 데이터 특성을 가진 새 샘플을 생성할 수 있으므로 데이터 샘플 수를 늘리는 목적을 달성한다. 훈련 세트 데이터 증강 방법은 보편적으로 적용할 수 있다. 데이터 증가는 그림 3에 나와 있으며 이미지 크기는 150 × 150 픽셀로 균일하게 처리된다.

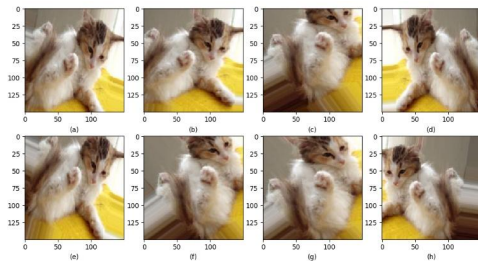


Fig. 3. Data augmentation example: (a) original image, (b) horizontal offset, (c) vertical offset after rotation, (d) mirror image, (e) cropping, (f) vertical offset, (g) rotation, (h) mirroring and scaling.

IV. EXPERIMENT

이 연구의 모든 실험은 Kaggle Dogs vs Cats 데이터 세트로 수행된다. CNN 개선 전후에 실험을 수행하였다. 32 이미지의 CNN 배치 크기가 사용되었으며 반복당 100단계 및 총 30에포크가 있는 RMSProp 옵티마이저가 사용되었으며 테스트의 손실률과 최상의 정확도는 각각 0.4423 및 0.8094이다. 훈련 및 예측 총 시간은 3274.3초이다.

초입니다. 향상된 CNN은 16의 반복 배치 크기로 구성되고 각 epoch는 50 epoch에서 64 반복이다. 얻은 테스트의 손실률과 최고의 정확도는 각각 1.3210과 0.9175이다. 또한 데이터 전처리 시간은 333.2초이다. 훈련 및 예측 시간은 253.5초이다. 손실률 및 정확도 통계는 그림 3에 나와 있다.

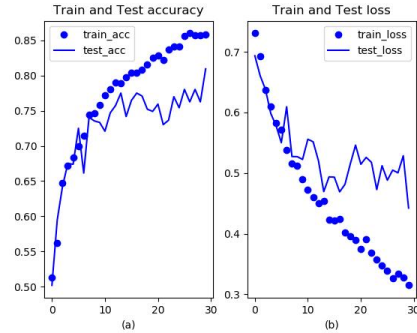


Fig. 5. CNN experiment results: (a) Accuracy statistics for training and testing, (b) Loss rate statistics for training and testing.

V. Conclusions

본 논문은 동물 이미지 분류를 위한 작은 데이터 세트를 기반으로 개선된 딥 러닝 방법을 제시했습니다. Kaggle Dogs vs Cats 데이터 세트에서 두 가지 범주의 실험이 수행되었습니다. Kaggle Dogs vs Cat 데이터 세트는 레이블이 있는 개와 고양이 사진에 대한 25,000개의 훈련 세트와 레이블이 없는 개와 고양이 사진에 대한 12500개의 테스트 세트로 구성됩니다.

REFERENCES

[1] Kaggle, Competition of creation an algorithm to distinguish dogs from cats homepage, [Online]. 2013 Available: <https://www.kaggle.com/c/dogs-vs-cats>