패션 의류 영상 분류 방법

고톱수렌 이칭허를러[O], 신성윤[*], 이현창[**]
[O]군산대학교 컴퓨터정보통신공학부,
[*]군산대학교 컴퓨터정보통신공학부,
[**]원광대학교 디지털컴퓨터공학과
e-mail: ichkog108@gmail.com[O], s3397220@kunsan.ac.kr[*], hclglory@wku.ac.kr[**]

A Method for Fashion Clothing Image Classification

Gotovsuren Ichinkhorloo[O], Seong-Yoon Shin[*], Hyun-Chang Lee[**]
[O]School of Computer Inf. & Communication Eng., Kunsan National University,
[*]School of Computer Inf. & Communication Eng., Kunsan National University,
[**]Dept. of Digital Contents Eng, Wonkwang University

● 요 약 ●

우리는 패션 의류 이미지의 빠르고 정확한 분류를 달성하기 위해 최적화 된 동적 감쇠 학습률과 개선 된 모델 구조를 갖춘 딥 러닝 모델을 기반으로 하는 새로운 방법을 제안했습니다. 우리는 Fashion-MNIST 데이터 셋에 대해 제안 된 모델을 사용하여 실험을 수행하고 이를 CNN, LeNet, LSTM 및 BiLSTM의 방법과 비교했습니다.

키워드: 패션 의류 영상(fashion clothing image), 딥 러닝(deep learning), 학습률(learning rate)

I. Introduction

최근 컴퓨터 기술의 급속한 발전으로 인공 지능 기술에 대한 연구가 획기적인 발전을 이루었습니다. 그중에서도 머신 러닝으로 대표되는 인공 지능 알고리즘의 연구와 응용이 급속한 발전을 이루었다 [1].

II. Methodology

1. Artificial Neural Networks

캐스케이드를 통해 입력에서 출력으로 완전히 연결된 퍼셉트론인 인공 신경망 (ANN)을 기반으로 하는 DL 모델 Fig. 1은 ANN 구조를 보여준다. 인공 뉴런의 개념은 McCulloch와 Pitts에 의해 처음 제안되었고 1958 년에 perceptron의 개념으로 발전했다.

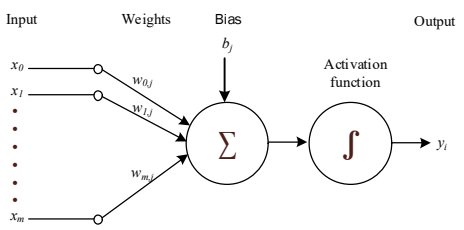

Fig. 1. The unit of Artificial Neural Network.

2. Convolutional Neural Network

CNN 또는 ConvNet이라고하는 Convolutional Neural Network는 컴퓨터 비전 및 음성 인식과 같은 많은 작업에 널리 사용된다. 2012 ImageNet Challenge에서 처음으로 깊은 CNN을 적용하여 인상적인 결과를 달성했다.

III. Proposed Method

1. Learning Rate Optimized Method

학습률은 신경망 훈련 및 학습 능력이며, 신경망에서 조정하기 위해 가장 까다로운 글로벌 매개 변수이다. 학습률이 너무 크게 설정되면 모델 교육의 손실률이 진동하거나 수렴되지 않는다. 학습률이 너무 작 으면 교육 과정이 크게 증가한다. 일반적으로 학습 속도 시작은 신경망 모델에서 0.1에서 선택한 다음 0.01, 0.001과 같이 지수 적으로 감소한다. 학습 속도는 단계 크기, 즉 역 전파 알고리즘에서

$\eta : w^n \leftarrow w^n - \eta \frac{\partial L}{\partial W^n}$ 라고도 한다.

2. Model Structure

이 논문에서는 컨볼 루션 레이어와 최대 풀링 레이어의 조합을 채택하고 드롭 아웃을 0.25로 설정하여 활성화 기능으로 ReLU를 사용했습니다. 우리 모델에서는 이러한 두 가지 조합 모듈을 사용했다. 컨볼 루션 커널 크기의 첫 번째 계층의 크기는 3X3이고 보폭은 1이다. 풀링 크기는 보폭이 1 인 2X2이다. 두 번째 조합 모듈의 보폭은 모두 2이다. 컨볼 루션 레이어와 드롭 아웃을 0.5로 설정한다. 다음 계층은 완전 연결 계층으로 드롭 아웃을 0.5로 설정하고 분류 기능으로 Softmax를 사용한다. 마지막 레이어는 출력 레이어이며 분류 결과를 출력한다.

Ⅳ. Experiment

Fashion-MNIST는 Zalando의 기사 이미지 데이터 세트이다. 이 데이터 세트에는 60,000 개의 학습 세트와 10,000 개의 테스트 세트가 포함되어 있다. 각 예는 10 가지 범주의 레이블과 관련된 28x28 회색조 이미지이다. 또한 각 범주의 데이터 예의 수가 고르게 분포되어 있다.

Fig. 2 및 3은 개선 된 모델의 정확성이 CNN 모델에 비해 상당히 증가되었음을 나타낸다. 훈련 반복 횟수가 증가함에 따라 모델의 검증 정확도는 점차 훈련 정확도와 일치한다. 또한, Fig. 2는 훈련 초기에 CNN 모델의 손실률이 감소한 다음 점차 증가하여 모델이 수렴되지 않았음을 나타낸다. 대조적으로, Fig. 3에 도시 된 바와 같이, 개선 된 모델의 손실률은 훈련 시작 이후 꾸준히 감소하였으며, 이는 훈련의 손실률보다 적다. 10 번의 반복 후에는 안정화되었고 훈련과 손실률 사이의 차이는 거의 없었다. 이는 모델의 빠른 수렴 속도와 높은 최적화 성능을 나타낸다.

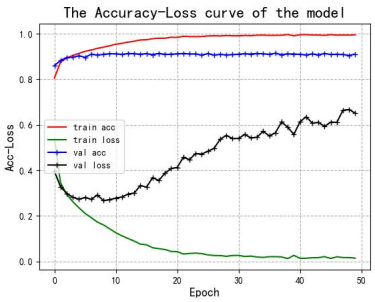

Fig. 2. Accuracy and loss rate curve of the CNN model

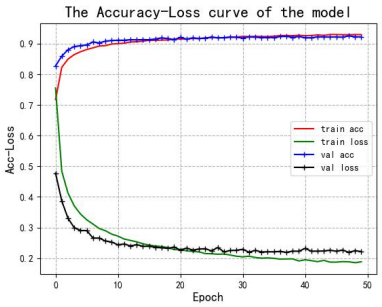

Fig. 3. Accuracy and loss rate curve of the optimized model

Ⅴ. Conclusions

본 연구에서는 학습률 및 모델 구조의 최적화에 중점을 둔 DL을 기반으로 한 패션 의류 이미지 인식 및 분류를 위한 새로운 방법을 제안했다. 훈련의 배치 크기에 따라 모델의 학습 속도를 동적으로 조정한다. CNN, LeNet, LSTM 및 BiLSTM 모델과 비교하여 제안된 모델의 실험 결과는 각각 4.5 %, 5.7 % 및 6.9 % 증가했다.

REFERENCES

[1] G. E. Hinton, R. R. Salakhutdinov, “Reducing the dimen sionality of data with neural networks,” Science, vol. 313, pp. 504-507, Jul. 28, 2006. DOI: 10.1126/science.1127647.