

새로운 패션 의류 이미지 분류

신성윤^{1*} · 이현창² · 신광성² · 김형진³ · 이재완¹

¹군산대학교 · ²원광대학교 · ³전북대학교

New Fashion Clothing Image Classification

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

우리는 패션 의류 이미지의 빠르고 정확한 분류를 달성하기 위해 최적화된 동적 붕괴 학습률과 개선된 모델 구조를 가진 딥 러닝 모델을 기반으로 하는 새로운 방법을 제안한다.

ABSTRACT

We propose a novel method based on a deep learning model with an optimized dynamic decay learning rate and improved model structure to achieve fast and accurate classification of fashion clothing images.

키워드

deep learning model, CNN, dynamic decay learning rate, fashion clothing images

I. 서 론

최근 몇 년 동안 컴퓨터 기술의 급속한 발전으로 인공 지능 기술에 대한 연구는 획기적인 진전을 이루었다. 그 중 머신러닝으로 대표되는 인공지능 알고리즘의 연구와 응용은 비약적인 발전을 이루었다[1]. 생산 및 생활 시나리오에서 점점 더 많이 사용되는 기계 학습. 예를 들어, 잘 알려진 AI 로봇 AlphaGo는 전문 바둑 선수가 되었다. 2016년에는 유명한 인간 전문 바둑 선수를 큰 점수로 이겼다[2][3].

설문 조사에 따르면 DL 모델의 주요 최적화 방법은 입력 데이터의 전처리, 모델 구조 최적화, 특징 추출 방법 최적화 및 모델 매개 변수 최적화에 중점을 둔다. 이 논문에서는 주로 모델 구조와 학습률의 최적화에 중점을 둔다.

II. 방법론

CNN 또는 ConvNet이라고 하는 Convolutional Neural Network는 컴퓨터 비전 및 음성 인식과 같은 많은 작업에서 널리 사용된다. 2012 ImageNet Challenge에서 Krichevsky et al.은 처음으로 심층 CNN을 적용하여 인상적인 결과를 얻었다. 그 이후로 deep CNN의 아키텍처 설계는 많은 연구자들이 기여하도록 끌어왔다. CNN은 일반적으로 입력 계층, 컨볼루션 계층, 풀링 계층, 완전 연결 계층 및 출력 계층으로 구성된다. 입력 레이어는 입력 벡터의 집합, 즉 신경망의 뉴런에 입력되는 벡터의 집합이다.

III. 제안방법

학습률은 신경망 훈련 및 학습 능력이며 신경망에서 조정하기 가장 까다로운 전역 매개변수이다. 학습률이 너무 크게 설정되면 모델 교육의 손실률이 진동하거나 수렴하지 않을 수도 있다. 학습

* corresponding author

률을 너무 작게 설정하면 훈련 과정이 크게 늘어난다. 일반적으로 학습률은 신경망 모델에서 0.1부터 시작하여 0.01, 0.001과 같이 기하급수적으로 감소한다. 학습률은 단계 크기라고도 한다. 즉, 역전

파 알고리즘 $\eta : w^n \leftarrow w^n - \eta \frac{\partial L}{\partial w^n}$ 에서 η 이다.

표 1은 우리가 제안한 학습률 최적화 알고리즘의 의사 코드(pseudo-code)를 나타낸다.

표 1. 동적 학습률 최적화 알고리즘의 의사 코드

Algorithm: Optimized Dynamic Learning Rate

Require: Initial learning rate α_0 , and the total iterations $global_step$

Require: Initialize parameter, decay rate ρ , hyper-parameter d .

- 1: Initialize accumulation variables $E[g^2]_i = E[\Delta\alpha^2]_{i=0}$
- 2: **For** $t = 1, 2, \dots, global_step$ **do**
- 3: Accumulate Gradient: $E[g^2]_i = \rho E[g^2]_{i-1} + (1-\rho)g_i^2$
- 4: Computer Update: $\Delta\alpha = RMS[\Delta\alpha]_{i-1} / RMS[g]_i \cdot g_i$
- 5: Accumulate update: $E[\Delta\alpha^2]_i = \rho E[\Delta\alpha^2]_{i-1} + (1-\rho)\Delta\alpha_i^2$
- 6: Computer learning rate: $\alpha_t = \alpha_{t-1} * e^{d * global_step}$
- 7: Learning rate update: $\alpha_{t+1} = \alpha_t + \Delta\alpha$
- 8: **End For**

IV. 실험

Table 2는 각 모델의 평가 정확도와 손실률을 보여준다. 최적화된 모델의 정확도가 92.87%로 가장 높고 손실률이 가장 낮은 0.20으로 감소한다는 결론을 내릴 수 있다. 표에서 우리가 제안한 방법이 다른 방법보다 성능이 우수함을 알 수 있다.

표 2. 최적화된 CNN 모델의 매개변수

Models	Accuracy (%)	Loss
CNN	89.10	0.37
LeNet	87.65	0.34
LSTM	87.15	0.36
BiLSTM	87.49	0.36
Ours	92.87	0.20

우리 모델의 Confusion Matrix는 그림 1과 같이 테스트 세트에 대한 예측 결과와 모델 분류를 직관적으로 반영합니다. 가장 좋은 예측은 "바지" 및 "가방" 범주이고 "샌들" 범주입니다. 최악의 결과는 "셔츠" 카테고리입니다. 이 때 "셔츠"와 "풀오버" 카테고리의 특징 정보가 매우 유사하다고 추정합니다. 개선된 모델은 "셔츠" 범주에서 예측 결과가 좋지 않지만 결과는 여전히 다른 모델보다 좋습니다.

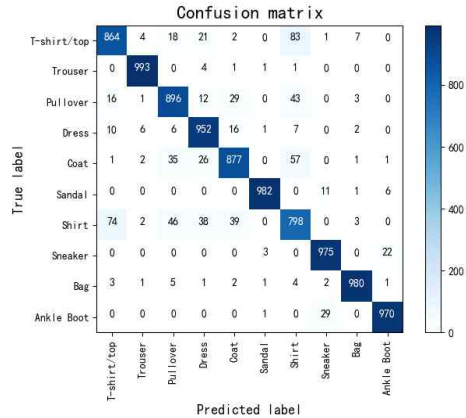


그림 1. 최적화된 모델의 정오분류표

V. 결론

본 논문에서는 학습률과 모델 구조의 최적화에 중점을 둔 DL 기반 패션 의류 이미지 인식 및 분류를 위한 새로운 방법을 제안했다. 훈련의 배치 크기에 따라 모델의 학습률을 동적으로 조정한다. convolution과 pooling 레이어의 조합 모듈을 설정함으로써 우리 모델은 이미지의 유용한 특징을 추출하는 뛰어난 능력을 가지고 있다.

Acknowledgement

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

References

- [1] G. E. Hinton, R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, pp. 504-507, Jul. 28, 2006.
- [2] David Silver, Aja Huang et al., "Mastering the game of Go with deep neural networks and tree search," Nature, vol. 529, pp. 484-489, 2016.
- [3] Michael C. Fu, "AlphaGo and Monte Carlo tree search: the simulation optimization perspective," 2016 Winter Simulation Conference (WSC), pp. 659-670, Dec. 11-14, 2016.