

MNIST 데이터셋에 Batch Normalization을 적용한 CNN의 정확도 향상에 관한 연구

이승빈, 이정수*

부산외국어대학교, *부산외국어대학교

dltmdqls98@naver.com, *rhee@bufs.ac.kr

A Study on the Accuracy Improvement of CNN with Batch Normalization Applied to the MNIST Dataset

Seungbin Lee, Jungsoo Rhee*

Busan University of Foreign Studies., *Busan University of Foreign Studies.

요약

본 논문은 MNIST 데이터셋을 활용한 손글씨 숫자 인식에서 합성곱 신경망(CNN)과 배치정규화(BN)를 결합한 모델을 제안한다. MNIST 데이터셋은 기계 학습 분야에서 손글씨 이미지 훈련 및 테스트용 학습 데이터로 꾸준히 사용되는 벤치마크로 인식되고 있다.

연구의 배경으로, 2015년 Ioffe와 Szegedy는 배치정규화 계층을 포함한 5계층 완전연결신경망을 통해 95.2%의 테스트 정확도를 달성하였다. 또한, 1998년 LeCun et al.은 CNN 모델인 LeNet-5를 사용하여 MNIST 데이터셋에서 99.05%의 테스트 정확도를 기록하였다. 본 연구의 목표는 LeCun et al.의 LeNet-5 모델의 성과를 뛰어넘는 것이다. 제안된 모델은 6계층 신경망 구조로 설계되었으며, 28x28 픽셀 이미지를 입력으로 받아 합성곱, 맥스 풀링, 완전연결 계층을 거쳐 처리한다. 특히 배치정규화 계층을 도입하여 학습의 안정성과 성능을 향상시켰다. 모델 구성은 30개의 필터와 필터 사이즈 5x5, 패딩 0, 스트라이드 1을 사용하였으며, ReLU 활성화 함수를 채택했다. 훈련 과정에서는 60,000개의 훈련 이미지와 10,000개의 테스트 이미지를 사용하였고, 미니배치 사이즈 100, 총 20 에포크, 학습률 0.1로 설정했다. 실험 결과, 제안된 모델은 99.22%의 테스트 정확도를 달성하여 LeNet-5의 99.05%를 상회하였으며, F1-score는 0.9919로 기록하여 모델의 성능을 입증했다. 또한, 본 논문에서 제안한 6계층 모델은 LeCun et al.의 LeNet-5(7계층 모델)와 Ji, Chun, Kim(10계층 모델)이 제안한 모델보다 더 단순한 구조로 모델의 효율성을 강조한다.

I. 서론

본 논문은 손글씨 이미지 분류 작업에서 Batch Normalization(BN)을 적용한 Convolutional Neural Network(CNN)의 성능을 평가하는 것을 목표로 한다. MNIST 데이터셋은 0~9까지의 숫자 손글씨 이미지로 구성된 표준 데이터셋으로, 다양한 신경망 모델의 성능 평가에 널리 사용되어 왔다.[1] 선행 연구에서 LeCun et al.은 CNN 기반의 LeNet-5 모델을 통해 99.05%의 정확도를 기록하였으며, Ioffe and Szegedy는 BN을 적용한 완전연결신경망(ANN)에서 95.2%의 정확도를 달성하였다.[2, 3] Han et al.은 MNIST 데이터셋과 잡음이 포함된 Kaggle 데이터셋에 CNN을 적용하여 테스트 정확도를 비교 분석하였다. 실제로 잡음을 입힌 Kaggle 데이터를 활용해 저자들은 93%의 테스트 정확도를 얻었다.[4] 참고로 Ji, Chun and Kim도 CNN과 BN을 결합하여 MNIST 데이터셋에서 학습하였으나 테스트 정확도가 99%로 LeNet-5에서 학습한 테스트 정확도 99.05%를 능가하지 못하였다.[5] 이러한 결과는 CNN에서의 BN의 효과를 명확하게 확인하지 못한 것으로 판단된다. 본 논문에서는 BN을 CNN에 결합하여, LeCun et al.의 논문의 결과보다 향상된 테스트 정확도를 보여주는 것을 목표로 한다. 이를 위해 제안된 CNN 모델에 SGD, Adam, AdaGrad, Momentum 네 가지 최적화 알고리즘을 모두 적용하여 훈련하고 테스트 정확도를 최고로 올려주는 최적화 알고리즘을 선택한다. 실험 결과 BN이 CNN과 결합된 경우에도 손글씨 분류 정확도가 향상되었으며, 최적의 성능을 보인 최적화 알고리즘을 도출한다.

Table 1. Test accuracy in previous studies

	CNN (kaggle)	LeNet-5	Ioffe's ANN with BN	ANN without BN	ANN with BN	CNN with BN (2018)
Author /Year	Han et al. /2019	LeCun et al. /1998	Ioffe & Szegedy /2015	Lee & Rhee /2023	Lee & Rhee /2023	Ji, Chun &Kim /2018
optimizer	Adam	SGD	SGD	Adam	Adam	-
test accuracy	93%	99.05%	95.2%	95.3%	96.47%	99%

이러한 결과는 BN의 효과가 CNN 구조에서도 긍정적인 영향을 미치며, 다양한 신경망 모델에서 BN의 활용 가능성을 제시한다.

II. 본론

LeNet-5는 32x32 크기의 손글씨 이미지 입력으로 시작하는 7계층의 합성곱 신경망(CNN)이다.

입력 데이터는 첫 번째 합성곱 계층(Conv1)에 유입된 후, subsampling 계층을 거쳐 두 번째 합성곱 계층(Conv2)으로 이동한다. 이후 추가적인 subsampling 2, 4계층과 합성곱 3, 5계층을 연속적으로 통과한 후, affine 계층을 거쳐 Euclidean Radial Basis Function과 SoftMax 계층으로 출력된다. 활성화 함수로는 tanh를 사용한다.[2]

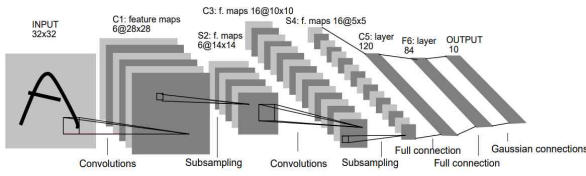


Fig. 1. LeNet-5 Model

본 논문에서는 CNN를 기반으로 BN을 결합한 CNN 모델을 제안한다. 이 제안된 모델은 MNIST 데이터셋을 활용하여 학습되었으며, Conv1, Pooling1, Affine1, BN, Affine2, SoftMax 계층을 포함한 6계층 구조로 구성된다. BN 계층은 데이터의 분포를 정규화하고 학습의 안정성을 향상시킨다.



Fig. 2. A proposed CNN Model equipped with BN

실험 과정은 다음과 같이 구성되었다. 첫째, MNIST 데이터셋은 60,000개의 훈련 이미지와 10,000개의 테스트 이미지로 구성되어 있으며, 각 이미지는 28x28 픽셀 크기의 원본 2D 형태로 입력된다. 둘째, 모델 구성은 30개의 필터와 필터 크기 5x5, 패딩 0, 스트라이드 1을 사용하였으며, ReLU 활성화 함수를 채택한다. SimpleConvNet 클래스를 사용하여 BN 계층이 포함된 CNN을 정의하였으며, 이 신경망은 20회의 에포크 동안 훈련된다. 훈련 중에는 100개의 미니배치를 사용하였고, 최적화 알고리즘으로는 Momentum을 적용하며, 학습률(learning rate)은 0.1로 설정한다. 매 에포크마다 1,000개의 테스트 샘플을 사용하여 모델의 성능을 평가한다. 마지막으로, 실험 결과 제안된 BN 결합 CNN 모델은 99.22%의 최종 테스트 정확도를 달성하였으며, 이는 BN의 효과를 통해 모델의 성능이 향상되었음을 보여준다. 참고로 본 논문에서 제안한 모델의 F1-score는 0.9919이다.

```
train loss:0.0001983261751535189
train loss:0.00028800496027501183
train loss:0.000581073381310562
train loss:0.0001455770924940003
train loss:0.00011600161758890328
train loss:0.00011693504041668307
train loss:0.0006261966099540035
train loss:0.0009741439282814275
train loss:0.0003936188987058515
train loss:0.00010789308929091645
train loss:0.0017537801845362633
train loss:0.00010572955246906924
train loss:0.000864874223131342
train loss:0.0004531861675933094
train loss:0.0006370124297017961
train loss:0.001236684964875127
train loss:0.00047680388804889373
===== Final Test Accuracy =====
test acc:0.9922
```

Fig. 3. A final test accuracy

III. 결론

본 논문에서는 CNN과 BN의 결합 효과와 중요성을 강조하고, MNIST 데이터셋에서 다양한 최적화 기술을 사용하는 신경망의 학습 과정에서 BN의 역할을 연구한다. 실제로 Fig. 3에서와 같이 BN을 적용한 1개의 합성곱계층을 가진 CNN에서 테스트정확도는 99.22%로 LeCun et al.가 제안한 LeNet-5에서의 테스트 정확도를 상회하는 것으로 나타난다. 더 나아가 위의 Table 2에서 보여지듯이 제안한 CNN의 모델 구성이 더 간단하다는 것도 이 논문의 결과로 알 수 있다.

Table 2. Detailed Comparison of LeNet-5 and the proposed CNN with BN

model	input	hidden1	hidden2	hidden3	hidden4	hidden5	hidden6	output
LeNet-5	mnist	conv	Sub sampling	conv	Sub sampling	conv	affine	ReLU, softmax
CNN with BN	mnist	conv	pooling	affine	BN	affine	-	softmax

Table 3. Detailed comparison of CNN with BN (2018) and the proposed CNN with BN.

model	input	hidden1	hidden2	hidden3	hidden4	hidden5	hidden6	hidden7	hidden8	hidden9	output
CNN with BN(2018)	mnist	conv	BN	maxpooling	conv	BN	maxpooling	fc625	BN	maxpooling	fc10
CNN with BN	mnist	conv	pooling	affine	BN	affine	-	-	-	-	softmax

참고로 Table 3는 서론에서 언급한 Ji, Chun and Kim이 제안한 모델과 비교한 표이다. 제안한 모델은 10계층으로 이루어져있어, 본 논문에서 제안한 모델과의 차이를 보인다.

ACKNOWLEDGMENT

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재 4.0 사업의 연구결과로 수행되었음” (IITP-2024-2020-0-01825)

참 고 문 헌

- [1] Lee, S. B., and Rhee, J. S. "Applying Batch Normalization to the MNIST Dataset," Quantitative Bio-Science, vol. 42, no. 2, pp. 133-137, Nov. 2023.
- [2] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.
- [3] Ioffe, S., and Szegedy, C. "Batch normalization: Accelerating deep network training by reducing internal covariate shift," Proceedings of the International Conference on Machine Learning, Proceedings of Machine Learning Research, vol. 37, pp. 448-456, Feb. 2015.
- [4] Han, M. R., Kim, S. H., Lee, S. I., Kim, K. H., and Kim, J. B. "A Study on the AI Based MNIST Handwriting Recognition," Journal of The Korea Society of Information Technology Policy & Management, vol. 11, no. 6, pp. 1509-1517, Dec. 2019.
- [5] Ji, M. G., Chun, J. C., and Kim, N. G. "An Improved Image Classification Using Batch Normalization and CNN," Journal of Internet Computing and Services, vol. 19, no. 3, pp. 35-42, June. 2018.