

합성곱 신경망 네트워크 구조 변화에 따른 숫자 인식을 비교

이종찬* · 김영현 · 송특섭

목원대학교

Comparison of Number Recognition Rates According to Changes in Convolutional Neural Structure

Jong-Chan Lee* · Young-Hyun Kim · Teuk-Seob Song

Mokwon University

E-mail : hellomias@naver.com

요 약

딥러닝을 적용한 기술 중 숫자 인식으로 예를 들 수 있다. 숫자 인식을 통하여 여러 분야에서 활용이 되고 있다. 숫자 인식을 가능하게 한 알고리즘 중 합성곱 신경망이 있다. 합성곱 신경망은 다양한 데이터들을 인식하는 데 사용되고 있다. MNIST 숫자 데이터셋을 활용하여 합성곱 신경망 구현 과정 중 깊게 레이어층을 쌓을수록 성능향상을 기대해볼 수 있다. 본 논문에서는 합성곱 레이어를 추가함으로써 성능향상을 76.96%에서 98.87%의 정확도가 산출되어 약 21.91%의 정확도가 향상됨을 확인하였다.

ABSTRACT

Digit recognition is one of the applications of deep learning, which appears in many fields. CNN network enables us to recognize handwritten digits. Also, It can process various types of data. As we stack more layers in CNN network, we expect more performance improvements. In this paper, we added a convolution layer. as a result, we achieved an accuracy improvement from 76.96% to 98.87%, which is a nearly 21.81% increase.

키워드

CNN, Machine learning, MNIST

1. 서 론

딥러닝이란 여러 데이터를 특정한 계층을 쌓아 특징을 학습하는 알고리즘이다. 10년간 하드웨어의 비약적인 발전으로 인하여 복잡한 연산 과정을 구현해나가는 인공지능 알고리즘을 효율적으로 수행할 수 있게 되었다. 딥러닝의 장점으로는 학습시킬 데이터들이 많아지고 학습시킬 단계 계층을 깊게 쌓을수록 성능향상을 기대할 수 있다는 점이다. 딥러닝 기술을 접목하게 시킨 활용 분야로는 문자나 생체 (홍채, 지문, 얼굴)인식, 자율주행 자동

차, 증강현실 등 다양한 분야에서 응용되고 있다 [1,2,3].

딥러닝을 적용한 기술 중 숫자 인식을 예를 들 수 있다. 숫자 인식의 활용 분야로는 자동차 차량 번호를 인식시키는 기능이 있다. 과거 우체국에서 사람들이 손으로 적은 우편번호를 인식 후 분류하여 배달하는 것에 인력을 소모하는 작업을 생략할 수 있는 등의 편의를 주고 있다. 숫자 인식의 방법으로는 대표적으로 인공신경망, 합성곱 신경망 형태 등이 있다 [2,3]. 인공신경망은 입력받은 데이터값과 각각의 연결에 가중값을 이용하여 입력의 합을 계산 후 시그모이드 함수를 이용하여 예측값을 도출한다. 합성곱 신경망은 합성곱 연산을 통하

* speaker

여 특징을 추출하여 활성화 함수를 적용하여 형태를 예측하는 역할을 한다. 본 논문에서는 숫자 필기체 데이터 세트 MNIST를 활용하여 합성곱 신경망의 네트워크 구조를 변경하고 학습 결과를 비교할 것이다 [5].

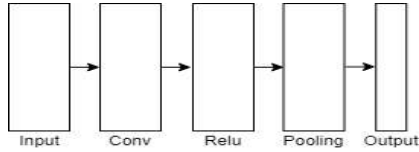


그림 1. CNN 실행 과정

II. 합성곱 신경망 네트워크를 통한 숫자 인식

합성곱 신경망은 합성곱 연산을 사용하는 인경 신경망 중 하나이고 시각적 이미지를 분류하는 데 쓰인다. 합성곱 신경망의 합성곱 단계에서는 여러 필터가 존재한다. 합성곱 단계의 처리 방법으로는 입력데이터에 필터를 일정 간격(스트라이크)도 이동하면서 입력데이터에 적용한 후 입력과 필터에 어울리는 요소와 곱한 뒤 특징의 뚜렷한 형태를 추출하기 위하여 맥 스폰링 과정을 거친다. 맥 스폰링이란 특정한 영역에서 일정한 간격으로 이동하면서 큰 값을 추출하는 방식이다. 각종 필터 작업이 마무리되면 그 총합을 구한 뒤 분류 층(완결 연결 층)에 저장된다. 손글씨 인식을 위해 CNN을 구성한 네트워크 구조로는 다음과 같다. 입력의 데이터 값은 MINST 28*28 흑백 손글씨 숫자 이미지이며 입력된 값에 넣을 첫 번째 콘볼루션 층에서는 필터값을 3으로 설정하였고 합성곱에 사용되는 필터의 크기로는 3으로 설정 후 스트라이드를 1로 구성하였다. 다음으로는 맥 스폰링을 3*3으로 설정하고 스트라이드를 2칸씩 연산하게 설정하였다. 첫 번째 실험 결과로는 정확도가 0.7696으로 10,000개의 이미지 중 7,696개가 성공하였다.

표 1. 실험1 네트워크 구조

구분	Convolution_1	MaxPooling_1
실험1	filter : 3 kernel_size : 5*5 stride : 1	pool_size : 3*3 stride : 2*2

표 2. 실험2 모델 구조

구분	Convolution_1	MaxPooling_1
실험2	filter : 5 kernel_size : 5*5 stride : 1	pool_size : 3*3 stride : 2*2
	Convolution_2	MaxPooling_2
	filter : 3 kernel_size : 5*5 stride : 1	kernel_size : 3*3 stride : 2

두 번째 실험에서는 이전에 적용한 구조를 덧붙여 중첩 적분 층과 풀링 층의 구조를 추가하였다. 최종적으로 완결 연결 층으로 활성화 함수를 적용하여 분류하였으며 연산을 시킨 결과 정확도는 0.9887으로 9887개를 예측하는 데 성공하였다.

표 3. 반복 횟수 5회 시행 결과

구분	네트워크 구조 변경 전	네트워크 구조 변경 후
정확도 (예측/비교)	7696/10000	9887/10000

III. 결 론

실험 결과 합성곱을 한번 거친 모델과 두 번 거친 모델의 성능 차이로 약 21.91%의 정확도 향상을 확인하였다. 앞으로의 과제는 모델의 성능향상을 위하여 더 단계 층을 더욱 깊게 층을 쌓고 하드웨어의 발전으로 성능의 발전을 기대하는 것이다.

Acknowledgement

이 논문은 2018년도 정부(교육부)의 재원으로 한 연구재단의 지원을 받아 수행된 기초연구사업임 (No.NRF-2018R1D1A1B07048675)

References

[1] E. S. Lee, E.R. Jeong, "Deep Learning based Frame Synchronization Using Convolutional Neural Network," *J. Korean Institute of Information and communication Engineering*. vol. 24, No. 4, 504-507, 2020

[2] Lee, S., Kim, H. J., & Kim, M.. "Implementing a Depth Map Generation Algorithm by Convolutional Neural Network." *Journal of Broadcast Engineering*.

- 23(1), 3-10.
- [3] D.-G. Lee, "Classification of Trucks using Convolutional Neural Network," *Journal of Convergence for Information Technology*, Vol. 8, No. 6, pp. 375-380, 2018.
 - [4] K.-M. Park and C.-O. Bae, "A Study on Fire Detection in Ship Engine Rooms Using Convolutional Neural Network," *Journal of the Korean Society of Marine Environment and Safety*, Vol. 25, No. 4, pp. 476-481, 2019.
 - [5] G.-M. Choi and Y.-J. Jeong, "Efficient Iris Recognition using Deep-Learning Convolution Neural Network," *The Journal of the Korea institute of electronic communication sciences*, Vol. 15, No. 3, pp. 521-526, 2020.