

효과적인 필기체 인식을 위한 인공지능 알고리즘

김승주, 이재영, 한종기

*safe586@naver.com, **lgy321456@gmail.com, ***hjk@sejong.edu

Artificial Intelligence Algorithms for Identification of Handwriting

*Seung-Ju Kim, **Jae-Yung Lee, ***Jong-Ki Han

Sejong University

요약

최근 스마트폰, PC, 태블릿 같은 전자기기들이 발전하면서 기계를 통해 소통하는 시대가 왔다. 기계와 소통하기 위해 우리가 사용하는 문자를 인식하는 것은 중요한 일이다. 이런 전자기기들이 문자, 영상인식을 해야 할 필요성이 더욱 증가함에 따라 머신러닝의 중요성이 대두되었다. 머신러닝은 컴퓨터의 학습을 위해 알고리즘과 기술을 개발하는 분야를 말한다. 머신러닝의 기법과 관련된 알고리즘의 종류는 수없이 많다. 그 중에서도 Neural Network는 사람의 뇌 신경구조를 토대로 착안하여 네트워크를 만들고 이를 학습에 이용한 머신러닝 기법이다. 이런 인공지능 알고리즘인 Neural Network 구조를 바탕으로 특징을 추출하여 학습을 하는 Convolution Neural Network 기법의 사용이 늘고 있다. 본 논문에서는 Neural Network와 Convolution Neural Network의 알고리즘을 이용한 필기체 인식 실험을 하고 그 내용을 비교하였다.

1. 서론

21세기에 컴퓨터와 네트워크의 발전으로 어디에서나 전자기기를 사용할 수 있는 시대가 왔다. 따라서 컴퓨터가 자동적으로 문자, 기호, 이미지를 인식해야 하는 문제를 해결해야 했다. 전자기기에서 영상을 인식하는 문제는 많이 발전해 왔고, 더불어 의사소통의 기본인 문자인 식문제에 대해서도 많은 머신러닝 기법들이 사용되었고 그 실험 또한 많이 진행되어 왔다. 우리는 문자를 기반으로 소통하는 데 같은 문자임에도 불구하고 사람마다 필기체에 차이가 있기 때문에 서로 다른 글씨에 대한 인식 문제의 해결이 필요해 졌다.

본 논문에서는 머신러닝 기법 중 하나인 인공지능 알고리즘 Neural Network와 Convolution Neural Network를 가지고 필기체 인식을 다루고 그 실험 결과를 비교하였다.

본 논문의 2장에서는 Neural Network를 이용한 필기체 인식 방법에 대해서 설명하고, 3장에서는 Convolutional Neural Network를 이용한 필기체 인식 방법에 대해서 설명한다. 4장에서 제시되는 실험 결과에서는 NN과 CNN을 이용한 필기체 인식 실험의 성능을 분석하고, 간단한 결론을 5장에서 제시한다.

2. Neural Network를 이용한 필기체 인식

Neural Network를 통한 인식은 두 가지 방향으로 이루어져 있다. input data를 넣어 output data를 출력해내는 과정인 forward 과정과

output data와 desire data 간의 차이 즉, error를 계산하여 줄이는 방향으로 back-propagation을 통해 학습하는 두 가지 과정으로 이루어져 있다. 우선 그 중 output data를 출력하는 forward 과정은 여러 단계의 layer를 사용하며 weight와 activation 함수에 따라 각 layer의 output이 만들어 진다. 이러한 각 layer의 함수는 수식 (1)과 같이 나타낼 수 있다. [1]

$$output = f\left(\sum_i w_i x_i + b\right) = f(W^T X + b) \quad (1)$$

수식(1)에서 x_i 은 input 하나를 나타내고 w_i 는 x_i 과 다음 layer의 neuron 과의 weight를 의미한다. W 는 weight들의 vector 값을 의미하고 X 는 input들의 vector 값을 의미한다. b 는 bias이며 f 는 activation function을 의미한다.

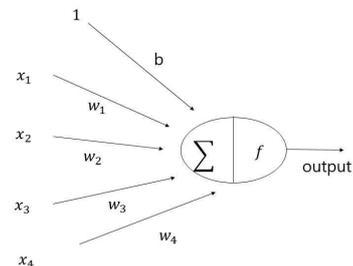


그림 1. 인공 뉴런 구조

1) 이 논문은 2015년도 미래창조과학부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2015R1A2A2A01006193).

연락처: 한종기

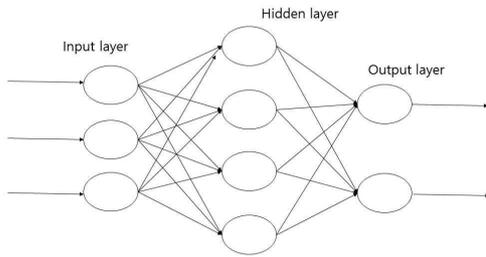


그림 2. 계층화된 Neural Network의 구조

그림 1에서 input data를 받아 weight를 곱해준 뒤 합을 구해 activation function을 통과시키는 인공뉴런의 기본 구조를 나타내었다. 그림 2에서는 인공뉴런이 모여 계층화된 Neural Network의 구조를 나타내었다.

3. Convolution Neural Network를 이용한 필기체 인식

Neural Network는 input data를 넣으면 output data를 출력해내는 구조이다. 이러한 구조는 input data간의 연관관계를 처음에 고려하지 않은 방식으로 오직 업데이트되는 weight에 대해서만 의존적인 방식이다. 이러한 문제를 해결하기 위해 input data를 넣기 전에 미리 전처리를 해서 사용하는 방법도 고안되었지만, 이보다는 input data 자체를 넣어서 필터를 사용하여 convolution 과정을 거쳐 서로 연관성을 찾아 output data를 산출해 내는 Convolution Neural Network의 성능이 좋아 인기를 끌고 있다.

Convolution Neural Network는 input data의 연관성을 고려한 만큼 Neural Network 보다 복잡할 뿐만 아니라, 3가지의 단계로 이루어져 있다. 우선 그림 3을 통해 Convolution Neural Network의 전체적인 구조를 나타내었다. [2]

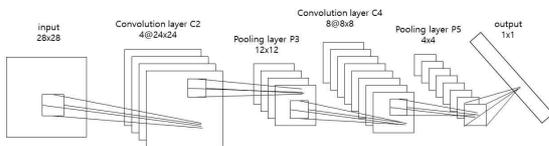


그림 3. Convolution Neural Network의 구조

(1) 1단계

Convolution 단계로 필터를 사용하여 input data의 특징을 추출해서 feature map을 만든다.

$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$
$\frac{1}{12}$	$\frac{4}{12}$	$\frac{1}{12}$
$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

$-\frac{1}{9}$	$-\frac{1}{9}$	$-\frac{1}{9}$
$-\frac{1}{9}$	1	$-\frac{1}{9}$
$-\frac{1}{9}$	$-\frac{1}{9}$	$-\frac{1}{9}$

그림 4. 필터의 예시 (왼쪽 가우시안 필터, 오른쪽 샤프닝 필터)

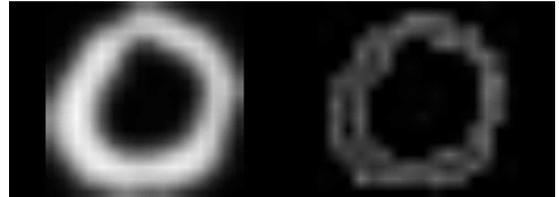


그림 5. 숫자 0의 feature map 예시 (왼쪽 가우시안 필터, 오른쪽 샤프닝 필터)

이러한 필터를 통해 input data 특히 영상 data에서의 특징 추출이 용이하다. edge, 저주파, 고주파 성분등 여러 가지 특징 정보를 알아낼 수 있다.

(2) 2단계

1 단계를 통해 feature map을 얻었으며, 2 단계는 이러한 feature map들의 크기 정보를 줄이는 max-pooling 단계이다. 2X2 크기의 영역을 잡아 가장 큰 값을 뽑아 추출해 낸다. 이 과정을 통해 3번째 단계에서의 neural network로 들어가는 input data의 양을 줄일 뿐만 아니라 영향을 주는 큰 값 하나를 제외한 나머지를 반영하지 않으므로 복잡도가 낮게 작업을 수행할 수 있다.

(3) 3단계

본 단계에서는 앞에서의 Neural Network와 같은 과정으로 진행된다. Convolution 단계를 거쳐 온 데이터 들을 input data로 하여 Neural Network의 단계를 진행하고 output을 얻어 낸 뒤, error를 구해 back-propagation방법을 통해 학습시킨다.

4. 실험 결과

손글씨 데이터 베이스를 이용해 Neural Network를 이용한 손글씨 인식률과 Convolution Neural Network를 이용한 손글씨 인식률을 비교하여 막대그래프로 작성하였다. 이러한 실험 결과를 통해 Convolution Neural Network를 사용했을 때의 인식률이 조금 더 좋은 것을 알 수 있다.

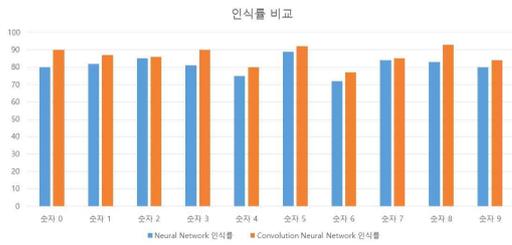


그림 6. 숫자에 따른 인공지능 알고리즘 인식률 비교

5. 결론

본 논문에서 필기체 인식을 인공지능 알고리즘을 통해 실험해 보았다. 두 가지의 방법으로 실험해 보았는데, Neural Network는 단순히 input data를 입력받아 output를 산출해 내고 학습시키는 알고리즘이다. 이런 과정은 기계 내부적인 weight 값에 의존적이므로 비효율적이라고 할 수 있다. 반면 Convolution Neural Network는 input data를 받아 스스로 필기체의 특징을 인식하여 output을 산출해 내고 학습시키는 알고리즘이다. 특징인식이라는 이점 덕분에 Convolution Neural Network의 인식 성능이 좋은 것을 직접 구현하여 실험을 통해 증명하였다.

참 고 문 헌

- [1] M. Wu, L. Chen, "Image recognition based on deep learning," Chinese Automation Congress, pp. 542 - 546, 11, 2015
- [2] C. Razafimandimby, V. Loscri, A. M. Vegni, "A neural network and IoT based scheme for performance assessment in internet of robotic things," 2016 IEEE Intern. Conf. on IoT Design and Implemt., pp. 241 - 246, 4, 2016