

Convolutional neural network를 이용한 눈동자 모션인식 시스템 구현

이승준 · 허승원 · 이희빈 · 유윤섭
한경대학교 전기전자제어공학과, IITC

Implementation to eye motion tracking system using convolutional neural network

Seung Jun Lee · Seung Won Heo · Hee Bin Lee · Yun Seop Yu

Dept. of Electrical, Electronic, Control Eng. and IITC, Hankyong National University

E-mail : tmdwns_4534@naver.com

요 약

본 논문은 몸을 움직이지 못하는 루게릭병 환자들을 위해 눈동자를 추적하여 의사소통 시스템에 필요한 눈동자의 위치를 파악해주는 인공신경망 설계에 대해 소개한다. Tensorflow를 이용해 신경망 생성 및 학습하고 학습된 신경망을 통하여 눈동자의 위치를 파악한다. 본 논문에서는 컨볼루션계층 2단계와 완전연결계층 2단계로 구성된 Convolution Neural Network(CNN)을 사용해서 구현했다.

ABSTRACT

An artificial neural network design that traces the pupil for the disables suffering from Lou Gehrig disease is introduced. It grasps the position of the pupil required for the communication system. Tensorflow is used for generating and learning the neural network, and the pupil position is determined through the learned neural network. Convolution neural network(CNN) which consists of 2 stages of convolution layer and 2 layers of complete connection layer is implemented for the system.

키워드

인공신경망, Tensorflow, CNN, 눈동자 추적

1. 서 론

인공 신경망은 생물학의 신경망에서 영감을 얻어 만들어진 통계학적 학습 알고리즘이다[1]. 최근에 인공 신경망은 음성 인식, 영상 처리, 데이터 분석 등의 여러 분야에 사용되고 있다. 루게릭병 환자와 같이 몸을 움직이지 못하는 환자를 위해 눈동자를 추적하여 의사소통 시스템에 대한 연구가 활발히 진행되고 있다[2]. 이 시스템에서 눈동자의 움직임을 인공신경망을 이용한 연구는 발표가 매우 적다 [3].

본 논문은 인공 신경망을 이용해서 눈동자의 움직임을 추적하는 시스템 구현에 대해서 소개한다.

II. 본 론

1) 신경망 구성

몸을 자유롭게 움직이지 못하고 말을 못 하는

환자와 의사소통하기 위해서 눈동자 움직임을 추적하는 시스템을 인공신경망을 이용해서 구현한다. 이 시스템에서 인공신경망은 컨볼루션 신경망(convolutional neural network; CNN)이 사용된다. CNN 계층은 가로와 세로의 크기가 5인 필터에 의해 합성곱 연산을 진행하고 sigmoid와 relu 활성화함수로 가공된다. 그 후에 2x2 max pooling을 거치고 완전연결계층으로 입력된다. 첫 번째 완전연결계층에 입력되는 크기는 1024이고 마지막으로 출력될 때의 크기는 3이다. 그리고 마지막 계층의 활성화함수는 softmax 함수이다. 그림 1은 CNN 계층별 블록도를 나타낸다. 첫 번째 컨볼루션 계층으로 240x320 크기의 그레이 변환된 이미지가 입력으로 들어가고 합성곱과 relu함수로 가공 후에 2x2 max pooling을 거쳐 다음 컨볼루션 계층의 입력으로 들어간다. 두 번째 컨볼루션계층에서도 첫 번째 계층과 같은 과정을 거치고 첫 번째 완전연결계층의 입력으로 들어가는데, 이때 3차원 데이터를 1차원으로

로 바꿔 주어야 한다. 첫 번째 완전연결계층에서 가중치와 입력의 곱과 바이어스값의 합인 가중합을 sigmoid함수로 가공해 다음 계층으로 넘겨주고 마지막 계층에서는 가중합을 softmax함수로 가공해줘서 마지막 출력을 얻는다.

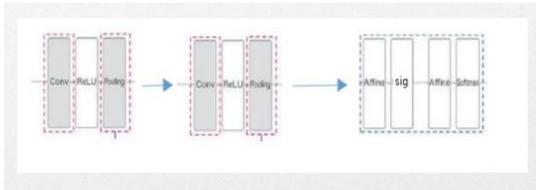


그림 1. CNN 계층별 블록도

2) 학습데이터

이 신경망에 학습데이터로 쓰이는 데이터들은 그림 2와 같이 눈동자가 오른쪽, 가운데, 왼쪽 각각 10개씩 총 30개의 이미지를 학습데이터로 사용하였고 이미지의 크기는 240x320이다. 또한 학습 알고리즘으로 ADAM 최적화 알고리즘을 사용하였다.



그림 2. 눈동자 추적 사진(오른쪽, 가운데, 왼쪽)

III. 결 과

본 인공지능망을 학습하는데 30개 이미지가 사용되었다. 총 10번 학습한 학습데이터로부터 얻은 가중치들로 구성된 신경망으로 30개의 이미지(왼쪽 10개, 오른쪽 10개, 가운데 10개)를 테스트한 결과는 표 1과 같다. 30개의 이미지 모두가 정확히 눈동자 위치를 검출하였다.

표 1. 눈동자 위치 추적 정확도

	왼쪽	가운데	오른쪽	총합
샘플갯수	10	10	10	30
검출갯수	10	10	10	30
검출율	100%	100%	100%	100%

IV. 결 론

본 논문에서 눈동자의 위치를 파악해주는 컨볼루션계층이 2단계, 완전연결계층이 2단계로 구성된 인공지능망을 소개하였다. Tensorflow 라이브러리를 이용해 인공지능망을 구성하고, 연산에 필요한 함수를 사용했다. 또한 240x320의 그레이 변환된 이미지 30개를 총 10회 학습을 하였고

학습 알고리즘으로는 ADAM 최적화 알고리즘을 사용하였다.

본 논문에서 소개한 인공지능망으로 눈동자 위치를 인식한 결과를 보면, 30개 이미지를 테스트했을 때 모두 정확히 검출되는 성능을 확인했다.

참고문헌

- [1] A. K. Jain, J. Mao, and K. M. Mohiuddin, *Computer*, vol. 29, no. 3, pp. 31-44, Mar 1996.
- [2] S. Dongre and S. Patil, *International Journal of Computer Science and Mobile Computing*, vol. 4, no. 7, pp. 156 - 160, 2015.
- [3] C. Gou, Y. Wu K. Wang, F.-Y. Wang, and Q. Ji, *Pattern Recognition*, vol. 67, pp. 23-31, 2017.