

딥러닝 네트워크 조사 보고서

학번: 201984-010010

이름: 김 동 훈

MLP(Multi-Layer-Perceptron, 다층 퍼셉트론, 이하 MLP)

MLP에 대한 개념을 이해하기 위해서는 먼저 퍼셉트론(Perceptron, 이하 퍼셉트론)에 대한 이해가 먼저 필요하다. 퍼셉트론은 프랑크 로젠블라트(Frank Rosenblatt)가 1957년에 고안한 알고리즘이다. 퍼셉트론은 '인공 뉴런'으로 불리기도 하는데 다수의 신호를 입력으로 받아 하나의 신호를 출력하는 구조를 지닌다.

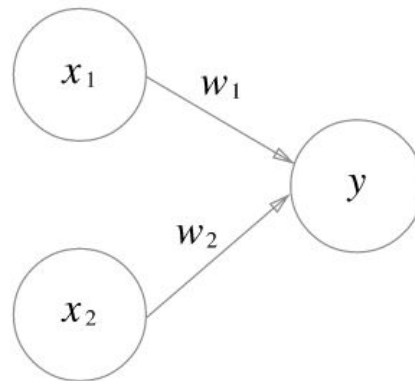


그림 1. 입력이 2개인 퍼셉트론

그림 1.은 입력으로 2개의 신호를 받은 퍼셉트론의 예이다. x_1 과 x_2 는 입력 신호, y 는 출력 신호, w_1 과 w_2 는 가중치를 뜻한다. 원은 뉴런 혹은 노드라고 부르고 입력 신호가 뉴런에 보내질 때 각각의 고유한 가중치가 곱해진다. 뉴런에서 보내온 신호의 총합이 정해진 한계(임계값)를 넘어설 때 1을 출력하는 구조이다. 퍼셉트론은 복수의 입력 신호 각각에 고유한 가중치를 부여하고 가중치는 각 신호가 결과에 주는 영향력을 조절하는 요소로 작용한다. 아래와 같이 퍼셉트론을 이용하면 AND 게이트, OR 게이트, NAND 게이트와 같은 단순 논리 회로 구현이 가능하다.

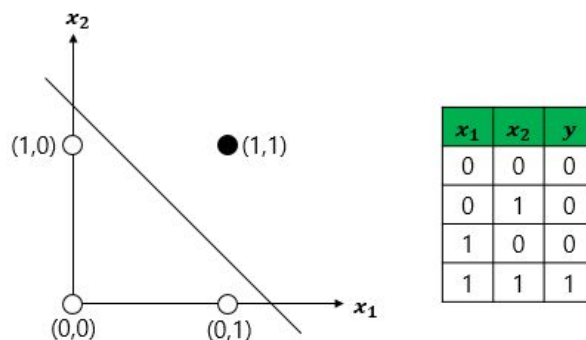


그림 2.AND 게이트

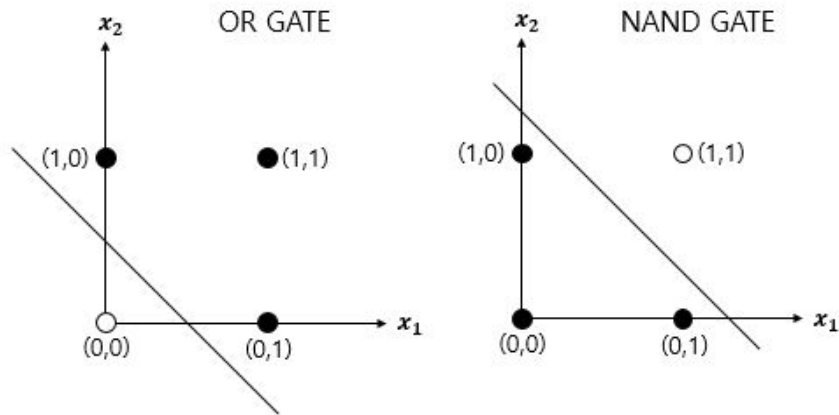


그림 3.OR 게이트

그림 4.NAND 게이트

위와 같이 퍼셉트론으로 구현된 AND 게이트, OR 게이트, NAND 게이트에서 중요한 점은 퍼셉트론의 구조는 모두 동일하고 세 가지 게이트에서 다른 것은 매개변수(가중치와 임계값)의 값만 다르다는 것이다. 하지만 퍼셉트론은 직선 하나로 나눈 영역만 표현할 수 있다는 XOR 게이트 구현이 불가능한 한계가 있다.

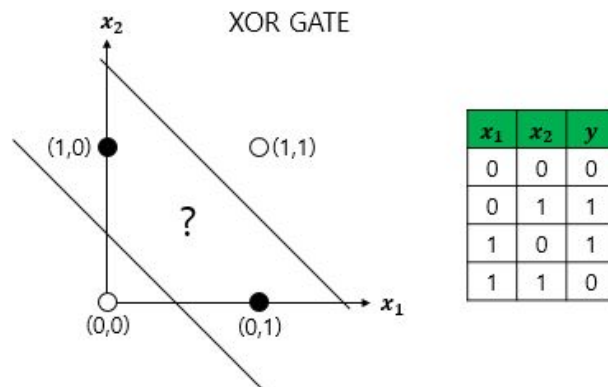


그림 5.퍼셉트론 한계

XOR 게이트를 구현하려면 아래와 같이 직선이라는 제약을 없애면 구현이 가능하다.

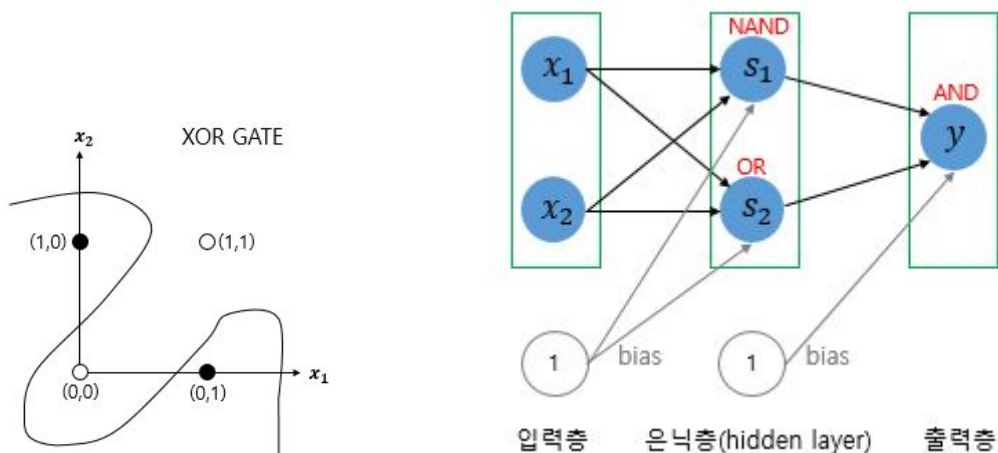


그림 6.XOR 게이트

그림 7.XOR 퍼셉트론

XOR는 그림 7.과 같은 다층의 네트워크 구조이다. 이처럼 층이 여러 개인 퍼셉트론을 다층 퍼셉트론 즉 MLP라고 한다. 단층 퍼셉트론으로 표현하지 못한 것을 층을 늘린

MLP를 이용하면 표현할 수 있는 것이다. 이러한 논리로 XOR 문제보다 더욱 복잡한 문제도 MLP의 중간 은닉층을 더 추가하고 적절한 가중치를 찾아 해결할 수 있다.

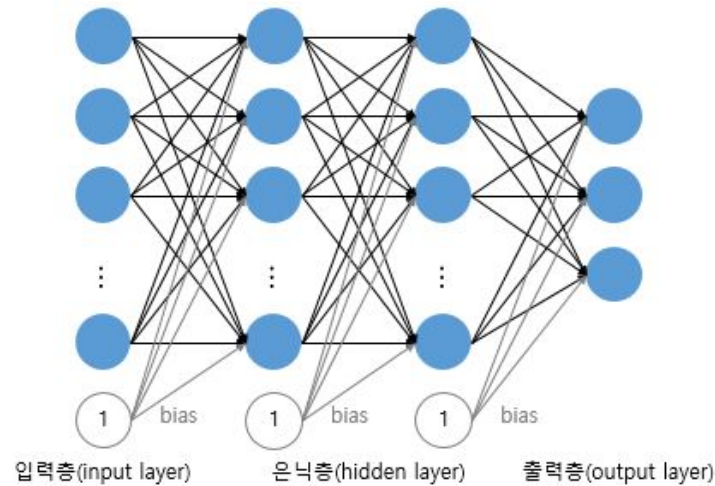


그림 8.다층 퍼셉트론

CNN(Convolutional Neural Network, 합성곱 신경망, 이하 CNN)

CNN은 이미지 인식과 음성 인식 등 다양한 곳에서 사용되는데, 특히 이미지 인식 분야에서 딥러닝을 활용한 기법은 거의 모두 CNN을 기초로 구현되었다. CNN은 신경망과 같이 계층을 조합하여 만들 수 있다. 다만 합성곱 계층(Convolutional Layer, 이하 합성곱 계층)과 풀링 계층(Pooling Layer, 이하 풀링 계층)이 추가되고 패딩(Padding, 이하 패딩), 스트라이드(Stride, 이하 스트라이드) 용어가 등장한다.

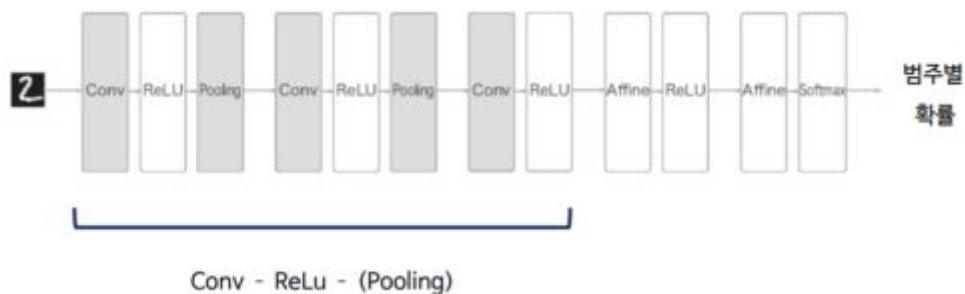


그림 9.합성곱 계층과 풀링 계층

합성곱 계층에서는 합성곱 연산을 처리한다. 합성곱 연산은 이미지 처리에서 말하는 필터 연산에 해당한다.

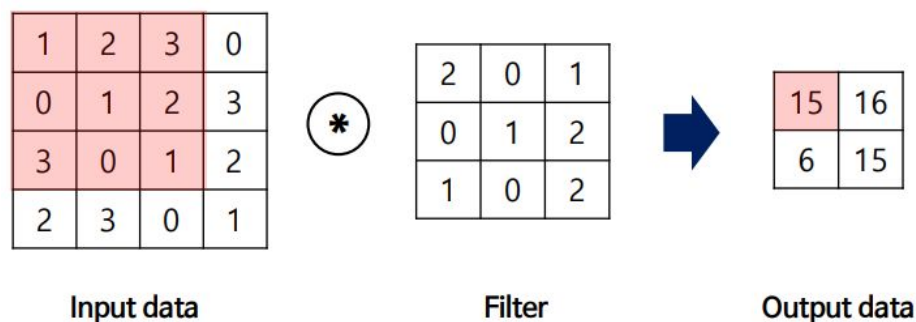


그림 10.합성곱 연산

합성곱 연산은 그림 10.과 같이 단일 곱셈-누산(fused multiply-add, FMA) 기법을 이용하여 필터의 윈도우(Window)를 일정한 간격으로 이동하며 입력 데이터에 적용한다.

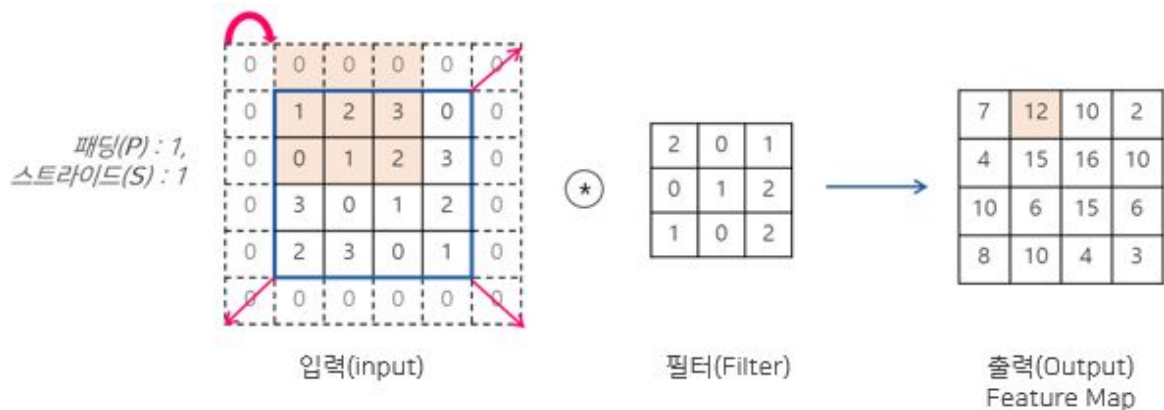


그림 11.패딩

합성곱 연산을 수행하기 전에 입력 데이터 주변을 특정값으로 채우는 것을 패딩이라고 하는데 합성곱 연산에서 자주 이용하는 기법이다. 필터를 적용하는 위치 간격을 스트라이드라고 한다. 스트라이드를 키우면 출력 크기는 작아진다.

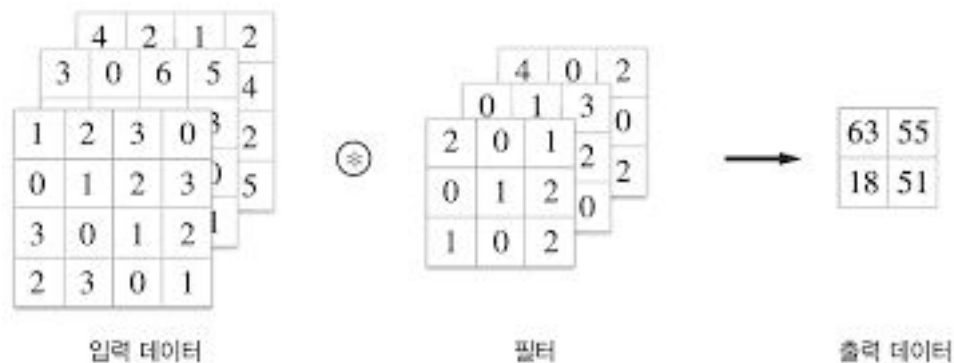


그림 12.3차원 데이터의 합성곱 연산

그림12.와 같이 3차원 데이터의 합성곱 연산이 가능한데 주의할 점은 입력 데이터의 채널 수와 필터의 수가 같아야 한다는 것이다. 3차원 합성곱 연산은 데이터와 필터를 직육면체 블록으로 생각하면 쉽다.

폴링은 세로.가로 방향의 공간을 줄이는 연산이다.

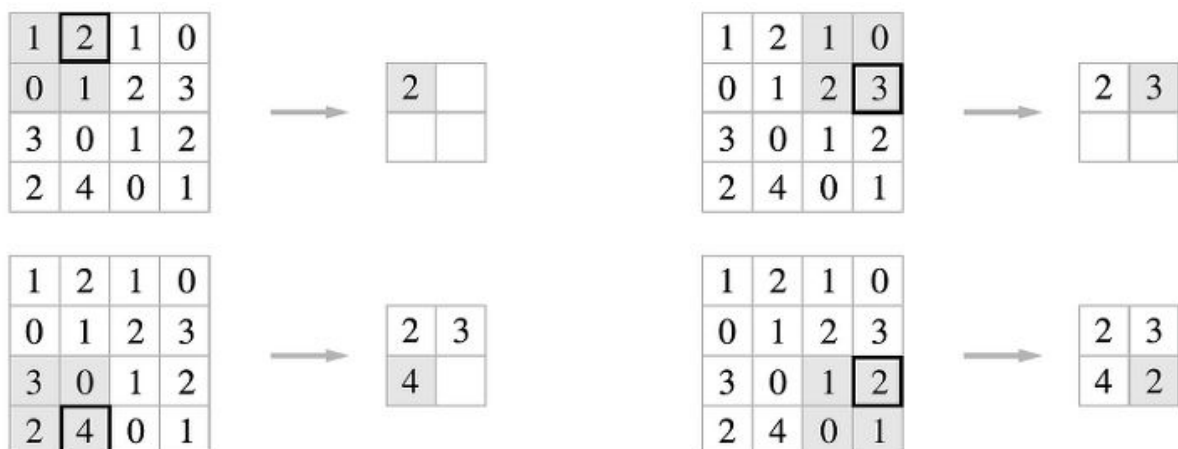


그림 13.최대 풀링(Max Pooling)

풀링은 최대 풀링과 평균 풀링 등이 있으며 그림 13.은 2X2 최대 풀링을 스트라이드 2로 처리하는 순서이다. 풀링의 특징은 아래와 같은 특징을 가진다.

- 학습해야 할 매개변수가 없다.
- 채널 수가 변하지 않는다.
- 입력의 변화에 영향을 적게 받는다.

대표적인 CNN은 1998년에 제안된 손글씨 숫자를 인식 네트워크인 LeNet과 2012년에 발표된 AlexNet이 있다.

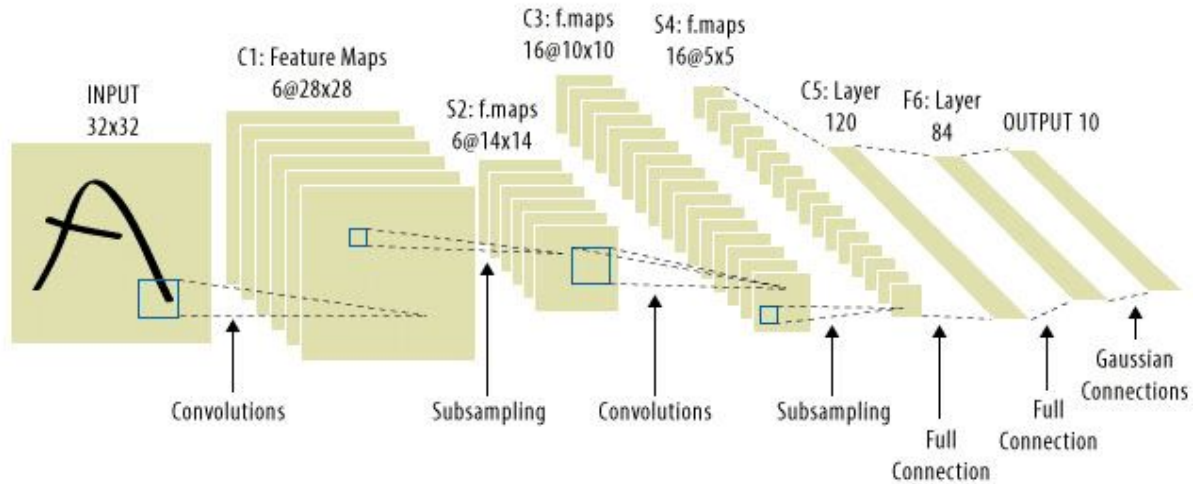


그림 14.LeNet

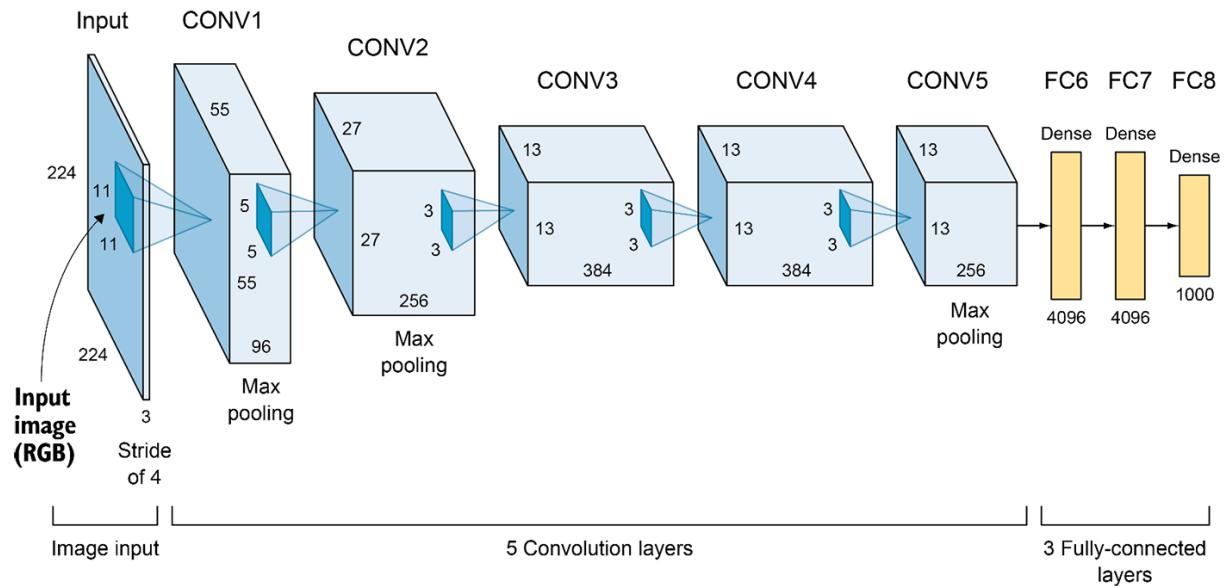


그림 15.AlexNet

참고문헌

[1] 조석희, 조규철 (2019). Deep CNN 기반의 얼굴인식 시스템 개발. 한국경영과학회 학술대회논문집, 2750-2752

[2] 이용규 (2014). RBM과 MLP를 이용한 얼굴 분류 시스템. 한국정보과학회 학술발표논문집, 607-609

- [3] 박준연, 김민재, 고준영, 김재현 (2018). CNN 구조 얼굴 식별의 정확도 향상을 위한 최적화. 한국통신학회 학술대회논문집, 387-388
- [4] 최인규, 안하은, 송혁, 고민수, 유지상 (2016). CNN 기반의 얼굴 표정 인식. 한국방송미디어공학회 학술발표대회 논문집, 320-321
- [5] 조석희, 조규철 (2019). Deep CNN 기반의 얼굴인식 시스템 개발. 한국경영과학회 학술대회논문집, 2750-2752
- [6] 이명오, 윤의녕, 고승현, 조근식 (2019). 특징 연결과 깊이별 분리 컨볼루션을 이용한 효율적인 얼굴 감정인식 CNN. 한국정보과학회 학술발표논문집, 754-756