

# 수학기반 인공지능프로그래밍

## (44630-01)

### 과제 4

제출일자: 10월 30일 수요일

담당교수: 이덕우 ([dwoolee@kmu.ac.kr](mailto:dwoolee@kmu.ac.kr))

#### 컴퓨터 비전(Computer Vision)

머신 러닝과 신경망을 사용하여 디지털 이미지, 비디오 및 기타 시각적 입력에서 의미있는 정보를 추출하고 결함이나 문제를 발견했을 때 권장 사항을 제시하거나 조치를 취하도록 컴퓨터와 시스템을 학습시킵니다. AI가 컴퓨터를 생각할 수 있게 만든다면, 컴퓨터 비전은 컴퓨터가 보고 관찰하며 이해할 수 있게 만듭니다.

컴퓨터 비전에서는 콘볼루션 신경망(Convolutional Neural Network, CNN)이 널리 사용됩니다.

##### 콘볼루션 신경망(CNN)

합성곱 신경망이라고도 하며 CNN은 시각적 영상을 분석하는 데 사용되는 다층의 '피드-포워드'적인 인공신경망의 한 종류이다. 세가지 주요 계층인 콘볼루션 계층, 풀링 계층, 완전 연결 계층으로 이루어져 있습니다.

##### 콘볼루션 계층

CNN의 핵심 빌딩 블록이며 대부분의 계산이 이루어지는 곳입니다. 입력 데이터, 필터 및 기능 맵과 같은 몇 가지 구성 요소가 필요합니다. 입력이 3d의 픽셀 매트릭스로 구성된 컬러 이미지라고 가정한다면 이는 입력이 이미지의 RGB에 해당하는 높이, 너비 및 깊이의 3차원을 갖게 됨을 의미합니다. 또한 이미지의 수용 필드를 가로 질러 이동하여 기능이 있는지 확인하는 기능 감지기(커널 또는 필터라고도 함)가 있습니다. 이 프로세스를 컨볼루션이라고 합니다.

기능 감지기는 이미지의 일부를 나타내는 가중치의 2차원(2-D) 배열입니다. 크기는 다양할 수 있지만 필터 크기는 일반적으로 3x3 행렬이며, 이것은 또한 수용 필드의 크기를 결정합니다. 그런 다음 필터가 이미지 영역에 적용되고 입력 픽셀과 필터 사이에서 점곱이 계산됩니다. 그런 다음 이 점곱은 출력 배열로 공급됩니다. 그런 다음 필터가 한 칸씩 이동하면서 커널이 이미지 전체를 스캔할 때까지 이 프로세스를 반복합니다. 입력과 필터로부터의 연속된 점곱으로 이루어진 최종 출력을 기능 맵, 활성화 맵 또는 컨볼빙된 기능이라고 합니다.

---

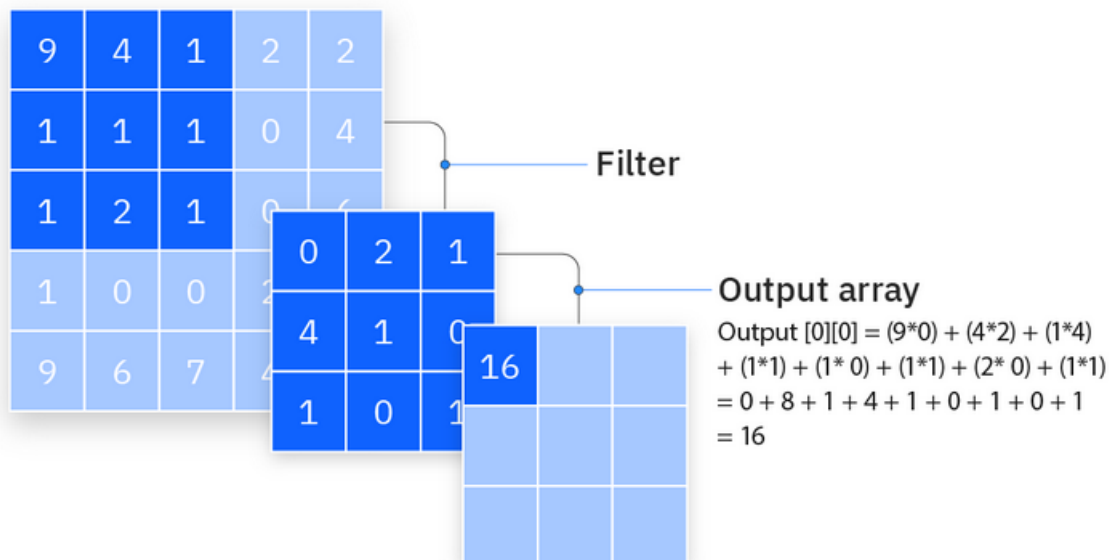
<sup>1</sup> <https://wikidocs.net/176148>, 순방향, 단방향

특징 감지기의 가중치는 이미지를 가로질러 이동할 때 고정된 상태로 유지되며, 이를 파라미터 공유라고도 합니다. 가중치 값과 같은 일부 파라미터는 역전파 및 경사하강법 과정을 통해 훈련 중에 조정됩니다. 그러나 신경망 학습이 시작되기 전에 설정해야 하는 아웃풋의 볼륨 크기에 영향을 주는 세 가지 하이퍼 파라미터가 있습니다. 여기에는 다음이 포함됩니다.

1. **필터의 개수**는 아웃풋의 깊이에 영향을 줍니다. 예를 들어, 세 개의 서로 다른 필터는 세 개의 서로 다른 기능 맵을 생성하여 3의 깊이를 만듭니다.
2. **스트라이드**는 커널이 인풋 행렬 위로 이동하는 거리 또는 픽셀 수입니다. 스트라이드 값이 2 이상인 경우는 드물지만 스트라이드가 클수록 아웃풋이 작아집니다.
3. **제로 패딩**은 일반적으로 필터가 인풋 이미지에 맞지 않을 때 사용됩니다. 이렇게 하면 인풋 행렬을 벗어나는 모든 요소가 0으로 설정되어 더 크거나 동일한 크기의 아웃풋이 생성됩니다. 패딩에는 세 가지 유형이 있습니다.
  - **유효한 패딩**: 패딩 없음이라고도 합니다. 이 경우 차원이 정렬되지 않으면 마지막 합성곱이 삭제됩니다.
  - **동일한 패딩**: 이 패딩은 아웃풋 계층이 인풋 계층과 동일한 크기를 갖도록 보장합니다.
  - **전체 패딩**: 이 유형의 패딩은 인풋 테두리에 0을 추가하여 아웃풋의 크기를 늘립니다.

각 컨볼루션 작업 후 CNN은 기능 맵에 ReLU(Rectified Linear Unit) 변환을 적용하여 모델에 비선형성을 도입합니다.

Input image



풀링 계층과 완전 연결 계층은 분량상 생략하겠습니다.

## 추천 시스템(Recommendation Systems)

추천 시스템은 사용자의 선호도를 분석하여 콘텐츠를 추천하는 시스템으로, 행렬 분해(Matrix Factorization), 신경망 기반 협업 필터링, 강화 학습 등이 자주 사용됩니다.

### 행렬 분해(Matrix Factorization)

행렬을 특정한 구조를 가진 다른 행렬의 곱으로 나타내는 것을 의미한다. 행렬 분해는 선형 방정식의 해를 구하거나, 행렬 계산을 효율적으로 하거나, 행렬의 특정 구조를 밝히는 등의 목적으로 사용된다.

이 외에도 아이템 간의 유사도를 계산하여 가장 유사한 항목을 추천하는 방식을 추천 알고리즘에서 사용합니다. 이를 위한 수학적 이론(원리)으로는 유클리드 거리, 코사인 유사도 등을 이용합니다. 각 사용자의 프로필이나 아이템의 속성을 벡터화 하여 가장 가까운 벡터를 찾습니다.

패턴 인식에서 K-최근접 이웃 알고리즘(줄여서 k-NN)은 분류나 회귀에 사용되는 비모수 방식이다. 두 경우 모두 입력이 특징 공간 내 k 개의 가장 가까운 훈련 데이터로 구성되어 있다. 출력은 k-NN 이 분류로 사용되었는지 또는 회귀로 사용되었는지에 따라 다르다.

k-NN 은 함수가 오직 지역적으로 근사하고 모든 계산이 분류될 때까지 연기되는 인스턴스 기반 학습 또는 게으른 학습의 일종이다. k-NN 알고리즘은 가장 간단한 기계학습 알고리즘에 속한다.

## 출처

### - 컴퓨터 비전

<https://www.ibm.com/kr-ko/topics/computer-vision>

[https://ko.wikipedia.org/wiki/합성곱\\_신경망](https://ko.wikipedia.org/wiki/합성곱_신경망)

<https://wikidocs.net/176148>

<https://www.ibm.com/kr-ko/topics/convolutional-neural-networks>

### - 추천 시스템

<https://tech.kakao.com/posts/463>

[https://ko.wikipedia.org/wiki/행렬\\_분해](https://ko.wikipedia.org/wiki/행렬_분해)

<https://vivi-world.tistory.com/136>

[https://ko.wikipedia.org/wiki/K-최근접\\_이웃\\_알고리즘](https://ko.wikipedia.org/wiki/K-최근접_이웃_알고리즘)