

Chapter 9. 합성곱 신경망

인공지능학과 오혜주

9.1 The Convolution Operation 합성곱 연산

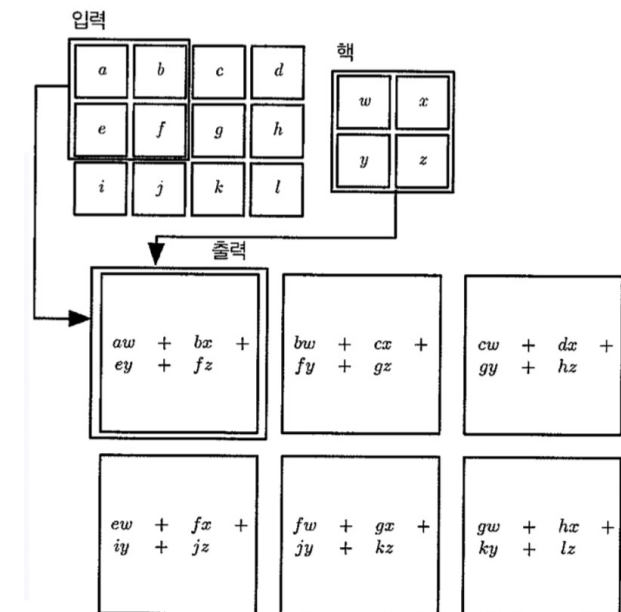
Convolution(합성곱) : 실수값을 받는 두 함수에 관한 연산

→ 두 함수를 합성하는 연산 $s(t) = (x * w)(t)$. x : input(입력), w : kernel(핵)

1) 한 번에 여러 축에 적용

2) 가환적

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n).$$



9.2 Motivation

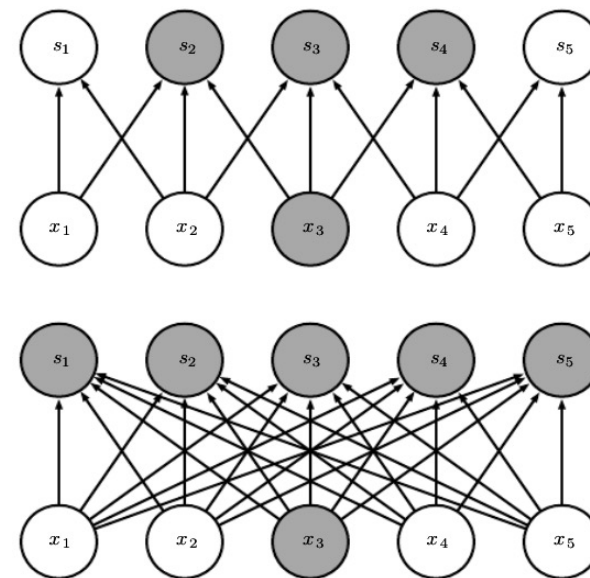
1) Sparse Interaction (희소 상호작용)

- kernel이 input보다 작기 때문

➔ 메모리 사용량 감소, 통계적 효율성 증가, 연산 수 감소

- 합성곱을 사용하여 출력의 연결 수를 k 로 제한

➔ 매개변수 $k*m$ 개, 실행시간 $O(k*n)$

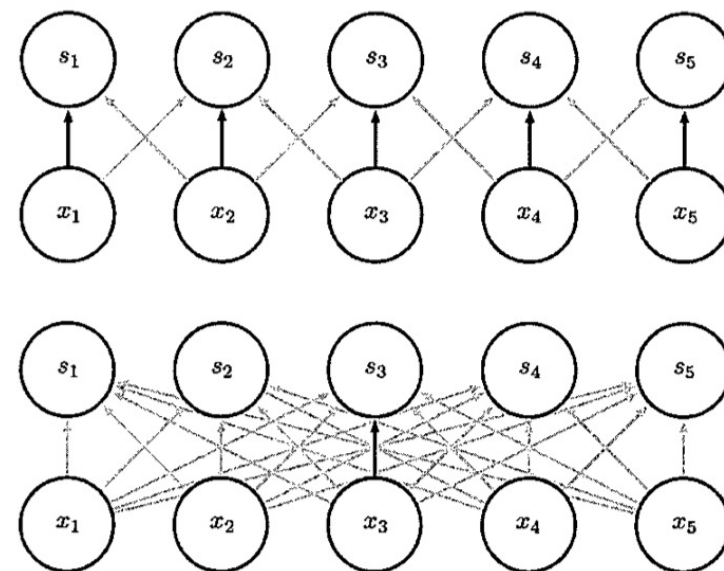


9.2 Motivation

2) Parameter Sharing (매개변수 공유)

: 둘 이상의 함수에 같은 매개변수를 사용

- kernel의 각 성분은 입력의 모든 곳에 쓰임



3) Equivariance to translation (이동에 대한 등변성)

- 등변성 : 입력이 변하면 출력도 같은 방식으로 변함 $f(g(x)) = g(f(x))$

9.3 풀링

- 합성곱 신경망의 한 층은 세 단계로 작동, 그 중 세번째 작업
- 특정 위치에서의 신경망의 출력을 근처 출력들의 요약통계량으로 대체
ex) max pooling (최댓값 풀링)
- Invariance to Translation (이동에 대한 불변성)
 - ➔ 특징의 존재 여부 자체가 중요할 때 유용
 - ➔ 신경망의 통계적 효율성이 극대화

9.4 무한히 강한 사전분포로서의 합성곱과 풀링

사전분포 : 모형의 매개변수들에 관한 하나의 확률분포

합성곱 신경망 → 가중치들에 관해 무한히 강한 사전분포가 있는
신경망

이 관점에서 합성곱 신경망을 바라보았을 때의 통찰
→ 합성곱과 풀링이 과소적합을 일으킬 수 있다.

9.5 기본 합성곱 함수의 여러 변형

신경망에 쓰이는 합성곱의 몇 가지 성질

1. 여러 합성곱을 병렬로 적용하는 하나의 연산을 말함
2. 합성곱에서의 입력이 그냥 실숫값이 아니라, 여러 개의 관측값들로 이루어진 벡터들의 격자무늬형 데이터
3. 다채널 합성곱을 사용, 각 연산의 출력 채널 수와 입력 채널 수가 같을 때만 교환법칙을 만족

9.5 기본 합성곱 함수의 여러 변형

4. downsampling : 핵의 일부 위치를 건너뛰고 출력하는 것

5. 입력에 암묵적으로 0을 채워서 입력 텐서를 더 넓게 만듦

→ valid convolution : 0 채우기를 아예 사용하지 않기

→ same convolution : 출력과 입력이 같은 정도가 되도록 0을 채움

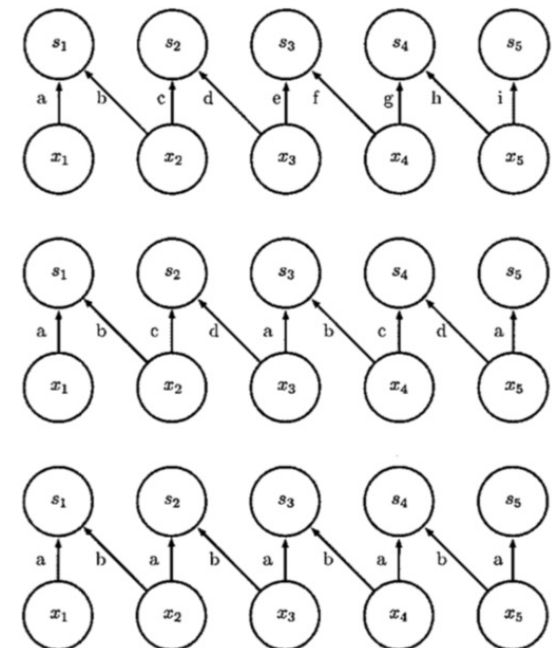
→ full convolution : 모든 픽셀이 k번 영향을 끼치도록 충분히 많은 0을 채움

9.5 기본 합성곱 함수의 여러 변형

6. tiled convolution : 공간을 나아가는 과정에서 순환되는 커널들의 집합을 학습

7. 출력 기울기로 커널의 기울기 계산 가능

8. 합성곱을 하나의 행렬 곱셈으로 서술 가능
이 관점 → 전치행렬을 곱하는 연산



9.6 구조적 출력

- 합성곱 신경망은 분류 문제에서 고차원 구조적 객체를 출력하게 할 수 있음

이때, 출력 평면이 입력 평면보다 작은 문제 발생 가능

- 이미지의 픽셀별 라벨 결정의 전략 :

이미지 표지들을 잠정적으로 추측 => 이웃 픽셀과의 상호작용을 통해 정교하게 다듬기

9.7 자료 형식

표 9.1: 합성곱 신경망에 사용할 수 있는 여러 자료 형식의 예.

	단일 채널	다중 채널
1차원	음향 파형 자료: 유일한 축은 시간에 해당한다. 시간을 이산화하고, 각 시간 단계에서 한 번만 파형(waveform)의 진폭을 측정한다.	골격 애니메이션 자료: 3차원 컴퓨터 렌더링 캐릭터의 골격 애니메이션(skeleton animation)은 시간에 따라 '골격(뼈대)'의 자세를 변경해서 생성한다. 각 시점(time point)에서 캐릭터의 자세는 캐릭터 골격의 각 관절의 각도들로 서술된다. 합성곱 신경망에 공급하는 자료의 각 채널은 한 관절의 한 축의 각도에 해당한다.
2차원	푸리에 변환으로 처리한 음향 자료: 음향 파형 자료를, 행들이 서로 다른 주파수(진동수)들에 대응되고 열들이 서로 다른 시점들에 대응되는 하나의 2차원 텐서로 변환할 수 있다. 시간 축에 대한 합성곱을 적용하면 모형은 시간의 이동에 대한 등변성을 가진다. 주파수 축에 대한 합성곱을 적용하면 모형은 주파수에 대한 등변성을 가지므로, 모형은 같은 선율을 다른 옥타브로 연주한 음향 자료에 대해 동일한(신경망 출력의 높이만 다른) 표현을 산출한다.	색상 이미지 자료: 한 채널은 픽셀의 적색광 세기를, 다른 한 채널은 녹색광 세기를, 나머지 한 채널은 청색광 세기를 담는다. 합성곱 핵은 이미지의 수평축과 수직축 모두를 따라 이동하며, 모형은 두 방향 모두에 대해 이동에 대한 등변성을 가진다.
3차원	부피(입체) 자료: 이런 종류의 자료는 흔히 CT 스캔 같은 의학 영상 기술에서 나온다.	색상 동영상 자료: 한 축은 시간, 다른 한 축은 동영상 프레임의 세로 방향(높이), 나머지 한 축은 동영상 프레임의 가로 방향(너비)에 대응된다.

9.8 효율적인 합성곱 알고리즘

- 분리 가능 : d 차원 커널을 d 개의 벡터들의 외적으로 표현할 수 있을 때, 그러한 커널을 분리 가능이라고 함

이 때 보통의 합성곱 계산은 효율적이지 않음

9.9 무작위 특징 또는 비지도 특징 학습

합성곱 신경망 훈련의 비용을 줄이는 방법

= 지도 훈련을 요구하지 않는 합성곱 커널을 얻는 방법

- 커널들을 무작위로 초기화
- 커널을 사람이 직접 설계
- 비지도 판정조건을 이용해 커널을 직접 학습

9.9 무작위 특징 또는 비지도 특징 학습

- greedy 층별 사전훈련 : 첫 층만 따로 훈련해 모든 특징을 한 번만 추출한 후, 그 특징으로 둘째 층을 따로 훈련

ex) 합성곱 심층 믿음망

9.10 합성곱 신경망의 신경과학적 근거

- 신경과학자 데이비드 허블, 토르스텐 비셀의 연구
 - 고양이 앞에 놓인 이미지에 대해 고양이 뇌의 뉴런들이 반응하는 방식 관찰
 - ➔ 뇌의 기능의 다양한 측면을 특징짓는 데 도움을 줌
 - cartoon view으로 살펴본 뇌의 기능 : 1차 시각피질 = V1
 1. 하나의 공간적 맵 형태로 배치
 2. 다수의 단순세포들로 구성
 3. 다수의 복합세포도 존재

9.10 합성곱 신경망의 신경과학적 근거

- 합성곱 신경망과 포유류의 시각체계의 차이점

1. 사람의 눈은 대체로 해상도가 아주 낮으나 합성곱 신경망은 고해상도 사진 이미지를 입력받음
2. 사람의 시각체계는 다른 감각과 통합되어 있으나 합성곱 신경망은 시각만 다룸
3. 사람의 시각체계는 물체를 인식하는 것 이상의 일을 수행하지만 합성곱 신경망은 아직 응용이 어려움
4. V1은 feedback에 큰 영향을 받지만 신경망 모형의 연구는 부족
5. 순방향 IT 피질 발화속도가 포착하는 정보가 합성곱 신경망이 학습하는 특징과 비슷하긴 하지만, 중간 계산이 비슷한지는 모름

9.11 합성곱 신경망으로 본 심층 학습의 역사

- 합성곱 신경망 : 초기 심층 모형 중 하나
 - 1990년 AT&T 연구진이 수표를 판독하는 합성곱 신경망 개발
 - 합성곱 신경망에 기초한 모형들이 다른 기계학습 대회와 컴퓨터 시각 대회에서 우승
- 역전파로 훈련해서 잘 작동하게 만들 수 있었던 최초의 심층 신경망 중 하나
- 합성곱 신경망은 명백한 격자 형태의 위상구조를 가진 자료에 맞게 신경망을 특수화하는, 그 자료를 아주 큰 규모로 확장할 수 있는 수단 제공