

딥러닝의 통계적이해

7강. 합성곱 신경망의 기초(2)

1. 합성곱 신경망의 기본구조
2. 합성곱 신경망 모형의 발전

한국방송통신대 이금희 교수

오늘의 **학습목표**

1. 합성곱 신경망의 기본구조를 이해한다.
2. 합성곱 신경망 모형의 발전을 이해한다.

1. 합성곱 신경망의 기본구조

1. 합성곱 신경망의 기본구조

합성곱 신경망

- ◆ 합성곱 신경망(Convolutional Neural Network, CNN)의 층은 높이, 너비와 깊이로 구성된 3차원 구조
 - 깊이는 채널(channel) 또는 특징 맵(feature map)에 해당

1. 합성곱 신경망의 기본구조

합성곱 신경망의 구조와 학습

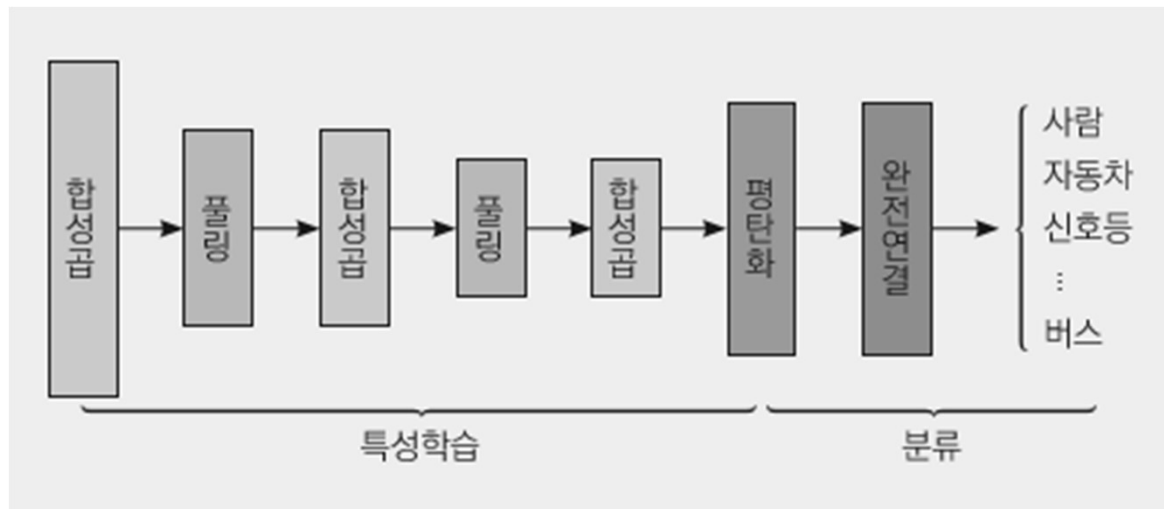
- ◆ 합성곱 신경망에서 합성곱 연산으로 특성(feature)을 파악하고 완전연결 신경망을 연결하여 분류 및 예측
 - 30×30 이미지 데이터를 900개로 펼친 후 합성곱 연산, 활성화 함수, 풀링을 연속적 적용, 특성학습



1. 합성곱 신경망의 기본구조

합성곱 신경망의 구조

- ◆ 합성곱층(활성화 함수 포함) + 풀링층 + 합성곱층 + 풀링층 + 완전연결 신경망(은닉층들 + 출력층)으로 구성



1. 합성곱 신경망의 기본구조

합성곱 신경망의 학습

- ◆ 합성곱 신경망에서 풀링층을 제외한 모든 층의 가중치들은 오차역전파법으로 구함
 - 오차역전파법 학습 : 학습 전 합성곱층의 필터들의 값들이 임의의 값 → 학습 후 필터값은 변해서 필터마다 특색 있게 변함
 - 합성곱층 필터의 값들은 데이터를 학습하여 구해짐

1. 합성곱 신경망의 기본구조

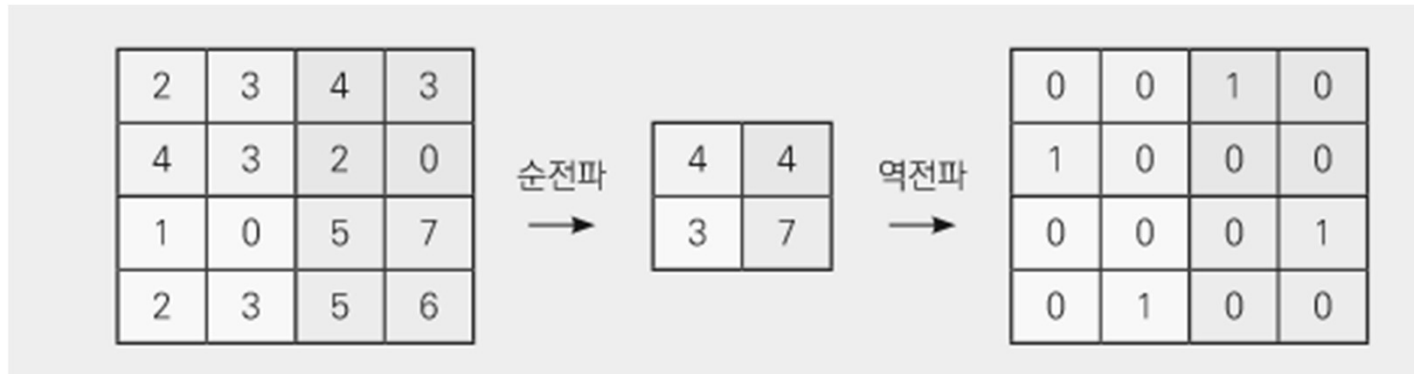
합성곱 신경망의 학습

- ◆ 풀링층 : 순전파는 문제없이 계산, 역전파 계산시 문제 발생
 - 순전파 : 최대 풀링을 통해 4×4 데이터가 2×2 데이터가 됨
 - 역전파 : 4×4 데이터로 업 샘플링(up sampling)
→ 최댓값이 있는 위치는 1, 나머지는 0으로 지정



1. 합성곱 신경망의 기본구조

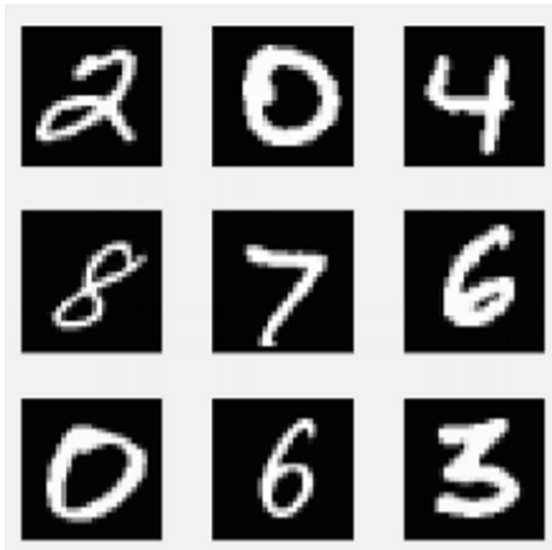
풀링층의 역전파



1. 합성곱 신경망의 기본구조

MNIST 데이터

- ◆ 28×28 크기의 손글씨 흑백 이미지, 60,000개의 훈련데이터와 10,000개의 시험데이터로 구성



1. 합성곱 신경망의 기본구조

완전연결 신경망

- ◆ Tensorflow-Keras로 MNIST 데이터를 1개의 은닉층이 있는 완전연결 신경망으로 적합한 결과

```
-----  
Layer (type)                 Output Shape                 Param #  
-----  
dense_1 (Dense)              (None, 512)                  401920  
-----  
dense_2 (Dense)              (None, 10)                   5130  
-----  
Total params: 407,050  
Trainable params: 407,050  
Non-trainable params: 0  
-----
```

1. 합성곱 신경망의 기본구조

완전연결 신경망

- ◆ 입력 데이터는 $28 \times 28 = 784$ 크기의 이미지
- ◆ 은닉층(dense_1) 1개 : 512개의 뉴런
 - 가중치의 수 : $(28 \times 28 + 1) \times 512 = 401,920$ 개
- ◆ 숫자 10개, 은닉층 512 뉴런에서 출력층까지 연결
 - $(512 + 1) \times 10 = 5,130$ 개의 가중치
- ◆ 총 모수의 수 : 407,050개

1. 합성곱 신경망의 기본구조

합성곱 신경망

- ◆ Tensorflow-Keras로 MNIST 합성곱 신경망 추정
→ 3개 합성곱층(풀링층 포함) 적용 + 완전연결 신경망

| Layer (type) | Output Shape | Param # |
|--------------------------------|--------------------|---------|
| conv2d_1 (Conv2D) | (None, 26, 26, 32) | 320 |
| max_pooling2d_1 (MaxPooling2D) | (None, 13, 13, 32) | 0 |
| conv2d_2 (Conv2D) | (None, 11, 11, 64) | 18496 |
| max_pooling2d_2 (MaxPooling2D) | (None, 5, 5, 64) | 0 |
| conv2d_3 (Conv2D) | (None, 3, 3, 64) | 36928 |
| flatten_1 (Flatten) | (None, 576) | 0 |
| dense_1 (Dense) | (None, 64) | 36928 |
| dense_2 (Dense) | (None, 10) | 650 |

=====

Total params: 93,322
Trainable params: 93,322
Non-trainable params: 0

1. 합성곱 신경망의 기본구조

합성곱 신경망

- ◆ 입력 데이터 : $28 \times 28 = 784$ 크기의 이미지
- ◆ 합성곱 층(conv2d_1) : 3×3 합성곱 필터 → 출력 이미지
크기 : $(28 - 3 + 1) \times (28 - 3 + 1) = 26 \times 26$
 - 필터 32개 : 가중치 수 $(3 \times 3 + 1) \times 32 = 320$ 개
- ◆ 2×2 최대 풀링(max_pooling2d_1) 적용
→ $13 \times 13 \times 32$ 크기 출력데이터

1. 합성곱 신경망의 기본구조

합성곱 신경망

- ◆ 3×3 합성곱 필터(conv2d_2) 64개 적용
→ 데이터 크기가 $11 \times 11 \times 64$
 - 가중치 수 $(3 \times 3 \times 32 + 1) \times 64 = 18,496$ 개
- ◆ 최대풀링(max_pooling2d_2) : 데이터 크기 $5 \times 5 \times 64$
- ◆ 3×3 합성곱 적용 : 데이터 크기 $3 \times 3 \times 64$
 - 가중치 수 $(3 \times 3 \times 64 + 1) \times 64 = 36,928$ 개



1. 합성곱 신경망의 기본구조

합성곱 신경망

- ◆ 데이터 $3 \times 3 \times 64$ 를 평탄화 하여 크기 576의 벡터로 전환
 - 64개 뉴런을 가진 은닉층(dense_1)과 10개 뉴런의 출력층(dense_2)을 추가
 - 합성곱 신경망
- ◆ 은닉층으로 가는 가중치의 수 : $(576 + 1) \times 64 = 36,928$ 개
- ◆ 출력층으로 가는 가중치의 수 : $(64 + 1) \times 10 = 650$ 개
- ◆ 가중치의 수 : 93,322개

1. 합성곱 신경망의 기본구조

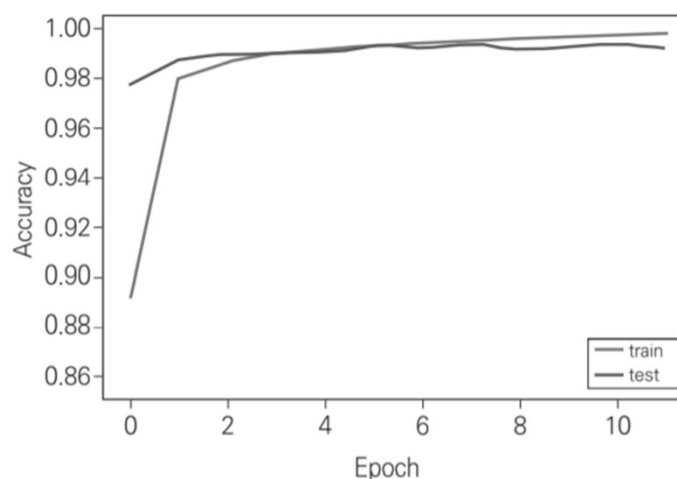
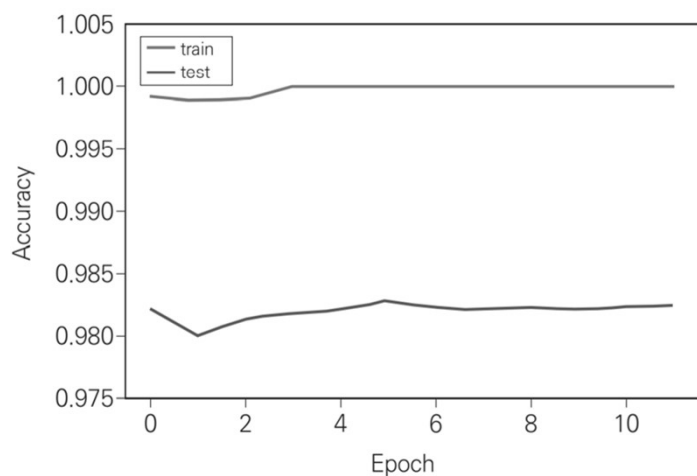
완전연결 신경망과 합성곱 신경망

- ◆ 완전연결 신경망과 합성곱 신경망의 가중치수
 - 완전연결 신경망 : 은닉층의 수 한 개, 가중치 수 407,050개
 - 합성곱 신경망 : 가중치를 공유함으로써 가중치의 수 93,322개(23%)

1. 합성곱 신경망의 기본구조

완전연결 신경망과 합성곱 신경망

- ◆ MNIST 훈련데이터 6만개, 검증(시험)데이터 1만개에 대해
완전연결 신경망과 합성곱 신경망으로 12 에포크 학습
 - 완전연결 신경망 : 정확도 98.24%
 - 합성곱 신경망 : 정확도 99.36%



1. 합성곱 신경망의 기본구조

LeNet-5

- ◆ 1998년 르쿤(Y. Lecun) 연구팀 LeNet-5의 구조
 - 합성곱 필터와 풀링 반복 후 평탄화 : 크기 120 벡터
→ 완전연결망 추가, 출력층 활성화함수 RBF(Radial Basis Function) 이용
 - 최근 마지막 활성화 함수로 RBF보다 소프트맥스 함수 사용
 - 평균 풀링이 아닌 최대풀링을 이용

1. 합성곱 신경망의 기본구조

LeNet-5의 구조

| 층(layer) | | 특성 맵 | 크기 | 필터의 커널 크기 | 스트라이드 | 활성화 함수 |
|----------|-------|------|-------|-----------|-------|--------|
| 입력 | 이미지 | 1 | 32×32 | – | – | – |
| Con1 | 합성곱 | 6 | 28×28 | 5×5 | 1 | tanh |
| P1 | 평균 풀링 | 6 | 14×14 | 2×2 | 2 | tanh |
| Con2 | 합성곱 | 16 | 10×10 | 5×5 | 1 | tanh |
| P2 | 평균 풀링 | 16 | 5×5 | 2×2 | 2 | tanh |
| Con3 | 합성곱 | 120 | 1×1 | 5×5 | 1 | tanh |
| FC | 완전연결 | – | 84 | – | – | tanh |
| 출력 | 완전연결 | – | 10 | – | – | RBF |

2. 합성곱 신경망 모형의 발전

2. 합성곱 신경망 모형의 발전

이미지넷 이미지인식 경진대회

- ◆ 딥러닝 모형(깊은 층 합성곱 신경망)이 이미지넷 이미지 인식 경진대회(ImageNet Large Scale Visual Recognition Challenge)에 참가, 수상 → 딥러닝 모형이 확산

2. 합성곱 신경망 모형의 발전

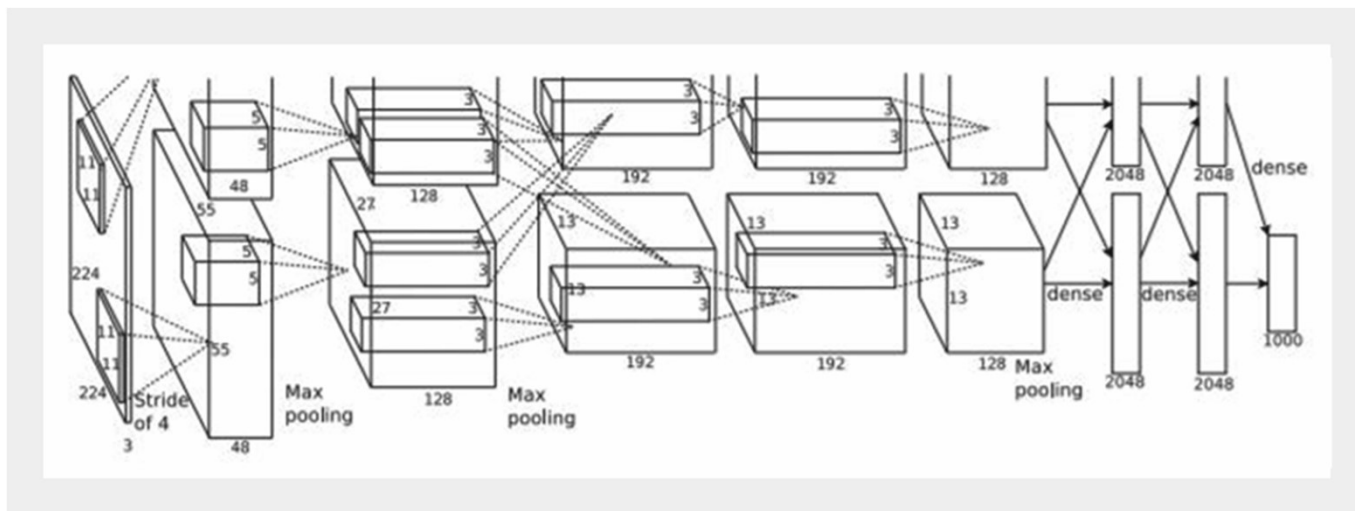
AlexNet

- ◆ 2012년 힌튼(G. Hinton) 팀은 LeNet-5와 유사, 층이 더 깊은 딥러닝 모형 AlexNet을 작성
 - 이미지넷 이미지인식 경진대회 AlexNet은 오분류율 16.4%
 - 기존 머신러닝 방법(오분류율 25.8%) 압도

2. 합성곱 신경망 모형의 발전

AlexNet

- ◆ 구조 : 병렬 컴퓨팅 위해 2개 구조로 나누어져 있음
 - 입력 데이터 크기 $227 \times 227 \times 3$ 의 컬러 이미지
 - 출력층 마지막 층의 크기 : 1,000개



출처 : Krizhevsky et al. (2012)

2. 합성곱 신경망 모형의 발전

AlexNet

- ◆ 가중치의 총수는 62,369,152개 : 합성곱층 가중치 수는 전체의 6%, 완전연결층 가중치 수는 전체 94%

| 층(layer) | | 특성 맵 | 크기 | 필터의 커널 크기 | 스트라이드 | 활성화 함수 |
|----------|------|------|-----------|--------------|-------|-----------|
| 입력 | 이미지 | 1 | 224×224×3 | - | - | - |
| Conv1 | 합성곱 | 96 | 55×55×96 | 11×11 | 4 | ReLU |
| Pool1 | 맥스풀링 | 96 | 27×27×96 | 3×3 | 2 | - |
| Conv2 | 합성곱 | 256 | 27×27×256 | 5×5 | 1 | ReLU |
| Pool2 | 맥스풀링 | 256 | 13×13×256 | 3×3 | 2 | - |
| Conv3 | 합성곱 | 384 | 13×13×384 | 3×3 | 1 | ReLU |
| Conv4 | 합성곱 | 384 | 13×13×384 | 3×3 | 1 | ReLU |
| Conv5 | 합성곱 | 256 | 13×13×256 | 3×3 | 1 | ReLU |
| Pool3 | 맥스풀링 | 256 | 6×6×256 | 3×3 | 2 | ReLU |
| FC6 | 완전연결 | - | 4096 | - | - | ReLU |
| FC7 | 완전연결 | - | 4096 | - | - | ReLU |
| 출력 | 완전연결 | - | 1000 | - | - | Softmax |

2. 합성곱 신경망 모형의 발전

AlexNet의 새로운 시도

- ◆ 활성화 함수 : \tanh 함수 대신 ReLU 함수, 데이터 증강
 - 과대적합 방지 위해 완전연결층에서 드롭아웃 적용
 - 모멘텀(Momentum) 확률적 경사 하강법 이용
- ◆ 7개 신경망의 앙상블을 통해 추가적 성능 향상
 - GPU를 이용한 병렬 계산 → 딥러닝 GPU의 활용도를 높임
 - 지역반응정규화 적용

2. 합성곱 신경망 모형의 발전

ZFNet

- ◆ 2013년 이미지넷 경진대회에서 우승한 딥러닝 모형
 - 신경망 구조는 AlexNet과 동일
 - Conv1에서 ZFNet은 스트라이드 2인 7×7 합성곱 필터 이용 → 합성곱 필터 개수를 AlexNet 보다 2배로 늘림
 - ZFNet은 AlexNet과 동일한 구조에서 일부 수정 → 오분류율 11.7%



2. 합성곱 신경망 모형의 발전

VGGNet

- ◆ VGGNet은 옥스퍼드 대 딥러닝 모형 :
2014년 경진대회에서 GoogLeNet에 이은 2등 → 단순한
구조로 많이 활용
 - VGG16과 VGG19 등
 - 프로그램 구현 용이은 용이하나, 가중치의 수가
지나치게 많아 계산량 많음
- 층을 깊게 쌓는 것이 층별 뉴런의 수를 늘려서 넓게
하는 것보다 효율적임을 보임



2. 합성곱 신경망 모형의 발전

VGGNet

- ◆ VGG16 구조 : 16개의 층(최대 풀링 제외) 구성
 - 3×3 합성곱 필터 + 3×3 최대 풀링을 연속 적용
 - 활성화 함수 : ReLU 함수
 - 합성곱 필터(특성 맵) 수 블록별로 64, 128, 512
 - 가중치 수 : 약 1억3천8백만 개(AlexNet의 2배)
 - 75%가 마지막 합성곱층을 완전연결층에 연결하는 데서 발생

2. 합성곱 신경망 모형의 발전

GoogLeNet

- ◆ GoogLeNet은 2014년 ILSVRC에서 오분류율 6.7%로 1위
 - GoogLeNet은 22층으로 설계된 합성곱 신경망
 - 인셉션(inception) 모듈 이용, 가중치 수 6백만 개
 - 분류시 완전연결망을 이용하지 않음



2. 합성곱 신경망 모형의 발전

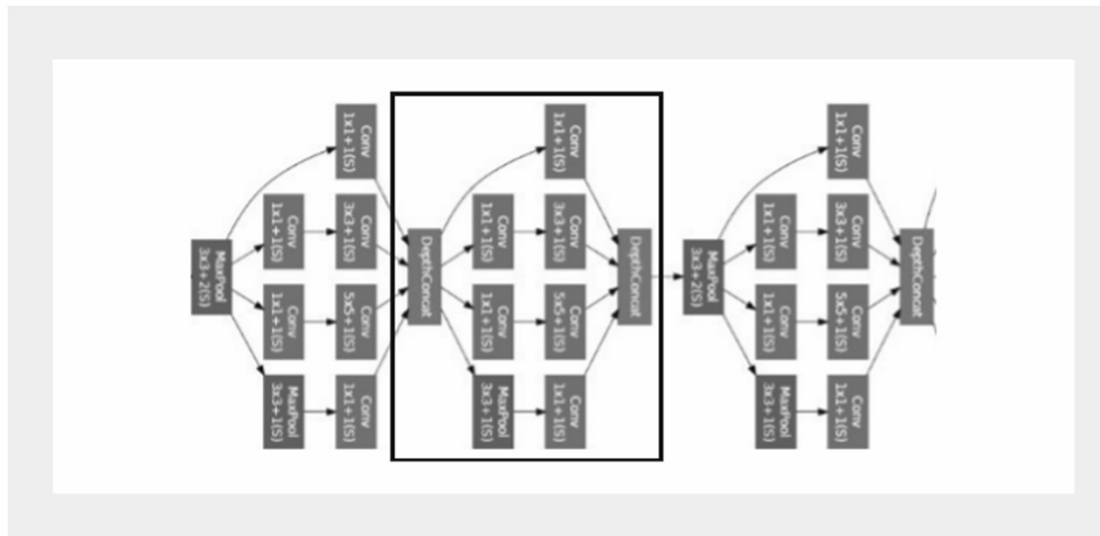
GoogLeNet

- ◆ 입력데이터로부터 중간 결과 2개, 최종결과 1개
 - 손실함수는 최종 결과의 손실과 중간 결과의 손실을 가중
 - 전체 손실함수 = 최종 손실함수 + $0.3 \times$ 중간결과1 손실함수 + $0.3 \times$ 중간결과2 손실함수
 - 오차역전파법으로 가중치를 구함
 - 최종 예측값은 최종 결과만 사용

2. 합성곱 신경망 모형의 발전

GoogLeNet

- ◆ 9개의 인셉션(inception) 모듈이 포함된 신경망
 - 인셉션 : 신경망 속에 신경망을 다시 만든 것



2. 합성곱 신경망 모형의 발전

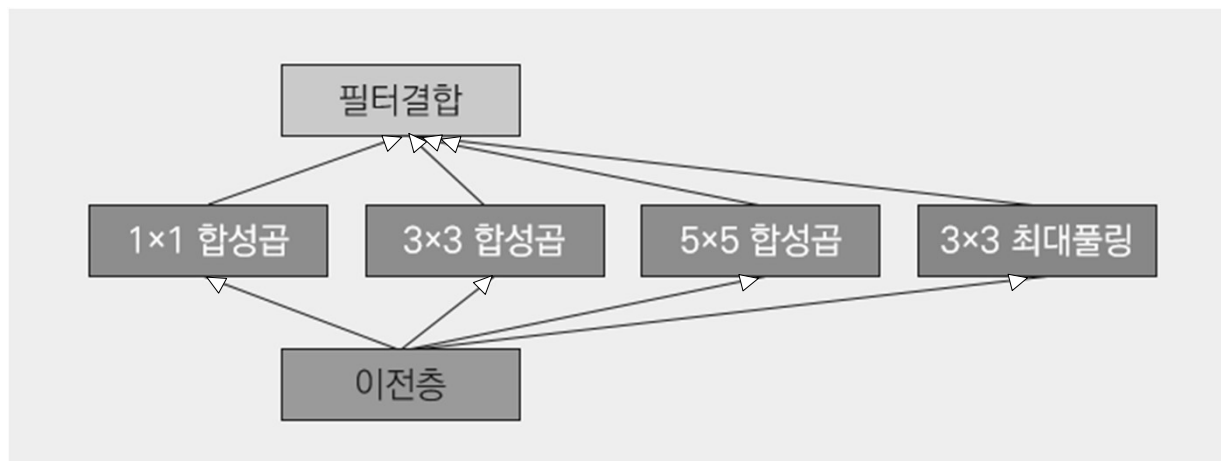
GoogLeNet

- ◆ 인셉션 모듈은 합성곱 필터로 구성된 작은 신경망
 - 크기가 작은 합성곱 필터는 이미지의 세밀한 부분을,
크기가 큰 합성곱 필터는 이미지를 크게 살펴봄
 - 인셉션 모듈은 여러 크기의 합성곱 필터 적용 후 이를 결합
→ 이미지의 여러 특성을 살펴봄

2. 합성곱 신경망 모형의 발전

GoogLeNet

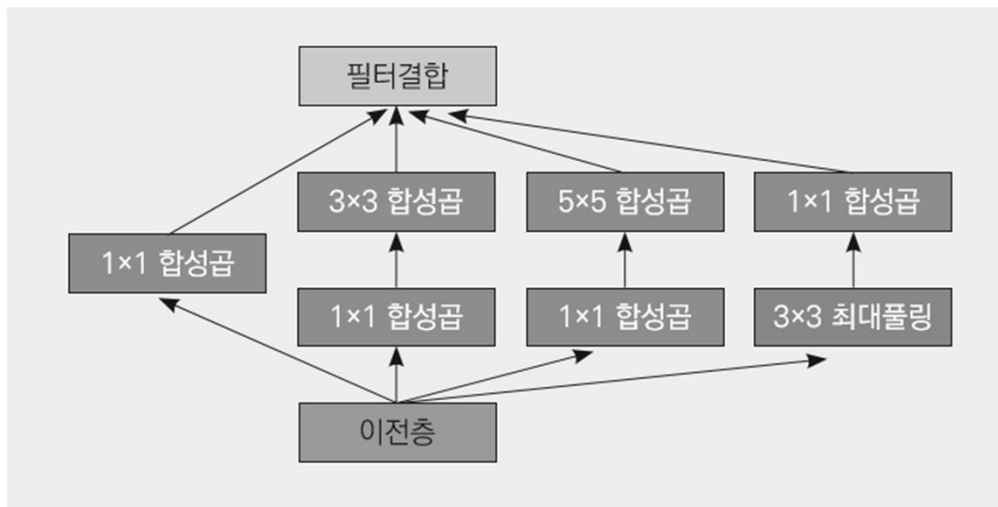
- ◆ 1×1 , 3×3 , 5×5 의 합성곱 필터와 3×3 맥스풀링의 결과를 결합한 인셉션 모듈
→ 여러 합성곱 연산으로 계산량 많음`



2. 합성곱 신경망 모형의 발전

GoogLeNet

- ◆ 1×1 합성곱 필터를 여러 합성곱 필터에 적용한 병목 연산



2. 합성곱 신경망 모형의 발전

GoogLeNet

- ◆ 마지막 출력층 : 완전연결층이 아닌 평균풀링 → 계산량을 줄였음
 - GoogLeNet은 22층의 딥러닝임에도 불구하고
가중치의 수는 600만개 수준(AlexNet의 1/10수준)
 - 오분류율은 6.7%를 기록

2. 합성곱 신경망 모형의 발전

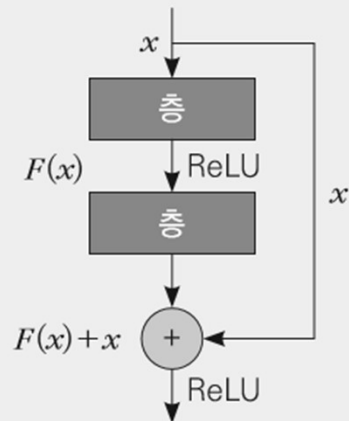
ResNet

- ◆ 히(K. He) 등은 152개 층의 ResNet을 제안 → 이 모형으로 2015년 이미지넷 경진대회 3.6%로 우승(인간 5%)
 - ResNet은 잔차 학습과 스킵 연결을 통해 층을 깊이 쌓으면서 성능이 좋아지는 방법 제안

2. 합성곱 신경망 모형의 발전

ResNet

- ◆ 잔차학습 : 합성곱층을 거치지 않은 x 를 더한 $H(x) = F(x) + x$ 가 다음층에 연결



2. 합성곱 신경망 모형의 발전

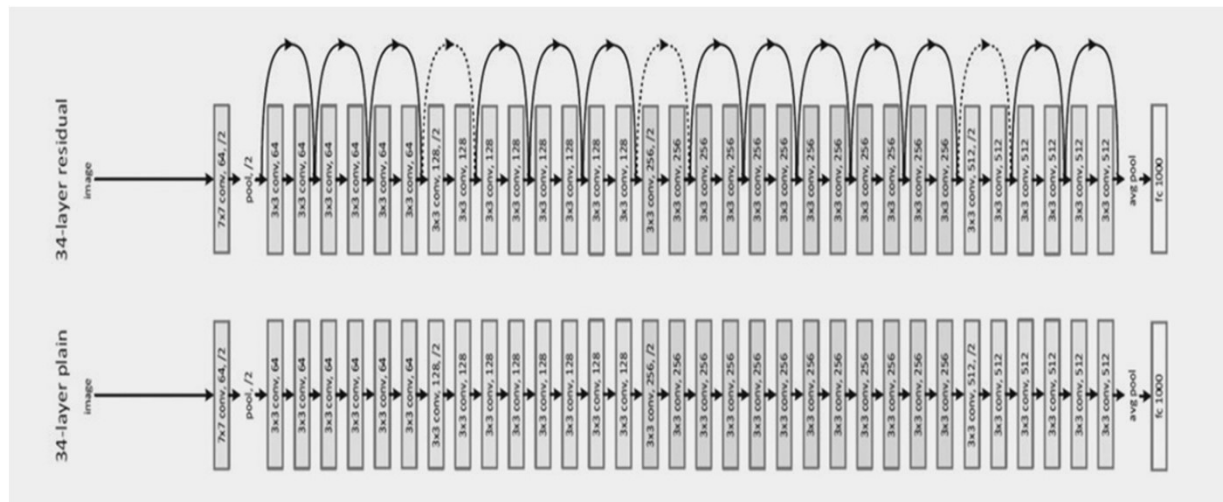
ResNet

- ◆ ResNet의 은닉층 : $h^{(l+2)} := a(z^{(l+2)} + h^{(l)})$
 $h^{(l)}$ 값 일반적으로 0 아님 → 오차역전파 경사소실 없음
- ◆ ResNet은 다른 딥러닝 모형과 달리 층수가 증가해도
잔차학습을 통해 예측오차가 낮아지는 경향

2. 합성곱 신경망 모형의 발전

ResNet

- ◆ 34층 ResNet의 구조 : 잔차학습의 스킵연결을 반복 적용
→ 기본 구조는 GoLeNet 유사



2. 합성곱 신경망 모형의 발전

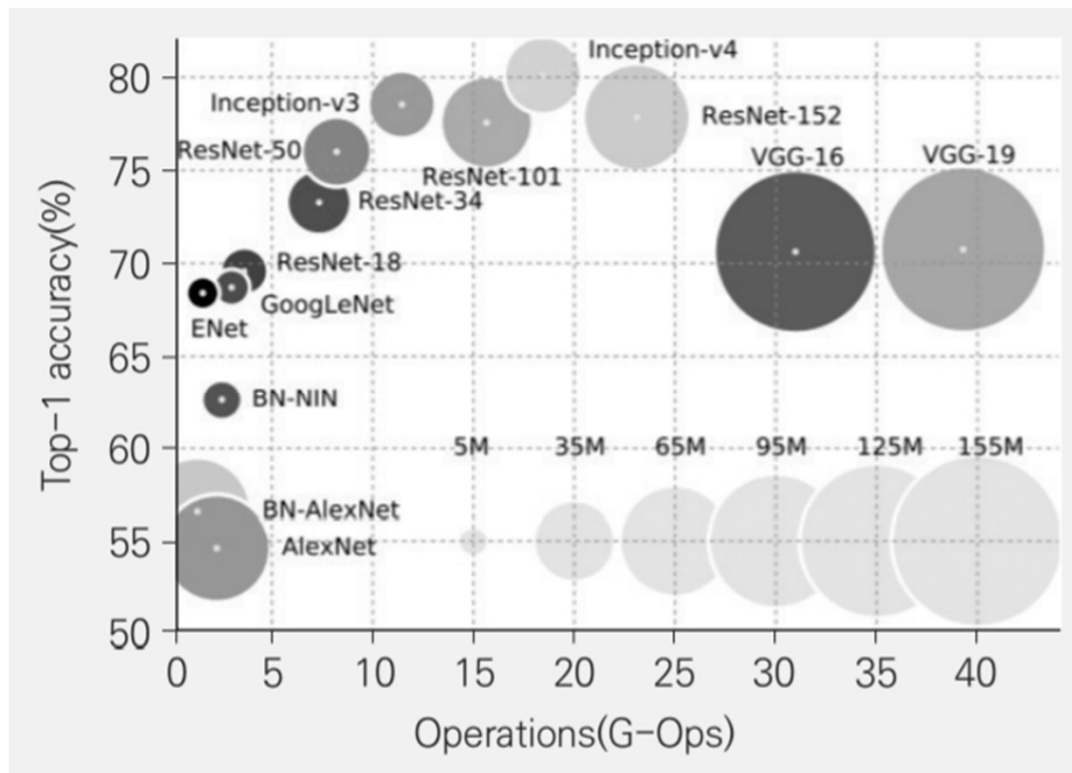
ResNet

◆ ResNet의 특징

- 3×3 합성곱 필터 이용, 완전연결망 이용하지 않음
- 배치정규화 사용, 드롭아웃을 하지 않음
- 스트라이드 2를 이용, 최대풀링을 이용하지 않음
- 스킵 연결은 2개의 합성곱 층마다 진행

2. 합성곱 신경망 모형의 발전

합성곱 딥러닝 모형들의 비교



학습정리

- ✓ 합성곱 신경망은 합성곱 연산을 통해 이미지의 특징을 추출하고 완전연결망의 지도학습으로 이미지를 분류하는 모형이다.
- ✓ 이미지넷 이미지인식 경진대회에서의 대표적인 합성곱 신경망으로는 AlexNet, GoogLeNet, VGG16, ResNet 등이 있다.

딥러닝의 통계적이해

다음시간안내

8강.

합성곱 신경망의 응용