

Machine Learning

5장 CNN의 응용과 이전학습(transfer learning)

고려대학교 통계학과
박유성



Contents

01 사람의 성별을 구분하는 CNN

02 이전학습(Transfer learning) I

03 Image Generator를 이용한 CNN과 자료증대

04 이전학습 II

CNN의 응용

- CNN은 컴퓨터 시각(computer vision)에 뛰어난 성능을 가진 딥러닝 모형이다.
- CNN의 층은 convolution과 pooling (주로 max-pooling)으로 구성되어 있다.
- convolution은 두 가지 중요한 특성을 가지고 있다. 첫 번째 특성은 위치이동불변(translation invariant)으로, 예를 들어 사진의 상단에서 발견된 동물의 귀와 같은 패턴은 위치에 관계없이 동일한 패턴이면 동일하게 인식하는 특성을 말한다.
- convolution의 두 번째 특성은 계층적 패턴인식(hierarchical pattern recognition)이다.

CNN의 응용

- 입력층과 출력층 사이에 다수의 convolution 층이 존재하면 입력층에 가까운 convolution 층은 이미지의 경계선, 모서리, 가장자리 등을 찾아내고
- 이어지는 convolution 층은 이러한 기초 엣지(edge)를 이용하여 귀, 눈, 입 등의 패턴을 찾아낸다.
- convolution 층이 출력층에 가까울수록 점차적으로 사람의 모습이나 동물의 모습에 가까워지는 구조를 가지고 있다.

CNN의 응용

- convolution이 적용되는 이미지 자료는 (표본수, 높이, 넓이, 채널수)로 구성되어 있다.
- 컬러이미지이면 RGB (red, green, blue)이므로 채널수는 3이 되며 흑백일 경우 채널 수는 1이 된다.
- 그러므로 CNN 사용되는 하나의 표본은 3D 텐서이며
- 이러한 자료에 적용되는 kernel window (3,3) or (5,5)를 주로 사용
- max-pooling(kernel window)는 보통 (2,2) 또는 (3,3)으로 픽셀의 크기를 1/2 또는 1/3로 줄임.
- 정보손실 방지를 위해 channe수는 증가시킨다(일반적으로 32로 출발하여 2배수)

CNN의 응용

제 5장의 내용

1. 이미지 자료의 자료준비과정을 논의하고
2. 충분한 이미지 자료가 없을 때 자료를 증대(augment)하거나 기존의 빅데이터로 이미 구축된 모형을 이용하는 방법을 논의한다.
3. 자료의 사전준비과정은 ImageDatGenerator 클래스를 이용해야 할 경우와 그렇지 않은 경우로 나누어 각각 논의할 것이다.
4. 1,000개의 이미지를 가지고 있는 140만개 이미지 자료(ImageNet data set)에 적용하여 구축한 VGG, ResNet, Inception, Xception 모형 중 VGG16 모형을 이용하여 우리가 보유하고 있는 작은 규모의 이미지 자료를 어떻게 분류하는지를 논의하게 될 것이다.

01 사람의 성별을 구분하는 CNN

- Convolution과 pooling을 이용하여 남녀 이미지 사진의 성별을 구분
- 자료가 2,000개에 불과하므로 과대적합이 발생하게 된다.
- 이는 기존의 빅데이터로 이미 학습된 CNN모형으로부터의 이전학습 필요성을 보여줌.
- CNN모형의 아키텍처는 입력층→convolution층→출력층으로 구성
- 입력층에는 3D텐서 자료가 입력되고 convolution 층에서는 convolution과 pooling을 반복적으로 사용하여 이미지를 줄이고 채널수는 늘린다.
- 최종적으로 MLP를 이용하여 노드가 1개인 1D 텐서자료를 출력하여 출력층에 입력으로 전달한다.
- 출력층은 이항분류이면 노드가 1개인 0D 텐서를 출력하고 다항분류이면 노드가 분류개수 만큼 인 1D 텐서를 출력한다.

02 이전학습 I (transfer Learning)

- 이전학습은 빅데이터를 이용하여 이미 학습된 딥러닝 모델을 이용하여 소규모 자료의 분석성능을 높이고자 할 때 사용하는 학습방법이다.
- 특히 CNN의 경우, 보유하고 있는 소규모 이미지자료와 이미 구축된 CNN 모델과 큰 관련이 없더라도 예상외의 성과를 이룰 수 있다.
- 근본적인 이유는 입력층에 가까운 convolution 층은 이미지의 엣지등 일반화가 가능한 기초적인 이미지 특성을 추출하기 때문이다.
- 이전학습은 출력층 직전까지의 모든 convolution층을 이용하고 MLP 층을 새롭게 도입하여 MLP층부터 출력층까지 새롭게 학습시키는 경우
- 출력층 이전의 일부 convolution 층부터 출력층까지 새롭게 학습시키는 경우로 구분할 수 있다.
- 이 절에서는 비교적 간단한 전자의 경우를 다루고자 한다.

03 Image Generator를 이용한 CNN과 자료증대(data augmentation)

- 개와 고양이의 이미지 자료를 이용하여 개와 고양이를 분류하는 CNN 모형
- 다음과 같이 다섯 단계를 거쳐 자료분석의 성능을 향상시키고자 한다.
- 첫 번째로 2,000개의 개와 고양이(각각 1,000개) 이미지 자료만을 이용하여 CNN 모형을 만들고,
- 두 번째로 자료증대를 통해 성능개선을 논의할 것이다.
- 세 번째로 이전학습으로 이미지자료 식별성능을 향상시킨 후,
- 네 번째로 자료증대와 결합하고,
- 다섯 번째로 VGG16 아키텍처의 일부 convolution 층을 재 학습하는 미세조정 (fine-tuning)을 논의할 것이다.

04 이전학습 II

- 우리가 가지고 있는 2,000개의 개와 고양이 식별성능을 향상시키기 위해
- 이미 사용한 VGG16에 이전학습을 적용하고 미세조정(fine-tuning) 방법에 대해 논의하고자 한다.
- 이전학습은 이미 논의한 바와 같이, 2,000개의 개와 고양이 이미지를 이미 추정(학습)된 모수를 가진 VGG16에 단순 적용한 후 산출된 convolution 출력값을 입력자료로 이용한 바 있다.
- 또 다른 방법은 VGG16 모형에서 추정된 모수의 일부를 고정(freeze)하고 VGG16모형에 새로운 히든층(convolution 층 또는 MLP 층)을 추가하여 추가된 은닉층의 모수를 추정하는 이전학습모형이 있다.
- 앞의 이전학습과 비교하여 이러한 이전학습의 장점은 자료증대기법을 사용할 수 있다는데 있다.

Q & A