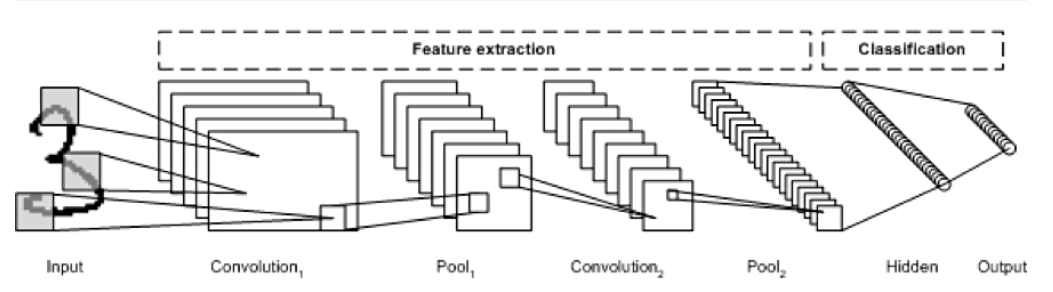
고급 소프트웨어 실습I 8주차 과제

20150325 김정수

1. Convolutional Neural Network에 대해 Convolution과 Pooling에 중점을 두고 설명하세요(1page)

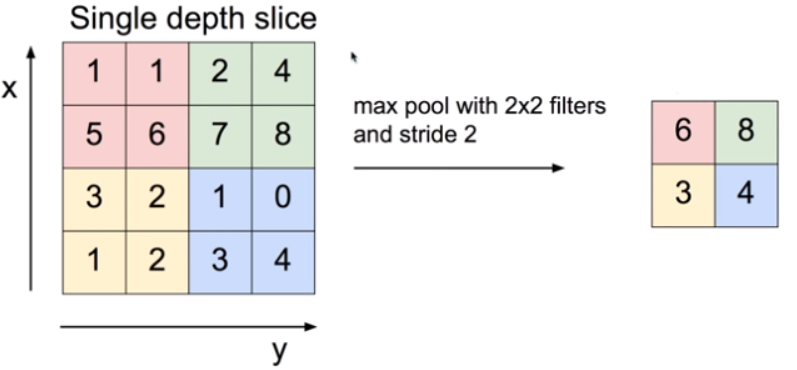


Convolutional neural network(이하 cnn)은, convolution이라는 작업이 들어가는 nn(신경망)을 의미한다. Cnn은 convolution layer와 pooling layer을 겹겹이 쌓는 형태로 구성되며, 각각은 ‘입력 데이터에 필터를 적용해, 활성화 함수를 적용’하거나, ‘추출한 이미지의 특징을 모으고 강화’하는 역할을 맡는다. 따라서, cnn을 이해하는 데에는 convolution과 pool의 의미를 파악하는 것이 중요하다고 할 수 있다.

그렇다면 convolution이 의미하는 바는 무엇인가? 다음과 같은 예를 들어보자. 7\*7사이즈의 이미지가 있다. 이를 3\*3사이즈의 필터를 통해 그 이미지의 데이터를 추출한다고 가정하자. 그렇다면 이 필터를 (1,1)부터 (3,3)까지에 맞춰놓고, 옮기는 양(=stride)를 1로 놓으면, ((5,5),(7,7))을 향해 필터를 옮겨가며 이미지를 훑을 수 있을 것이다. 이 과정을 필터의 아랫부분이 이미지의 끝에 다다를 때까지, 총 5\*5=25번을 하게 된다.

이 때, output size는 (N-F)/stride+1(단 N은 이미지, F는 필터) = (7-3)/1+1=5가 나온다. 이러한 output이 곧 convolution layer이다. 위의 과정에서 필터 1개를 통해, layer을 1개 획득하였다. 만일 여러 개의 필터를 두고, 여러 개의 레이어를 획득한다면 이것이 곧 convolution layers가 될 것이다. 가령 위의 과정에서 layer을 6개 얻었다고 하자. 그럼 이 size는 5이니, 5\*5에 6개, 즉 5\*5\*6의 convolution layers를 얻게 되는 것이다. 이러한 레이어들을 얻는 과정을 곧 convolution이라고 한다. 단, 이러한 작업을 거치면서 이미지의 크기가 작아지고 가장자리의 픽셀의 정보가 사라지는데, 이는 가장자리의 특정값으로 설정된 픽셀을 추가하는 zero-padding을 통해 해결해야 한다.

Pooling은 이렇게 얻어낸 convolution layer들의 resize를 통해 layer을 얻는 작업을 의미한다. 이를 통해 layer의 크기를 줄이거나 특정 데이터를 강조할 수 있는데, 최댓값을 출력하는 max pooling이나 평균값을 출력하는 mean pooling 등 여러가지가 있다. Max pooling을 예로 들면,

이와 같이 stride가 2인 필터에서 (1, 1, 5, 6)에서 1을, 다음으로 (2, 4, 7, 8)에서 8을 뽑는 등 가장 큰 값을 뽑아내는 것이 max pooling이라 할 수 있다. 만약 pooling을 쓰지 않는다면, 사이즈가 96\*96인 이미지를 400개의 필터로 convolution한 8\*8\*400 크기의 convolution layers가 있다면, (96–8+1)\* (96–8+1)=7921개의 feature가, 400개의 layer에 걸쳐 있는 것이니 총 \*89\*400=3,168,400개의 feature가 존재해 그 데이터의 양이 방대해질 것이다. 하지만 이러한 pooling을 통해서라면 이미지의 특징을 살리면서도 데이터를 줄일 수 있다.

최종적으로, 이러한 방식, 즉 여러 층으로 conversion을 한 다음, 비선형 활성함수(ex. Relu)와 pooling을 하여, 원본 이미지에 필터링 기법을 적용한 다음 필터링 된 이미지에 대해 분류 연산이 수행되도록 구성하여 그 결과를 출력하는 것이 cnn기법이라 할 수 있다.