고급 소프트웨어 실습I 8주차 과제

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

Convolutional Neural Network(CNN)은 기존 Fully connected Neural Network와 비교하여각 레이어의 입출력 데이터의 형상을 유지하고, 이미지의 공간 정보를 유지하면서 인접 이미지와의 특징을 효과적으로 인식할 뿐더러 필터를 공유 파라미터로 사용하기 때문에 일반 인공 신경망에 비해 학습되는 파라미터의 수가 매우 적다. CNN은 크게 1) 이미지의 특징을 추출하는 부분과 2) 클래스를 분류하는 부분으로 나누어진다. 그 중 전자인 이미지의 특징을 추출하는 부분은 Convolutional layer와 Pooling layer가 겹쳐 쌓인 형태이다. 그리고 마지막에 이미지 분류를 위한 Fully Connected Layer가 추가된다. 이외에도 이미지의 특징을 추출하는 부분과 이미지를 분류하는 부분 사이에 이미지 형태의 데이터를 배열 형태로 만드는 Flatten Layer도 존재한다. 해당 과제에서는 CNN의 이미지 특징 추출 과정 중 가장 특징적인 요소인 Convolution과 Pooling을 중점적으로 다룰 것이다.

Convolution은 이미지보다 작은 필터를 통하여 이미지의 부분마다의 값을 도출해낸다. 예를 들어 이미지의 크기가 7X7이고, 필터가 3X3이라면 필터를 옮겨가며(stride) 전체 이미지를 훑고난 후 5X5 크기의 layer를 얻는다. 이 때 필터에는 각각의 weight가 존재하여 해당 픽셀의 입력값과 weight를 곱연산을 한다. 그리고 위의 예시를 바탕으로 레이어의 크기를 공식화하면 다음과 같다.

원본 크기 : N X N

필터 크기 : F X F

레이어 크기 : K X K ,

강의 자료에 의하면 필터는 항상 입력의 부피, 즉 폭을 동일하게 유지시킨다.

앞선 Convolutional Layer의 출력 데이터를 입력으로 받는 Pooling Layer는 출력 데이터의 크기를 줄이거나 특정 데이터를 강조하는 용도로 사용된다. 즉 앞서 추출된 특징의 공간적인 위치를 파악하는 데 도움을 준다. 그리고 용도에 따라 적절한 pooling 값을 찾아야하는데, 주로 max pooling 방식을 활용한다. 이는 지정 필터 내 가장 큰 값을 추출하는 방식이다. 이 때의 크기는 필터 사이즈로 나누면 된다. 예를 들어 입력 데이터의 크기가 16 X 12 X 40이고 pooling layer 사이즈가 2X2이라면, 출력 데이터 크기는 (16/2 = 8, 12/2= 6, 40)이다.