고급 소프트웨어 실습I 9주차 과제

20171672 이정원

1. Convolutional Neural Network의 Backpropagation에 대해 설명하시오. (1 page)

기존의 인공신경망처럼 CNN 또한 Backpropagation 알고리즘을 이용하여 학습시킨다. 스트라이드를 1로 가정했을 경우, CNN의 l번째 합성곱 계층의 출력 이미지 X의 k번째 채널에 대한 i번째 행, j번째 열의 요소는 다음과 같이 계산된다.

텍스트이(가) 표시된 사진

자동 생성된 설명- 식 (1)

위의 식에서 W(l)k은 l번째 convolutional 층의 k번째 필터, F(l)H과 F(l)W은 각각 l번째 convolutional 층 필터의 높이와 너비를 의미하고, f는 ReLU와 같은 비선형 활성 함수이다. CNN은 convolution 연산 시에 (i,j)에 (m,n)을 빼는 것이 아니라 (m,n)을 더해가면서 필터를 적용하여 필터 자체를 학습한다.

일반적인 인공신경망에 대해 Backpropagation 알고리즘 기반의 gradient descent method을 이용한 학습에서 인공신경망의 가중치 W는 아래와 같이 계산된다:

텍스트이(가) 표시된 사진

자동 생성된 설명- 식(2)

이 식에서 L은 loss function, η는 learning rate이다. CNN에서는 일반적인 인공신경망의 가중치가 합성곱 계층의 필터에 해당하기 때문에 CNN을 학습시킨다는 것은 필터들의 각 요소를 학습하는 것과 같다. 이를 위해 CNN의 l번째 convolutional 층에서 k번째 필터의 i,j번째 요소인 W(l)k,i,j에 대해 손실 함수 L을 편미분하면 다음과 같다:

텍스트이(가) 표시된 사진

자동 생성된 설명- 식 (3)

식 3을 구하기 위해, 다음 식4, 식5을 계산한다.

텍스트이(가) 표시된 사진

자동 생성된 설명- 식 (4)

텍스트이(가) 표시된 사진

자동 생성된 설명- 식 (5)

식3에 식4, 식5을 대입하여 W(l)k,i,j에 대한 L의 gradient를 계산할 수 있으며, 이를 이용하여 식2와 같이 weight를 업데이트한다. Bias를 의미하는 b 또한 weight와 비슷한 방식으로 업데이트한다.