

Deep learning Report

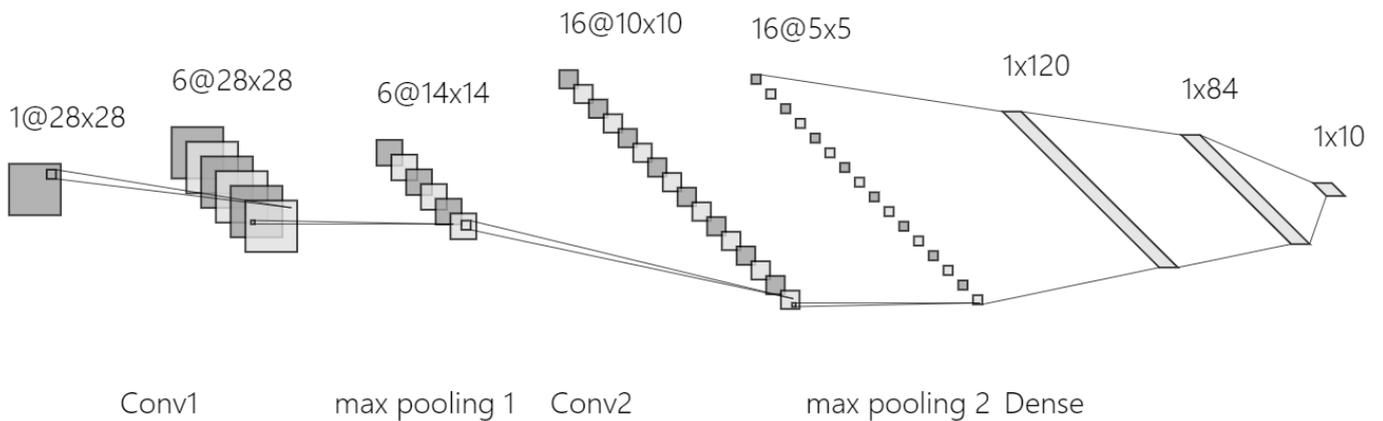
MNIST classification

데이터사이언스학과

20510076 김인조

1. Model Architecture and number of parameter

1-1. Lenet 5



1. Convolution : kernel size = 5, kernel num = 6, padding = same, activation = relu

- Parameters = $(5 \times 5 \times 1) \times 6 = 150$

2. Max pooling : kernel size, stride = 2

3. Convolution : kernel size = 5, kernel num = 16, padding = valid, activation = relu

- Parameters = $(5 \times 5 \times 6) \times 16 = 2,400$

4. Max pooling : kernel size, stride = 2

5. FCN

- Parameters = $(5 \times 5 \times 16) \times 120 = 48,000$
- Parameters = $120 \times 84 = 10,080$
- Parameters = $84 \times 10 = 840$

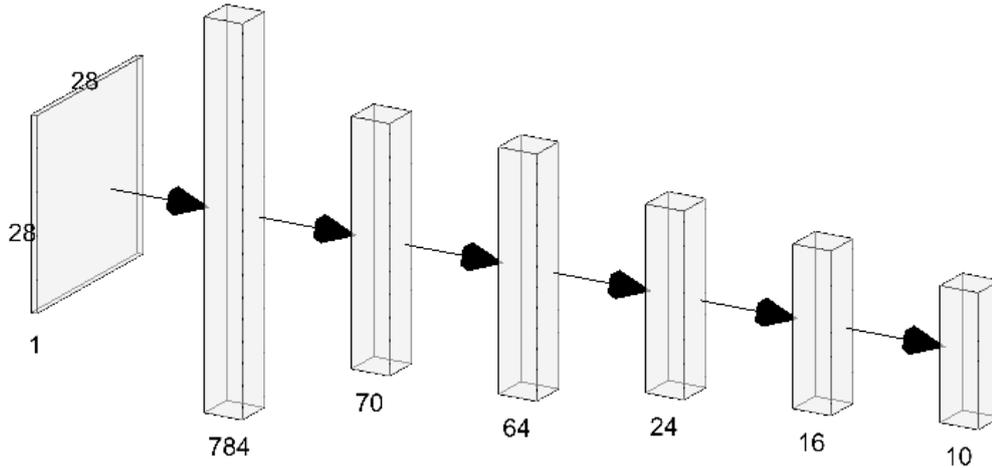
Total parameters = $61,470 \times 2(\text{backpropagation}) = 122,940$

1-2. Regularized Lenet 5

사용한 정규화 방법

1. Cost function에 weight term 추가 → 0.001
2. FCN layer에 dropout 추가 → rate : 0.3

1-3. Custom model



1. input size = 28x28, output size = 70, dropout = 0.5, activation = relu

- Parameters = $(28 \times 28) \times 70 = 54,880$

2. input size = 70, output size = 64, dropout = 0.5, activation = relu

- Parameters = $70 \times 64 = 4,480$

3. input size = 64, output size = 24, dropout = 0.5, activation = relu

- Parameters = $64 \times 24 = 1,536$

4. input size = 24, output size = 16, dropout = 0.5, activation = relu

- Parameters = $24 \times 16 = 384$

5. input size = 16, output size = 10, activation = softmax

- Parameters = $16 \times 10 = 160$

Total parameters = $61,440 \times 2(\text{backpropagation}) = 122,880$

2. Experiment & Result

2-1. 실험 조건

Epoch = 10

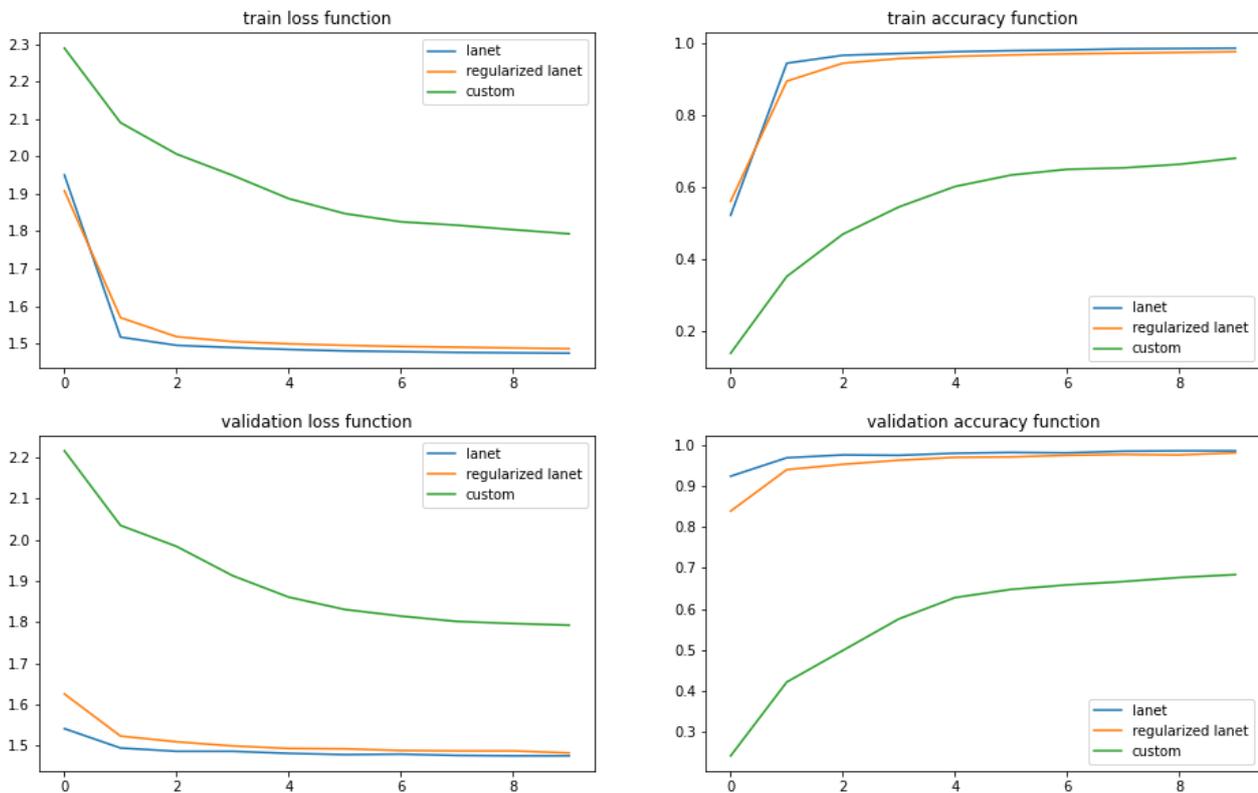
Cost function = cross entropy loss

Optimizer = SGD with learning rate 0.01, momentum 0.9

2-2. 실험 결과

실험 결과 epoch 15 내외에서 saturation됨

본 레포트에서는 각 모델 별 성능 비교를 위해 epoch를 10으로 한 결과를 첨부



실험 결과 바닐라 lenet 모델이 Regularized 모델에 비해 더 높은 정확도를 보임과 동시에 더 빠르게 saturation됨을 확인

Custom 모델은 전체적으로 다른 두 lenet 모델에 비해 performance가 떨어짐을 확인

이는 validation set에서도 완벽히 동일한 결과로 overfitting 없이 적절히 학습되었음을 알 수 있음

처음 제안된 Lenet model은 당시의 MNIST data를 기준으로 99.05%의 정확도를 달성

본 연구의 경우도 vanilla model은 98.6%, regularized model은 98.1%의 정확도를 얻어 비슷한 결과를 얻음

그러나 본 연구는 epoch를 10으로 하였기에 이를 적절하게 조절하면 99%까지도 도달가능할 것으로 생각됨