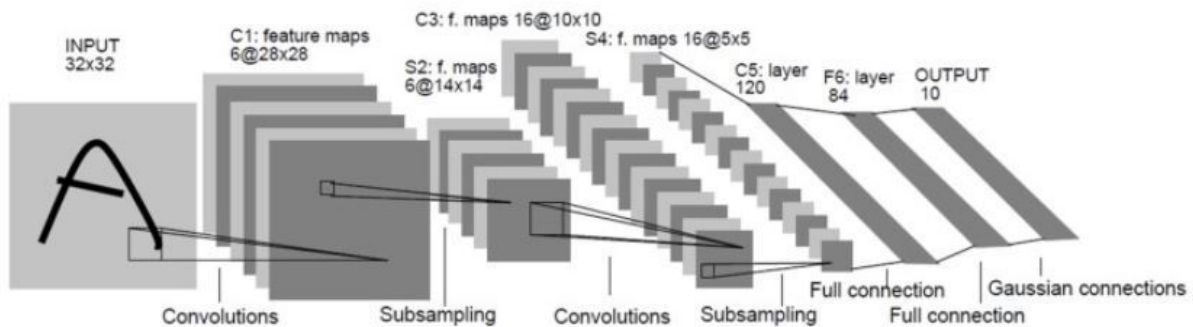


MNIST classification assignment

19512087 데이터사이언스학과 정영상

1. the number of model parameters

(1) LeNet-5 model



Layer (type)	Output Shape	Param
Conv2d - 1	[-1, 6, 28, 28]	156
Tanh - 2	[-1, 6, 28, 28]	0
AvgPool2d - 3	[-1, 6, 14, 14]	0
Conv2d - 4	[-1, 16, 10, 10]	2,416
Tanh - 5	[-1, 16, 10, 10]	0
AvgPool2d - 6	[-1, 16, 5, 5]	0
Linear - 7	[-1, 120]	48,120
Tanh - 8	[-1, 120]	0
Linear - 9	[-1, 84]	10,164
Tanh - 10	[-1, 84]	0
Linear - 11	[-1, 10]	850
Softmax - 12	[-1, 10]	0

Total params: 61,706

Forward/backward pass Total params: 123,412

Training time: 1279.52s

Epoch: 10

Optimizer: SGD

Batch size: 32

(2) Custom model

Layer (type)	Output Shape	Param
Linear – 1	[-1, 64]	50,240
Dropout – 2	[-1, 64]	0
ReLU – 3	[-1, 64]	0
Linear – 4	[-1, 64]	4,160
Dropout – 5	[-1, 64]	0
ReLU – 6	[-1, 64]	0
Linear – 7	[-1, 64]	4,160
Dropout – 8	[-1, 64]	0
ReLU – 9	[-1, 64]	0
Linear – 10	[-1, 32]	2,080
Dropout – 11	[-1, 32]	0
ReLU – 12	[-1, 32]	0
Linear – 13	[-1, 16]	528
Dropout – 14	[-1, 16]	0
ReLU – 15	[-1, 16]	0
Linear – 16	[-1, 16]	272
Dropout – 17	[-1, 16]	0
ReLU – 18	[-1, 16]	0
Linear – 19	[-1, 10]	170
Dropout – 20	[-1, 10]	0
ReLU – 21	[-1, 10]	0
Linear – 22	[-1, 10]	110
Softmax – 23	[-1, 10]	0

Total params: 61,720

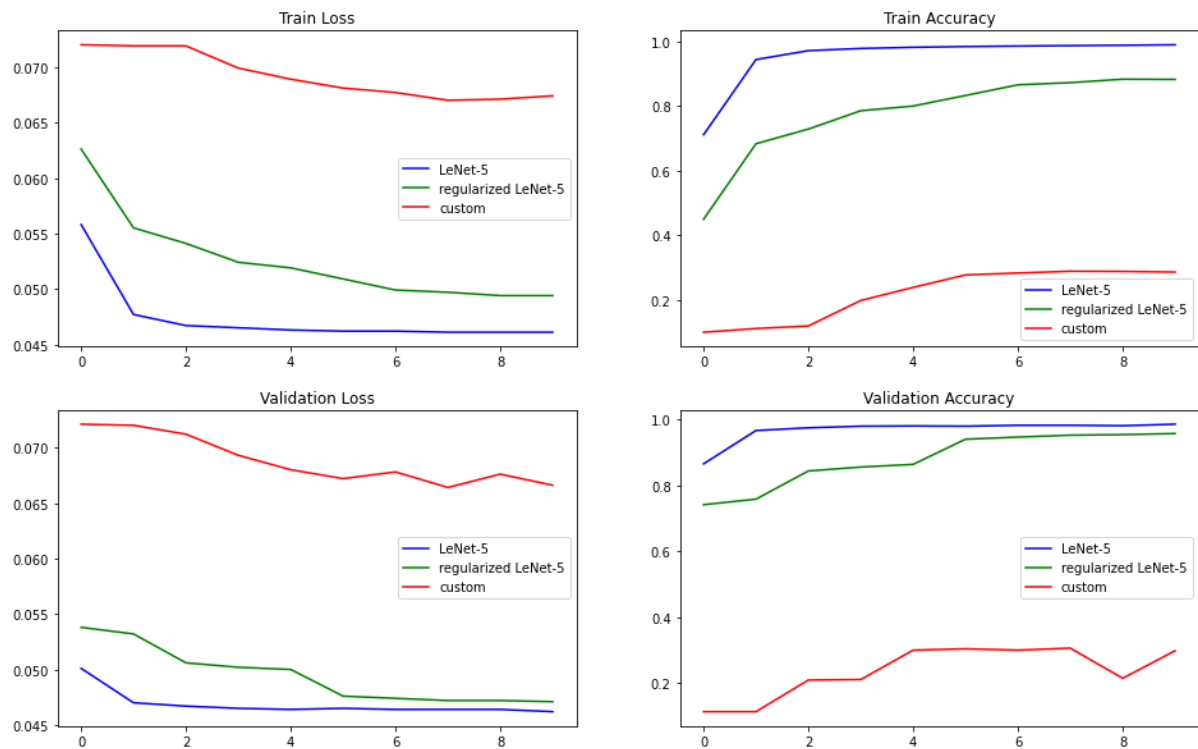
Forward/backward pass Total params: 123,440

Training time: 1209.29s

Epoch: 10

Optimizer: SGD

Batch size: 32



2. Average loss value and accuracy

(1) LeNet-5

Average training accuracy = [0.7111, 0.9439, 0.9715, 0.9783, 0.9819, 0.9842, 0.9858, 0.9870, 0.9882, 0.9897]

Average training loss = [0.0558, 0.0477, 0.0467, 0.0465, 0.0463, 0.0462, 0.0462, 0.0461, 0.0461, 0.0461]

Average validation accuracy = [0.8653, 0.9664, 0.9747, 0.9796, 0.9803, 0.9796, 0.9823, 0.9823, 0.9810, 0.9856]

Average validation loss = [0.0501, 0.0470, 0.0467, 0.0465, 0.0464, 0.0465, 0.0464, 0.0464, 0.0464, 0.0462]

(2) Custom model

Average training accuracy = [0.0996, 0.1110, 0.1190, 0.1978, 0.2384, 0.2771, 0.2829, 0.2885, 0.2879, 0.2859]

Average training loss = [0.0720, 0.0719, 0.0719, 0.0699, 0.0689, 0.0681, 0.0677, 0.0670, 0.0671, 0.0674]

Average validation accuracy = [0.1135, 0.1135, 0.2094, 0.2109, 0.3001, 0.3042, 0.3001, 0.3061, 0.2148, 0.2983]

Average validation loss = [0.0721, 0.0720, 0.0712, 0.0693, 0.0680, 0.0672, 0.0678, 0.0664, 0.0676, 0.0666]

(3) Regularized LeNet-5

Average training accuracy = [0.4493, 0.6832, 0.7287, 0.7854, 0.8000, 0.8326, 0.8657, 0.8724, 0.8831, 0.8824]

Average training loss = [0.0626, 0.0555, 0.0541, 0.0524, 0.0519, 0.0509, 0.0499, 0.0497, 0.0494, 0.0494]

Average validation accuracy = [0.7414, 0.7582, 0.8437, 0.8556, 0.8637, 0.9401, 0.9465, 0.9523, 0.9540, 0.9573]

Average validation loss = [0.0538, 0.0532, 0.0506, 0.0502, 0.0500, 0.0476, 0.0474, 0.0472, 0.0472, 0.0471]

3. Comparison LeNet-5 and Custom MLP

평균 loss값은 custom model이 가장 높게 나왔으며 epoch을 늘릴 경우 regularization model의 loss는 점점 낮아질 것으로 예상된다. 평균 accuracy의 경우 LeNet-5 모델이 가장 높았으며 마찬가지로 epoch을 늘릴 경우 regularization model의 accuracy가 증가할 것으로 보인다.

4. Regularization techniques

(1) Regularized LeNet5

ReLU activation function

Maxpooling

Dropout / $p = 0.5$

L2 Regularization(weight decay / 0.001)

Total params: 61,706

Forward/backward pass Total params: 123,412

Training time: 1470.43s