

Mnist 성능 향상 보고서

ICT 융합공학부 202002514 함상규

본 실험에서 최고 성적은 CNN 구조를 활용하여 얻은 train loss: 0.008452, train acc: 0.988, test acc: 0.986였습니다.

Twolayernet을 기반으로 한 실험에서 최고 결과는 다음과 같습니다:

구조: 4층 Affine 계층

Optimizer: RMSprop

활성화 함수: ReLU

초기화 방법: He 초기화

추가 기법: 학습률 스케줄러, 배치 정규화

Hidden Size: 15

Batch Size: 256

초기 학습률: 0.019

최종 학습률: 0.001

이 설정으로 얻은 최고 성적은 다음과 같습니다:

Learning Rate: 0.0013

Train Accuracy: 0.9900

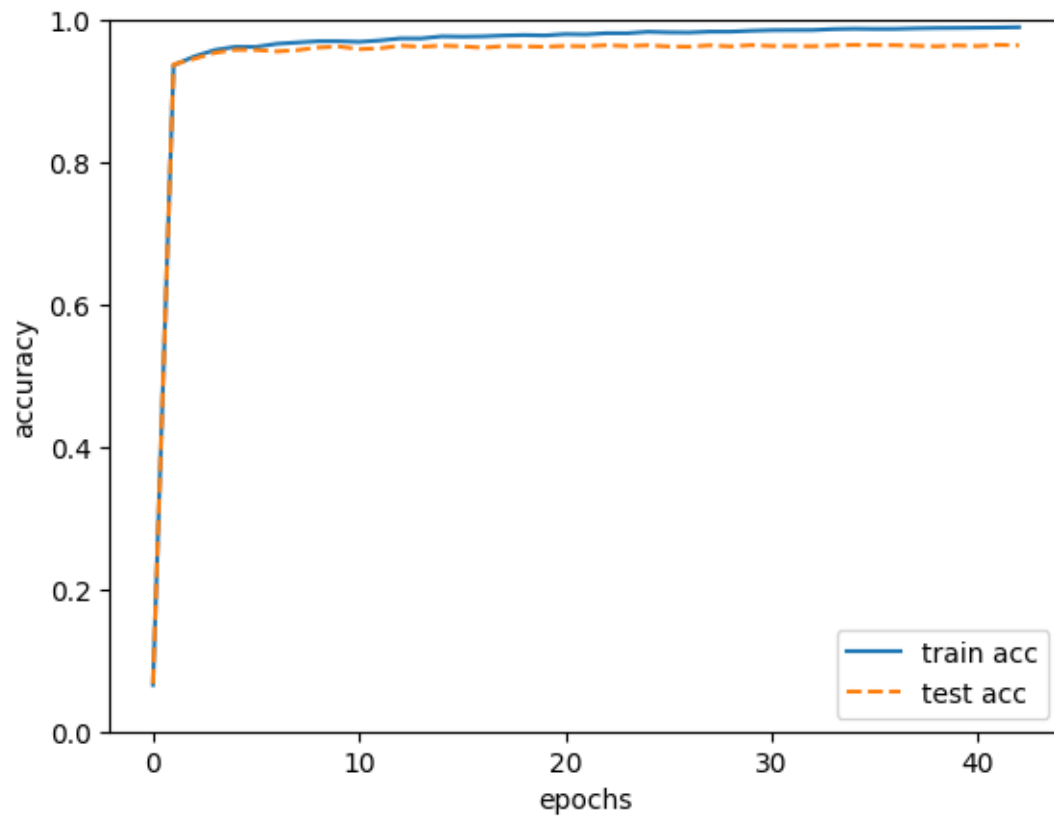
Test Accuracy: 0.9644

하지만 과적합 문제를 보여 L2 규제와 드롭아웃 등을 시도했으나, 과적합은 해결되었으나 전체적인 정확도가 떨어졌습니다. 따라서 배치 정규화와 학습률 스케줄러를 제외한 방법들은 제외하였습니다.

학습 진행 결과

Iteration	Learning Rate	Train Acc	Test Acc
0	0.01900	0.06538	0.06710
234	0.01858	0.93658	0.93760
468	0.01816	0.94878	0.94560
702	0.01774	0.95792	0.95400
936	0.01732	0.96242	0.95830
1170	0.01689	0.96230	0.95800
1404	0.01647	0.96667	0.95600
1638	0.01605	0.96848	0.95780
1872	0.01563	0.97032	0.96140
2106	0.01521	0.97033	0.96290
2340	0.01479	0.96937	0.95920
2574	0.01437	0.97137	0.96030
2808	0.01395	0.97425	0.96410
3042	0.01352	0.97418	0.96250
3276	0.01310	0.97713	0.96410
3510	0.01268	0.97673	0.96310
3744	0.01226	0.97710	0.96120
3978	0.01184	0.97832	0.96350
4212	0.01142	0.97897	0.96300
4446	0.01100	0.97840	0.96250
4680	0.01058	0.98032	0.96370
4914	0.01015	0.97990	0.96310
5148	0.00973	0.98162	0.96480
5382	0.00931	0.98153	0.96340
5616	0.00889	0.98360	0.96460
5850	0.00847	0.98283	0.96290
6084	0.00805	0.98260	0.96260
6318	0.00763	0.98382	0.96450
6552	0.00721	0.98378	0.96300
6786	0.00679	0.98500	0.96480
7020	0.00636	0.98580	0.96360
7254	0.00594	0.98592	0.96350
7488	0.00552	0.98600	0.96330
7722	0.00510	0.98735	0.96440
7956	0.00468	0.98783	0.96510

Iteration	Learning Rate	Train Acc	Test Acc
8190	0.00426	0.98755	0.96490
8424	0.00384	0.98765	0.96470
8658	0.00342	0.98840	0.96390
8892	0.00299	0.98885	0.96310
9126	0.00257	0.98900	0.96440
9360	0.00215	0.98930	0.96350
9594	0.00173	0.98957	0.96520
9828	0.00131	0.99002	0.96440



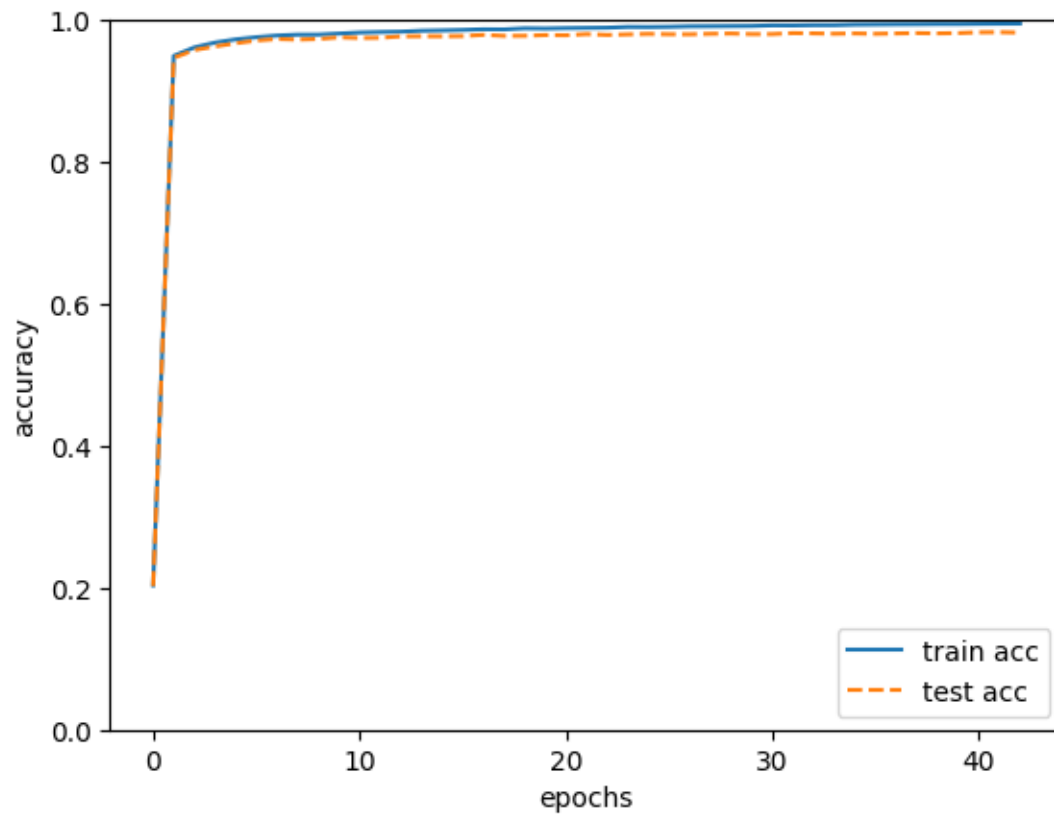
Hidden Size 증가 실험

보너스 점수를 얻기 위해 hidden size를 150으로 증가시킨 후 실험을 진행한 결과, 테스트 정확도가 98.2% 이상의 결과는 얻을 수 있었습니다.

Iteration	Learning Rate	Train Acc	Test Acc
0	0.01900	0.20337	0.20250
234	0.01858	0.94948	0.94680
468	0.01816	0.96145	0.95770
702	0.01774	0.96822	0.96280
936	0.01732	0.97275	0.96730

Iteration	Learning Rate	Train Acc	Test Acc
1170	0.01689	0.97582	0.97110
1404	0.01647	0.97815	0.97330
1638	0.01605	0.97938	0.97250
1872	0.01563	0.97945	0.97330
2106	0.01521	0.98092	0.97560
2340	0.01479	0.98240	0.97470
2574	0.01437	0.98308	0.97500
2808	0.01395	0.98358	0.97700
3042	0.01352	0.98492	0.97720
3276	0.01310	0.98542	0.97710
3510	0.01268	0.98605	0.97740
3744	0.01226	0.98698	0.97910
3978	0.01184	0.98683	0.97770
4212	0.01142	0.98832	0.97780
4446	0.01100	0.98807	0.97880
4680	0.01058	0.98858	0.97850
4914	0.01015	0.98890	0.98020
5148	0.00973	0.98920	0.97900
5382	0.00931	0.99032	0.97980
5616	0.00889	0.99018	0.98050
5850	0.00847	0.99052	0.98000
6084	0.00805	0.99118	0.97990
6318	0.00763	0.99132	0.98060
6552	0.00721	0.99148	0.98100
6786	0.00679	0.99165	0.98020
7020	0.00636	0.99248	0.98020
7254	0.00594	0.99242	0.98160
7488	0.00552	0.99282	0.98120
7722	0.00510	0.99285	0.98070
7956	0.00468	0.99358	0.98110
8190	0.00426	0.99378	0.98060
8424	0.00384	0.99398	0.98110
8658	0.00342	0.99410	0.98160
8892	0.00299	0.99430	0.98120
9126	0.00257	0.99432	0.98160
9360	0.00215	0.99460	0.98270
9594	0.00173	0.99478	0.98290

Iteration	Learning Rate	Train Acc	Test Acc
9828	0.00131	0.99492	0.98230



마지막으로 CNN 구조의 결과는 최대 98.6%의 결과를 보여주었습니다.

Epoch	Train Loss	Train Acc	Test Acc
1	0.21909	0.92500	0.92400
2	0.13883	0.95800	0.95500
3	0.08771	0.96500	0.96100
4	0.06967	0.97500	0.97000
5	0.07947	0.97700	0.97300
6	0.11081	0.98100	0.97100
7	0.02575	0.98200	0.97800
8	0.02083	0.98500	0.98400
9	0.03401	0.98500	0.98400
10	0.02184	0.98300	0.98000
11	0.00845	0.98800	0.98600

