

```
Python 3.7.7 (default, Mar 26 2020, 15:48:22)
Type "copyright", "credits" or "license" for more information.
```

```
IPython 7.13.0 -- An enhanced Interactive Python.
```

```
In [1]:      '/home/administrator/Documents/OU/Deep_learning/project/DeepLearning_PlantDiseases/Scripts/
overfit_study.py'      ='./home/administrator/Documents/OU/Deep_learning/project/DeepLearning_PlantDiseases/Scripts'
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
0           NaN       NaN       NaN       NaN       NaN       NaN
1           NaN       NaN       NaN       NaN       NaN       NaN
2           NaN       NaN       NaN       NaN       NaN       NaN
3           NaN       NaN       NaN       NaN       NaN       NaN
4           NaN       NaN       NaN       NaN       NaN       NaN
5           NaN       NaN       NaN       NaN       NaN       NaN
6           NaN       NaN       NaN       NaN       NaN       NaN
7           NaN       NaN       NaN       NaN       NaN       NaN
8           NaN       NaN       NaN       NaN       NaN       NaN
9           NaN       NaN       NaN       NaN       NaN       NaN
10          NaN       NaN       NaN       NaN       NaN       NaN
11          NaN       NaN       NaN       NaN       NaN       NaN
12          NaN       NaN       NaN       NaN       NaN       NaN
13          NaN       NaN       NaN       NaN       NaN       NaN
14          NaN       NaN       NaN       NaN       NaN       NaN
15          NaN       NaN       NaN       NaN       NaN       NaN
16          NaN       NaN       NaN       NaN       NaN       NaN
17          NaN       NaN       NaN       NaN       NaN       NaN
18          NaN       NaN       NaN       NaN       NaN       NaN
19          NaN       NaN       NaN       NaN       NaN       NaN
20          NaN       NaN       NaN       NaN       NaN       NaN
21          NaN       NaN       NaN       NaN       NaN       NaN
22          NaN       NaN       NaN       NaN       NaN       NaN
23          NaN       NaN       NaN       NaN       NaN       NaN
24          NaN       NaN       NaN       NaN       NaN       NaN
25          NaN       NaN       NaN       NaN       NaN       NaN
26          NaN       NaN       NaN       NaN       NaN       NaN
27          NaN       NaN       NaN       NaN       NaN       NaN
28          NaN       NaN       NaN       NaN       NaN       NaN
29          NaN       NaN       NaN       NaN       NaN       NaN
30          NaN       NaN       NaN       NaN       NaN       NaN
```

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

```
109      NaN      NaN      NaN      NaN      NaN      NaN
110      NaN      NaN      NaN      NaN      NaN      NaN
111      NaN      NaN      NaN      NaN      NaN      NaN
112      NaN      NaN      NaN      NaN      NaN      NaN
113      NaN      NaN      NaN      NaN      NaN      NaN
114      NaN      NaN      NaN      NaN      NaN      NaN
115      NaN      NaN      NaN      NaN      NaN      NaN
116      NaN      NaN      NaN      NaN      NaN      NaN
117      NaN      NaN      NaN      NaN      NaN      NaN
118      NaN      NaN      NaN      NaN      NaN      NaN
119      NaN      NaN      NaN      NaN      NaN      NaN
Directory ovft_intermediate_models already exists
-----
TRAINING from scratch

Targeting alexnet with 39 classes
-----
Resizing input images to max of (224, 224)
Transferring models to GPU(s)
/home/administrator/anaconda3/envs/torch_nvidia_390/lib/python3.7/site-packages/torchvision/transforms/
transforms.py:562: UserWarning: The use of the transforms.RandomSizedCrop transform is deprecated, please use
transforms.RandomResizedCrop instead.
    warnings.warn("The use of the transforms.RandomSizedCrop transform is deprecated, " +
/home/administrator/anaconda3/envs/torch_nvidia_390/lib/python3.7/site-packages/torchvision/transforms/
transforms.py:187: UserWarning: The use of the transforms.Scale transform is deprecated, please use
transforms.Resize instead.
    warnings.warn("The use of the transforms.Scale transform is deprecated, "
Training...
[1,   30] loss: 3.663
[1,   60] loss: 3.654
[1,   90] loss: 3.644
[1,  120] loss: 3.634
[1,  150] loss: 3.623
[1,  180] loss: 3.603
[1,  210] loss: 3.533
[1,  240] loss: 3.392
[1,  270] loss: 3.413
[1,  300] loss: 3.404
[1,  330] loss: 3.371
```

```
[1, 360] loss: 3.373
[1, 390] loss: 3.350
[1, 420] loss: 3.327
[1, 450] loss: 3.293
[1, 480] loss: 3.321
[1, 510] loss: 3.219
[1, 540] loss: 3.186
[1, 570] loss: 3.215
[1, 600] loss: 3.225
[1, 630] loss: 3.185
[1, 660] loss: 3.258
[1, 690] loss: 3.204
[1, 720] loss: 3.142
[1, 750] loss: 3.215
[1, 780] loss: 3.154
[1, 810] loss: 3.126
[1, 840] loss: 3.137
[1, 870] loss: 3.127
[1, 900] loss: 3.110
[1, 930] loss: 3.091
[1, 960] loss: 3.048
[1, 990] loss: 3.096
[1, 1020] loss: 3.063
[1, 1050] loss: 3.051
[1, 1080] loss: 3.037
[1, 1110] loss: 3.031
[1, 1140] loss: 3.001
[1, 1170] loss: 2.914
[1, 1200] loss: 2.958
[1, 1230] loss: 2.971
[1, 1260] loss: 2.909
[1, 1290] loss: 2.760
[1, 1320] loss: 2.851
[1, 1350] loss: 2.924
[1, 1380] loss: 2.889
[1, 1410] loss: 2.970
[1, 1440] loss: 2.860
[1, 1470] loss: 2.847
[1, 1500] loss: 2.879
```

```
[1, 1530] loss: 2.887
[1, 1560] loss: 2.758
[1, 1590] loss: 2.807
[1, 1620] loss: 2.700
[1, 1650] loss: 2.717
[1, 1680] loss: 2.729
[1, 1710] loss: 2.820
[1, 1740] loss: 2.757
[1, 1770] loss: 2.720
[1, 1800] loss: 2.619
[1, 1830] loss: 2.839
[1, 1860] loss: 2.614
[1, 1890] loss: 2.543
[1, 1920] loss: 2.787
[1, 1950] loss: 2.610
[1, 1980] loss: 2.697
[1, 2010] loss: 2.543
[1, 2040] loss: 2.527
[1, 2070] loss: 2.608
[1, 2100] loss: 2.597
[1, 2130] loss: 2.550
[1, 2160] loss: 2.583
[1, 2190] loss: 2.506
ovft_intermediate_models/epoch0_scratch.pt saved.
```

evaluating on training set...

Accuracy on test images: 0.307297

epoch 0 accuracy on train set is: 0.3072973464194838

evaluating on training set takes 65 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN

9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN

48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN

87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.326790

epoch 0 accuracy on test set is: 0.326790258087968

evaluating on test set takes 16 seconds.

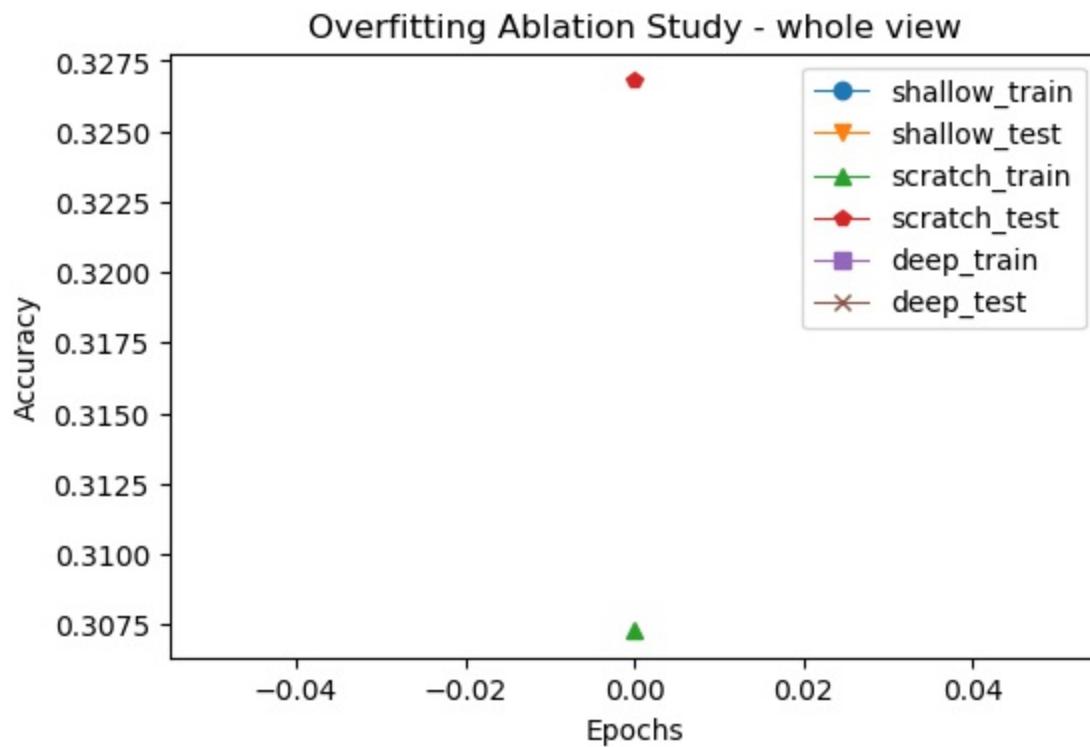
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN

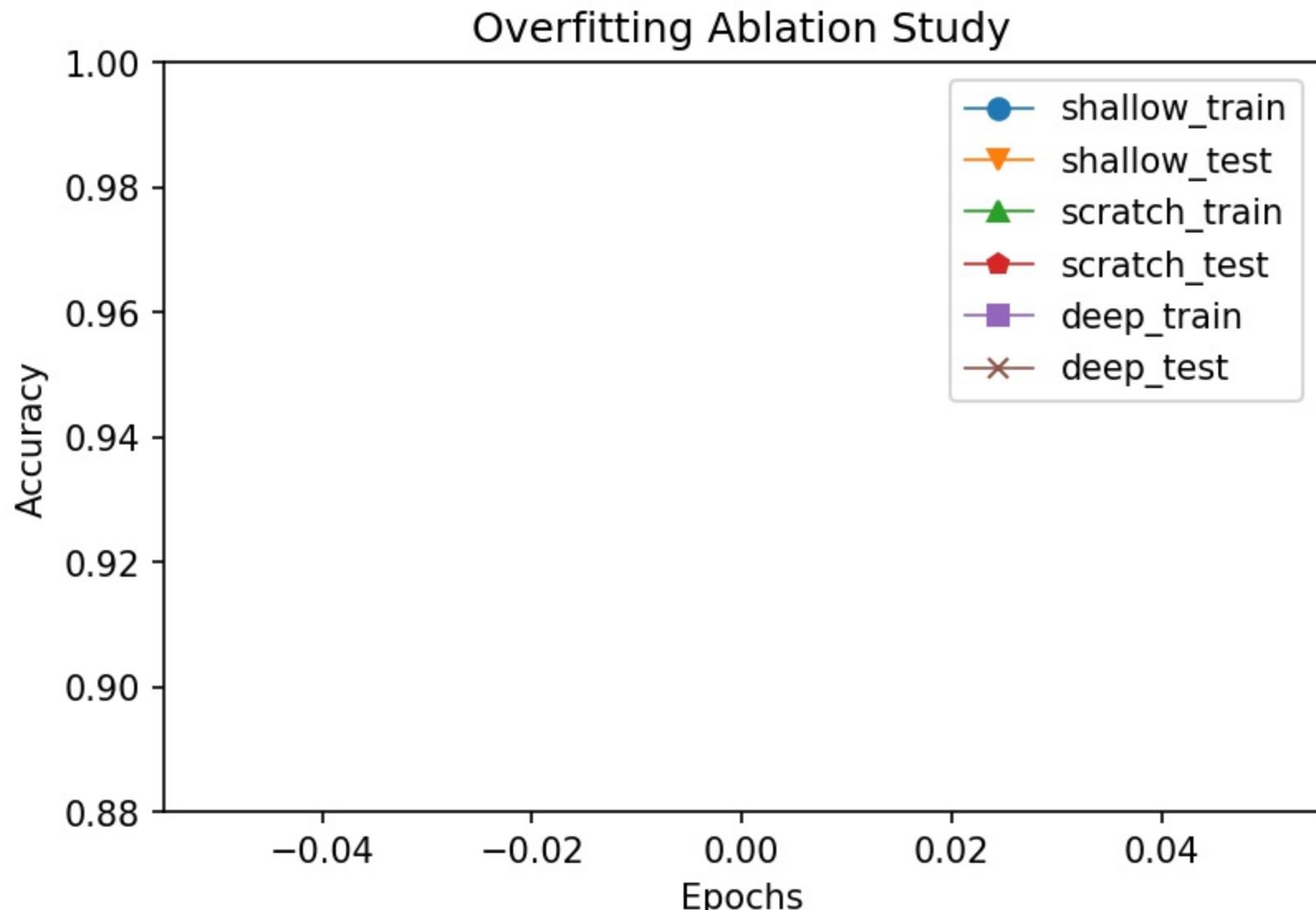
1	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN

40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN

79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN

118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 0 takes 329 seconds.
```

```
[2,    30] loss: 2.552
```

```
[2,    60] loss: 2.518
```

```
[2,    90] loss: 2.496
[2,   120] loss: 2.537
[2,   150] loss: 2.459
[2,   180] loss: 2.527
[2,   210] loss: 2.630
[2,   240] loss: 2.387
[2,   270] loss: 2.453
[2,   300] loss: 2.426
[2,   330] loss: 2.515
[2,   360] loss: 2.391
[2,   390] loss: 2.321
[2,   420] loss: 2.465
[2,   450] loss: 2.341
[2,   480] loss: 2.287
[2,   510] loss: 2.275
[2,   540] loss: 2.316
[2,   570] loss: 2.225
[2,   600] loss: 2.362
[2,   630] loss: 2.361
[2,   660] loss: 2.315
[2,   690] loss: 2.280
[2,   720] loss: 2.369
[2,   750] loss: 2.132
[2,   780] loss: 2.178
[2,   810] loss: 2.259
[2,   840] loss: 2.250
[2,   870] loss: 2.208
[2,   900] loss: 2.088
[2,   930] loss: 2.354
[2,   960] loss: 2.108
[2,   990] loss: 2.078
[2,  1020] loss: 2.118
[2,  1050] loss: 2.316
[2,  1080] loss: 2.079
[2,  1110] loss: 2.024
[2,  1140] loss: 2.032
[2,  1170] loss: 1.992
[2,  1200] loss: 2.150
[2,  1230] loss: 2.048
```

```
[2, 1260] loss: 2.295
[2, 1290] loss: 2.116
[2, 1320] loss: 2.086
[2, 1350] loss: 1.975
[2, 1380] loss: 1.978
[2, 1410] loss: 1.964
[2, 1440] loss: 1.893
[2, 1470] loss: 2.053
[2, 1500] loss: 2.037
[2, 1530] loss: 1.996
[2, 1560] loss: 1.988
[2, 1590] loss: 1.979
[2, 1620] loss: 2.023
[2, 1650] loss: 1.921
[2, 1680] loss: 2.007
[2, 1710] loss: 2.058
[2, 1740] loss: 1.832
[2, 1770] loss: 1.838
[2, 1800] loss: 1.898
[2, 1830] loss: 2.044
[2, 1860] loss: 1.881
[2, 1890] loss: 1.893
[2, 1920] loss: 2.040
[2, 1950] loss: 1.761
[2, 1980] loss: 1.838
[2, 2010] loss: 1.972
[2, 2040] loss: 1.865
[2, 2070] loss: 1.825
[2, 2100] loss: 1.781
[2, 2130] loss: 1.742
[2, 2160] loss: 1.662
[2, 2190] loss: 1.844
eval intermediate_models/epoch1_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.467444
epoch 1 accuracy on train set is: 0.4674436568520538
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.501363

epoch 1 accuracy on test set is: 0.5013631406761178

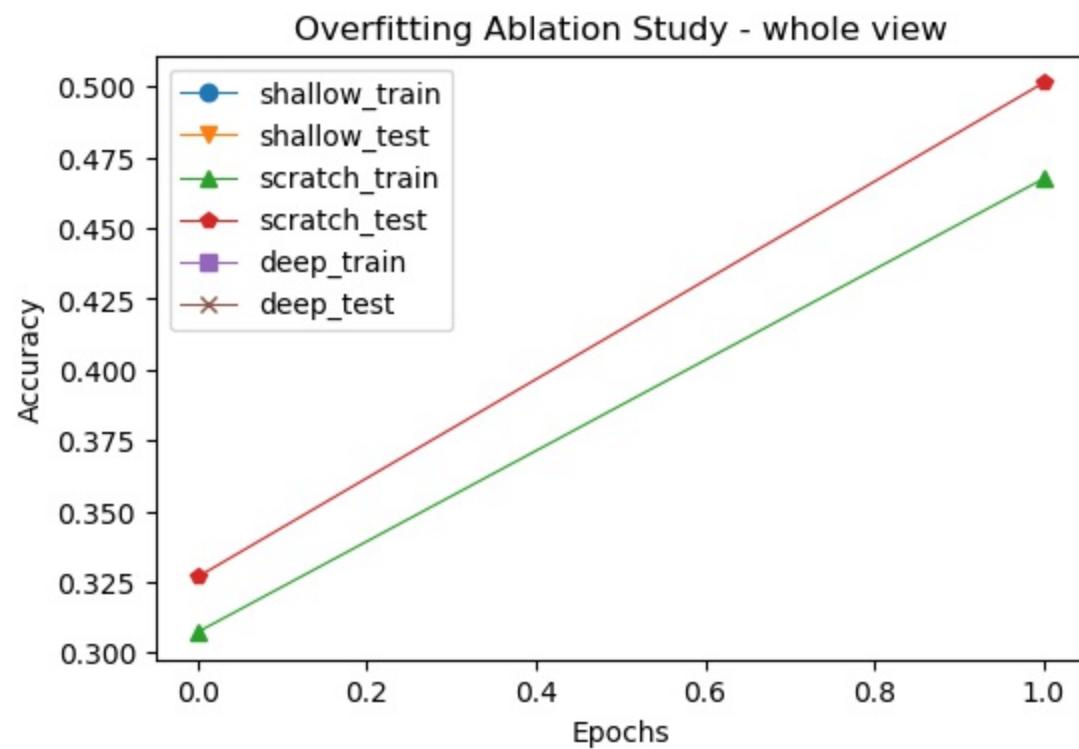
evaluating on test set takes 16 seconds.

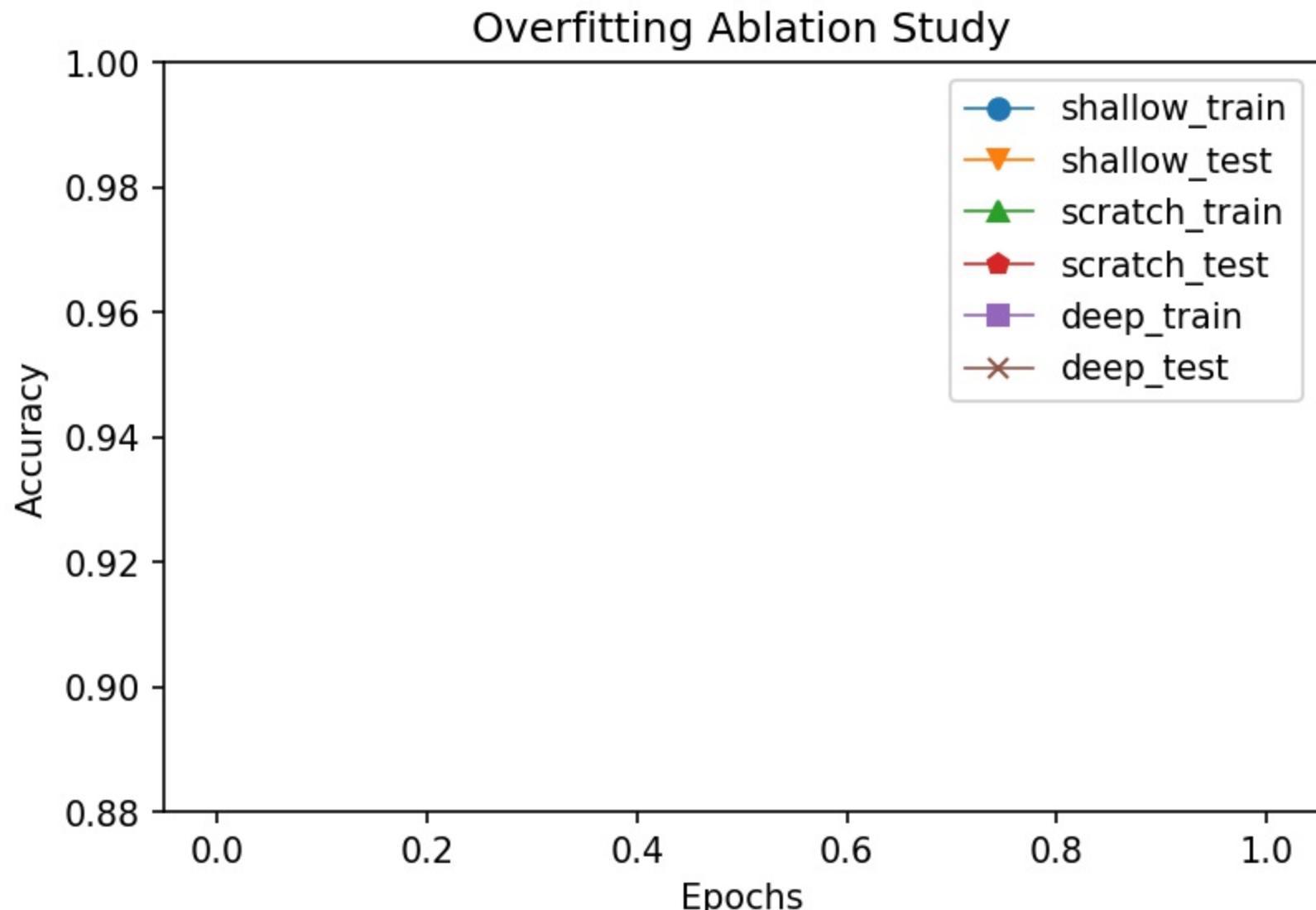
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 1 takes 320 seconds.
```

```
[3,    30] loss: 1.814
```

```
[3,    60] loss: 1.748
```

```
[3,    90] loss: 1.764
[3,   120] loss: 1.742
[3,   150] loss: 1.786
[3,   180] loss: 1.688
[3,   210] loss: 1.629
[3,   240] loss: 1.651
[3,   270] loss: 1.759
[3,   300] loss: 1.708
[3,   330] loss: 1.707
[3,   360] loss: 1.686
[3,   390] loss: 1.597
[3,   420] loss: 1.731
[3,   450] loss: 1.790
[3,   480] loss: 1.756
[3,   510] loss: 1.718
[3,   540] loss: 1.591
[3,   570] loss: 1.675
[3,   600] loss: 1.619
[3,   630] loss: 1.524
[3,   660] loss: 1.655
[3,   690] loss: 1.528
[3,   720] loss: 1.561
[3,   750] loss: 1.485
[3,   780] loss: 1.577
[3,   810] loss: 1.628
[3,   840] loss: 1.542
[3,   870] loss: 1.438
[3,   900] loss: 1.644
[3,   930] loss: 1.604
[3,   960] loss: 1.661
[3,   990] loss: 1.488
[3,  1020] loss: 1.560
[3,  1050] loss: 1.509
[3,  1080] loss: 1.389
[3,  1110] loss: 1.470
[3,  1140] loss: 1.565
[3,  1170] loss: 1.476
[3,  1200] loss: 1.476
[3,  1230] loss: 1.515
```

```
[3, 1260] loss: 1.424
[3, 1290] loss: 1.367
[3, 1320] loss: 1.444
[3, 1350] loss: 1.533
[3, 1380] loss: 1.496
[3, 1410] loss: 1.534
[3, 1440] loss: 1.458
[3, 1470] loss: 1.410
[3, 1500] loss: 1.453
[3, 1530] loss: 1.486
[3, 1560] loss: 1.449
[3, 1590] loss: 1.491
[3, 1620] loss: 1.315
[3, 1650] loss: 1.401
[3, 1680] loss: 1.256
[3, 1710] loss: 1.321
[3, 1740] loss: 1.376
[3, 1770] loss: 1.300
[3, 1800] loss: 1.323
[3, 1830] loss: 1.410
[3, 1860] loss: 1.169
[3, 1890] loss: 1.295
[3, 1920] loss: 1.364
[3, 1950] loss: 1.370
[3, 1980] loss: 1.278
[3, 2010] loss: 1.434
[3, 2040] loss: 1.381
[3, 2070] loss: 1.274
[3, 2100] loss: 1.268
[3, 2130] loss: 1.452
[3, 2160] loss: 1.310
[3, 2190] loss: 1.327
eval intermediate_models/epoch2_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.650900
epoch 2 accuracy on train set is: 0.6508996728462377
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.720738

epoch 2 accuracy on test set is: 0.7207379134860051

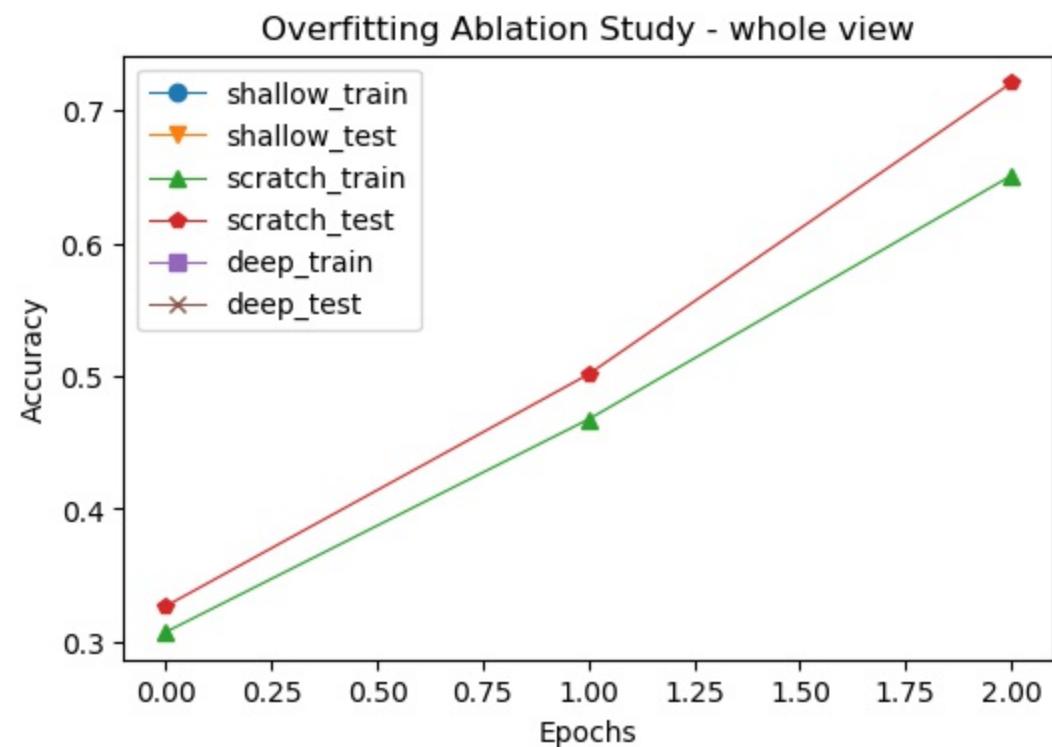
evaluating on test set takes 16 seconds.

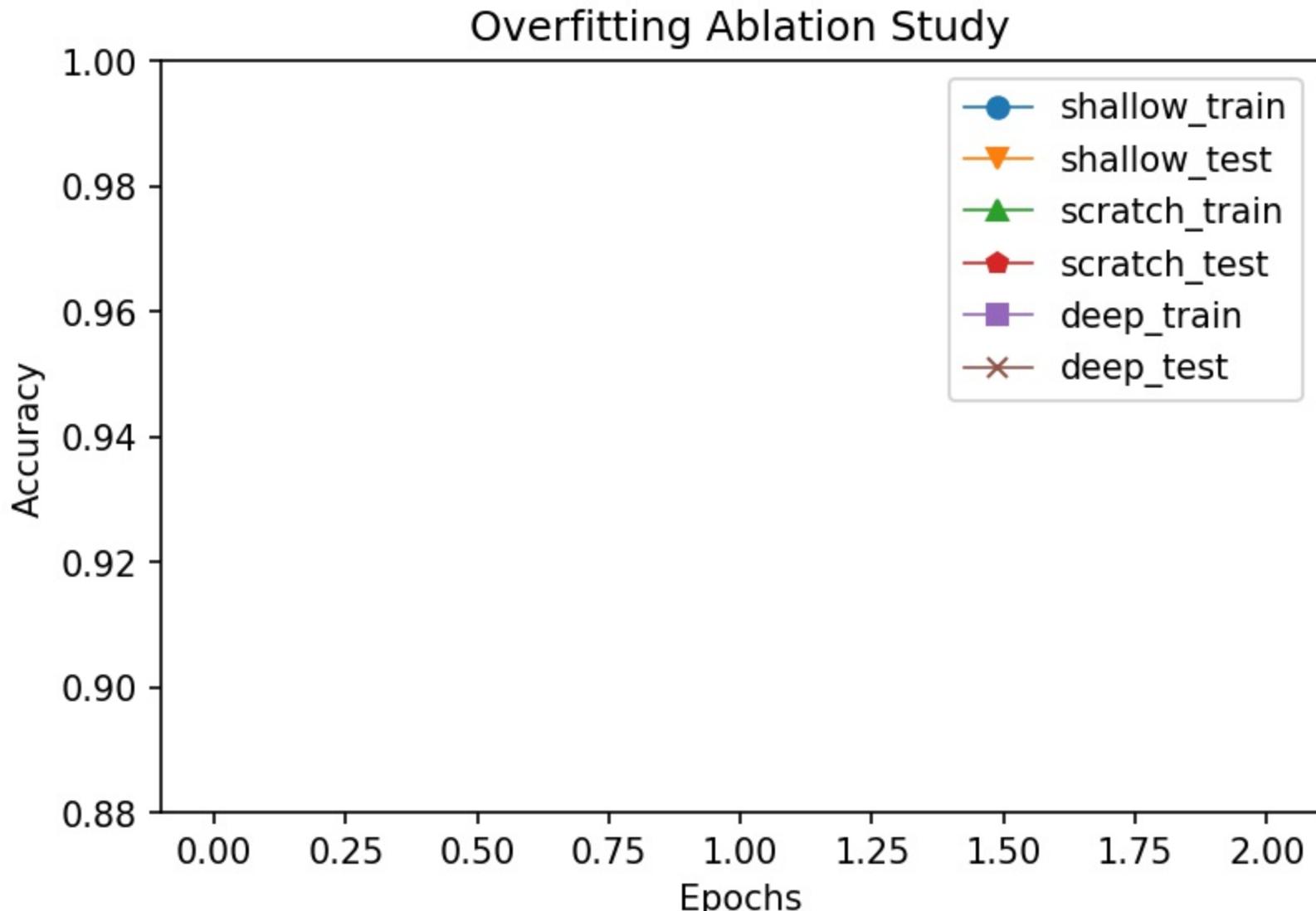
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 2 takes 319 seconds.

[4, 30] loss: 1.270

[4, 60] loss: 1.267

```
[4,    90] loss: 1.272
[4,   120] loss: 1.207
[4,   150] loss: 1.171
[4,   180] loss: 1.363
[4,   210] loss: 1.261
[4,   240] loss: 1.297
[4,   270] loss: 1.132
[4,   300] loss: 1.121
[4,   330] loss: 1.226
[4,   360] loss: 1.218
[4,   390] loss: 1.160
[4,   420] loss: 1.242
[4,   450] loss: 1.291
[4,   480] loss: 1.195
[4,   510] loss: 1.134
[4,   540] loss: 1.198
[4,   570] loss: 1.298
[4,   600] loss: 1.095
[4,   630] loss: 1.160
[4,   660] loss: 1.197
[4,   690] loss: 1.156
[4,   720] loss: 1.141
[4,   750] loss: 1.069
[4,   780] loss: 1.181
[4,   810] loss: 1.155
[4,   840] loss: 1.101
[4,   870] loss: 1.094
[4,   900] loss: 1.121
[4,   930] loss: 1.008
[4,   960] loss: 1.025
[4,   990] loss: 1.036
[4,  1020] loss: 1.077
[4,  1050] loss: 1.181
[4,  1080] loss: 1.073
[4,  1110] loss: 1.039
[4,  1140] loss: 1.081
[4,  1170] loss: 0.999
[4,  1200] loss: 0.989
[4,  1230] loss: 0.999
```

```
[4, 1260] loss: 1.086
[4, 1290] loss: 1.011
[4, 1320] loss: 1.036
[4, 1350] loss: 1.015
[4, 1380] loss: 1.197
[4, 1410] loss: 1.061
[4, 1440] loss: 0.983
[4, 1470] loss: 0.978
[4, 1500] loss: 1.148
[4, 1530] loss: 0.948
[4, 1560] loss: 0.953
[4, 1590] loss: 1.104
[4, 1620] loss: 0.973
[4, 1650] loss: 0.991
[4, 1680] loss: 1.009
[4, 1710] loss: 0.921
[4, 1740] loss: 1.046
[4, 1770] loss: 0.951
[4, 1800] loss: 1.039
[4, 1830] loss: 1.021
[4, 1860] loss: 1.012
[4, 1890] loss: 1.012
[4, 1920] loss: 0.952
[4, 1950] loss: 0.927
[4, 1980] loss: 0.997
[4, 2010] loss: 0.974
[4, 2040] loss: 0.989
[4, 2070] loss: 0.975
[4, 2100] loss: 0.998
[4, 2130] loss: 1.157
[4, 2160] loss: 1.086
[4, 2190] loss: 1.021
eval intermediate_models/epoch3_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.742480
epoch 3 accuracy on train set is: 0.7424800072700836
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.819066

epoch 3 accuracy on test set is: 0.819065794256634

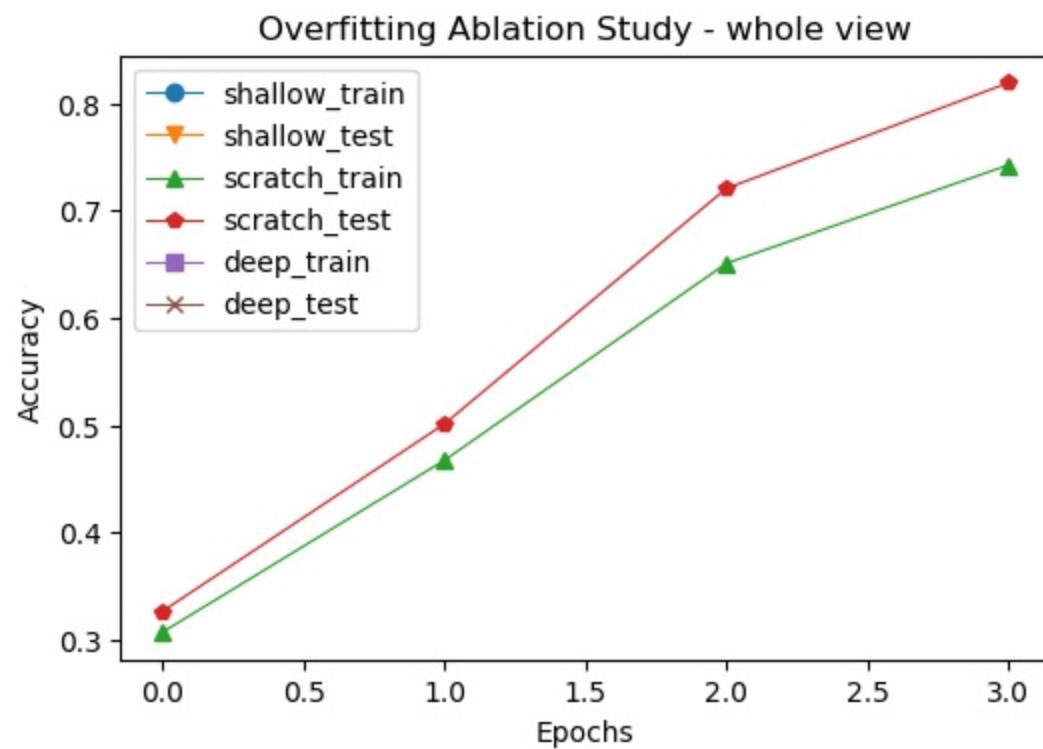
evaluating on test set takes 16 seconds.

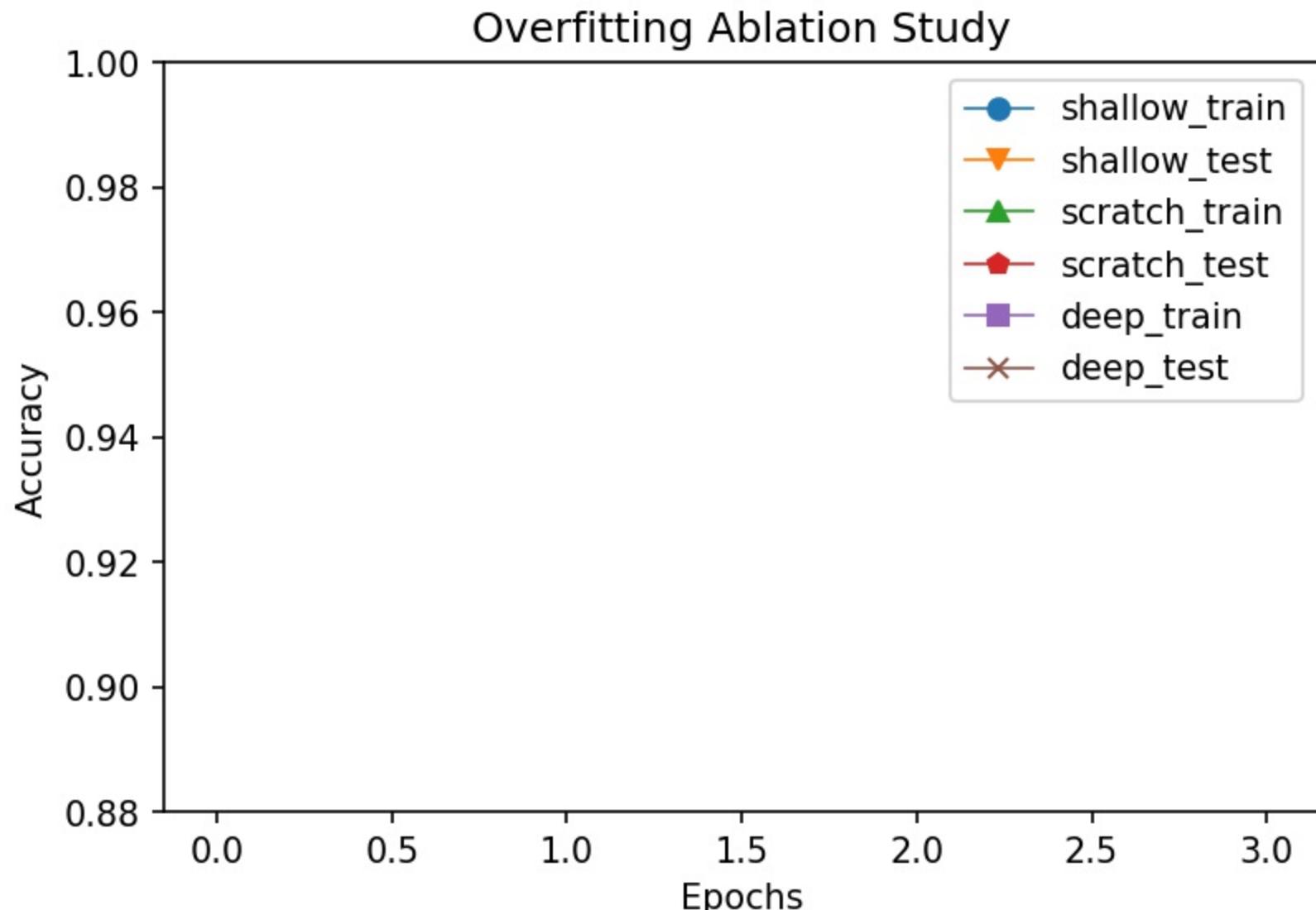
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 3 takes 320 seconds.

[5, 30] loss: 0.863

[5, 60] loss: 0.923

```
[5,    90] loss: 0.832
[5,   120] loss: 0.943
[5,   150] loss: 0.985
[5,   180] loss: 1.023
[5,   210] loss: 0.876
[5,   240] loss: 0.991
[5,   270] loss: 0.873
[5,   300] loss: 0.937
[5,   330] loss: 1.066
[5,   360] loss: 0.960
[5,   390] loss: 1.048
[5,   420] loss: 0.883
[5,   450] loss: 0.907
[5,   480] loss: 0.957
[5,   510] loss: 0.936
[5,   540] loss: 0.867
[5,   570] loss: 1.021
[5,   600] loss: 0.816
[5,   630] loss: 0.939
[5,   660] loss: 0.889
[5,   690] loss: 0.944
[5,   720] loss: 0.978
[5,   750] loss: 0.897
[5,   780] loss: 0.883
[5,   810] loss: 0.980
[5,   840] loss: 0.813
[5,   870] loss: 0.872
[5,   900] loss: 0.850
[5,   930] loss: 0.904
[5,   960] loss: 0.845
[5,   990] loss: 0.868
[5,  1020] loss: 0.822
[5,  1050] loss: 0.830
[5,  1080] loss: 0.849
[5,  1110] loss: 0.810
[5,  1140] loss: 0.806
[5,  1170] loss: 0.890
[5,  1200] loss: 0.778
[5,  1230] loss: 0.776
```

```
[5, 1260] loss: 0.724
[5, 1290] loss: 0.807
[5, 1320] loss: 0.758
[5, 1350] loss: 0.866
[5, 1380] loss: 0.896
[5, 1410] loss: 0.739
[5, 1440] loss: 0.862
[5, 1470] loss: 0.822
[5, 1500] loss: 0.812
[5, 1530] loss: 0.745
[5, 1560] loss: 0.803
[5, 1590] loss: 0.915
[5, 1620] loss: 0.775
[5, 1650] loss: 0.837
[5, 1680] loss: 0.764
[5, 1710] loss: 0.754
[5, 1740] loss: 0.881
[5, 1770] loss: 0.785
[5, 1800] loss: 0.790
[5, 1830] loss: 0.760
[5, 1860] loss: 0.773
[5, 1890] loss: 0.821
[5, 1920] loss: 0.803
[5, 1950] loss: 0.790
[5, 1980] loss: 0.787
[5, 2010] loss: 0.808
[5, 2040] loss: 0.725
[5, 2070] loss: 0.783
[5, 2100] loss: 0.872
[5, 2130] loss: 0.721
[5, 2160] loss: 0.744
[5, 2190] loss: 0.765
eval intermediate_models/epoch4_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.790008
epoch 4 accuracy on train set is: 0.7900081788440567
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.838332

epoch 4 accuracy on test set is: 0.8383315158124318

evaluating on test set takes 16 seconds.

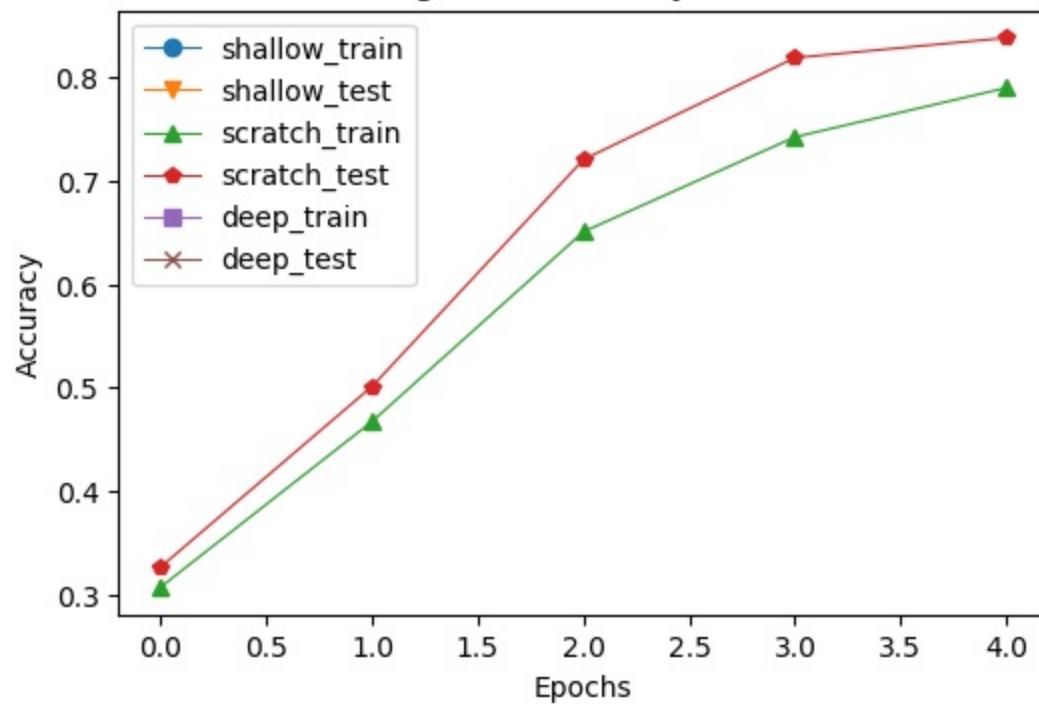
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

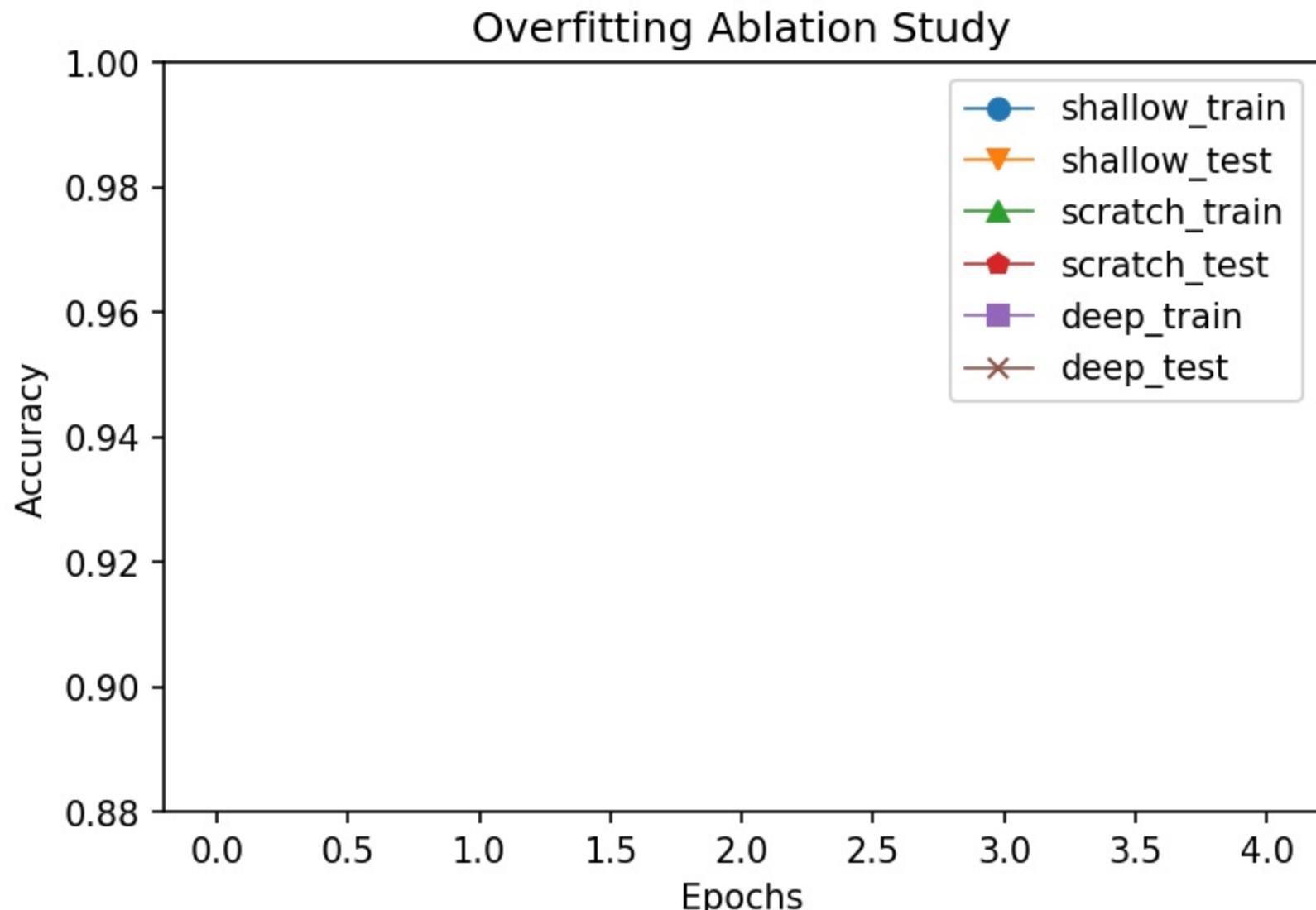
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view





epoch 4 takes 318 seconds.

[6, 30] loss: 0.830

[6, 60] loss: 0.797

```
[6,    90] loss: 0.789
[6,   120] loss: 0.667
[6,   150] loss: 0.769
[6,   180] loss: 0.686
[6,   210] loss: 0.791
[6,   240] loss: 0.738
[6,   270] loss: 0.712
[6,   300] loss: 0.733
[6,   330] loss: 0.794
[6,   360] loss: 0.726
[6,   390] loss: 0.653
[6,   420] loss: 0.697
[6,   450] loss: 0.752
[6,   480] loss: 0.713
[6,   510] loss: 0.621
[6,   540] loss: 0.796
[6,   570] loss: 0.735
[6,   600] loss: 0.756
[6,   630] loss: 0.632
[6,   660] loss: 0.735
[6,   690] loss: 0.734
[6,   720] loss: 0.767
[6,   750] loss: 0.784
[6,   780] loss: 0.626
[6,   810] loss: 0.690
[6,   840] loss: 0.737
[6,   870] loss: 0.698
[6,   900] loss: 0.713
[6,   930] loss: 0.662
[6,   960] loss: 0.616
[6,   990] loss: 0.614
[6,  1020] loss: 0.714
[6,  1050] loss: 0.693
[6,  1080] loss: 0.755
[6,  1110] loss: 0.670
[6,  1140] loss: 0.692
[6,  1170] loss: 0.644
[6,  1200] loss: 0.727
[6,  1230] loss: 0.660
```

```
[6, 1260] loss: 0.667
[6, 1290] loss: 0.750
[6, 1320] loss: 0.773
[6, 1350] loss: 0.760
[6, 1380] loss: 0.754
[6, 1410] loss: 0.686
[6, 1440] loss: 0.767
[6, 1470] loss: 0.610
[6, 1500] loss: 0.619
[6, 1530] loss: 0.728
[6, 1560] loss: 0.696
[6, 1590] loss: 0.673
[6, 1620] loss: 0.621
[6, 1650] loss: 0.685
[6, 1680] loss: 0.745
[6, 1710] loss: 0.731
[6, 1740] loss: 0.644
[6, 1770] loss: 0.652
[6, 1800] loss: 0.700
[6, 1830] loss: 0.595
[6, 1860] loss: 0.667
[6, 1890] loss: 0.761
[6, 1920] loss: 0.603
[6, 1950] loss: 0.639
[6, 1980] loss: 0.621
[6, 2010] loss: 0.649
[6, 2040] loss: 0.670
[6, 2070] loss: 0.699
[6, 2100] loss: 0.662
[6, 2130] loss: 0.664
[6, 2160] loss: 0.608
[6, 2190] loss: 0.586
eval intermediate_models/epoch5_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.832652
epoch 5 accuracy on train set is: 0.8326517629952744
evaluating on training set takes 65 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.893675

epoch 5 accuracy on test set is: 0.8936750272628136

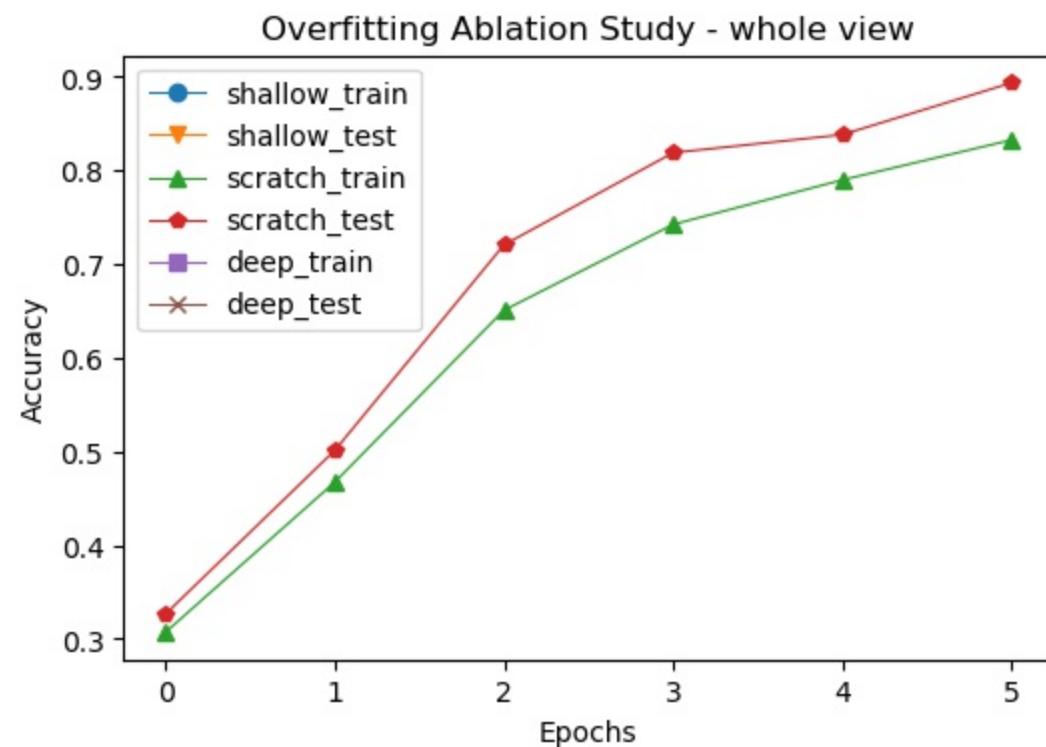
evaluating on test set takes 17 seconds.

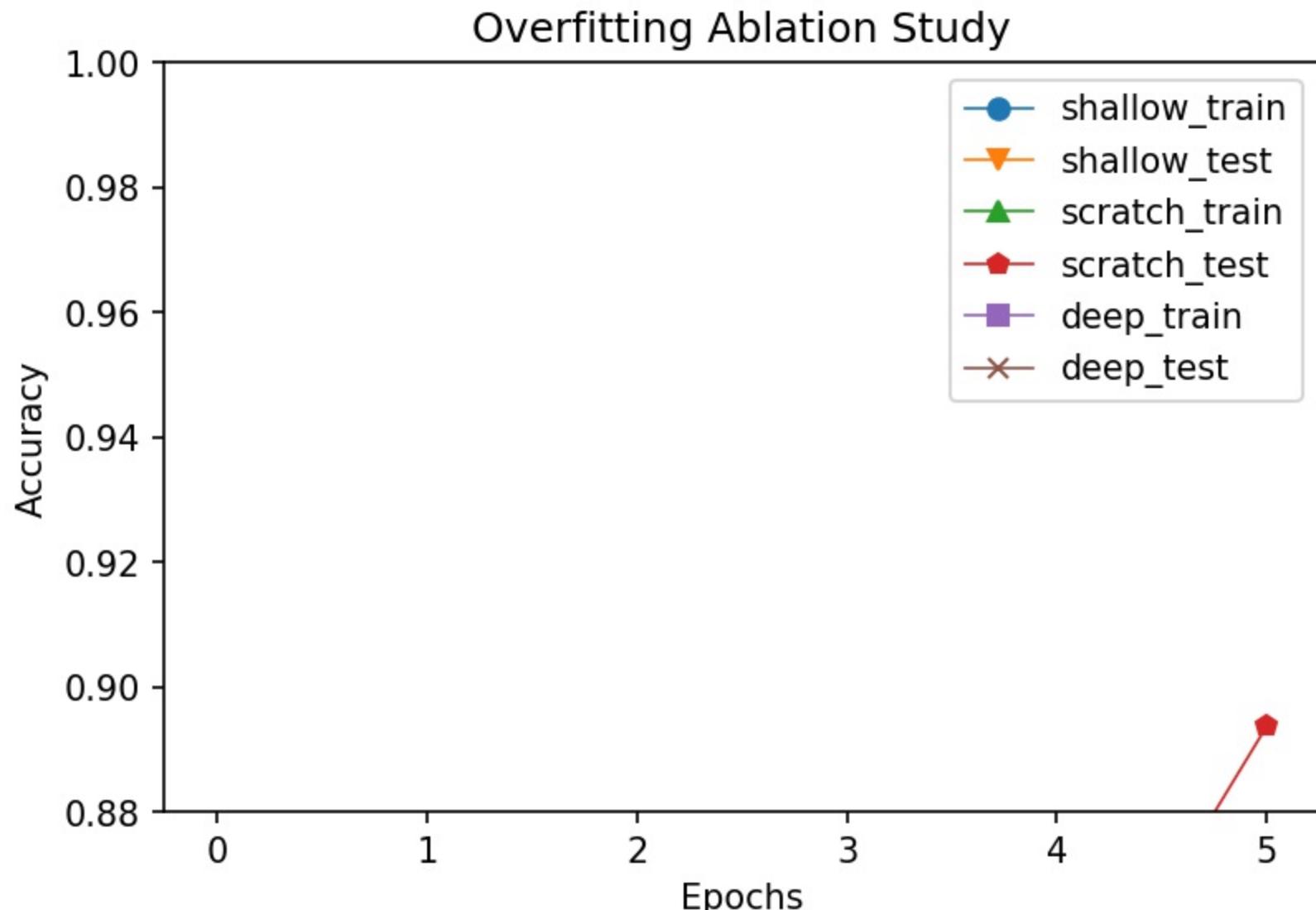
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 5 takes 321 seconds.

[7, 30] loss: 0.601

[7, 60] loss: 0.672

```
[7,    90] loss: 0.707
[7,   120] loss: 0.699
[7,   150] loss: 0.632
[7,   180] loss: 0.665
[7,   210] loss: 0.703
[7,   240] loss: 0.531
[7,   270] loss: 0.577
[7,   300] loss: 0.695
[7,   330] loss: 0.657
[7,   360] loss: 0.557
[7,   390] loss: 0.582
[7,   420] loss: 0.622
[7,   450] loss: 0.592
[7,   480] loss: 0.563
[7,   510] loss: 0.572
[7,   540] loss: 0.644
[7,   570] loss: 0.628
[7,   600] loss: 0.585
[7,   630] loss: 0.606
[7,   660] loss: 0.517
[7,   690] loss: 0.565
[7,   720] loss: 0.567
[7,   750] loss: 0.618
[7,   780] loss: 0.653
[7,   810] loss: 0.558
[7,   840] loss: 0.643
[7,   870] loss: 0.501
[7,   900] loss: 0.582
[7,   930] loss: 0.699
[7,   960] loss: 0.625
[7,   990] loss: 0.686
[7,  1020] loss: 0.576
[7,  1050] loss: 0.598
[7,  1080] loss: 0.621
[7,  1110] loss: 0.632
[7,  1140] loss: 0.646
[7,  1170] loss: 0.591
[7,  1200] loss: 0.551
[7,  1230] loss: 0.462
```

```
[7, 1260] loss: 0.600
[7, 1290] loss: 0.584
[7, 1320] loss: 0.600
[7, 1350] loss: 0.492
[7, 1380] loss: 0.602
[7, 1410] loss: 0.569
[7, 1440] loss: 0.679
[7, 1470] loss: 0.610
[7, 1500] loss: 0.535
[7, 1530] loss: 0.660
[7, 1560] loss: 0.604
[7, 1590] loss: 0.544
[7, 1620] loss: 0.568
[7, 1650] loss: 0.591
[7, 1680] loss: 0.600
[7, 1710] loss: 0.577
[7, 1740] loss: 0.510
[7, 1770] loss: 0.544
[7, 1800] loss: 0.553
[7, 1830] loss: 0.505
[7, 1860] loss: 0.570
[7, 1890] loss: 0.521
[7, 1920] loss: 0.509
[7, 1950] loss: 0.498
[7, 1980] loss: 0.578
[7, 2010] loss: 0.474
[7, 2040] loss: 0.479
[7, 2070] loss: 0.610
[7, 2100] loss: 0.590
[7, 2130] loss: 0.479
[7, 2160] loss: 0.522
[7, 2190] loss: 0.480
eval intermediate_models/epoch6_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.851531
epoch 6 accuracy on train set is: 0.8515312613595056
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.904398

epoch 6 accuracy on test set is: 0.9043984005816067

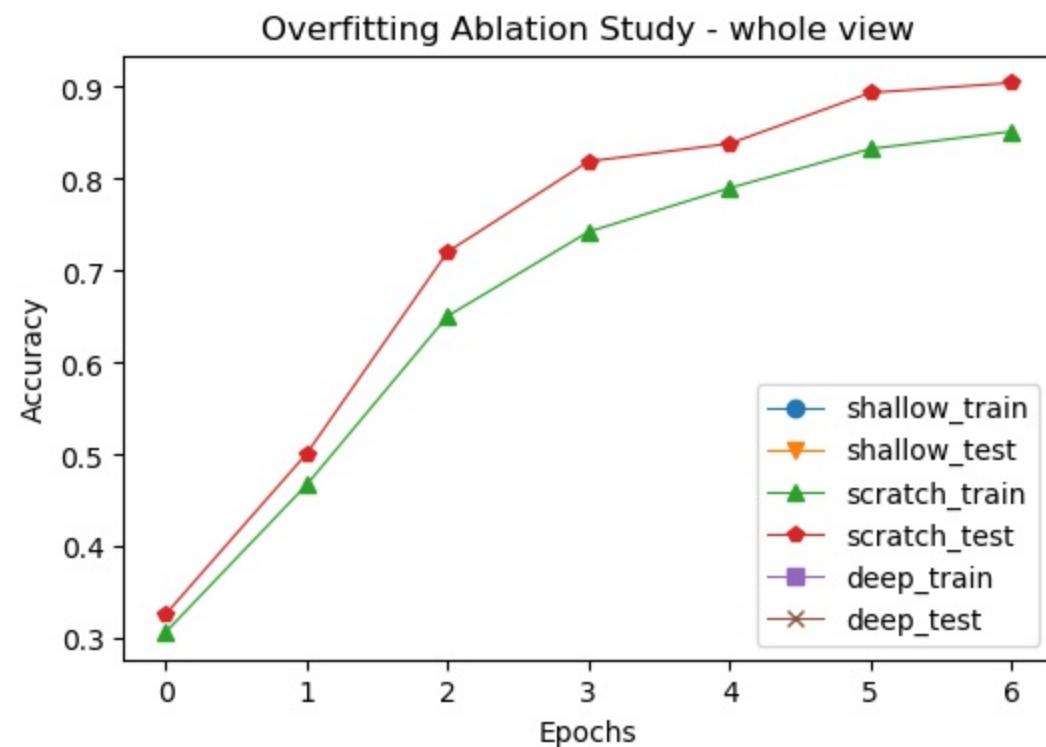
evaluating on test set takes 16 seconds.

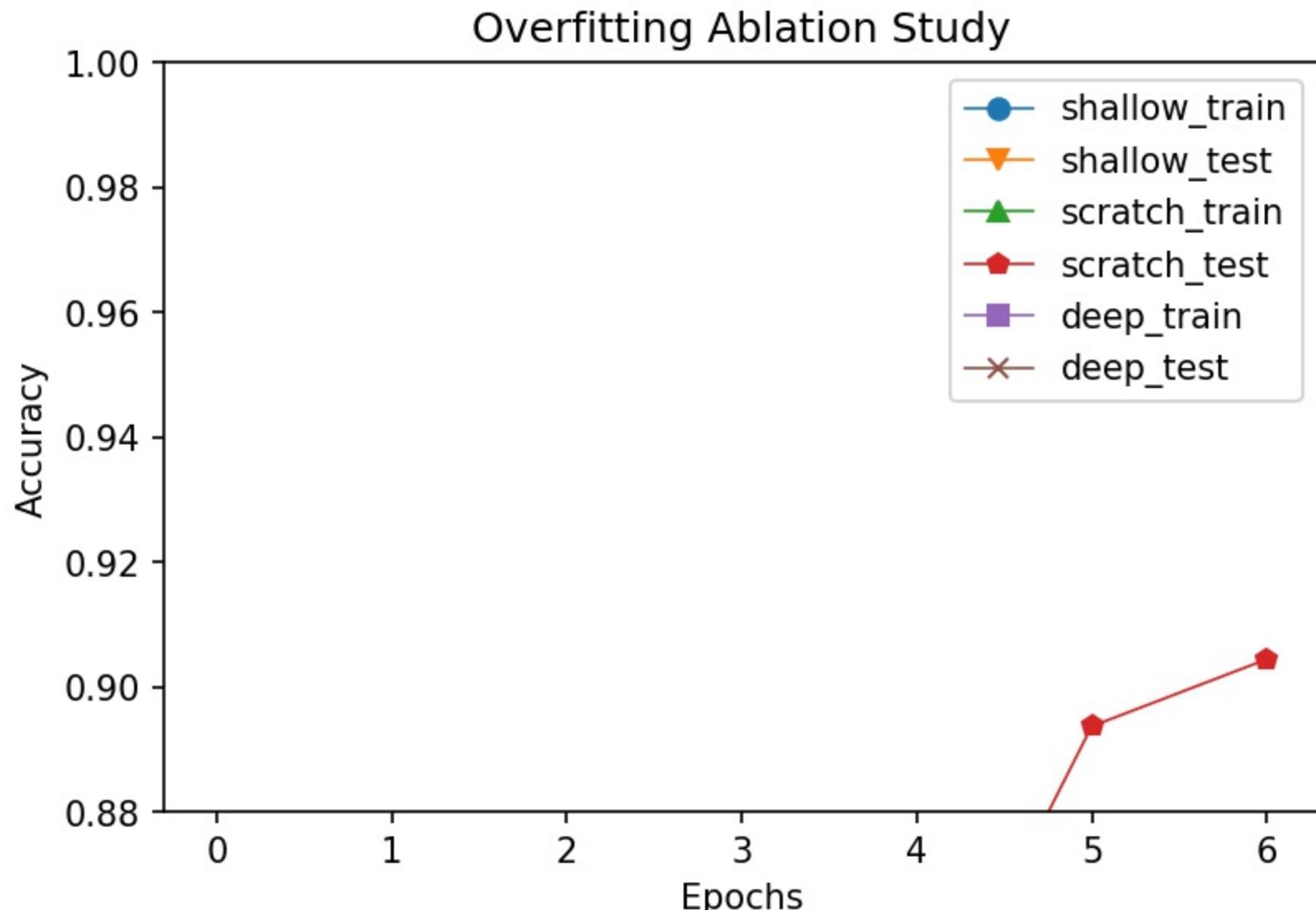
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 6 takes 323 seconds.

[8, 30] loss: 0.461

[8, 60] loss: 0.488

```
[8,    90] loss: 0.497
[8,   120] loss: 0.569
[8,   150] loss: 0.600
[8,   180] loss: 0.501
[8,   210] loss: 0.588
[8,   240] loss: 0.558
[8,   270] loss: 0.534
[8,   300] loss: 0.552
[8,   330] loss: 0.473
[8,   360] loss: 0.518
[8,   390] loss: 0.464
[8,   420] loss: 0.577
[8,   450] loss: 0.539
[8,   480] loss: 0.480
[8,   510] loss: 0.533
[8,   540] loss: 0.639
[8,   570] loss: 0.502
[8,   600] loss: 0.511
[8,   630] loss: 0.468
[8,   660] loss: 0.577
[8,   690] loss: 0.555
[8,   720] loss: 0.467
[8,   750] loss: 0.605
[8,   780] loss: 0.544
[8,   810] loss: 0.535
[8,   840] loss: 0.641
[8,   870] loss: 0.535
[8,   900] loss: 0.488
[8,   930] loss: 0.511
[8,   960] loss: 0.564
[8,   990] loss: 0.446
[8,  1020] loss: 0.498
[8,  1050] loss: 0.523
[8,  1080] loss: 0.540
[8,  1110] loss: 0.540
[8,  1140] loss: 0.494
[8,  1170] loss: 0.553
[8,  1200] loss: 0.526
[8,  1230] loss: 0.541
```

```
[8, 1260] loss: 0.403
[8, 1290] loss: 0.582
[8, 1320] loss: 0.555
[8, 1350] loss: 0.500
[8, 1380] loss: 0.494
[8, 1410] loss: 0.549
[8, 1440] loss: 0.512
[8, 1470] loss: 0.531
[8, 1500] loss: 0.499
[8, 1530] loss: 0.555
[8, 1560] loss: 0.499
[8, 1590] loss: 0.585
[8, 1620] loss: 0.576
[8, 1650] loss: 0.435
[8, 1680] loss: 0.474
[8, 1710] loss: 0.469
[8, 1740] loss: 0.508
[8, 1770] loss: 0.422
[8, 1800] loss: 0.472
[8, 1830] loss: 0.417
[8, 1860] loss: 0.466
[8, 1890] loss: 0.475
[8, 1920] loss: 0.536
[8, 1950] loss: 0.477
[8, 1980] loss: 0.453
[8, 2010] loss: 0.540
[8, 2040] loss: 0.483
[8, 2070] loss: 0.407
[8, 2100] loss: 0.487
[8, 2130] loss: 0.477
[8, 2160] loss: 0.483
[8, 2190] loss: 0.459
eval intermediate_models/epoch7_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.857256
epoch 7 accuracy on train set is: 0.8572564521992003
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.903308

epoch 7 accuracy on test set is: 0.9033078880407125

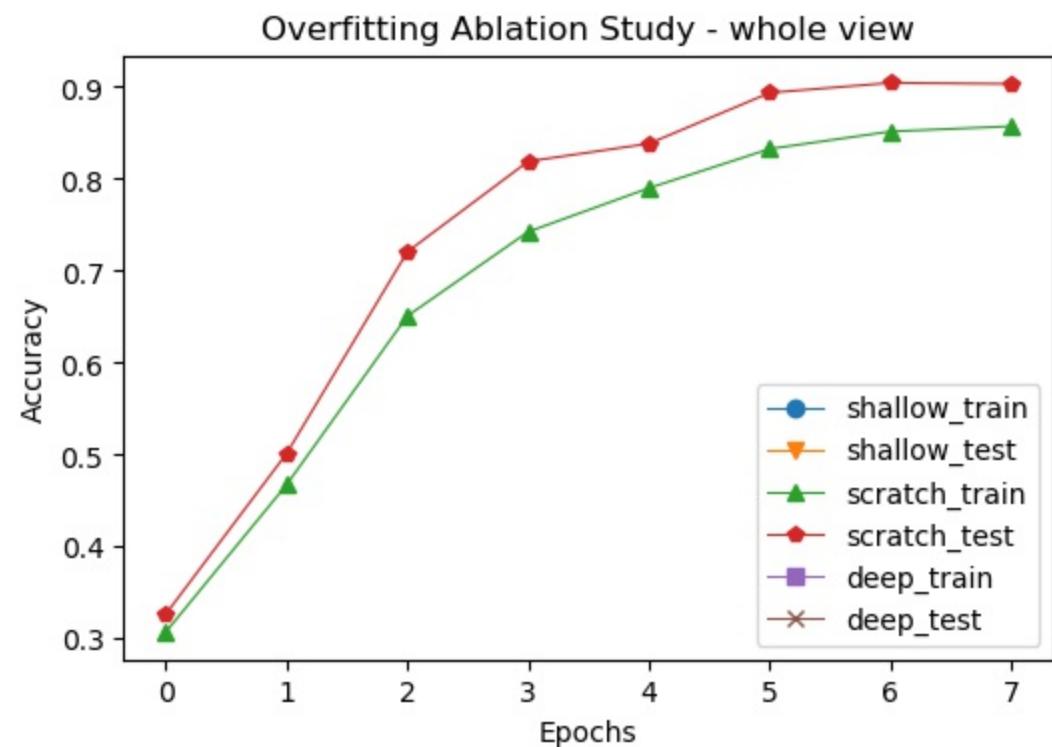
evaluating on test set takes 16 seconds.

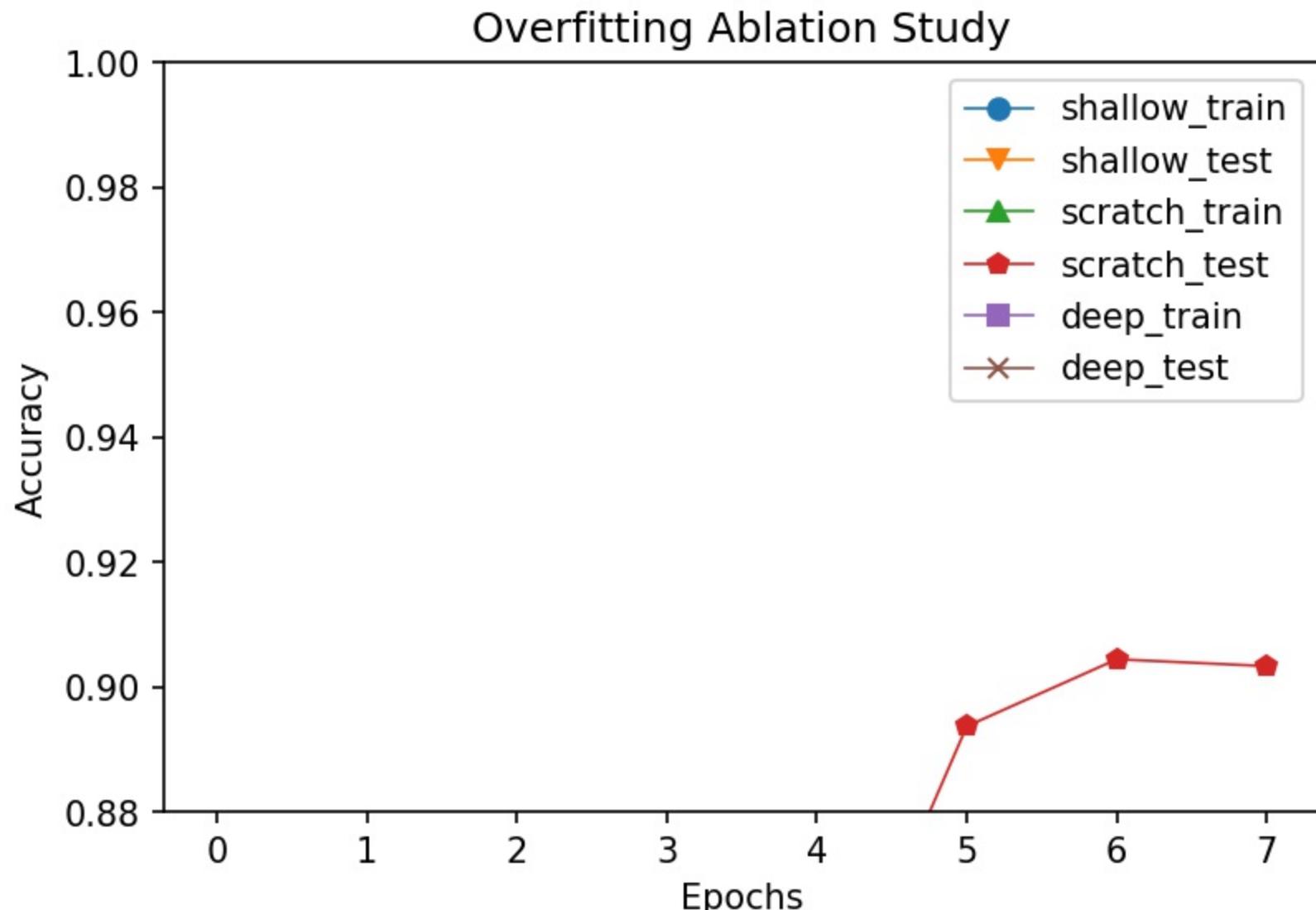
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 7 takes 319 seconds.

[9, 30] loss: 0.568

[9, 60] loss: 0.537

```
[9,    90] loss: 0.413
[9,   120] loss: 0.485
[9,   150] loss: 0.495
[9,   180] loss: 0.508
[9,   210] loss: 0.460
[9,   240] loss: 0.395
[9,   270] loss: 0.454
[9,   300] loss: 0.465
[9,   330] loss: 0.491
[9,   360] loss: 0.416
[9,   390] loss: 0.457
[9,   420] loss: 0.467
[9,   450] loss: 0.393
[9,   480] loss: 0.477
[9,   510] loss: 0.422
[9,   540] loss: 0.391
[9,   570] loss: 0.480
[9,   600] loss: 0.454
[9,   630] loss: 0.422
[9,   660] loss: 0.534
[9,   690] loss: 0.499
[9,   720] loss: 0.492
[9,   750] loss: 0.464
[9,   780] loss: 0.505
[9,   810] loss: 0.433
[9,   840] loss: 0.493
[9,   870] loss: 0.474
[9,   900] loss: 0.460
[9,   930] loss: 0.424
[9,   960] loss: 0.536
[9,   990] loss: 0.389
[9,  1020] loss: 0.403
[9,  1050] loss: 0.411
[9,  1080] loss: 0.446
[9,  1110] loss: 0.622
[9,  1140] loss: 0.446
[9,  1170] loss: 0.418
[9,  1200] loss: 0.481
[9,  1230] loss: 0.440
```

```
[9, 1260] loss: 0.441
[9, 1290] loss: 0.451
[9, 1320] loss: 0.400
[9, 1350] loss: 0.453
[9, 1380] loss: 0.387
[9, 1410] loss: 0.466
[9, 1440] loss: 0.473
[9, 1470] loss: 0.441
[9, 1500] loss: 0.426
[9, 1530] loss: 0.462
[9, 1560] loss: 0.467
[9, 1590] loss: 0.503
[9, 1620] loss: 0.450
[9, 1650] loss: 0.406
[9, 1680] loss: 0.488
[9, 1710] loss: 0.484
[9, 1740] loss: 0.439
[9, 1770] loss: 0.420
[9, 1800] loss: 0.431
[9, 1830] loss: 0.414
[9, 1860] loss: 0.423
[9, 1890] loss: 0.475
[9, 1920] loss: 0.454
[9, 1950] loss: 0.458
[9, 1980] loss: 0.333
[9, 2010] loss: 0.389
[9, 2040] loss: 0.480
[9, 2070] loss: 0.480
[9, 2100] loss: 0.392
[9, 2130] loss: 0.402
[9, 2160] loss: 0.327
[9, 2190] loss: 0.460
eval intermediate_models/epoch8_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.863050
epoch 8 accuracy on train set is: 0.8630498000727008
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.918484

epoch 8 accuracy on test set is: 0.9184841875681571

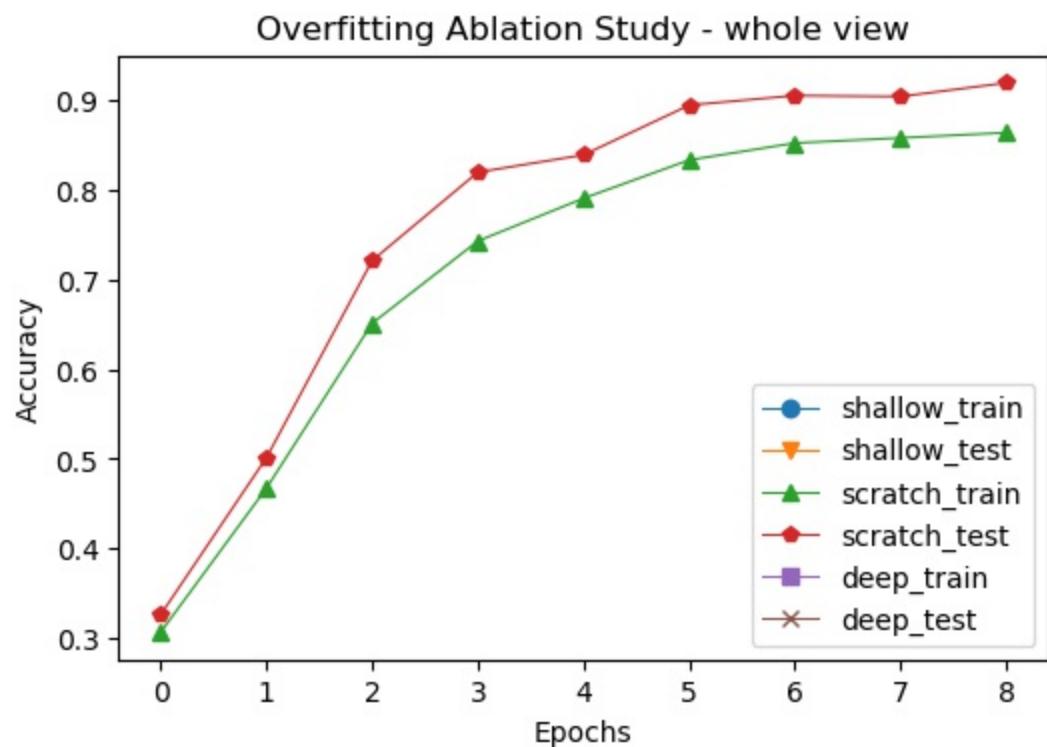
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

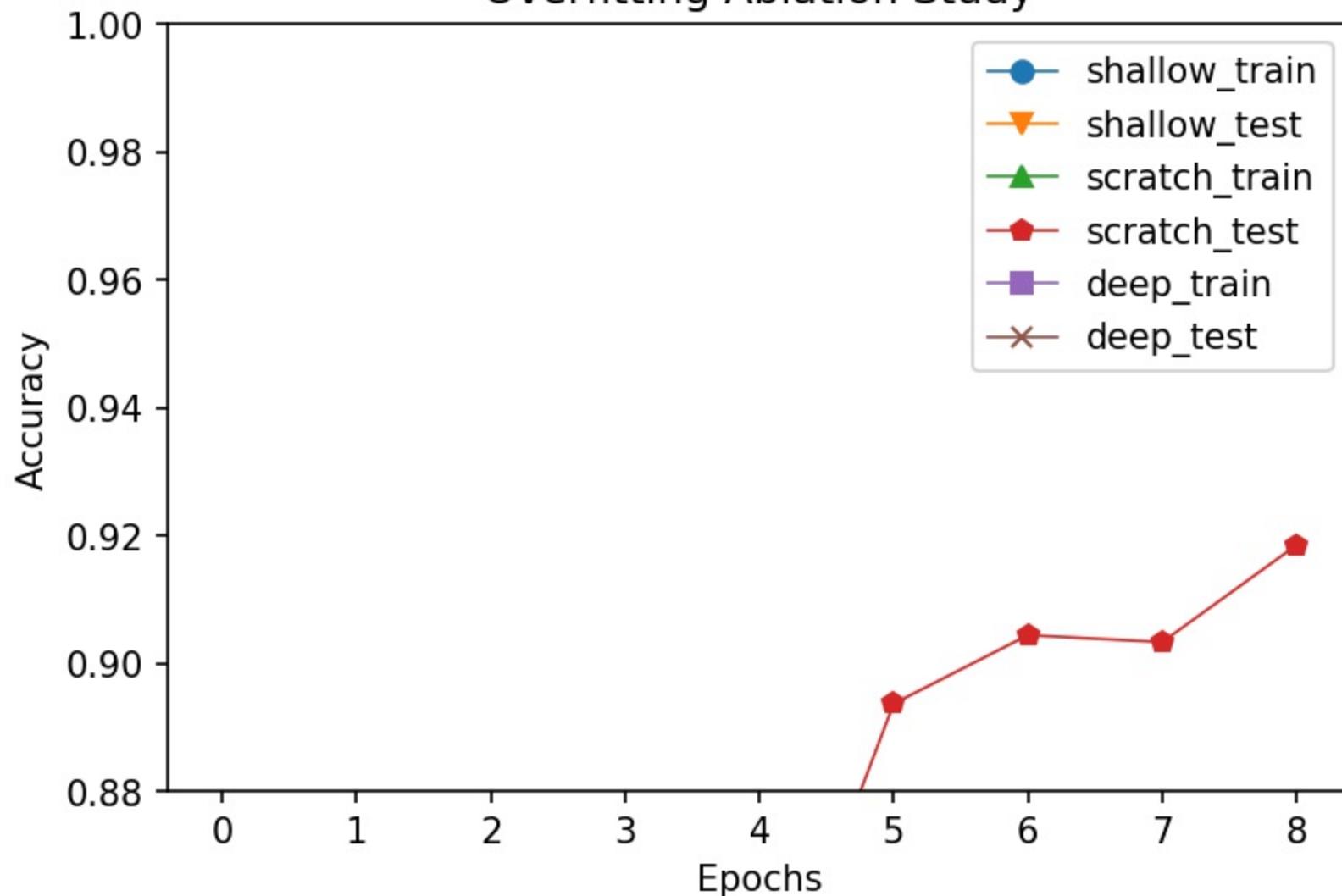
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 8 takes 318 seconds.

[10, 30] loss: 0.459

[10, 60] loss: 0.417

```
[10,    90] loss: 0.370
[10,   120] loss: 0.437
[10,   150] loss: 0.363
[10,   180] loss: 0.510
[10,   210] loss: 0.459
[10,   240] loss: 0.475
[10,   270] loss: 0.326
[10,   300] loss: 0.391
[10,   330] loss: 0.440
[10,   360] loss: 0.349
[10,   390] loss: 0.353
[10,   420] loss: 0.432
[10,   450] loss: 0.410
[10,   480] loss: 0.432
[10,   510] loss: 0.469
[10,   540] loss: 0.503
[10,   570] loss: 0.400
[10,   600] loss: 0.398
[10,   630] loss: 0.452
[10,   660] loss: 0.353
[10,   690] loss: 0.379
[10,   720] loss: 0.463
[10,   750] loss: 0.391
[10,   780] loss: 0.485
[10,   810] loss: 0.432
[10,   840] loss: 0.436
[10,   870] loss: 0.360
[10,   900] loss: 0.411
[10,   930] loss: 0.420
[10,   960] loss: 0.382
[10,   990] loss: 0.446
[10,  1020] loss: 0.322
[10,  1050] loss: 0.415
[10,  1080] loss: 0.376
[10,  1110] loss: 0.410
[10,  1140] loss: 0.399
[10,  1170] loss: 0.429
[10,  1200] loss: 0.454
[10,  1230] loss: 0.423
```

```
[10, 1260] loss: 0.336
[10, 1290] loss: 0.469
[10, 1320] loss: 0.411
[10, 1350] loss: 0.432
[10, 1380] loss: 0.449
[10, 1410] loss: 0.494
[10, 1440] loss: 0.430
[10, 1470] loss: 0.464
[10, 1500] loss: 0.367
[10, 1530] loss: 0.380
[10, 1560] loss: 0.328
[10, 1590] loss: 0.365
[10, 1620] loss: 0.398
[10, 1650] loss: 0.393
[10, 1680] loss: 0.418
[10, 1710] loss: 0.442
[10, 1740] loss: 0.400
[10, 1770] loss: 0.311
[10, 1800] loss: 0.352
[10, 1830] loss: 0.358
[10, 1860] loss: 0.358
[10, 1890] loss: 0.395
[10, 1920] loss: 0.396
[10, 1950] loss: 0.385
[10, 1980] loss: 0.426
[10, 2010] loss: 0.367
[10, 2040] loss: 0.387
[10, 2070] loss: 0.414
[10, 2100] loss: 0.423
[10, 2130] loss: 0.342
[10, 2160] loss: 0.392
[10, 2190] loss: 0.455
eval intermediate_models/epoch9_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.902467
epoch 9 accuracy on train set is: 0.9024672846237731
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.947383

epoch 9 accuracy on test set is: 0.9473827699018539

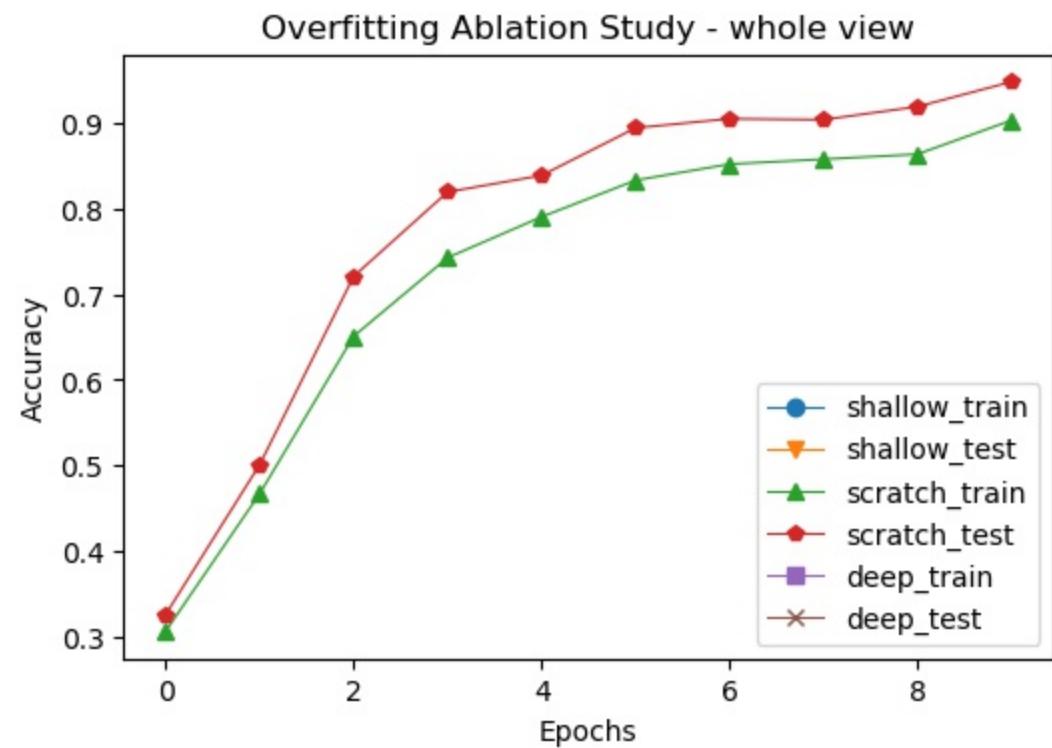
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

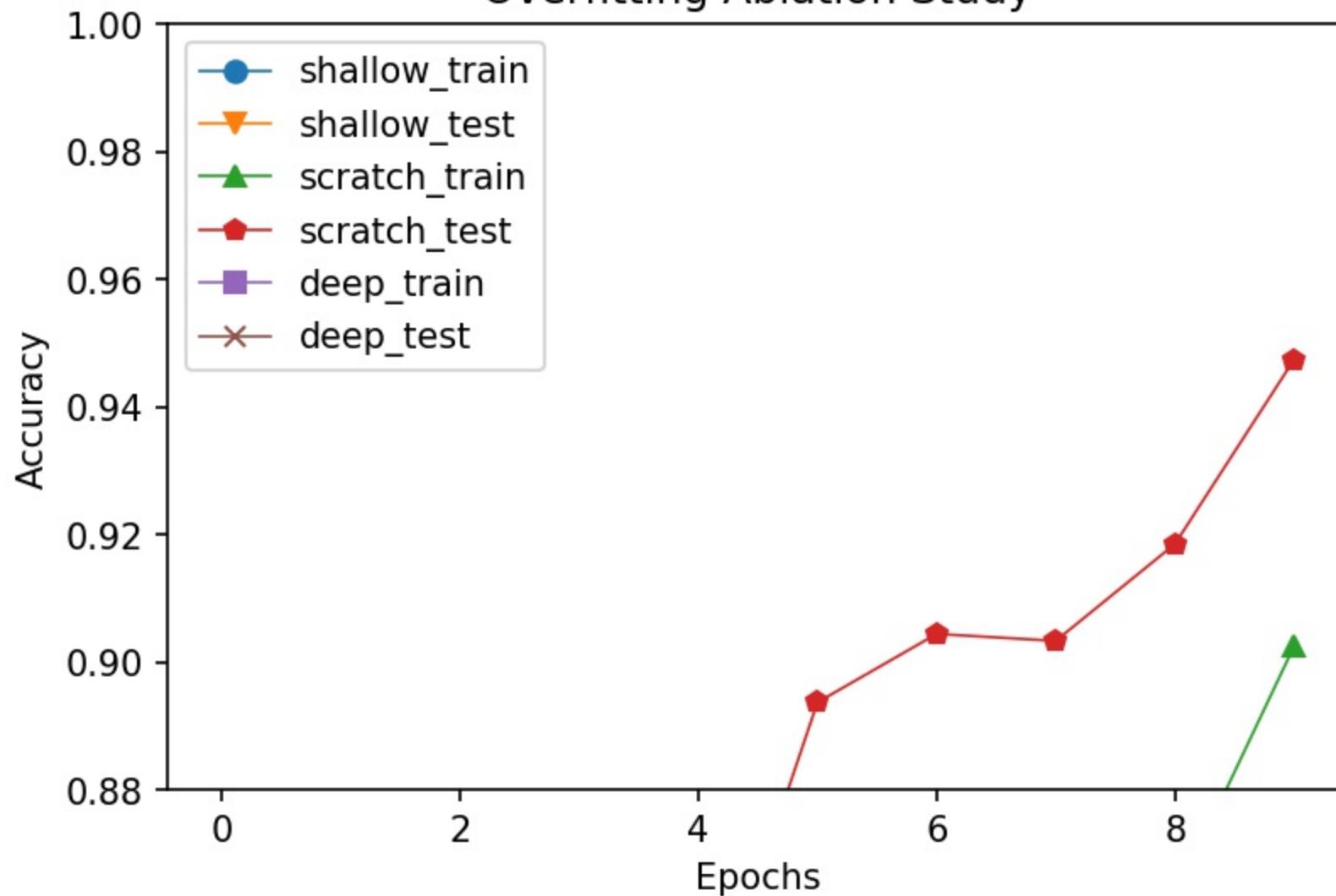
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 9 takes 317 seconds.

[11, 30] loss: 0.301

[11, 60] loss: 0.430

```
[11,    90] loss: 0.371
[11,   120] loss: 0.370
[11,   150] loss: 0.388
[11,   180] loss: 0.346
[11,   210] loss: 0.314
[11,   240] loss: 0.382
[11,   270] loss: 0.359
[11,   300] loss: 0.337
[11,   330] loss: 0.297
[11,   360] loss: 0.366
[11,   390] loss: 0.407
[11,   420] loss: 0.395
[11,   450] loss: 0.422
[11,   480] loss: 0.401
[11,   510] loss: 0.368
[11,   540] loss: 0.355
[11,   570] loss: 0.327
[11,   600] loss: 0.363
[11,   630] loss: 0.319
[11,   660] loss: 0.315
[11,   690] loss: 0.424
[11,   720] loss: 0.317
[11,   750] loss: 0.409
[11,   780] loss: 0.437
[11,   810] loss: 0.362
[11,   840] loss: 0.365
[11,   870] loss: 0.367
[11,   900] loss: 0.440
[11,   930] loss: 0.445
[11,   960] loss: 0.395
[11,   990] loss: 0.398
[11,  1020] loss: 0.331
[11,  1050] loss: 0.292
[11,  1080] loss: 0.340
[11,  1110] loss: 0.377
[11,  1140] loss: 0.426
[11,  1170] loss: 0.301
[11,  1200] loss: 0.347
[11,  1230] loss: 0.413
```

```
[11, 1260] loss: 0.310
[11, 1290] loss: 0.377
[11, 1320] loss: 0.374
[11, 1350] loss: 0.381
[11, 1380] loss: 0.275
[11, 1410] loss: 0.329
[11, 1440] loss: 0.348
[11, 1470] loss: 0.396
[11, 1500] loss: 0.406
[11, 1530] loss: 0.378
[11, 1560] loss: 0.333
[11, 1590] loss: 0.369
[11, 1620] loss: 0.321
[11, 1650] loss: 0.377
[11, 1680] loss: 0.394
[11, 1710] loss: 0.353
[11, 1740] loss: 0.348
[11, 1770] loss: 0.371
[11, 1800] loss: 0.382
[11, 1830] loss: 0.338
[11, 1860] loss: 0.449
[11, 1890] loss: 0.388
[11, 1920] loss: 0.349
[11, 1950] loss: 0.364
[11, 1980] loss: 0.354
[11, 2010] loss: 0.370
[11, 2040] loss: 0.395
[11, 2070] loss: 0.290
[11, 2100] loss: 0.321
[11, 2130] loss: 0.306
[11, 2160] loss: 0.357
[11, 2190] loss: 0.370
evf intermediate_models/epoch10_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.911964
epoch 10 accuracy on train set is: 0.9119638313340603
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.955834

epoch 10 accuracy on test set is: 0.9558342420937841

evaluating on test set takes 16 seconds.

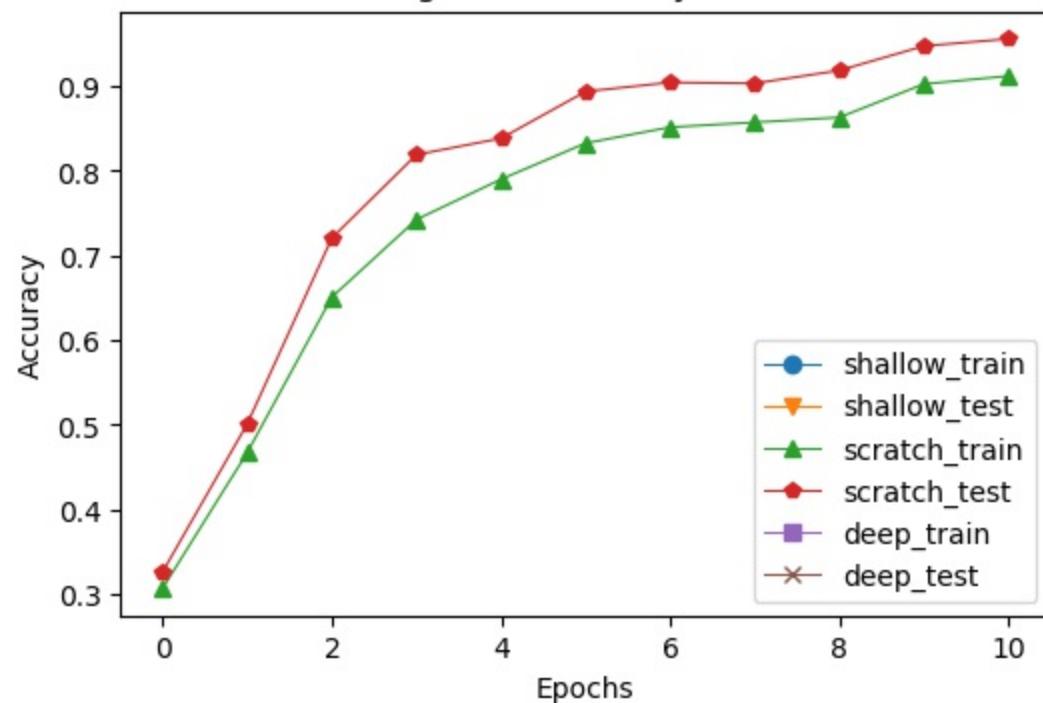
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

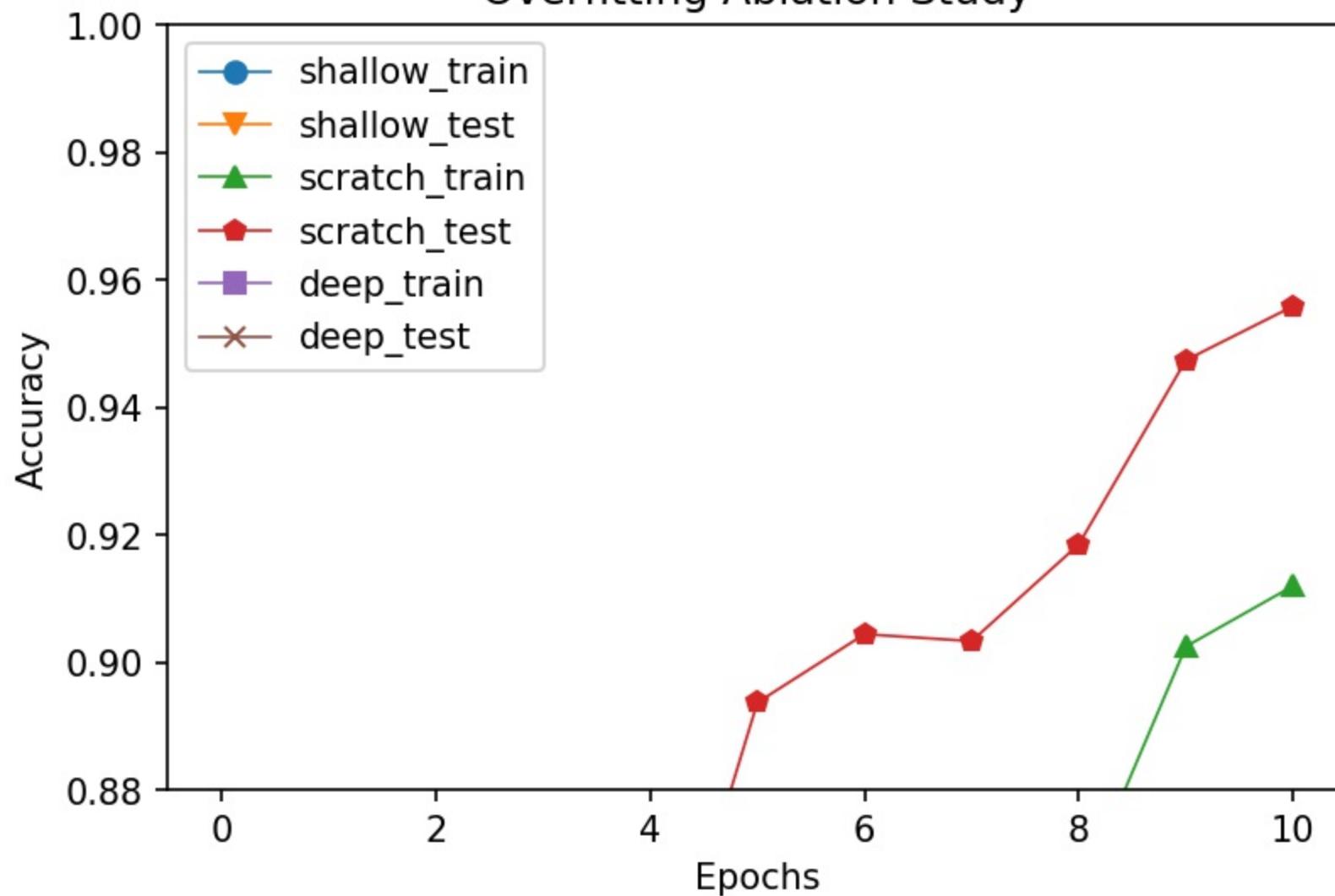
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view



## Overfitting Ablation Study



epoch 10 takes 318 seconds.

[12, 30] loss: 0.410

[12, 60] loss: 0.340

```
[12,    90] loss: 0.353
[12,   120] loss: 0.352
[12,   150] loss: 0.326
[12,   180] loss: 0.326
[12,   210] loss: 0.333
[12,   240] loss: 0.397
[12,   270] loss: 0.339
[12,   300] loss: 0.363
[12,   330] loss: 0.321
[12,   360] loss: 0.348
[12,   390] loss: 0.311
[12,   420] loss: 0.354
[12,   450] loss: 0.312
[12,   480] loss: 0.363
[12,   510] loss: 0.373
[12,   540] loss: 0.374
[12,   570] loss: 0.395
[12,   600] loss: 0.368
[12,   630] loss: 0.370
[12,   660] loss: 0.306
[12,   690] loss: 0.312
[12,   720] loss: 0.356
[12,   750] loss: 0.340
[12,   780] loss: 0.313
[12,   810] loss: 0.385
[12,   840] loss: 0.368
[12,   870] loss: 0.283
[12,   900] loss: 0.319
[12,   930] loss: 0.327
[12,   960] loss: 0.359
[12,   990] loss: 0.320
[12,  1020] loss: 0.354
[12,  1050] loss: 0.335
[12,  1080] loss: 0.361
[12,  1110] loss: 0.343
[12,  1140] loss: 0.336
[12,  1170] loss: 0.370
[12,  1200] loss: 0.400
[12,  1230] loss: 0.302
```

```
[12, 1260] loss: 0.337
[12, 1290] loss: 0.299
[12, 1320] loss: 0.317
[12, 1350] loss: 0.325
[12, 1380] loss: 0.387
[12, 1410] loss: 0.360
[12, 1440] loss: 0.407
[12, 1470] loss: 0.272
[12, 1500] loss: 0.374
[12, 1530] loss: 0.305
[12, 1560] loss: 0.338
[12, 1590] loss: 0.325
[12, 1620] loss: 0.303
[12, 1650] loss: 0.357
[12, 1680] loss: 0.268
[12, 1710] loss: 0.329
[12, 1740] loss: 0.355
[12, 1770] loss: 0.364
[12, 1800] loss: 0.327
[12, 1830] loss: 0.297
[12, 1860] loss: 0.335
[12, 1890] loss: 0.399
[12, 1920] loss: 0.384
[12, 1950] loss: 0.359
[12, 1980] loss: 0.304
[12, 2010] loss: 0.291
[12, 2040] loss: 0.259
[12, 2070] loss: 0.334
[12, 2100] loss: 0.287
[12, 2130] loss: 0.286
[12, 2160] loss: 0.302
[12, 2190] loss: 0.315
eval intermediate_models/epoch11_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.922574
epoch 11 accuracy on train set is: 0.9225736095965104
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.954653

epoch 11 accuracy on test set is: 0.9546528535078154

evaluating on test set takes 16 seconds.

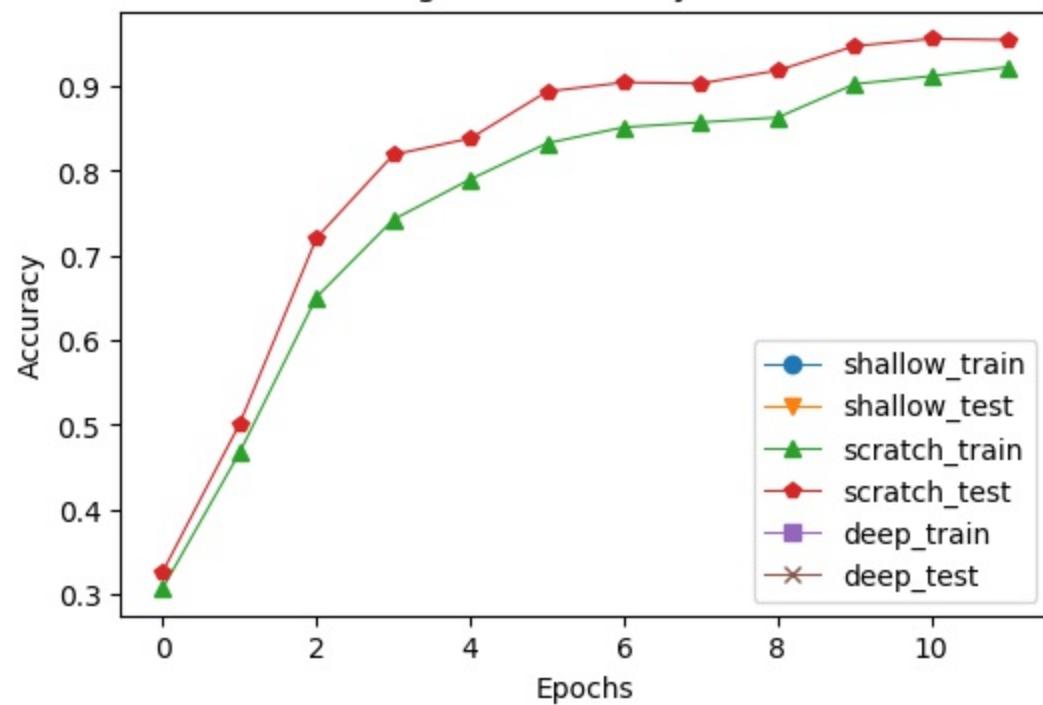
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

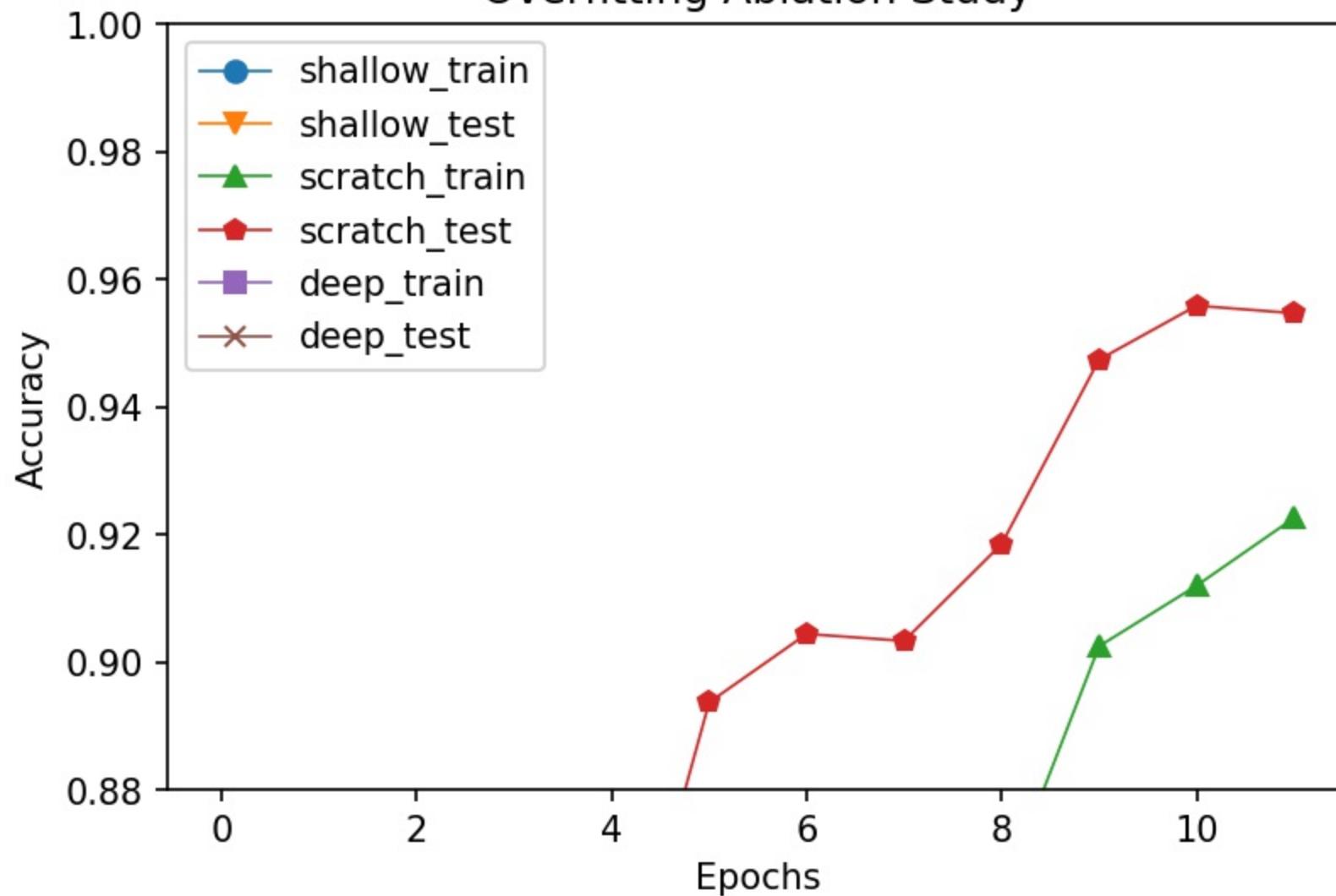
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view



## Overfitting Ablation Study



```
epoch 11 takes 318 seconds.  
[13,      30] loss: 0.336  
[13,      60] loss: 0.282
```

```
[13,    90] loss: 0.322
[13,   120] loss: 0.282
[13,   150] loss: 0.250
[13,   180] loss: 0.339
[13,   210] loss: 0.310
[13,   240] loss: 0.316
[13,   270] loss: 0.259
[13,   300] loss: 0.341
[13,   330] loss: 0.438
[13,   360] loss: 0.336
[13,   390] loss: 0.310
[13,   420] loss: 0.332
[13,   450] loss: 0.369
[13,   480] loss: 0.323
[13,   510] loss: 0.267
[13,   540] loss: 0.307
[13,   570] loss: 0.332
[13,   600] loss: 0.301
[13,   630] loss: 0.325
[13,   660] loss: 0.343
[13,   690] loss: 0.280
[13,   720] loss: 0.341
[13,   750] loss: 0.314
[13,   780] loss: 0.317
[13,   810] loss: 0.314
[13,   840] loss: 0.293
[13,   870] loss: 0.307
[13,   900] loss: 0.244
[13,   930] loss: 0.364
[13,   960] loss: 0.315
[13,   990] loss: 0.287
[13,  1020] loss: 0.307
[13,  1050] loss: 0.262
[13,  1080] loss: 0.339
[13,  1110] loss: 0.358
[13,  1140] loss: 0.265
[13,  1170] loss: 0.233
[13,  1200] loss: 0.297
[13,  1230] loss: 0.298
```

```
[13, 1260] loss: 0.269
[13, 1290] loss: 0.266
[13, 1320] loss: 0.292
[13, 1350] loss: 0.314
[13, 1380] loss: 0.312
[13, 1410] loss: 0.334
[13, 1440] loss: 0.330
[13, 1470] loss: 0.305
[13, 1500] loss: 0.304
[13, 1530] loss: 0.295
[13, 1560] loss: 0.258
[13, 1590] loss: 0.278
[13, 1620] loss: 0.312
[13, 1650] loss: 0.346
[13, 1680] loss: 0.333
[13, 1710] loss: 0.268
[13, 1740] loss: 0.341
[13, 1770] loss: 0.334
[13, 1800] loss: 0.299
[13, 1830] loss: 0.290
[13, 1860] loss: 0.284
[13, 1890] loss: 0.290
[13, 1920] loss: 0.313
[13, 1950] loss: 0.277
[13, 1980] loss: 0.250
[13, 2010] loss: 0.320
[13, 2040] loss: 0.383
[13, 2070] loss: 0.366
[13, 2100] loss: 0.343
[13, 2130] loss: 0.307
[13, 2160] loss: 0.341
[13, 2190] loss: 0.299
eval intermediate_models/epoch12_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.924414
epoch 12 accuracy on train set is: 0.9244138495092693
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.958288

epoch 12 accuracy on test set is: 0.9582878953107961

evaluating on test set takes 16 seconds.

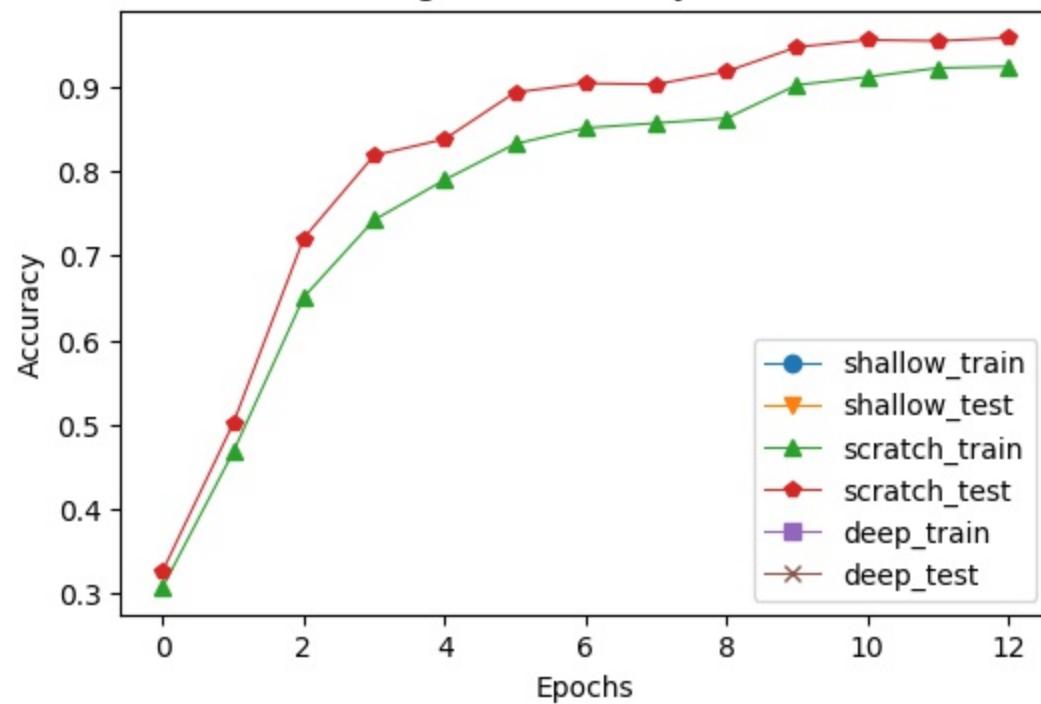
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

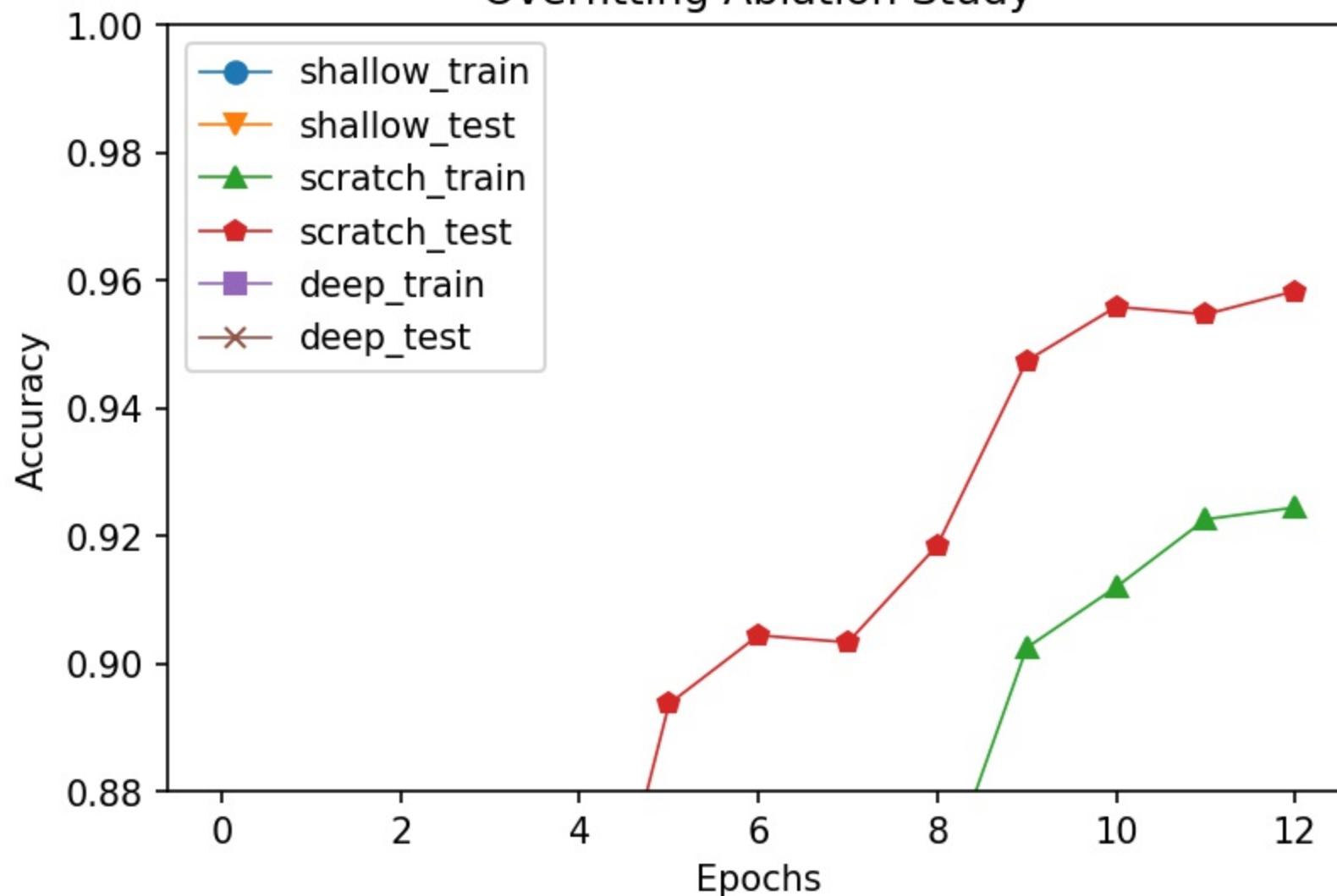
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view



## Overfitting Ablation Study



epoch 12 takes 317 seconds.

[14, 30] loss: 0.255

[14, 60] loss: 0.312

```
[14,    90] loss: 0.318
[14,   120] loss: 0.328
[14,   150] loss: 0.253
[14,   180] loss: 0.249
[14,   210] loss: 0.269
[14,   240] loss: 0.245
[14,   270] loss: 0.220
[14,   300] loss: 0.294
[14,   330] loss: 0.272
[14,   360] loss: 0.327
[14,   390] loss: 0.332
[14,   420] loss: 0.314
[14,   450] loss: 0.248
[14,   480] loss: 0.253
[14,   510] loss: 0.245
[14,   540] loss: 0.286
[14,   570] loss: 0.287
[14,   600] loss: 0.229
[14,   630] loss: 0.277
[14,   660] loss: 0.311
[14,   690] loss: 0.250
[14,   720] loss: 0.319
[14,   750] loss: 0.290
[14,   780] loss: 0.304
[14,   810] loss: 0.301
[14,   840] loss: 0.275
[14,   870] loss: 0.341
[14,   900] loss: 0.322
[14,   930] loss: 0.320
[14,   960] loss: 0.262
[14,   990] loss: 0.258
[14,  1020] loss: 0.340
[14,  1050] loss: 0.342
[14,  1080] loss: 0.265
[14,  1110] loss: 0.264
[14,  1140] loss: 0.260
[14,  1170] loss: 0.319
[14,  1200] loss: 0.347
[14,  1230] loss: 0.232
```

```
[14, 1260] loss: 0.300
[14, 1290] loss: 0.247
[14, 1320] loss: 0.256
[14, 1350] loss: 0.256
[14, 1380] loss: 0.386
[14, 1410] loss: 0.283
[14, 1440] loss: 0.274
[14, 1470] loss: 0.292
[14, 1500] loss: 0.310
[14, 1530] loss: 0.261
[14, 1560] loss: 0.296
[14, 1590] loss: 0.315
[14, 1620] loss: 0.276
[14, 1650] loss: 0.292
[14, 1680] loss: 0.262
[14, 1710] loss: 0.246
[14, 1740] loss: 0.258
[14, 1770] loss: 0.298
[14, 1800] loss: 0.300
[14, 1830] loss: 0.291
[14, 1860] loss: 0.310
[14, 1890] loss: 0.307
[14, 1920] loss: 0.201
[14, 1950] loss: 0.270
[14, 1980] loss: 0.279
[14, 2010] loss: 0.263
[14, 2040] loss: 0.310
[14, 2070] loss: 0.252
[14, 2100] loss: 0.306
[14, 2130] loss: 0.257
[14, 2160] loss: 0.278
[14, 2190] loss: 0.200
eval intermediate_models/epoch13_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.927435
epoch 13 accuracy on train set is: 0.9274354780079971
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.960287

epoch 13 accuracy on test set is: 0.9602871683024354

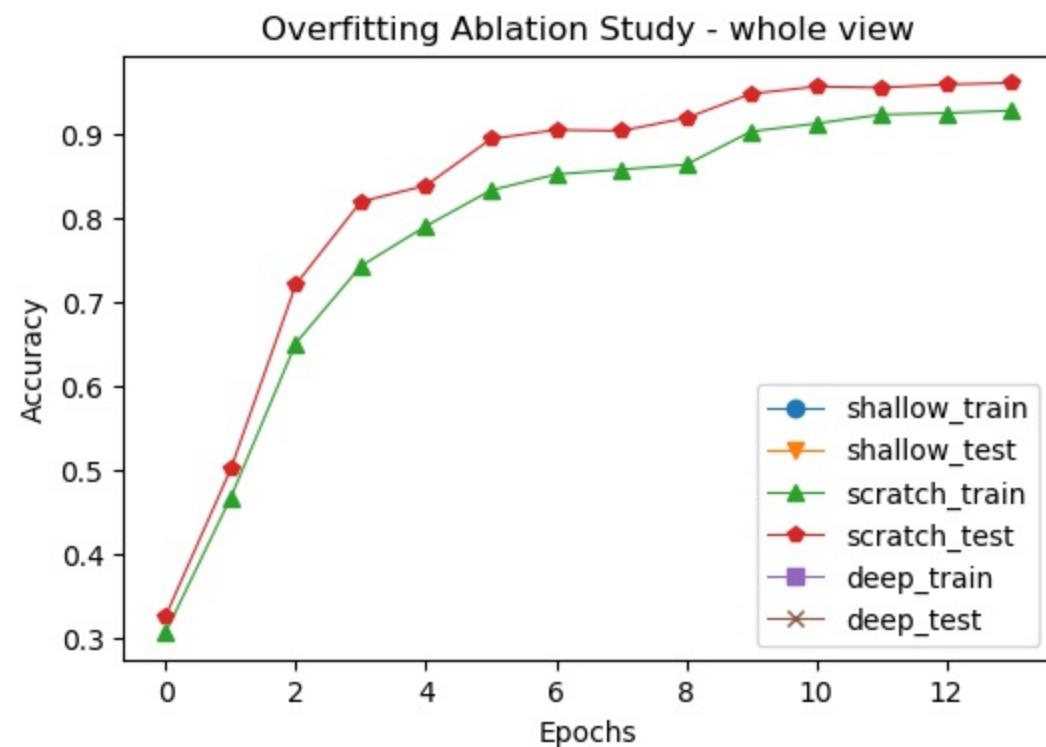
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

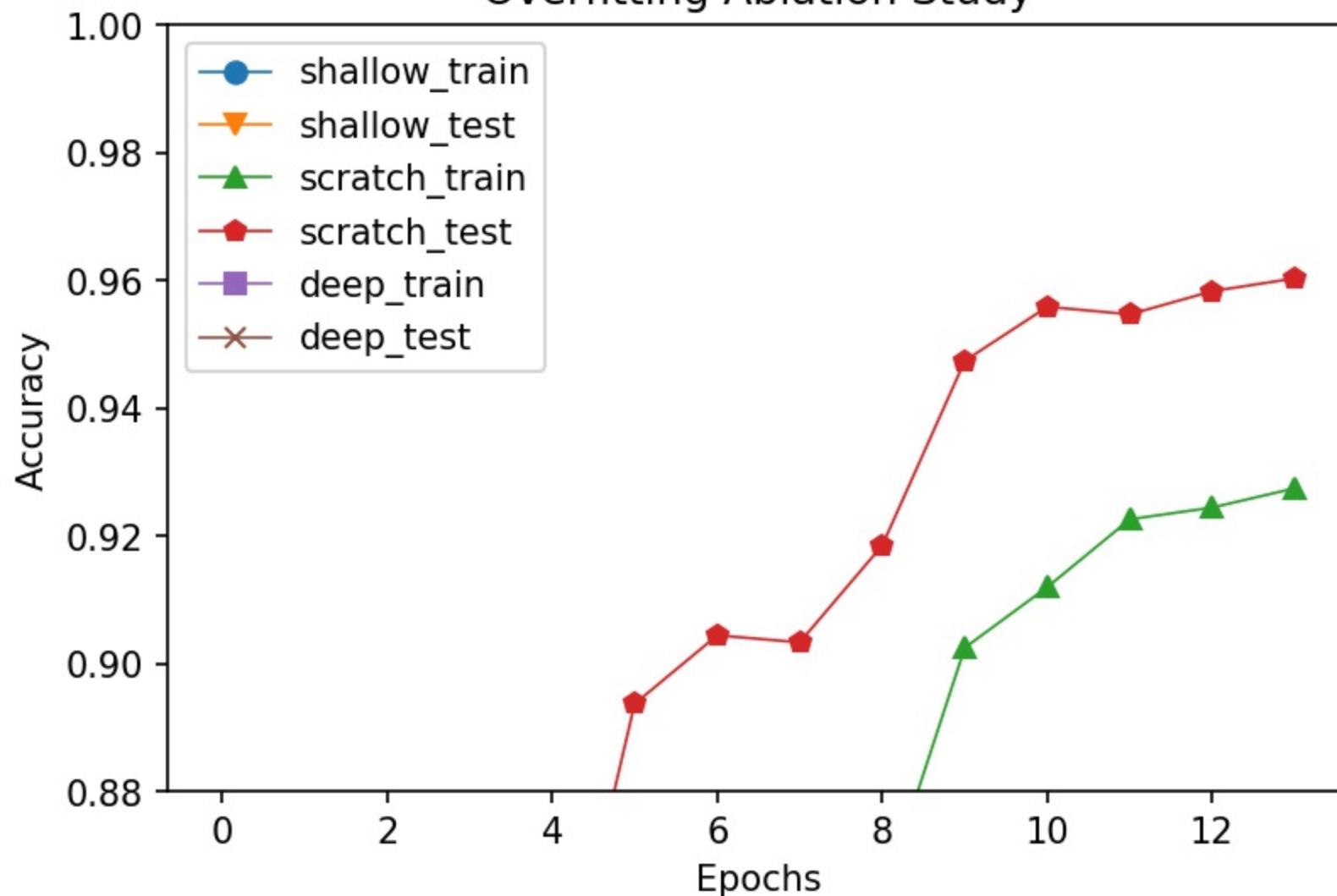
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 13 takes 318 seconds.

[15, 30] loss: 0.189

[15, 60] loss: 0.213

```
[15,    90] loss: 0.267
[15,   120] loss: 0.294
[15,   150] loss: 0.279
[15,   180] loss: 0.293
[15,   210] loss: 0.255
[15,   240] loss: 0.296
[15,   270] loss: 0.293
[15,   300] loss: 0.274
[15,   330] loss: 0.310
[15,   360] loss: 0.213
[15,   390] loss: 0.273
[15,   420] loss: 0.365
[15,   450] loss: 0.320
[15,   480] loss: 0.272
[15,   510] loss: 0.272
[15,   540] loss: 0.267
[15,   570] loss: 0.226
[15,   600] loss: 0.295
[15,   630] loss: 0.287
[15,   660] loss: 0.305
[15,   690] loss: 0.253
[15,   720] loss: 0.297
[15,   750] loss: 0.261
[15,   780] loss: 0.292
[15,   810] loss: 0.316
[15,   840] loss: 0.250
[15,   870] loss: 0.264
[15,   900] loss: 0.279
[15,   930] loss: 0.267
[15,   960] loss: 0.239
[15,   990] loss: 0.237
[15,  1020] loss: 0.334
[15,  1050] loss: 0.325
[15,  1080] loss: 0.321
[15,  1110] loss: 0.236
[15,  1140] loss: 0.241
[15,  1170] loss: 0.335
[15,  1200] loss: 0.276
[15,  1230] loss: 0.232
```

```
[15, 1260] loss: 0.274
[15, 1290] loss: 0.243
[15, 1320] loss: 0.275
[15, 1350] loss: 0.304
[15, 1380] loss: 0.338
[15, 1410] loss: 0.269
[15, 1440] loss: 0.259
[15, 1470] loss: 0.260
[15, 1500] loss: 0.321
[15, 1530] loss: 0.259
[15, 1560] loss: 0.246
[15, 1590] loss: 0.204
[15, 1620] loss: 0.303
[15, 1650] loss: 0.235
[15, 1680] loss: 0.279
[15, 1710] loss: 0.227
[15, 1740] loss: 0.327
[15, 1770] loss: 0.229
[15, 1800] loss: 0.273
[15, 1830] loss: 0.231
[15, 1860] loss: 0.297
[15, 1890] loss: 0.250
[15, 1920] loss: 0.244
[15, 1950] loss: 0.282
[15, 1980] loss: 0.298
[15, 2010] loss: 0.310
[15, 2040] loss: 0.364
[15, 2070] loss: 0.201
[15, 2100] loss: 0.236
[15, 2130] loss: 0.258
[15, 2160] loss: 0.210
[15, 2190] loss: 0.294
eval intermediate_models/epoch14_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.934251
epoch 14 accuracy on train set is: 0.934251181388586
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.969920

epoch 14 accuracy on test set is: 0.9699200290803345

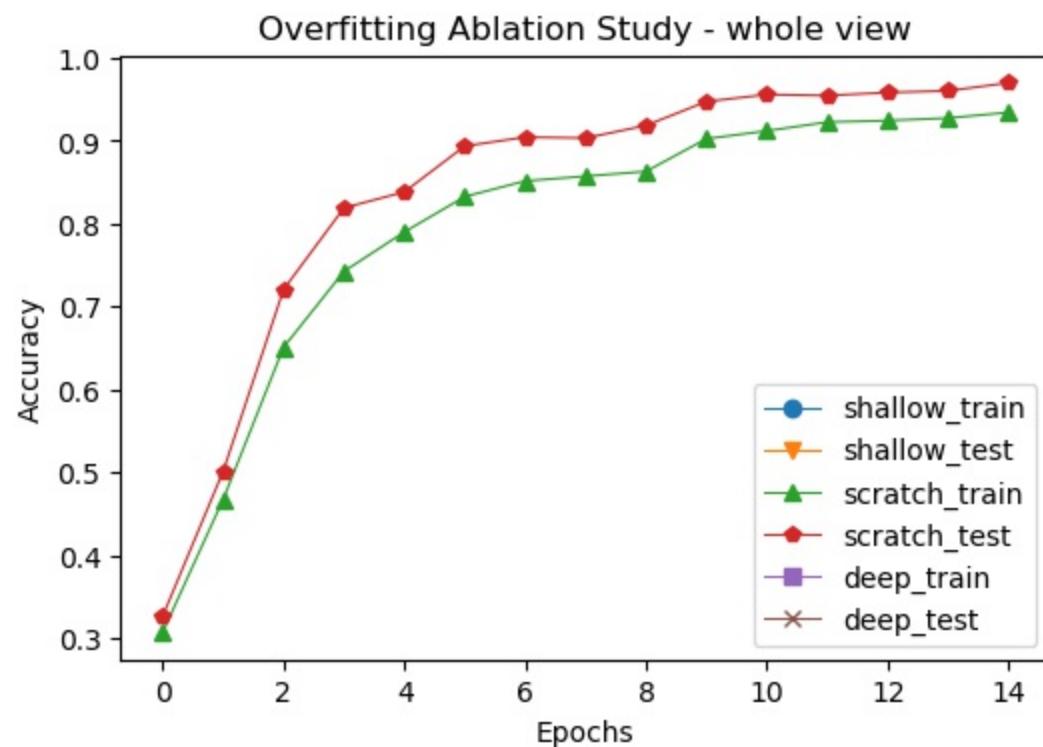
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

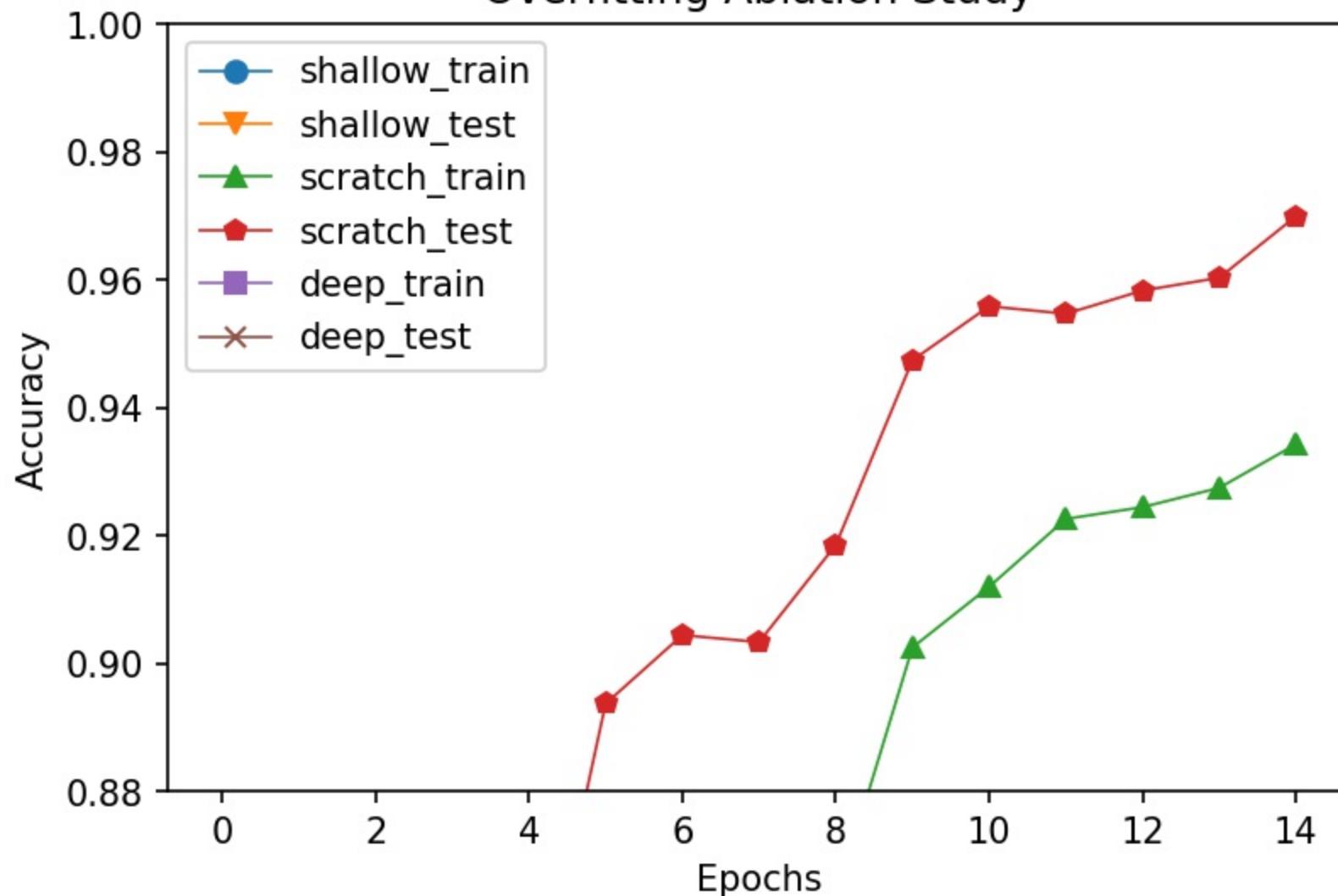
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 14 takes 318 seconds.

[16, 30] loss: 0.242

[16, 60] loss: 0.242

```
[16,    90] loss: 0.299
[16,   120] loss: 0.277
[16,   150] loss: 0.215
[16,   180] loss: 0.195
[16,   210] loss: 0.265
[16,   240] loss: 0.271
[16,   270] loss: 0.234
[16,   300] loss: 0.284
[16,   330] loss: 0.280
[16,   360] loss: 0.208
[16,   390] loss: 0.297
[16,   420] loss: 0.284
[16,   450] loss: 0.217
[16,   480] loss: 0.269
[16,   510] loss: 0.301
[16,   540] loss: 0.217
[16,   570] loss: 0.289
[16,   600] loss: 0.255
[16,   630] loss: 0.263
[16,   660] loss: 0.276
[16,   690] loss: 0.266
[16,   720] loss: 0.231
[16,   750] loss: 0.288
[16,   780] loss: 0.275
[16,   810] loss: 0.283
[16,   840] loss: 0.274
[16,   870] loss: 0.240
[16,   900] loss: 0.263
[16,   930] loss: 0.245
[16,   960] loss: 0.269
[16,   990] loss: 0.278
[16,  1020] loss: 0.217
[16,  1050] loss: 0.254
[16,  1080] loss: 0.279
[16,  1110] loss: 0.270
[16,  1140] loss: 0.267
[16,  1170] loss: 0.247
[16,  1200] loss: 0.245
[16,  1230] loss: 0.299
```

```
[16, 1260] loss: 0.277
[16, 1290] loss: 0.238
[16, 1320] loss: 0.252
[16, 1350] loss: 0.213
[16, 1380] loss: 0.201
[16, 1410] loss: 0.259
[16, 1440] loss: 0.262
[16, 1470] loss: 0.248
[16, 1500] loss: 0.249
[16, 1530] loss: 0.226
[16, 1560] loss: 0.228
[16, 1590] loss: 0.260
[16, 1620] loss: 0.300
[16, 1650] loss: 0.285
[16, 1680] loss: 0.220
[16, 1710] loss: 0.189
[16, 1740] loss: 0.263
[16, 1770] loss: 0.252
[16, 1800] loss: 0.199
[16, 1830] loss: 0.282
[16, 1860] loss: 0.316
[16, 1890] loss: 0.234
[16, 1920] loss: 0.211
[16, 1950] loss: 0.275
[16, 1980] loss: 0.268
[16, 2010] loss: 0.292
[16, 2040] loss: 0.260
[16, 2070] loss: 0.277
[16, 2100] loss: 0.302
[16, 2130] loss: 0.268
[16, 2160] loss: 0.255
[16, 2190] loss: 0.217
evf intermediate_models/epoch15_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.942112
epoch 15 accuracy on train set is: 0.9421119592875318
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.971374

epoch 15 accuracy on test set is: 0.9713740458015268

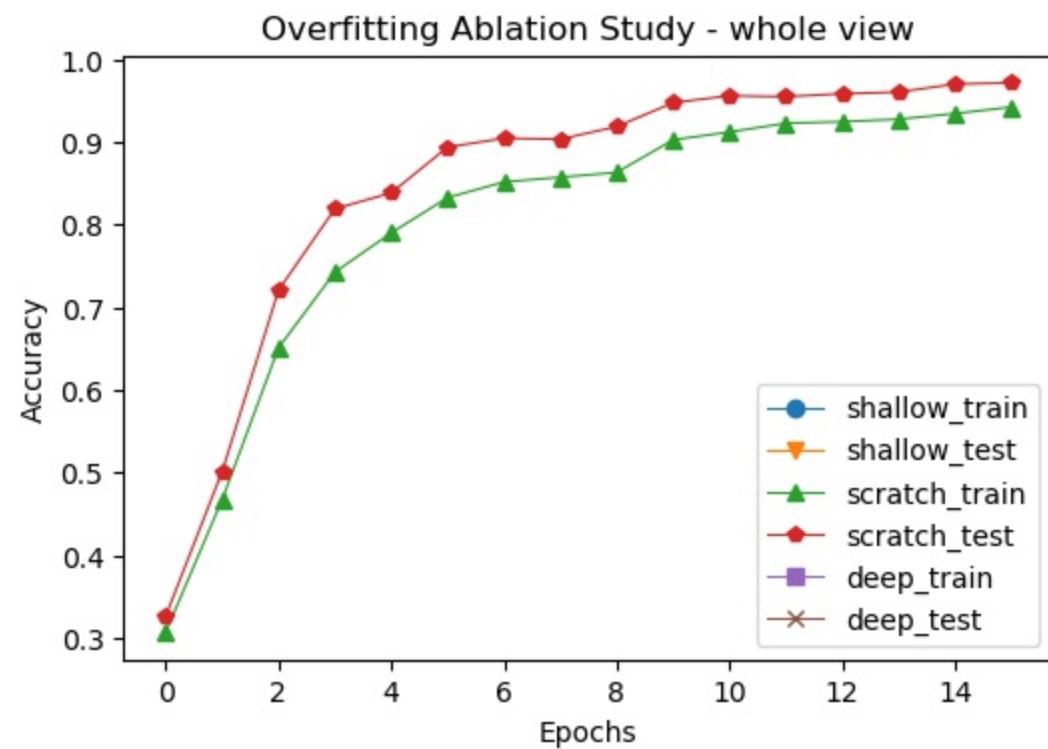
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

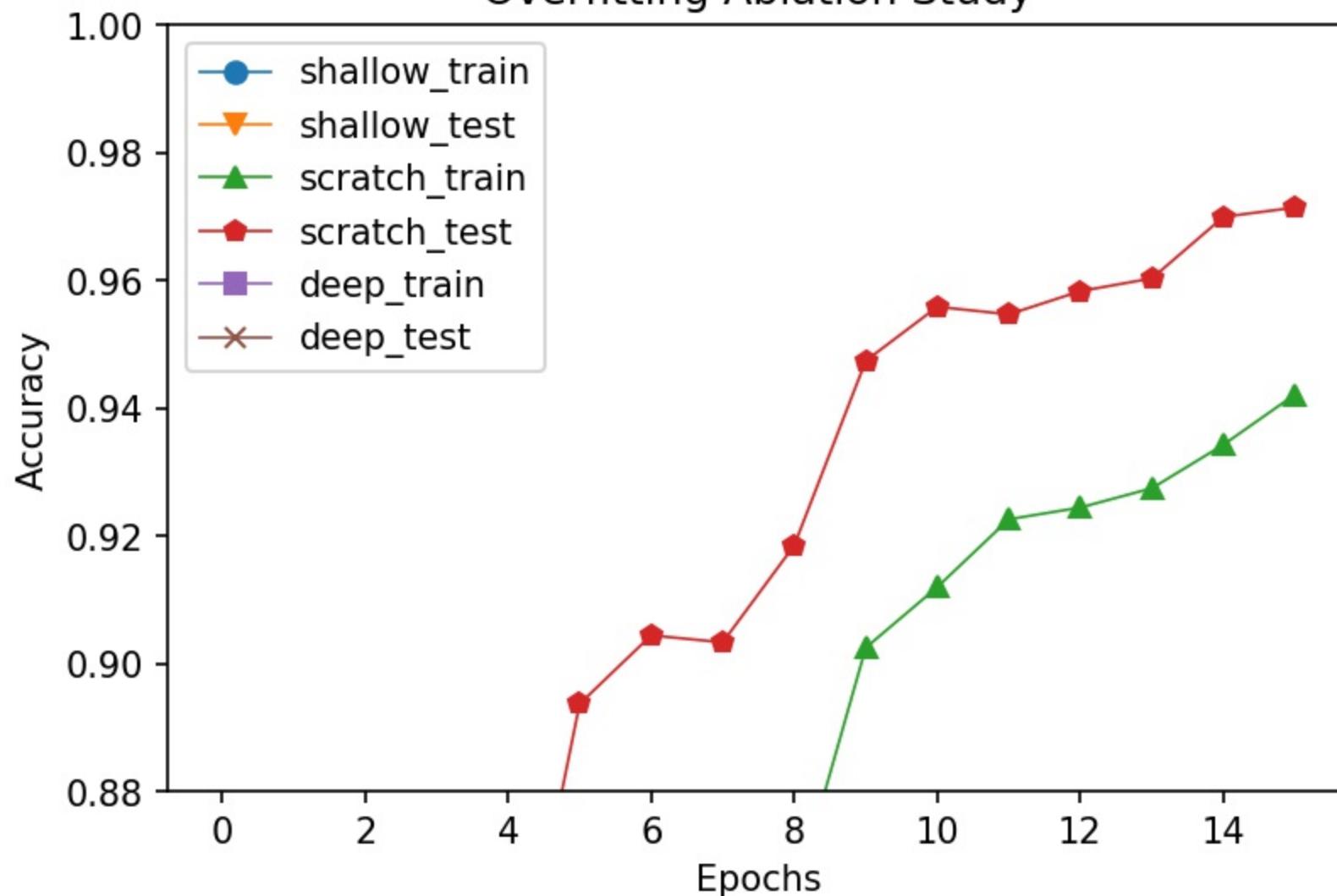
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 15 takes 317 seconds.

[17, 30] loss: 0.267

[17, 60] loss: 0.202

```
[17,    90] loss: 0.235
[17,   120] loss: 0.293
[17,   150] loss: 0.239
[17,   180] loss: 0.215
[17,   210] loss: 0.216
[17,   240] loss: 0.287
[17,   270] loss: 0.250
[17,   300] loss: 0.308
[17,   330] loss: 0.255
[17,   360] loss: 0.271
[17,   390] loss: 0.242
[17,   420] loss: 0.225
[17,   450] loss: 0.215
[17,   480] loss: 0.247
[17,   510] loss: 0.239
[17,   540] loss: 0.249
[17,   570] loss: 0.178
[17,   600] loss: 0.211
[17,   630] loss: 0.224
[17,   660] loss: 0.332
[17,   690] loss: 0.273
[17,   720] loss: 0.249
[17,   750] loss: 0.231
[17,   780] loss: 0.262
[17,   810] loss: 0.249
[17,   840] loss: 0.248
[17,   870] loss: 0.272
[17,   900] loss: 0.300
[17,   930] loss: 0.242
[17,   960] loss: 0.237
[17,   990] loss: 0.232
[17,  1020] loss: 0.250
[17,  1050] loss: 0.214
[17,  1080] loss: 0.236
[17,  1110] loss: 0.250
[17,  1140] loss: 0.301
[17,  1170] loss: 0.281
[17,  1200] loss: 0.222
[17,  1230] loss: 0.241
```

```
[17, 1260] loss: 0.298
[17, 1290] loss: 0.188
[17, 1320] loss: 0.194
[17, 1350] loss: 0.283
[17, 1380] loss: 0.281
[17, 1410] loss: 0.336
[17, 1440] loss: 0.196
[17, 1470] loss: 0.189
[17, 1500] loss: 0.237
[17, 1530] loss: 0.247
[17, 1560] loss: 0.228
[17, 1590] loss: 0.284
[17, 1620] loss: 0.249
[17, 1650] loss: 0.210
[17, 1680] loss: 0.212
[17, 1710] loss: 0.268
[17, 1740] loss: 0.243
[17, 1770] loss: 0.253
[17, 1800] loss: 0.167
[17, 1830] loss: 0.198
[17, 1860] loss: 0.240
[17, 1890] loss: 0.185
[17, 1920] loss: 0.260
[17, 1950] loss: 0.184
[17, 1980] loss: 0.183
[17, 2010] loss: 0.221
[17, 2040] loss: 0.190
[17, 2070] loss: 0.206
[17, 2100] loss: 0.236
[17, 2130] loss: 0.294
[17, 2160] loss: 0.262
[17, 2190] loss: 0.271
eval intermediate_models/epoch16_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.945020
epoch 16 accuracy on train set is: 0.9450199927299164
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.976009

epoch 16 accuracy on test set is: 0.9760087241003271

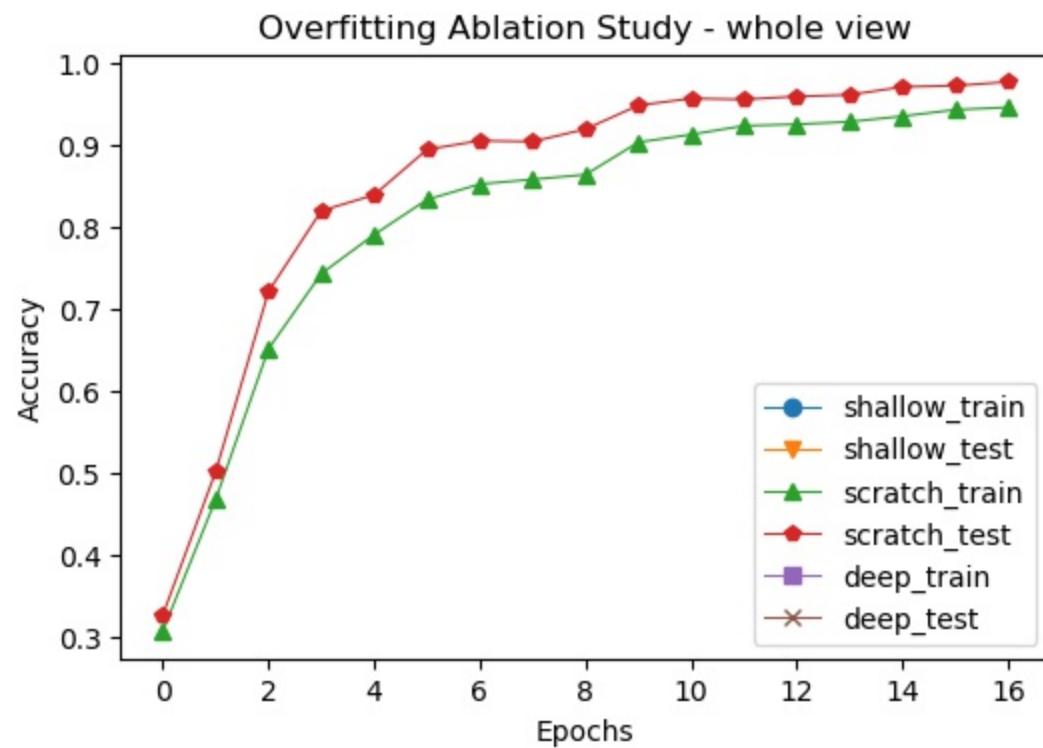
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

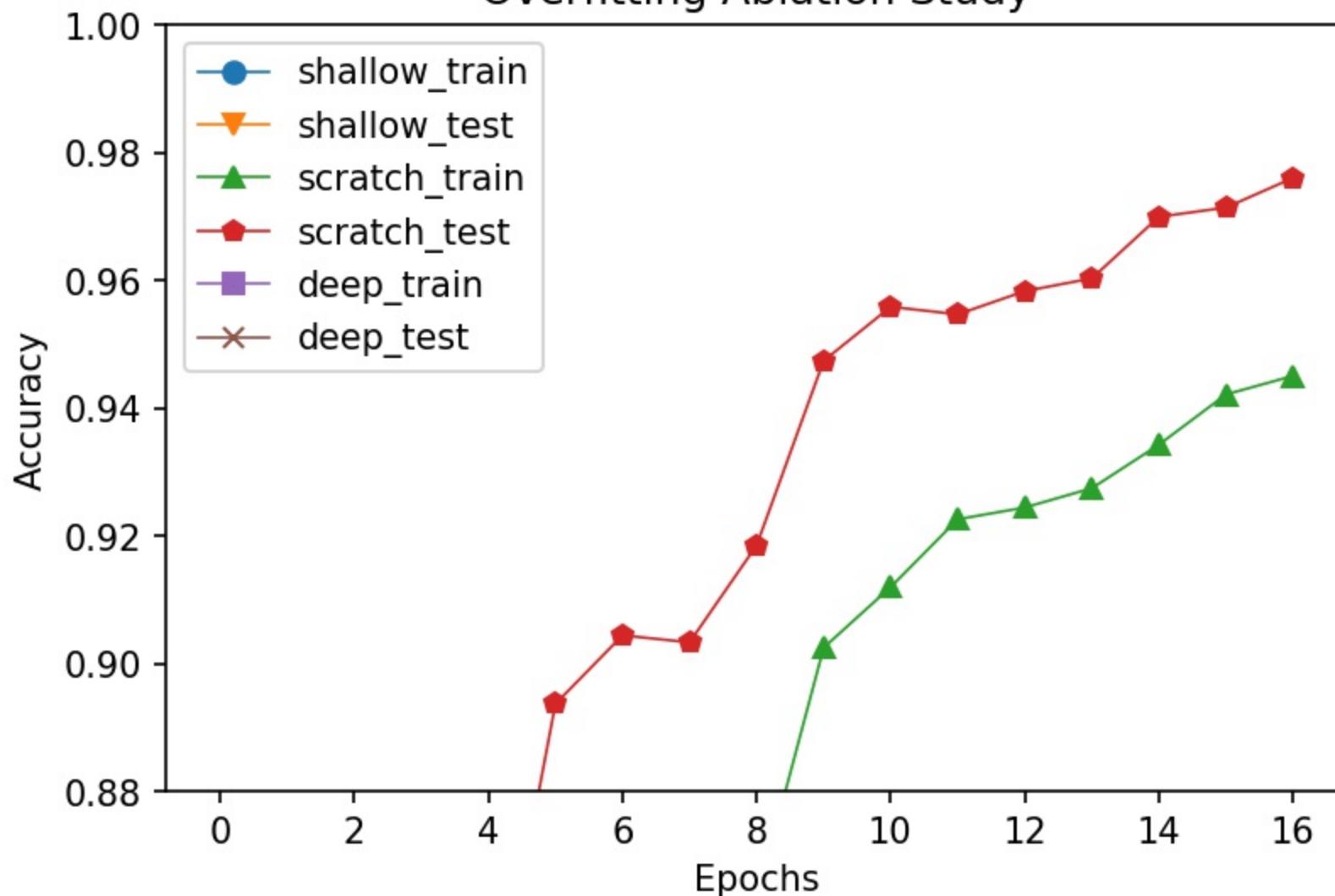
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 16 takes 318 seconds.

[18, 30] loss: 0.225

[18, 60] loss: 0.227

```
[18,    90] loss: 0.201
[18,   120] loss: 0.222
[18,   150] loss: 0.233
[18,   180] loss: 0.207
[18,   210] loss: 0.260
[18,   240] loss: 0.276
[18,   270] loss: 0.224
[18,   300] loss: 0.175
[18,   330] loss: 0.159
[18,   360] loss: 0.162
[18,   390] loss: 0.224
[18,   420] loss: 0.264
[18,   450] loss: 0.277
[18,   480] loss: 0.222
[18,   510] loss: 0.234
[18,   540] loss: 0.276
[18,   570] loss: 0.262
[18,   600] loss: 0.144
[18,   630] loss: 0.213
[18,   660] loss: 0.228
[18,   690] loss: 0.245
[18,   720] loss: 0.248
[18,   750] loss: 0.230
[18,   780] loss: 0.282
[18,   810] loss: 0.184
[18,   840] loss: 0.200
[18,   870] loss: 0.226
[18,   900] loss: 0.255
[18,   930] loss: 0.324
[18,   960] loss: 0.215
[18,   990] loss: 0.203
[18,  1020] loss: 0.202
[18,  1050] loss: 0.192
[18,  1080] loss: 0.221
[18,  1110] loss: 0.233
[18,  1140] loss: 0.216
[18,  1170] loss: 0.234
[18,  1200] loss: 0.269
[18,  1230] loss: 0.232
```

```
[18, 1260] loss: 0.289
[18, 1290] loss: 0.224
[18, 1320] loss: 0.243
[18, 1350] loss: 0.247
[18, 1380] loss: 0.154
[18, 1410] loss: 0.259
[18, 1440] loss: 0.250
[18, 1470] loss: 0.180
[18, 1500] loss: 0.252
[18, 1530] loss: 0.198
[18, 1560] loss: 0.201
[18, 1590] loss: 0.208
[18, 1620] loss: 0.217
[18, 1650] loss: 0.249
[18, 1680] loss: 0.239
[18, 1710] loss: 0.217
[18, 1740] loss: 0.260
[18, 1770] loss: 0.351
[18, 1800] loss: 0.210
[18, 1830] loss: 0.211
[18, 1860] loss: 0.286
[18, 1890] loss: 0.215
[18, 1920] loss: 0.171
[18, 1950] loss: 0.240
[18, 1980] loss: 0.274
[18, 2010] loss: 0.229
[18, 2040] loss: 0.235
[18, 2070] loss: 0.224
[18, 2100] loss: 0.217
[18, 2130] loss: 0.185
[18, 2160] loss: 0.192
[18, 2190] loss: 0.261
eval intermediate_models/epoch17_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.935955
epoch 17 accuracy on train set is: 0.9359551072337332
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.973555

epoch 17 accuracy on test set is: 0.9735550708833152

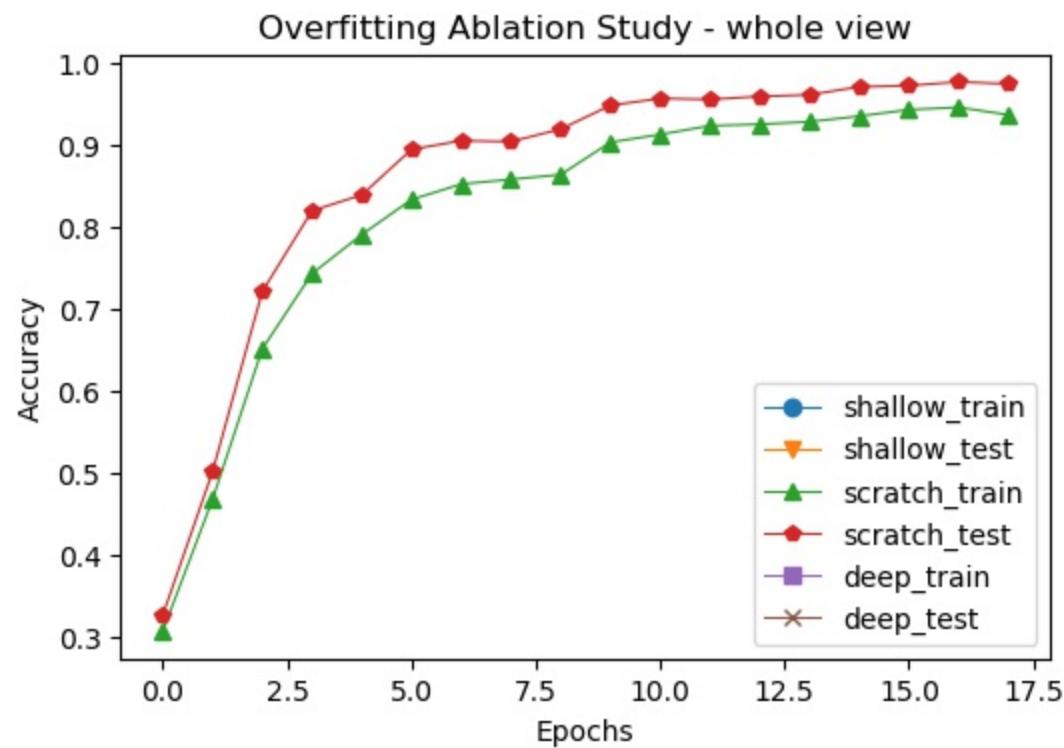
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

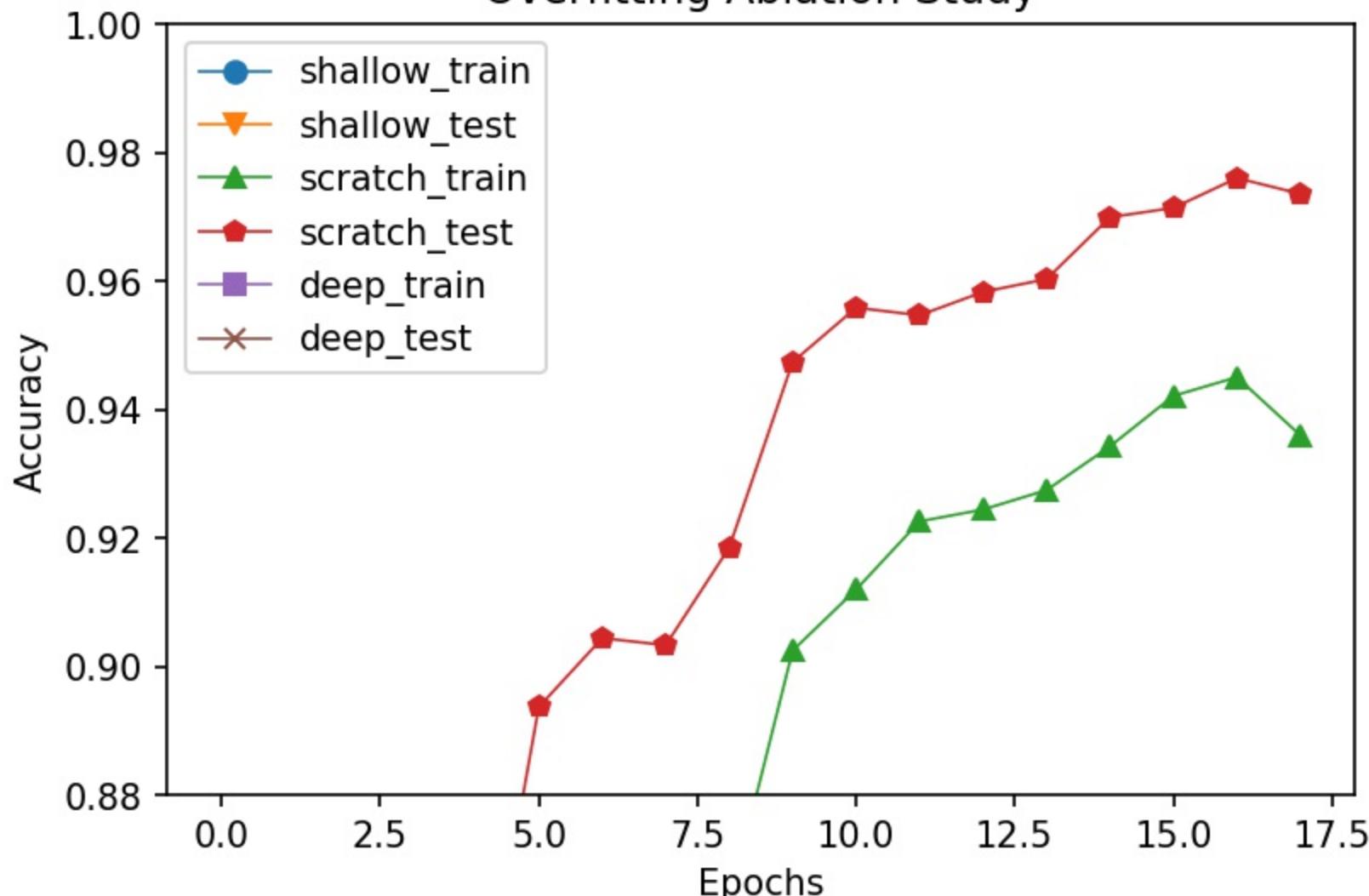
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 17 takes 318 seconds.

[19, 30] loss: 0.251

[19, 60] loss: 0.210

```
[19,    90] loss: 0.210
[19,   120] loss: 0.201
[19,   150] loss: 0.207
[19,   180] loss: 0.177
[19,   210] loss: 0.269
[19,   240] loss: 0.257
[19,   270] loss: 0.203
[19,   300] loss: 0.249
[19,   330] loss: 0.216
[19,   360] loss: 0.230
[19,   390] loss: 0.190
[19,   420] loss: 0.211
[19,   450] loss: 0.207
[19,   480] loss: 0.167
[19,   510] loss: 0.257
[19,   540] loss: 0.228
[19,   570] loss: 0.240
[19,   600] loss: 0.213
[19,   630] loss: 0.215
[19,   660] loss: 0.229
[19,   690] loss: 0.257
[19,   720] loss: 0.201
[19,   750] loss: 0.212
[19,   780] loss: 0.288
[19,   810] loss: 0.195
[19,   840] loss: 0.194
[19,   870] loss: 0.167
[19,   900] loss: 0.182
[19,   930] loss: 0.256
[19,   960] loss: 0.214
[19,   990] loss: 0.175
[19,  1020] loss: 0.218
[19,  1050] loss: 0.200
[19,  1080] loss: 0.215
[19,  1110] loss: 0.144
[19,  1140] loss: 0.243
[19,  1170] loss: 0.239
[19,  1200] loss: 0.233
[19,  1230] loss: 0.203
```

```
[19, 1260] loss: 0.270
[19, 1290] loss: 0.210
[19, 1320] loss: 0.175
[19, 1350] loss: 0.263
[19, 1380] loss: 0.229
[19, 1410] loss: 0.225
[19, 1440] loss: 0.269
[19, 1470] loss: 0.172
[19, 1500] loss: 0.187
[19, 1530] loss: 0.191
[19, 1560] loss: 0.275
[19, 1590] loss: 0.229
[19, 1620] loss: 0.228
[19, 1650] loss: 0.219
[19, 1680] loss: 0.207
[19, 1710] loss: 0.254
[19, 1740] loss: 0.242
[19, 1770] loss: 0.230
[19, 1800] loss: 0.235
[19, 1830] loss: 0.260
[19, 1860] loss: 0.223
[19, 1890] loss: 0.222
[19, 1920] loss: 0.189
[19, 1950] loss: 0.266
[19, 1980] loss: 0.224
[19, 2010] loss: 0.217
[19, 2040] loss: 0.242
[19, 2070] loss: 0.214
[19, 2100] loss: 0.197
[19, 2130] loss: 0.158
[19, 2160] loss: 0.160
[19, 2190] loss: 0.239
eval intermediate_models/epoch18_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.950109
epoch 18 accuracy on train set is: 0.9501090512540894
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.975282

epoch 18 accuracy on test set is: 0.975281715739731

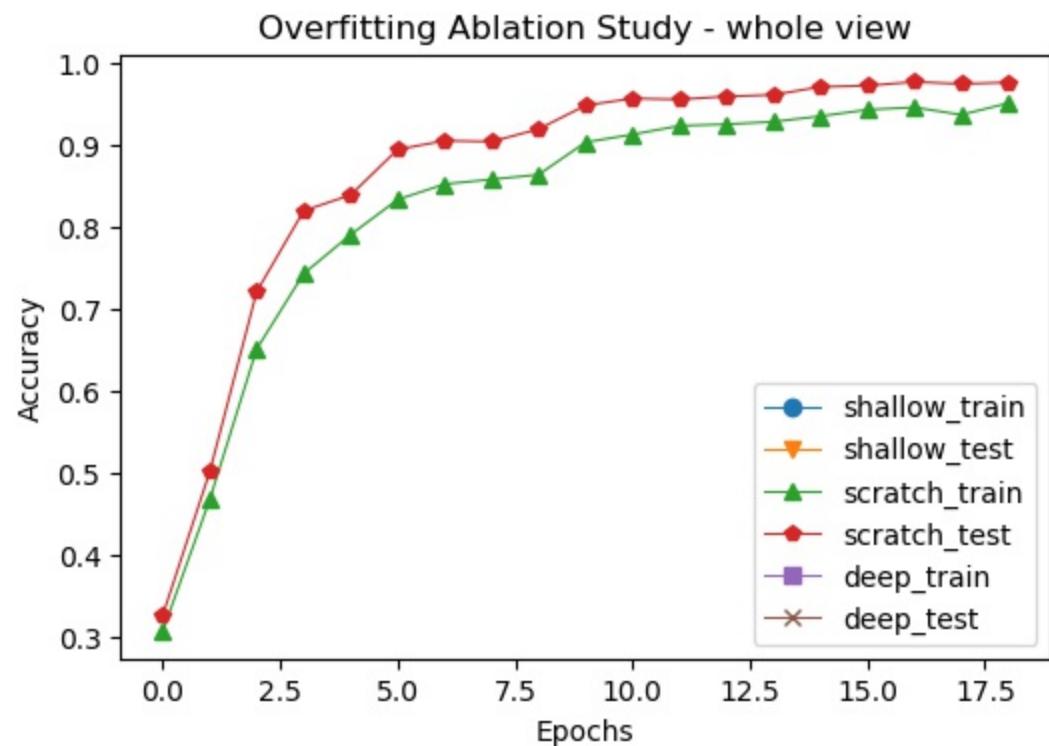
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

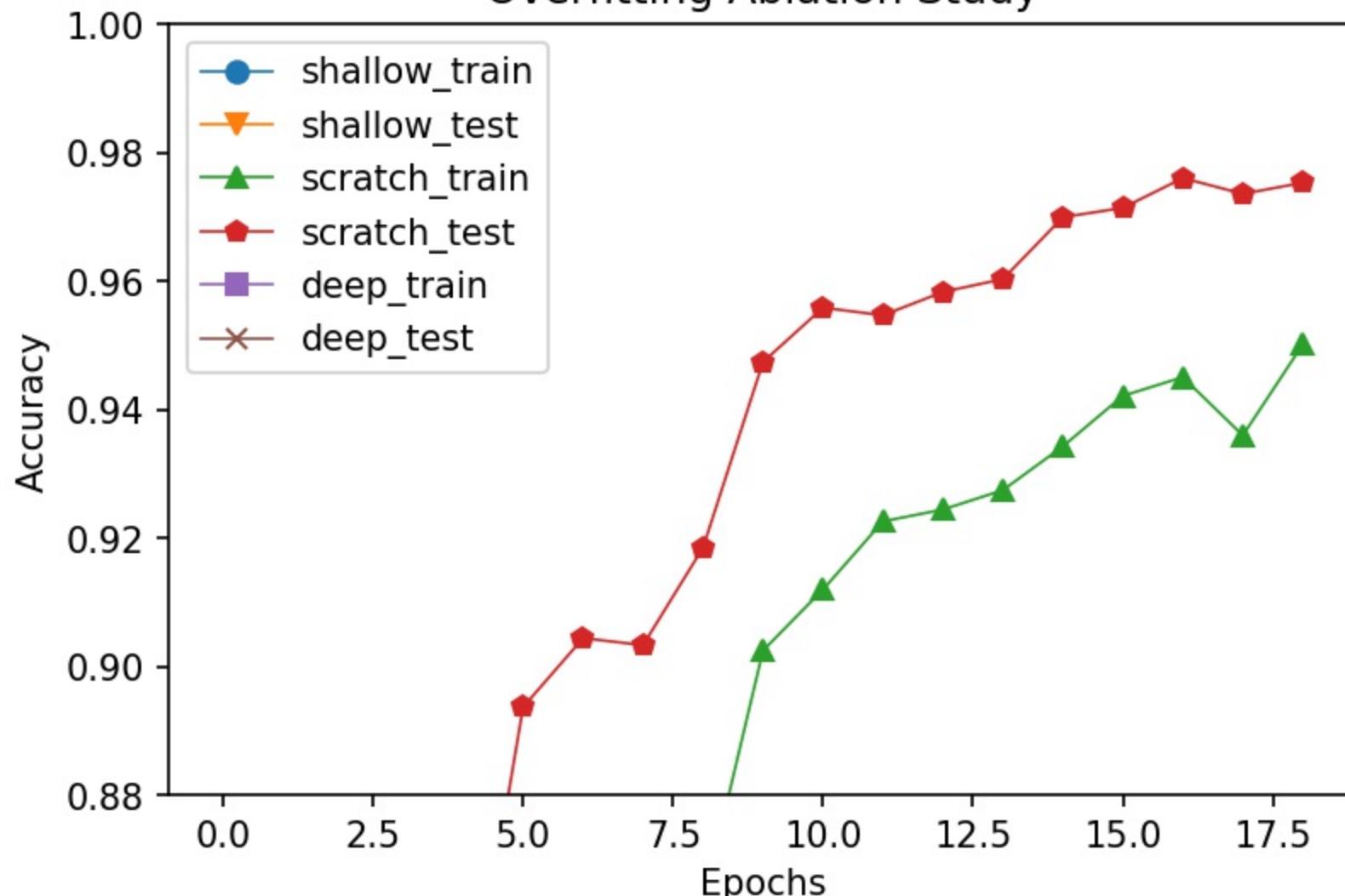
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 18 takes 317 seconds.

[20, 30] loss: 0.161

[20, 60] loss: 0.217

```
[20,    90] loss: 0.195
[20,   120] loss: 0.194
[20,   150] loss: 0.247
[20,   180] loss: 0.308
[20,   210] loss: 0.214
[20,   240] loss: 0.234
[20,   270] loss: 0.200
[20,   300] loss: 0.241
[20,   330] loss: 0.217
[20,   360] loss: 0.209
[20,   390] loss: 0.255
[20,   420] loss: 0.286
[20,   450] loss: 0.203
[20,   480] loss: 0.233
[20,   510] loss: 0.184
[20,   540] loss: 0.206
[20,   570] loss: 0.274
[20,   600] loss: 0.201
[20,   630] loss: 0.272
[20,   660] loss: 0.196
[20,   690] loss: 0.182
[20,   720] loss: 0.183
[20,   750] loss: 0.242
[20,   780] loss: 0.262
[20,   810] loss: 0.236
[20,   840] loss: 0.185
[20,   870] loss: 0.177
[20,   900] loss: 0.204
[20,   930] loss: 0.224
[20,   960] loss: 0.153
[20,   990] loss: 0.227
[20,  1020] loss: 0.241
[20,  1050] loss: 0.212
[20,  1080] loss: 0.181
[20,  1110] loss: 0.229
[20,  1140] loss: 0.180
[20,  1170] loss: 0.160
[20,  1200] loss: 0.175
[20,  1230] loss: 0.211
```

```
[20, 1260] loss: 0.180
[20, 1290] loss: 0.163
[20, 1320] loss: 0.203
[20, 1350] loss: 0.259
[20, 1380] loss: 0.245
[20, 1410] loss: 0.218
[20, 1440] loss: 0.229
[20, 1470] loss: 0.214
[20, 1500] loss: 0.213
[20, 1530] loss: 0.197
[20, 1560] loss: 0.198
[20, 1590] loss: 0.200
[20, 1620] loss: 0.212
[20, 1650] loss: 0.178
[20, 1680] loss: 0.236
[20, 1710] loss: 0.244
[20, 1740] loss: 0.226
[20, 1770] loss: 0.216
[20, 1800] loss: 0.234
[20, 1830] loss: 0.216
[20, 1860] loss: 0.199
[20, 1890] loss: 0.235
[20, 1920] loss: 0.185
[20, 1950] loss: 0.179
[20, 1980] loss: 0.252
[20, 2010] loss: 0.241
[20, 2040] loss: 0.262
[20, 2070] loss: 0.207
[20, 2100] loss: 0.237
[20, 2130] loss: 0.200
[20, 2160] loss: 0.214
[20, 2190] loss: 0.188
evf intermediate_models/epoch19_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.938023
epoch 19 accuracy on train set is: 0.9380225372591785
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.973282

epoch 19 accuracy on test set is: 0.9732824427480916

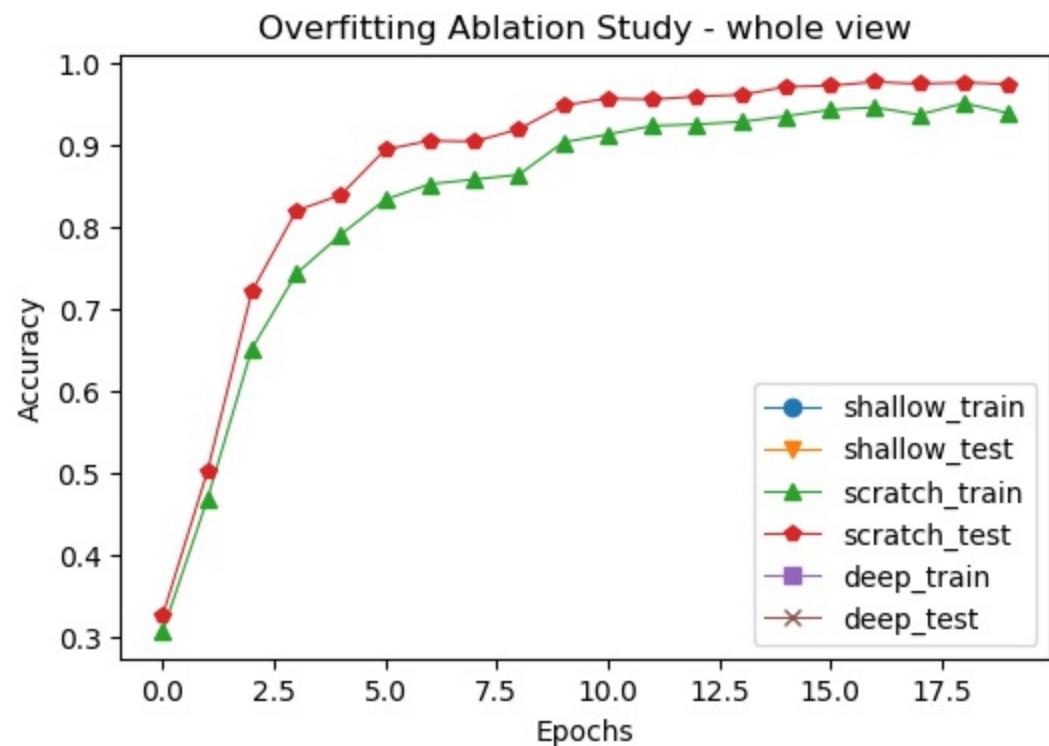
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

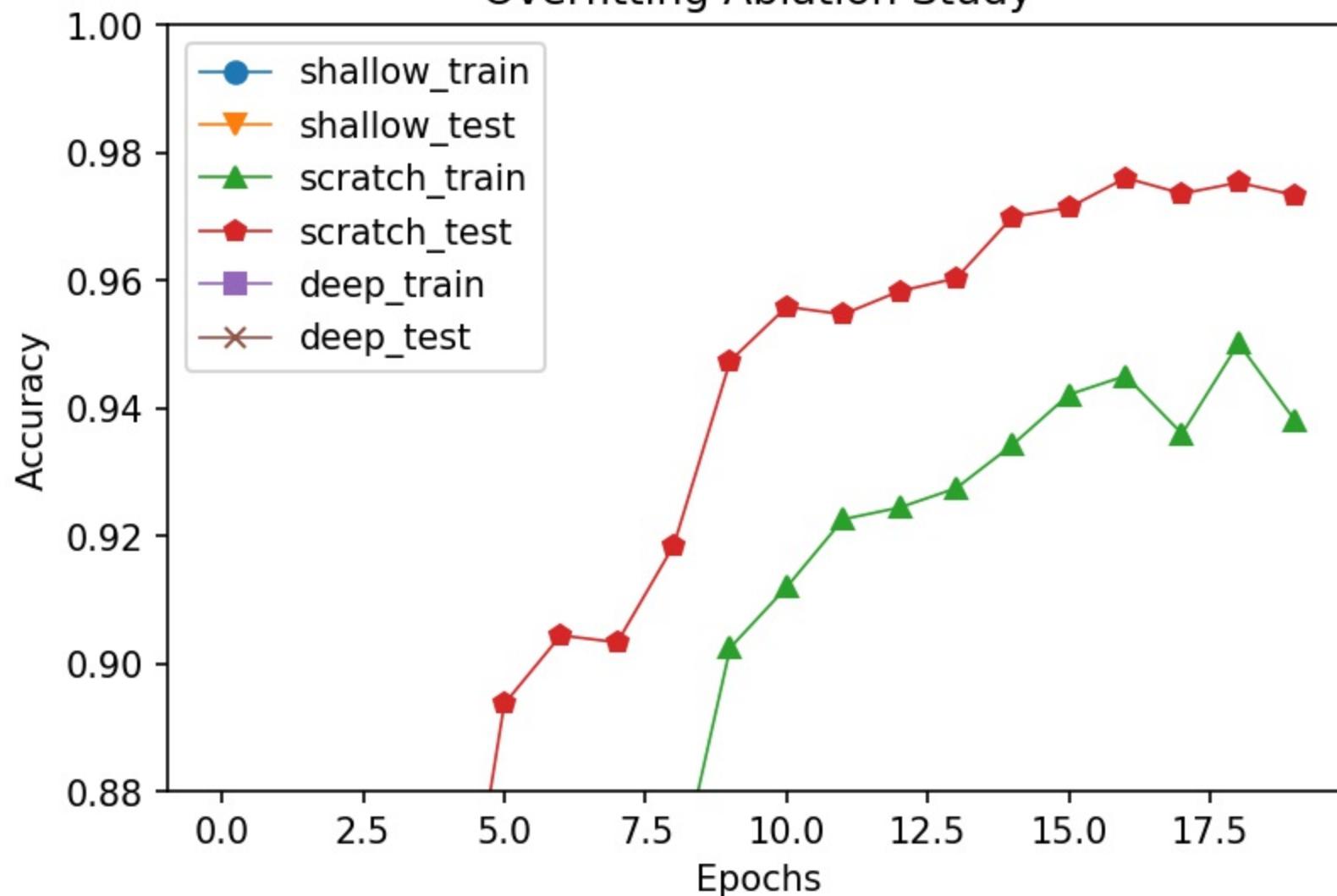
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 19 takes 318 seconds.

[21, 30] loss: 0.218

[21, 60] loss: 0.153

```
[21,    90] loss: 0.183
[21,   120] loss: 0.227
[21,   150] loss: 0.228
[21,   180] loss: 0.207
[21,   210] loss: 0.232
[21,   240] loss: 0.156
[21,   270] loss: 0.196
[21,   300] loss: 0.174
[21,   330] loss: 0.165
[21,   360] loss: 0.140
[21,   390] loss: 0.191
[21,   420] loss: 0.213
[21,   450] loss: 0.230
[21,   480] loss: 0.199
[21,   510] loss: 0.201
[21,   540] loss: 0.218
[21,   570] loss: 0.227
[21,   600] loss: 0.196
[21,   630] loss: 0.163
[21,   660] loss: 0.180
[21,   690] loss: 0.228
[21,   720] loss: 0.209
[21,   750] loss: 0.221
[21,   780] loss: 0.148
[21,   810] loss: 0.182
[21,   840] loss: 0.229
[21,   870] loss: 0.164
[21,   900] loss: 0.190
[21,   930] loss: 0.231
[21,   960] loss: 0.165
[21,   990] loss: 0.179
[21,  1020] loss: 0.246
[21,  1050] loss: 0.186
[21,  1080] loss: 0.188
[21,  1110] loss: 0.220
[21,  1140] loss: 0.228
[21,  1170] loss: 0.211
[21,  1200] loss: 0.260
[21,  1230] loss: 0.171
```

```
[21, 1260] loss: 0.260
[21, 1290] loss: 0.159
[21, 1320] loss: 0.214
[21, 1350] loss: 0.180
[21, 1380] loss: 0.138
[21, 1410] loss: 0.170
[21, 1440] loss: 0.220
[21, 1470] loss: 0.226
[21, 1500] loss: 0.203
[21, 1530] loss: 0.253
[21, 1560] loss: 0.210
[21, 1590] loss: 0.279
[21, 1620] loss: 0.197
[21, 1650] loss: 0.244
[21, 1680] loss: 0.172
[21, 1710] loss: 0.155
[21, 1740] loss: 0.187
[21, 1770] loss: 0.219
[21, 1800] loss: 0.197
[21, 1830] loss: 0.208
[21, 1860] loss: 0.177
[21, 1890] loss: 0.269
[21, 1920] loss: 0.174
[21, 1950] loss: 0.212
[21, 1980] loss: 0.159
[21, 2010] loss: 0.207
[21, 2040] loss: 0.147
[21, 2070] loss: 0.168
[21, 2100] loss: 0.193
[21, 2130] loss: 0.167
[21, 2160] loss: 0.232
[21, 2190] loss: 0.214
eval intermediate_models/epoch20_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.952245
epoch 20 accuracy on train set is: 0.9522446383133406
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.978917

epoch 20 accuracy on test set is: 0.9789167575427118

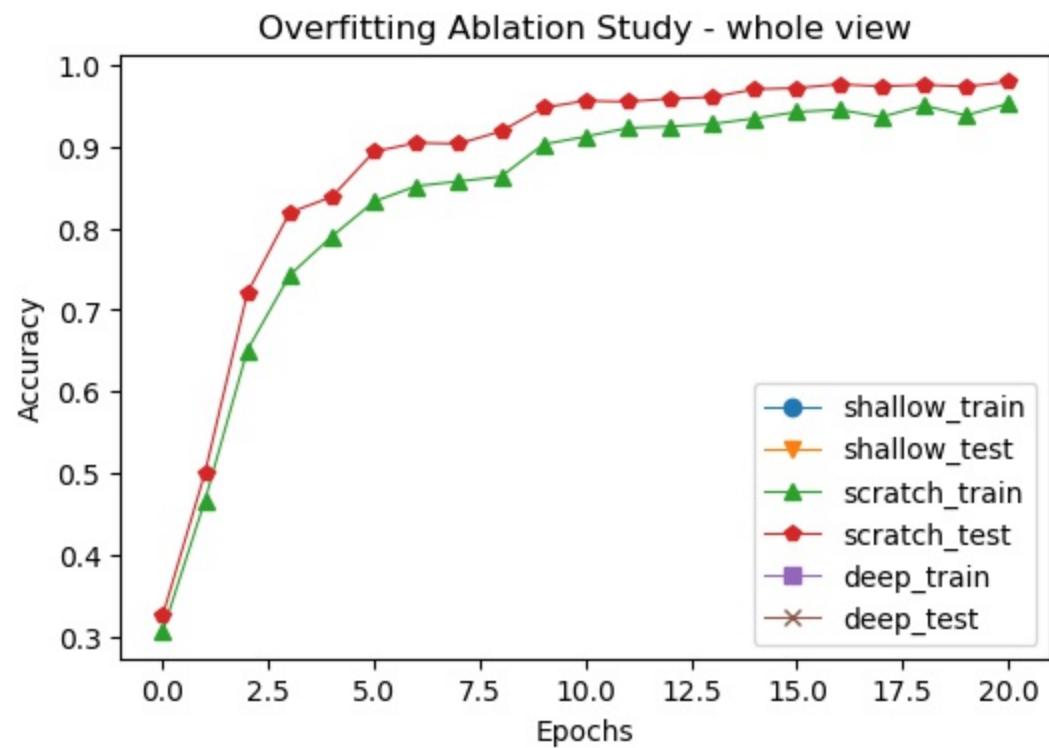
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

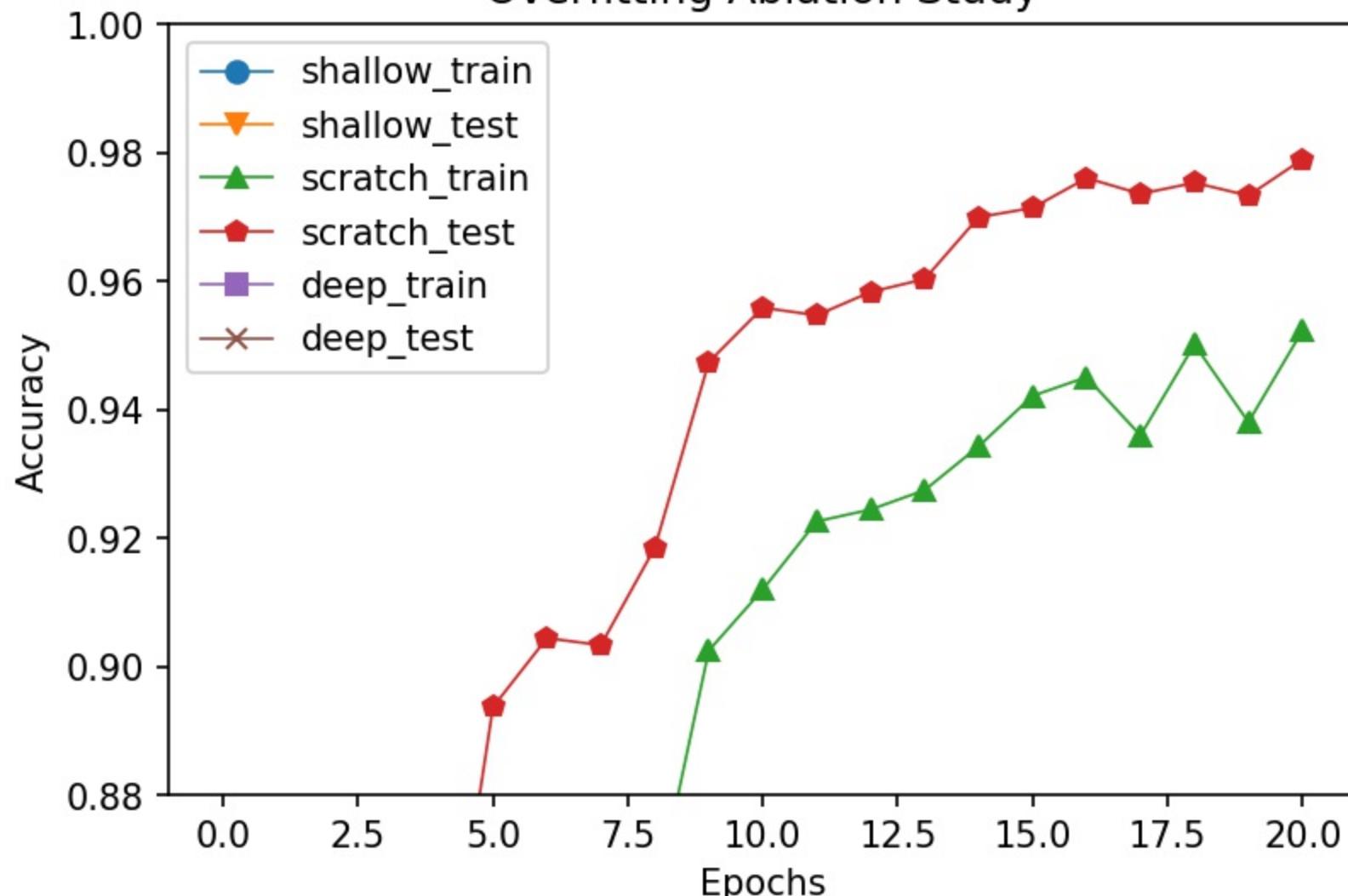
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 20 takes 318 seconds.

[22, 30] loss: 0.223

[22, 60] loss: 0.190

```
[22,    90] loss: 0.160
[22,   120] loss: 0.214
[22,   150] loss: 0.169
[22,   180] loss: 0.214
[22,   210] loss: 0.187
[22,   240] loss: 0.204
[22,   270] loss: 0.168
[22,   300] loss: 0.211
[22,   330] loss: 0.187
[22,   360] loss: 0.223
[22,   390] loss: 0.187
[22,   420] loss: 0.203
[22,   450] loss: 0.256
[22,   480] loss: 0.193
[22,   510] loss: 0.224
[22,   540] loss: 0.212
[22,   570] loss: 0.169
[22,   600] loss: 0.192
[22,   630] loss: 0.228
[22,   660] loss: 0.219
[22,   690] loss: 0.236
[22,   720] loss: 0.206
[22,   750] loss: 0.144
[22,   780] loss: 0.195
[22,   810] loss: 0.206
[22,   840] loss: 0.190
[22,   870] loss: 0.151
[22,   900] loss: 0.169
[22,   930] loss: 0.161
[22,   960] loss: 0.177
[22,   990] loss: 0.209
[22,  1020] loss: 0.193
[22,  1050] loss: 0.150
[22,  1080] loss: 0.230
[22,  1110] loss: 0.238
[22,  1140] loss: 0.228
[22,  1170] loss: 0.176
[22,  1200] loss: 0.204
[22,  1230] loss: 0.230
```

```
[22, 1260] loss: 0.164
[22, 1290] loss: 0.149
[22, 1320] loss: 0.252
[22, 1350] loss: 0.188
[22, 1380] loss: 0.275
[22, 1410] loss: 0.194
[22, 1440] loss: 0.231
[22, 1470] loss: 0.197
[22, 1500] loss: 0.176
[22, 1530] loss: 0.169
[22, 1560] loss: 0.212
[22, 1590] loss: 0.189
[22, 1620] loss: 0.234
[22, 1650] loss: 0.239
[22, 1680] loss: 0.181
[22, 1710] loss: 0.166
[22, 1740] loss: 0.182
[22, 1770] loss: 0.186
[22, 1800] loss: 0.175
[22, 1830] loss: 0.140
[22, 1860] loss: 0.130
[22, 1890] loss: 0.295
[22, 1920] loss: 0.241
[22, 1950] loss: 0.196
[22, 1980] loss: 0.201
[22, 2010] loss: 0.212
[22, 2040] loss: 0.181
[22, 2070] loss: 0.175
[22, 2100] loss: 0.163
[22, 2130] loss: 0.203
[22, 2160] loss: 0.187
[22, 2190] loss: 0.189
evf intermediate_models/epoch21_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.953335
epoch 21 accuracy on train set is: 0.9533351508542348
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.980825

epoch 21 accuracy on test set is: 0.9808251544892767

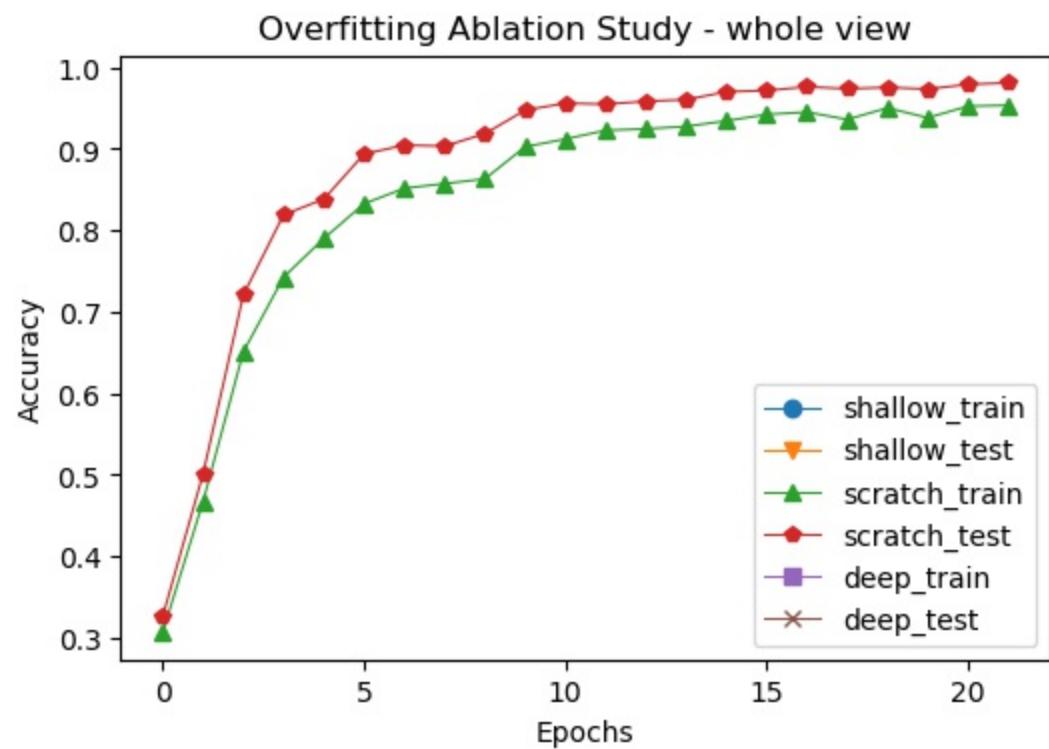
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

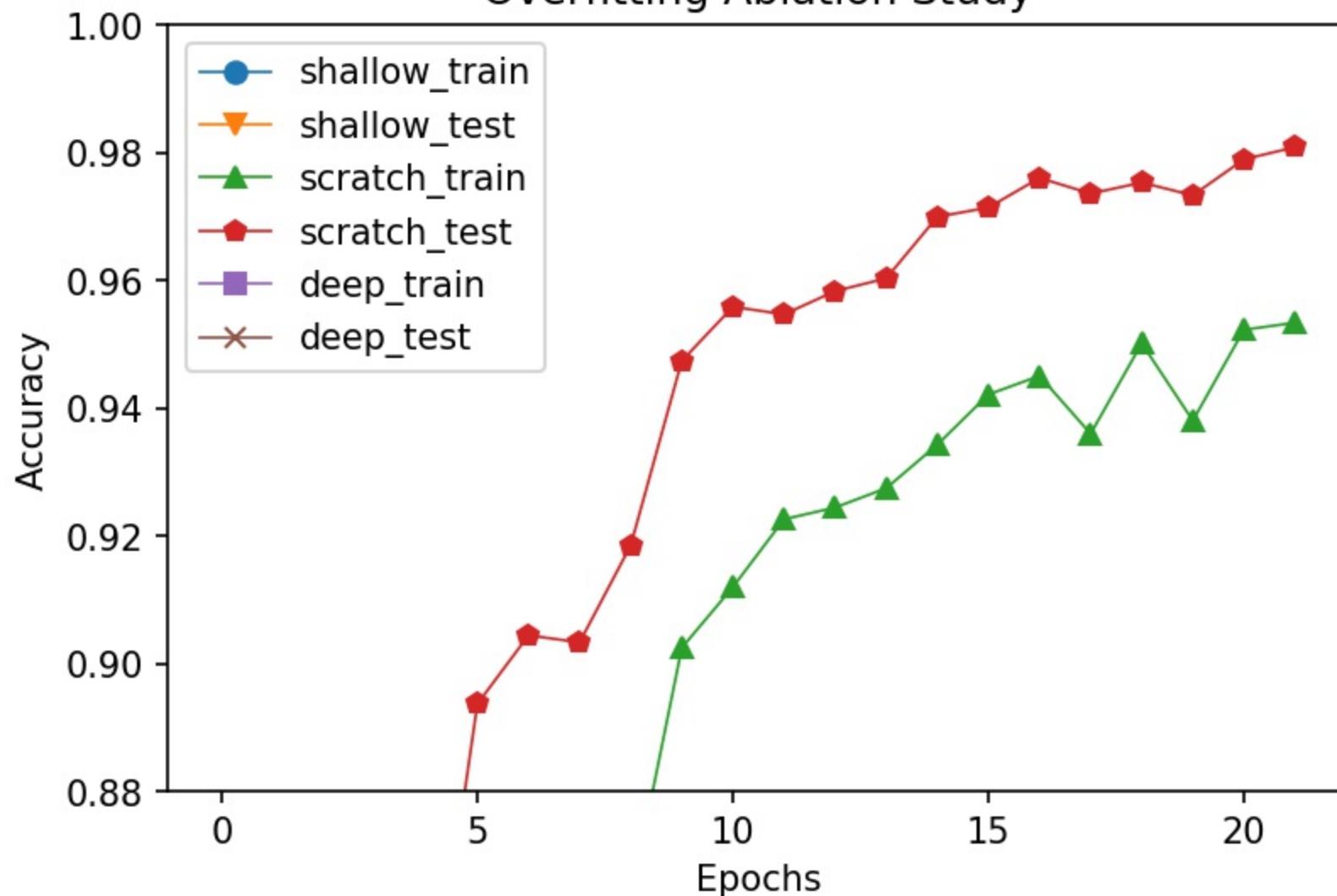
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 21 takes 317 seconds.

[23, 30] loss: 0.153

[23, 60] loss: 0.150

```
[23,    90] loss: 0.164
[23,   120] loss: 0.176
[23,   150] loss: 0.203
[23,   180] loss: 0.210
[23,   210] loss: 0.147
[23,   240] loss: 0.121
[23,   270] loss: 0.147
[23,   300] loss: 0.231
[23,   330] loss: 0.216
[23,   360] loss: 0.224
[23,   390] loss: 0.179
[23,   420] loss: 0.212
[23,   450] loss: 0.186
[23,   480] loss: 0.204
[23,   510] loss: 0.222
[23,   540] loss: 0.246
[23,   570] loss: 0.218
[23,   600] loss: 0.174
[23,   630] loss: 0.184
[23,   660] loss: 0.186
[23,   690] loss: 0.172
[23,   720] loss: 0.155
[23,   750] loss: 0.146
[23,   780] loss: 0.175
[23,   810] loss: 0.173
[23,   840] loss: 0.086
[23,   870] loss: 0.164
[23,   900] loss: 0.178
[23,   930] loss: 0.241
[23,   960] loss: 0.218
[23,   990] loss: 0.186
[23,  1020] loss: 0.185
[23,  1050] loss: 0.171
[23,  1080] loss: 0.204
[23,  1110] loss: 0.198
[23,  1140] loss: 0.151
[23,  1170] loss: 0.183
[23,  1200] loss: 0.153
[23,  1230] loss: 0.242
```

```
[23, 1260] loss: 0.219
[23, 1290] loss: 0.192
[23, 1320] loss: 0.186
[23, 1350] loss: 0.144
[23, 1380] loss: 0.152
[23, 1410] loss: 0.216
[23, 1440] loss: 0.188
[23, 1470] loss: 0.202
[23, 1500] loss: 0.140
[23, 1530] loss: 0.129
[23, 1560] loss: 0.195
[23, 1590] loss: 0.153
[23, 1620] loss: 0.185
[23, 1650] loss: 0.158
[23, 1680] loss: 0.226
[23, 1710] loss: 0.182
[23, 1740] loss: 0.151
[23, 1770] loss: 0.185
[23, 1800] loss: 0.206
[23, 1830] loss: 0.189
[23, 1860] loss: 0.204
[23, 1890] loss: 0.172
[23, 1920] loss: 0.138
[23, 1950] loss: 0.173
[23, 1980] loss: 0.169
[23, 2010] loss: 0.190
[23, 2040] loss: 0.234
[23, 2070] loss: 0.224
[23, 2100] loss: 0.181
[23, 2130] loss: 0.160
[23, 2160] loss: 0.167
[23, 2190] loss: 0.177
eval intermediate_models/epoch22_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.958152
epoch 22 accuracy on train set is: 0.9581515812431843
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.981461

epoch 22 accuracy on test set is: 0.9814612868047983

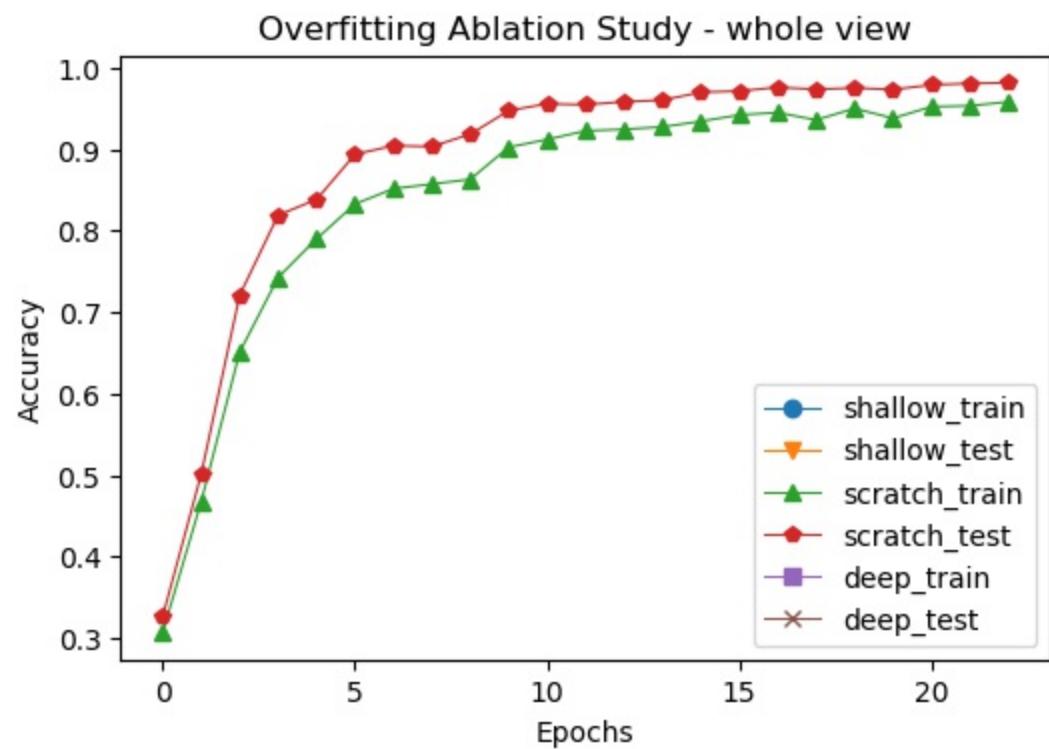
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

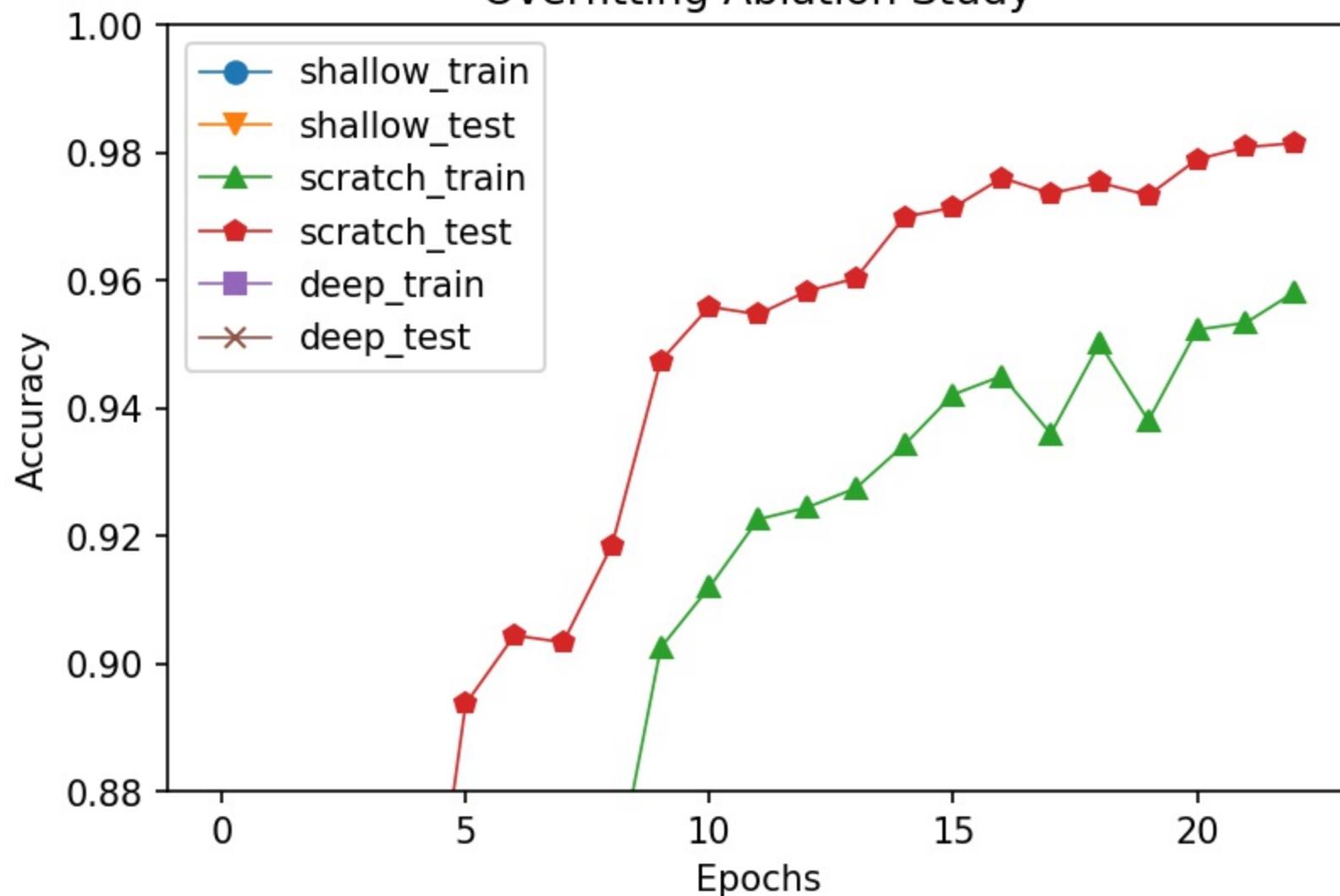
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 22 takes 318 seconds.

[24, 30] loss: 0.165

[24, 60] loss: 0.155

```
[24,    90] loss: 0.175
[24,   120] loss: 0.203
[24,   150] loss: 0.184
[24,   180] loss: 0.137
[24,   210] loss: 0.221
[24,   240] loss: 0.214
[24,   270] loss: 0.197
[24,   300] loss: 0.237
[24,   330] loss: 0.207
[24,   360] loss: 0.193
[24,   390] loss: 0.150
[24,   420] loss: 0.162
[24,   450] loss: 0.130
[24,   480] loss: 0.154
[24,   510] loss: 0.123
[24,   540] loss: 0.182
[24,   570] loss: 0.171
[24,   600] loss: 0.201
[24,   630] loss: 0.188
[24,   660] loss: 0.197
[24,   690] loss: 0.120
[24,   720] loss: 0.185
[24,   750] loss: 0.213
[24,   780] loss: 0.197
[24,   810] loss: 0.178
[24,   840] loss: 0.155
[24,   870] loss: 0.223
[24,   900] loss: 0.184
[24,   930] loss: 0.185
[24,   960] loss: 0.238
[24,   990] loss: 0.186
[24,  1020] loss: 0.174
[24,  1050] loss: 0.192
[24,  1080] loss: 0.173
[24,  1110] loss: 0.185
[24,  1140] loss: 0.139
[24,  1170] loss: 0.187
[24,  1200] loss: 0.224
[24,  1230] loss: 0.167
```

```
[24, 1260] loss: 0.174
[24, 1290] loss: 0.197
[24, 1320] loss: 0.169
[24, 1350] loss: 0.197
[24, 1380] loss: 0.186
[24, 1410] loss: 0.152
[24, 1440] loss: 0.186
[24, 1470] loss: 0.171
[24, 1500] loss: 0.135
[24, 1530] loss: 0.186
[24, 1560] loss: 0.162
[24, 1590] loss: 0.133
[24, 1620] loss: 0.144
[24, 1650] loss: 0.164
[24, 1680] loss: 0.198
[24, 1710] loss: 0.204
[24, 1740] loss: 0.155
[24, 1770] loss: 0.182
[24, 1800] loss: 0.210
[24, 1830] loss: 0.170
[24, 1860] loss: 0.197
[24, 1890] loss: 0.160
[24, 1920] loss: 0.171
[24, 1950] loss: 0.142
[24, 1980] loss: 0.188
[24, 2010] loss: 0.182
[24, 2040] loss: 0.198
[24, 2070] loss: 0.210
[24, 2100] loss: 0.215
[24, 2130] loss: 0.189
[24, 2160] loss: 0.151
[24, 2190] loss: 0.219
eval intermediate_models/epoch23_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.950314
epoch 23 accuracy on train set is: 0.950313522355507
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.976917

epoch 23 accuracy on test set is: 0.9769174845510723

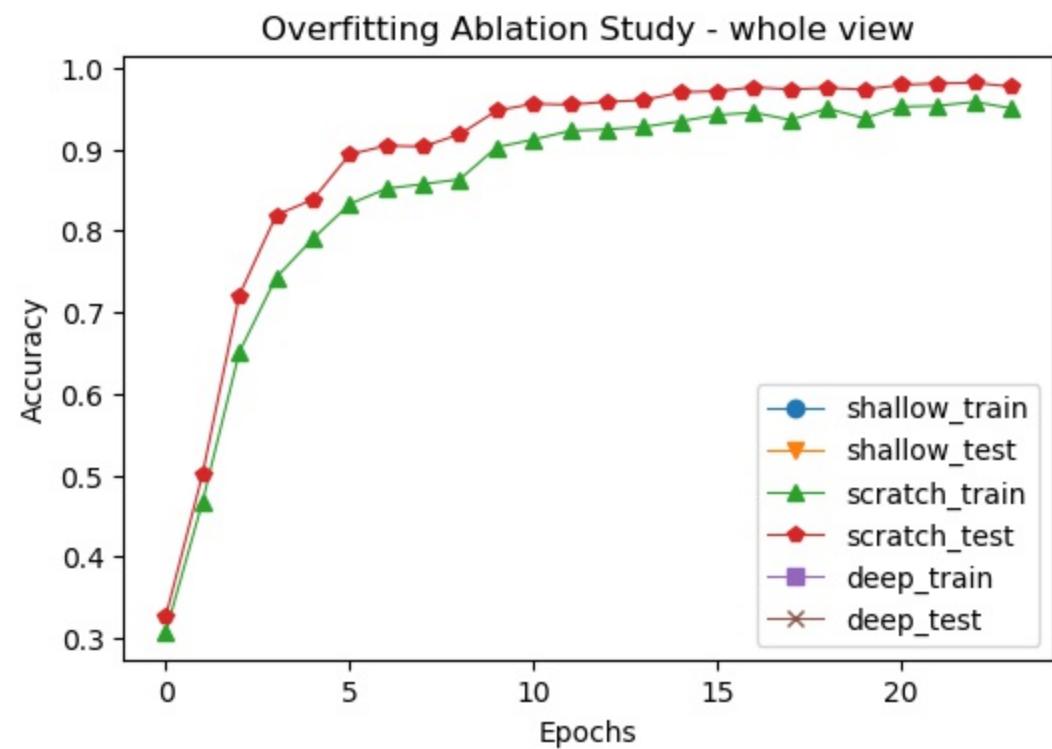
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

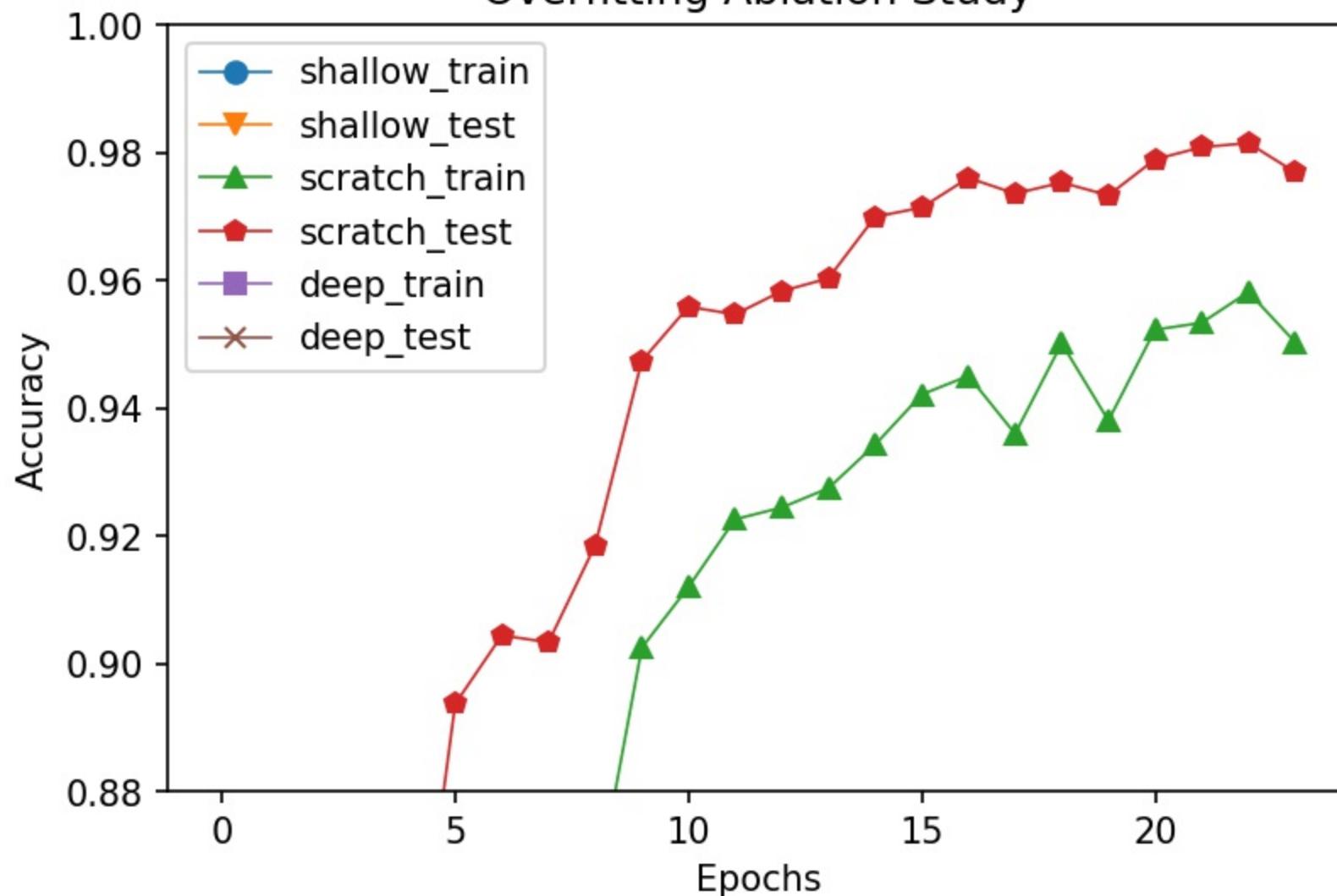
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 23 takes 318 seconds.

[25, 30] loss: 0.172

[25, 60] loss: 0.174

```
[25,    90] loss: 0.195
[25,   120] loss: 0.212
[25,   150] loss: 0.198
[25,   180] loss: 0.177
[25,   210] loss: 0.166
[25,   240] loss: 0.204
[25,   270] loss: 0.211
[25,   300] loss: 0.153
[25,   330] loss: 0.161
[25,   360] loss: 0.157
[25,   390] loss: 0.189
[25,   420] loss: 0.196
[25,   450] loss: 0.181
[25,   480] loss: 0.201
[25,   510] loss: 0.142
[25,   540] loss: 0.170
[25,   570] loss: 0.236
[25,   600] loss: 0.219
[25,   630] loss: 0.169
[25,   660] loss: 0.231
[25,   690] loss: 0.177
[25,   720] loss: 0.171
[25,   750] loss: 0.185
[25,   780] loss: 0.167
[25,   810] loss: 0.192
[25,   840] loss: 0.168
[25,   870] loss: 0.153
[25,   900] loss: 0.193
[25,   930] loss: 0.146
[25,   960] loss: 0.181
[25,   990] loss: 0.132
[25,  1020] loss: 0.181
[25,  1050] loss: 0.145
[25,  1080] loss: 0.155
[25,  1110] loss: 0.195
[25,  1140] loss: 0.173
[25,  1170] loss: 0.168
[25,  1200] loss: 0.189
[25,  1230] loss: 0.177
```

```
[25, 1260] loss: 0.241
[25, 1290] loss: 0.164
[25, 1320] loss: 0.159
[25, 1350] loss: 0.152
[25, 1380] loss: 0.165
[25, 1410] loss: 0.220
[25, 1440] loss: 0.166
[25, 1470] loss: 0.188
[25, 1500] loss: 0.137
[25, 1530] loss: 0.163
[25, 1560] loss: 0.136
[25, 1590] loss: 0.222
[25, 1620] loss: 0.179
[25, 1650] loss: 0.175
[25, 1680] loss: 0.176
[25, 1710] loss: 0.216
[25, 1740] loss: 0.174
[25, 1770] loss: 0.251
[25, 1800] loss: 0.140
[25, 1830] loss: 0.122
[25, 1860] loss: 0.145
[25, 1890] loss: 0.178
[25, 1920] loss: 0.182
[25, 1950] loss: 0.141
[25, 1980] loss: 0.142
[25, 2010] loss: 0.144
[25, 2040] loss: 0.140
[25, 2070] loss: 0.186
[25, 2100] loss: 0.148
[25, 2130] loss: 0.195
[25, 2160] loss: 0.195
[25, 2190] loss: 0.108
eval intermediate_models/epoch24_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.959378
epoch 24 accuracy on train set is: 0.9593784078516903
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.978099

epoch 24 accuracy on test set is: 0.978098873137041

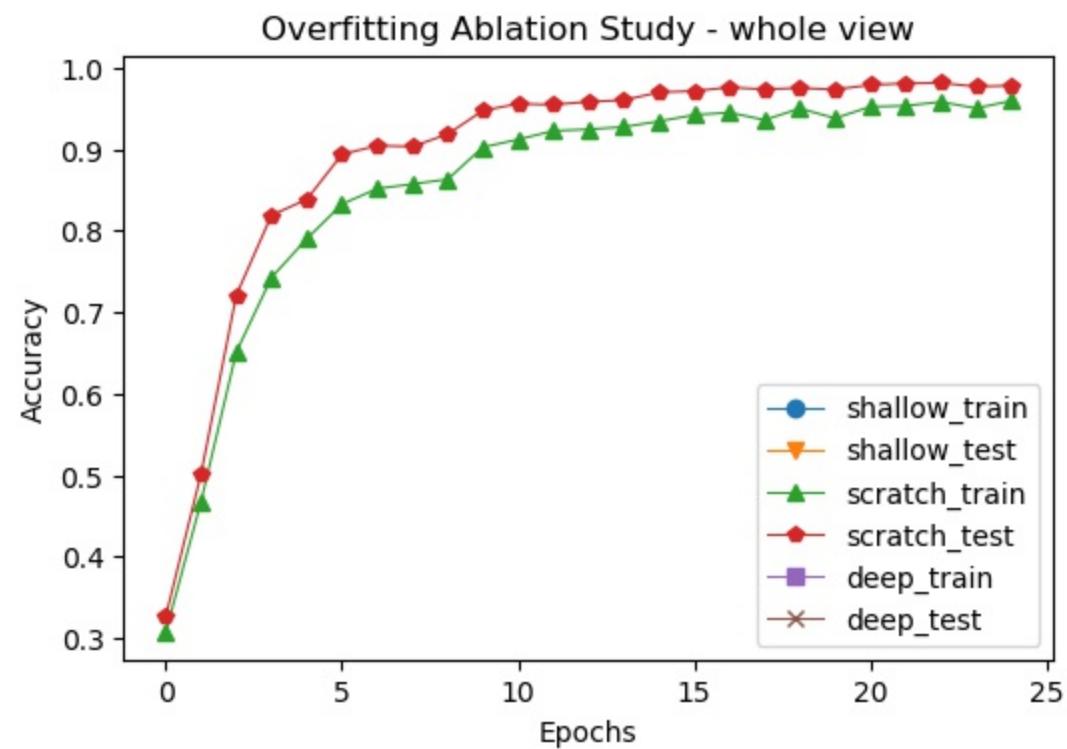
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

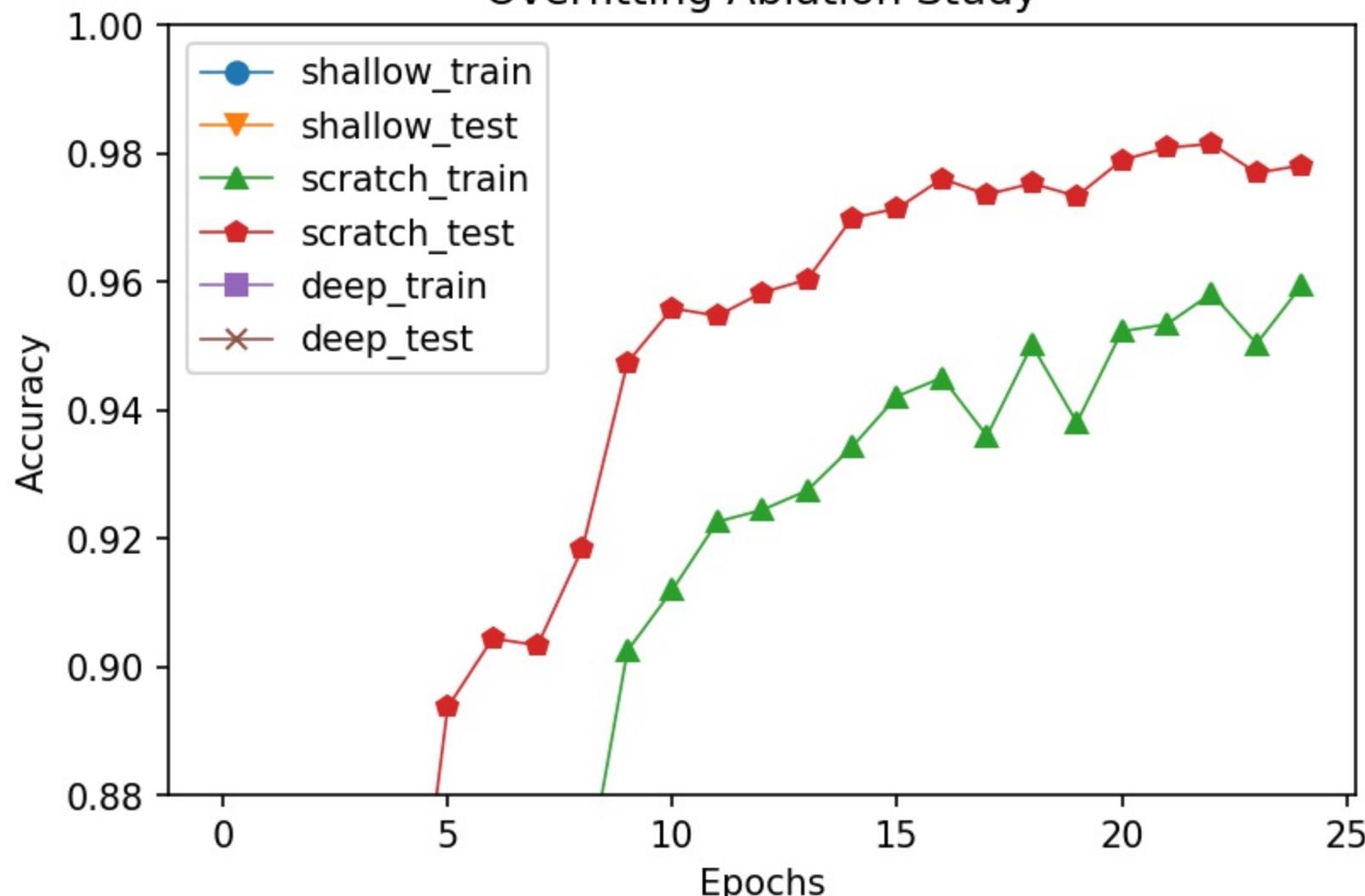
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 24 takes 317 seconds.

[26, 30] loss: 0.139

[26, 60] loss: 0.159

```
[26,    90] loss: 0.229
[26,   120] loss: 0.190
[26,   150] loss: 0.160
[26,   180] loss: 0.198
[26,   210] loss: 0.185
[26,   240] loss: 0.158
[26,   270] loss: 0.175
[26,   300] loss: 0.140
[26,   330] loss: 0.138
[26,   360] loss: 0.108
[26,   390] loss: 0.183
[26,   420] loss: 0.170
[26,   450] loss: 0.175
[26,   480] loss: 0.174
[26,   510] loss: 0.190
[26,   540] loss: 0.132
[26,   570] loss: 0.156
[26,   600] loss: 0.196
[26,   630] loss: 0.149
[26,   660] loss: 0.115
[26,   690] loss: 0.181
[26,   720] loss: 0.171
[26,   750] loss: 0.188
[26,   780] loss: 0.142
[26,   810] loss: 0.157
[26,   840] loss: 0.137
[26,   870] loss: 0.165
[26,   900] loss: 0.125
[26,   930] loss: 0.225
[26,   960] loss: 0.111
[26,   990] loss: 0.176
[26,  1020] loss: 0.116
[26,  1050] loss: 0.142
[26,  1080] loss: 0.178
[26,  1110] loss: 0.183
[26,  1140] loss: 0.197
[26,  1170] loss: 0.164
[26,  1200] loss: 0.190
[26,  1230] loss: 0.164
```

```
[26, 1260] loss: 0.148
[26, 1290] loss: 0.146
[26, 1320] loss: 0.201
[26, 1350] loss: 0.167
[26, 1380] loss: 0.139
[26, 1410] loss: 0.161
[26, 1440] loss: 0.193
[26, 1470] loss: 0.172
[26, 1500] loss: 0.218
[26, 1530] loss: 0.190
[26, 1560] loss: 0.173
[26, 1590] loss: 0.163
[26, 1620] loss: 0.193
[26, 1650] loss: 0.172
[26, 1680] loss: 0.159
[26, 1710] loss: 0.203
[26, 1740] loss: 0.127
[26, 1770] loss: 0.181
[26, 1800] loss: 0.173
[26, 1830] loss: 0.124
[26, 1860] loss: 0.150
[26, 1890] loss: 0.187
[26, 1920] loss: 0.214
[26, 1950] loss: 0.172
[26, 1980] loss: 0.182
[26, 2010] loss: 0.190
[26, 2040] loss: 0.118
[26, 2070] loss: 0.197
[26, 2100] loss: 0.174
[26, 2130] loss: 0.150
[26, 2160] loss: 0.174
[26, 2190] loss: 0.174
eval intermediate_models/epoch25_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.951086
epoch 25 accuracy on train set is: 0.9510859687386405
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.973737

epoch 25 accuracy on test set is: 0.9737368229734642

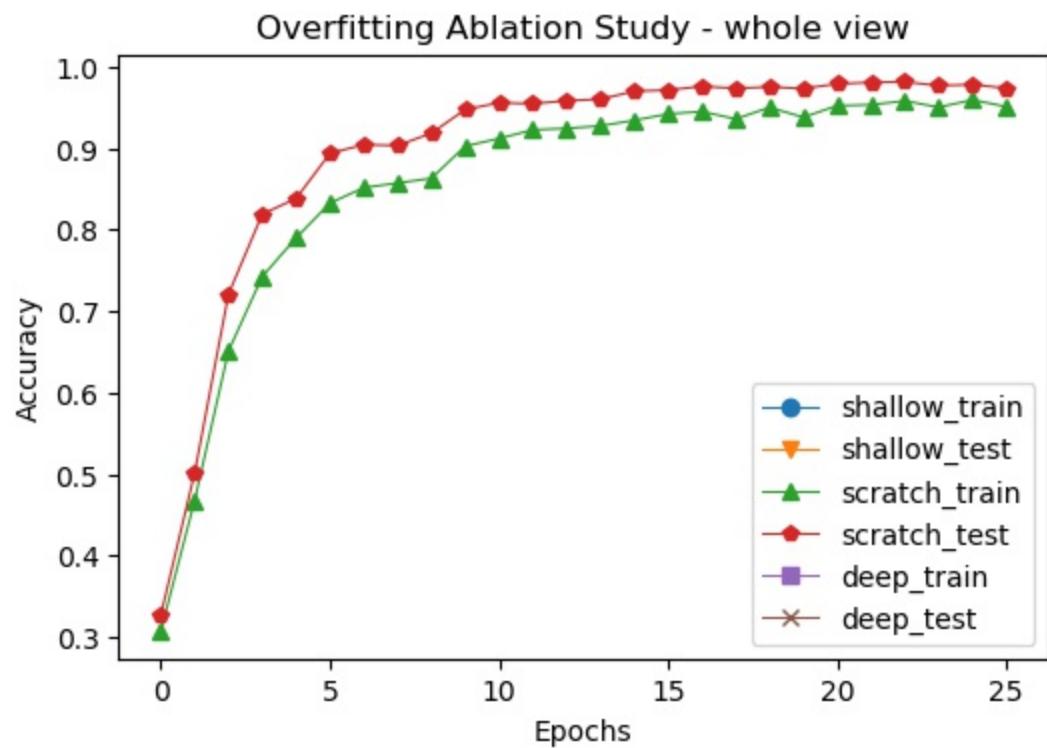
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

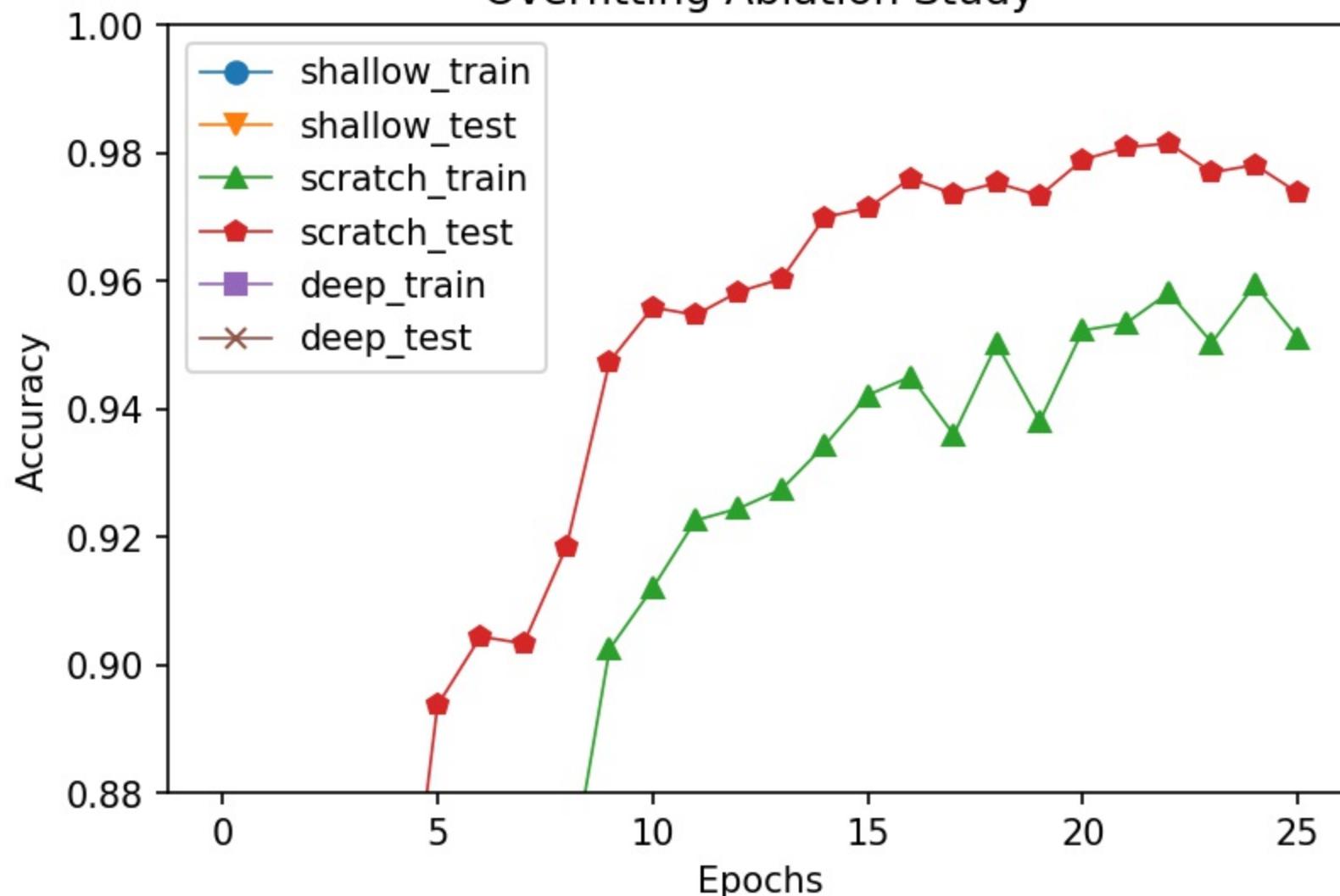
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 25 takes 317 seconds.

[27, 30] loss: 0.186

[27, 60] loss: 0.166

```
[27,    90] loss: 0.164
[27,   120] loss: 0.205
[27,   150] loss: 0.146
[27,   180] loss: 0.132
[27,   210] loss: 0.169
[27,   240] loss: 0.147
[27,   270] loss: 0.196
[27,   300] loss: 0.168
[27,   330] loss: 0.150
[27,   360] loss: 0.207
[27,   390] loss: 0.173
[27,   420] loss: 0.128
[27,   450] loss: 0.162
[27,   480] loss: 0.195
[27,   510] loss: 0.146
[27,   540] loss: 0.150
[27,   570] loss: 0.151
[27,   600] loss: 0.193
[27,   630] loss: 0.159
[27,   660] loss: 0.153
[27,   690] loss: 0.142
[27,   720] loss: 0.185
[27,   750] loss: 0.168
[27,   780] loss: 0.178
[27,   810] loss: 0.156
[27,   840] loss: 0.196
[27,   870] loss: 0.158
[27,   900] loss: 0.183
[27,   930] loss: 0.119
[27,   960] loss: 0.173
[27,   990] loss: 0.154
[27,  1020] loss: 0.180
[27,  1050] loss: 0.118
[27,  1080] loss: 0.131
[27,  1110] loss: 0.170
[27,  1140] loss: 0.176
[27,  1170] loss: 0.119
[27,  1200] loss: 0.134
[27,  1230] loss: 0.136
```

```
[27, 1260] loss: 0.155
[27, 1290] loss: 0.157
[27, 1320] loss: 0.140
[27, 1350] loss: 0.207
[27, 1380] loss: 0.203
[27, 1410] loss: 0.161
[27, 1440] loss: 0.131
[27, 1470] loss: 0.128
[27, 1500] loss: 0.175
[27, 1530] loss: 0.171
[27, 1560] loss: 0.185
[27, 1590] loss: 0.167
[27, 1620] loss: 0.143
[27, 1650] loss: 0.122
[27, 1680] loss: 0.171
[27, 1710] loss: 0.180
[27, 1740] loss: 0.177
[27, 1770] loss: 0.197
[27, 1800] loss: 0.139
[27, 1830] loss: 0.142
[27, 1860] loss: 0.230
[27, 1890] loss: 0.134
[27, 1920] loss: 0.124
[27, 1950] loss: 0.199
[27, 1980] loss: 0.155
[27, 2010] loss: 0.187
[27, 2040] loss: 0.139
[27, 2070] loss: 0.237
[27, 2100] loss: 0.170
[27, 2130] loss: 0.182
[27, 2160] loss: 0.151
[27, 2190] loss: 0.134
eval intermediate_models/epoch26_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.960742
epoch 26 accuracy on train set is: 0.960741548527808
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.981461

epoch 26 accuracy on test set is: 0.9814612868047983

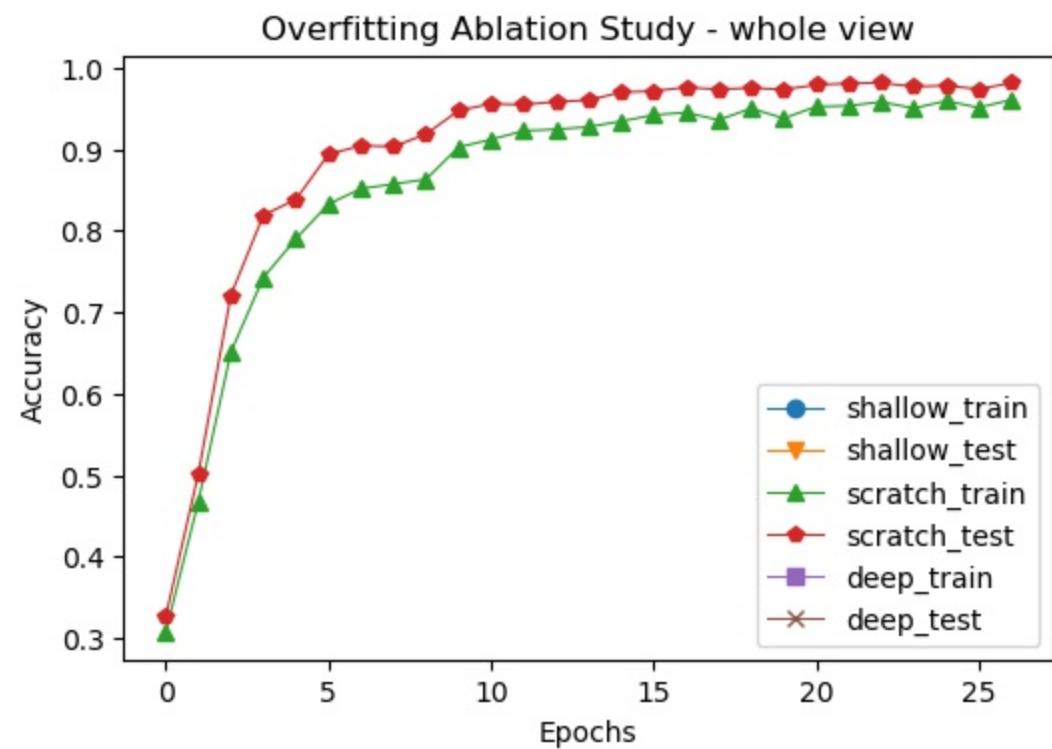
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

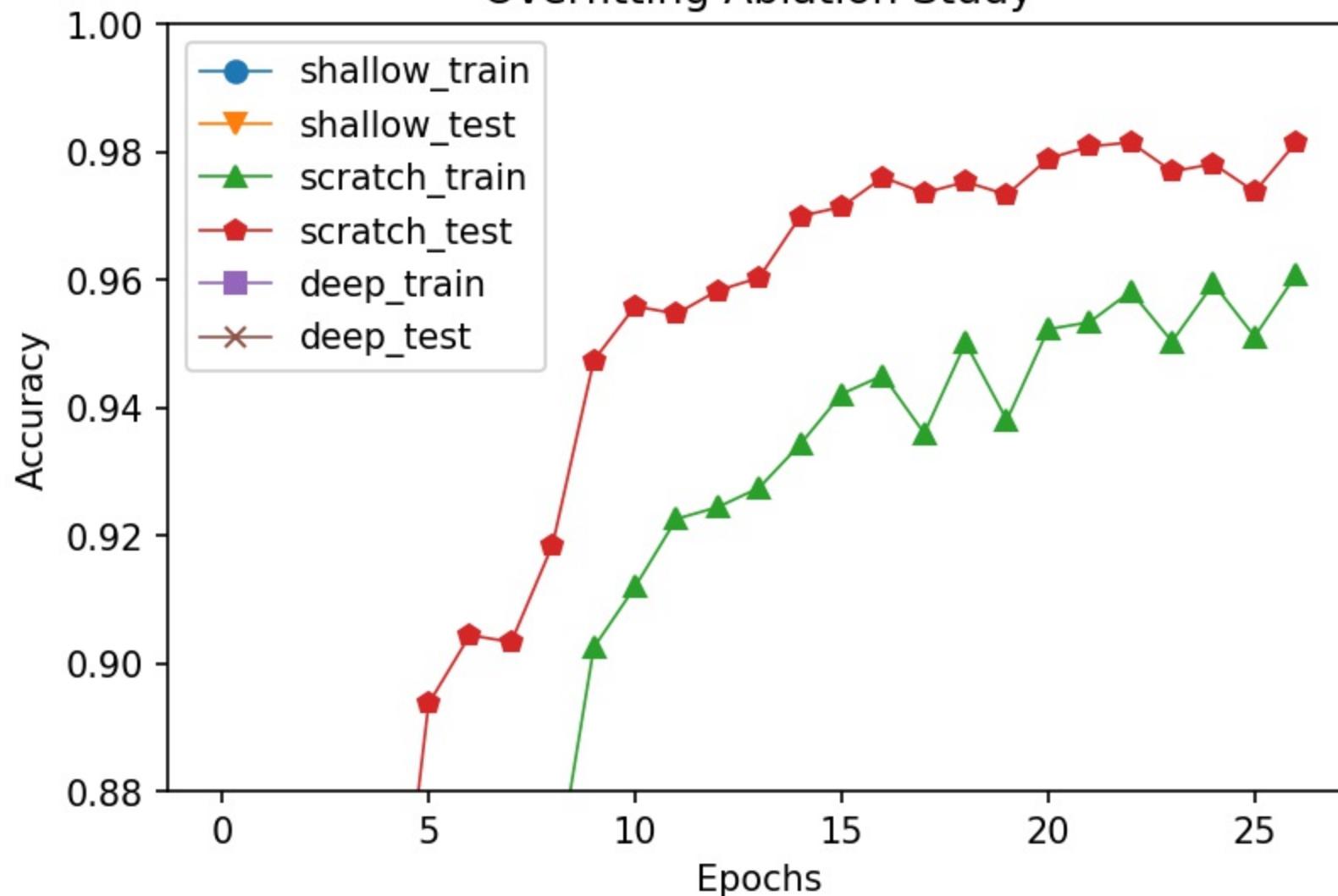
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



```
epoch 26 takes 318 seconds.  
[28,      30] loss: 0.127  
[28,      60] loss: 0.176
```

```
[28,    90] loss: 0.142
[28,   120] loss: 0.122
[28,   150] loss: 0.145
[28,   180] loss: 0.164
[28,   210] loss: 0.134
[28,   240] loss: 0.213
[28,   270] loss: 0.149
[28,   300] loss: 0.146
[28,   330] loss: 0.141
[28,   360] loss: 0.143
[28,   390] loss: 0.137
[28,   420] loss: 0.128
[28,   450] loss: 0.145
[28,   480] loss: 0.174
[28,   510] loss: 0.192
[28,   540] loss: 0.201
[28,   570] loss: 0.176
[28,   600] loss: 0.143
[28,   630] loss: 0.154
[28,   660] loss: 0.157
[28,   690] loss: 0.109
[28,   720] loss: 0.206
[28,   750] loss: 0.117
[28,   780] loss: 0.142
[28,   810] loss: 0.183
[28,   840] loss: 0.153
[28,   870] loss: 0.195
[28,   900] loss: 0.204
[28,   930] loss: 0.112
[28,   960] loss: 0.101
[28,   990] loss: 0.161
[28,  1020] loss: 0.154
[28,  1050] loss: 0.153
[28,  1080] loss: 0.135
[28,  1110] loss: 0.135
[28,  1140] loss: 0.128
[28,  1170] loss: 0.165
[28,  1200] loss: 0.147
[28,  1230] loss: 0.172
```

```
[28, 1260] loss: 0.181
[28, 1290] loss: 0.206
[28, 1320] loss: 0.157
[28, 1350] loss: 0.167
[28, 1380] loss: 0.147
[28, 1410] loss: 0.126
[28, 1440] loss: 0.156
[28, 1470] loss: 0.204
[28, 1500] loss: 0.102
[28, 1530] loss: 0.198
[28, 1560] loss: 0.181
[28, 1590] loss: 0.163
[28, 1620] loss: 0.125
[28, 1650] loss: 0.153
[28, 1680] loss: 0.158
[28, 1710] loss: 0.156
[28, 1740] loss: 0.196
[28, 1770] loss: 0.119
[28, 1800] loss: 0.150
[28, 1830] loss: 0.162
[28, 1860] loss: 0.143
[28, 1890] loss: 0.177
[28, 1920] loss: 0.155
[28, 1950] loss: 0.190
[28, 1980] loss: 0.228
[28, 2010] loss: 0.185
[28, 2040] loss: 0.221
[28, 2070] loss: 0.120
[28, 2100] loss: 0.141
[28, 2130] loss: 0.183
[28, 2160] loss: 0.150
[28, 2190] loss: 0.178
evf intermediate_models/epoch27_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.951768
epoch 27 accuracy on train set is: 0.9517675390766994
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.977008

epoch 27 accuracy on test set is: 0.9770083605961468

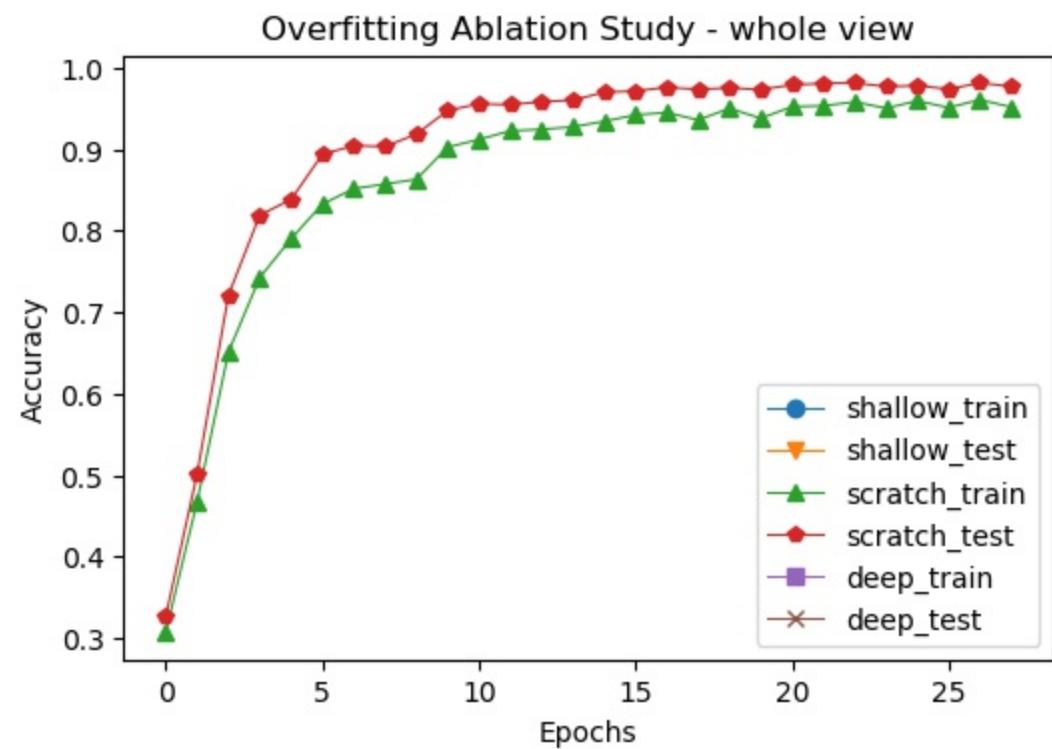
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

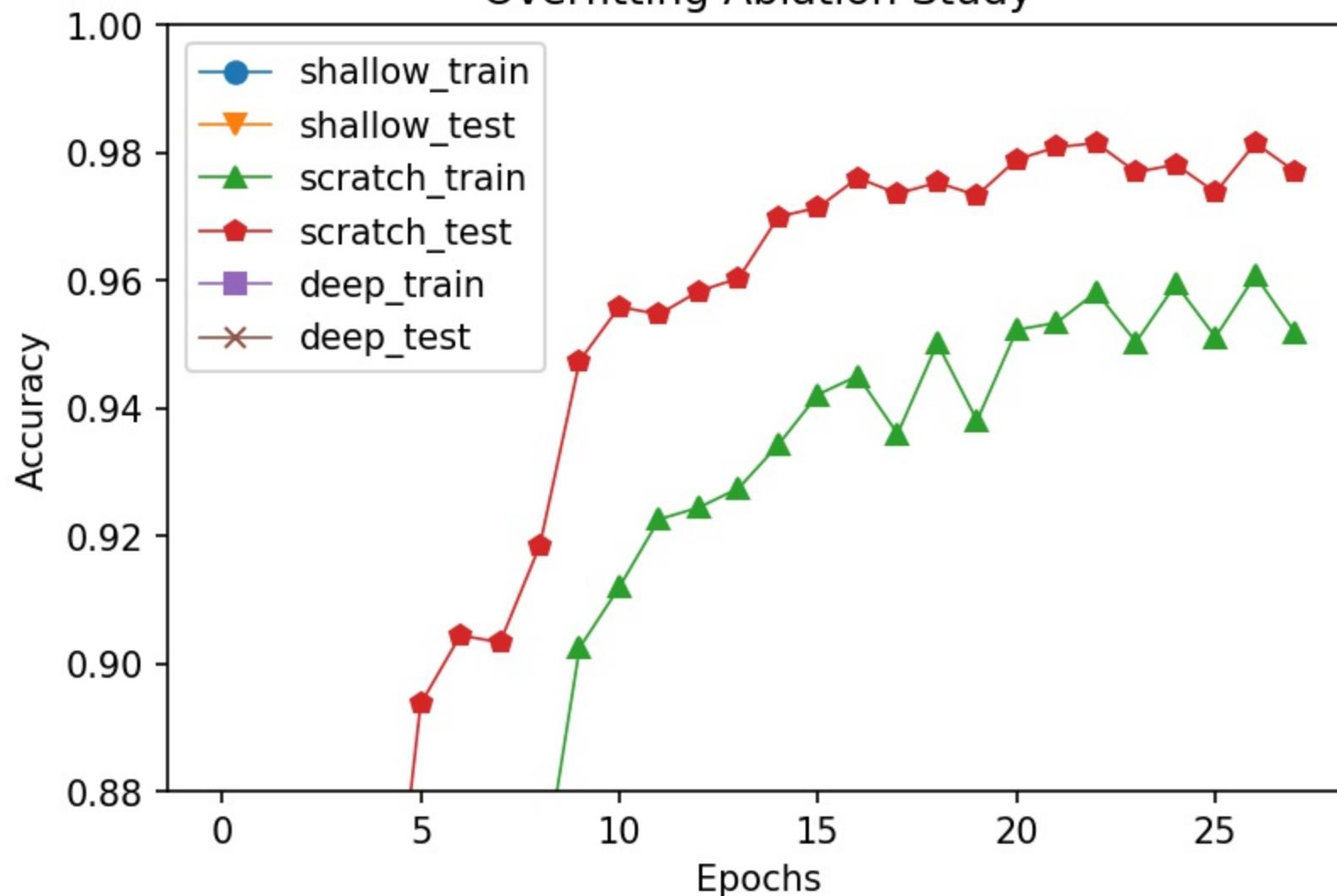
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 27 takes 318 seconds.

[29, 30] loss: 0.209

[29, 60] loss: 0.130

```
[29,    90] loss: 0.154
[29,   120] loss: 0.153
[29,   150] loss: 0.114
[29,   180] loss: 0.108
[29,   210] loss: 0.157
[29,   240] loss: 0.164
[29,   270] loss: 0.168
[29,   300] loss: 0.153
[29,   330] loss: 0.146
[29,   360] loss: 0.171
[29,   390] loss: 0.141
[29,   420] loss: 0.135
[29,   450] loss: 0.123
[29,   480] loss: 0.090
[29,   510] loss: 0.168
[29,   540] loss: 0.181
[29,   570] loss: 0.103
[29,   600] loss: 0.146
[29,   630] loss: 0.135
[29,   660] loss: 0.150
[29,   690] loss: 0.160
[29,   720] loss: 0.188
[29,   750] loss: 0.161
[29,   780] loss: 0.160
[29,   810] loss: 0.155
[29,   840] loss: 0.154
[29,   870] loss: 0.193
[29,   900] loss: 0.173
[29,   930] loss: 0.163
[29,   960] loss: 0.194
[29,   990] loss: 0.158
[29,  1020] loss: 0.190
[29,  1050] loss: 0.170
[29,  1080] loss: 0.192
[29,  1110] loss: 0.153
[29,  1140] loss: 0.207
[29,  1170] loss: 0.171
[29,  1200] loss: 0.167
[29,  1230] loss: 0.157
```

```
[29, 1260] loss: 0.136
[29, 1290] loss: 0.121
[29, 1320] loss: 0.161
[29, 1350] loss: 0.145
[29, 1380] loss: 0.148
[29, 1410] loss: 0.163
[29, 1440] loss: 0.140
[29, 1470] loss: 0.157
[29, 1500] loss: 0.144
[29, 1530] loss: 0.147
[29, 1560] loss: 0.151
[29, 1590] loss: 0.172
[29, 1620] loss: 0.140
[29, 1650] loss: 0.130
[29, 1680] loss: 0.149
[29, 1710] loss: 0.198
[29, 1740] loss: 0.194
[29, 1770] loss: 0.151
[29, 1800] loss: 0.161
[29, 1830] loss: 0.124
[29, 1860] loss: 0.109
[29, 1890] loss: 0.143
[29, 1920] loss: 0.096
[29, 1950] loss: 0.144
[29, 1980] loss: 0.156
[29, 2010] loss: 0.132
[29, 2040] loss: 0.184
[29, 2070] loss: 0.129
[29, 2100] loss: 0.162
[29, 2130] loss: 0.135
[29, 2160] loss: 0.165
[29, 2190] loss: 0.155
eval intermediate_models/epoch28_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.964490
epoch 28 accuracy on train set is: 0.9644901853871319
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.981552

epoch 28 accuracy on test set is: 0.9815521628498728

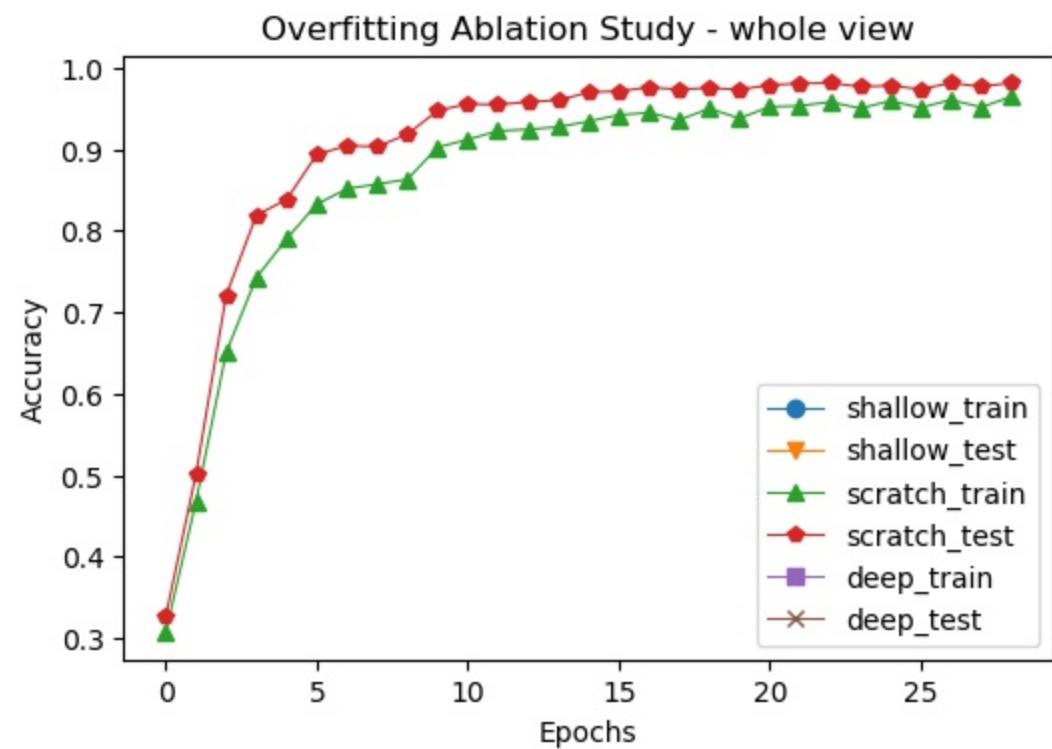
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

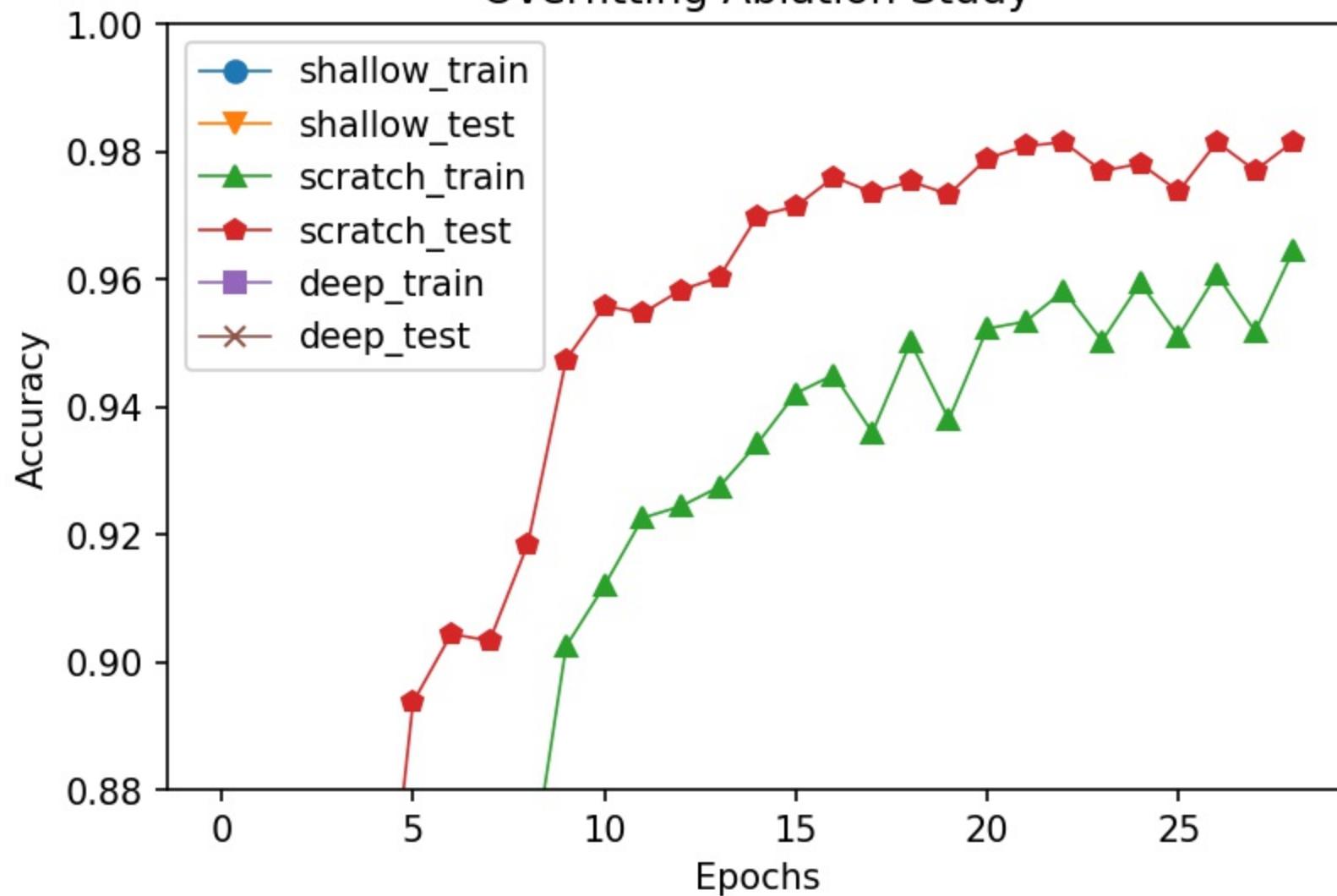
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 28 takes 318 seconds.

[30, 30] loss: 0.179

[30, 60] loss: 0.144

```
[30,    90] loss: 0.123
[30,   120] loss: 0.182
[30,   150] loss: 0.095
[30,   180] loss: 0.176
[30,   210] loss: 0.156
[30,   240] loss: 0.156
[30,   270] loss: 0.190
[30,   300] loss: 0.143
[30,   330] loss: 0.149
[30,   360] loss: 0.158
[30,   390] loss: 0.162
[30,   420] loss: 0.171
[30,   450] loss: 0.164
[30,   480] loss: 0.197
[30,   510] loss: 0.157
[30,   540] loss: 0.141
[30,   570] loss: 0.137
[30,   600] loss: 0.129
[30,   630] loss: 0.159
[30,   660] loss: 0.187
[30,   690] loss: 0.134
[30,   720] loss: 0.149
[30,   750] loss: 0.115
[30,   780] loss: 0.177
[30,   810] loss: 0.155
[30,   840] loss: 0.114
[30,   870] loss: 0.164
[30,   900] loss: 0.158
[30,   930] loss: 0.116
[30,   960] loss: 0.115
[30,   990] loss: 0.165
[30,  1020] loss: 0.127
[30,  1050] loss: 0.180
[30,  1080] loss: 0.163
[30,  1110] loss: 0.162
[30,  1140] loss: 0.126
[30,  1170] loss: 0.123
[30,  1200] loss: 0.222
[30,  1230] loss: 0.180
```

```
[30, 1260] loss: 0.144
[30, 1290] loss: 0.194
[30, 1320] loss: 0.124
[30, 1350] loss: 0.115
[30, 1380] loss: 0.156
[30, 1410] loss: 0.114
[30, 1440] loss: 0.138
[30, 1470] loss: 0.088
[30, 1500] loss: 0.183
[30, 1530] loss: 0.115
[30, 1560] loss: 0.139
[30, 1590] loss: 0.132
[30, 1620] loss: 0.178
[30, 1650] loss: 0.181
[30, 1680] loss: 0.157
[30, 1710] loss: 0.139
[30, 1740] loss: 0.133
[30, 1770] loss: 0.111
[30, 1800] loss: 0.114
[30, 1830] loss: 0.160
[30, 1860] loss: 0.165
[30, 1890] loss: 0.121
[30, 1920] loss: 0.164
[30, 1950] loss: 0.112
[30, 1980] loss: 0.115
[30, 2010] loss: 0.100
[30, 2040] loss: 0.182
[30, 2070] loss: 0.125
[30, 2100] loss: 0.173
[30, 2130] loss: 0.181
[30, 2160] loss: 0.175
[30, 2190] loss: 0.137
eval intermediate_models/epoch29_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.955675
epoch 29 accuracy on train set is: 0.9556752090149037
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.980098

epoch 29 accuracy on test set is: 0.9800981461286805

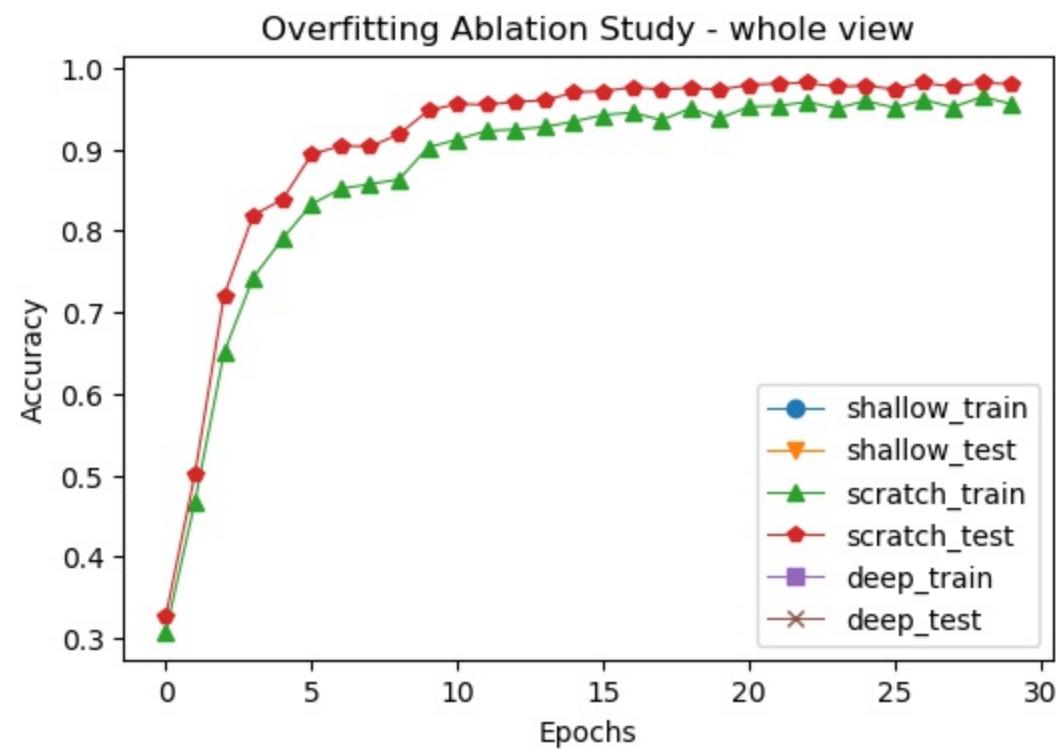
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	NaN	NaN	NaN	NaN

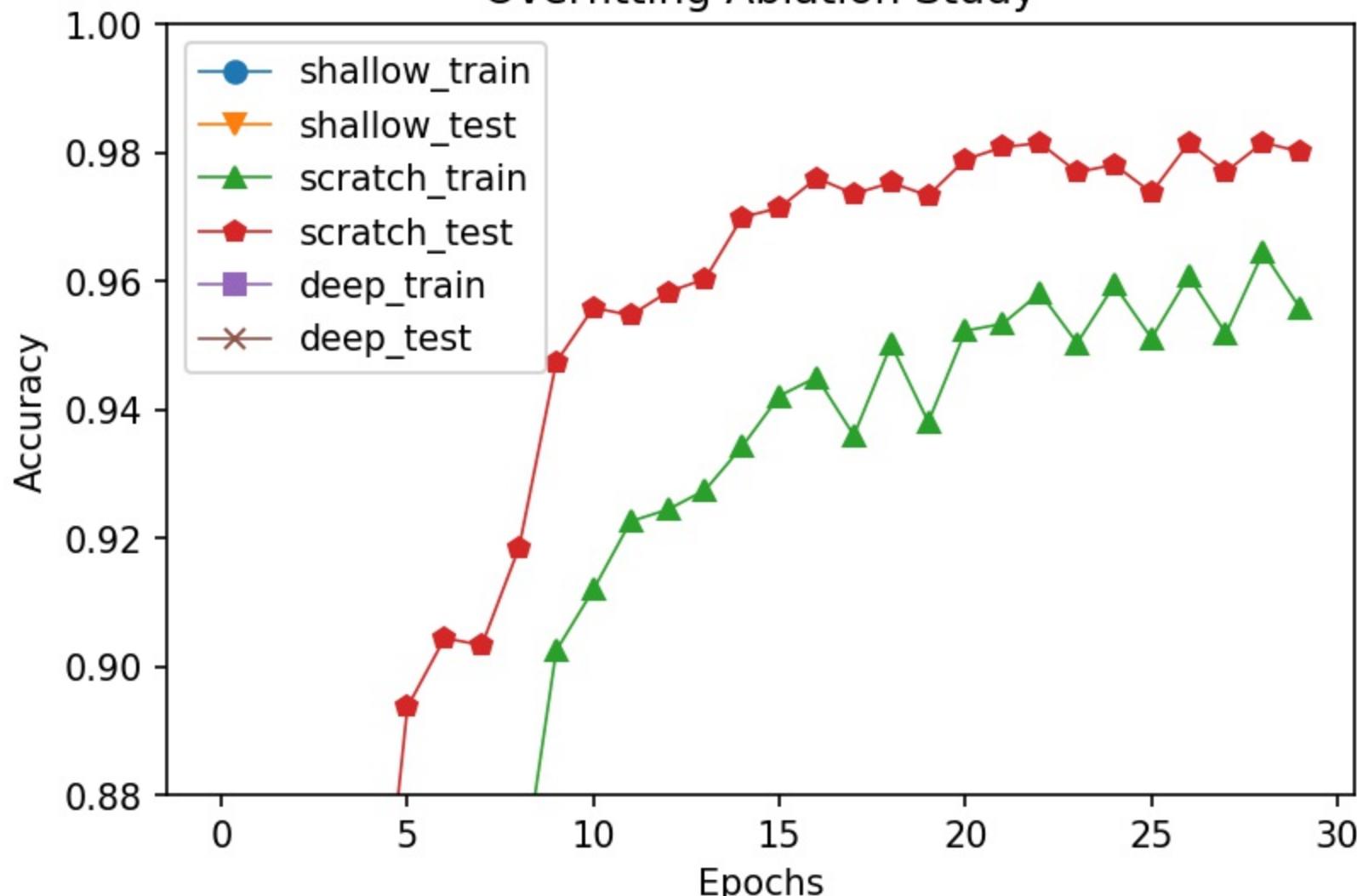
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 29 takes 317 seconds.

[31, 30] loss: 0.149

[31, 60] loss: 0.213

```
[31,    90] loss: 0.125
[31,   120] loss: 0.125
[31,   150] loss: 0.120
[31,   180] loss: 0.118
[31,   210] loss: 0.168
[31,   240] loss: 0.140
[31,   270] loss: 0.126
[31,   300] loss: 0.136
[31,   330] loss: 0.105
[31,   360] loss: 0.179
[31,   390] loss: 0.137
[31,   420] loss: 0.108
[31,   450] loss: 0.144
[31,   480] loss: 0.121
[31,   510] loss: 0.131
[31,   540] loss: 0.188
[31,   570] loss: 0.164
[31,   600] loss: 0.149
[31,   630] loss: 0.137
[31,   660] loss: 0.122
[31,   690] loss: 0.160
[31,   720] loss: 0.164
[31,   750] loss: 0.108
[31,   780] loss: 0.106
[31,   810] loss: 0.133
[31,   840] loss: 0.149
[31,   870] loss: 0.145
[31,   900] loss: 0.186
[31,   930] loss: 0.161
[31,   960] loss: 0.120
[31,   990] loss: 0.123
[31,  1020] loss: 0.116
[31,  1050] loss: 0.112
[31,  1080] loss: 0.129
[31,  1110] loss: 0.149
[31,  1140] loss: 0.139
[31,  1170] loss: 0.130
[31,  1200] loss: 0.149
[31,  1230] loss: 0.119
```

```
[31, 1260] loss: 0.168
[31, 1290] loss: 0.171
[31, 1320] loss: 0.126
[31, 1350] loss: 0.134
[31, 1380] loss: 0.164
[31, 1410] loss: 0.183
[31, 1440] loss: 0.126
[31, 1470] loss: 0.152
[31, 1500] loss: 0.133
[31, 1530] loss: 0.178
[31, 1560] loss: 0.121
[31, 1590] loss: 0.172
[31, 1620] loss: 0.148
[31, 1650] loss: 0.170
[31, 1680] loss: 0.106
[31, 1710] loss: 0.122
[31, 1740] loss: 0.119
[31, 1770] loss: 0.127
[31, 1800] loss: 0.114
[31, 1830] loss: 0.108
[31, 1860] loss: 0.168
[31, 1890] loss: 0.186
[31, 1920] loss: 0.110
[31, 1950] loss: 0.187
[31, 1980] loss: 0.160
[31, 2010] loss: 0.180
[31, 2040] loss: 0.138
[31, 2070] loss: 0.133
[31, 2100] loss: 0.159
[31, 2130] loss: 0.101
[31, 2160] loss: 0.182
[31, 2190] loss: 0.158
eval intermediate_models/epoch30_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.965581
epoch 30 accuracy on train set is: 0.9655806979280261
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	NaN	NaN	NaN
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.984097

epoch 30 accuracy on test set is: 0.9840966921119593

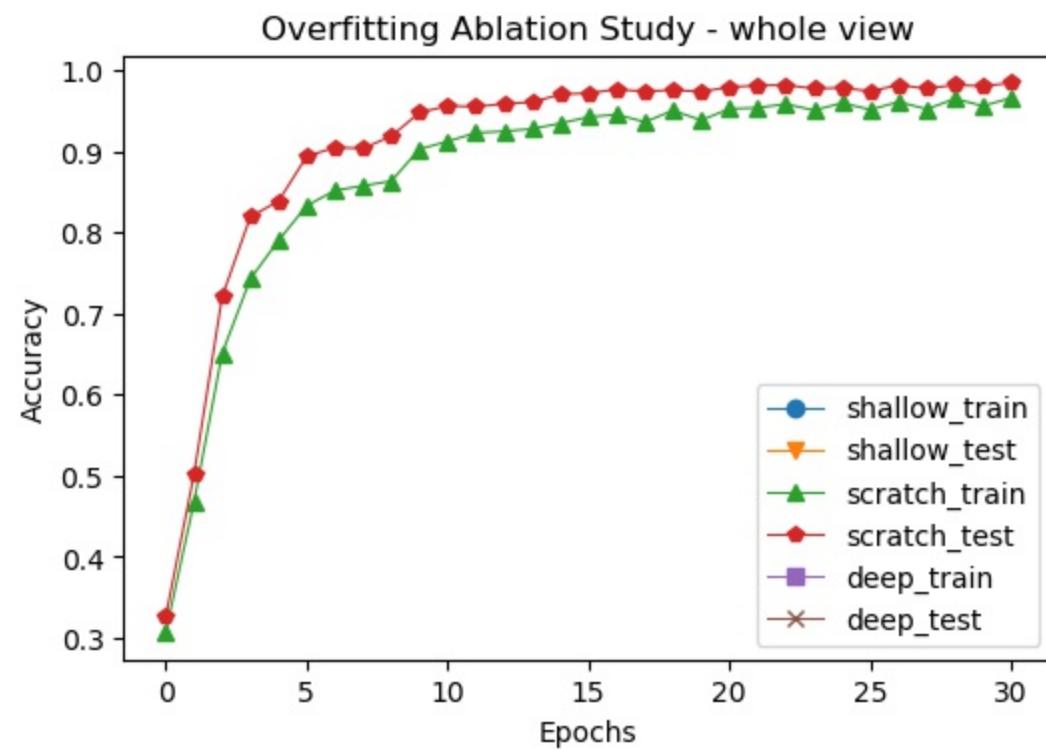
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

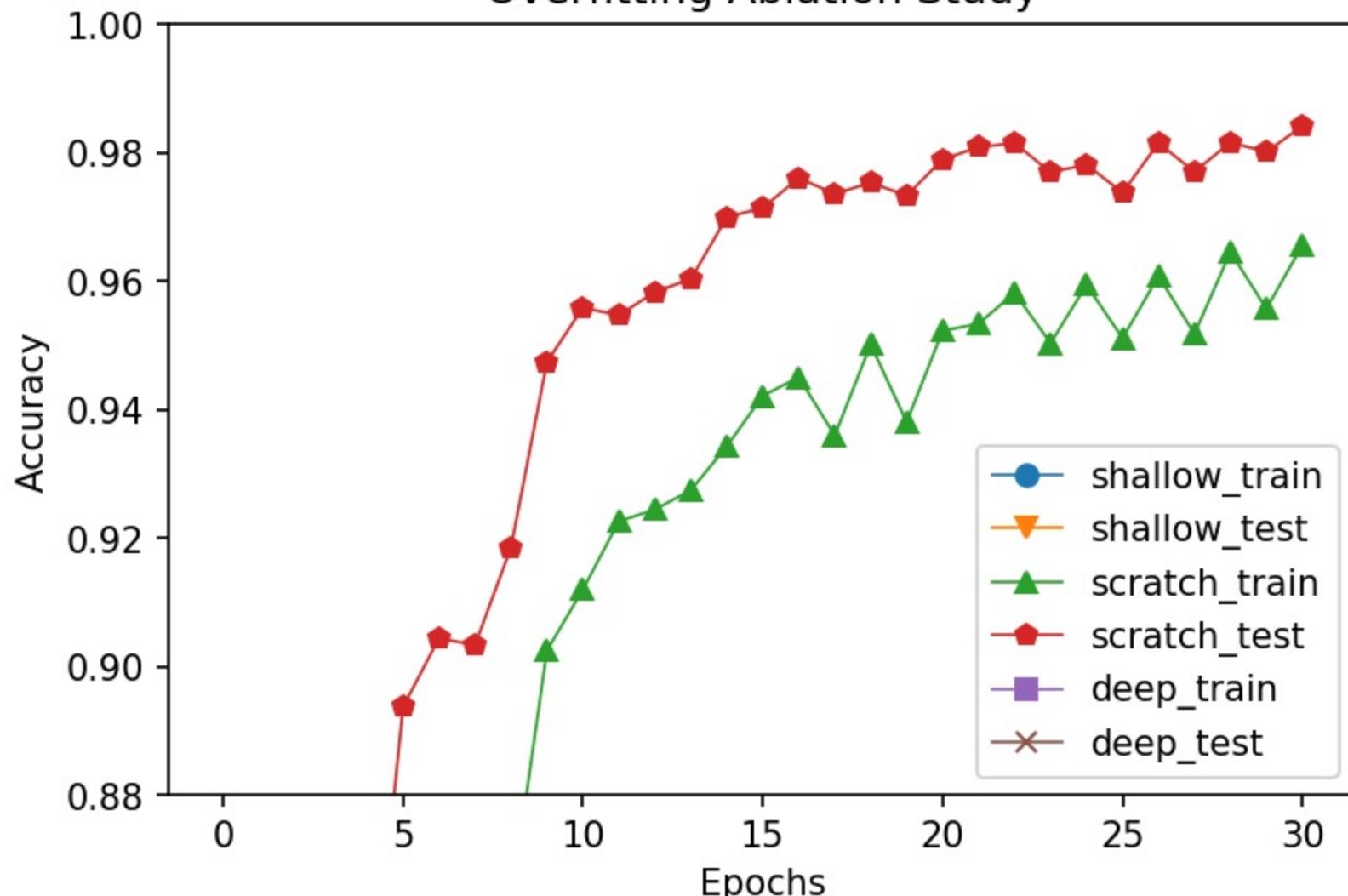
31	NaN	NaN	NaN	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 30 takes 317 seconds.

[32, 30] loss: 0.085

[32, 60] loss: 0.153

```
[32,    90] loss: 0.107
[32,   120] loss: 0.137
[32,   150] loss: 0.177
[32,   180] loss: 0.151
[32,   210] loss: 0.143
[32,   240] loss: 0.195
[32,   270] loss: 0.107
[32,   300] loss: 0.135
[32,   330] loss: 0.139
[32,   360] loss: 0.134
[32,   390] loss: 0.112
[32,   420] loss: 0.119
[32,   450] loss: 0.113
[32,   480] loss: 0.106
[32,   510] loss: 0.134
[32,   540] loss: 0.133
[32,   570] loss: 0.129
[32,   600] loss: 0.145
[32,   630] loss: 0.124
[32,   660] loss: 0.146
[32,   690] loss: 0.111
[32,   720] loss: 0.103
[32,   750] loss: 0.171
[32,   780] loss: 0.103
[32,   810] loss: 0.149
[32,   840] loss: 0.139
[32,   870] loss: 0.167
[32,   900] loss: 0.114
[32,   930] loss: 0.105
[32,   960] loss: 0.141
[32,   990] loss: 0.181
[32,  1020] loss: 0.135
[32,  1050] loss: 0.115
[32,  1080] loss: 0.145
[32,  1110] loss: 0.095
[32,  1140] loss: 0.138
[32,  1170] loss: 0.134
[32,  1200] loss: 0.140
[32,  1230] loss: 0.131
```

```
[32, 1260] loss: 0.165
[32, 1290] loss: 0.174
[32, 1320] loss: 0.121
[32, 1350] loss: 0.127
[32, 1380] loss: 0.133
[32, 1410] loss: 0.129
[32, 1440] loss: 0.156
[32, 1470] loss: 0.099
[32, 1500] loss: 0.128
[32, 1530] loss: 0.118
[32, 1560] loss: 0.163
[32, 1590] loss: 0.159
[32, 1620] loss: 0.122
[32, 1650] loss: 0.139
[32, 1680] loss: 0.151
[32, 1710] loss: 0.156
[32, 1740] loss: 0.176
[32, 1770] loss: 0.162
[32, 1800] loss: 0.190
[32, 1830] loss: 0.125
[32, 1860] loss: 0.123
[32, 1890] loss: 0.148
[32, 1920] loss: 0.150
[32, 1950] loss: 0.127
[32, 1980] loss: 0.149
[32, 2010] loss: 0.156
[32, 2040] loss: 0.151
[32, 2070] loss: 0.136
[32, 2100] loss: 0.130
[32, 2130] loss: 0.154
[32, 2160] loss: 0.124
[32, 2190] loss: 0.099
eval intermediate_models/epoch31_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.965535
epoch 31 accuracy on train set is: 0.9655352599054889
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	NaN	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.983733

epoch 31 accuracy on test set is: 0.9837331879316612

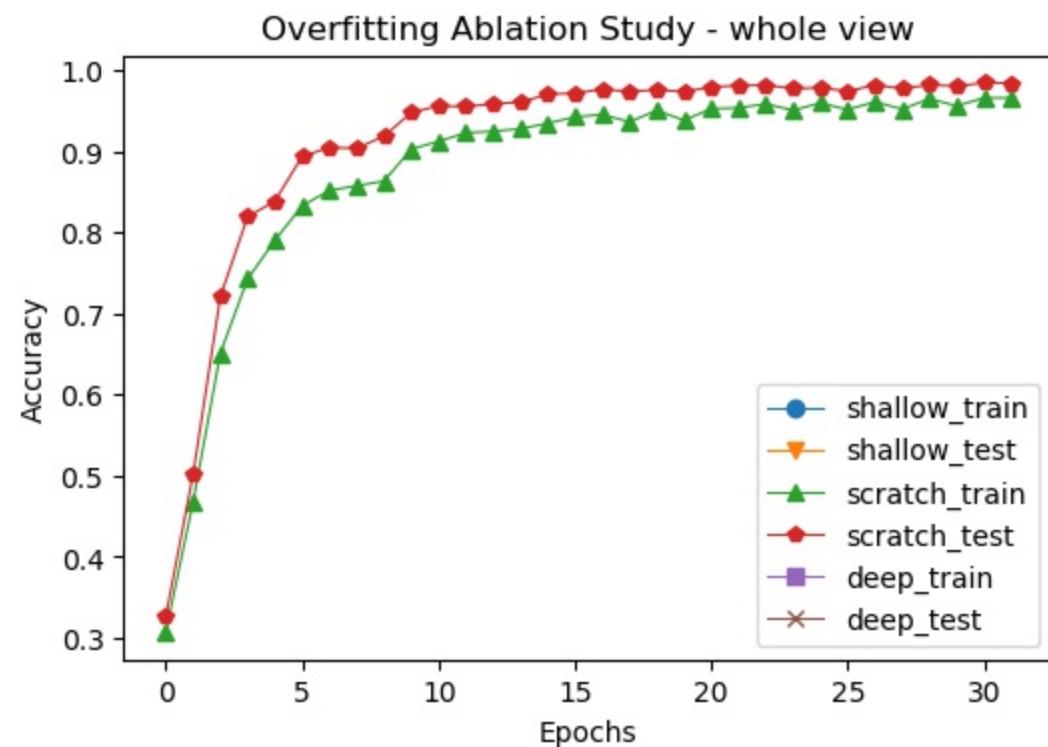
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

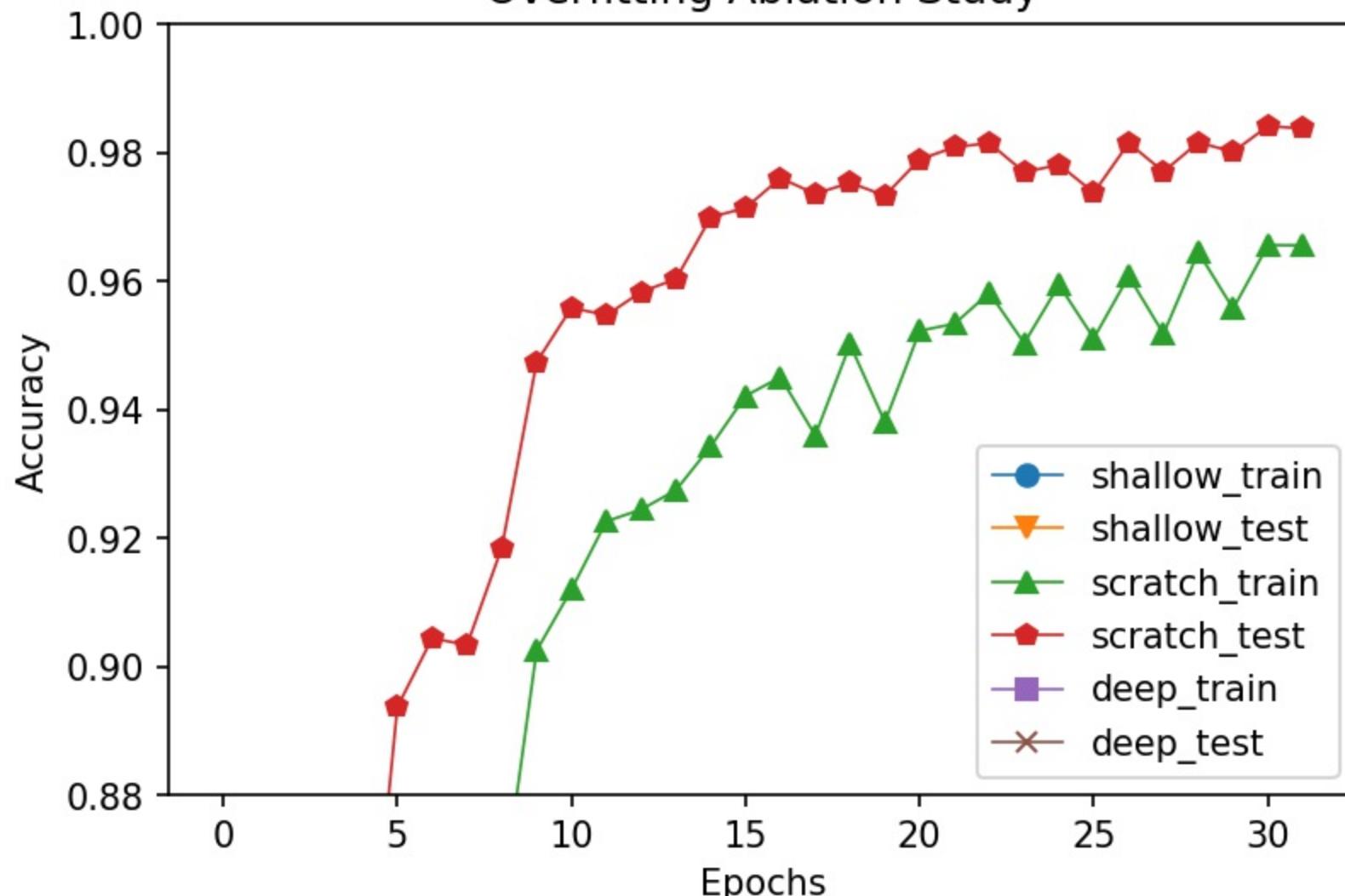
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	NaN	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 31 takes 317 seconds.

[33, 30] loss: 0.143

[33, 60] loss: 0.164

```
[33,    90] loss: 0.178
[33,   120] loss: 0.205
[33,   150] loss: 0.141
[33,   180] loss: 0.117
[33,   210] loss: 0.114
[33,   240] loss: 0.184
[33,   270] loss: 0.112
[33,   300] loss: 0.150
[33,   330] loss: 0.157
[33,   360] loss: 0.140
[33,   390] loss: 0.181
[33,   420] loss: 0.137
[33,   450] loss: 0.117
[33,   480] loss: 0.097
[33,   510] loss: 0.115
[33,   540] loss: 0.155
[33,   570] loss: 0.145
[33,   600] loss: 0.119
[33,   630] loss: 0.171
[33,   660] loss: 0.103
[33,   690] loss: 0.139
[33,   720] loss: 0.154
[33,   750] loss: 0.165
[33,   780] loss: 0.112
[33,   810] loss: 0.145
[33,   840] loss: 0.153
[33,   870] loss: 0.097
[33,   900] loss: 0.130
[33,   930] loss: 0.118
[33,   960] loss: 0.115
[33,   990] loss: 0.131
[33,  1020] loss: 0.168
[33,  1050] loss: 0.172
[33,  1080] loss: 0.154
[33,  1110] loss: 0.175
[33,  1140] loss: 0.145
[33,  1170] loss: 0.179
[33,  1200] loss: 0.162
[33,  1230] loss: 0.100
```

```
[33, 1260] loss: 0.108
[33, 1290] loss: 0.175
[33, 1320] loss: 0.106
[33, 1350] loss: 0.132
[33, 1380] loss: 0.148
[33, 1410] loss: 0.167
[33, 1440] loss: 0.174
[33, 1470] loss: 0.166
[33, 1500] loss: 0.159
[33, 1530] loss: 0.160
[33, 1560] loss: 0.145
[33, 1590] loss: 0.118
[33, 1620] loss: 0.169
[33, 1650] loss: 0.146
[33, 1680] loss: 0.121
[33, 1710] loss: 0.105
[33, 1740] loss: 0.123
[33, 1770] loss: 0.157
[33, 1800] loss: 0.125
[33, 1830] loss: 0.148
[33, 1860] loss: 0.145
[33, 1890] loss: 0.124
[33, 1920] loss: 0.112
[33, 1950] loss: 0.087
[33, 1980] loss: 0.119
[33, 2010] loss: 0.156
[33, 2040] loss: 0.167
[33, 2070] loss: 0.126
[33, 2100] loss: 0.154
[33, 2130] loss: 0.160
[33, 2160] loss: 0.112
[33, 2190] loss: 0.104
eval intermediate_models/epoch32_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.971192
epoch 32 accuracy on train set is: 0.9711922937113777
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	NaN	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988731

epoch 32 accuracy on test set is: 0.9887313704107598

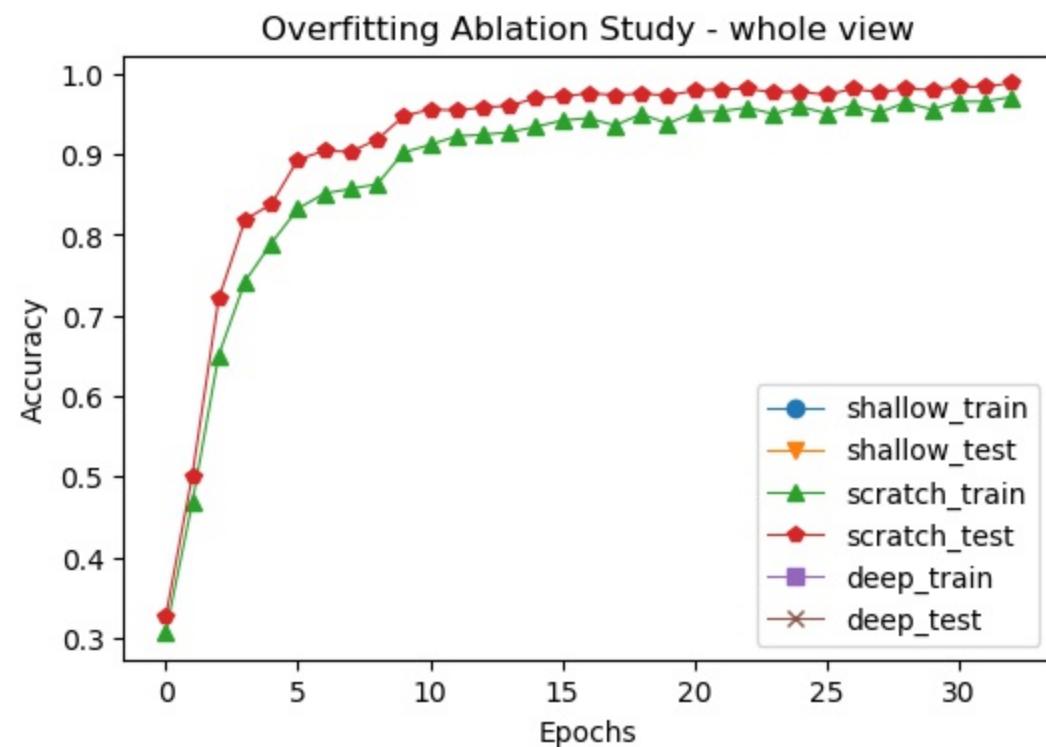
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

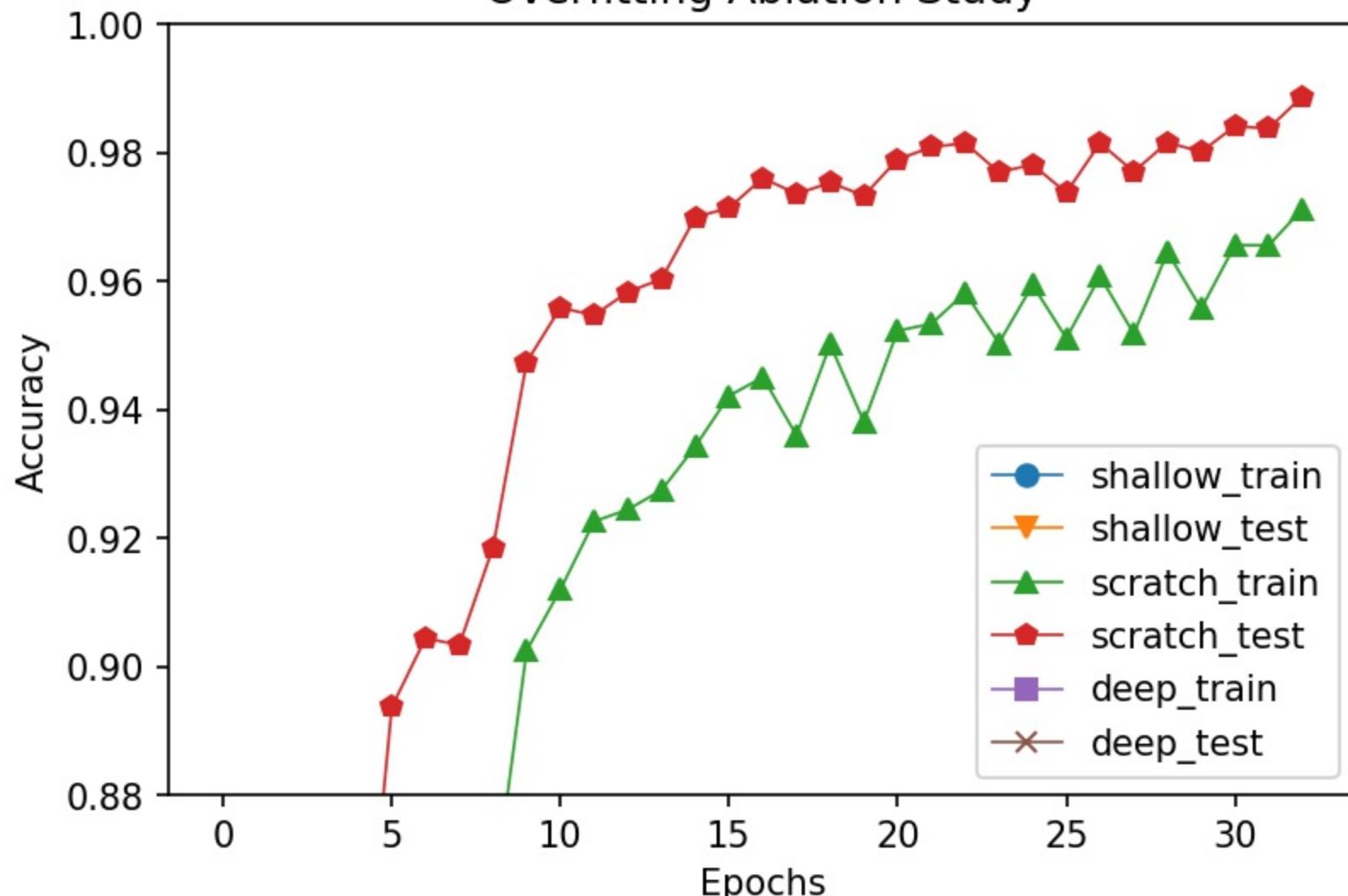
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	NaN	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 32 takes 317 seconds.

[34, 30] loss: 0.146

[34, 60] loss: 0.131

```
[34,    90] loss: 0.146
[34,   120] loss: 0.197
[34,   150] loss: 0.126
[34,   180] loss: 0.127
[34,   210] loss: 0.126
[34,   240] loss: 0.146
[34,   270] loss: 0.170
[34,   300] loss: 0.135
[34,   330] loss: 0.127
[34,   360] loss: 0.172
[34,   390] loss: 0.133
[34,   420] loss: 0.154
[34,   450] loss: 0.171
[34,   480] loss: 0.137
[34,   510] loss: 0.172
[34,   540] loss: 0.167
[34,   570] loss: 0.142
[34,   600] loss: 0.183
[34,   630] loss: 0.118
[34,   660] loss: 0.158
[34,   690] loss: 0.144
[34,   720] loss: 0.169
[34,   750] loss: 0.160
[34,   780] loss: 0.180
[34,   810] loss: 0.152
[34,   840] loss: 0.123
[34,   870] loss: 0.121
[34,   900] loss: 0.121
[34,   930] loss: 0.144
[34,   960] loss: 0.184
[34,   990] loss: 0.177
[34,  1020] loss: 0.131
[34,  1050] loss: 0.149
[34,  1080] loss: 0.120
[34,  1110] loss: 0.171
[34,  1140] loss: 0.132
[34,  1170] loss: 0.149
[34,  1200] loss: 0.118
[34,  1230] loss: 0.198
```

```
[34, 1260] loss: 0.168
[34, 1290] loss: 0.137
[34, 1320] loss: 0.160
[34, 1350] loss: 0.106
[34, 1380] loss: 0.089
[34, 1410] loss: 0.188
[34, 1440] loss: 0.099
[34, 1470] loss: 0.117
[34, 1500] loss: 0.089
[34, 1530] loss: 0.136
[34, 1560] loss: 0.119
[34, 1590] loss: 0.153
[34, 1620] loss: 0.115
[34, 1650] loss: 0.101
[34, 1680] loss: 0.204
[34, 1710] loss: 0.132
[34, 1740] loss: 0.156
[34, 1770] loss: 0.118
[34, 1800] loss: 0.091
[34, 1830] loss: 0.121
[34, 1860] loss: 0.102
[34, 1890] loss: 0.105
[34, 1920] loss: 0.139
[34, 1950] loss: 0.124
[34, 1980] loss: 0.156
[34, 2010] loss: 0.158
[34, 2040] loss: 0.147
[34, 2070] loss: 0.116
[34, 2100] loss: 0.085
[34, 2130] loss: 0.141
[34, 2160] loss: 0.121
[34, 2190] loss: 0.141
eval intermediate_models/epoch33_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.966603
epoch 33 accuracy on train set is: 0.9666030534351145
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	NaN	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.984006

epoch 33 accuracy on test set is: 0.9840058160668848

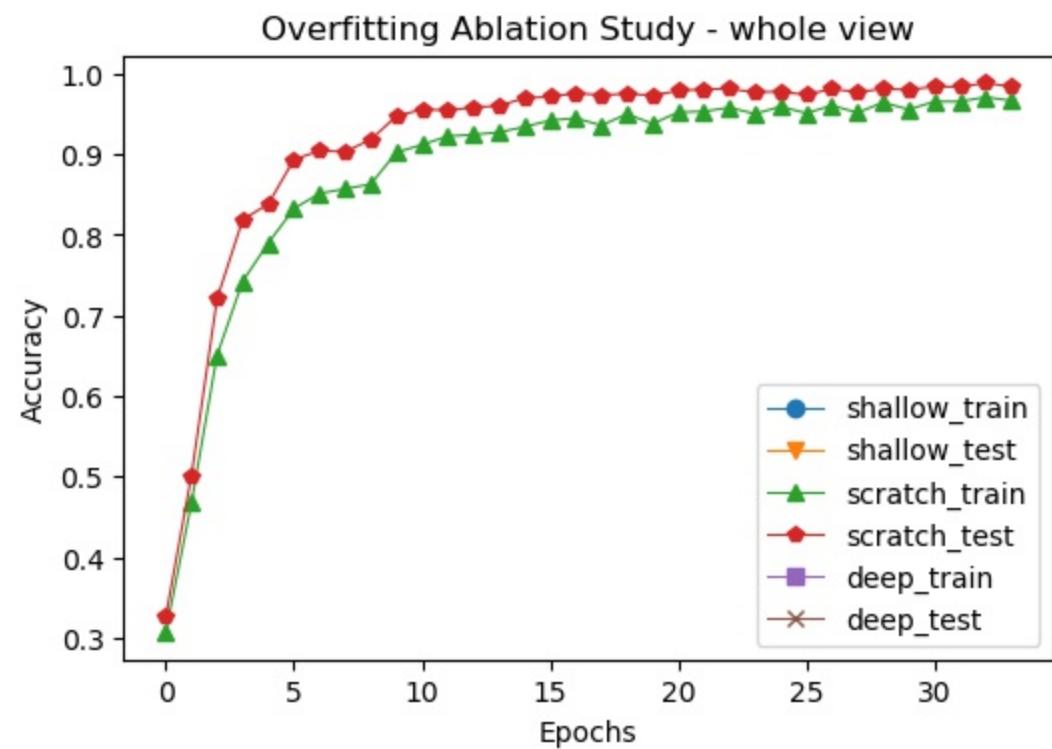
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

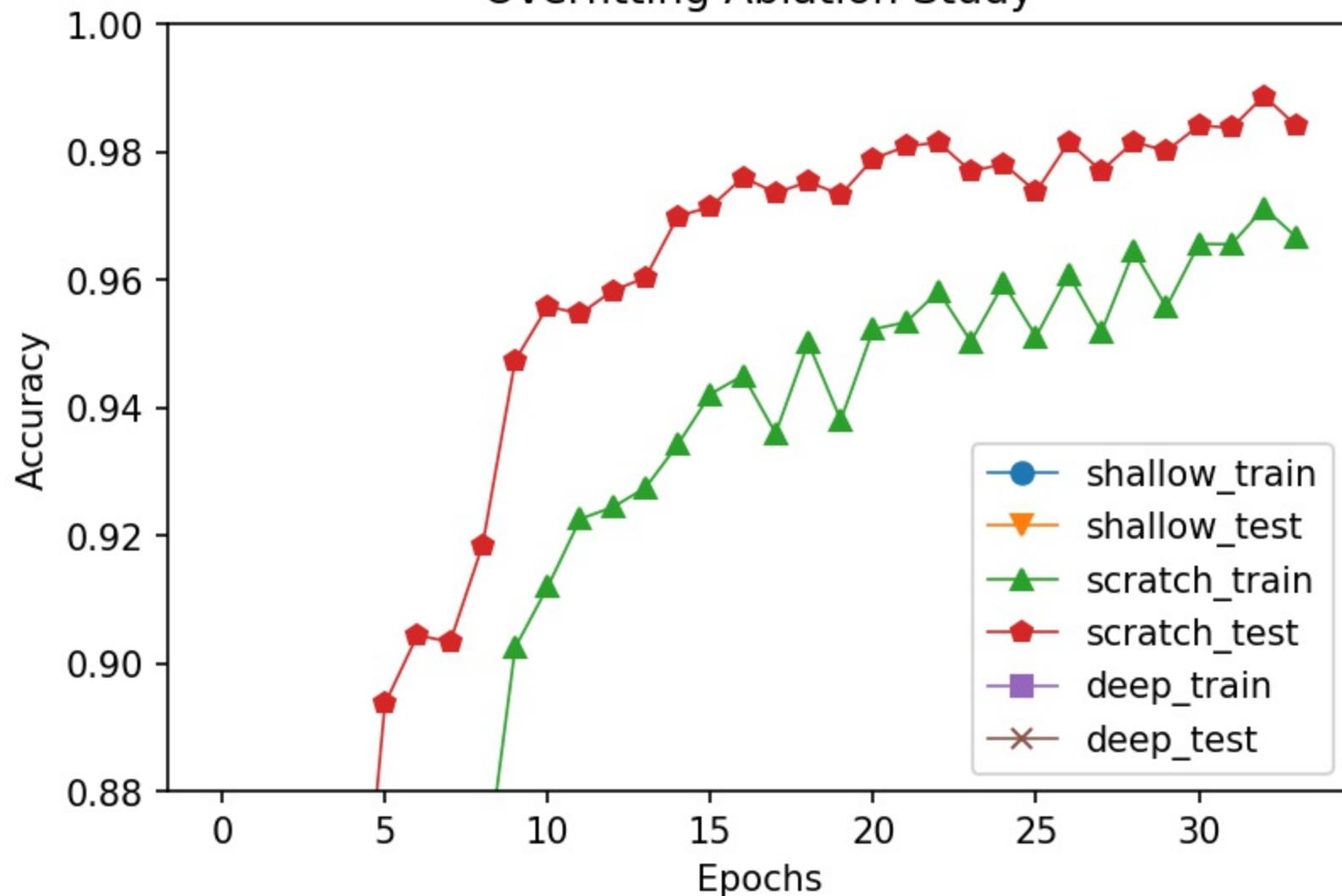
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	NaN	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 33 takes 317 seconds.

[35, 30] loss: 0.160

[35, 60] loss: 0.145

```
[35,    90] loss: 0.144
[35,   120] loss: 0.128
[35,   150] loss: 0.123
[35,   180] loss: 0.096
[35,   210] loss: 0.105
[35,   240] loss: 0.142
[35,   270] loss: 0.148
[35,   300] loss: 0.119
[35,   330] loss: 0.132
[35,   360] loss: 0.122
[35,   390] loss: 0.121
[35,   420] loss: 0.104
[35,   450] loss: 0.114
[35,   480] loss: 0.120
[35,   510] loss: 0.091
[35,   540] loss: 0.180
[35,   570] loss: 0.152
[35,   600] loss: 0.148
[35,   630] loss: 0.134
[35,   660] loss: 0.072
[35,   690] loss: 0.107
[35,   720] loss: 0.127
[35,   750] loss: 0.131
[35,   780] loss: 0.162
[35,   810] loss: 0.120
[35,   840] loss: 0.145
[35,   870] loss: 0.131
[35,   900] loss: 0.100
[35,   930] loss: 0.125
[35,   960] loss: 0.124
[35,   990] loss: 0.135
[35,  1020] loss: 0.119
[35,  1050] loss: 0.129
[35,  1080] loss: 0.138
[35,  1110] loss: 0.136
[35,  1140] loss: 0.162
[35,  1170] loss: 0.161
[35,  1200] loss: 0.092
[35,  1230] loss: 0.155
```

```
[35, 1260] loss: 0.138
[35, 1290] loss: 0.123
[35, 1320] loss: 0.128
[35, 1350] loss: 0.147
[35, 1380] loss: 0.141
[35, 1410] loss: 0.124
[35, 1440] loss: 0.113
[35, 1470] loss: 0.088
[35, 1500] loss: 0.095
[35, 1530] loss: 0.136
[35, 1560] loss: 0.116
[35, 1590] loss: 0.151
[35, 1620] loss: 0.177
[35, 1650] loss: 0.124
[35, 1680] loss: 0.100
[35, 1710] loss: 0.140
[35, 1740] loss: 0.138
[35, 1770] loss: 0.174
[35, 1800] loss: 0.191
[35, 1830] loss: 0.093
[35, 1860] loss: 0.090
[35, 1890] loss: 0.125
[35, 1920] loss: 0.146
[35, 1950] loss: 0.151
[35, 1980] loss: 0.117
[35, 2010] loss: 0.134
[35, 2040] loss: 0.097
[35, 2070] loss: 0.135
[35, 2100] loss: 0.136
[35, 2130] loss: 0.171
[35, 2160] loss: 0.145
[35, 2190] loss: 0.132
eval intermediate_models/epoch34_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.968534
epoch 34 accuracy on train set is: 0.968534169392948
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	NaN	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.986732

epoch 34 accuracy on test set is: 0.9867320974191203

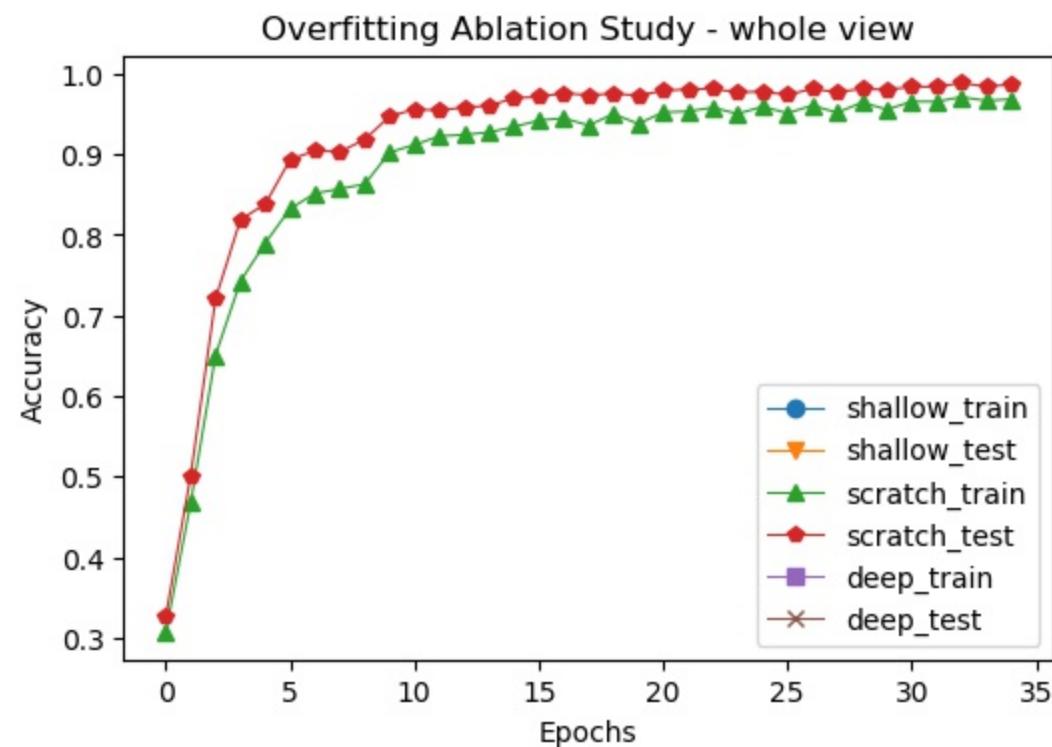
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

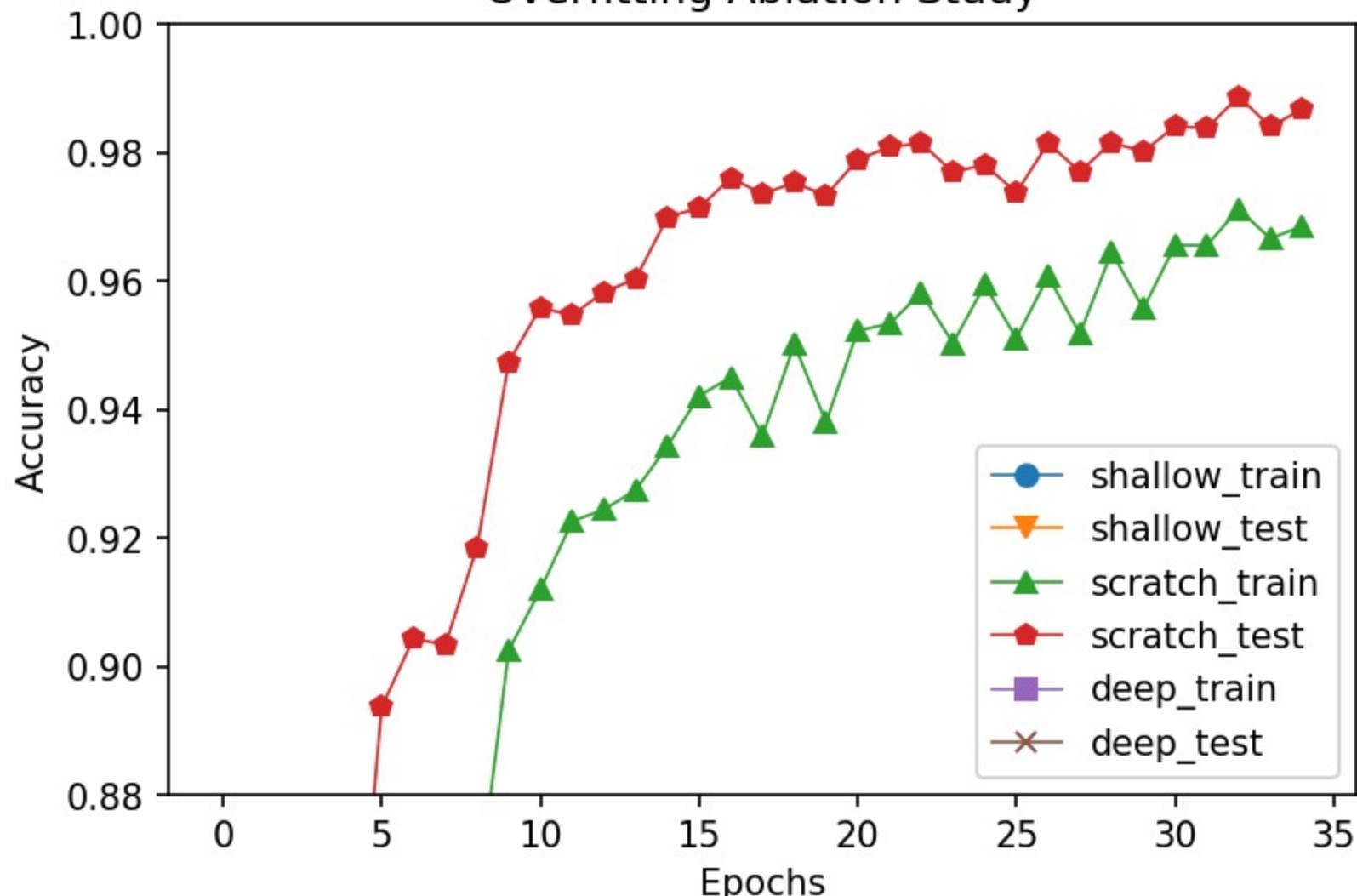
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	NaN	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 34 takes 317 seconds.

[36, 30] loss: 0.162

[36, 60] loss: 0.162

```
[36,    90] loss: 0.112
[36,   120] loss: 0.127
[36,   150] loss: 0.163
[36,   180] loss: 0.101
[36,   210] loss: 0.116
[36,   240] loss: 0.128
[36,   270] loss: 0.101
[36,   300] loss: 0.121
[36,   330] loss: 0.122
[36,   360] loss: 0.179
[36,   390] loss: 0.197
[36,   420] loss: 0.162
[36,   450] loss: 0.161
[36,   480] loss: 0.106
[36,   510] loss: 0.159
[36,   540] loss: 0.093
[36,   570] loss: 0.158
[36,   600] loss: 0.165
[36,   630] loss: 0.152
[36,   660] loss: 0.110
[36,   690] loss: 0.116
[36,   720] loss: 0.154
[36,   750] loss: 0.090
[36,   780] loss: 0.116
[36,   810] loss: 0.117
[36,   840] loss: 0.144
[36,   870] loss: 0.109
[36,   900] loss: 0.146
[36,   930] loss: 0.111
[36,   960] loss: 0.160
[36,   990] loss: 0.087
[36,  1020] loss: 0.076
[36,  1050] loss: 0.129
[36,  1080] loss: 0.104
[36,  1110] loss: 0.128
[36,  1140] loss: 0.133
[36,  1170] loss: 0.127
[36,  1200] loss: 0.160
[36,  1230] loss: 0.121
```

```
[36, 1260] loss: 0.134
[36, 1290] loss: 0.077
[36, 1320] loss: 0.104
[36, 1350] loss: 0.119
[36, 1380] loss: 0.125
[36, 1410] loss: 0.140
[36, 1440] loss: 0.159
[36, 1470] loss: 0.125
[36, 1500] loss: 0.116
[36, 1530] loss: 0.094
[36, 1560] loss: 0.097
[36, 1590] loss: 0.077
[36, 1620] loss: 0.116
[36, 1650] loss: 0.107
[36, 1680] loss: 0.136
[36, 1710] loss: 0.124
[36, 1740] loss: 0.120
[36, 1770] loss: 0.144
[36, 1800] loss: 0.181
[36, 1830] loss: 0.175
[36, 1860] loss: 0.119
[36, 1890] loss: 0.137
[36, 1920] loss: 0.115
[36, 1950] loss: 0.148
[36, 1980] loss: 0.143
[36, 2010] loss: 0.137
[36, 2040] loss: 0.109
[36, 2070] loss: 0.108
[36, 2100] loss: 0.097
[36, 2130] loss: 0.115
[36, 2160] loss: 0.148
[36, 2190] loss: 0.173
eval intermediate_models/epoch35_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.964786
epoch 35 accuracy on train set is: 0.9647855325336241
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	NaN	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.983006

epoch 35 accuracy on test set is: 0.9830061795710651

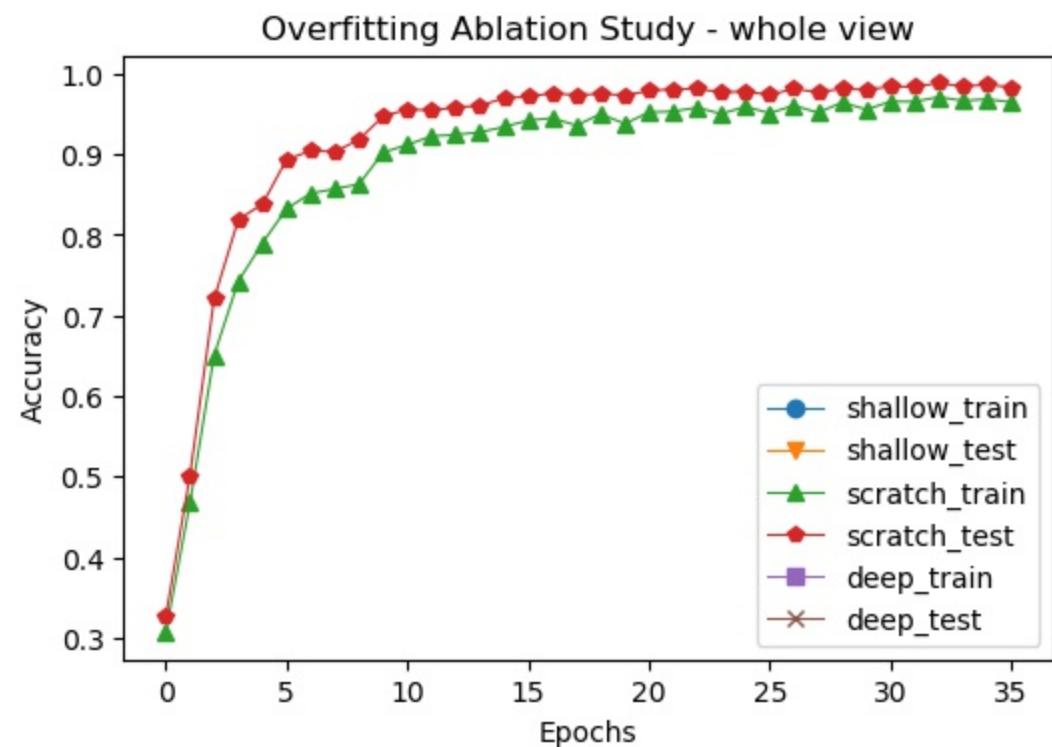
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

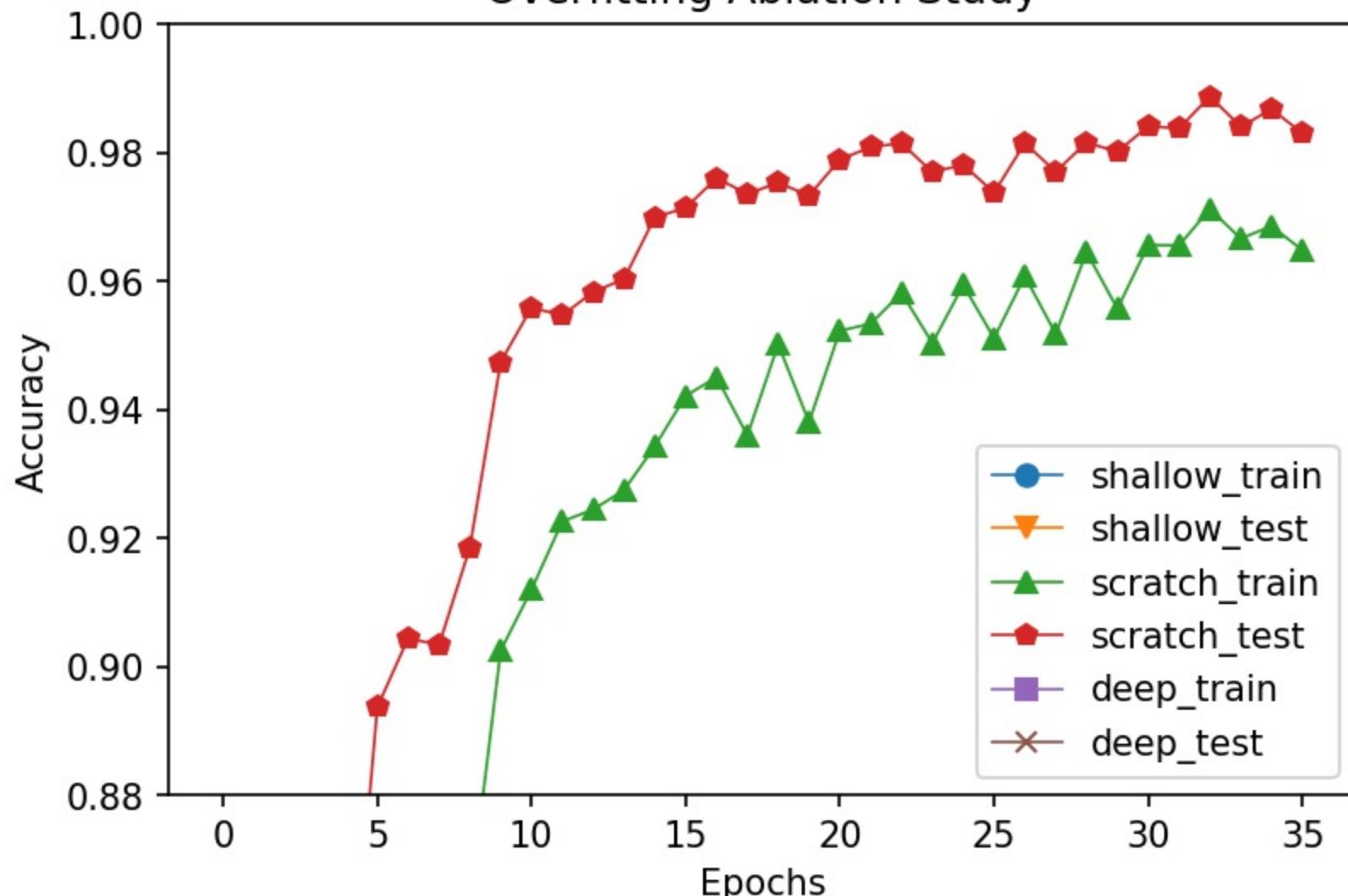
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 35 takes 317 seconds.

[37, 30] loss: 0.156

[37, 60] loss: 0.092

```
[37,    90] loss: 0.104
[37,   120] loss: 0.118
[37,   150] loss: 0.172
[37,   180] loss: 0.129
[37,   210] loss: 0.112
[37,   240] loss: 0.139
[37,   270] loss: 0.134
[37,   300] loss: 0.173
[37,   330] loss: 0.080
[37,   360] loss: 0.086
[37,   390] loss: 0.114
[37,   420] loss: 0.180
[37,   450] loss: 0.174
[37,   480] loss: 0.134
[37,   510] loss: 0.108
[37,   540] loss: 0.131
[37,   570] loss: 0.119
[37,   600] loss: 0.121
[37,   630] loss: 0.127
[37,   660] loss: 0.115
[37,   690] loss: 0.187
[37,   720] loss: 0.152
[37,   750] loss: 0.142
[37,   780] loss: 0.128
[37,   810] loss: 0.116
[37,   840] loss: 0.147
[37,   870] loss: 0.109
[37,   900] loss: 0.101
[37,   930] loss: 0.133
[37,   960] loss: 0.180
[37,   990] loss: 0.108
[37,  1020] loss: 0.179
[37,  1050] loss: 0.154
[37,  1080] loss: 0.126
[37,  1110] loss: 0.084
[37,  1140] loss: 0.106
[37,  1170] loss: 0.101
[37,  1200] loss: 0.110
[37,  1230] loss: 0.145
```

```
[37, 1260] loss: 0.100
[37, 1290] loss: 0.122
[37, 1320] loss: 0.176
[37, 1350] loss: 0.096
[37, 1380] loss: 0.121
[37, 1410] loss: 0.121
[37, 1440] loss: 0.125
[37, 1470] loss: 0.087
[37, 1500] loss: 0.110
[37, 1530] loss: 0.104
[37, 1560] loss: 0.099
[37, 1590] loss: 0.114
[37, 1620] loss: 0.141
[37, 1650] loss: 0.119
[37, 1680] loss: 0.133
[37, 1710] loss: 0.114
[37, 1740] loss: 0.109
[37, 1770] loss: 0.124
[37, 1800] loss: 0.150
[37, 1830] loss: 0.133
[37, 1860] loss: 0.140
[37, 1890] loss: 0.092
[37, 1920] loss: 0.130
[37, 1950] loss: 0.095
[37, 1980] loss: 0.149
[37, 2010] loss: 0.134
[37, 2040] loss: 0.115
[37, 2070] loss: 0.118
[37, 2100] loss: 0.136
[37, 2130] loss: 0.136
[37, 2160] loss: 0.085
[37, 2190] loss: 0.104
eval intermediate_models/epoch36_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.971828
epoch 36 accuracy on train set is: 0.9718284260268993
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988822

epoch 36 accuracy on test set is: 0.9888222464558343

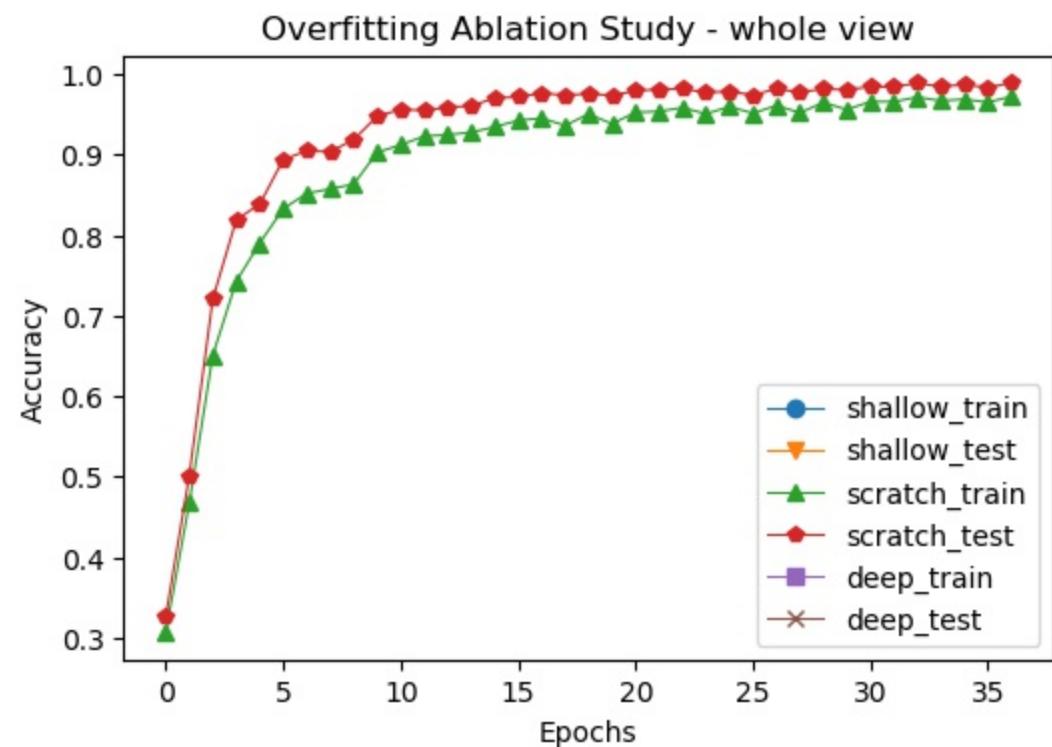
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

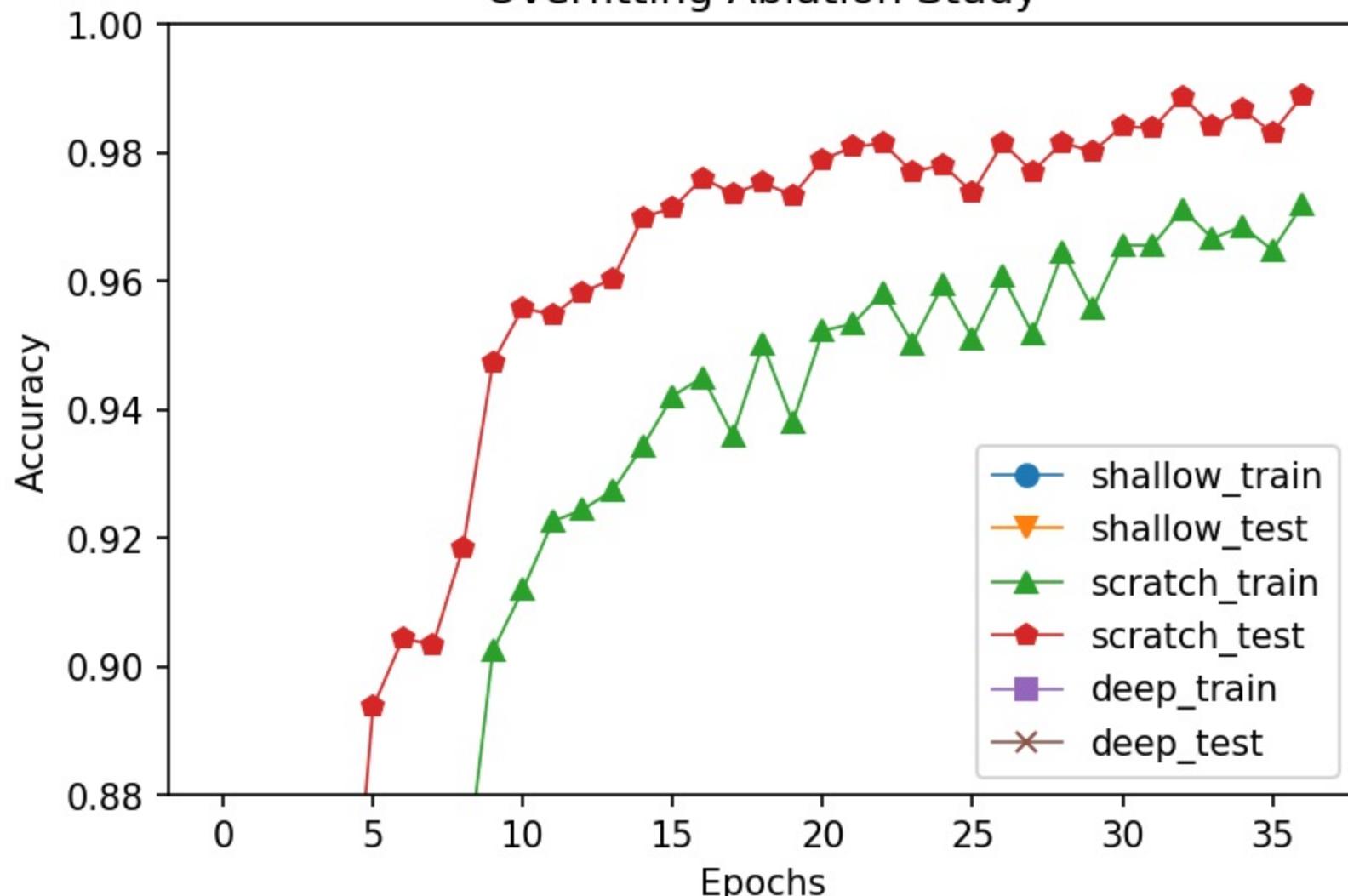
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	NaN	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 36 takes 317 seconds.

[38, 30] loss: 0.087

[38, 60] loss: 0.132

```
[38,    90] loss: 0.095
[38,   120] loss: 0.093
[38,   150] loss: 0.100
[38,   180] loss: 0.120
[38,   210] loss: 0.144
[38,   240] loss: 0.110
[38,   270] loss: 0.109
[38,   300] loss: 0.120
[38,   330] loss: 0.101
[38,   360] loss: 0.113
[38,   390] loss: 0.154
[38,   420] loss: 0.158
[38,   450] loss: 0.124
[38,   480] loss: 0.099
[38,   510] loss: 0.132
[38,   540] loss: 0.146
[38,   570] loss: 0.130
[38,   600] loss: 0.113
[38,   630] loss: 0.116
[38,   660] loss: 0.114
[38,   690] loss: 0.092
[38,   720] loss: 0.150
[38,   750] loss: 0.105
[38,   780] loss: 0.118
[38,   810] loss: 0.114
[38,   840] loss: 0.124
[38,   870] loss: 0.113
[38,   900] loss: 0.133
[38,   930] loss: 0.117
[38,   960] loss: 0.116
[38,   990] loss: 0.110
[38,  1020] loss: 0.112
[38,  1050] loss: 0.141
[38,  1080] loss: 0.132
[38,  1110] loss: 0.142
[38,  1140] loss: 0.155
[38,  1170] loss: 0.134
[38,  1200] loss: 0.079
[38,  1230] loss: 0.111
```

```
[38, 1260] loss: 0.093
[38, 1290] loss: 0.128
[38, 1320] loss: 0.098
[38, 1350] loss: 0.119
[38, 1380] loss: 0.146
[38, 1410] loss: 0.135
[38, 1440] loss: 0.142
[38, 1470] loss: 0.111
[38, 1500] loss: 0.124
[38, 1530] loss: 0.093
[38, 1560] loss: 0.111
[38, 1590] loss: 0.112
[38, 1620] loss: 0.077
[38, 1650] loss: 0.089
[38, 1680] loss: 0.135
[38, 1710] loss: 0.164
[38, 1740] loss: 0.128
[38, 1770] loss: 0.106
[38, 1800] loss: 0.124
[38, 1830] loss: 0.097
[38, 1860] loss: 0.115
[38, 1890] loss: 0.099
[38, 1920] loss: 0.169
[38, 1950] loss: 0.176
[38, 1980] loss: 0.102
[38, 2010] loss: 0.122
[38, 2040] loss: 0.129
[38, 2070] loss: 0.132
[38, 2100] loss: 0.090
[38, 2130] loss: 0.111
[38, 2160] loss: 0.082
[38, 2190] loss: 0.146
eval intermediate_models/epoch37_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.972192
epoch 37 accuracy on train set is: 0.9721919302071974
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	NaN	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988095

epoch 37 accuracy on test set is: 0.9880952380952381

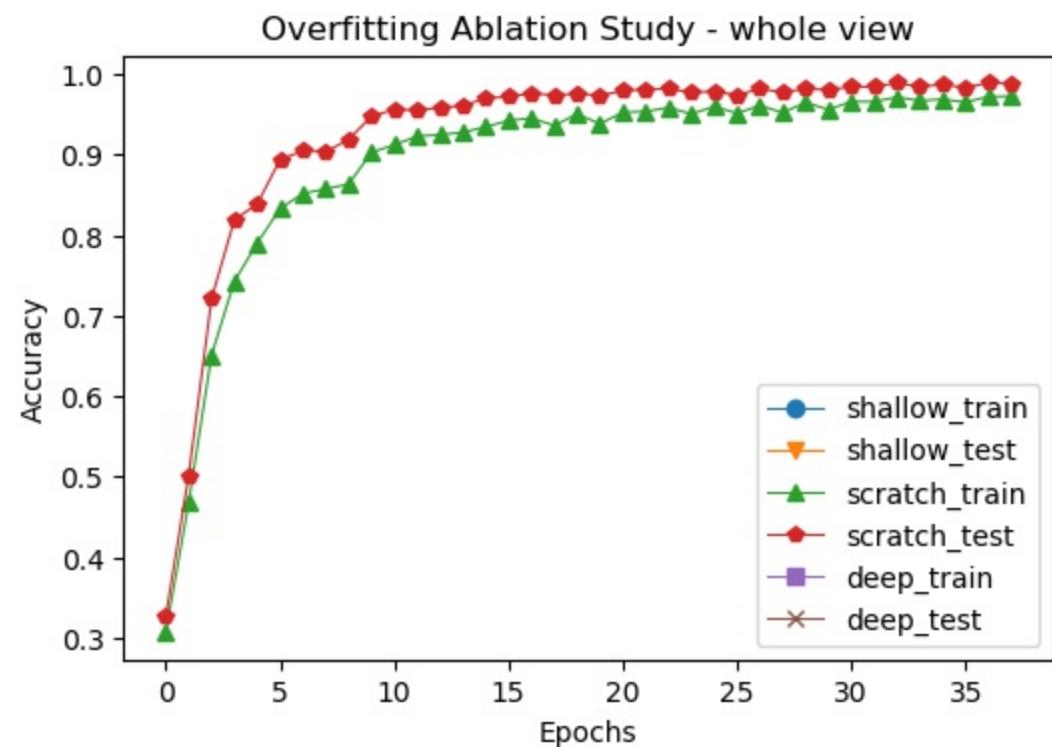
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

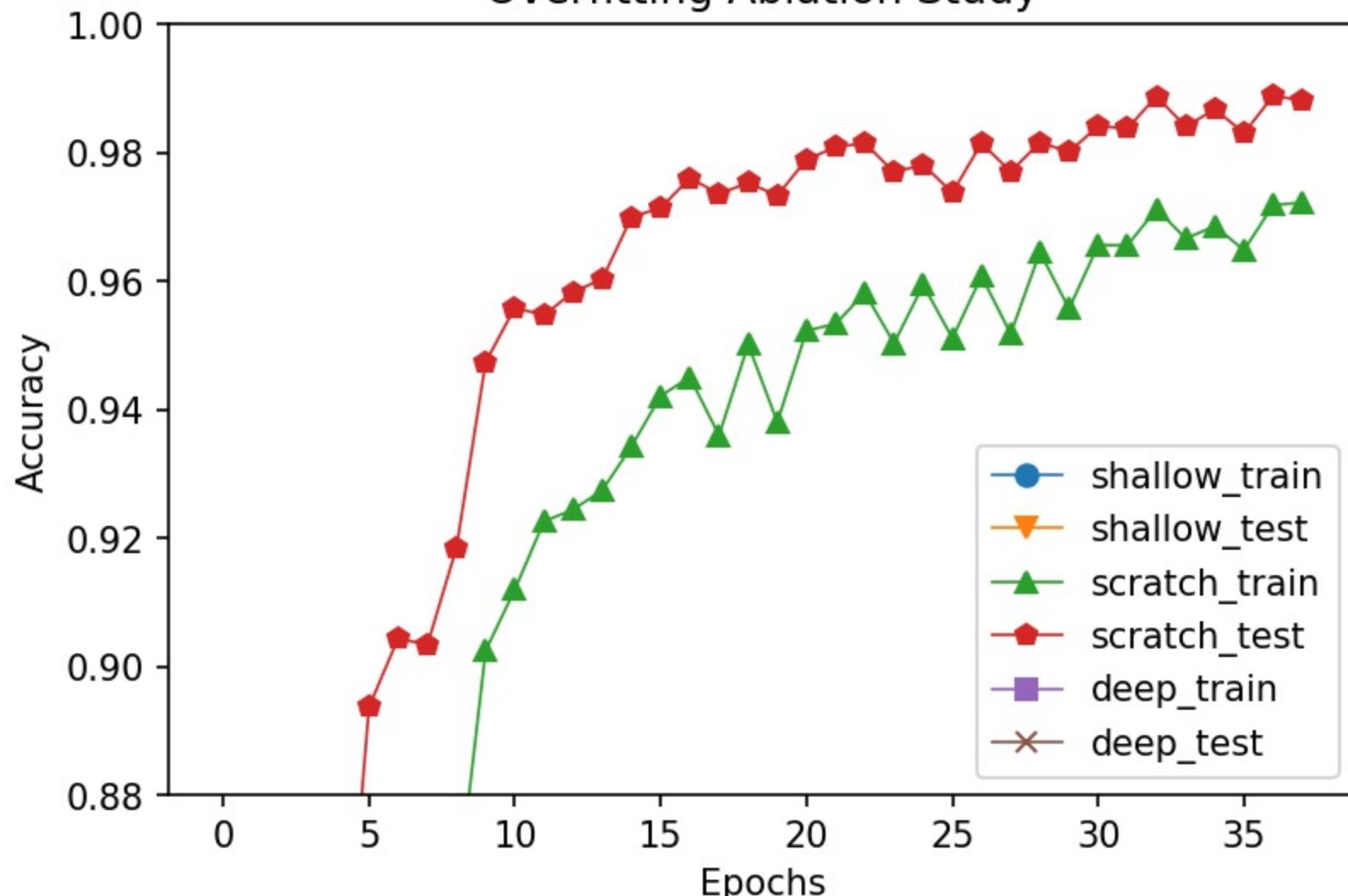
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	NaN	NaN	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 37 takes 317 seconds.

[39, 30] loss: 0.105

[39, 60] loss: 0.137

```
[39,    90] loss: 0.119
[39,   120] loss: 0.112
[39,   150] loss: 0.077
[39,   180] loss: 0.108
[39,   210] loss: 0.165
[39,   240] loss: 0.098
[39,   270] loss: 0.139
[39,   300] loss: 0.182
[39,   330] loss: 0.112
[39,   360] loss: 0.117
[39,   390] loss: 0.117
[39,   420] loss: 0.147
[39,   450] loss: 0.139
[39,   480] loss: 0.169
[39,   510] loss: 0.102
[39,   540] loss: 0.090
[39,   570] loss: 0.086
[39,   600] loss: 0.119
[39,   630] loss: 0.114
[39,   660] loss: 0.096
[39,   690] loss: 0.148
[39,   720] loss: 0.121
[39,   750] loss: 0.118
[39,   780] loss: 0.096
[39,   810] loss: 0.097
[39,   840] loss: 0.143
[39,   870] loss: 0.086
[39,   900] loss: 0.128
[39,   930] loss: 0.106
[39,   960] loss: 0.170
[39,   990] loss: 0.168
[39,  1020] loss: 0.109
[39,  1050] loss: 0.105
[39,  1080] loss: 0.113
[39,  1110] loss: 0.120
[39,  1140] loss: 0.098
[39,  1170] loss: 0.130
[39,  1200] loss: 0.074
[39,  1230] loss: 0.138
```

```
[39, 1260] loss: 0.182
[39, 1290] loss: 0.150
[39, 1320] loss: 0.088
[39, 1350] loss: 0.100
[39, 1380] loss: 0.137
[39, 1410] loss: 0.129
[39, 1440] loss: 0.102
[39, 1470] loss: 0.151
[39, 1500] loss: 0.117
[39, 1530] loss: 0.163
[39, 1560] loss: 0.170
[39, 1590] loss: 0.104
[39, 1620] loss: 0.127
[39, 1650] loss: 0.116
[39, 1680] loss: 0.069
[39, 1710] loss: 0.117
[39, 1740] loss: 0.158
[39, 1770] loss: 0.144
[39, 1800] loss: 0.113
[39, 1830] loss: 0.116
[39, 1860] loss: 0.085
[39, 1890] loss: 0.074
[39, 1920] loss: 0.114
[39, 1950] loss: 0.091
[39, 1980] loss: 0.236
[39, 2010] loss: 0.115
[39, 2040] loss: 0.118
[39, 2070] loss: 0.107
[39, 2100] loss: 0.109
[39, 2130] loss: 0.112
[39, 2160] loss: 0.109
[39, 2190] loss: 0.164
eval intermediate_models/epoch38_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.967671
epoch 38 accuracy on train set is: 0.96767084696474
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	NaN	NaN	NaN

39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.983915

epoch 38 accuracy on test set is: 0.9839149400218102

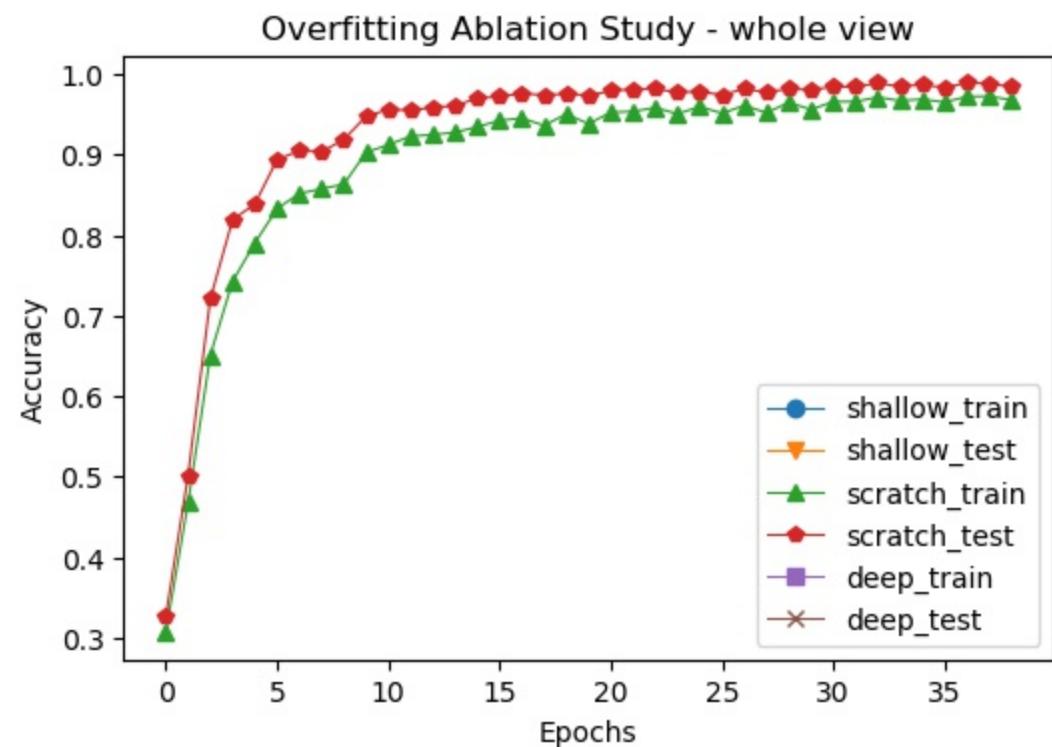
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

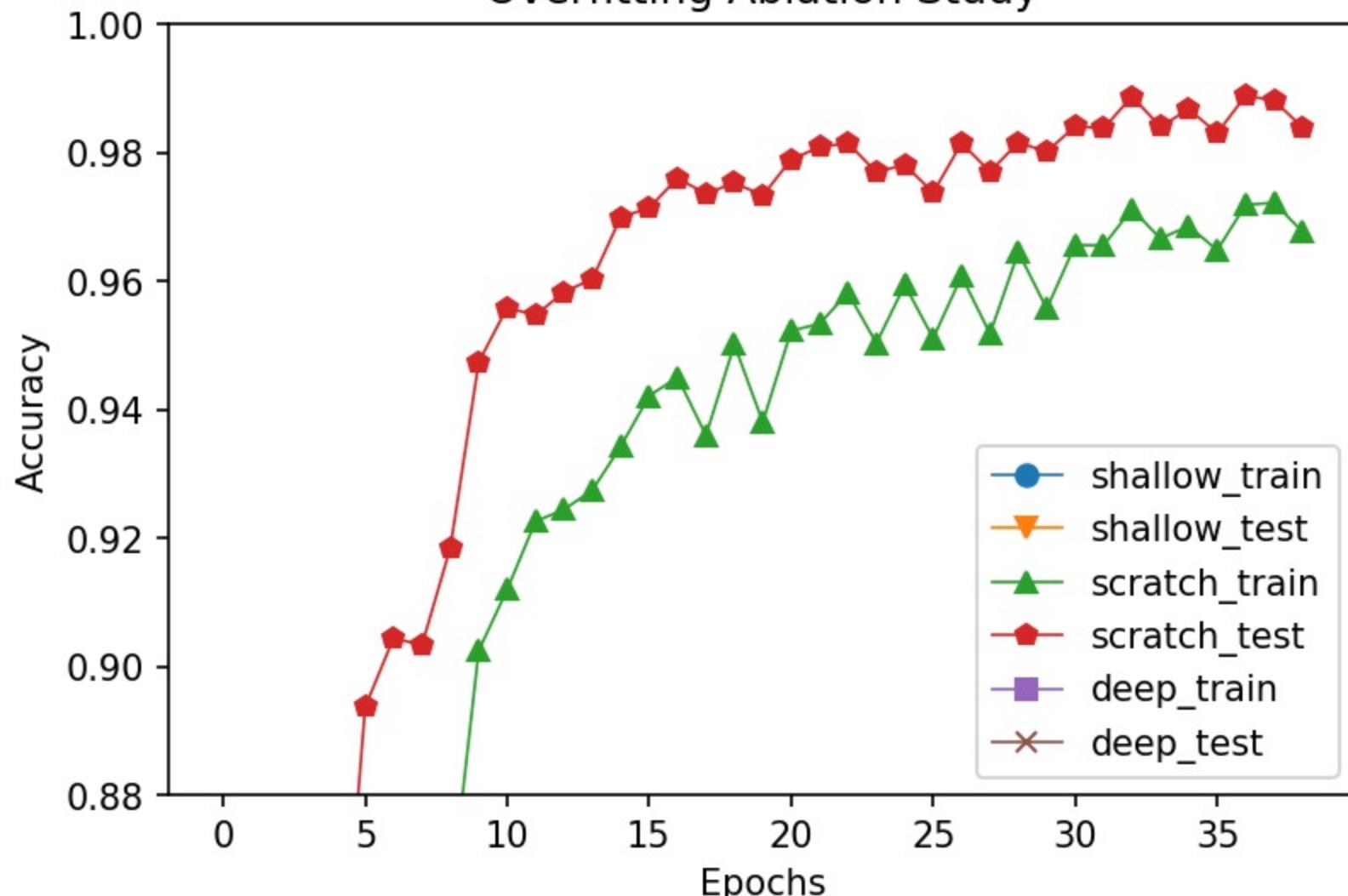
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	NaN	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 38 takes 317 seconds.

[40, 30] loss: 0.128

[40, 60] loss: 0.099

```
[40,    90] loss: 0.100
[40,   120] loss: 0.110
[40,   150] loss: 0.125
[40,   180] loss: 0.110
[40,   210] loss: 0.106
[40,   240] loss: 0.117
[40,   270] loss: 0.092
[40,   300] loss: 0.092
[40,   330] loss: 0.095
[40,   360] loss: 0.161
[40,   390] loss: 0.105
[40,   420] loss: 0.130
[40,   450] loss: 0.137
[40,   480] loss: 0.125
[40,   510] loss: 0.086
[40,   540] loss: 0.103
[40,   570] loss: 0.103
[40,   600] loss: 0.215
[40,   630] loss: 0.223
[40,   660] loss: 0.134
[40,   690] loss: 0.080
[40,   720] loss: 0.124
[40,   750] loss: 0.165
[40,   780] loss: 0.121
[40,   810] loss: 0.119
[40,   840] loss: 0.154
[40,   870] loss: 0.154
[40,   900] loss: 0.142
[40,   930] loss: 0.128
[40,   960] loss: 0.115
[40,   990] loss: 0.097
[40,  1020] loss: 0.105
[40,  1050] loss: 0.090
[40,  1080] loss: 0.120
[40,  1110] loss: 0.074
[40,  1140] loss: 0.139
[40,  1170] loss: 0.140
[40,  1200] loss: 0.133
[40,  1230] loss: 0.135
```

```
[40, 1260] loss: 0.111
[40, 1290] loss: 0.144
[40, 1320] loss: 0.112
[40, 1350] loss: 0.122
[40, 1380] loss: 0.111
[40, 1410] loss: 0.130
[40, 1440] loss: 0.119
[40, 1470] loss: 0.118
[40, 1500] loss: 0.110
[40, 1530] loss: 0.117
[40, 1560] loss: 0.140
[40, 1590] loss: 0.086
[40, 1620] loss: 0.139
[40, 1650] loss: 0.116
[40, 1680] loss: 0.098
[40, 1710] loss: 0.108
[40, 1740] loss: 0.121
[40, 1770] loss: 0.098
[40, 1800] loss: 0.128
[40, 1830] loss: 0.123
[40, 1860] loss: 0.088
[40, 1890] loss: 0.145
[40, 1920] loss: 0.130
[40, 1950] loss: 0.129
[40, 1980] loss: 0.134
[40, 2010] loss: 0.131
[40, 2040] loss: 0.134
[40, 2070] loss: 0.126
[40, 2100] loss: 0.141
[40, 2130] loss: 0.148
[40, 2160] loss: 0.122
[40, 2190] loss: 0.118
eval intermediate_models/epoch39_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.975168
epoch 39 accuracy on train set is: 0.9751681206833879
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	NaN	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.989095

epoch 39 accuracy on test set is: 0.9890948745910578

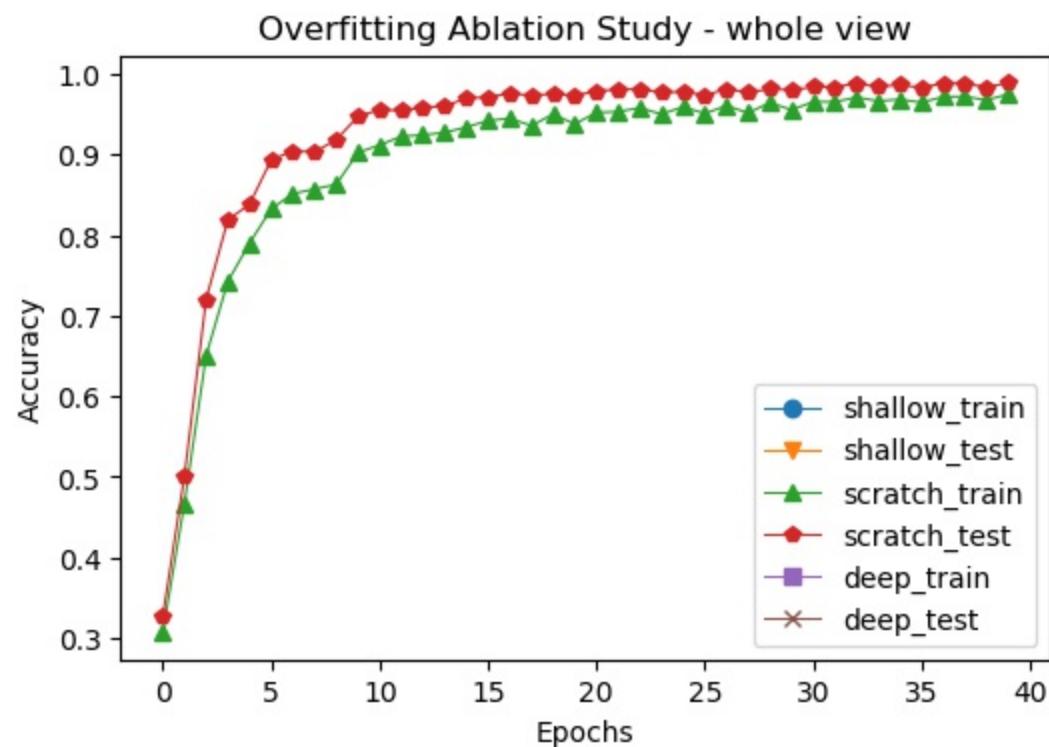
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

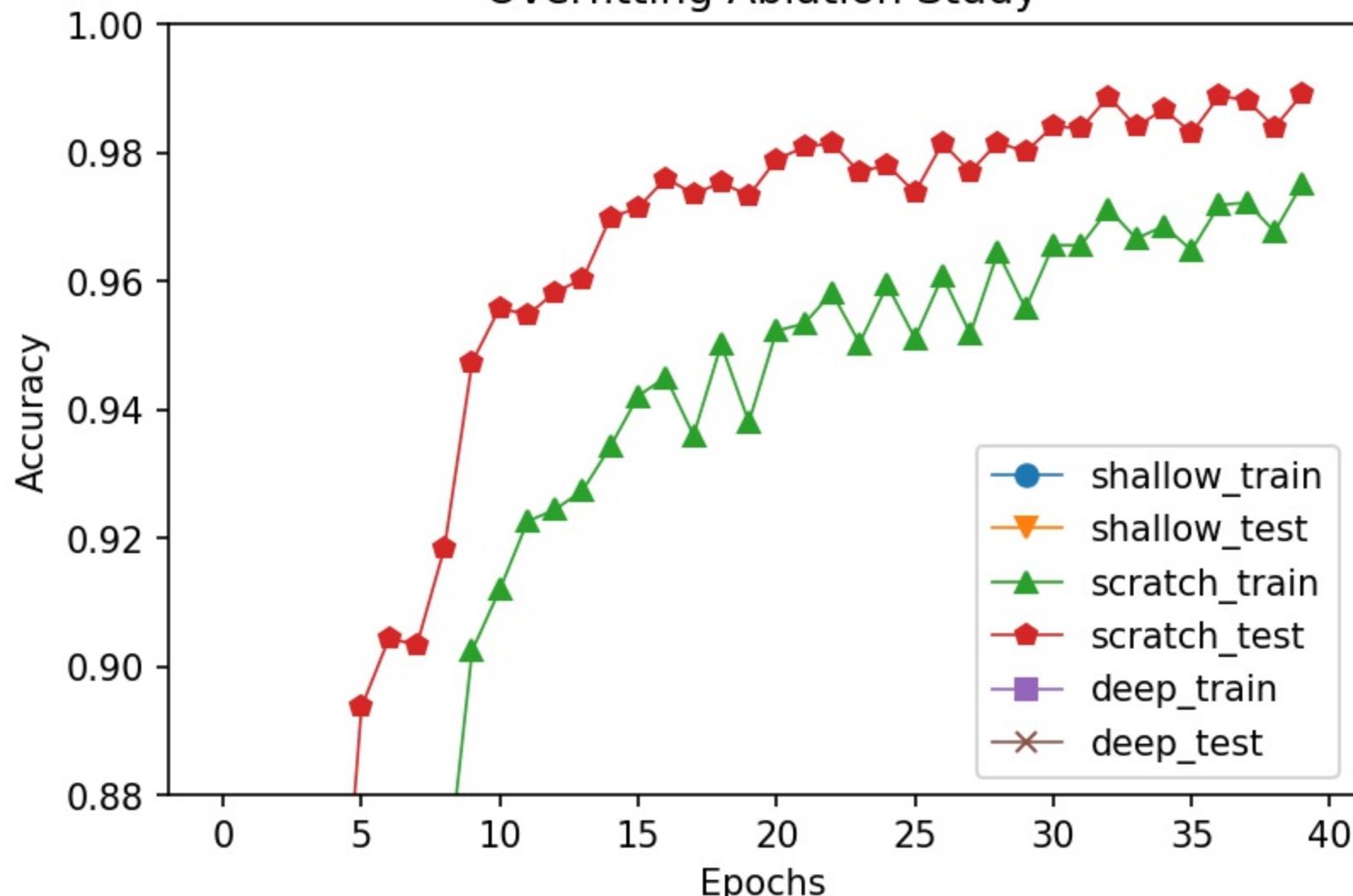
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 39 takes 317 seconds.

[41, 30] loss: 0.141

[41, 60] loss: 0.089

```
[41,    90] loss: 0.114
[41,   120] loss: 0.099
[41,   150] loss: 0.123
[41,   180] loss: 0.092
[41,   210] loss: 0.100
[41,   240] loss: 0.102
[41,   270] loss: 0.136
[41,   300] loss: 0.138
[41,   330] loss: 0.131
[41,   360] loss: 0.121
[41,   390] loss: 0.102
[41,   420] loss: 0.135
[41,   450] loss: 0.097
[41,   480] loss: 0.121
[41,   510] loss: 0.115
[41,   540] loss: 0.123
[41,   570] loss: 0.091
[41,   600] loss: 0.092
[41,   630] loss: 0.166
[41,   660] loss: 0.108
[41,   690] loss: 0.126
[41,   720] loss: 0.107
[41,   750] loss: 0.112
[41,   780] loss: 0.083
[41,   810] loss: 0.093
[41,   840] loss: 0.146
[41,   870] loss: 0.124
[41,   900] loss: 0.119
[41,   930] loss: 0.120
[41,   960] loss: 0.126
[41,   990] loss: 0.147
[41,  1020] loss: 0.138
[41,  1050] loss: 0.137
[41,  1080] loss: 0.117
[41,  1110] loss: 0.092
[41,  1140] loss: 0.093
[41,  1170] loss: 0.129
[41,  1200] loss: 0.090
[41,  1230] loss: 0.119
```

```
[41, 1260] loss: 0.115
[41, 1290] loss: 0.098
[41, 1320] loss: 0.147
[41, 1350] loss: 0.116
[41, 1380] loss: 0.131
[41, 1410] loss: 0.108
[41, 1440] loss: 0.136
[41, 1470] loss: 0.089
[41, 1500] loss: 0.106
[41, 1530] loss: 0.132
[41, 1560] loss: 0.090
[41, 1590] loss: 0.107
[41, 1620] loss: 0.119
[41, 1650] loss: 0.112
[41, 1680] loss: 0.127
[41, 1710] loss: 0.105
[41, 1740] loss: 0.126
[41, 1770] loss: 0.113
[41, 1800] loss: 0.122
[41, 1830] loss: 0.115
[41, 1860] loss: 0.072
[41, 1890] loss: 0.087
[41, 1920] loss: 0.094
[41, 1950] loss: 0.088
[41, 1980] loss: 0.084
[41, 2010] loss: 0.130
[41, 2040] loss: 0.129
[41, 2070] loss: 0.107
[41, 2100] loss: 0.103
[41, 2130] loss: 0.098
[41, 2160] loss: 0.136
[41, 2190] loss: 0.131
eval intermediate_models/epoch40_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.975941
epoch 40 accuracy on train set is: 0.9759405670665212
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988822

epoch 40 accuracy on test set is: 0.9888222464558343

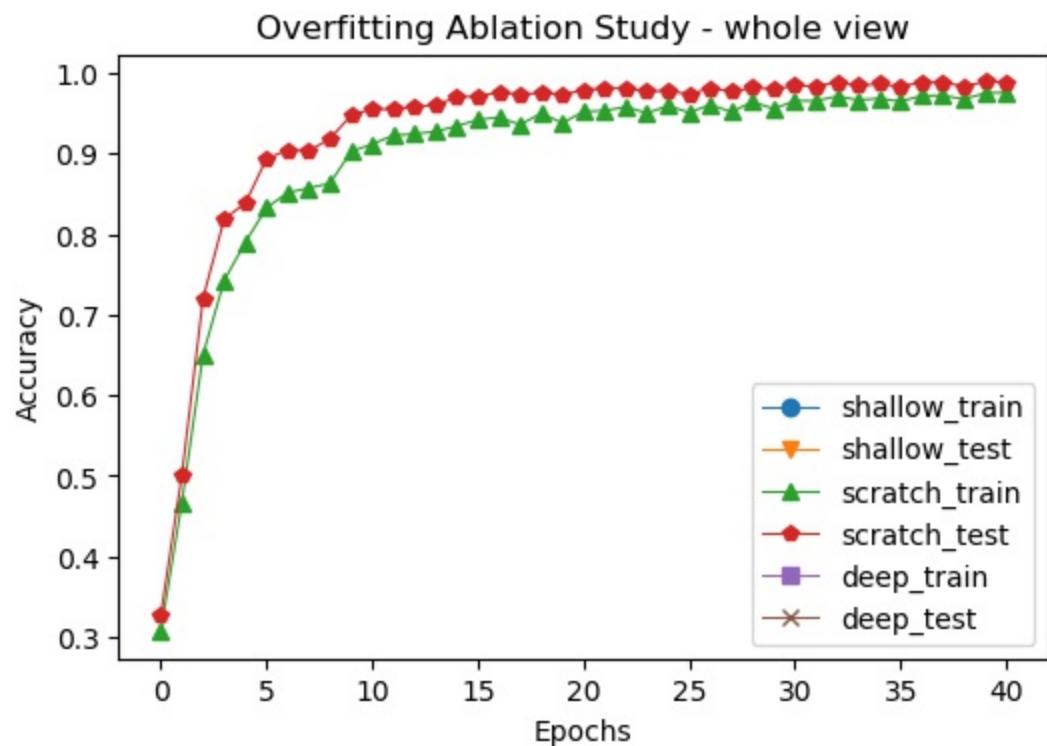
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

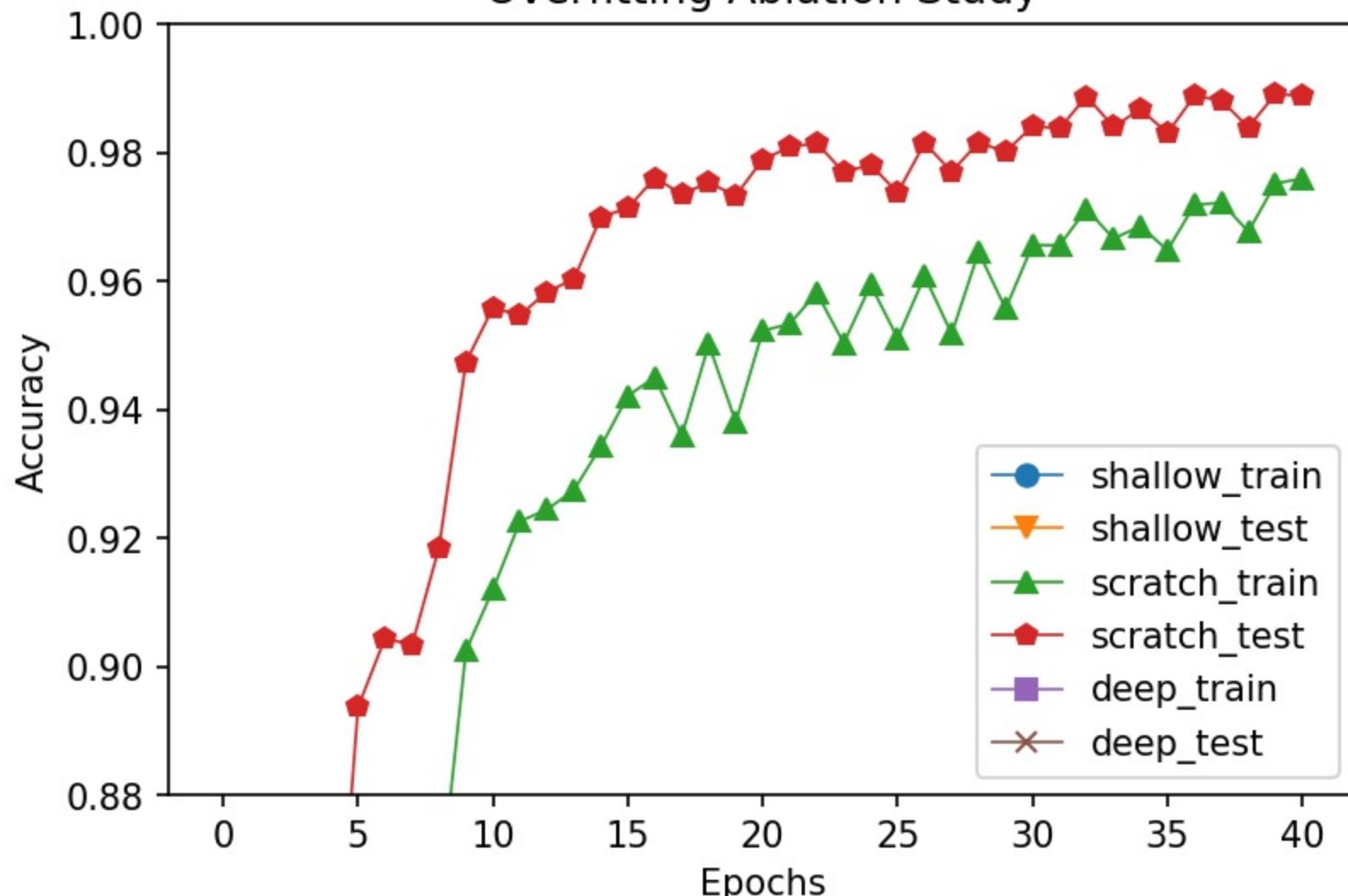
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	NaN	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 40 takes 317 seconds.

[42, 30] loss: 0.103

[42, 60] loss: 0.117

```
[42,    90] loss: 0.087
[42,   120] loss: 0.099
[42,   150] loss: 0.164
[42,   180] loss: 0.115
[42,   210] loss: 0.097
[42,   240] loss: 0.100
[42,   270] loss: 0.112
[42,   300] loss: 0.097
[42,   330] loss: 0.141
[42,   360] loss: 0.148
[42,   390] loss: 0.109
[42,   420] loss: 0.108
[42,   450] loss: 0.124
[42,   480] loss: 0.107
[42,   510] loss: 0.099
[42,   540] loss: 0.135
[42,   570] loss: 0.086
[42,   600] loss: 0.144
[42,   630] loss: 0.167
[42,   660] loss: 0.083
[42,   690] loss: 0.072
[42,   720] loss: 0.092
[42,   750] loss: 0.097
[42,   780] loss: 0.155
[42,   810] loss: 0.143
[42,   840] loss: 0.099
[42,   870] loss: 0.092
[42,   900] loss: 0.105
[42,   930] loss: 0.128
[42,   960] loss: 0.139
[42,   990] loss: 0.119
[42,  1020] loss: 0.109
[42,  1050] loss: 0.163
[42,  1080] loss: 0.109
[42,  1110] loss: 0.120
[42,  1140] loss: 0.163
[42,  1170] loss: 0.124
[42,  1200] loss: 0.139
[42,  1230] loss: 0.109
```

```
[42, 1260] loss: 0.128
[42, 1290] loss: 0.087
[42, 1320] loss: 0.133
[42, 1350] loss: 0.086
[42, 1380] loss: 0.123
[42, 1410] loss: 0.112
[42, 1440] loss: 0.133
[42, 1470] loss: 0.104
[42, 1500] loss: 0.132
[42, 1530] loss: 0.121
[42, 1560] loss: 0.102
[42, 1590] loss: 0.109
[42, 1620] loss: 0.112
[42, 1650] loss: 0.123
[42, 1680] loss: 0.094
[42, 1710] loss: 0.089
[42, 1740] loss: 0.113
[42, 1770] loss: 0.164
[42, 1800] loss: 0.145
[42, 1830] loss: 0.072
[42, 1860] loss: 0.179
[42, 1890] loss: 0.097
[42, 1920] loss: 0.116
[42, 1950] loss: 0.133
[42, 1980] loss: 0.098
[42, 2010] loss: 0.103
[42, 2040] loss: 0.156
[42, 2070] loss: 0.089
[42, 2100] loss: 0.139
[42, 2130] loss: 0.112
[42, 2160] loss: 0.176
[42, 2190] loss: 0.115
eval intermediate_models/epoch41_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.968125
epoch 41 accuracy on train set is: 0.9681252271901127
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	NaN	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.986278

epoch 41 accuracy on test set is: 0.9862777171937477

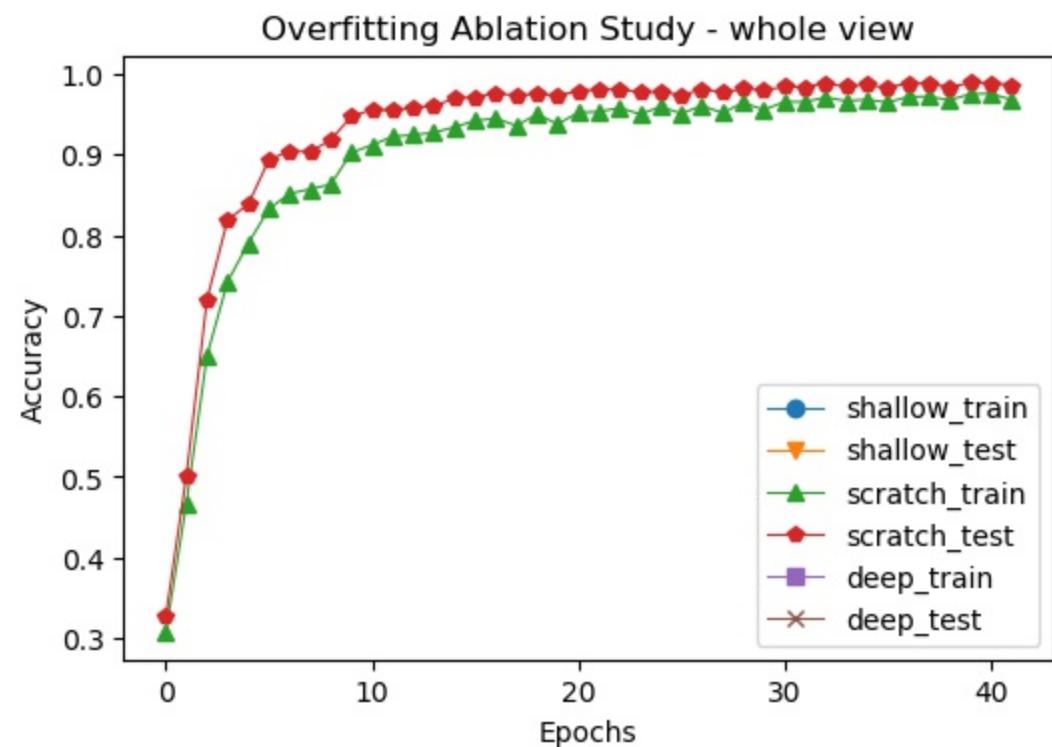
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

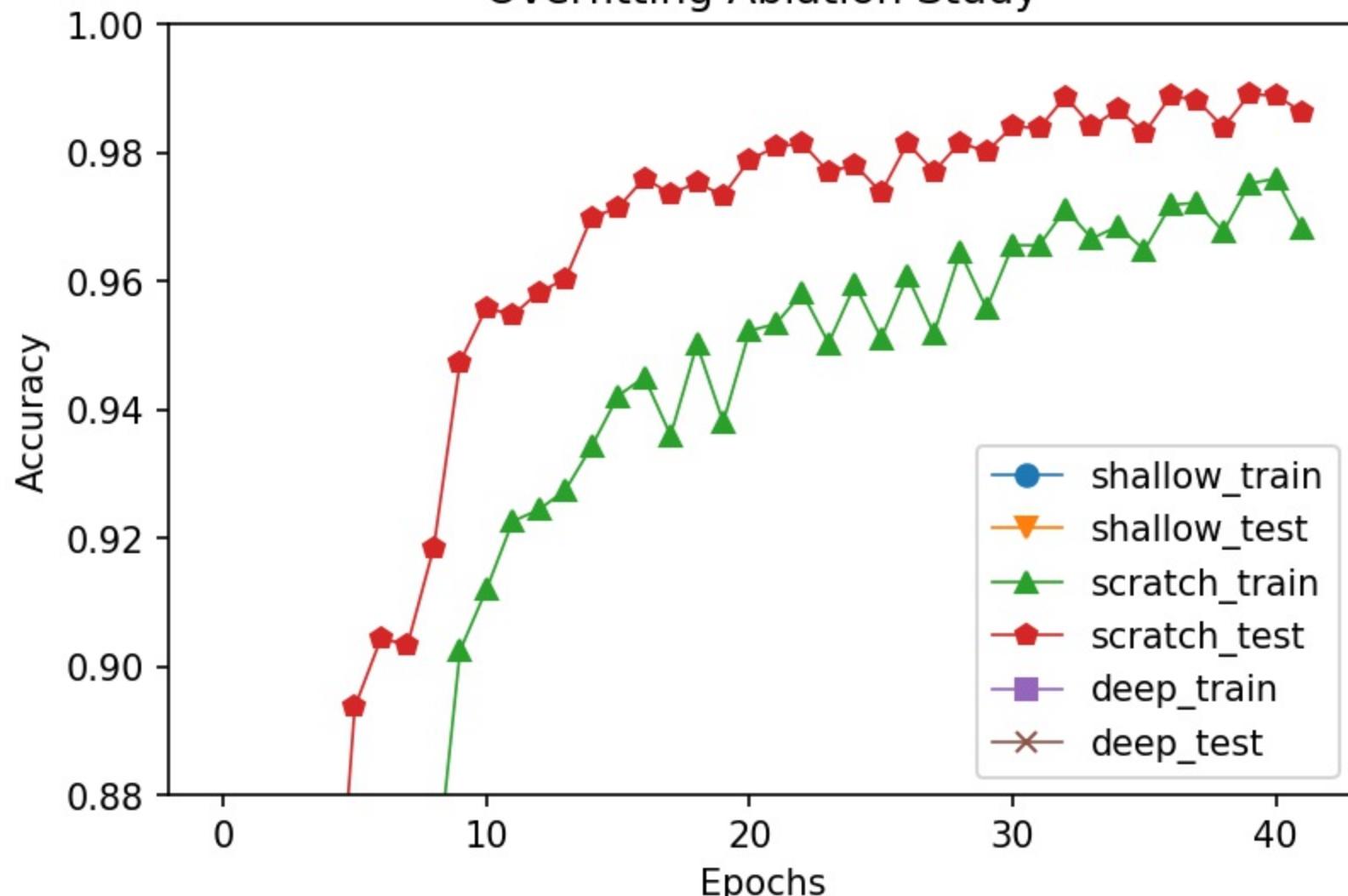
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	NaN	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 41 takes 316 seconds.

[43, 30] loss: 0.097

[43, 60] loss: 0.109

```
[43,    90] loss: 0.086
[43,   120] loss: 0.133
[43,   150] loss: 0.127
[43,   180] loss: 0.068
[43,   210] loss: 0.203
[43,   240] loss: 0.109
[43,   270] loss: 0.080
[43,   300] loss: 0.103
[43,   330] loss: 0.086
[43,   360] loss: 0.131
[43,   390] loss: 0.117
[43,   420] loss: 0.152
[43,   450] loss: 0.124
[43,   480] loss: 0.131
[43,   510] loss: 0.149
[43,   540] loss: 0.091
[43,   570] loss: 0.087
[43,   600] loss: 0.146
[43,   630] loss: 0.114
[43,   660] loss: 0.119
[43,   690] loss: 0.085
[43,   720] loss: 0.107
[43,   750] loss: 0.125
[43,   780] loss: 0.086
[43,   810] loss: 0.101
[43,   840] loss: 0.126
[43,   870] loss: 0.118
[43,   900] loss: 0.127
[43,   930] loss: 0.091
[43,   960] loss: 0.119
[43,   990] loss: 0.090
[43,  1020] loss: 0.101
[43,  1050] loss: 0.093
[43,  1080] loss: 0.108
[43,  1110] loss: 0.095
[43,  1140] loss: 0.095
[43,  1170] loss: 0.094
[43,  1200] loss: 0.107
[43,  1230] loss: 0.105
```

```
[43, 1260] loss: 0.108
[43, 1290] loss: 0.134
[43, 1320] loss: 0.109
[43, 1350] loss: 0.125
[43, 1380] loss: 0.094
[43, 1410] loss: 0.093
[43, 1440] loss: 0.102
[43, 1470] loss: 0.134
[43, 1500] loss: 0.094
[43, 1530] loss: 0.140
[43, 1560] loss: 0.105
[43, 1590] loss: 0.124
[43, 1620] loss: 0.116
[43, 1650] loss: 0.135
[43, 1680] loss: 0.116
[43, 1710] loss: 0.100
[43, 1740] loss: 0.112
[43, 1770] loss: 0.111
[43, 1800] loss: 0.178
[43, 1830] loss: 0.178
[43, 1860] loss: 0.167
[43, 1890] loss: 0.110
[43, 1920] loss: 0.083
[43, 1950] loss: 0.103
[43, 1980] loss: 0.107
[43, 2010] loss: 0.113
[43, 2040] loss: 0.095
[43, 2070] loss: 0.118
[43, 2100] loss: 0.124
[43, 2130] loss: 0.098
[43, 2160] loss: 0.113
[43, 2190] loss: 0.096
eval intermediate_models/epoch42_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.968307
epoch 42 accuracy on train set is: 0.9683069792802618
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	NaN	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.983733

epoch 42 accuracy on test set is: 0.9837331879316612

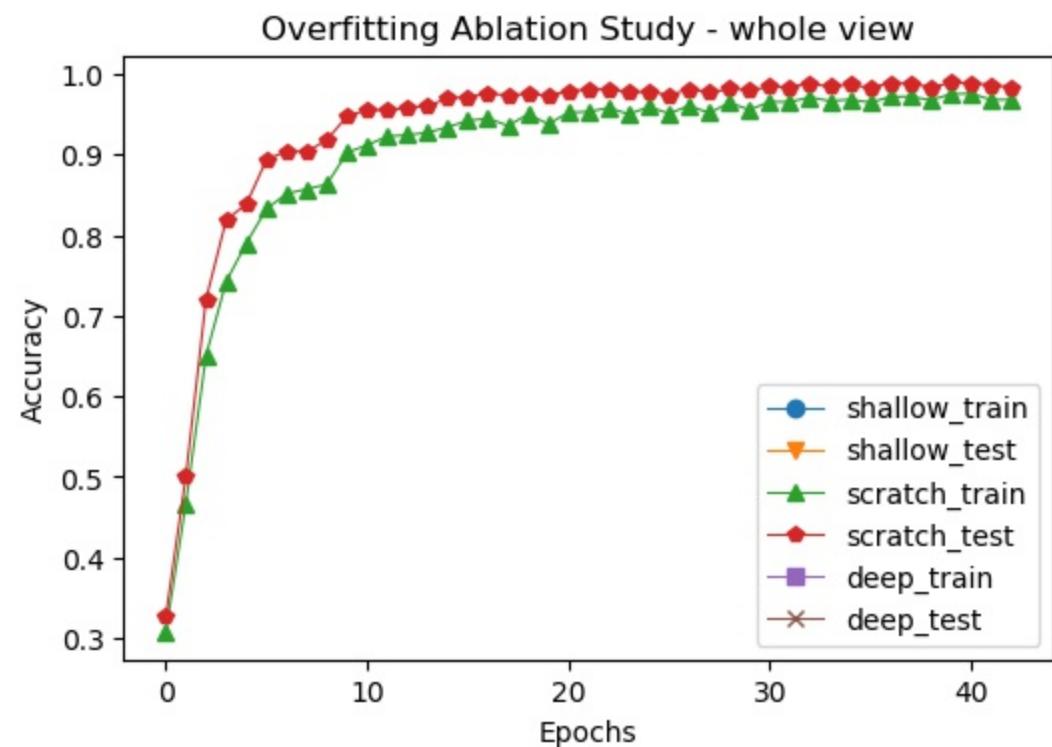
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

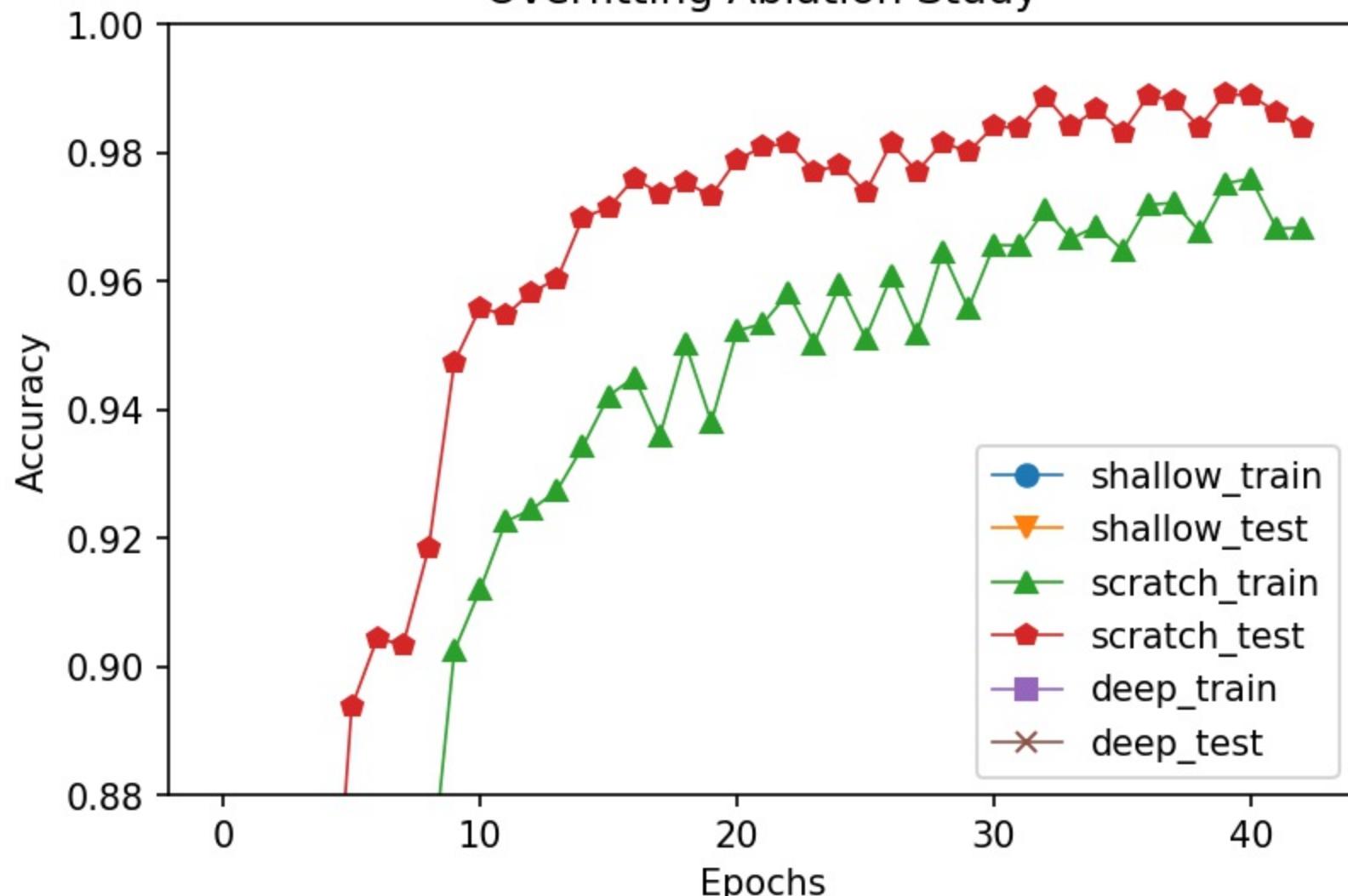
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	NaN	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 42 takes 316 seconds.

[44, 30] loss: 0.120

[44, 60] loss: 0.113

```
[44,    90] loss: 0.120
[44,   120] loss: 0.143
[44,   150] loss: 0.103
[44,   180] loss: 0.111
[44,   210] loss: 0.133
[44,   240] loss: 0.115
[44,   270] loss: 0.096
[44,   300] loss: 0.089
[44,   330] loss: 0.104
[44,   360] loss: 0.127
[44,   390] loss: 0.110
[44,   420] loss: 0.101
[44,   450] loss: 0.101
[44,   480] loss: 0.094
[44,   510] loss: 0.145
[44,   540] loss: 0.107
[44,   570] loss: 0.073
[44,   600] loss: 0.080
[44,   630] loss: 0.094
[44,   660] loss: 0.099
[44,   690] loss: 0.136
[44,   720] loss: 0.139
[44,   750] loss: 0.120
[44,   780] loss: 0.117
[44,   810] loss: 0.160
[44,   840] loss: 0.110
[44,   870] loss: 0.107
[44,   900] loss: 0.125
[44,   930] loss: 0.112
[44,   960] loss: 0.135
[44,   990] loss: 0.117
[44,  1020] loss: 0.094
[44,  1050] loss: 0.102
[44,  1080] loss: 0.105
[44,  1110] loss: 0.116
[44,  1140] loss: 0.099
[44,  1170] loss: 0.157
[44,  1200] loss: 0.136
[44,  1230] loss: 0.120
```

```
[44, 1260] loss: 0.113
[44, 1290] loss: 0.114
[44, 1320] loss: 0.125
[44, 1350] loss: 0.070
[44, 1380] loss: 0.105
[44, 1410] loss: 0.122
[44, 1440] loss: 0.092
[44, 1470] loss: 0.129
[44, 1500] loss: 0.078
[44, 1530] loss: 0.096
[44, 1560] loss: 0.125
[44, 1590] loss: 0.109
[44, 1620] loss: 0.108
[44, 1650] loss: 0.074
[44, 1680] loss: 0.128
[44, 1710] loss: 0.149
[44, 1740] loss: 0.106
[44, 1770] loss: 0.090
[44, 1800] loss: 0.072
[44, 1830] loss: 0.121
[44, 1860] loss: 0.123
[44, 1890] loss: 0.124
[44, 1920] loss: 0.094
[44, 1950] loss: 0.118
[44, 1980] loss: 0.070
[44, 2010] loss: 0.118
[44, 2040] loss: 0.103
[44, 2070] loss: 0.120
[44, 2100] loss: 0.126
[44, 2130] loss: 0.097
[44, 2160] loss: 0.108
[44, 2190] loss: 0.108
eval intermediate_models/epoch43_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.972896
epoch 43 accuracy on train set is: 0.9728962195565249
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	NaN	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988550

epoch 43 accuracy on test set is: 0.9885496183206107

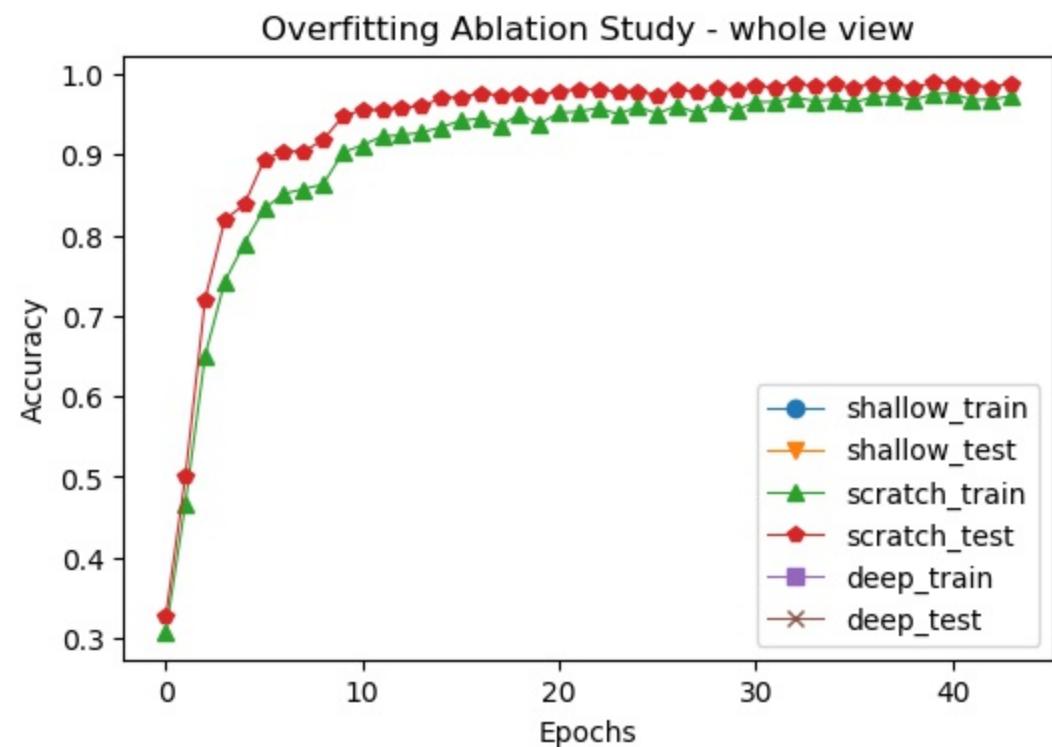
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

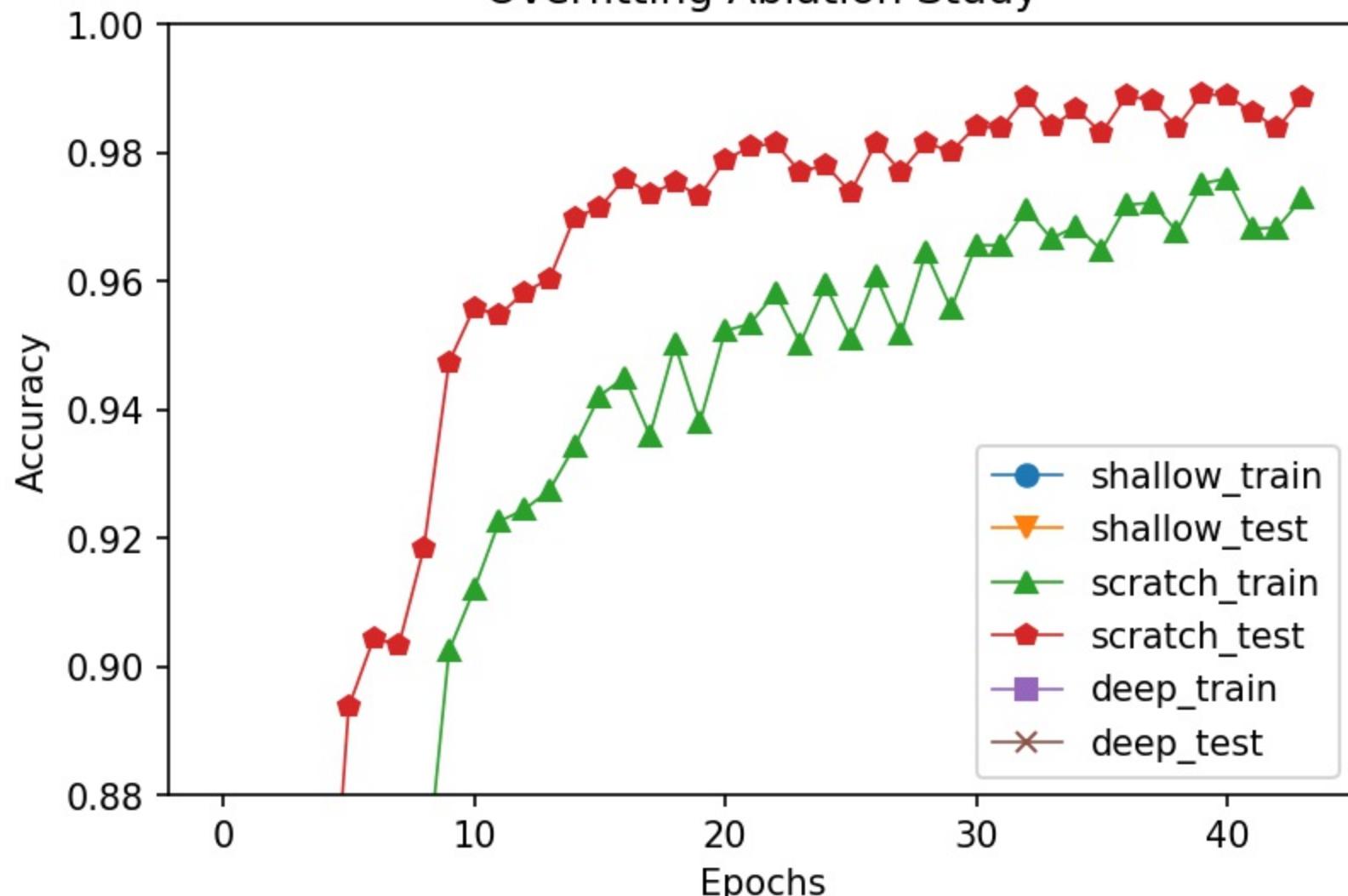
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	NaN	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 43 takes 316 seconds.

[45, 30] loss: 0.106

[45, 60] loss: 0.094

```
[45,    90] loss: 0.124
[45,   120] loss: 0.165
[45,   150] loss: 0.101
[45,   180] loss: 0.116
[45,   210] loss: 0.096
[45,   240] loss: 0.112
[45,   270] loss: 0.129
[45,   300] loss: 0.146
[45,   330] loss: 0.097
[45,   360] loss: 0.088
[45,   390] loss: 0.104
[45,   420] loss: 0.077
[45,   450] loss: 0.097
[45,   480] loss: 0.072
[45,   510] loss: 0.098
[45,   540] loss: 0.109
[45,   570] loss: 0.128
[45,   600] loss: 0.136
[45,   630] loss: 0.122
[45,   660] loss: 0.088
[45,   690] loss: 0.118
[45,   720] loss: 0.117
[45,   750] loss: 0.085
[45,   780] loss: 0.115
[45,   810] loss: 0.097
[45,   840] loss: 0.116
[45,   870] loss: 0.111
[45,   900] loss: 0.115
[45,   930] loss: 0.087
[45,   960] loss: 0.135
[45,   990] loss: 0.133
[45,  1020] loss: 0.133
[45,  1050] loss: 0.095
[45,  1080] loss: 0.061
[45,  1110] loss: 0.121
[45,  1140] loss: 0.128
[45,  1170] loss: 0.117
[45,  1200] loss: 0.102
[45,  1230] loss: 0.086
```

```
[45, 1260] loss: 0.079
[45, 1290] loss: 0.080
[45, 1320] loss: 0.106
[45, 1350] loss: 0.134
[45, 1380] loss: 0.117
[45, 1410] loss: 0.074
[45, 1440] loss: 0.139
[45, 1470] loss: 0.139
[45, 1500] loss: 0.104
[45, 1530] loss: 0.148
[45, 1560] loss: 0.103
[45, 1590] loss: 0.088
[45, 1620] loss: 0.111
[45, 1650] loss: 0.108
[45, 1680] loss: 0.094
[45, 1710] loss: 0.081
[45, 1740] loss: 0.109
[45, 1770] loss: 0.084
[45, 1800] loss: 0.125
[45, 1830] loss: 0.084
[45, 1860] loss: 0.084
[45, 1890] loss: 0.127
[45, 1920] loss: 0.103
[45, 1950] loss: 0.066
[45, 1980] loss: 0.123
[45, 2010] loss: 0.095
[45, 2040] loss: 0.137
[45, 2070] loss: 0.118
[45, 2100] loss: 0.088
[45, 2130] loss: 0.086
[45, 2160] loss: 0.096
[45, 2190] loss: 0.096
eval intermediate_models/epoch44_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.974895
epoch 44 accuracy on train set is: 0.9748954925481643
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	NaN	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990731

epoch 44 accuracy on test set is: 0.9907306434023991

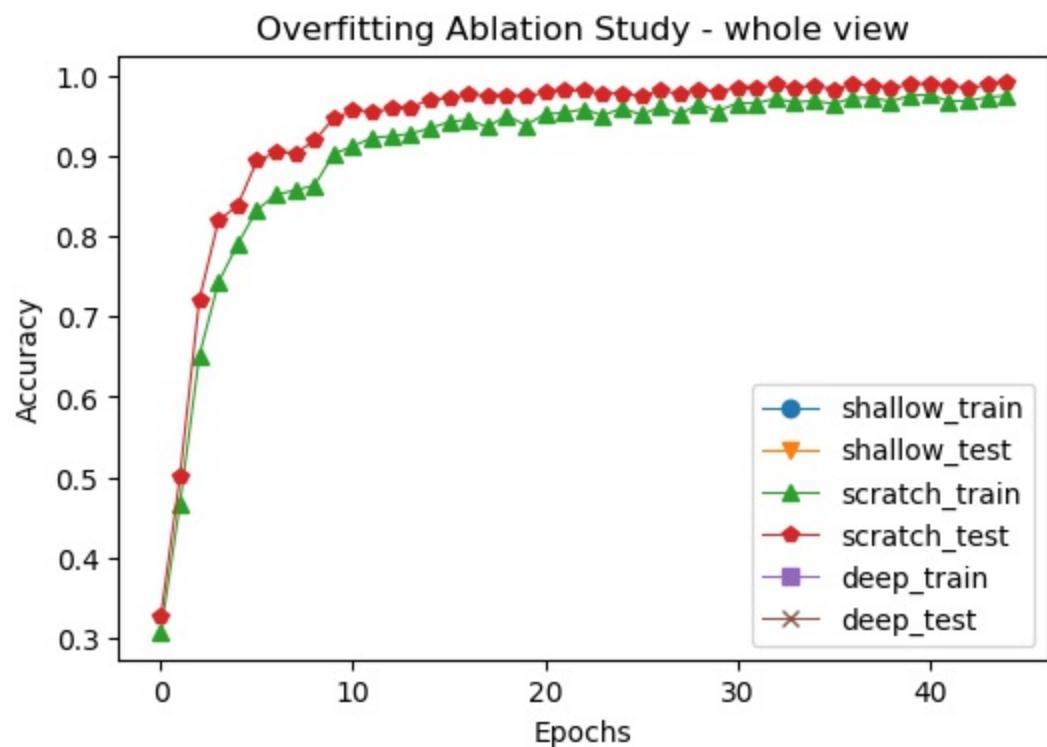
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

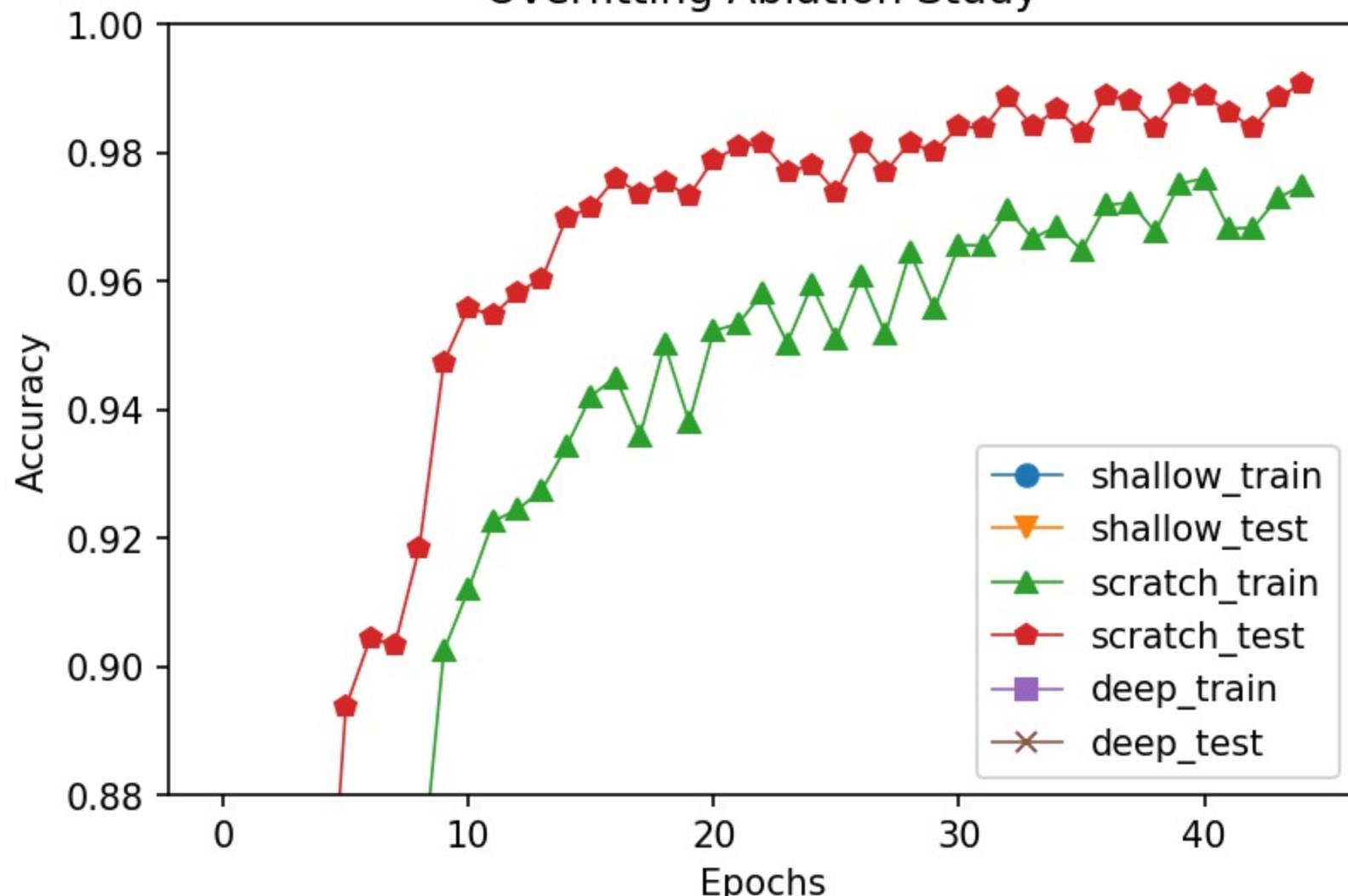
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	NaN	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 44 takes 316 seconds.

[46, 30] loss: 0.098

[46, 60] loss: 0.085

```
[46,    90] loss: 0.095
[46,   120] loss: 0.129
[46,   150] loss: 0.092
[46,   180] loss: 0.103
[46,   210] loss: 0.118
[46,   240] loss: 0.088
[46,   270] loss: 0.124
[46,   300] loss: 0.115
[46,   330] loss: 0.100
[46,   360] loss: 0.102
[46,   390] loss: 0.128
[46,   420] loss: 0.102
[46,   450] loss: 0.097
[46,   480] loss: 0.095
[46,   510] loss: 0.111
[46,   540] loss: 0.128
[46,   570] loss: 0.090
[46,   600] loss: 0.079
[46,   630] loss: 0.124
[46,   660] loss: 0.091
[46,   690] loss: 0.087
[46,   720] loss: 0.154
[46,   750] loss: 0.076
[46,   780] loss: 0.095
[46,   810] loss: 0.114
[46,   840] loss: 0.111
[46,   870] loss: 0.125
[46,   900] loss: 0.101
[46,   930] loss: 0.102
[46,   960] loss: 0.072
[46,   990] loss: 0.097
[46,  1020] loss: 0.110
[46,  1050] loss: 0.077
[46,  1080] loss: 0.084
[46,  1110] loss: 0.125
[46,  1140] loss: 0.113
[46,  1170] loss: 0.119
[46,  1200] loss: 0.115
[46,  1230] loss: 0.142
```

```
[46, 1260] loss: 0.112
[46, 1290] loss: 0.105
[46, 1320] loss: 0.115
[46, 1350] loss: 0.156
[46, 1380] loss: 0.115
[46, 1410] loss: 0.137
[46, 1440] loss: 0.115
[46, 1470] loss: 0.077
[46, 1500] loss: 0.099
[46, 1530] loss: 0.172
[46, 1560] loss: 0.173
[46, 1590] loss: 0.082
[46, 1620] loss: 0.123
[46, 1650] loss: 0.074
[46, 1680] loss: 0.102
[46, 1710] loss: 0.100
[46, 1740] loss: 0.114
[46, 1770] loss: 0.097
[46, 1800] loss: 0.134
[46, 1830] loss: 0.122
[46, 1860] loss: 0.100
[46, 1890] loss: 0.104
[46, 1920] loss: 0.136
[46, 1950] loss: 0.110
[46, 1980] loss: 0.104
[46, 2010] loss: 0.095
[46, 2040] loss: 0.095
[46, 2070] loss: 0.058
[46, 2100] loss: 0.123
[46, 2130] loss: 0.095
[46, 2160] loss: 0.144
[46, 2190] loss: 0.127
eval intermediate_models/epoch45_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.971669
epoch 45 accuracy on train set is: 0.9716693929480189
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	NaN	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.986823

epoch 45 accuracy on test set is: 0.9868229734641948

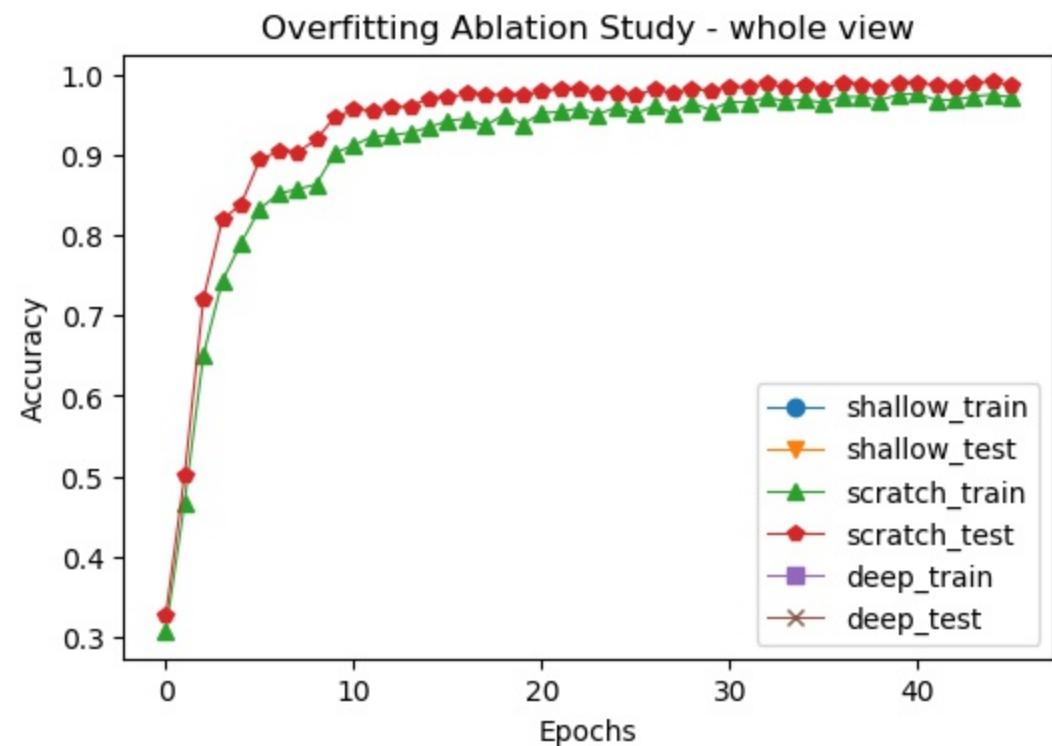
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

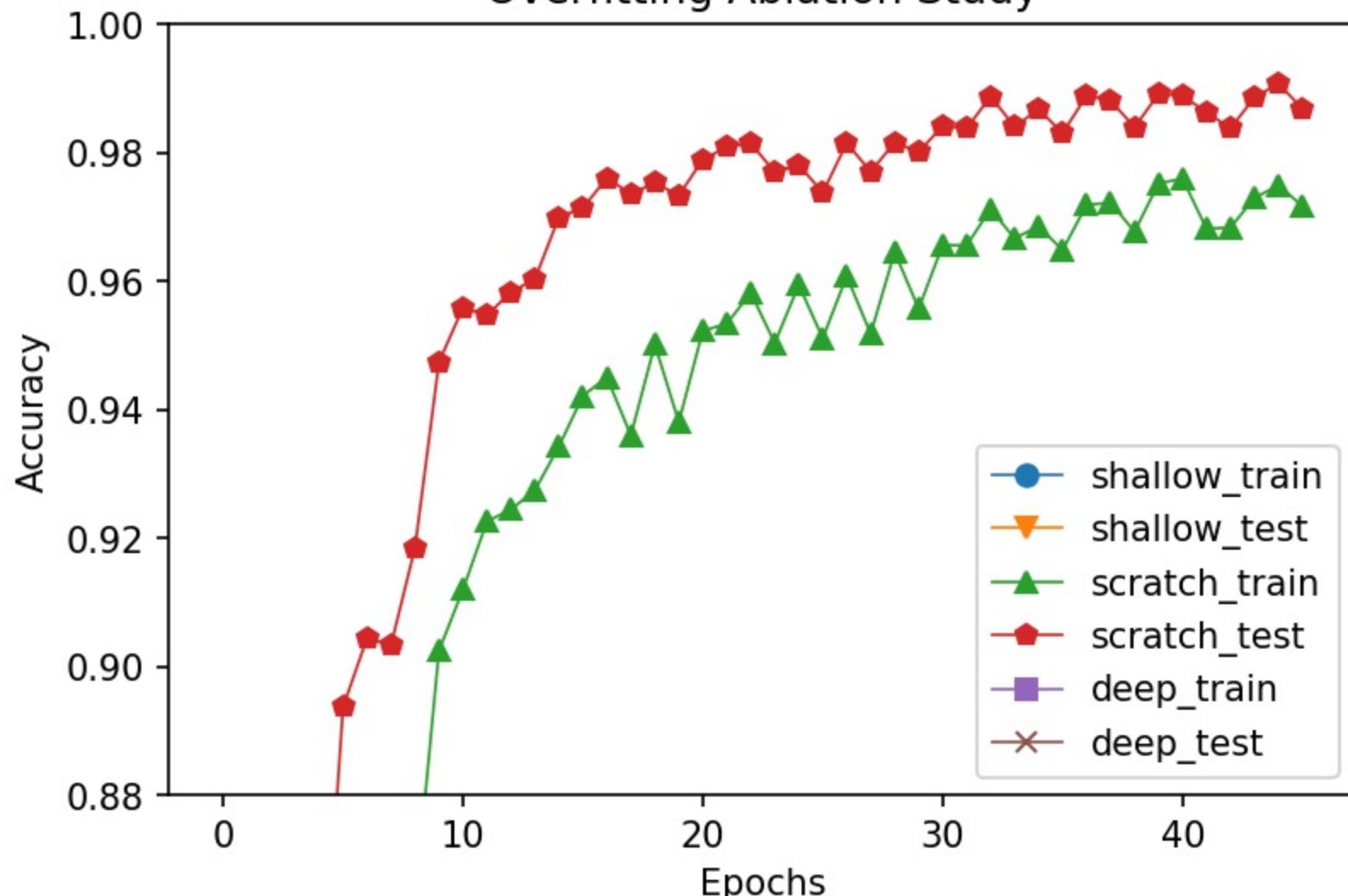
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	NaN	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 45 takes 316 seconds.

[47, 30] loss: 0.120

[47, 60] loss: 0.086

```
[47,    90] loss: 0.119
[47,   120] loss: 0.126
[47,   150] loss: 0.123
[47,   180] loss: 0.116
[47,   210] loss: 0.101
[47,   240] loss: 0.111
[47,   270] loss: 0.079
[47,   300] loss: 0.095
[47,   330] loss: 0.078
[47,   360] loss: 0.119
[47,   390] loss: 0.121
[47,   420] loss: 0.088
[47,   450] loss: 0.085
[47,   480] loss: 0.085
[47,   510] loss: 0.086
[47,   540] loss: 0.121
[47,   570] loss: 0.077
[47,   600] loss: 0.124
[47,   630] loss: 0.093
[47,   660] loss: 0.090
[47,   690] loss: 0.142
[47,   720] loss: 0.135
[47,   750] loss: 0.082
[47,   780] loss: 0.087
[47,   810] loss: 0.095
[47,   840] loss: 0.096
[47,   870] loss: 0.088
[47,   900] loss: 0.123
[47,   930] loss: 0.095
[47,   960] loss: 0.091
[47,   990] loss: 0.162
[47,  1020] loss: 0.090
[47,  1050] loss: 0.096
[47,  1080] loss: 0.108
[47,  1110] loss: 0.164
[47,  1140] loss: 0.134
[47,  1170] loss: 0.158
[47,  1200] loss: 0.099
[47,  1230] loss: 0.129
```

```
[47, 1260] loss: 0.122
[47, 1290] loss: 0.126
[47, 1320] loss: 0.097
[47, 1350] loss: 0.068
[47, 1380] loss: 0.128
[47, 1410] loss: 0.105
[47, 1440] loss: 0.136
[47, 1470] loss: 0.110
[47, 1500] loss: 0.119
[47, 1530] loss: 0.093
[47, 1560] loss: 0.114
[47, 1590] loss: 0.108
[47, 1620] loss: 0.133
[47, 1650] loss: 0.113
[47, 1680] loss: 0.101
[47, 1710] loss: 0.128
[47, 1740] loss: 0.119
[47, 1770] loss: 0.086
[47, 1800] loss: 0.084
[47, 1830] loss: 0.086
[47, 1860] loss: 0.113
[47, 1890] loss: 0.096
[47, 1920] loss: 0.099
[47, 1950] loss: 0.128
[47, 1980] loss: 0.088
[47, 2010] loss: 0.093
[47, 2040] loss: 0.112
[47, 2070] loss: 0.111
[47, 2100] loss: 0.149
[47, 2130] loss: 0.105
[47, 2160] loss: 0.117
[47, 2190] loss: 0.111
eval intermediate_models/epoch46_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.975941
epoch 46 accuracy on train set is: 0.9759405670665212
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	NaN	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.987732

epoch 46 accuracy on test set is: 0.9877317339149401

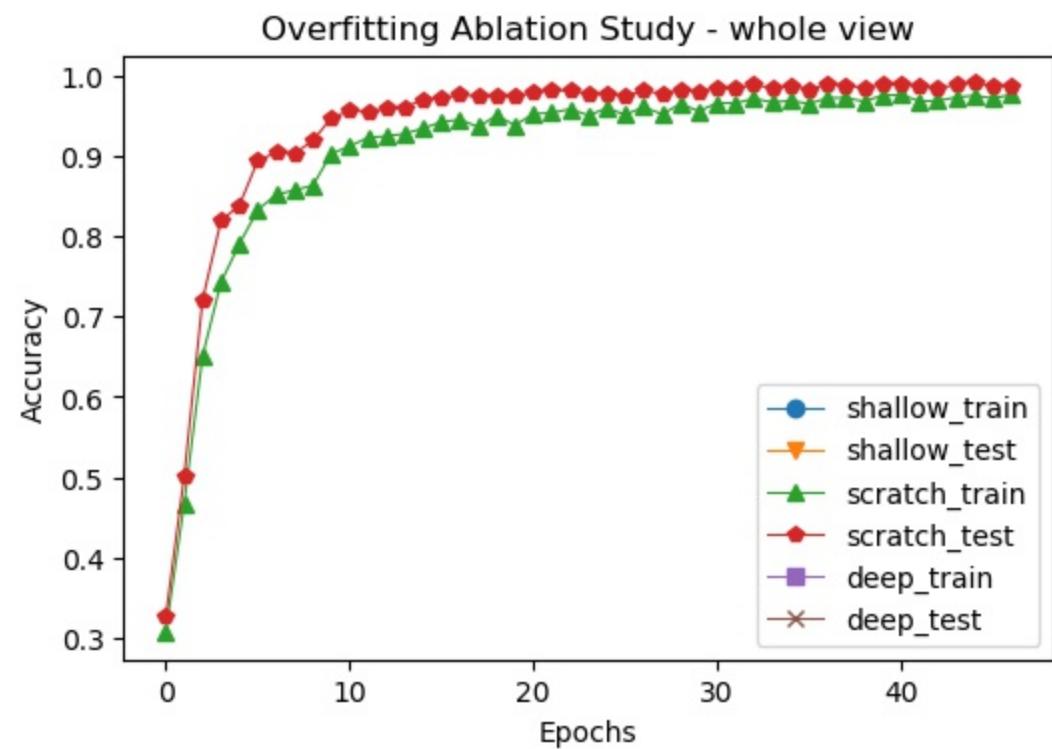
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

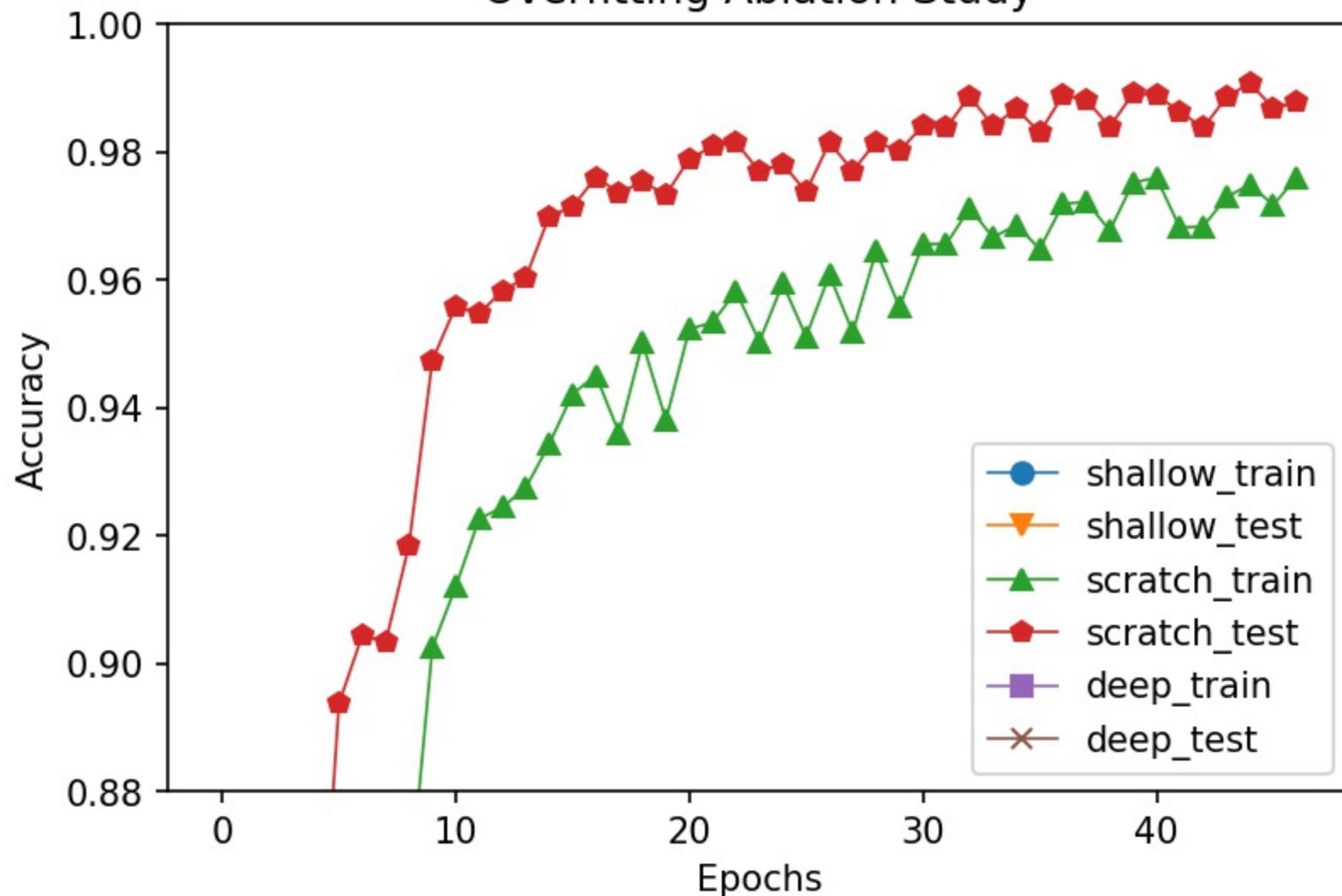
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	NaN	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 46 takes 317 seconds.

[48, 30] loss: 0.110

[48, 60] loss: 0.070

```
[48,    90] loss: 0.138
[48,   120] loss: 0.081
[48,   150] loss: 0.103
[48,   180] loss: 0.123
[48,   210] loss: 0.096
[48,   240] loss: 0.073
[48,   270] loss: 0.109
[48,   300] loss: 0.064
[48,   330] loss: 0.108
[48,   360] loss: 0.100
[48,   390] loss: 0.071
[48,   420] loss: 0.093
[48,   450] loss: 0.102
[48,   480] loss: 0.105
[48,   510] loss: 0.093
[48,   540] loss: 0.127
[48,   570] loss: 0.142
[48,   600] loss: 0.112
[48,   630] loss: 0.084
[48,   660] loss: 0.077
[48,   690] loss: 0.138
[48,   720] loss: 0.082
[48,   750] loss: 0.111
[48,   780] loss: 0.078
[48,   810] loss: 0.159
[48,   840] loss: 0.097
[48,   870] loss: 0.100
[48,   900] loss: 0.119
[48,   930] loss: 0.091
[48,   960] loss: 0.097
[48,   990] loss: 0.089
[48,  1020] loss: 0.105
[48,  1050] loss: 0.071
[48,  1080] loss: 0.086
[48,  1110] loss: 0.073
[48,  1140] loss: 0.106
[48,  1170] loss: 0.091
[48,  1200] loss: 0.113
[48,  1230] loss: 0.096
```

```
[48, 1260] loss: 0.081
[48, 1290] loss: 0.105
[48, 1320] loss: 0.123
[48, 1350] loss: 0.110
[48, 1380] loss: 0.082
[48, 1410] loss: 0.090
[48, 1440] loss: 0.088
[48, 1470] loss: 0.070
[48, 1500] loss: 0.080
[48, 1530] loss: 0.080
[48, 1560] loss: 0.102
[48, 1590] loss: 0.133
[48, 1620] loss: 0.098
[48, 1650] loss: 0.086
[48, 1680] loss: 0.096
[48, 1710] loss: 0.096
[48, 1740] loss: 0.093
[48, 1770] loss: 0.077
[48, 1800] loss: 0.076
[48, 1830] loss: 0.091
[48, 1860] loss: 0.101
[48, 1890] loss: 0.103
[48, 1920] loss: 0.101
[48, 1950] loss: 0.083
[48, 1980] loss: 0.117
[48, 2010] loss: 0.135
[48, 2040] loss: 0.077
[48, 2070] loss: 0.101
[48, 2100] loss: 0.073
[48, 2130] loss: 0.089
[48, 2160] loss: 0.138
[48, 2190] loss: 0.114
eval intermediate_models/epoch47_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.974123
epoch 47 accuracy on train set is: 0.9741230461650309
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	NaN	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.987186

epoch 47 accuracy on test set is: 0.987186477644493

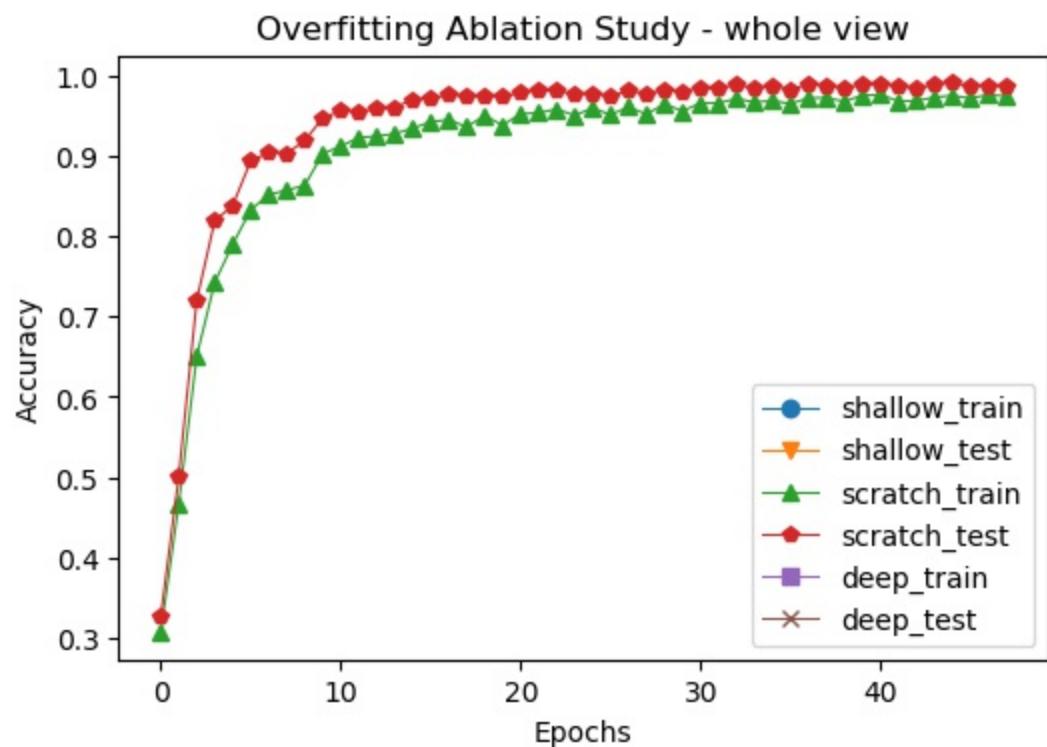
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

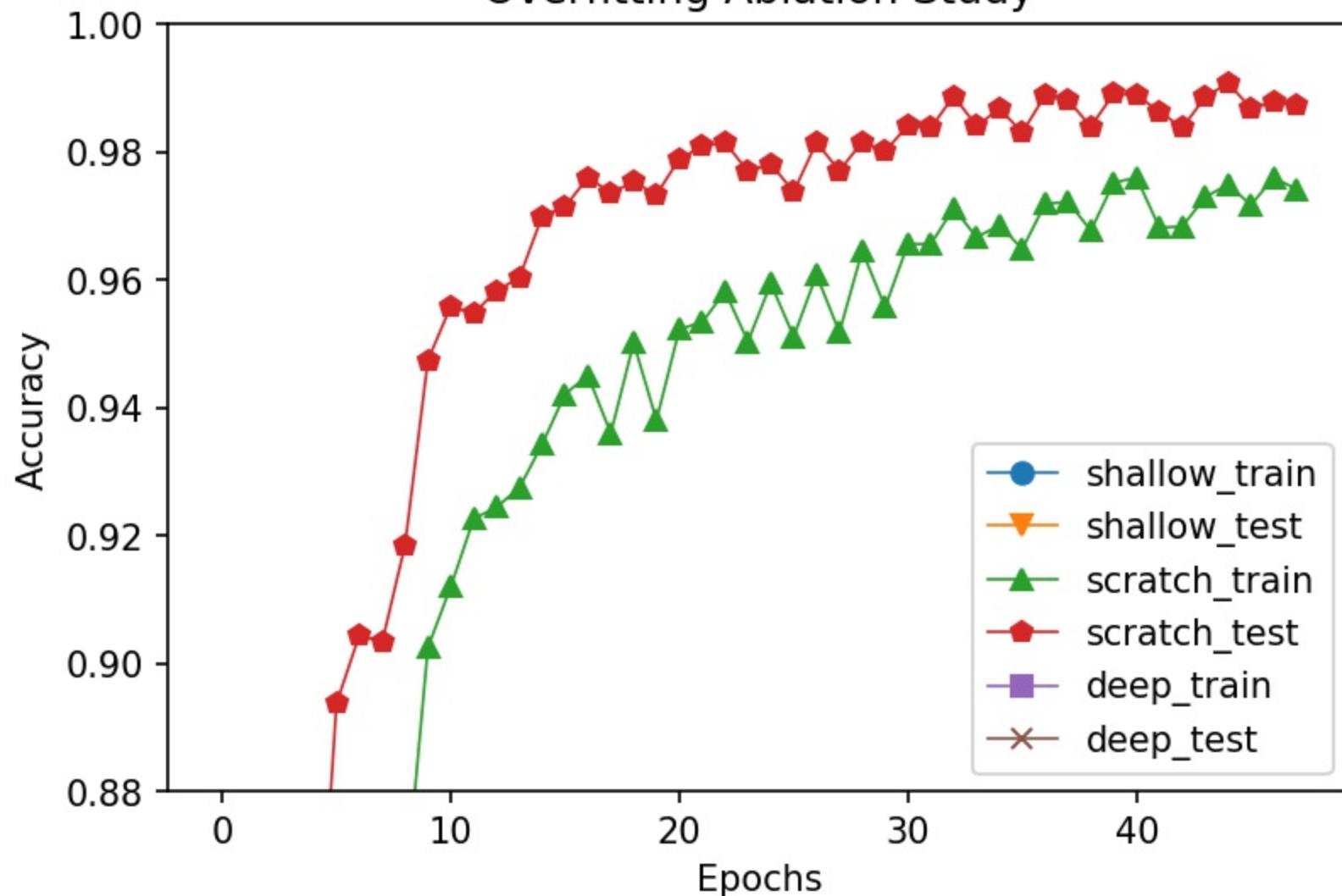
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	NaN	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 47 takes 316 seconds.

[49, 30] loss: 0.105

[49, 60] loss: 0.108

```
[49,    90] loss: 0.097
[49,   120] loss: 0.132
[49,   150] loss: 0.089
[49,   180] loss: 0.152
[49,   210] loss: 0.121
[49,   240] loss: 0.116
[49,   270] loss: 0.079
[49,   300] loss: 0.096
[49,   330] loss: 0.107
[49,   360] loss: 0.121
[49,   390] loss: 0.060
[49,   420] loss: 0.069
[49,   450] loss: 0.139
[49,   480] loss: 0.051
[49,   510] loss: 0.082
[49,   540] loss: 0.097
[49,   570] loss: 0.089
[49,   600] loss: 0.097
[49,   630] loss: 0.064
[49,   660] loss: 0.093
[49,   690] loss: 0.080
[49,   720] loss: 0.089
[49,   750] loss: 0.121
[49,   780] loss: 0.132
[49,   810] loss: 0.120
[49,   840] loss: 0.097
[49,   870] loss: 0.136
[49,   900] loss: 0.088
[49,   930] loss: 0.147
[49,   960] loss: 0.080
[49,   990] loss: 0.080
[49,  1020] loss: 0.076
[49,  1050] loss: 0.083
[49,  1080] loss: 0.125
[49,  1110] loss: 0.111
[49,  1140] loss: 0.112
[49,  1170] loss: 0.119
[49,  1200] loss: 0.106
[49,  1230] loss: 0.096
```

```
[49, 1260] loss: 0.114
[49, 1290] loss: 0.147
[49, 1320] loss: 0.127
[49, 1350] loss: 0.130
[49, 1380] loss: 0.104
[49, 1410] loss: 0.063
[49, 1440] loss: 0.089
[49, 1470] loss: 0.099
[49, 1500] loss: 0.133
[49, 1530] loss: 0.113
[49, 1560] loss: 0.080
[49, 1590] loss: 0.137
[49, 1620] loss: 0.099
[49, 1650] loss: 0.098
[49, 1680] loss: 0.102
[49, 1710] loss: 0.084
[49, 1740] loss: 0.100
[49, 1770] loss: 0.117
[49, 1800] loss: 0.084
[49, 1830] loss: 0.096
[49, 1860] loss: 0.087
[49, 1890] loss: 0.100
[49, 1920] loss: 0.080
[49, 1950] loss: 0.109
[49, 1980] loss: 0.095
[49, 2010] loss: 0.090
[49, 2040] loss: 0.095
[49, 2070] loss: 0.118
[49, 2100] loss: 0.112
[49, 2130] loss: 0.112
[49, 2160] loss: 0.073
[49, 2190] loss: 0.084
eval intermediate_models/epoch48_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.975304
epoch 48 accuracy on train set is: 0.9753044347509996
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	NaN	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.987005

epoch 48 accuracy on test set is: 0.9870047255543438

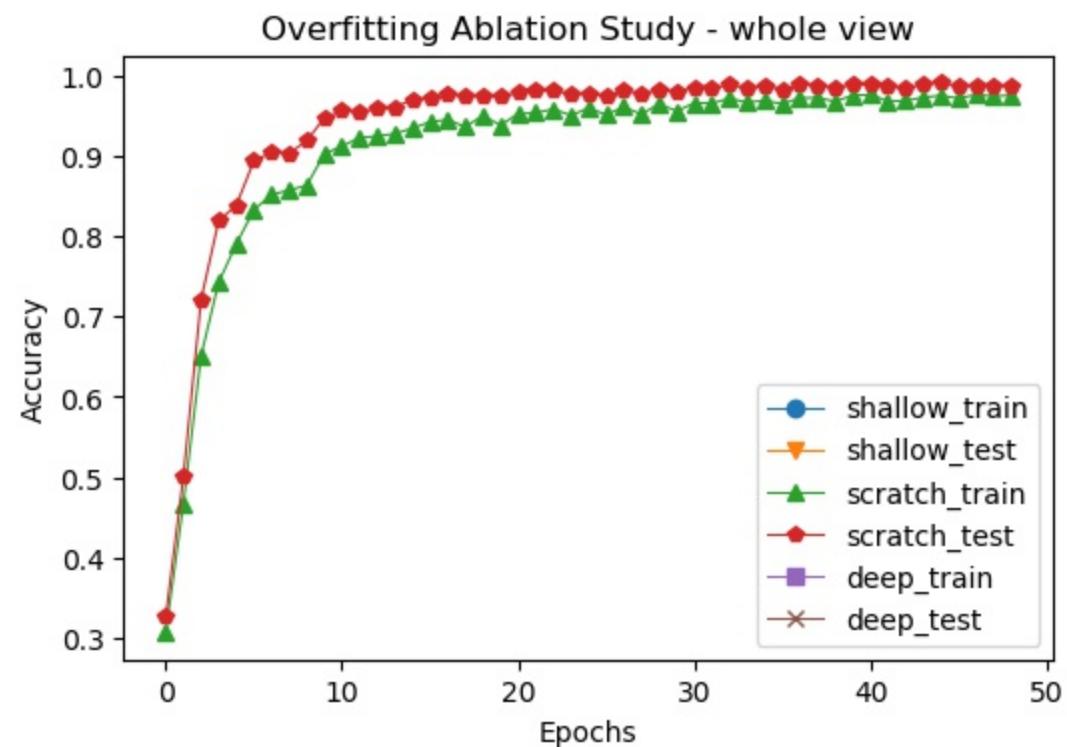
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

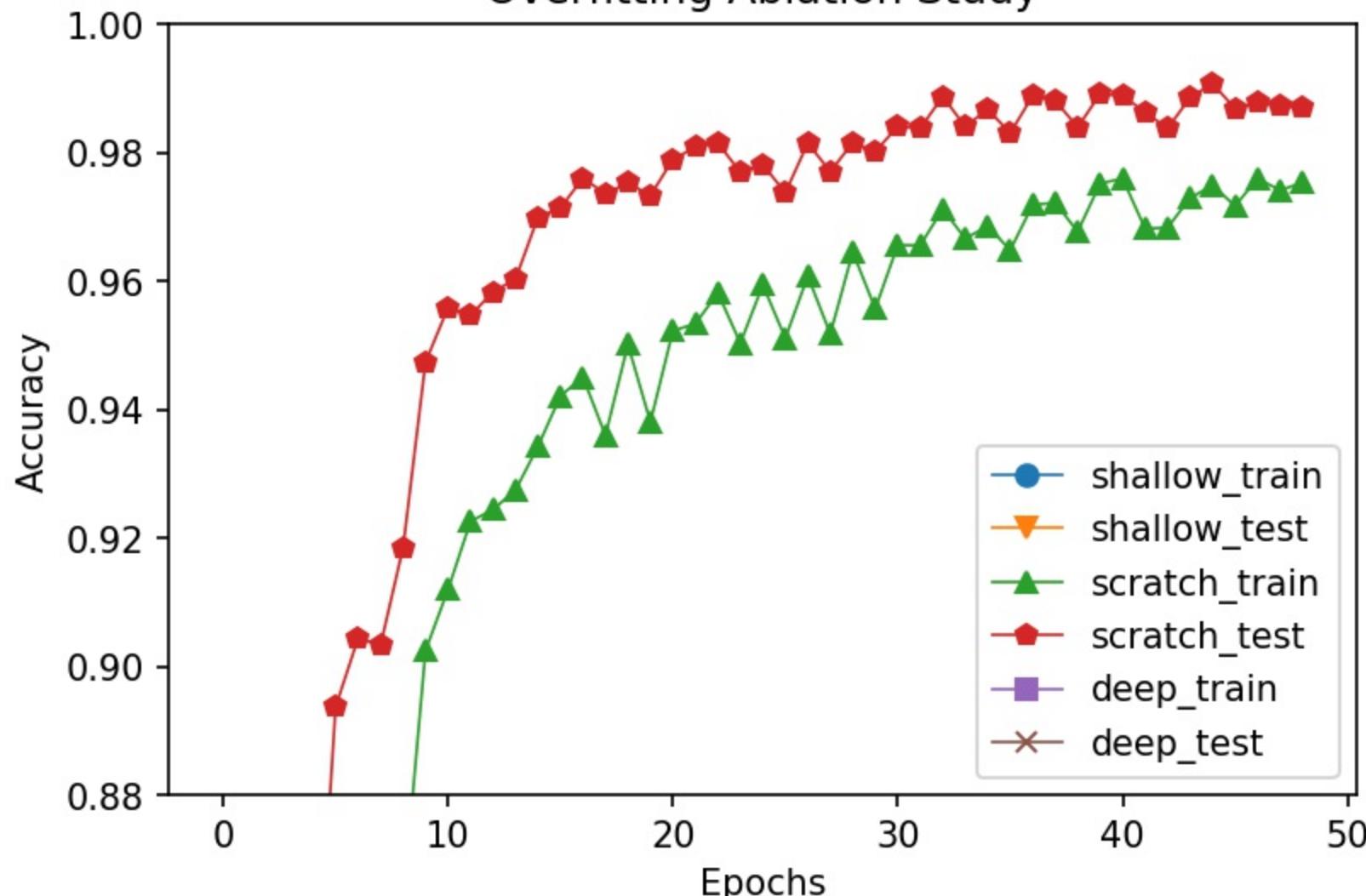
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	NaN	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 48 takes 316 seconds.

[50, 30] loss: 0.106

[50, 60] loss: 0.110

```
[50,    90] loss: 0.109
[50,   120] loss: 0.078
[50,   150] loss: 0.127
[50,   180] loss: 0.059
[50,   210] loss: 0.102
[50,   240] loss: 0.068
[50,   270] loss: 0.111
[50,   300] loss: 0.100
[50,   330] loss: 0.100
[50,   360] loss: 0.055
[50,   390] loss: 0.102
[50,   420] loss: 0.086
[50,   450] loss: 0.083
[50,   480] loss: 0.094
[50,   510] loss: 0.152
[50,   540] loss: 0.097
[50,   570] loss: 0.095
[50,   600] loss: 0.091
[50,   630] loss: 0.088
[50,   660] loss: 0.112
[50,   690] loss: 0.097
[50,   720] loss: 0.083
[50,   750] loss: 0.097
[50,   780] loss: 0.094
[50,   810] loss: 0.083
[50,   840] loss: 0.093
[50,   870] loss: 0.088
[50,   900] loss: 0.128
[50,   930] loss: 0.119
[50,   960] loss: 0.080
[50,   990] loss: 0.117
[50,  1020] loss: 0.099
[50,  1050] loss: 0.107
[50,  1080] loss: 0.104
[50,  1110] loss: 0.097
[50,  1140] loss: 0.104
[50,  1170] loss: 0.104
[50,  1200] loss: 0.096
[50,  1230] loss: 0.102
```

```
[50, 1260] loss: 0.137
[50, 1290] loss: 0.123
[50, 1320] loss: 0.086
[50, 1350] loss: 0.113
[50, 1380] loss: 0.114
[50, 1410] loss: 0.103
[50, 1440] loss: 0.083
[50, 1470] loss: 0.068
[50, 1500] loss: 0.093
[50, 1530] loss: 0.141
[50, 1560] loss: 0.101
[50, 1590] loss: 0.126
[50, 1620] loss: 0.092
[50, 1650] loss: 0.098
[50, 1680] loss: 0.095
[50, 1710] loss: 0.100
[50, 1740] loss: 0.068
[50, 1770] loss: 0.069
[50, 1800] loss: 0.078
[50, 1830] loss: 0.127
[50, 1860] loss: 0.126
[50, 1890] loss: 0.093
[50, 1920] loss: 0.084
[50, 1950] loss: 0.103
[50, 1980] loss: 0.083
[50, 2010] loss: 0.090
[50, 2040] loss: 0.083
[50, 2070] loss: 0.108
[50, 2100] loss: 0.110
[50, 2130] loss: 0.106
[50, 2160] loss: 0.091
[50, 2190] loss: 0.084
eval intermediate_models/epoch49_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.969443
epoch 49 accuracy on train set is: 0.9694429298436932
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	NaN	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.982370

epoch 49 accuracy on test set is: 0.9823700472555434

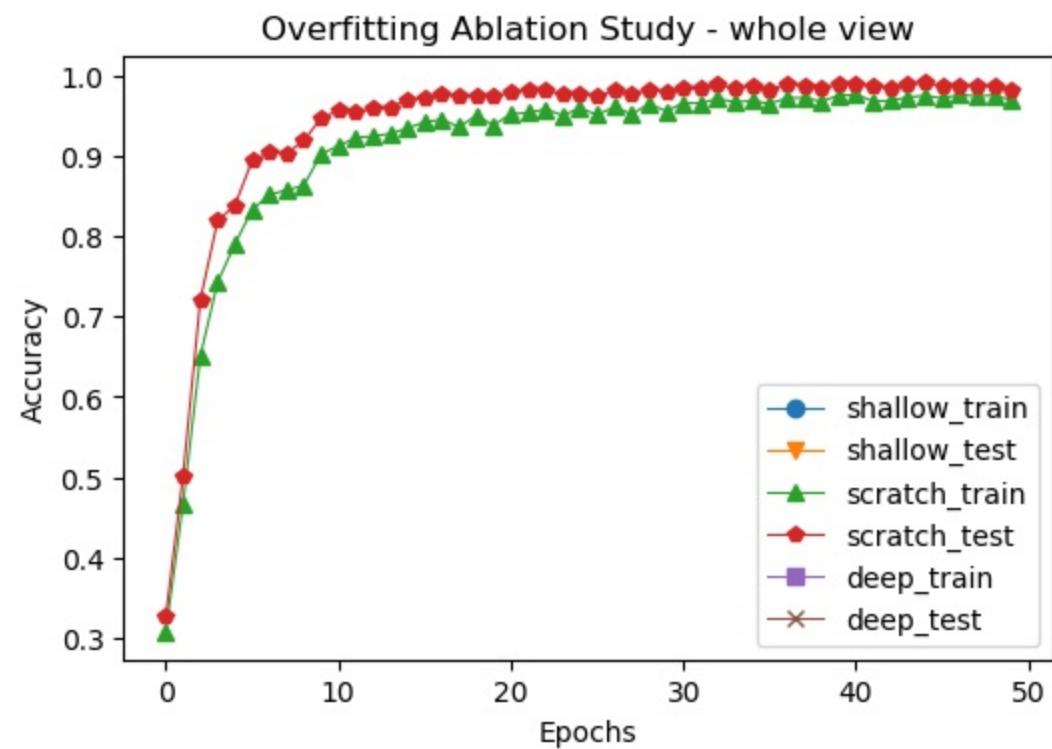
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

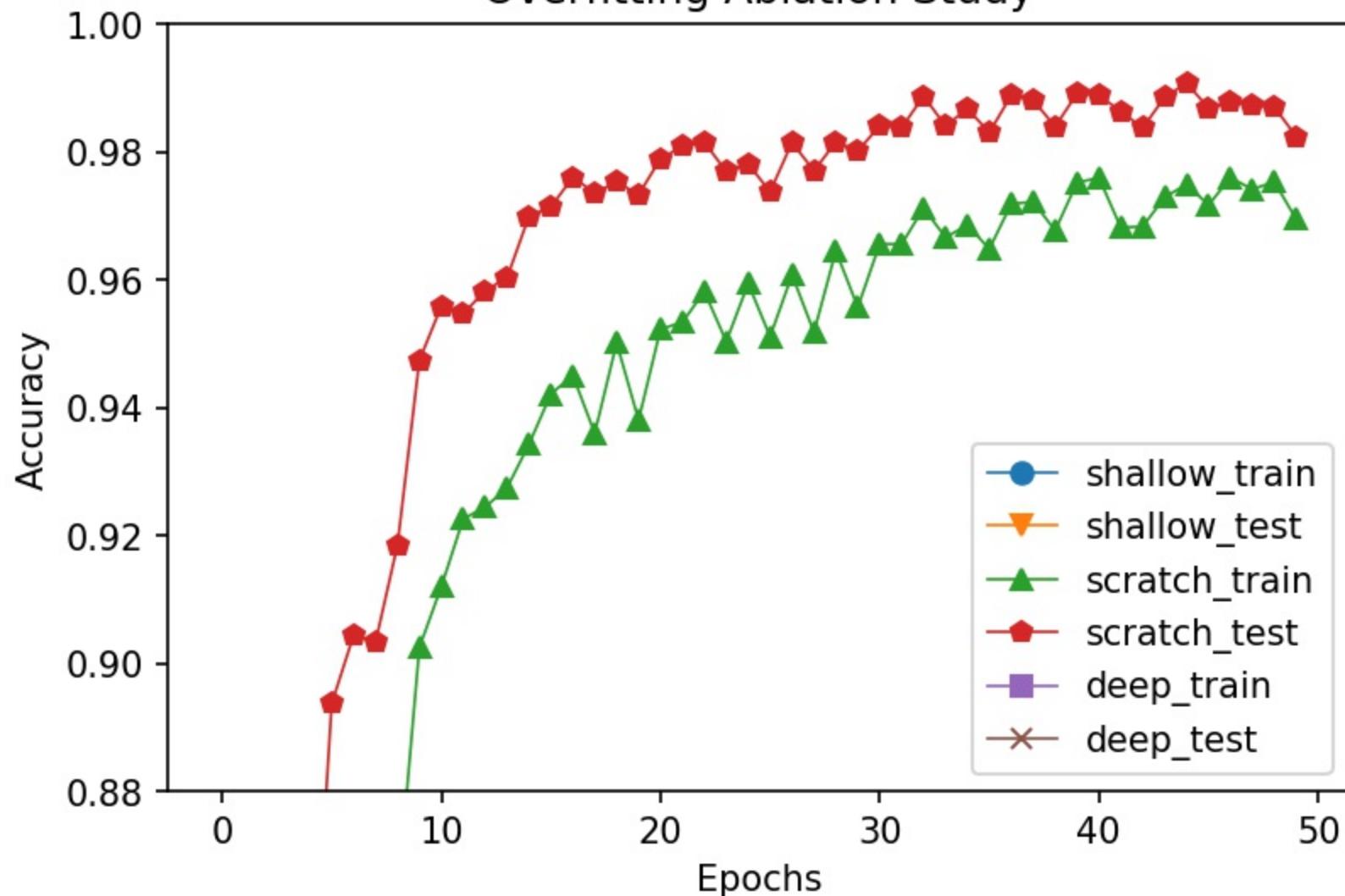
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	NaN	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 49 takes 317 seconds.

[51, 30] loss: 0.149

[51, 60] loss: 0.098

```
[51,    90] loss: 0.106
[51,   120] loss: 0.073
[51,   150] loss: 0.108
[51,   180] loss: 0.099
[51,   210] loss: 0.115
[51,   240] loss: 0.125
[51,   270] loss: 0.082
[51,   300] loss: 0.093
[51,   330] loss: 0.090
[51,   360] loss: 0.116
[51,   390] loss: 0.099
[51,   420] loss: 0.126
[51,   450] loss: 0.060
[51,   480] loss: 0.090
[51,   510] loss: 0.113
[51,   540] loss: 0.080
[51,   570] loss: 0.086
[51,   600] loss: 0.084
[51,   630] loss: 0.106
[51,   660] loss: 0.117
[51,   690] loss: 0.108
[51,   720] loss: 0.101
[51,   750] loss: 0.100
[51,   780] loss: 0.129
[51,   810] loss: 0.071
[51,   840] loss: 0.135
[51,   870] loss: 0.084
[51,   900] loss: 0.105
[51,   930] loss: 0.119
[51,   960] loss: 0.107
[51,   990] loss: 0.088
[51,  1020] loss: 0.069
[51,  1050] loss: 0.059
[51,  1080] loss: 0.123
[51,  1110] loss: 0.112
[51,  1140] loss: 0.129
[51,  1170] loss: 0.103
[51,  1200] loss: 0.077
[51,  1230] loss: 0.104
```

```
[51, 1260] loss: 0.093
[51, 1290] loss: 0.064
[51, 1320] loss: 0.098
[51, 1350] loss: 0.069
[51, 1380] loss: 0.081
[51, 1410] loss: 0.100
[51, 1440] loss: 0.108
[51, 1470] loss: 0.092
[51, 1500] loss: 0.095
[51, 1530] loss: 0.106
[51, 1560] loss: 0.102
[51, 1590] loss: 0.071
[51, 1620] loss: 0.097
[51, 1650] loss: 0.091
[51, 1680] loss: 0.122
[51, 1710] loss: 0.099
[51, 1740] loss: 0.070
[51, 1770] loss: 0.093
[51, 1800] loss: 0.081
[51, 1830] loss: 0.107
[51, 1860] loss: 0.081
[51, 1890] loss: 0.054
[51, 1920] loss: 0.099
[51, 1950] loss: 0.109
[51, 1980] loss: 0.080
[51, 2010] loss: 0.075
[51, 2040] loss: 0.088
[51, 2070] loss: 0.102
[51, 2100] loss: 0.105
[51, 2130] loss: 0.084
[51, 2160] loss: 0.059
[51, 2190] loss: 0.079
eval intermediate_models/epoch50_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.977008
epoch 50 accuracy on train set is: 0.9770083605961468
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	NaN	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990276

epoch 50 accuracy on test set is: 0.9902762631770266

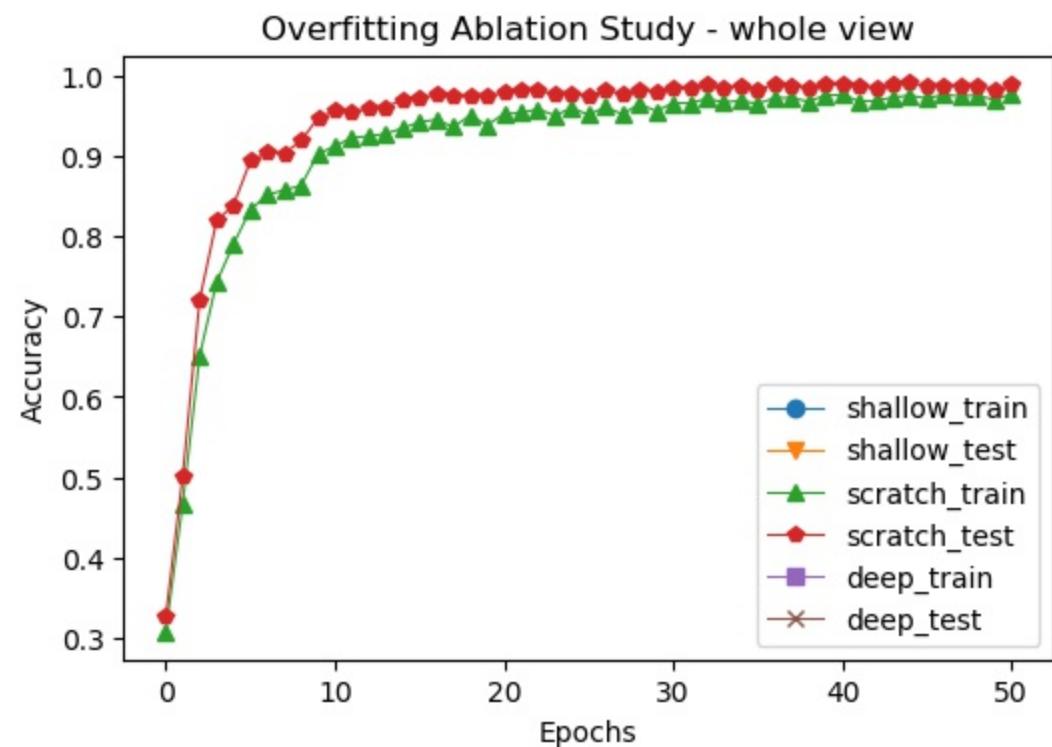
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

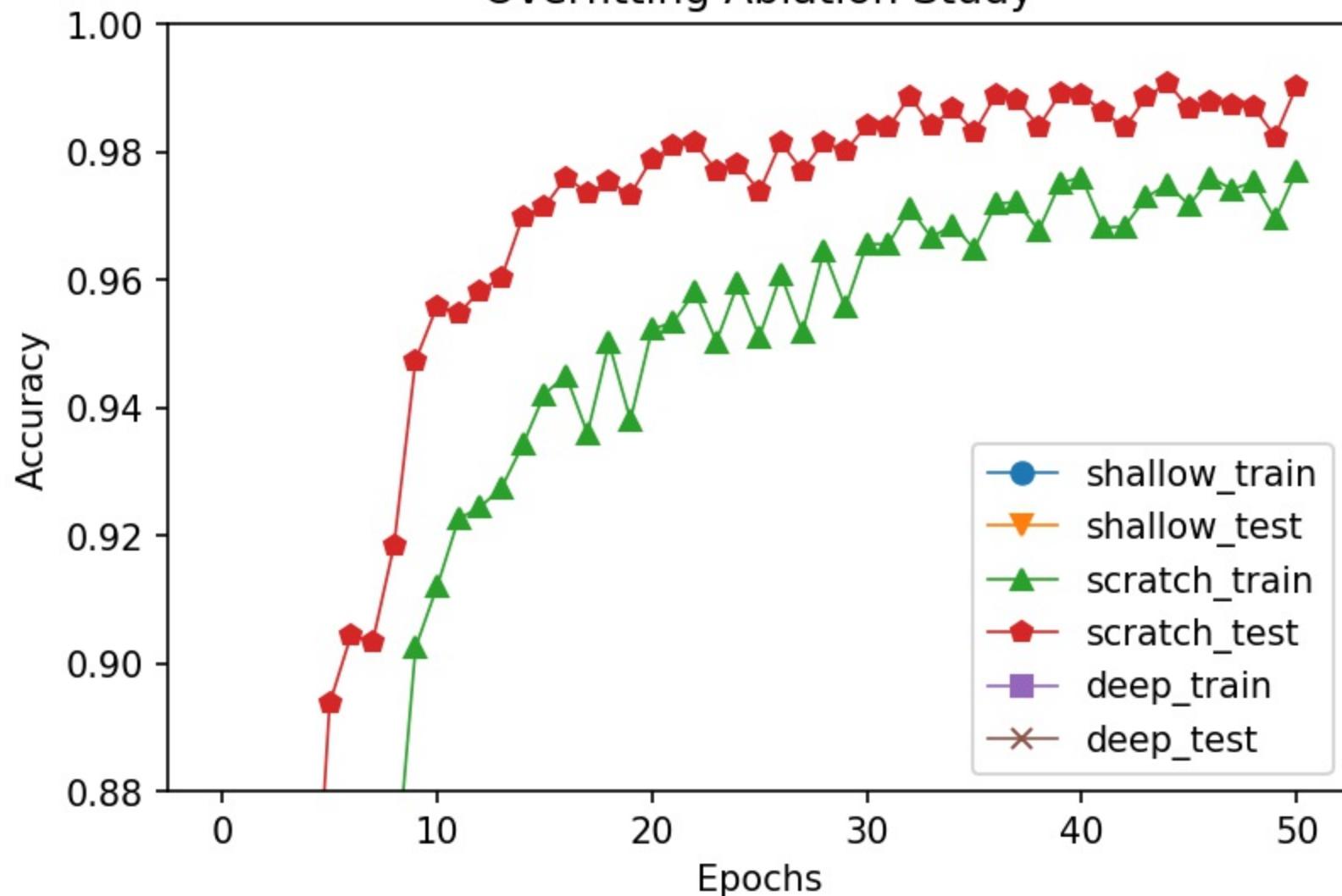
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	NaN	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 50 takes 316 seconds.

[52, 30] loss: 0.087

[52, 60] loss: 0.114

```
[52,    90] loss: 0.056
[52,   120] loss: 0.090
[52,   150] loss: 0.064
[52,   180] loss: 0.127
[52,   210] loss: 0.088
[52,   240] loss: 0.133
[52,   270] loss: 0.081
[52,   300] loss: 0.106
[52,   330] loss: 0.133
[52,   360] loss: 0.092
[52,   390] loss: 0.098
[52,   420] loss: 0.084
[52,   450] loss: 0.097
[52,   480] loss: 0.129
[52,   510] loss: 0.089
[52,   540] loss: 0.129
[52,   570] loss: 0.122
[52,   600] loss: 0.086
[52,   630] loss: 0.130
[52,   660] loss: 0.098
[52,   690] loss: 0.073
[52,   720] loss: 0.082
[52,   750] loss: 0.093
[52,   780] loss: 0.112
[52,   810] loss: 0.100
[52,   840] loss: 0.092
[52,   870] loss: 0.095
[52,   900] loss: 0.095
[52,   930] loss: 0.089
[52,   960] loss: 0.089
[52,   990] loss: 0.104
[52,  1020] loss: 0.093
[52,  1050] loss: 0.107
[52,  1080] loss: 0.105
[52,  1110] loss: 0.086
[52,  1140] loss: 0.075
[52,  1170] loss: 0.079
[52,  1200] loss: 0.102
[52,  1230] loss: 0.108
```

```
[52, 1260] loss: 0.075
[52, 1290] loss: 0.077
[52, 1320] loss: 0.100
[52, 1350] loss: 0.059
[52, 1380] loss: 0.117
[52, 1410] loss: 0.106
[52, 1440] loss: 0.084
[52, 1470] loss: 0.082
[52, 1500] loss: 0.086
[52, 1530] loss: 0.094
[52, 1560] loss: 0.098
[52, 1590] loss: 0.082
[52, 1620] loss: 0.135
[52, 1650] loss: 0.115
[52, 1680] loss: 0.107
[52, 1710] loss: 0.065
[52, 1740] loss: 0.106
[52, 1770] loss: 0.097
[52, 1800] loss: 0.096
[52, 1830] loss: 0.147
[52, 1860] loss: 0.063
[52, 1890] loss: 0.077
[52, 1920] loss: 0.083
[52, 1950] loss: 0.076
[52, 1980] loss: 0.119
[52, 2010] loss: 0.123
[52, 2040] loss: 0.084
[52, 2070] loss: 0.078
[52, 2100] loss: 0.087
[52, 2130] loss: 0.077
[52, 2160] loss: 0.090
[52, 2190] loss: 0.086
eval intermediate_models/epoch51_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.975782
epoch 51 accuracy on train set is: 0.9757815339876409
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	NaN	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.987732

epoch 51 accuracy on test set is: 0.9877317339149401

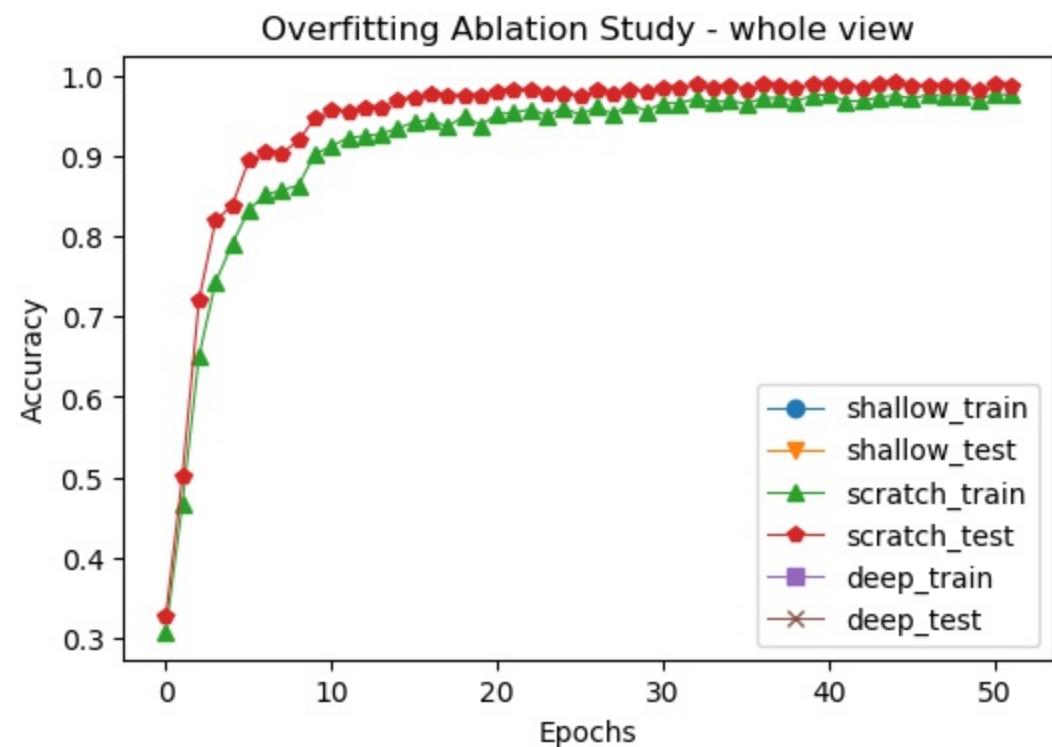
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

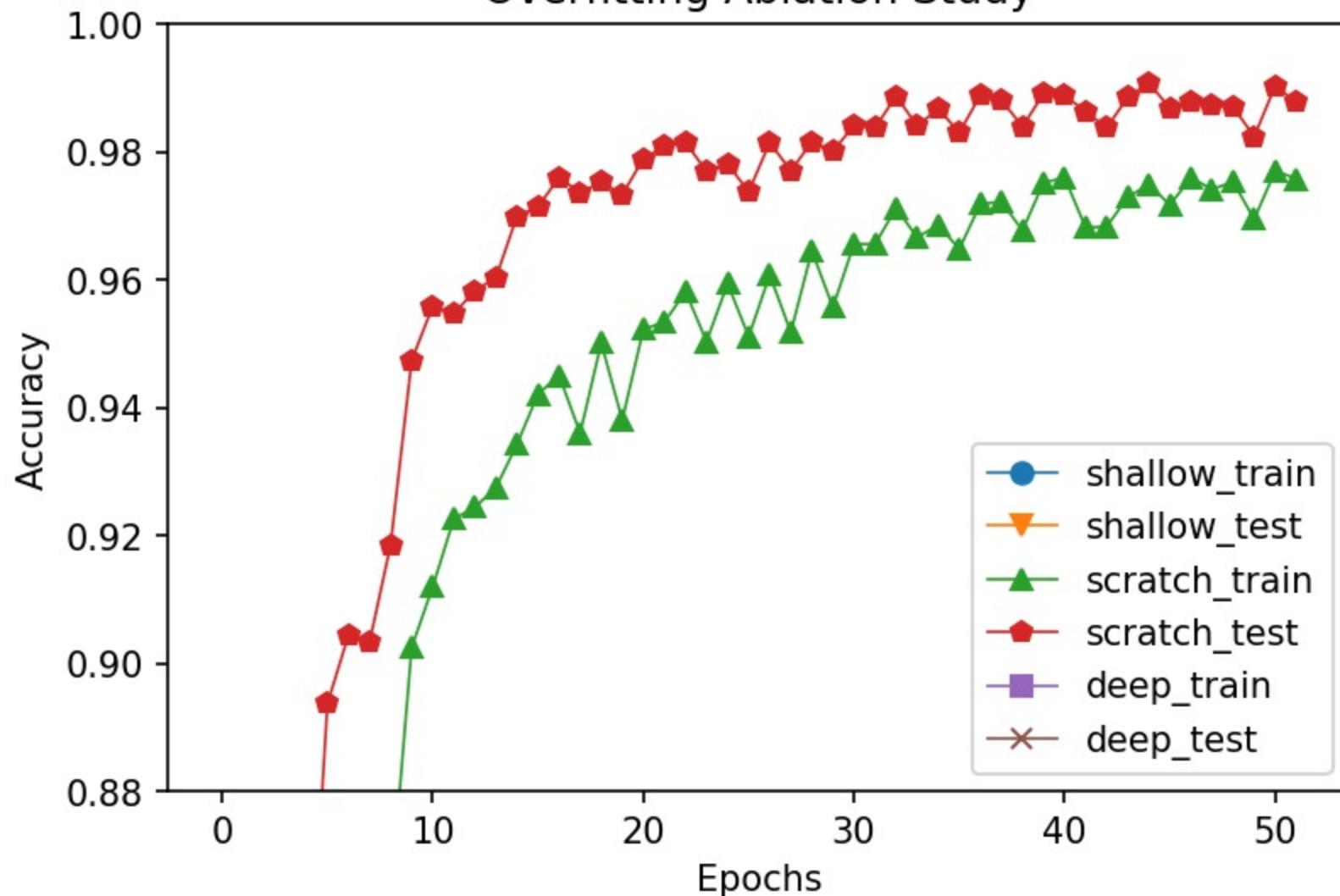
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	NaN	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 51 takes 316 seconds.

[53, 30] loss: 0.083

[53, 60] loss: 0.046

```
[53,    90] loss: 0.063
[53,   120] loss: 0.093
[53,   150] loss: 0.105
[53,   180] loss: 0.112
[53,   210] loss: 0.091
[53,   240] loss: 0.131
[53,   270] loss: 0.149
[53,   300] loss: 0.079
[53,   330] loss: 0.113
[53,   360] loss: 0.081
[53,   390] loss: 0.065
[53,   420] loss: 0.095
[53,   450] loss: 0.095
[53,   480] loss: 0.085
[53,   510] loss: 0.080
[53,   540] loss: 0.080
[53,   570] loss: 0.071
[53,   600] loss: 0.076
[53,   630] loss: 0.086
[53,   660] loss: 0.087
[53,   690] loss: 0.070
[53,   720] loss: 0.113
[53,   750] loss: 0.072
[53,   780] loss: 0.130
[53,   810] loss: 0.083
[53,   840] loss: 0.069
[53,   870] loss: 0.069
[53,   900] loss: 0.090
[53,   930] loss: 0.089
[53,   960] loss: 0.109
[53,   990] loss: 0.100
[53,  1020] loss: 0.100
[53,  1050] loss: 0.096
[53,  1080] loss: 0.103
[53,  1110] loss: 0.101
[53,  1140] loss: 0.089
[53,  1170] loss: 0.134
[53,  1200] loss: 0.082
[53,  1230] loss: 0.120
```

```
[53, 1260] loss: 0.130
[53, 1290] loss: 0.082
[53, 1320] loss: 0.103
[53, 1350] loss: 0.079
[53, 1380] loss: 0.106
[53, 1410] loss: 0.124
[53, 1440] loss: 0.133
[53, 1470] loss: 0.068
[53, 1500] loss: 0.109
[53, 1530] loss: 0.111
[53, 1560] loss: 0.089
[53, 1590] loss: 0.095
[53, 1620] loss: 0.090
[53, 1650] loss: 0.136
[53, 1680] loss: 0.067
[53, 1710] loss: 0.077
[53, 1740] loss: 0.111
[53, 1770] loss: 0.106
[53, 1800] loss: 0.069
[53, 1830] loss: 0.084
[53, 1860] loss: 0.104
[53, 1890] loss: 0.152
[53, 1920] loss: 0.058
[53, 1950] loss: 0.066
[53, 1980] loss: 0.064
[53, 2010] loss: 0.095
[53, 2040] loss: 0.079
[53, 2070] loss: 0.094
[53, 2100] loss: 0.073
[53, 2130] loss: 0.060
[53, 2160] loss: 0.106
[53, 2190] loss: 0.080
eval intermediate_models/epoch52_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.977031
epoch 52 accuracy on train set is: 0.9770310796074155
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	NaN	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992276

epoch 52 accuracy on test set is: 0.9922755361686659

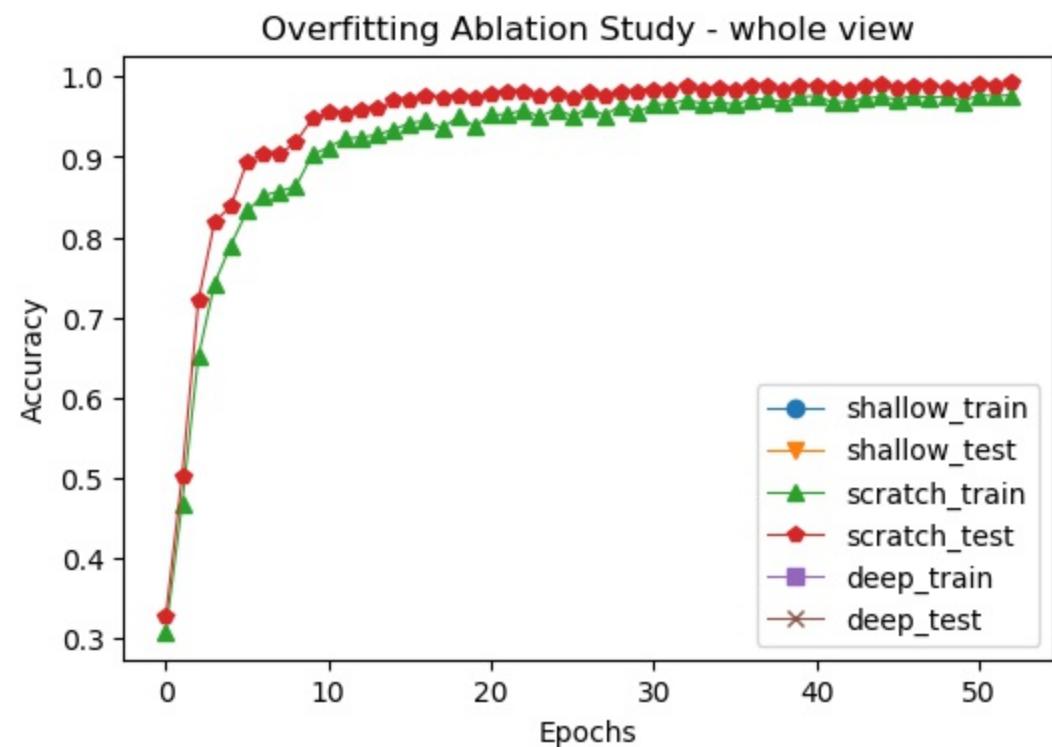
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

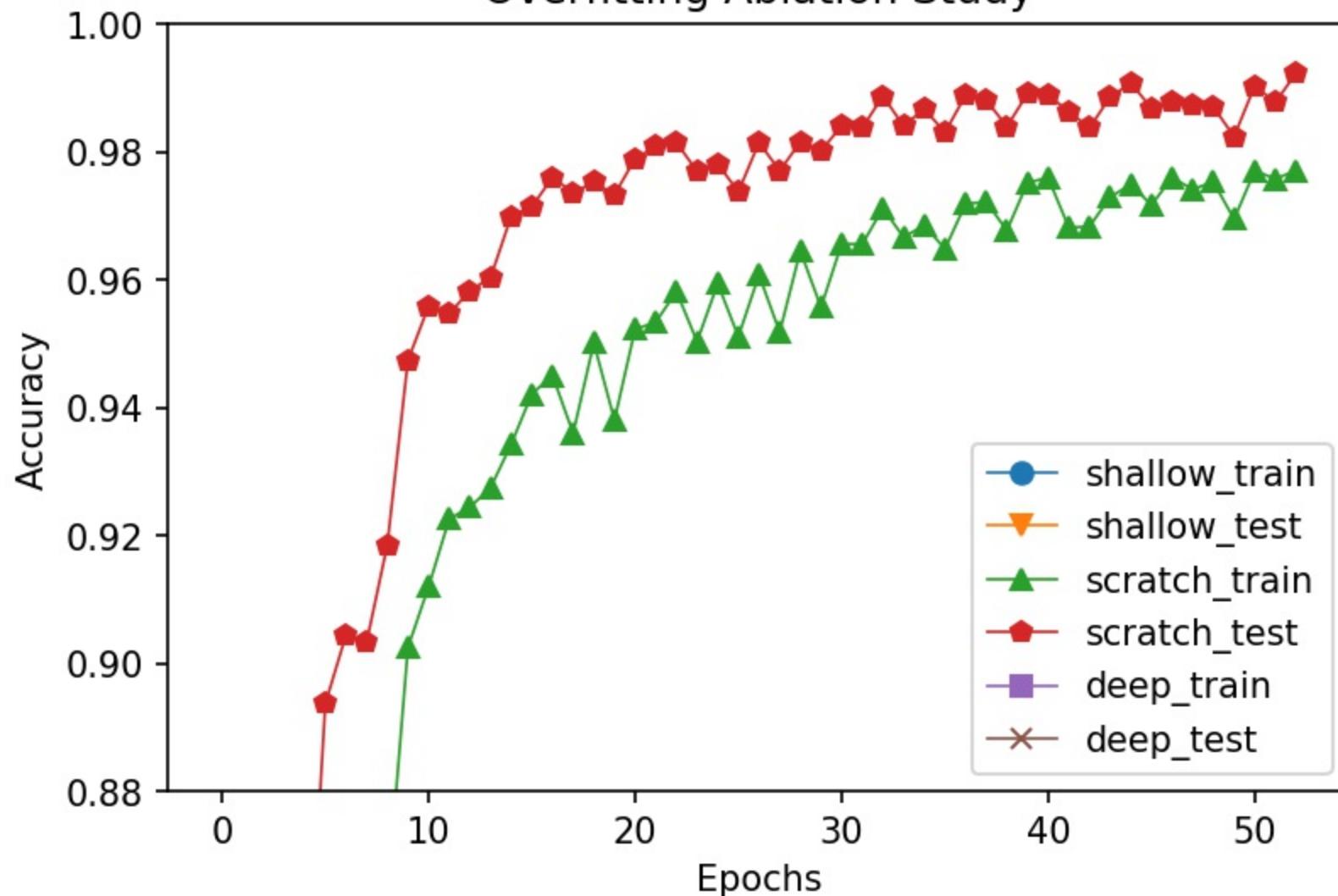
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	NaN	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 52 takes 316 seconds.

[54, 30] loss: 0.067

[54, 60] loss: 0.129

```
[54,    90] loss: 0.079
[54,   120] loss: 0.129
[54,   150] loss: 0.084
[54,   180] loss: 0.101
[54,   210] loss: 0.095
[54,   240] loss: 0.094
[54,   270] loss: 0.145
[54,   300] loss: 0.111
[54,   330] loss: 0.062
[54,   360] loss: 0.080
[54,   390] loss: 0.066
[54,   420] loss: 0.083
[54,   450] loss: 0.113
[54,   480] loss: 0.131
[54,   510] loss: 0.086
[54,   540] loss: 0.094
[54,   570] loss: 0.077
[54,   600] loss: 0.101
[54,   630] loss: 0.108
[54,   660] loss: 0.096
[54,   690] loss: 0.072
[54,   720] loss: 0.080
[54,   750] loss: 0.132
[54,   780] loss: 0.095
[54,   810] loss: 0.087
[54,   840] loss: 0.132
[54,   870] loss: 0.095
[54,   900] loss: 0.059
[54,   930] loss: 0.109
[54,   960] loss: 0.122
[54,   990] loss: 0.096
[54,  1020] loss: 0.079
[54,  1050] loss: 0.091
[54,  1080] loss: 0.098
[54,  1110] loss: 0.092
[54,  1140] loss: 0.100
[54,  1170] loss: 0.109
[54,  1200] loss: 0.071
[54,  1230] loss: 0.066
```

```
[54, 1260] loss: 0.078
[54, 1290] loss: 0.086
[54, 1320] loss: 0.098
[54, 1350] loss: 0.102
[54, 1380] loss: 0.098
[54, 1410] loss: 0.091
[54, 1440] loss: 0.141
[54, 1470] loss: 0.120
[54, 1500] loss: 0.069
[54, 1530] loss: 0.082
[54, 1560] loss: 0.082
[54, 1590] loss: 0.079
[54, 1620] loss: 0.113
[54, 1650] loss: 0.177
[54, 1680] loss: 0.080
[54, 1710] loss: 0.074
[54, 1740] loss: 0.063
[54, 1770] loss: 0.102
[54, 1800] loss: 0.130
[54, 1830] loss: 0.120
[54, 1860] loss: 0.081
[54, 1890] loss: 0.077
[54, 1920] loss: 0.080
[54, 1950] loss: 0.043
[54, 1980] loss: 0.140
[54, 2010] loss: 0.099
[54, 2040] loss: 0.085
[54, 2070] loss: 0.101
[54, 2100] loss: 0.108
[54, 2130] loss: 0.123
[54, 2160] loss: 0.070
[54, 2190] loss: 0.067
eval intermediate_models/epoch53_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.979053
epoch 53 accuracy on train set is: 0.9790530716103235
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	NaN	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.989004

epoch 53 accuracy on test set is: 0.9890039985459833

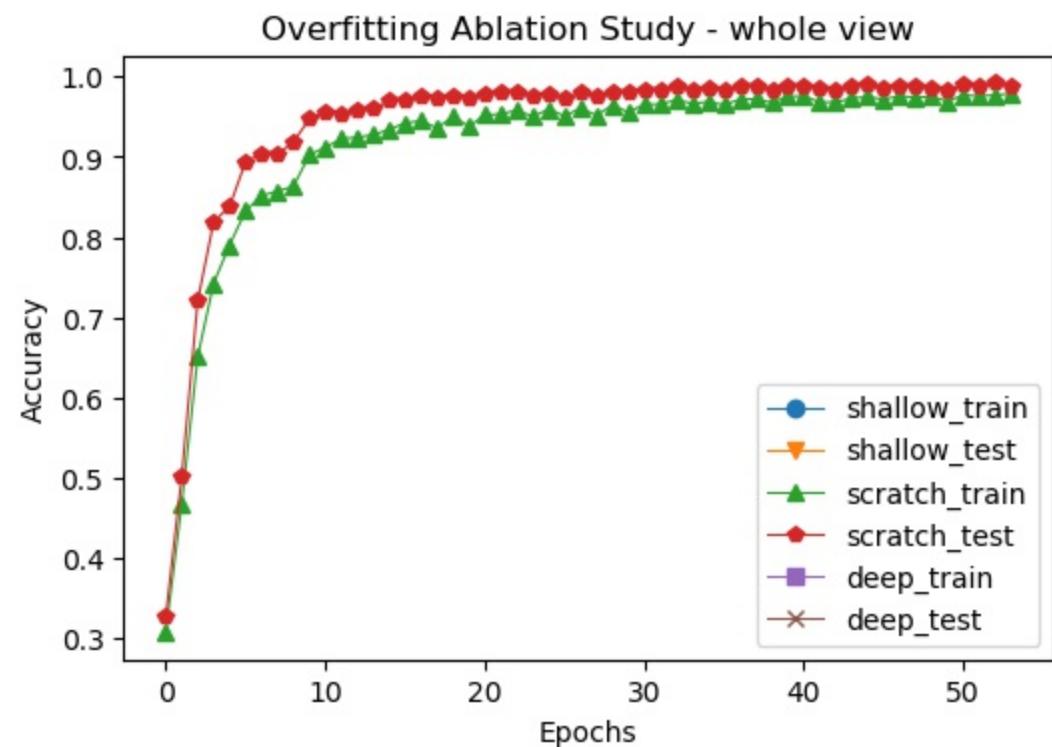
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

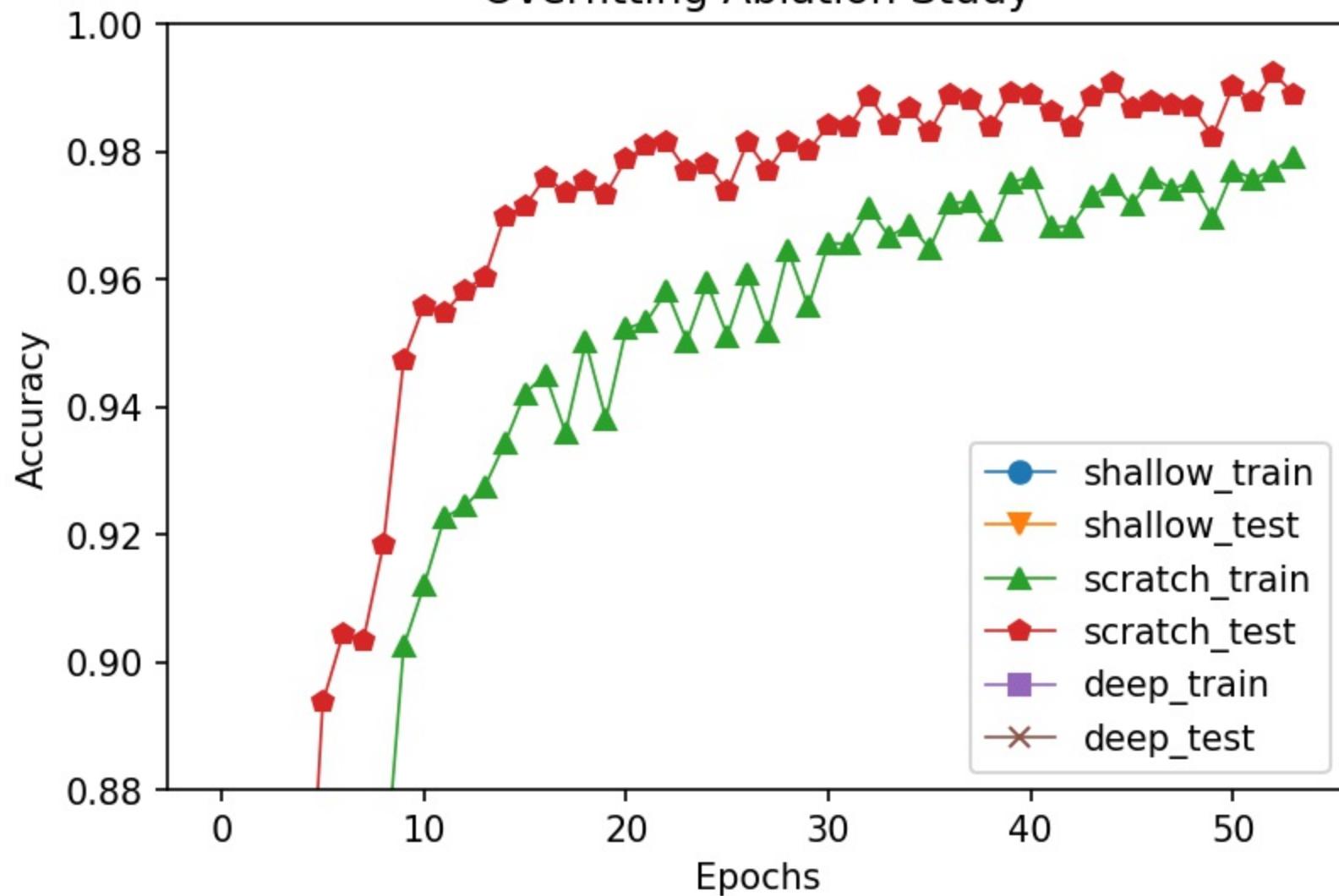
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	NaN	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 53 takes 316 seconds.

[55, 30] loss: 0.091

[55, 60] loss: 0.074

```
[55,    90] loss: 0.063
[55,   120] loss: 0.130
[55,   150] loss: 0.135
[55,   180] loss: 0.164
[55,   210] loss: 0.086
[55,   240] loss: 0.090
[55,   270] loss: 0.086
[55,   300] loss: 0.069
[55,   330] loss: 0.121
[55,   360] loss: 0.102
[55,   390] loss: 0.077
[55,   420] loss: 0.055
[55,   450] loss: 0.082
[55,   480] loss: 0.081
[55,   510] loss: 0.080
[55,   540] loss: 0.086
[55,   570] loss: 0.085
[55,   600] loss: 0.061
[55,   630] loss: 0.069
[55,   660] loss: 0.084
[55,   690] loss: 0.061
[55,   720] loss: 0.119
[55,   750] loss: 0.093
[55,   780] loss: 0.099
[55,   810] loss: 0.047
[55,   840] loss: 0.079
[55,   870] loss: 0.090
[55,   900] loss: 0.088
[55,   930] loss: 0.093
[55,   960] loss: 0.080
[55,   990] loss: 0.099
[55,  1020] loss: 0.137
[55,  1050] loss: 0.129
[55,  1080] loss: 0.107
[55,  1110] loss: 0.111
[55,  1140] loss: 0.087
[55,  1170] loss: 0.074
[55,  1200] loss: 0.101
[55,  1230] loss: 0.058
```

```
[55, 1260] loss: 0.108
[55, 1290] loss: 0.110
[55, 1320] loss: 0.126
[55, 1350] loss: 0.086
[55, 1380] loss: 0.090
[55, 1410] loss: 0.087
[55, 1440] loss: 0.102
[55, 1470] loss: 0.094
[55, 1500] loss: 0.074
[55, 1530] loss: 0.099
[55, 1560] loss: 0.128
[55, 1590] loss: 0.091
[55, 1620] loss: 0.098
[55, 1650] loss: 0.114
[55, 1680] loss: 0.118
[55, 1710] loss: 0.071
[55, 1740] loss: 0.092
[55, 1770] loss: 0.118
[55, 1800] loss: 0.064
[55, 1830] loss: 0.091
[55, 1860] loss: 0.096
[55, 1890] loss: 0.118
[55, 1920] loss: 0.084
[55, 1950] loss: 0.133
[55, 1980] loss: 0.117
[55, 2010] loss: 0.097
[55, 2040] loss: 0.125
[55, 2070] loss: 0.085
[55, 2100] loss: 0.092
[55, 2130] loss: 0.092
[55, 2160] loss: 0.145
[55, 2190] loss: 0.103
eval intermediate_models/epoch54_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.976100
epoch 54 accuracy on train set is: 0.9760996001454016
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	NaN	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.989549

epoch 54 accuracy on test set is: 0.9895492548164304

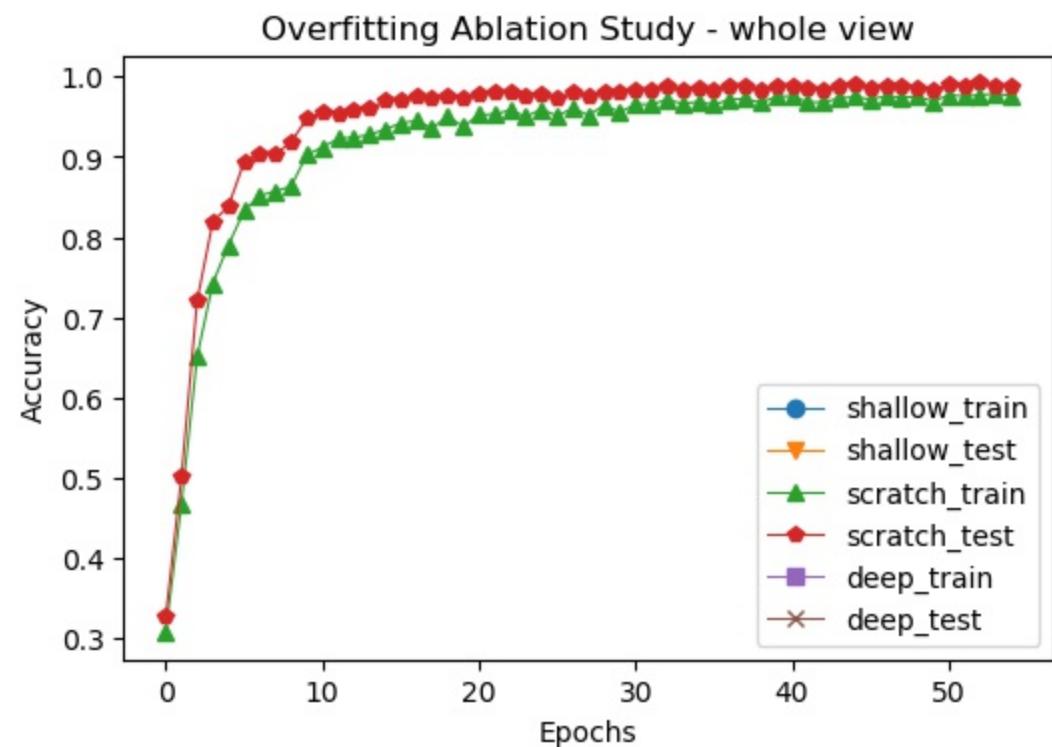
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

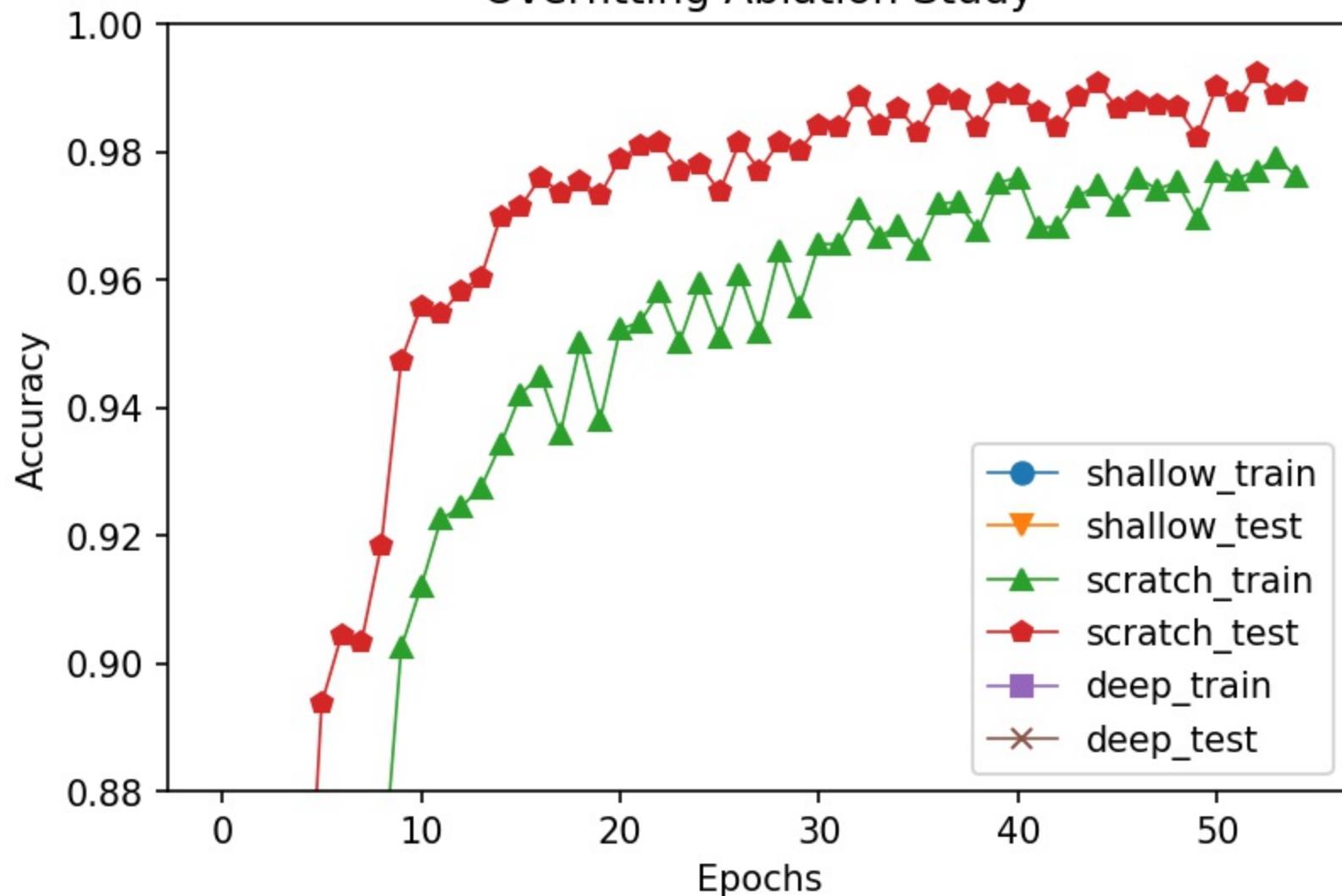
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	NaN	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 54 takes 316 seconds.

[56, 30] loss: 0.073

[56, 60] loss: 0.059

```
[56,    90] loss: 0.095
[56,   120] loss: 0.107
[56,   150] loss: 0.092
[56,   180] loss: 0.062
[56,   210] loss: 0.062
[56,   240] loss: 0.077
[56,   270] loss: 0.120
[56,   300] loss: 0.097
[56,   330] loss: 0.071
[56,   360] loss: 0.066
[56,   390] loss: 0.105
[56,   420] loss: 0.051
[56,   450] loss: 0.068
[56,   480] loss: 0.090
[56,   510] loss: 0.096
[56,   540] loss: 0.117
[56,   570] loss: 0.114
[56,   600] loss: 0.114
[56,   630] loss: 0.106
[56,   660] loss: 0.094
[56,   690] loss: 0.096
[56,   720] loss: 0.086
[56,   750] loss: 0.123
[56,   780] loss: 0.073
[56,   810] loss: 0.117
[56,   840] loss: 0.083
[56,   870] loss: 0.089
[56,   900] loss: 0.079
[56,   930] loss: 0.083
[56,   960] loss: 0.083
[56,   990] loss: 0.119
[56,  1020] loss: 0.116
[56,  1050] loss: 0.113
[56,  1080] loss: 0.089
[56,  1110] loss: 0.079
[56,  1140] loss: 0.096
[56,  1170] loss: 0.089
[56,  1200] loss: 0.093
[56,  1230] loss: 0.076
```

```
[56, 1260] loss: 0.065
[56, 1290] loss: 0.138
[56, 1320] loss: 0.060
[56, 1350] loss: 0.069
[56, 1380] loss: 0.090
[56, 1410] loss: 0.083
[56, 1440] loss: 0.110
[56, 1470] loss: 0.104
[56, 1500] loss: 0.099
[56, 1530] loss: 0.124
[56, 1560] loss: 0.148
[56, 1590] loss: 0.094
[56, 1620] loss: 0.100
[56, 1650] loss: 0.071
[56, 1680] loss: 0.097
[56, 1710] loss: 0.122
[56, 1740] loss: 0.080
[56, 1770] loss: 0.072
[56, 1800] loss: 0.096
[56, 1830] loss: 0.074
[56, 1860] loss: 0.138
[56, 1890] loss: 0.039
[56, 1920] loss: 0.075
[56, 1950] loss: 0.090
[56, 1980] loss: 0.110
[56, 2010] loss: 0.100
[56, 2040] loss: 0.088
[56, 2070] loss: 0.071
[56, 2100] loss: 0.139
[56, 2130] loss: 0.100
[56, 2160] loss: 0.094
[56, 2190] loss: 0.097
eval intermediate_models/epoch55_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.976509
epoch 55 accuracy on train set is: 0.976508542348237
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	NaN	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.987641

epoch 55 accuracy on test set is: 0.9876408578698656

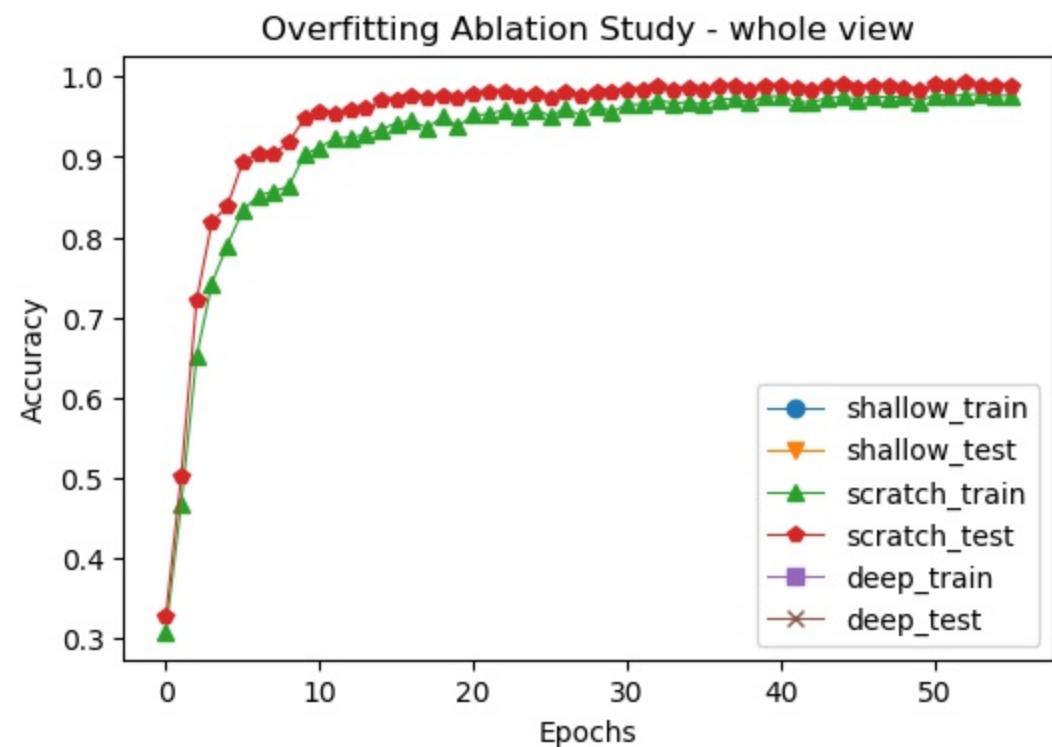
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

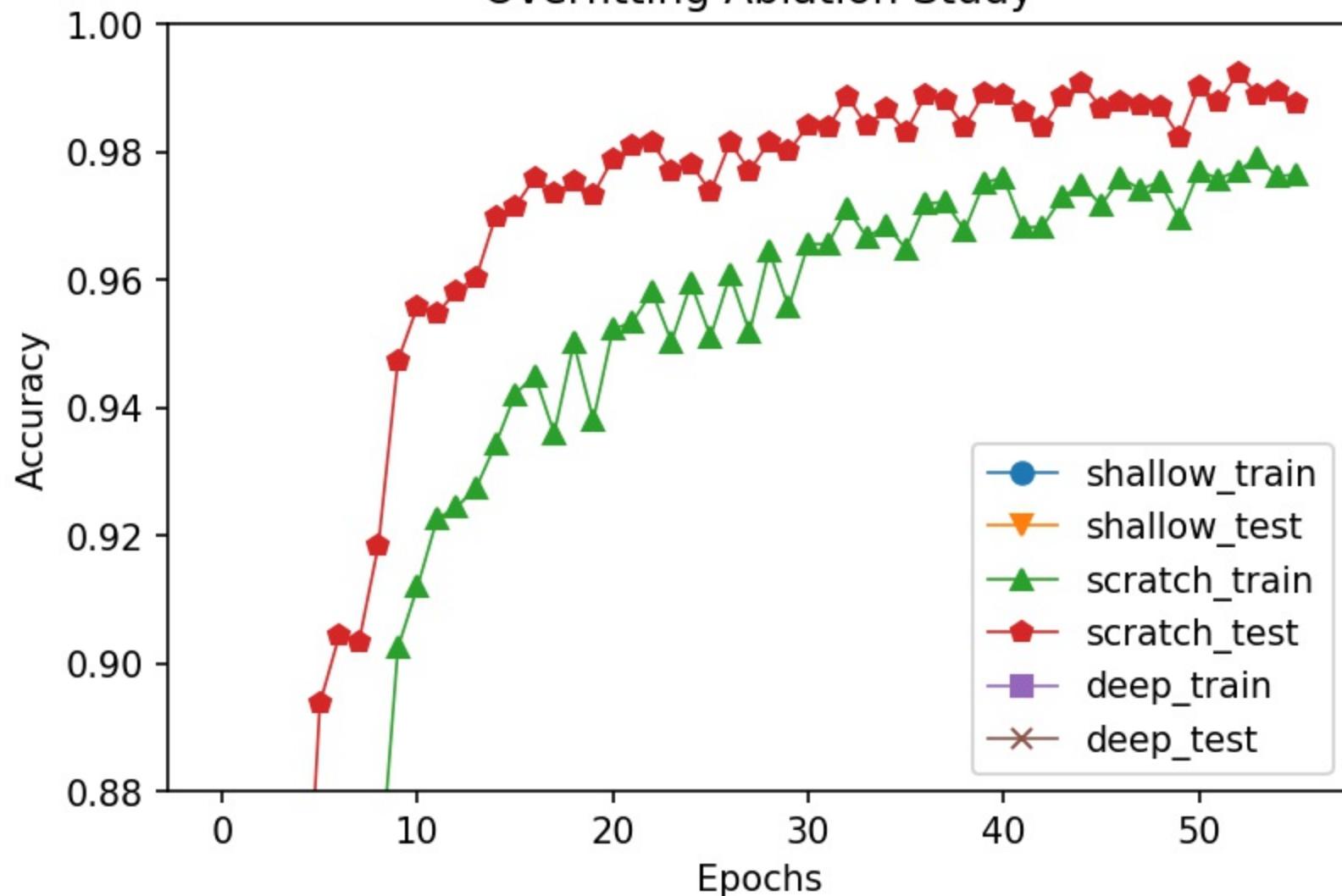
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 55 takes 316 seconds.

[57, 30] loss: 0.104

[57, 60] loss: 0.106

```
[57,    90] loss: 0.126
[57,   120] loss: 0.144
[57,   150] loss: 0.093
[57,   180] loss: 0.070
[57,   210] loss: 0.093
[57,   240] loss: 0.075
[57,   270] loss: 0.086
[57,   300] loss: 0.082
[57,   330] loss: 0.061
[57,   360] loss: 0.100
[57,   390] loss: 0.066
[57,   420] loss: 0.104
[57,   450] loss: 0.066
[57,   480] loss: 0.085
[57,   510] loss: 0.067
[57,   540] loss: 0.086
[57,   570] loss: 0.090
[57,   600] loss: 0.045
[57,   630] loss: 0.108
[57,   660] loss: 0.078
[57,   690] loss: 0.109
[57,   720] loss: 0.090
[57,   750] loss: 0.103
[57,   780] loss: 0.084
[57,   810] loss: 0.082
[57,   840] loss: 0.092
[57,   870] loss: 0.097
[57,   900] loss: 0.070
[57,   930] loss: 0.111
[57,   960] loss: 0.125
[57,   990] loss: 0.067
[57,  1020] loss: 0.098
[57,  1050] loss: 0.078
[57,  1080] loss: 0.091
[57,  1110] loss: 0.077
[57,  1140] loss: 0.082
[57,  1170] loss: 0.077
[57,  1200] loss: 0.127
[57,  1230] loss: 0.076
```

```
[57, 1260] loss: 0.063
[57, 1290] loss: 0.111
[57, 1320] loss: 0.103
[57, 1350] loss: 0.096
[57, 1380] loss: 0.068
[57, 1410] loss: 0.090
[57, 1440] loss: 0.083
[57, 1470] loss: 0.122
[57, 1500] loss: 0.067
[57, 1530] loss: 0.085
[57, 1560] loss: 0.098
[57, 1590] loss: 0.088
[57, 1620] loss: 0.096
[57, 1650] loss: 0.072
[57, 1680] loss: 0.100
[57, 1710] loss: 0.094
[57, 1740] loss: 0.106
[57, 1770] loss: 0.068
[57, 1800] loss: 0.116
[57, 1830] loss: 0.113
[57, 1860] loss: 0.097
[57, 1890] loss: 0.062
[57, 1920] loss: 0.091
[57, 1950] loss: 0.073
[57, 1980] loss: 0.127
[57, 2010] loss: 0.077
[57, 2040] loss: 0.066
[57, 2070] loss: 0.100
[57, 2100] loss: 0.117
[57, 2130] loss: 0.056
[57, 2160] loss: 0.067
[57, 2190] loss: 0.129
eval intermediate_models/epoch56_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.978190
epoch 56 accuracy on train set is: 0.9781897491821155
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990549

epoch 56 accuracy on test set is: 0.9905488913122501

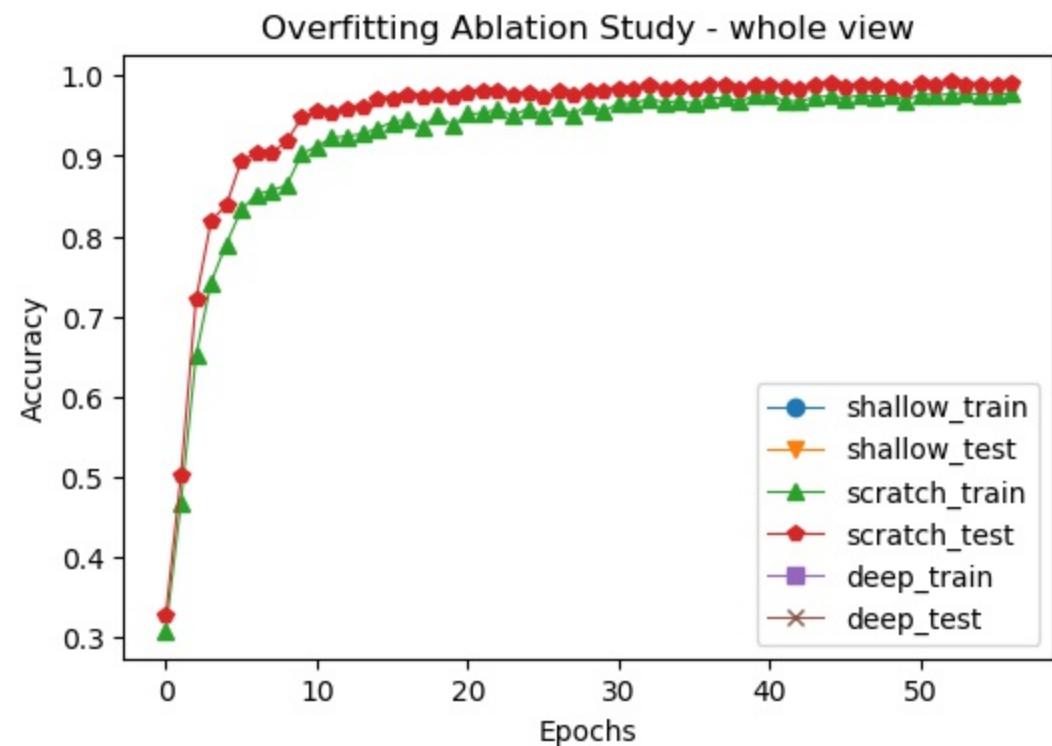
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

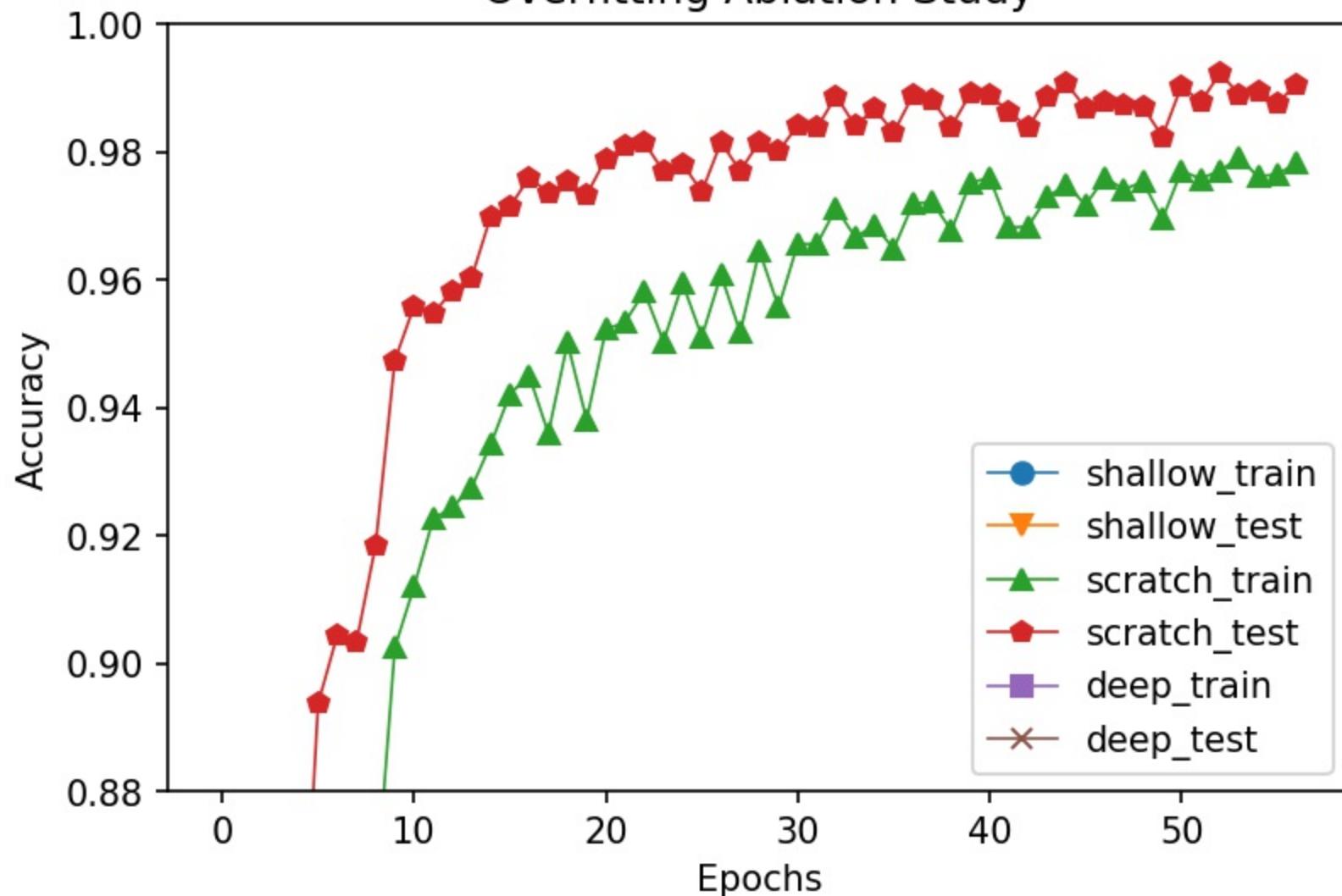
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	NaN	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 56 takes 316 seconds.

[58, 30] loss: 0.087

[58, 60] loss: 0.086

```
[58,    90] loss: 0.103
[58,   120] loss: 0.095
[58,   150] loss: 0.088
[58,   180] loss: 0.103
[58,   210] loss: 0.100
[58,   240] loss: 0.057
[58,   270] loss: 0.089
[58,   300] loss: 0.114
[58,   330] loss: 0.086
[58,   360] loss: 0.163
[58,   390] loss: 0.084
[58,   420] loss: 0.123
[58,   450] loss: 0.093
[58,   480] loss: 0.109
[58,   510] loss: 0.085
[58,   540] loss: 0.048
[58,   570] loss: 0.124
[58,   600] loss: 0.100
[58,   630] loss: 0.097
[58,   660] loss: 0.089
[58,   690] loss: 0.082
[58,   720] loss: 0.060
[58,   750] loss: 0.069
[58,   780] loss: 0.090
[58,   810] loss: 0.095
[58,   840] loss: 0.096
[58,   870] loss: 0.104
[58,   900] loss: 0.104
[58,   930] loss: 0.077
[58,   960] loss: 0.077
[58,   990] loss: 0.089
[58,  1020] loss: 0.069
[58,  1050] loss: 0.101
[58,  1080] loss: 0.096
[58,  1110] loss: 0.090
[58,  1140] loss: 0.111
[58,  1170] loss: 0.086
[58,  1200] loss: 0.075
[58,  1230] loss: 0.064
```

```
[58, 1260] loss: 0.042
[58, 1290] loss: 0.079
[58, 1320] loss: 0.086
[58, 1350] loss: 0.097
[58, 1380] loss: 0.133
[58, 1410] loss: 0.092
[58, 1440] loss: 0.155
[58, 1470] loss: 0.093
[58, 1500] loss: 0.086
[58, 1530] loss: 0.081
[58, 1560] loss: 0.082
[58, 1590] loss: 0.065
[58, 1620] loss: 0.065
[58, 1650] loss: 0.087
[58, 1680] loss: 0.095
[58, 1710] loss: 0.079
[58, 1740] loss: 0.126
[58, 1770] loss: 0.097
[58, 1800] loss: 0.116
[58, 1830] loss: 0.094
[58, 1860] loss: 0.072
[58, 1890] loss: 0.093
[58, 1920] loss: 0.054
[58, 1950] loss: 0.084
[58, 1980] loss: 0.106
[58, 2010] loss: 0.077
[58, 2040] loss: 0.125
[58, 2070] loss: 0.093
[58, 2100] loss: 0.081
[58, 2130] loss: 0.109
[58, 2160] loss: 0.084
[58, 2190] loss: 0.127
eval intermediate_models/epoch57_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.980802
epoch 57 accuracy on train set is: 0.980802435478008
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	NaN	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992185

epoch 57 accuracy on test set is: 0.9921846601235914

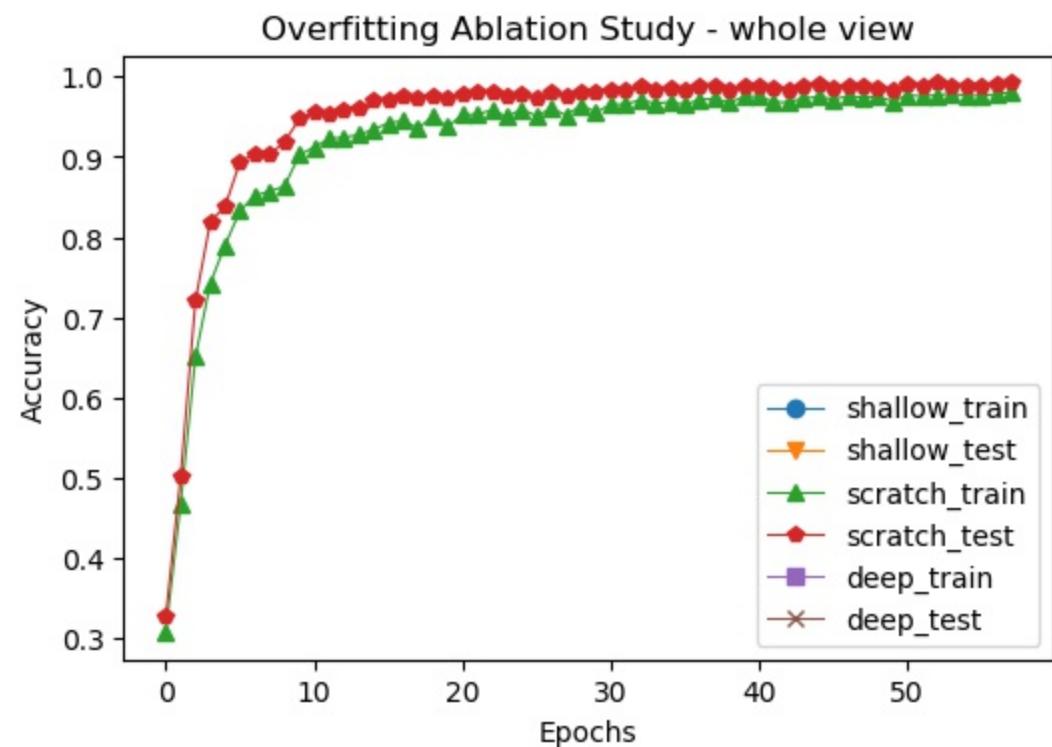
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

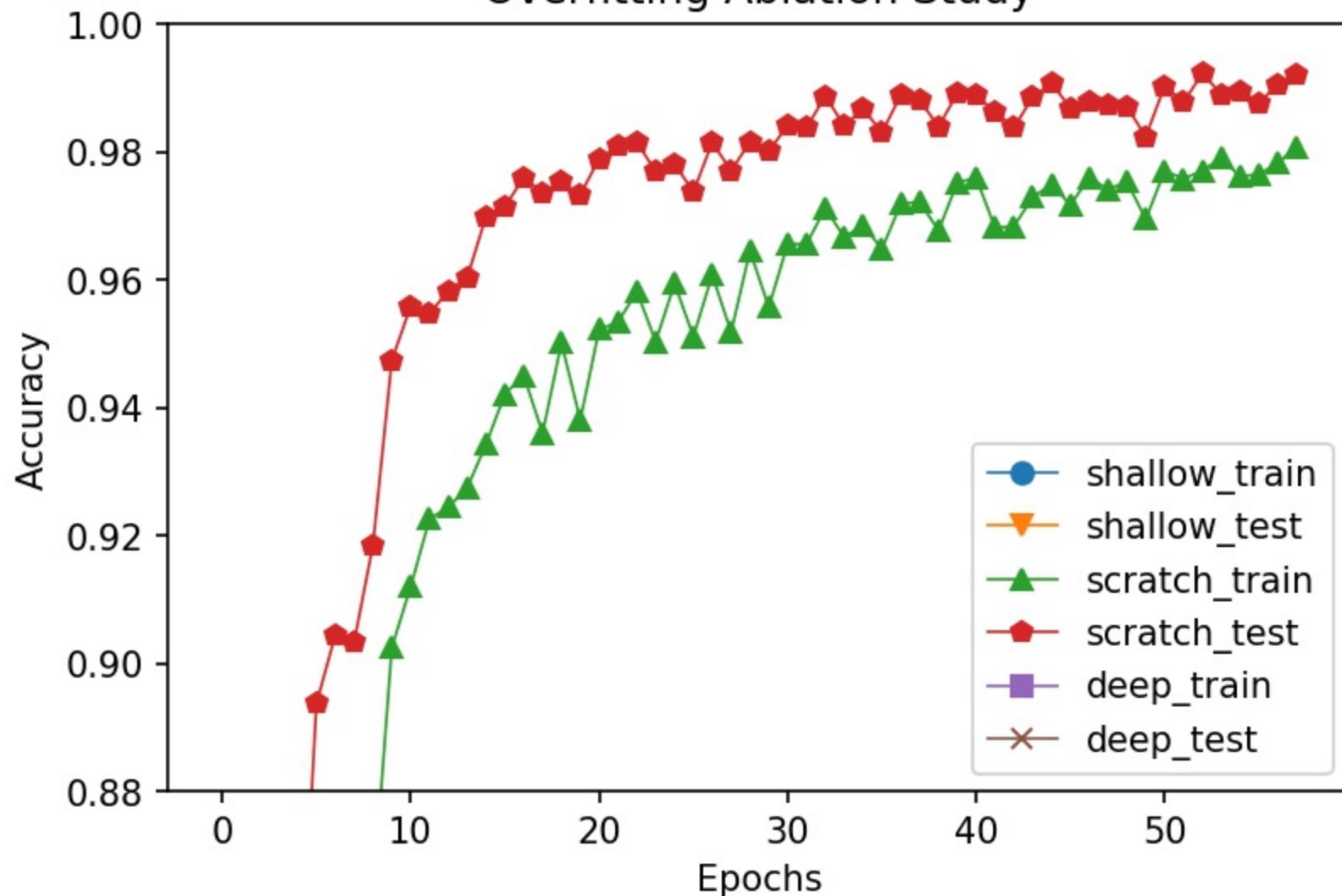
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	NaN	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 57 takes 316 seconds.

[59, 30] loss: 0.069

[59, 60] loss: 0.113

```
[59,    90] loss: 0.129
[59,   120] loss: 0.071
[59,   150] loss: 0.087
[59,   180] loss: 0.084
[59,   210] loss: 0.083
[59,   240] loss: 0.069
[59,   270] loss: 0.092
[59,   300] loss: 0.082
[59,   330] loss: 0.047
[59,   360] loss: 0.043
[59,   390] loss: 0.075
[59,   420] loss: 0.086
[59,   450] loss: 0.097
[59,   480] loss: 0.100
[59,   510] loss: 0.092
[59,   540] loss: 0.090
[59,   570] loss: 0.084
[59,   600] loss: 0.086
[59,   630] loss: 0.123
[59,   660] loss: 0.093
[59,   690] loss: 0.110
[59,   720] loss: 0.074
[59,   750] loss: 0.062
[59,   780] loss: 0.093
[59,   810] loss: 0.082
[59,   840] loss: 0.077
[59,   870] loss: 0.078
[59,   900] loss: 0.097
[59,   930] loss: 0.084
[59,   960] loss: 0.074
[59,   990] loss: 0.068
[59,  1020] loss: 0.095
[59,  1050] loss: 0.094
[59,  1080] loss: 0.088
[59,  1110] loss: 0.117
[59,  1140] loss: 0.105
[59,  1170] loss: 0.078
[59,  1200] loss: 0.091
[59,  1230] loss: 0.079
```

```
[59, 1260] loss: 0.114
[59, 1290] loss: 0.104
[59, 1320] loss: 0.076
[59, 1350] loss: 0.103
[59, 1380] loss: 0.092
[59, 1410] loss: 0.082
[59, 1440] loss: 0.094
[59, 1470] loss: 0.101
[59, 1500] loss: 0.133
[59, 1530] loss: 0.075
[59, 1560] loss: 0.128
[59, 1590] loss: 0.098
[59, 1620] loss: 0.064
[59, 1650] loss: 0.093
[59, 1680] loss: 0.107
[59, 1710] loss: 0.073
[59, 1740] loss: 0.069
[59, 1770] loss: 0.077
[59, 1800] loss: 0.105
[59, 1830] loss: 0.080
[59, 1860] loss: 0.105
[59, 1890] loss: 0.078
[59, 1920] loss: 0.066
[59, 1950] loss: 0.068
[59, 1980] loss: 0.074
[59, 2010] loss: 0.057
[59, 2040] loss: 0.077
[59, 2070] loss: 0.081
[59, 2100] loss: 0.074
[59, 2130] loss: 0.064
[59, 2160] loss: 0.081
[59, 2190] loss: 0.102
eval intermediate_models/epoch58_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.976713
epoch 58 accuracy on train set is: 0.9767130134496547
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	NaN	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988459

epoch 58 accuracy on test set is: 0.9884587422755362

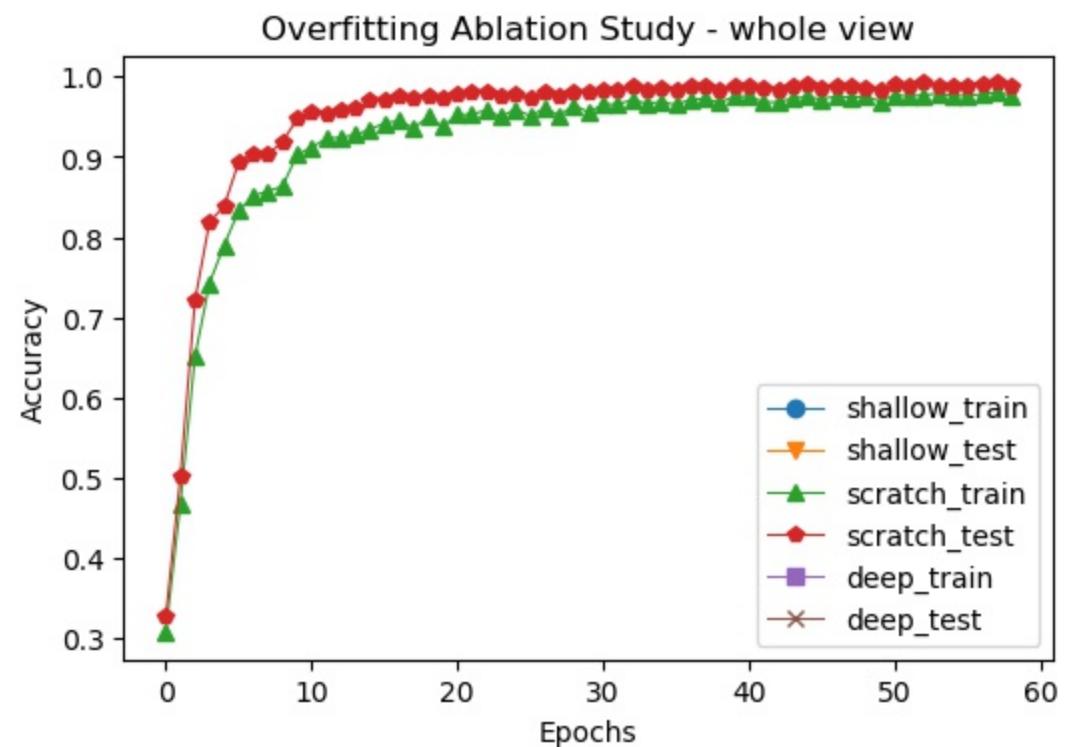
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

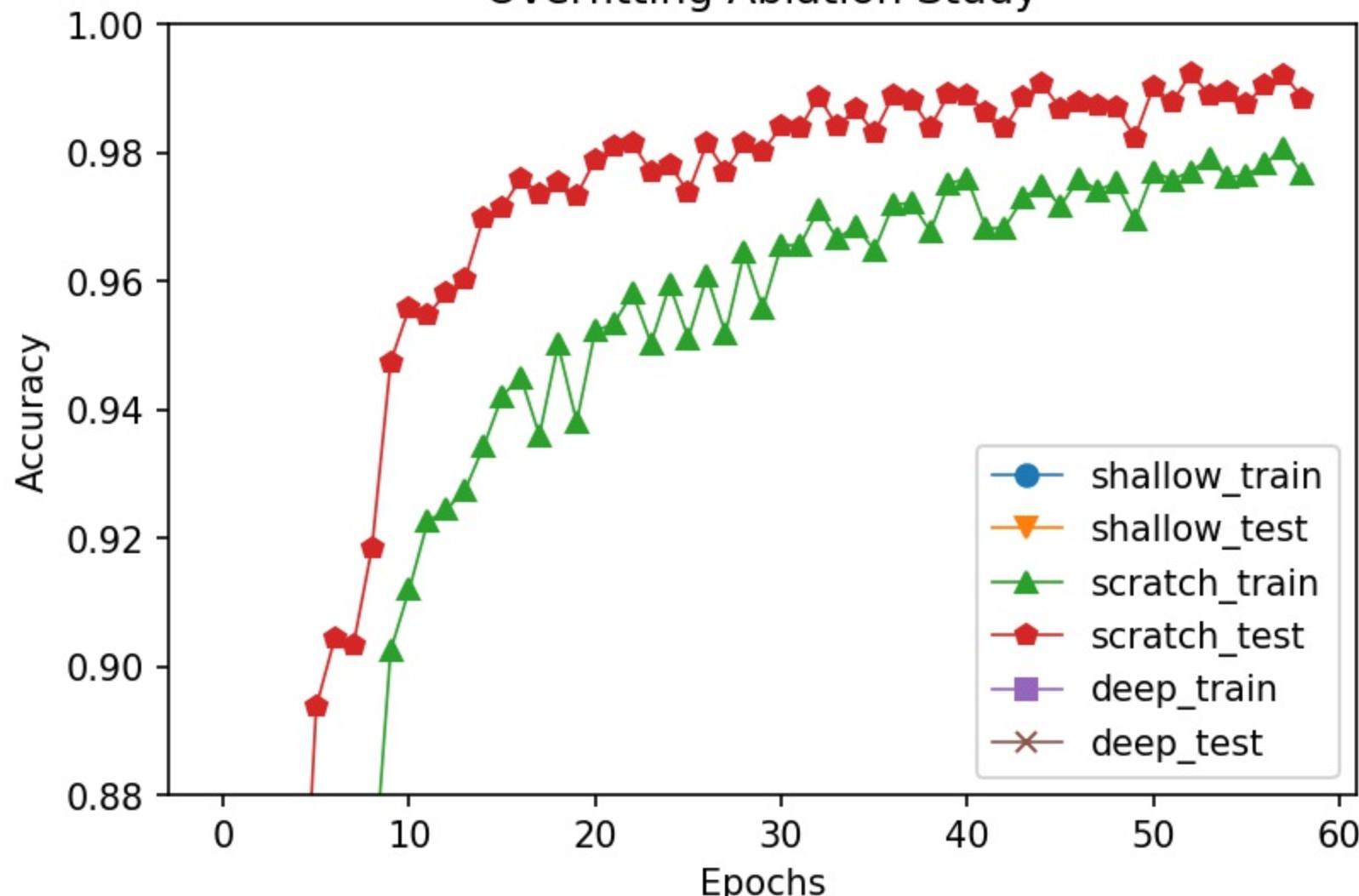
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 58 takes 316 seconds.

[60, 30] loss: 0.098

[60, 60] loss: 0.092

```
[60,    90] loss: 0.074
[60,   120] loss: 0.081
[60,   150] loss: 0.067
[60,   180] loss: 0.091
[60,   210] loss: 0.086
[60,   240] loss: 0.104
[60,   270] loss: 0.108
[60,   300] loss: 0.102
[60,   330] loss: 0.082
[60,   360] loss: 0.115
[60,   390] loss: 0.106
[60,   420] loss: 0.051
[60,   450] loss: 0.071
[60,   480] loss: 0.108
[60,   510] loss: 0.072
[60,   540] loss: 0.153
[60,   570] loss: 0.083
[60,   600] loss: 0.079
[60,   630] loss: 0.087
[60,   660] loss: 0.067
[60,   690] loss: 0.088
[60,   720] loss: 0.101
[60,   750] loss: 0.076
[60,   780] loss: 0.079
[60,   810] loss: 0.065
[60,   840] loss: 0.100
[60,   870] loss: 0.043
[60,   900] loss: 0.076
[60,   930] loss: 0.101
[60,   960] loss: 0.095
[60,   990] loss: 0.099
[60,  1020] loss: 0.078
[60,  1050] loss: 0.076
[60,  1080] loss: 0.068
[60,  1110] loss: 0.064
[60,  1140] loss: 0.088
[60,  1170] loss: 0.069
[60,  1200] loss: 0.065
[60,  1230] loss: 0.117
```

```
[60, 1260] loss: 0.076
[60, 1290] loss: 0.104
[60, 1320] loss: 0.129
[60, 1350] loss: 0.096
[60, 1380] loss: 0.087
[60, 1410] loss: 0.074
[60, 1440] loss: 0.091
[60, 1470] loss: 0.084
[60, 1500] loss: 0.074
[60, 1530] loss: 0.071
[60, 1560] loss: 0.093
[60, 1590] loss: 0.066
[60, 1620] loss: 0.066
[60, 1650] loss: 0.082
[60, 1680] loss: 0.105
[60, 1710] loss: 0.105
[60, 1740] loss: 0.068
[60, 1770] loss: 0.108
[60, 1800] loss: 0.146
[60, 1830] loss: 0.097
[60, 1860] loss: 0.080
[60, 1890] loss: 0.071
[60, 1920] loss: 0.080
[60, 1950] loss: 0.113
[60, 1980] loss: 0.059
[60, 2010] loss: 0.069
[60, 2040] loss: 0.099
[60, 2070] loss: 0.118
[60, 2100] loss: 0.089
[60, 2130] loss: 0.132
[60, 2160] loss: 0.093
[60, 2190] loss: 0.088
eval intermediate_models/epoch59_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.979099
epoch 59 accuracy on train set is: 0.9790985096328608
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990458

epoch 59 accuracy on test set is: 0.9904580152671756

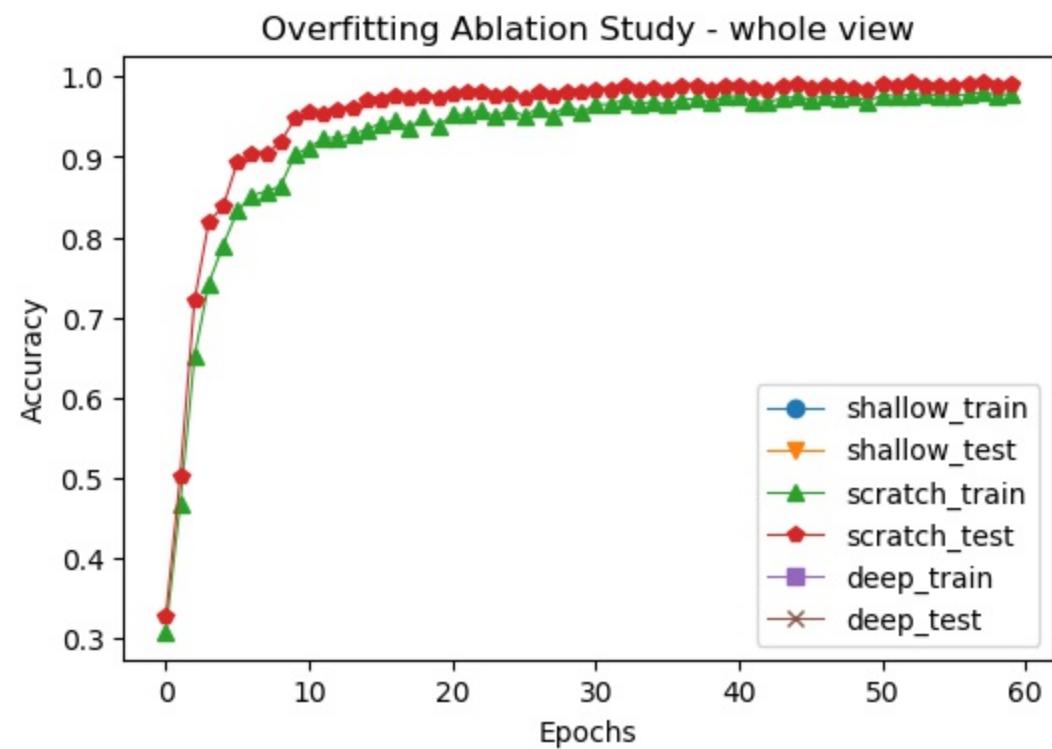
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

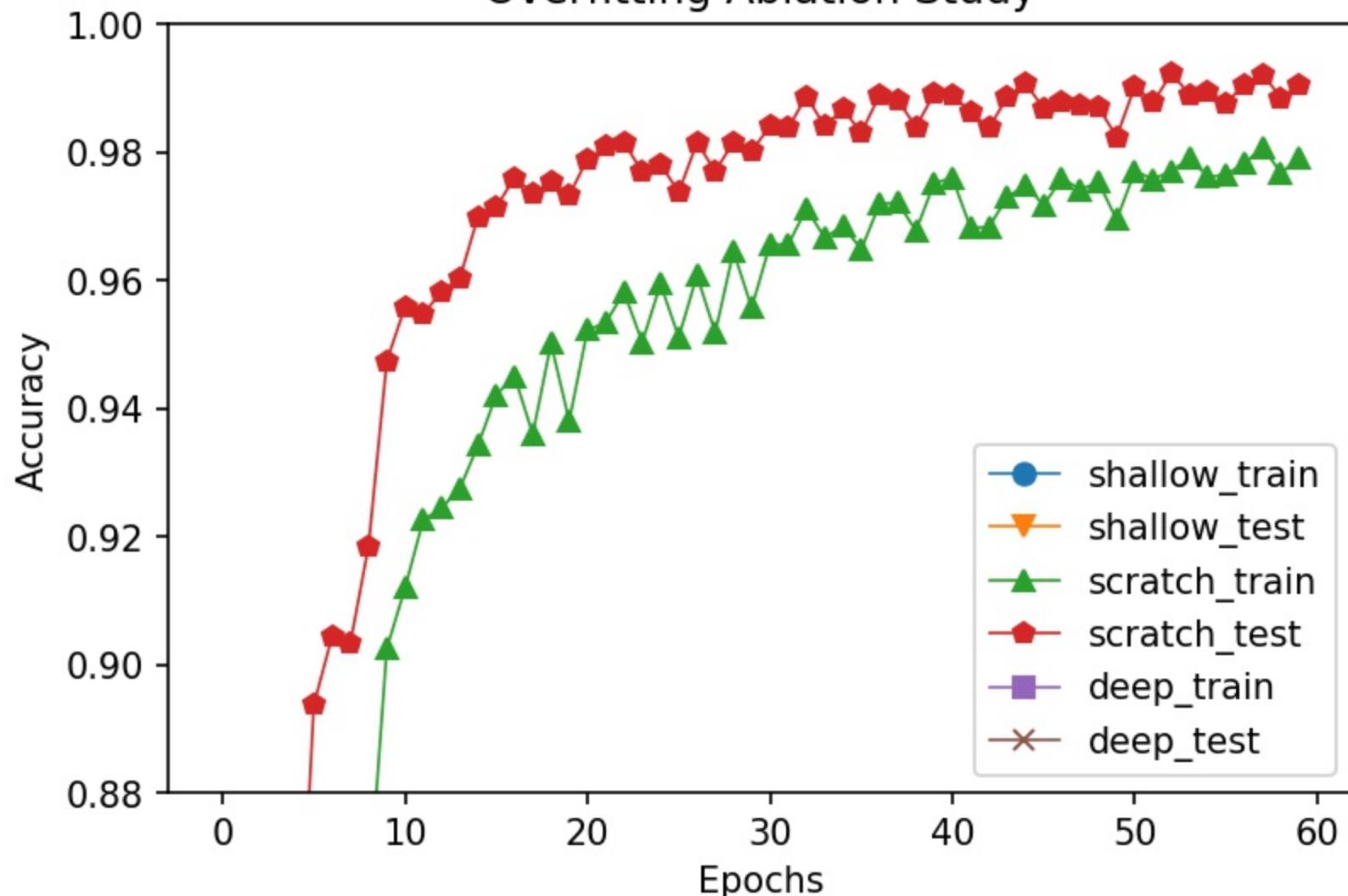
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 59 takes 316 seconds.

[61, 30] loss: 0.095

[61, 60] loss: 0.068

```
[61,    90] loss: 0.069
[61,   120] loss: 0.105
[61,   150] loss: 0.071
[61,   180] loss: 0.086
[61,   210] loss: 0.103
[61,   240] loss: 0.109
[61,   270] loss: 0.126
[61,   300] loss: 0.096
[61,   330] loss: 0.102
[61,   360] loss: 0.123
[61,   390] loss: 0.095
[61,   420] loss: 0.109
[61,   450] loss: 0.078
[61,   480] loss: 0.082
[61,   510] loss: 0.073
[61,   540] loss: 0.101
[61,   570] loss: 0.123
[61,   600] loss: 0.082
[61,   630] loss: 0.074
[61,   660] loss: 0.076
[61,   690] loss: 0.080
[61,   720] loss: 0.065
[61,   750] loss: 0.071
[61,   780] loss: 0.084
[61,   810] loss: 0.083
[61,   840] loss: 0.085
[61,   870] loss: 0.097
[61,   900] loss: 0.077
[61,   930] loss: 0.084
[61,   960] loss: 0.066
[61,   990] loss: 0.065
[61,  1020] loss: 0.062
[61,  1050] loss: 0.051
[61,  1080] loss: 0.100
[61,  1110] loss: 0.095
[61,  1140] loss: 0.065
[61,  1170] loss: 0.065
[61,  1200] loss: 0.084
[61,  1230] loss: 0.090
```

```
[61, 1260] loss: 0.067
[61, 1290] loss: 0.091
[61, 1320] loss: 0.087
[61, 1350] loss: 0.103
[61, 1380] loss: 0.096
[61, 1410] loss: 0.079
[61, 1440] loss: 0.081
[61, 1470] loss: 0.068
[61, 1500] loss: 0.065
[61, 1530] loss: 0.070
[61, 1560] loss: 0.084
[61, 1590] loss: 0.063
[61, 1620] loss: 0.057
[61, 1650] loss: 0.117
[61, 1680] loss: 0.118
[61, 1710] loss: 0.106
[61, 1740] loss: 0.073
[61, 1770] loss: 0.069
[61, 1800] loss: 0.101
[61, 1830] loss: 0.095
[61, 1860] loss: 0.134
[61, 1890] loss: 0.070
[61, 1920] loss: 0.084
[61, 1950] loss: 0.087
[61, 1980] loss: 0.098
[61, 2010] loss: 0.102
[61, 2040] loss: 0.077
[61, 2070] loss: 0.051
[61, 2100] loss: 0.094
[61, 2130] loss: 0.060
[61, 2160] loss: 0.064
[61, 2190] loss: 0.083
eval intermediate_models/epoch60_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.979939
epoch 60 accuracy on train set is: 0.9799391130498001
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990004

epoch 60 accuracy on test set is: 0.990003635041803

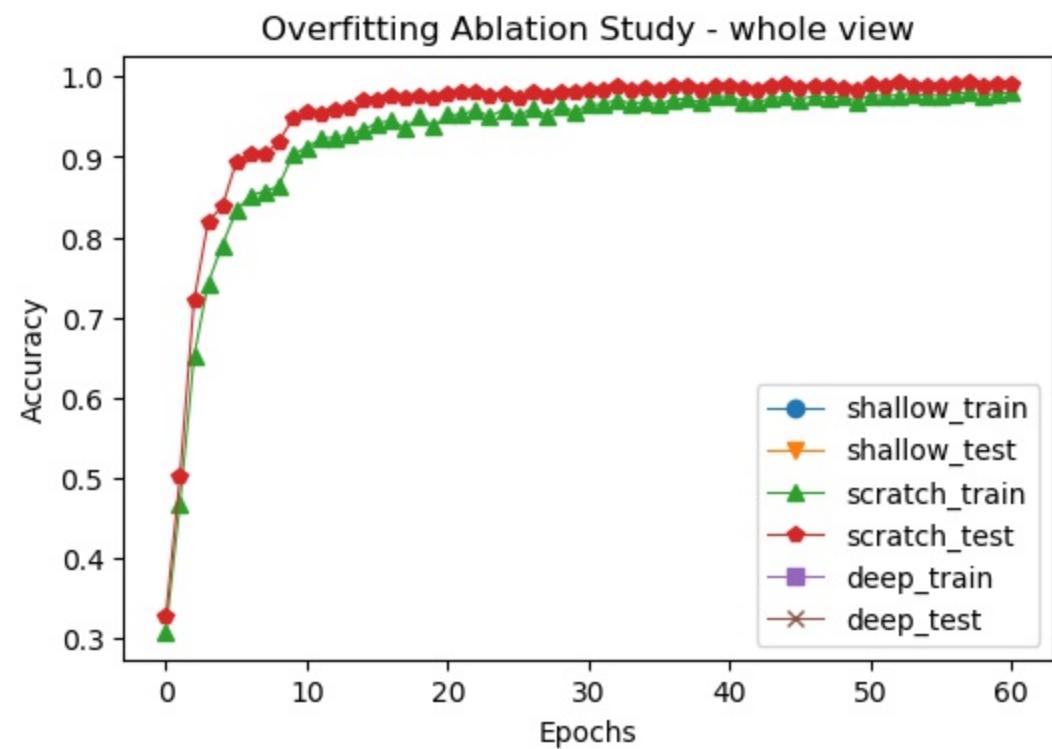
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

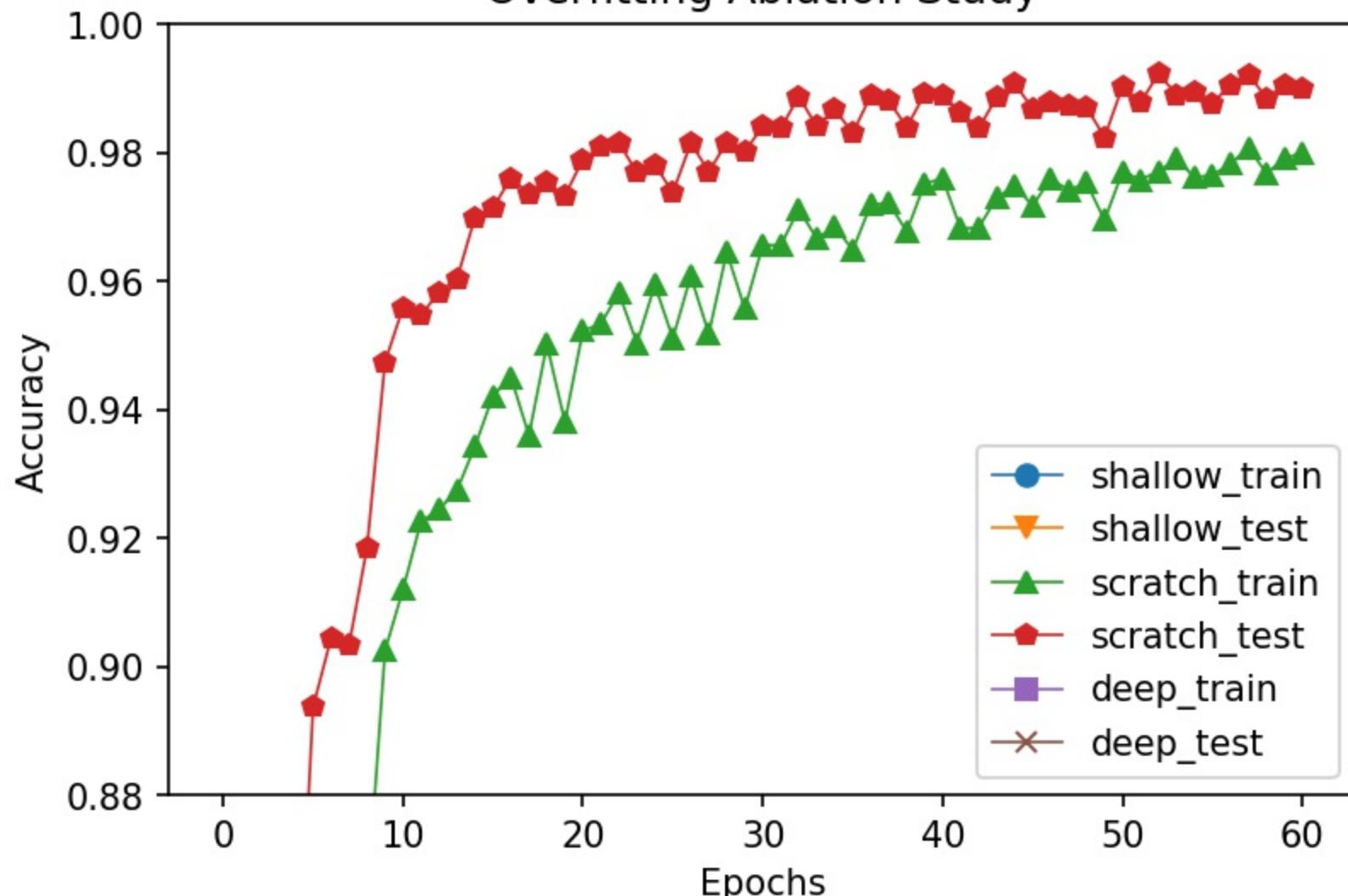
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 60 takes 316 seconds.

[62, 30] loss: 0.077

[62, 60] loss: 0.139

```
[62,    90] loss: 0.133
[62,   120] loss: 0.081
[62,   150] loss: 0.092
[62,   180] loss: 0.063
[62,   210] loss: 0.069
[62,   240] loss: 0.056
[62,   270] loss: 0.052
[62,   300] loss: 0.140
[62,   330] loss: 0.085
[62,   360] loss: 0.107
[62,   390] loss: 0.065
[62,   420] loss: 0.080
[62,   450] loss: 0.046
[62,   480] loss: 0.110
[62,   510] loss: 0.077
[62,   540] loss: 0.117
[62,   570] loss: 0.076
[62,   600] loss: 0.051
[62,   630] loss: 0.100
[62,   660] loss: 0.070
[62,   690] loss: 0.048
[62,   720] loss: 0.109
[62,   750] loss: 0.087
[62,   780] loss: 0.087
[62,   810] loss: 0.073
[62,   840] loss: 0.113
[62,   870] loss: 0.098
[62,   900] loss: 0.055
[62,   930] loss: 0.076
[62,   960] loss: 0.107
[62,   990] loss: 0.069
[62,  1020] loss: 0.107
[62,  1050] loss: 0.120
[62,  1080] loss: 0.067
[62,  1110] loss: 0.087
[62,  1140] loss: 0.126
[62,  1170] loss: 0.096
[62,  1200] loss: 0.071
[62,  1230] loss: 0.069
```

```
[62, 1260] loss: 0.069
[62, 1290] loss: 0.117
[62, 1320] loss: 0.081
[62, 1350] loss: 0.093
[62, 1380] loss: 0.058
[62, 1410] loss: 0.077
[62, 1440] loss: 0.079
[62, 1470] loss: 0.122
[62, 1500] loss: 0.056
[62, 1530] loss: 0.062
[62, 1560] loss: 0.061
[62, 1590] loss: 0.079
[62, 1620] loss: 0.063
[62, 1650] loss: 0.074
[62, 1680] loss: 0.079
[62, 1710] loss: 0.113
[62, 1740] loss: 0.062
[62, 1770] loss: 0.097
[62, 1800] loss: 0.098
[62, 1830] loss: 0.068
[62, 1860] loss: 0.075
[62, 1890] loss: 0.089
[62, 1920] loss: 0.115
[62, 1950] loss: 0.056
[62, 1980] loss: 0.104
[62, 2010] loss: 0.106
[62, 2040] loss: 0.072
[62, 2070] loss: 0.057
[62, 2100] loss: 0.095
[62, 2130] loss: 0.094
[62, 2160] loss: 0.088
[62, 2190] loss: 0.111
eval intermediate_models/epoch61_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.976440
epoch 61 accuracy on train set is: 0.9764403853144311
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990004

epoch 61 accuracy on test set is: 0.990003635041803

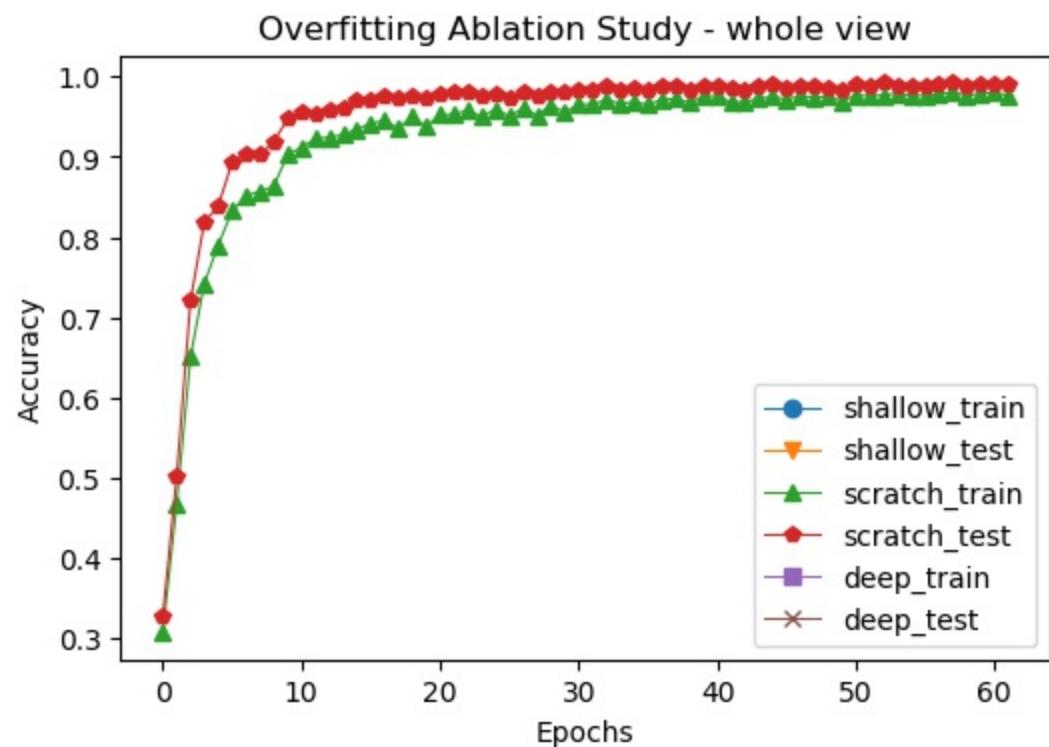
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

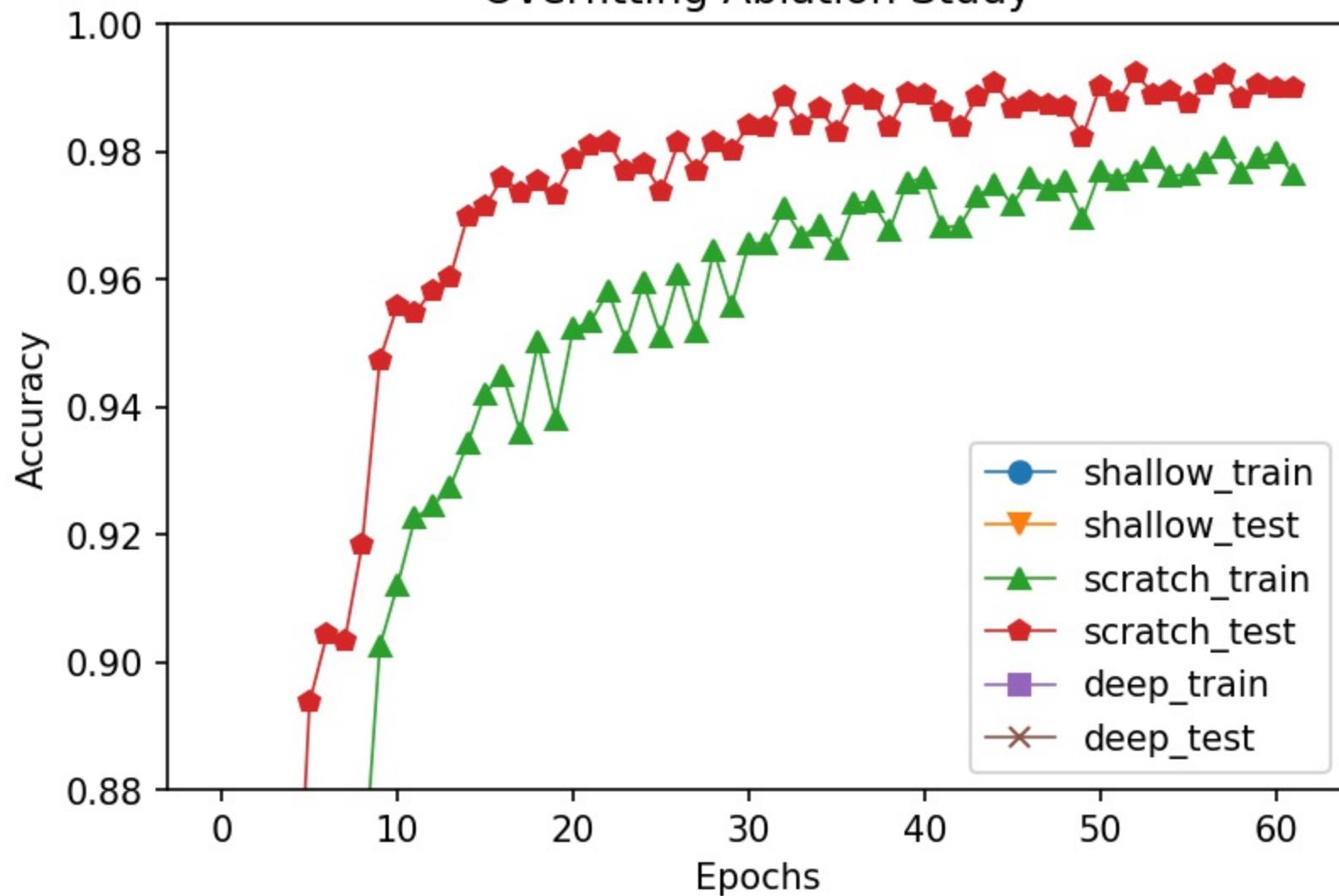
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 61 takes 316 seconds.

[63, 30] loss: 0.074

[63, 60] loss: 0.145

```
[63,    90] loss: 0.114
[63,   120] loss: 0.077
[63,   150] loss: 0.076
[63,   180] loss: 0.094
[63,   210] loss: 0.074
[63,   240] loss: 0.083
[63,   270] loss: 0.089
[63,   300] loss: 0.086
[63,   330] loss: 0.061
[63,   360] loss: 0.098
[63,   390] loss: 0.109
[63,   420] loss: 0.122
[63,   450] loss: 0.115
[63,   480] loss: 0.099
[63,   510] loss: 0.100
[63,   540] loss: 0.068
[63,   570] loss: 0.080
[63,   600] loss: 0.091
[63,   630] loss: 0.138
[63,   660] loss: 0.112
[63,   690] loss: 0.115
[63,   720] loss: 0.065
[63,   750] loss: 0.084
[63,   780] loss: 0.089
[63,   810] loss: 0.092
[63,   840] loss: 0.087
[63,   870] loss: 0.142
[63,   900] loss: 0.106
[63,   930] loss: 0.083
[63,   960] loss: 0.069
[63,   990] loss: 0.063
[63,  1020] loss: 0.101
[63,  1050] loss: 0.088
[63,  1080] loss: 0.103
[63,  1110] loss: 0.162
[63,  1140] loss: 0.113
[63,  1170] loss: 0.095
[63,  1200] loss: 0.104
[63,  1230] loss: 0.075
```

```
[63, 1260] loss: 0.115
[63, 1290] loss: 0.085
[63, 1320] loss: 0.073
[63, 1350] loss: 0.060
[63, 1380] loss: 0.070
[63, 1410] loss: 0.103
[63, 1440] loss: 0.070
[63, 1470] loss: 0.101
[63, 1500] loss: 0.075
[63, 1530] loss: 0.083
[63, 1560] loss: 0.072
[63, 1590] loss: 0.072
[63, 1620] loss: 0.096
[63, 1650] loss: 0.072
[63, 1680] loss: 0.112
[63, 1710] loss: 0.097
[63, 1740] loss: 0.105
[63, 1770] loss: 0.070
[63, 1800] loss: 0.068
[63, 1830] loss: 0.070
[63, 1860] loss: 0.091
[63, 1890] loss: 0.074
[63, 1920] loss: 0.080
[63, 1950] loss: 0.077
[63, 1980] loss: 0.071
[63, 2010] loss: 0.074
[63, 2040] loss: 0.102
[63, 2070] loss: 0.122
[63, 2100] loss: 0.075
[63, 2130] loss: 0.095
[63, 2160] loss: 0.110
[63, 2190] loss: 0.083
evf intermediate_models/epoch62_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.982075
epoch 62 accuracy on train set is: 0.9820747001090513
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992003

epoch 62 accuracy on test set is: 0.9920029080334424

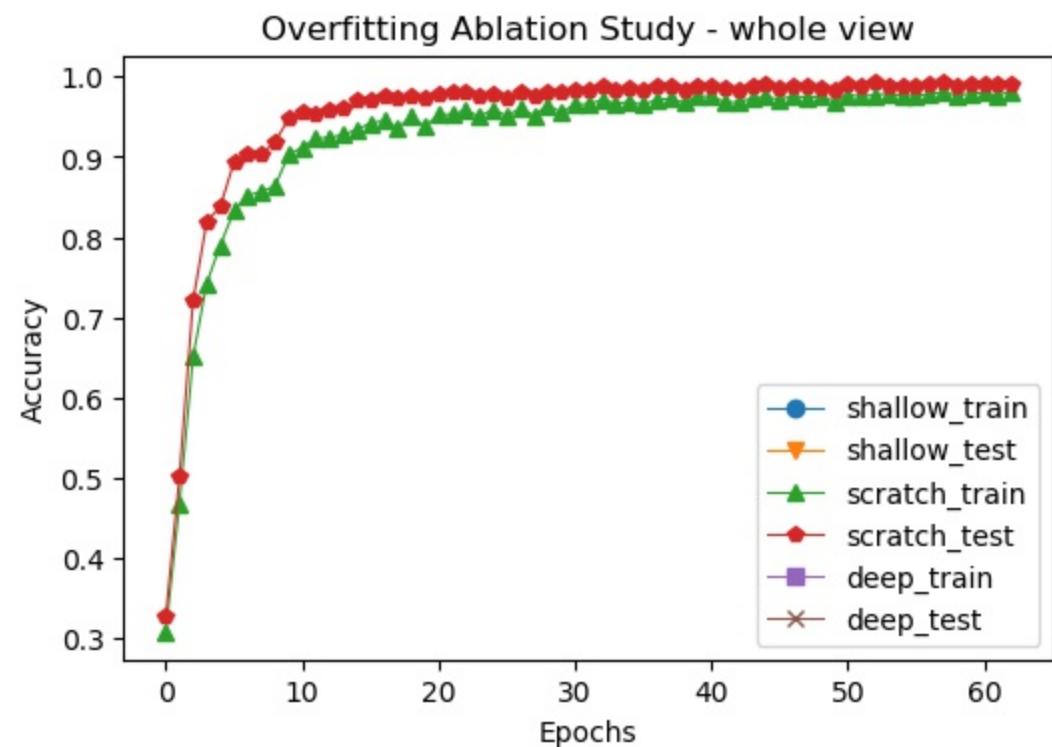
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

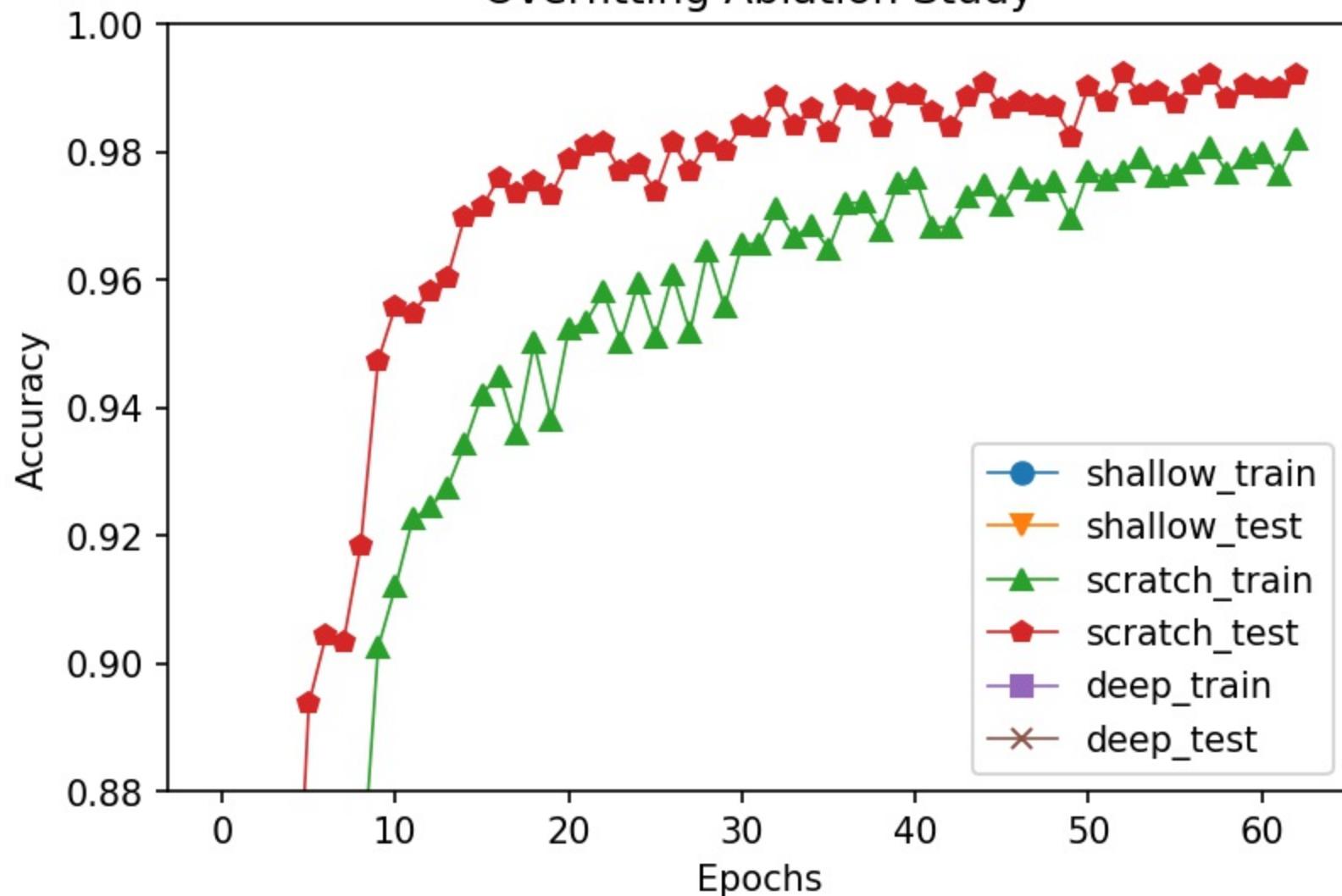
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 62 takes 316 seconds.

[64, 30] loss: 0.068

[64, 60] loss: 0.085

```
[64,    90] loss: 0.056
[64,   120] loss: 0.054
[64,   150] loss: 0.064
[64,   180] loss: 0.084
[64,   210] loss: 0.056
[64,   240] loss: 0.088
[64,   270] loss: 0.091
[64,   300] loss: 0.111
[64,   330] loss: 0.094
[64,   360] loss: 0.108
[64,   390] loss: 0.083
[64,   420] loss: 0.129
[64,   450] loss: 0.089
[64,   480] loss: 0.121
[64,   510] loss: 0.072
[64,   540] loss: 0.114
[64,   570] loss: 0.059
[64,   600] loss: 0.082
[64,   630] loss: 0.063
[64,   660] loss: 0.082
[64,   690] loss: 0.121
[64,   720] loss: 0.098
[64,   750] loss: 0.089
[64,   780] loss: 0.133
[64,   810] loss: 0.081
[64,   840] loss: 0.156
[64,   870] loss: 0.085
[64,   900] loss: 0.081
[64,   930] loss: 0.098
[64,   960] loss: 0.087
[64,   990] loss: 0.071
[64,  1020] loss: 0.115
[64,  1050] loss: 0.077
[64,  1080] loss: 0.089
[64,  1110] loss: 0.062
[64,  1140] loss: 0.072
[64,  1170] loss: 0.060
[64,  1200] loss: 0.116
[64,  1230] loss: 0.084
```

```
[64, 1260] loss: 0.122
[64, 1290] loss: 0.075
[64, 1320] loss: 0.085
[64, 1350] loss: 0.079
[64, 1380] loss: 0.113
[64, 1410] loss: 0.086
[64, 1440] loss: 0.075
[64, 1470] loss: 0.084
[64, 1500] loss: 0.113
[64, 1530] loss: 0.104
[64, 1560] loss: 0.077
[64, 1590] loss: 0.089
[64, 1620] loss: 0.110
[64, 1650] loss: 0.071
[64, 1680] loss: 0.080
[64, 1710] loss: 0.086
[64, 1740] loss: 0.101
[64, 1770] loss: 0.098
[64, 1800] loss: 0.075
[64, 1830] loss: 0.087
[64, 1860] loss: 0.069
[64, 1890] loss: 0.072
[64, 1920] loss: 0.070
[64, 1950] loss: 0.060
[64, 1980] loss: 0.079
[64, 2010] loss: 0.094
[64, 2040] loss: 0.112
[64, 2070] loss: 0.069
[64, 2100] loss: 0.060
[64, 2130] loss: 0.063
[64, 2160] loss: 0.049
[64, 2190] loss: 0.084
eval intermediate_models/epoch63_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.980348
epoch 63 accuracy on train set is: 0.9803480552526354
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991367

epoch 63 accuracy on test set is: 0.9913667757179208

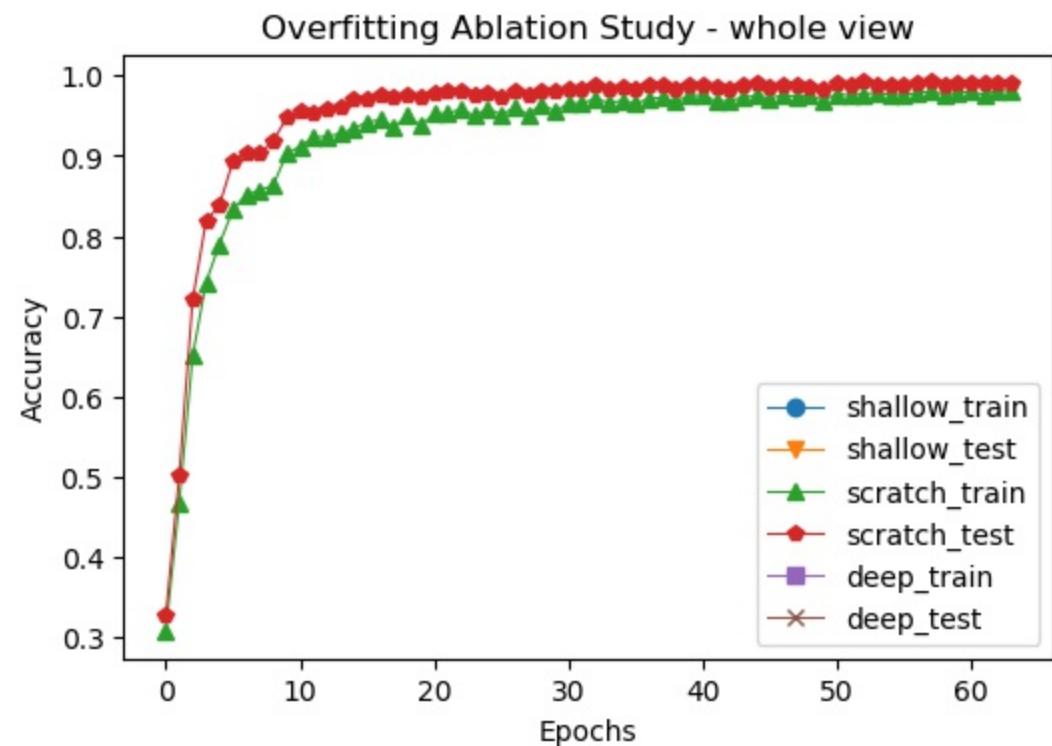
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

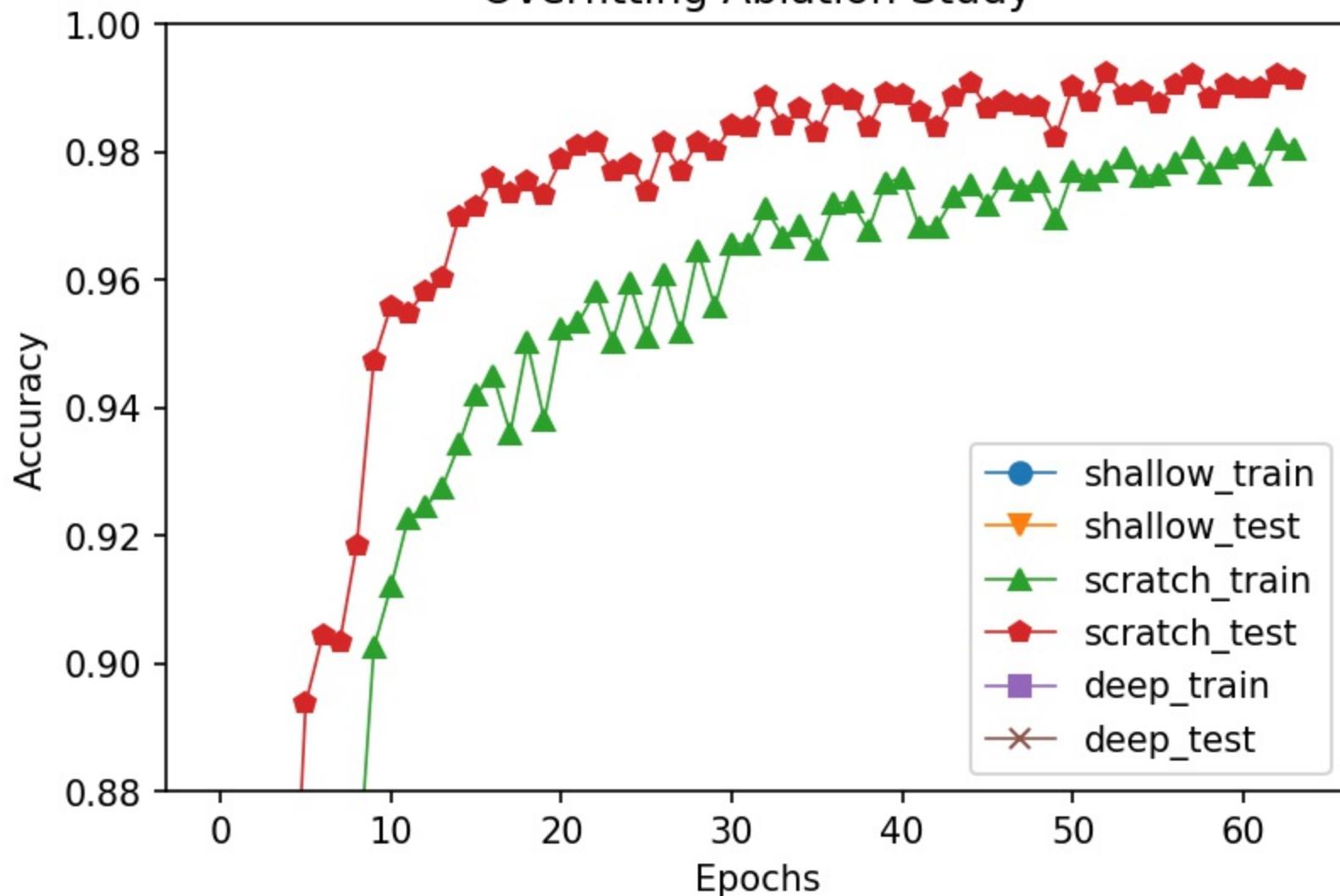
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 63 takes 316 seconds.

[65, 30] loss: 0.071

[65, 60] loss: 0.064

```
[65,    90] loss: 0.076
[65,   120] loss: 0.063
[65,   150] loss: 0.058
[65,   180] loss: 0.094
[65,   210] loss: 0.073
[65,   240] loss: 0.071
[65,   270] loss: 0.072
[65,   300] loss: 0.066
[65,   330] loss: 0.128
[65,   360] loss: 0.106
[65,   390] loss: 0.096
[65,   420] loss: 0.102
[65,   450] loss: 0.103
[65,   480] loss: 0.075
[65,   510] loss: 0.052
[65,   540] loss: 0.063
[65,   570] loss: 0.065
[65,   600] loss: 0.105
[65,   630] loss: 0.092
[65,   660] loss: 0.130
[65,   690] loss: 0.129
[65,   720] loss: 0.096
[65,   750] loss: 0.115
[65,   780] loss: 0.085
[65,   810] loss: 0.062
[65,   840] loss: 0.087
[65,   870] loss: 0.066
[65,   900] loss: 0.099
[65,   930] loss: 0.096
[65,   960] loss: 0.099
[65,   990] loss: 0.081
[65,  1020] loss: 0.080
[65,  1050] loss: 0.059
[65,  1080] loss: 0.072
[65,  1110] loss: 0.058
[65,  1140] loss: 0.071
[65,  1170] loss: 0.059
[65,  1200] loss: 0.109
[65,  1230] loss: 0.081
```

```
[65, 1260] loss: 0.096
[65, 1290] loss: 0.089
[65, 1320] loss: 0.089
[65, 1350] loss: 0.093
[65, 1380] loss: 0.075
[65, 1410] loss: 0.062
[65, 1440] loss: 0.075
[65, 1470] loss: 0.072
[65, 1500] loss: 0.094
[65, 1530] loss: 0.072
[65, 1560] loss: 0.078
[65, 1590] loss: 0.071
[65, 1620] loss: 0.060
[65, 1650] loss: 0.065
[65, 1680] loss: 0.107
[65, 1710] loss: 0.125
[65, 1740] loss: 0.059
[65, 1770] loss: 0.080
[65, 1800] loss: 0.038
[65, 1830] loss: 0.078
[65, 1860] loss: 0.095
[65, 1890] loss: 0.061
[65, 1920] loss: 0.076
[65, 1950] loss: 0.080
[65, 1980] loss: 0.078
[65, 2010] loss: 0.070
[65, 2040] loss: 0.083
[65, 2070] loss: 0.062
[65, 2100] loss: 0.082
[65, 2130] loss: 0.057
[65, 2160] loss: 0.065
[65, 2190] loss: 0.077
eval intermediate_models/epoch64_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.981325
epoch 64 accuracy on train set is: 0.9813249727371864
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992457

epoch 64 accuracy on test set is: 0.992457288258815

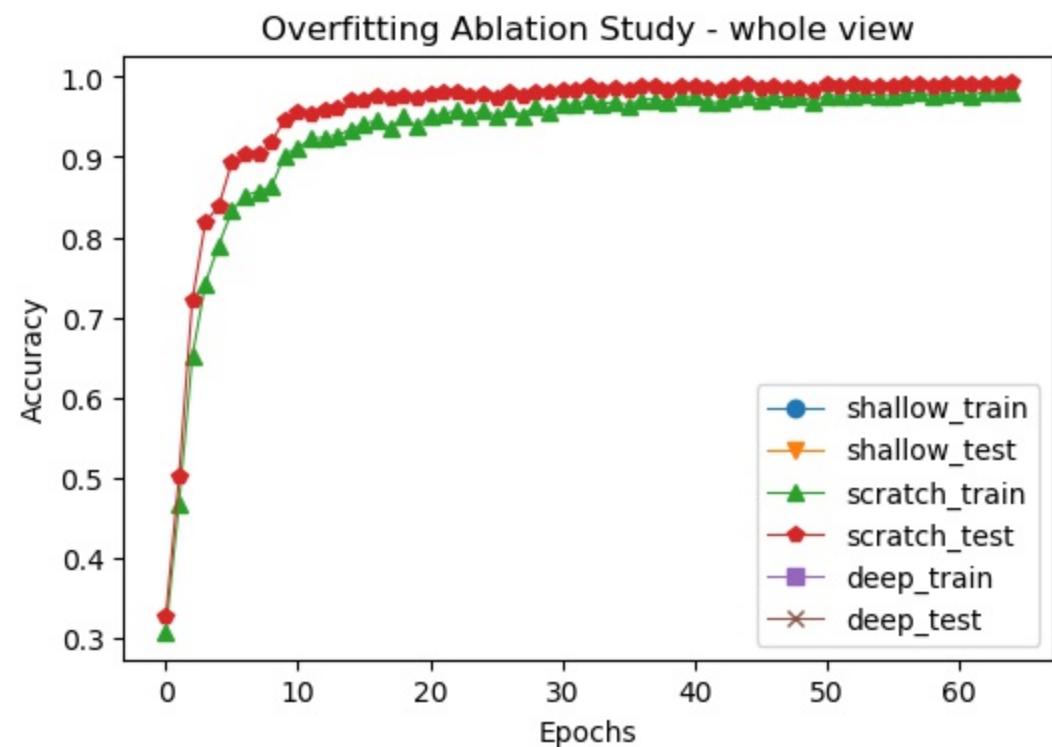
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

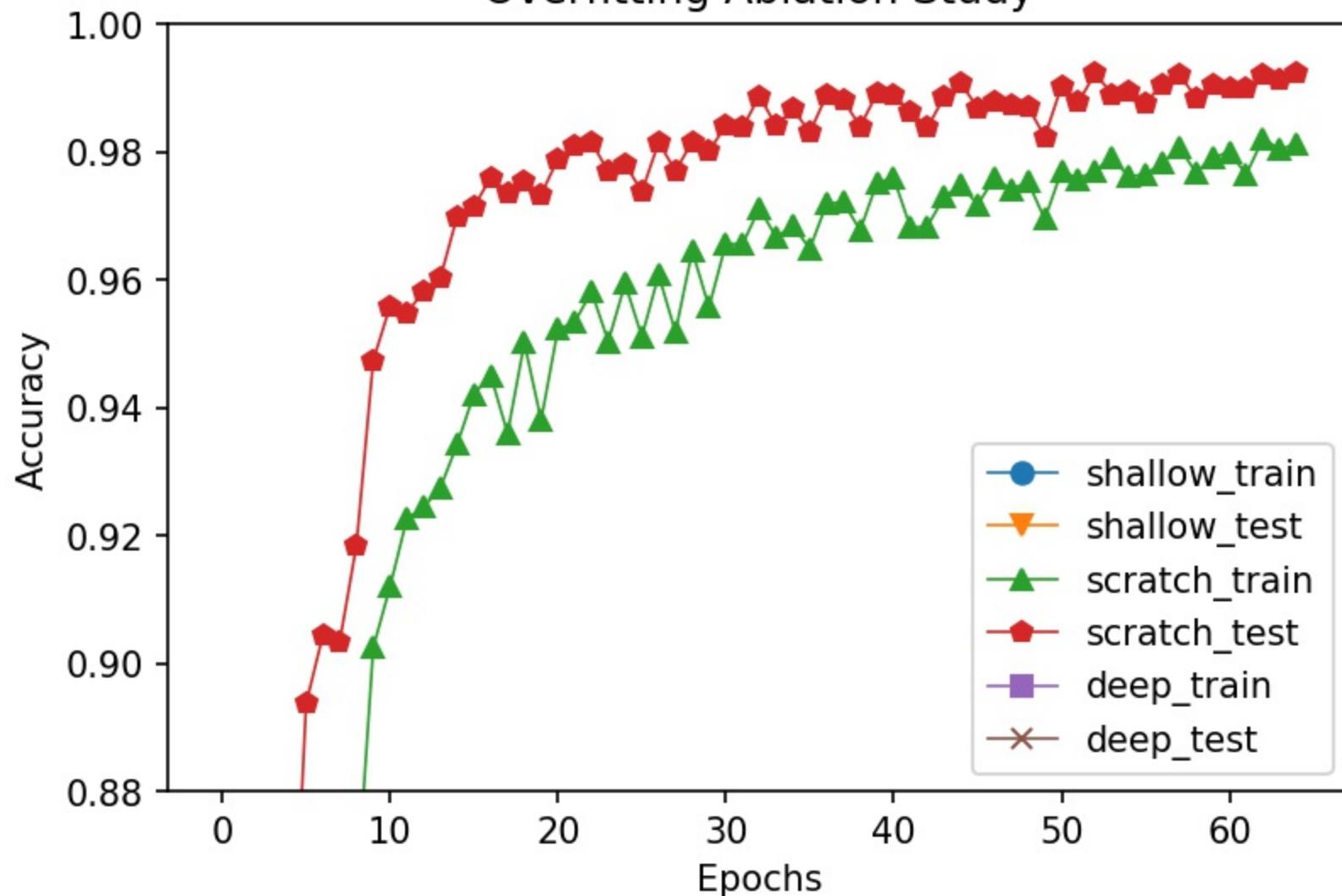
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 64 takes 317 seconds.

[66, 30] loss: 0.081

[66, 60] loss: 0.087

```
[66,    90] loss: 0.105
[66,   120] loss: 0.107
[66,   150] loss: 0.114
[66,   180] loss: 0.069
[66,   210] loss: 0.087
[66,   240] loss: 0.058
[66,   270] loss: 0.089
[66,   300] loss: 0.113
[66,   330] loss: 0.055
[66,   360] loss: 0.097
[66,   390] loss: 0.087
[66,   420] loss: 0.096
[66,   450] loss: 0.108
[66,   480] loss: 0.057
[66,   510] loss: 0.064
[66,   540] loss: 0.100
[66,   570] loss: 0.081
[66,   600] loss: 0.068
[66,   630] loss: 0.062
[66,   660] loss: 0.043
[66,   690] loss: 0.066
[66,   720] loss: 0.088
[66,   750] loss: 0.067
[66,   780] loss: 0.075
[66,   810] loss: 0.085
[66,   840] loss: 0.088
[66,   870] loss: 0.079
[66,   900] loss: 0.094
[66,   930] loss: 0.067
[66,   960] loss: 0.079
[66,   990] loss: 0.068
[66,  1020] loss: 0.065
[66,  1050] loss: 0.056
[66,  1080] loss: 0.092
[66,  1110] loss: 0.077
[66,  1140] loss: 0.048
[66,  1170] loss: 0.085
[66,  1200] loss: 0.053
[66,  1230] loss: 0.076
```

```
[66, 1260] loss: 0.109
[66, 1290] loss: 0.058
[66, 1320] loss: 0.093
[66, 1350] loss: 0.100
[66, 1380] loss: 0.083
[66, 1410] loss: 0.083
[66, 1440] loss: 0.119
[66, 1470] loss: 0.113
[66, 1500] loss: 0.131
[66, 1530] loss: 0.078
[66, 1560] loss: 0.097
[66, 1590] loss: 0.051
[66, 1620] loss: 0.062
[66, 1650] loss: 0.082
[66, 1680] loss: 0.068
[66, 1710] loss: 0.076
[66, 1740] loss: 0.124
[66, 1770] loss: 0.089
[66, 1800] loss: 0.109
[66, 1830] loss: 0.097
[66, 1860] loss: 0.064
[66, 1890] loss: 0.074
[66, 1920] loss: 0.071
[66, 1950] loss: 0.053
[66, 1980] loss: 0.093
[66, 2010] loss: 0.065
[66, 2040] loss: 0.087
[66, 2070] loss: 0.090
[66, 2100] loss: 0.070
[66, 2130] loss: 0.076
[66, 2160] loss: 0.085
[66, 2190] loss: 0.079
eval intermediate_models/epoch65_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.982734
epoch 65 accuracy on train set is: 0.9827335514358415
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992003

epoch 65 accuracy on test set is: 0.9920029080334424

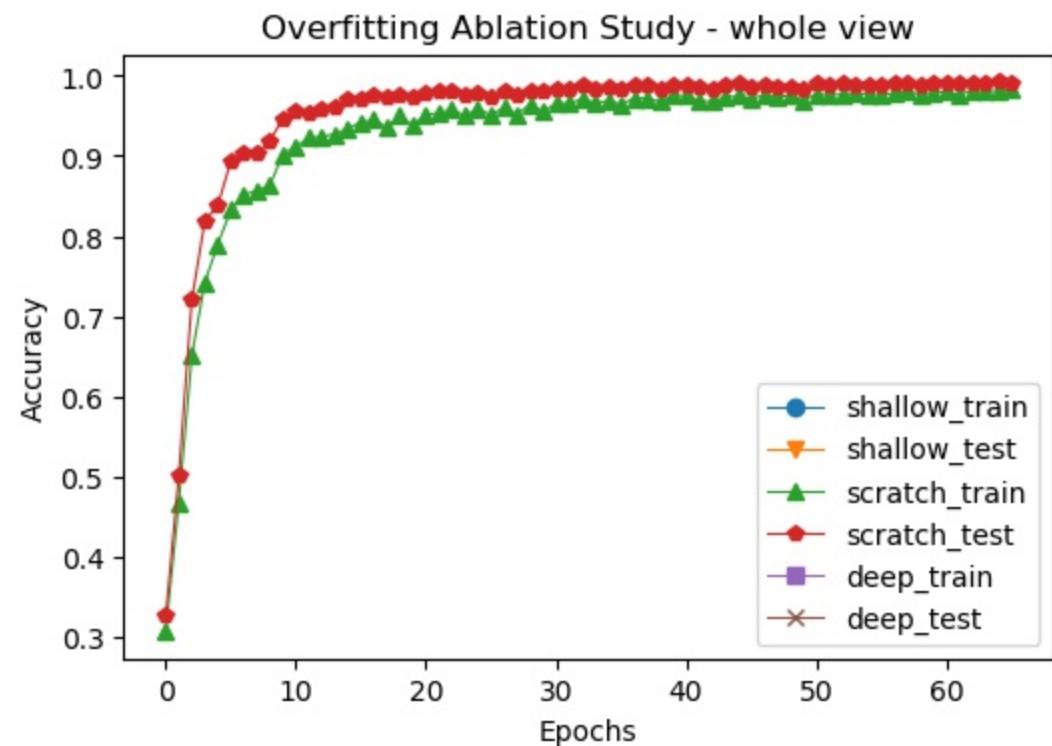
evaluating on test set takes 16 seconds.

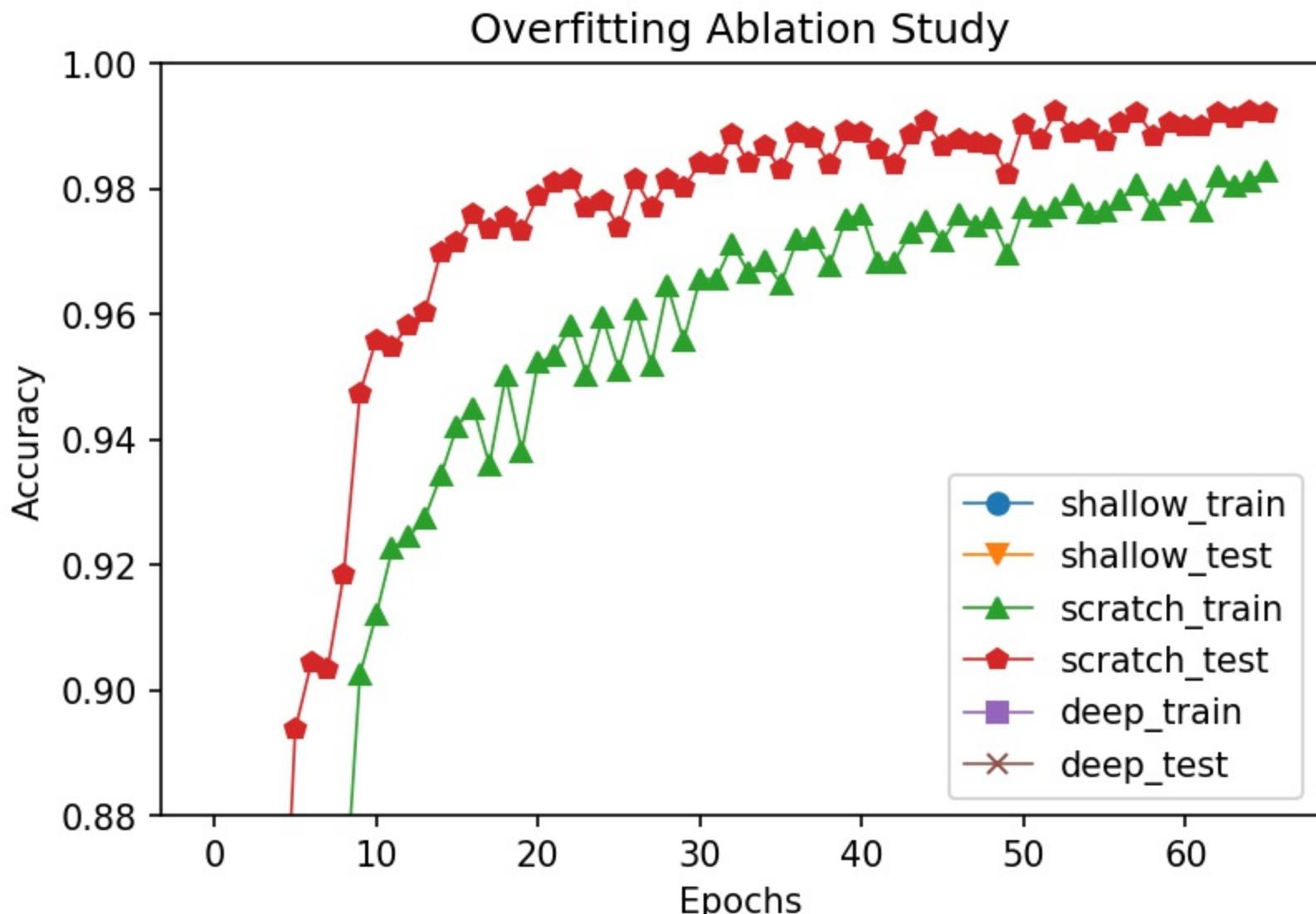
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 65 takes 317 seconds.

[67, 30] loss: 0.069

[67, 60] loss: 0.120

```
[67,    90] loss: 0.086
[67,   120] loss: 0.096
[67,   150] loss: 0.065
[67,   180] loss: 0.050
[67,   210] loss: 0.065
[67,   240] loss: 0.065
[67,   270] loss: 0.117
[67,   300] loss: 0.077
[67,   330] loss: 0.091
[67,   360] loss: 0.092
[67,   390] loss: 0.073
[67,   420] loss: 0.059
[67,   450] loss: 0.061
[67,   480] loss: 0.090
[67,   510] loss: 0.096
[67,   540] loss: 0.056
[67,   570] loss: 0.053
[67,   600] loss: 0.096
[67,   630] loss: 0.080
[67,   660] loss: 0.062
[67,   690] loss: 0.089
[67,   720] loss: 0.081
[67,   750] loss: 0.111
[67,   780] loss: 0.079
[67,   810] loss: 0.066
[67,   840] loss: 0.075
[67,   870] loss: 0.073
[67,   900] loss: 0.092
[67,   930] loss: 0.099
[67,   960] loss: 0.069
[67,   990] loss: 0.069
[67,  1020] loss: 0.096
[67,  1050] loss: 0.122
[67,  1080] loss: 0.043
[67,  1110] loss: 0.080
[67,  1140] loss: 0.054
[67,  1170] loss: 0.032
[67,  1200] loss: 0.135
[67,  1230] loss: 0.120
```

```
[67, 1260] loss: 0.108
[67, 1290] loss: 0.145
[67, 1320] loss: 0.071
[67, 1350] loss: 0.081
[67, 1380] loss: 0.130
[67, 1410] loss: 0.083
[67, 1440] loss: 0.075
[67, 1470] loss: 0.074
[67, 1500] loss: 0.087
[67, 1530] loss: 0.126
[67, 1560] loss: 0.086
[67, 1590] loss: 0.067
[67, 1620] loss: 0.093
[67, 1650] loss: 0.102
[67, 1680] loss: 0.078
[67, 1710] loss: 0.079
[67, 1740] loss: 0.061
[67, 1770] loss: 0.095
[67, 1800] loss: 0.079
[67, 1830] loss: 0.123
[67, 1860] loss: 0.053
[67, 1890] loss: 0.062
[67, 1920] loss: 0.049
[67, 1950] loss: 0.065
[67, 1980] loss: 0.064
[67, 2010] loss: 0.059
[67, 2040] loss: 0.075
[67, 2070] loss: 0.070
[67, 2100] loss: 0.066
[67, 2130] loss: 0.146
[67, 2160] loss: 0.096
[67, 2190] loss: 0.066
eval intermediate_models/epoch66_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.977849
epoch 66 accuracy on train set is: 0.9778489640130862
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988822

epoch 66 accuracy on test set is: 0.9888222464558343

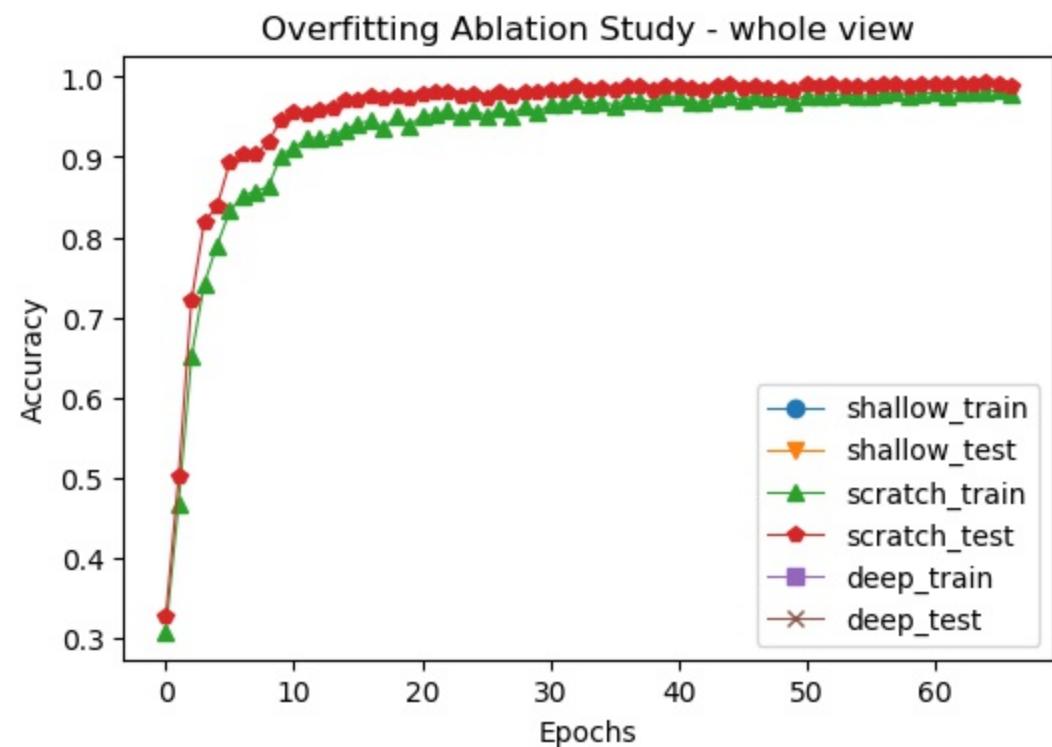
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

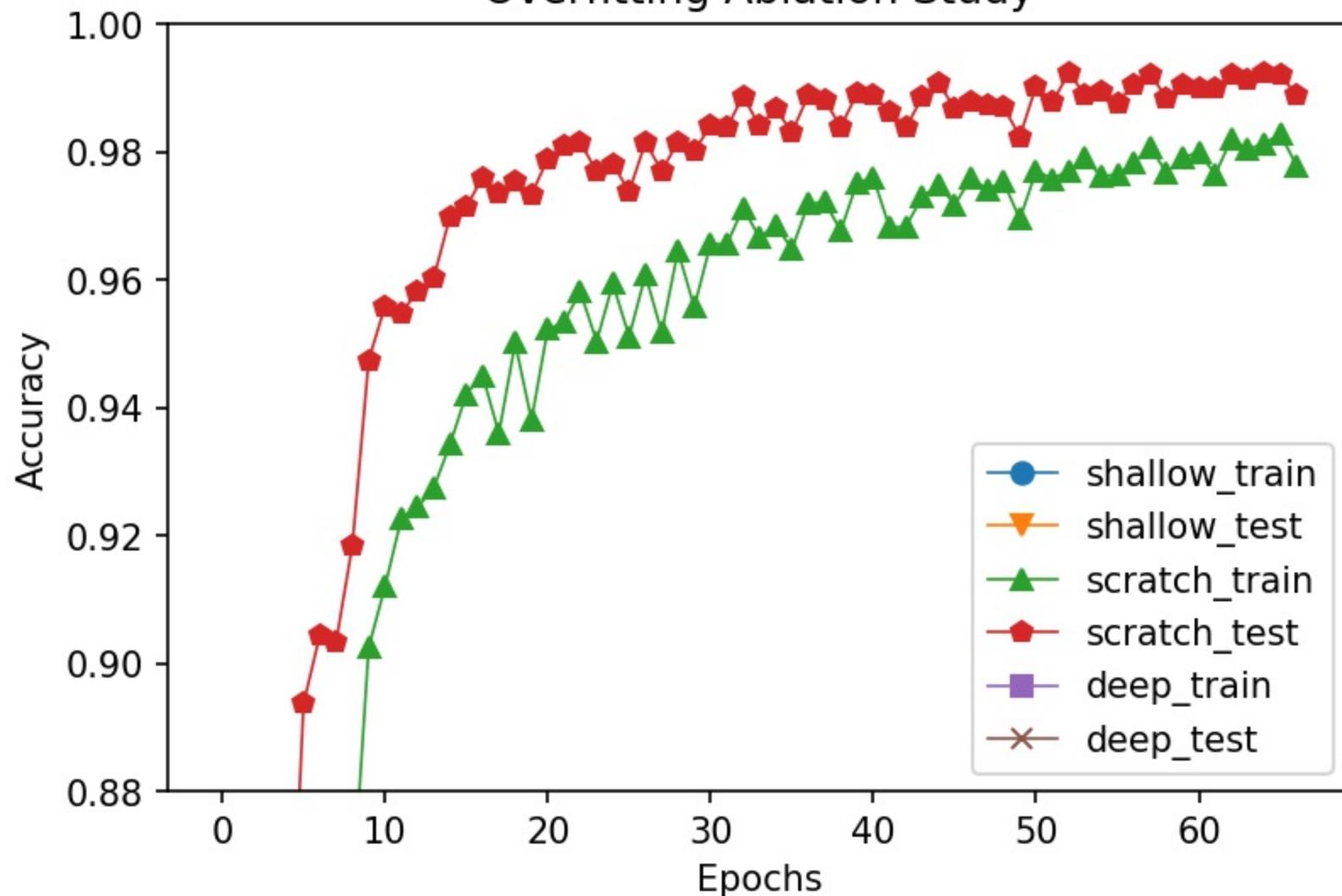
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 66 takes 317 seconds.

[68, 30] loss: 0.105

[68, 60] loss: 0.093

```
[68,    90] loss: 0.067
[68,   120] loss: 0.082
[68,   150] loss: 0.087
[68,   180] loss: 0.067
[68,   210] loss: 0.047
[68,   240] loss: 0.065
[68,   270] loss: 0.105
[68,   300] loss: 0.078
[68,   330] loss: 0.094
[68,   360] loss: 0.069
[68,   390] loss: 0.096
[68,   420] loss: 0.081
[68,   450] loss: 0.079
[68,   480] loss: 0.105
[68,   510] loss: 0.077
[68,   540] loss: 0.067
[68,   570] loss: 0.080
[68,   600] loss: 0.087
[68,   630] loss: 0.076
[68,   660] loss: 0.071
[68,   690] loss: 0.075
[68,   720] loss: 0.077
[68,   750] loss: 0.092
[68,   780] loss: 0.072
[68,   810] loss: 0.082
[68,   840] loss: 0.114
[68,   870] loss: 0.079
[68,   900] loss: 0.055
[68,   930] loss: 0.055
[68,   960] loss: 0.082
[68,   990] loss: 0.075
[68,  1020] loss: 0.126
[68,  1050] loss: 0.073
[68,  1080] loss: 0.056
[68,  1110] loss: 0.104
[68,  1140] loss: 0.088
[68,  1170] loss: 0.049
[68,  1200] loss: 0.062
[68,  1230] loss: 0.069
```

```
[68, 1260] loss: 0.048
[68, 1290] loss: 0.059
[68, 1320] loss: 0.040
[68, 1350] loss: 0.071
[68, 1380] loss: 0.059
[68, 1410] loss: 0.052
[68, 1440] loss: 0.117
[68, 1470] loss: 0.113
[68, 1500] loss: 0.104
[68, 1530] loss: 0.062
[68, 1560] loss: 0.075
[68, 1590] loss: 0.078
[68, 1620] loss: 0.079
[68, 1650] loss: 0.074
[68, 1680] loss: 0.071
[68, 1710] loss: 0.083
[68, 1740] loss: 0.044
[68, 1770] loss: 0.067
[68, 1800] loss: 0.079
[68, 1830] loss: 0.090
[68, 1860] loss: 0.085
[68, 1890] loss: 0.078
[68, 1920] loss: 0.069
[68, 1950] loss: 0.065
[68, 1980] loss: 0.091
[68, 2010] loss: 0.070
[68, 2040] loss: 0.083
[68, 2070] loss: 0.067
[68, 2100] loss: 0.114
[68, 2130] loss: 0.049
[68, 2160] loss: 0.058
[68, 2190] loss: 0.069
evf intermediate_models/epoch67_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.979621
epoch 67 accuracy on train set is: 0.9796210468920392
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992185

epoch 67 accuracy on test set is: 0.9921846601235914

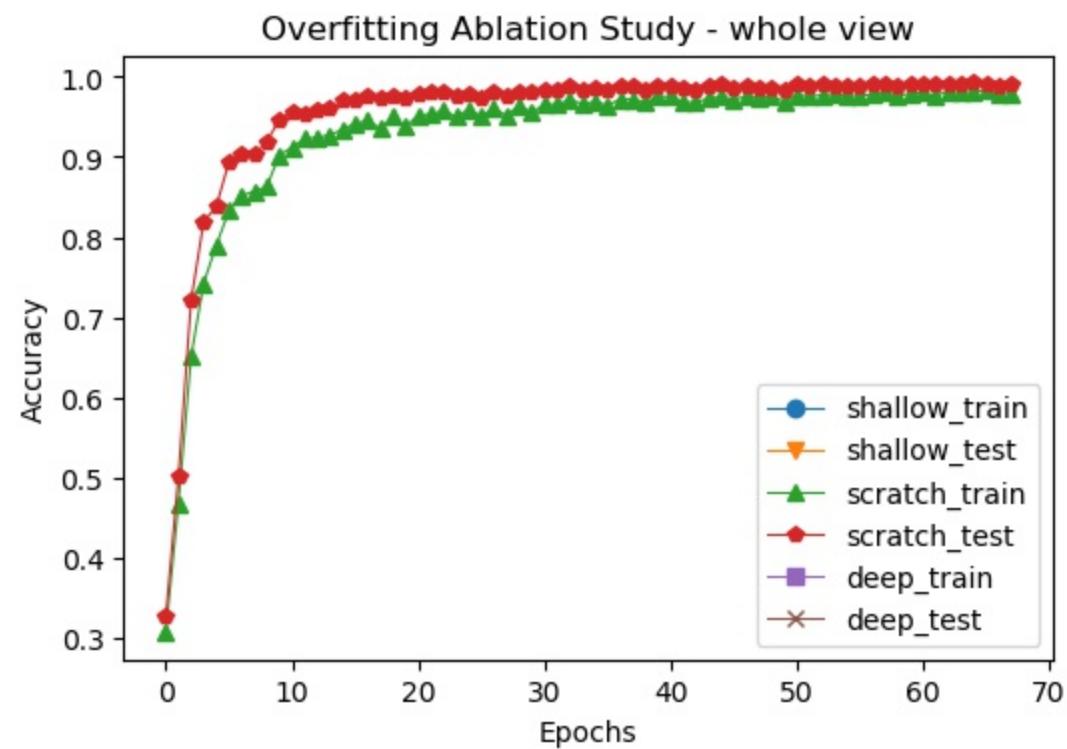
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

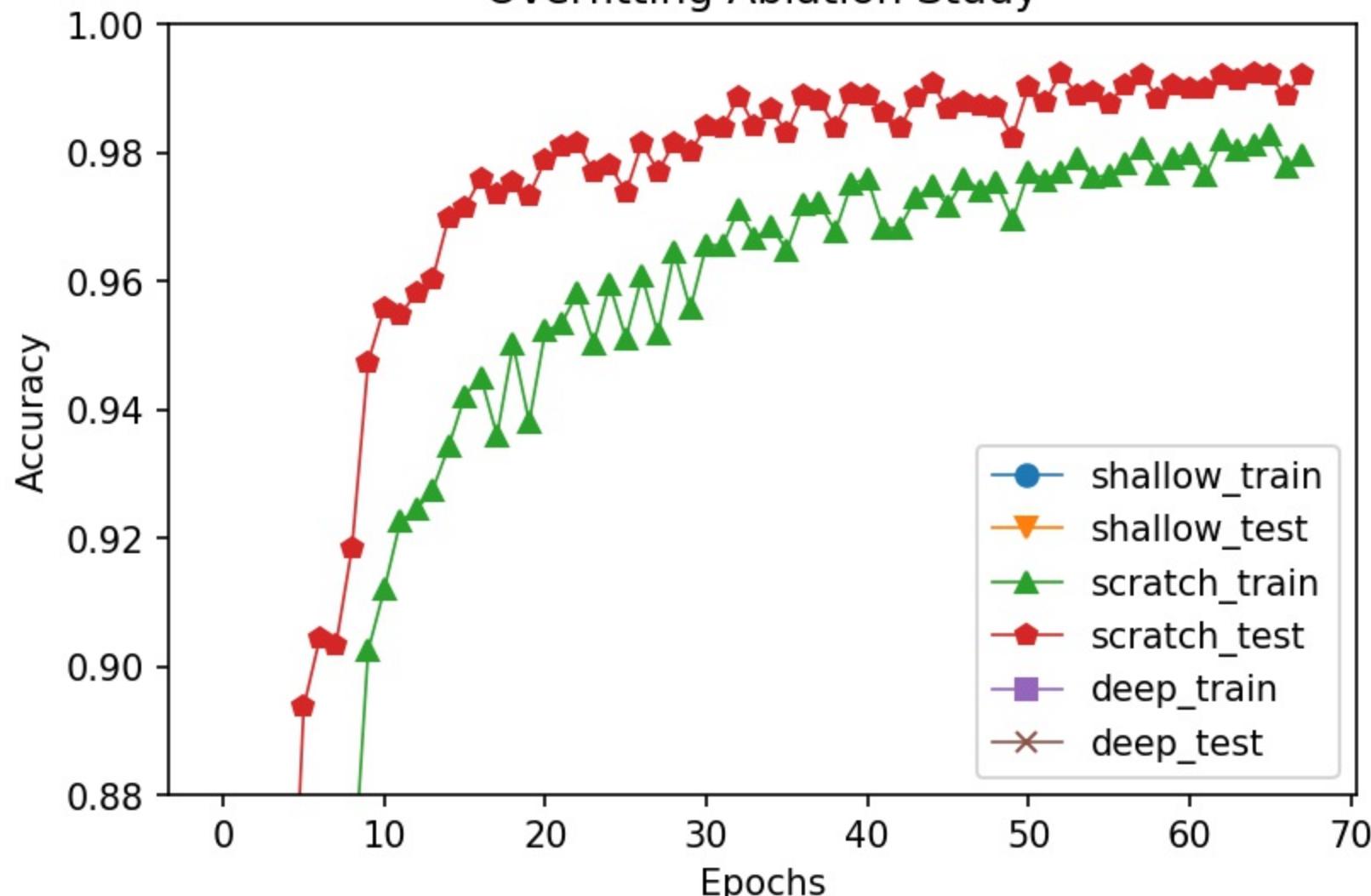
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 67 takes 317 seconds.

[69, 30] loss: 0.094

[69, 60] loss: 0.105

```
[69,    90] loss: 0.094
[69,   120] loss: 0.084
[69,   150] loss: 0.070
[69,   180] loss: 0.077
[69,   210] loss: 0.080
[69,   240] loss: 0.062
[69,   270] loss: 0.054
[69,   300] loss: 0.078
[69,   330] loss: 0.042
[69,   360] loss: 0.066
[69,   390] loss: 0.081
[69,   420] loss: 0.081
[69,   450] loss: 0.085
[69,   480] loss: 0.109
[69,   510] loss: 0.044
[69,   540] loss: 0.067
[69,   570] loss: 0.086
[69,   600] loss: 0.067
[69,   630] loss: 0.088
[69,   660] loss: 0.080
[69,   690] loss: 0.106
[69,   720] loss: 0.055
[69,   750] loss: 0.080
[69,   780] loss: 0.079
[69,   810] loss: 0.075
[69,   840] loss: 0.104
[69,   870] loss: 0.086
[69,   900] loss: 0.060
[69,   930] loss: 0.085
[69,   960] loss: 0.078
[69,   990] loss: 0.071
[69,  1020] loss: 0.073
[69,  1050] loss: 0.086
[69,  1080] loss: 0.100
[69,  1110] loss: 0.062
[69,  1140] loss: 0.101
[69,  1170] loss: 0.098
[69,  1200] loss: 0.107
[69,  1230] loss: 0.083
```

```
[69, 1260] loss: 0.096
[69, 1290] loss: 0.075
[69, 1320] loss: 0.081
[69, 1350] loss: 0.076
[69, 1380] loss: 0.088
[69, 1410] loss: 0.059
[69, 1440] loss: 0.068
[69, 1470] loss: 0.074
[69, 1500] loss: 0.062
[69, 1530] loss: 0.065
[69, 1560] loss: 0.155
[69, 1590] loss: 0.056
[69, 1620] loss: 0.078
[69, 1650] loss: 0.057
[69, 1680] loss: 0.071
[69, 1710] loss: 0.115
[69, 1740] loss: 0.077
[69, 1770] loss: 0.097
[69, 1800] loss: 0.089
[69, 1830] loss: 0.086
[69, 1860] loss: 0.127
[69, 1890] loss: 0.100
[69, 1920] loss: 0.072
[69, 1950] loss: 0.111
[69, 1980] loss: 0.078
[69, 2010] loss: 0.060
[69, 2040] loss: 0.079
[69, 2070] loss: 0.049
[69, 2100] loss: 0.091
[69, 2130] loss: 0.055
[69, 2160] loss: 0.059
[69, 2190] loss: 0.078
eval intermediate_models/epoch68_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.981007
epoch 68 accuracy on train set is: 0.9810069065794257
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992912

epoch 68 accuracy on test set is: 0.9929116684841875

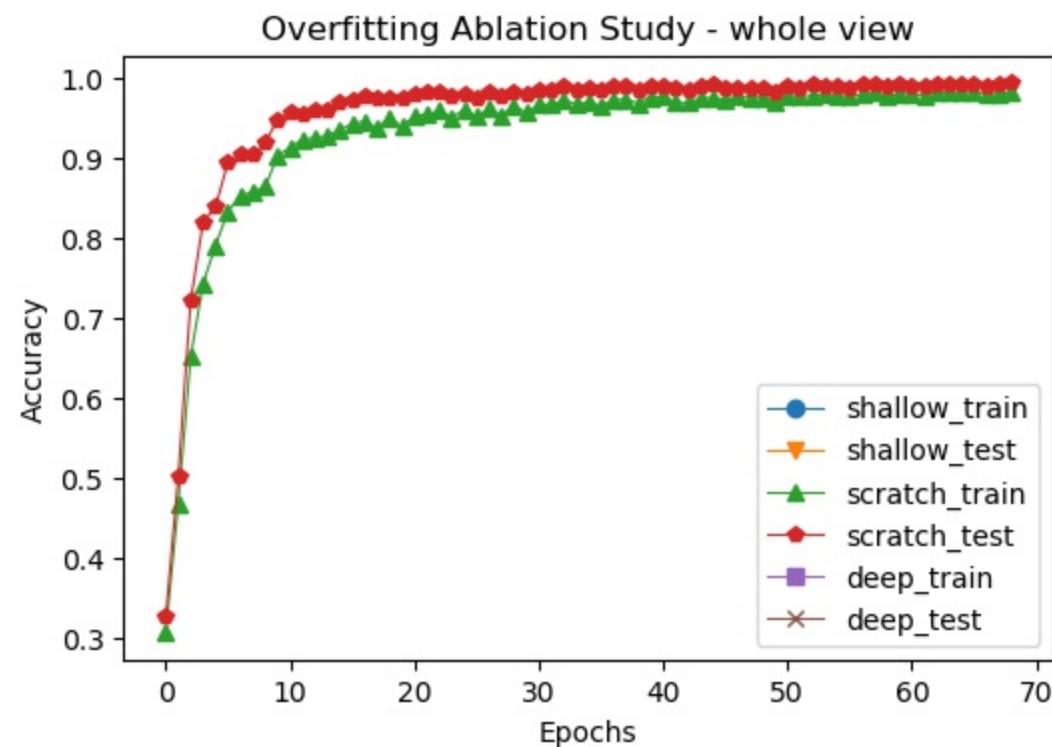
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

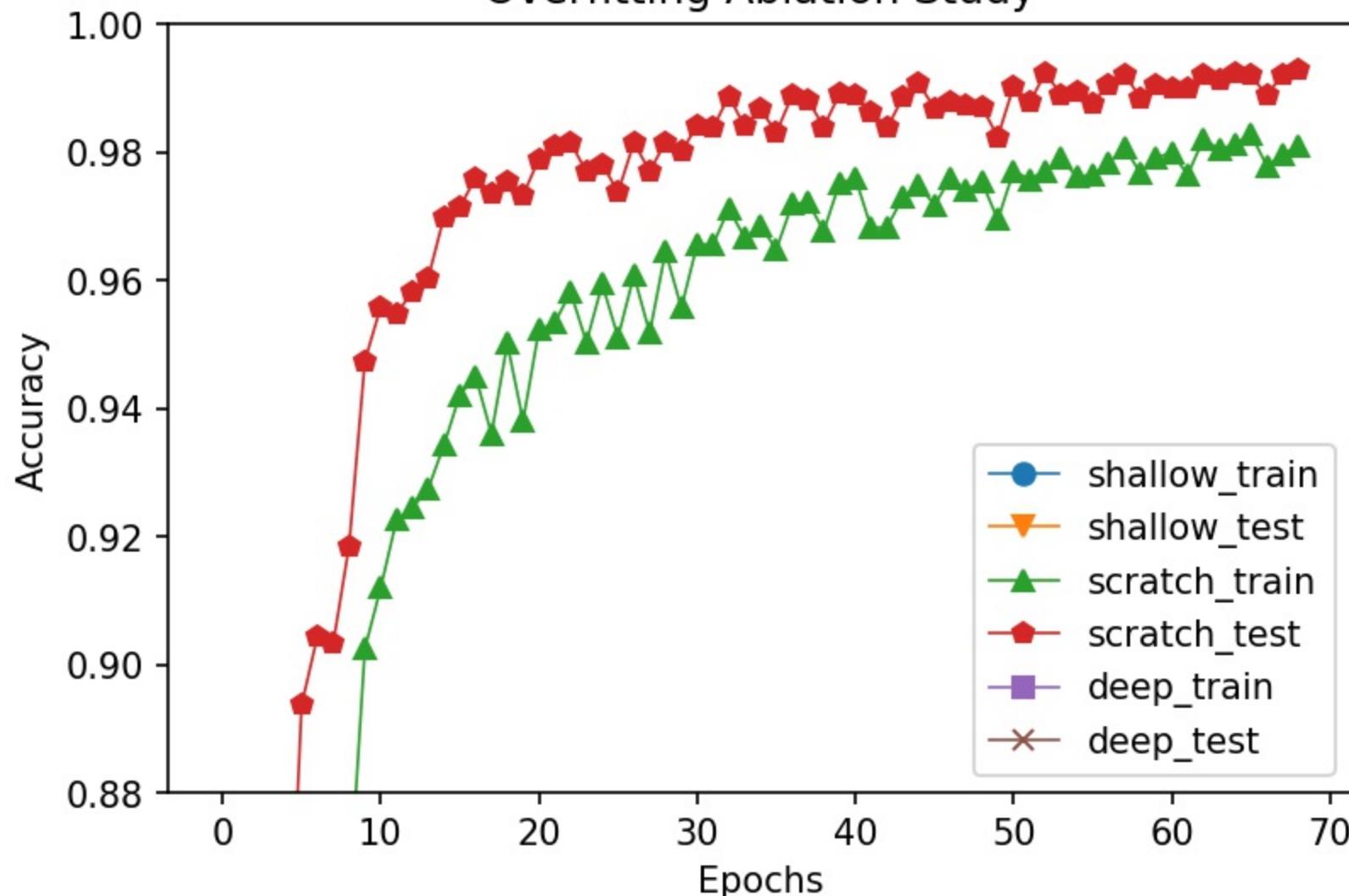
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 68 takes 316 seconds.

[70, 30] loss: 0.059

[70, 60] loss: 0.091

```
[70,    90] loss: 0.069
[70,   120] loss: 0.101
[70,   150] loss: 0.057
[70,   180] loss: 0.077
[70,   210] loss: 0.048
[70,   240] loss: 0.098
[70,   270] loss: 0.079
[70,   300] loss: 0.075
[70,   330] loss: 0.094
[70,   360] loss: 0.092
[70,   390] loss: 0.089
[70,   420] loss: 0.061
[70,   450] loss: 0.086
[70,   480] loss: 0.065
[70,   510] loss: 0.081
[70,   540] loss: 0.087
[70,   570] loss: 0.127
[70,   600] loss: 0.075
[70,   630] loss: 0.086
[70,   660] loss: 0.071
[70,   690] loss: 0.064
[70,   720] loss: 0.088
[70,   750] loss: 0.096
[70,   780] loss: 0.130
[70,   810] loss: 0.086
[70,   840] loss: 0.066
[70,   870] loss: 0.110
[70,   900] loss: 0.084
[70,   930] loss: 0.078
[70,   960] loss: 0.063
[70,   990] loss: 0.079
[70,  1020] loss: 0.063
[70,  1050] loss: 0.101
[70,  1080] loss: 0.071
[70,  1110] loss: 0.085
[70,  1140] loss: 0.080
[70,  1170] loss: 0.098
[70,  1200] loss: 0.122
[70,  1230] loss: 0.069
```

```
[70, 1260] loss: 0.071
[70, 1290] loss: 0.072
[70, 1320] loss: 0.085
[70, 1350] loss: 0.054
[70, 1380] loss: 0.085
[70, 1410] loss: 0.075
[70, 1440] loss: 0.092
[70, 1470] loss: 0.076
[70, 1500] loss: 0.056
[70, 1530] loss: 0.071
[70, 1560] loss: 0.040
[70, 1590] loss: 0.109
[70, 1620] loss: 0.070
[70, 1650] loss: 0.108
[70, 1680] loss: 0.083
[70, 1710] loss: 0.066
[70, 1740] loss: 0.067
[70, 1770] loss: 0.055
[70, 1800] loss: 0.097
[70, 1830] loss: 0.077
[70, 1860] loss: 0.071
[70, 1890] loss: 0.075
[70, 1920] loss: 0.072
[70, 1950] loss: 0.092
[70, 1980] loss: 0.092
[70, 2010] loss: 0.045
[70, 2040] loss: 0.085
[70, 2070] loss: 0.059
[70, 2100] loss: 0.079
[70, 2130] loss: 0.040
[70, 2160] loss: 0.049
[70, 2190] loss: 0.063
eval intermediate_models/epoch69_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.983915
epoch 69 accuracy on train set is: 0.9839149400218102
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992821

epoch 69 accuracy on test set is: 0.992820792439113

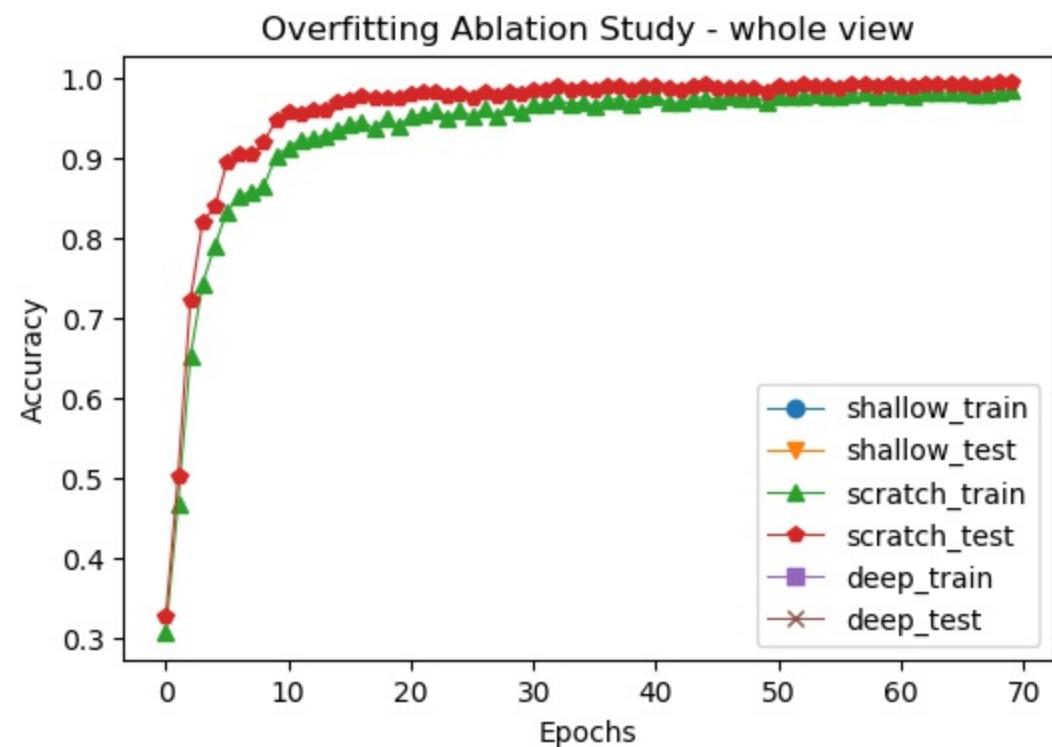
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

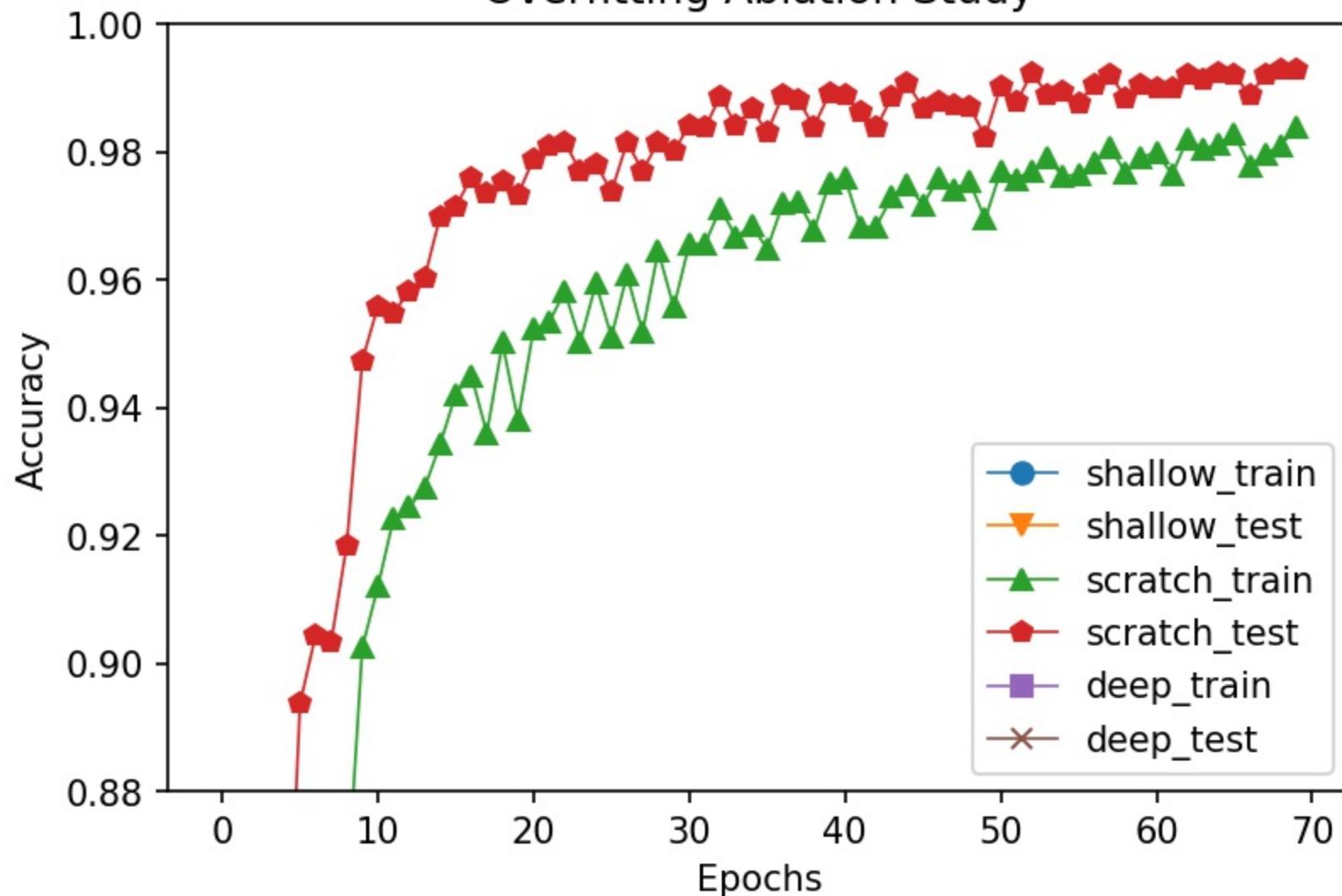
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 69 takes 315 seconds.

[71, 30] loss: 0.053

[71, 60] loss: 0.090

```
[71,    90] loss: 0.088
[71,   120] loss: 0.081
[71,   150] loss: 0.065
[71,   180] loss: 0.073
[71,   210] loss: 0.062
[71,   240] loss: 0.064
[71,   270] loss: 0.089
[71,   300] loss: 0.083
[71,   330] loss: 0.055
[71,   360] loss: 0.070
[71,   390] loss: 0.111
[71,   420] loss: 0.075
[71,   450] loss: 0.062
[71,   480] loss: 0.071
[71,   510] loss: 0.082
[71,   540] loss: 0.046
[71,   570] loss: 0.094
[71,   600] loss: 0.081
[71,   630] loss: 0.066
[71,   660] loss: 0.046
[71,   690] loss: 0.081
[71,   720] loss: 0.096
[71,   750] loss: 0.049
[71,   780] loss: 0.083
[71,   810] loss: 0.068
[71,   840] loss: 0.067
[71,   870] loss: 0.107
[71,   900] loss: 0.087
[71,   930] loss: 0.069
[71,   960] loss: 0.086
[71,   990] loss: 0.083
[71,  1020] loss: 0.054
[71,  1050] loss: 0.080
[71,  1080] loss: 0.081
[71,  1110] loss: 0.074
[71,  1140] loss: 0.061
[71,  1170] loss: 0.094
[71,  1200] loss: 0.093
[71,  1230] loss: 0.072
```

```
[71, 1260] loss: 0.099
[71, 1290] loss: 0.090
[71, 1320] loss: 0.089
[71, 1350] loss: 0.063
[71, 1380] loss: 0.074
[71, 1410] loss: 0.065
[71, 1440] loss: 0.072
[71, 1470] loss: 0.101
[71, 1500] loss: 0.060
[71, 1530] loss: 0.083
[71, 1560] loss: 0.094
[71, 1590] loss: 0.069
[71, 1620] loss: 0.084
[71, 1650] loss: 0.080
[71, 1680] loss: 0.076
[71, 1710] loss: 0.110
[71, 1740] loss: 0.065
[71, 1770] loss: 0.097
[71, 1800] loss: 0.046
[71, 1830] loss: 0.050
[71, 1860] loss: 0.092
[71, 1890] loss: 0.079
[71, 1920] loss: 0.101
[71, 1950] loss: 0.090
[71, 1980] loss: 0.098
[71, 2010] loss: 0.112
[71, 2040] loss: 0.077
[71, 2070] loss: 0.078
[71, 2100] loss: 0.091
[71, 2130] loss: 0.134
[71, 2160] loss: 0.082
[71, 2190] loss: 0.072
evf intermediate_models/epoch70_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.982938
epoch 70 accuracy on train set is: 0.9829380225372591
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991185

epoch 70 accuracy on test set is: 0.9911850236277717

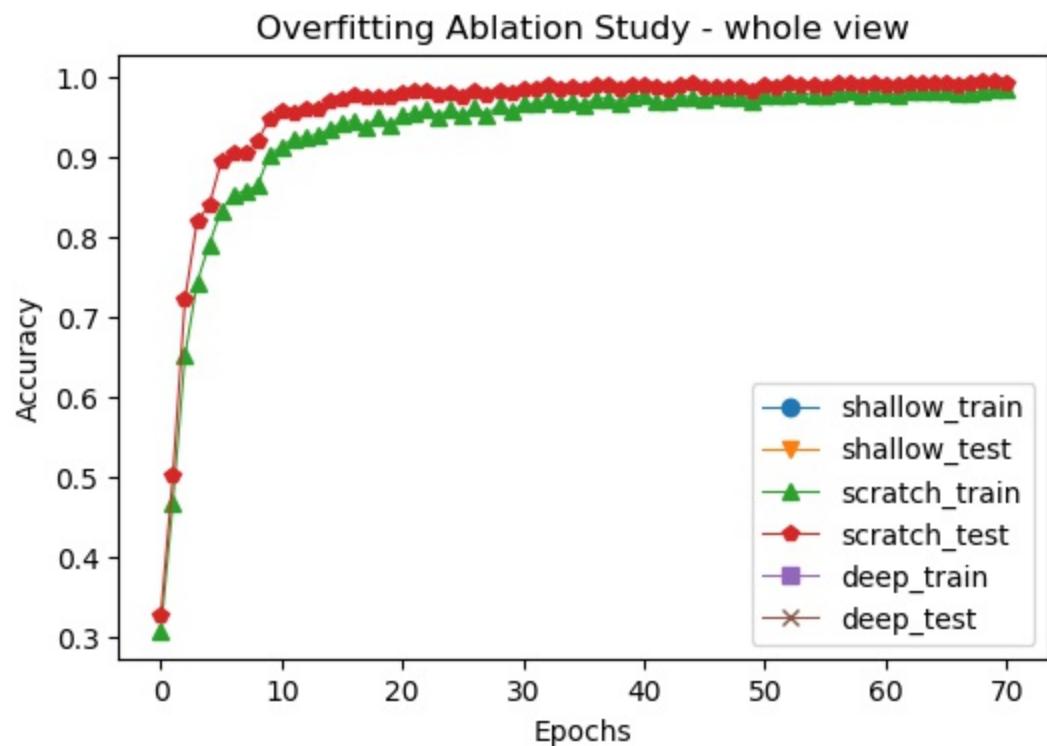
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

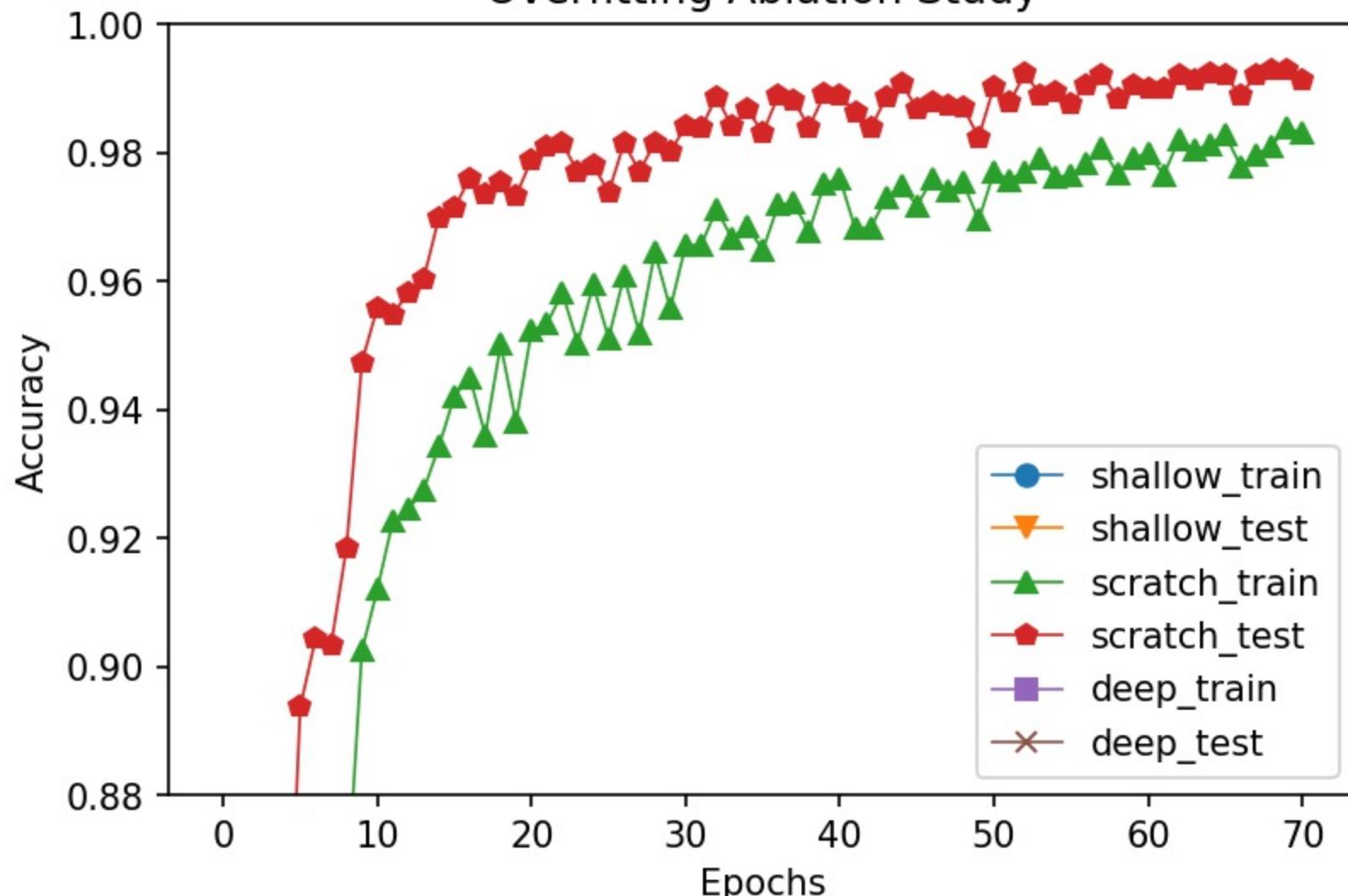
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 70 takes 316 seconds.

[72, 30] loss: 0.056

[72, 60] loss: 0.063

```
[72,    90] loss: 0.037
[72,   120] loss: 0.076
[72,   150] loss: 0.067
[72,   180] loss: 0.085
[72,   210] loss: 0.082
[72,   240] loss: 0.079
[72,   270] loss: 0.076
[72,   300] loss: 0.074
[72,   330] loss: 0.062
[72,   360] loss: 0.046
[72,   390] loss: 0.095
[72,   420] loss: 0.101
[72,   450] loss: 0.073
[72,   480] loss: 0.090
[72,   510] loss: 0.087
[72,   540] loss: 0.065
[72,   570] loss: 0.075
[72,   600] loss: 0.070
[72,   630] loss: 0.060
[72,   660] loss: 0.057
[72,   690] loss: 0.076
[72,   720] loss: 0.081
[72,   750] loss: 0.104
[72,   780] loss: 0.087
[72,   810] loss: 0.073
[72,   840] loss: 0.061
[72,   870] loss: 0.084
[72,   900] loss: 0.063
[72,   930] loss: 0.061
[72,   960] loss: 0.114
[72,   990] loss: 0.030
[72,  1020] loss: 0.064
[72,  1050] loss: 0.104
[72,  1080] loss: 0.058
[72,  1110] loss: 0.056
[72,  1140] loss: 0.055
[72,  1170] loss: 0.087
[72,  1200] loss: 0.065
[72,  1230] loss: 0.111
```

```
[72, 1260] loss: 0.055
[72, 1290] loss: 0.093
[72, 1320] loss: 0.110
[72, 1350] loss: 0.080
[72, 1380] loss: 0.042
[72, 1410] loss: 0.079
[72, 1440] loss: 0.081
[72, 1470] loss: 0.063
[72, 1500] loss: 0.073
[72, 1530] loss: 0.068
[72, 1560] loss: 0.082
[72, 1590] loss: 0.095
[72, 1620] loss: 0.067
[72, 1650] loss: 0.075
[72, 1680] loss: 0.042
[72, 1710] loss: 0.094
[72, 1740] loss: 0.084
[72, 1770] loss: 0.044
[72, 1800] loss: 0.102
[72, 1830] loss: 0.074
[72, 1860] loss: 0.048
[72, 1890] loss: 0.073
[72, 1920] loss: 0.088
[72, 1950] loss: 0.078
[72, 1980] loss: 0.103
[72, 2010] loss: 0.063
[72, 2040] loss: 0.068
[72, 2070] loss: 0.063
[72, 2100] loss: 0.063
[72, 2130] loss: 0.051
[72, 2160] loss: 0.059
[72, 2190] loss: 0.066
eval intermediate_models/epoch71_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.980371
epoch 71 accuracy on train set is: 0.9803707742639041
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.989913

epoch 71 accuracy on test set is: 0.9899127589967285

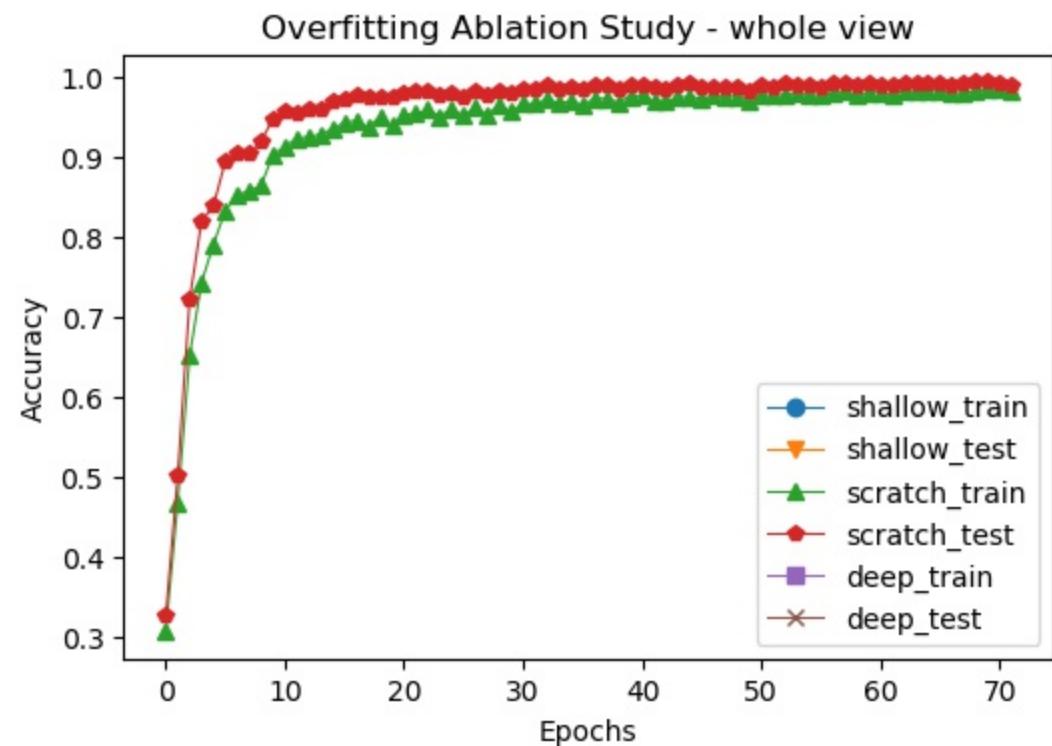
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

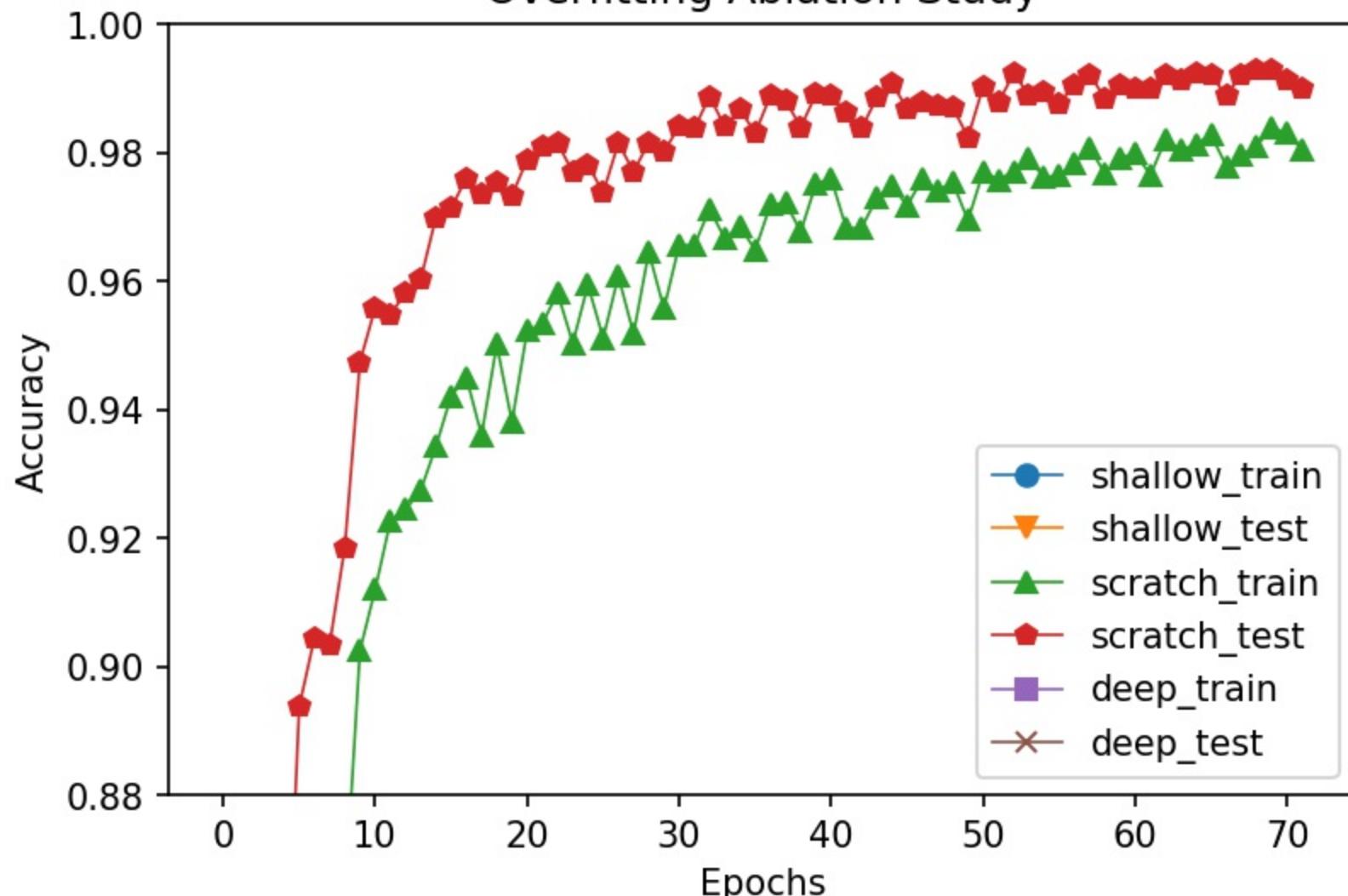
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 71 takes 316 seconds.

[73, 30] loss: 0.074

[73, 60] loss: 0.058

```
[73,    90] loss: 0.067
[73,   120] loss: 0.084
[73,   150] loss: 0.043
[73,   180] loss: 0.077
[73,   210] loss: 0.072
[73,   240] loss: 0.066
[73,   270] loss: 0.075
[73,   300] loss: 0.069
[73,   330] loss: 0.064
[73,   360] loss: 0.072
[73,   390] loss: 0.093
[73,   420] loss: 0.080
[73,   450] loss: 0.084
[73,   480] loss: 0.073
[73,   510] loss: 0.067
[73,   540] loss: 0.059
[73,   570] loss: 0.057
[73,   600] loss: 0.045
[73,   630] loss: 0.095
[73,   660] loss: 0.062
[73,   690] loss: 0.089
[73,   720] loss: 0.113
[73,   750] loss: 0.075
[73,   780] loss: 0.059
[73,   810] loss: 0.072
[73,   840] loss: 0.063
[73,   870] loss: 0.056
[73,   900] loss: 0.089
[73,   930] loss: 0.065
[73,   960] loss: 0.063
[73,   990] loss: 0.056
[73,  1020] loss: 0.067
[73,  1050] loss: 0.062
[73,  1080] loss: 0.080
[73,  1110] loss: 0.048
[73,  1140] loss: 0.084
[73,  1170] loss: 0.044
[73,  1200] loss: 0.045
[73,  1230] loss: 0.058
```

```
[73, 1260] loss: 0.060
[73, 1290] loss: 0.068
[73, 1320] loss: 0.070
[73, 1350] loss: 0.054
[73, 1380] loss: 0.077
[73, 1410] loss: 0.059
[73, 1440] loss: 0.090
[73, 1470] loss: 0.129
[73, 1500] loss: 0.054
[73, 1530] loss: 0.073
[73, 1560] loss: 0.106
[73, 1590] loss: 0.049
[73, 1620] loss: 0.080
[73, 1650] loss: 0.094
[73, 1680] loss: 0.066
[73, 1710] loss: 0.074
[73, 1740] loss: 0.056
[73, 1770] loss: 0.126
[73, 1800] loss: 0.082
[73, 1830] loss: 0.084
[73, 1860] loss: 0.104
[73, 1890] loss: 0.064
[73, 1920] loss: 0.105
[73, 1950] loss: 0.079
[73, 1980] loss: 0.087
[73, 2010] loss: 0.079
[73, 2040] loss: 0.085
[73, 2070] loss: 0.088
[73, 2100] loss: 0.066
[73, 2130] loss: 0.084
[73, 2160] loss: 0.068
[73, 2190] loss: 0.090
eval intermediate_models/epoch72_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.981302
epoch 72 accuracy on train set is: 0.9813022537259178
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992639

epoch 72 accuracy on test set is: 0.992639040348964

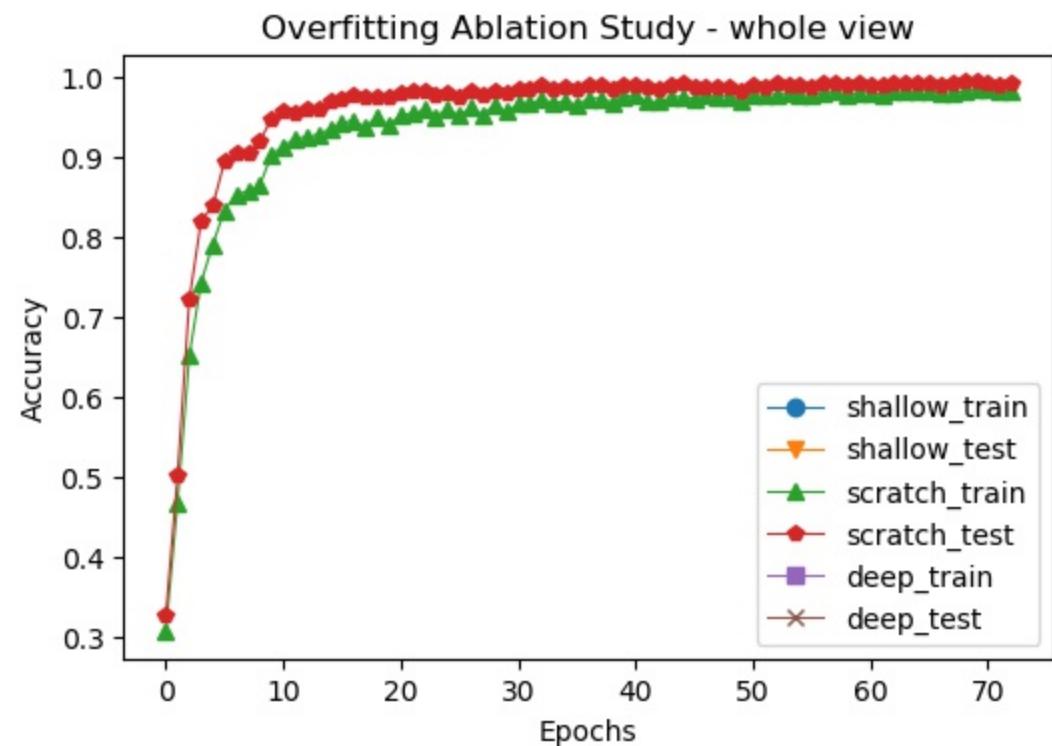
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

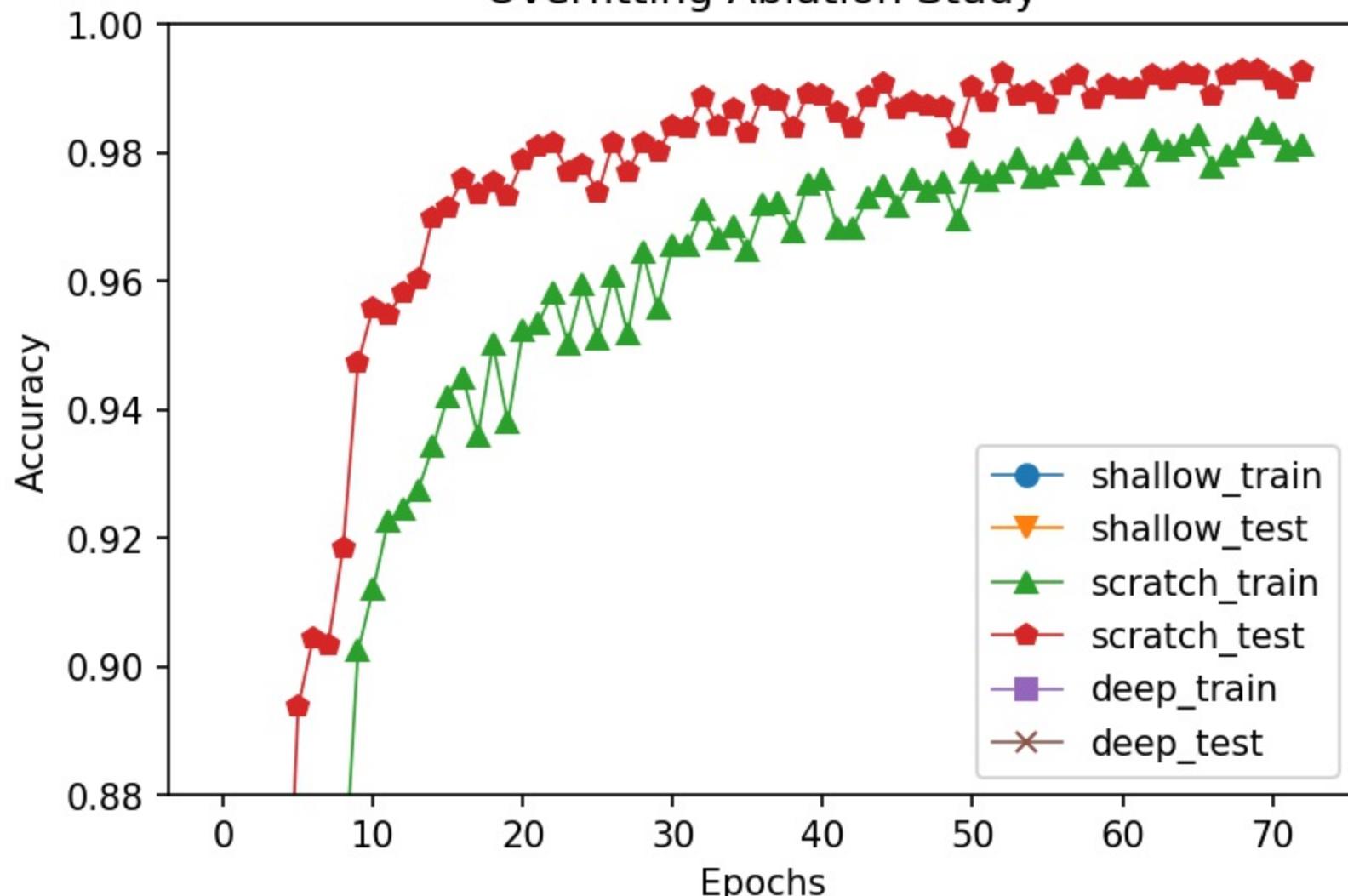
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 72 takes 316 seconds.

[74, 30] loss: 0.064

[74, 60] loss: 0.078

```
[74,    90] loss: 0.097
[74,   120] loss: 0.101
[74,   150] loss: 0.061
[74,   180] loss: 0.108
[74,   210] loss: 0.073
[74,   240] loss: 0.083
[74,   270] loss: 0.075
[74,   300] loss: 0.065
[74,   330] loss: 0.107
[74,   360] loss: 0.054
[74,   390] loss: 0.057
[74,   420] loss: 0.061
[74,   450] loss: 0.072
[74,   480] loss: 0.084
[74,   510] loss: 0.082
[74,   540] loss: 0.081
[74,   570] loss: 0.092
[74,   600] loss: 0.063
[74,   630] loss: 0.086
[74,   660] loss: 0.107
[74,   690] loss: 0.061
[74,   720] loss: 0.064
[74,   750] loss: 0.067
[74,   780] loss: 0.056
[74,   810] loss: 0.075
[74,   840] loss: 0.056
[74,   870] loss: 0.097
[74,   900] loss: 0.065
[74,   930] loss: 0.049
[74,   960] loss: 0.062
[74,   990] loss: 0.069
[74,  1020] loss: 0.062
[74,  1050] loss: 0.096
[74,  1080] loss: 0.140
[74,  1110] loss: 0.075
[74,  1140] loss: 0.065
[74,  1170] loss: 0.067
[74,  1200] loss: 0.078
[74,  1230] loss: 0.077
```

```
[74, 1260] loss: 0.080
[74, 1290] loss: 0.065
[74, 1320] loss: 0.056
[74, 1350] loss: 0.067
[74, 1380] loss: 0.096
[74, 1410] loss: 0.074
[74, 1440] loss: 0.052
[74, 1470] loss: 0.059
[74, 1500] loss: 0.075
[74, 1530] loss: 0.077
[74, 1560] loss: 0.085
[74, 1590] loss: 0.096
[74, 1620] loss: 0.081
[74, 1650] loss: 0.085
[74, 1680] loss: 0.059
[74, 1710] loss: 0.040
[74, 1740] loss: 0.039
[74, 1770] loss: 0.069
[74, 1800] loss: 0.050
[74, 1830] loss: 0.046
[74, 1860] loss: 0.093
[74, 1890] loss: 0.083
[74, 1920] loss: 0.073
[74, 1950] loss: 0.065
[74, 1980] loss: 0.070
[74, 2010] loss: 0.051
[74, 2040] loss: 0.056
[74, 2070] loss: 0.102
[74, 2100] loss: 0.086
[74, 2130] loss: 0.066
[74, 2160] loss: 0.104
[74, 2190] loss: 0.114
eval intermediate_models/epoch73_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.983938
epoch 73 accuracy on train set is: 0.9839376590330788
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991821

epoch 73 accuracy on test set is: 0.9918211559432933

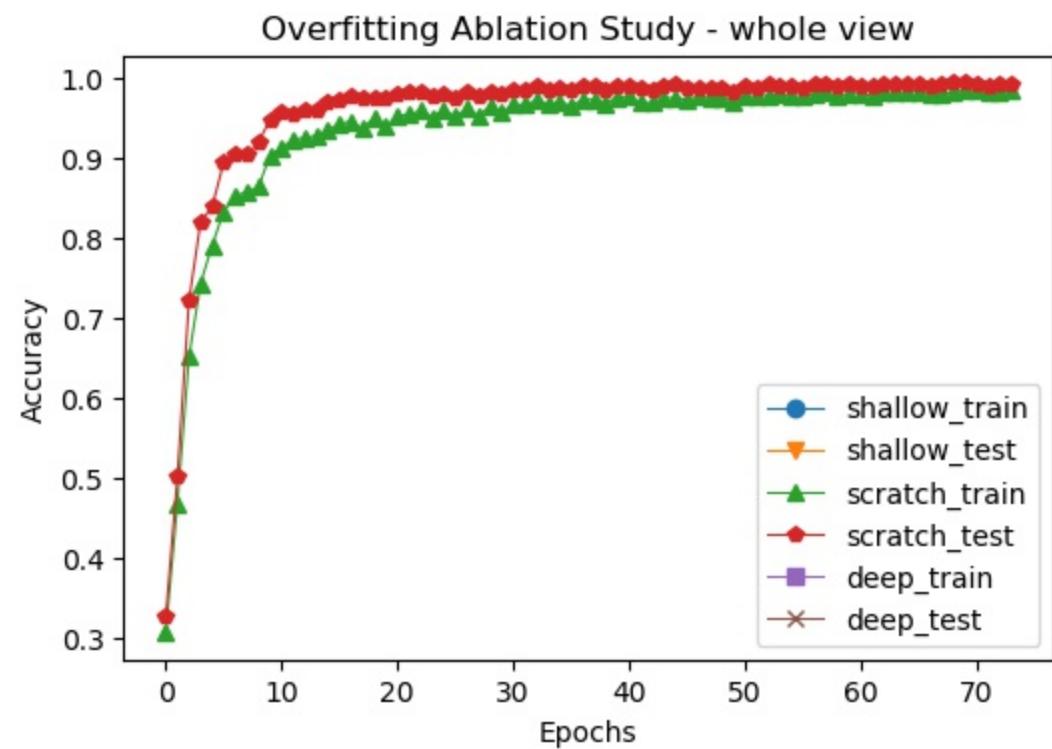
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

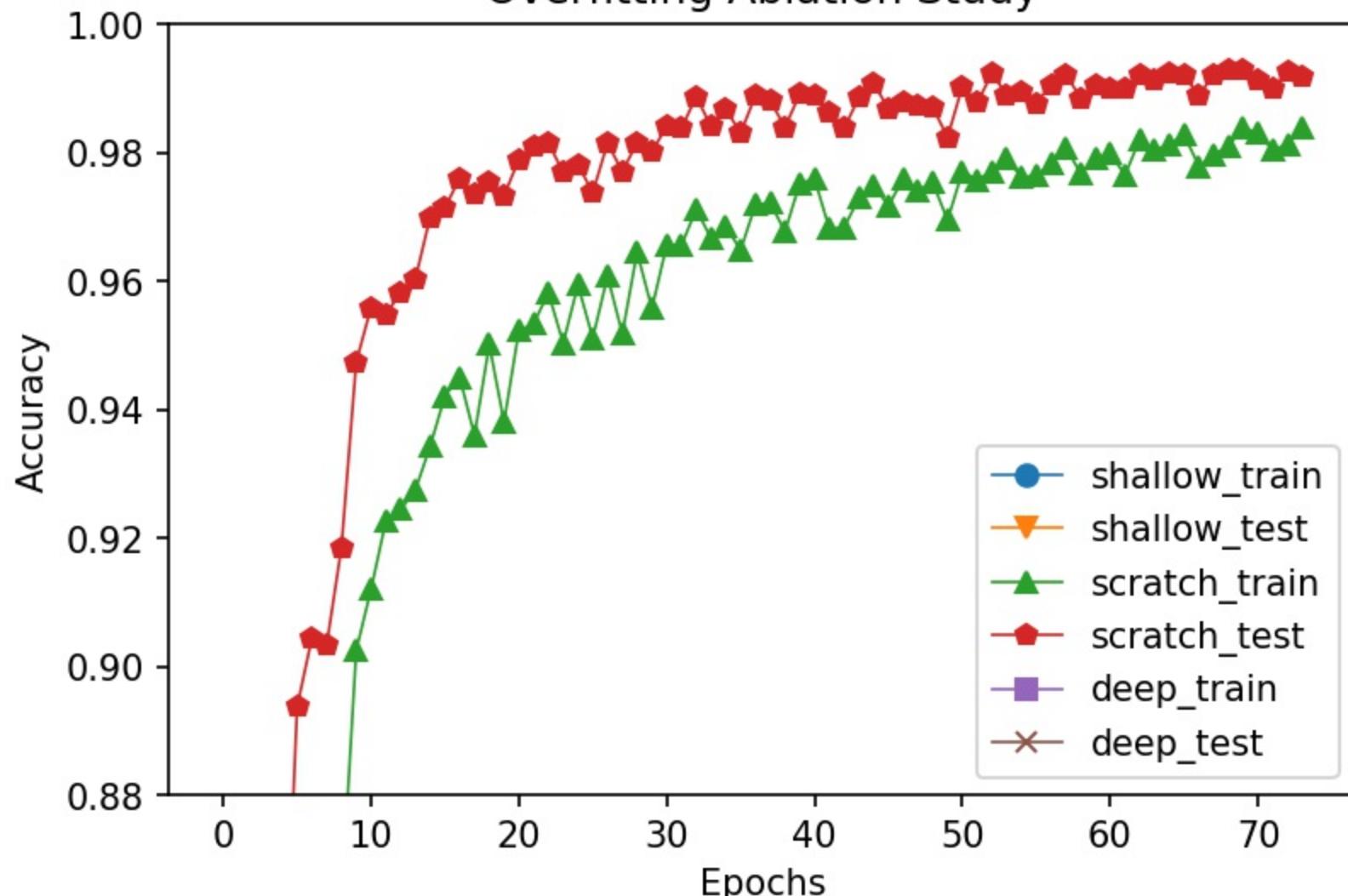
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 73 takes 315 seconds.

[75, 30] loss: 0.049

[75, 60] loss: 0.079

```
[75,    90] loss: 0.064
[75,   120] loss: 0.079
[75,   150] loss: 0.076
[75,   180] loss: 0.101
[75,   210] loss: 0.072
[75,   240] loss: 0.094
[75,   270] loss: 0.045
[75,   300] loss: 0.071
[75,   330] loss: 0.101
[75,   360] loss: 0.074
[75,   390] loss: 0.083
[75,   420] loss: 0.086
[75,   450] loss: 0.062
[75,   480] loss: 0.075
[75,   510] loss: 0.070
[75,   540] loss: 0.068
[75,   570] loss: 0.048
[75,   600] loss: 0.093
[75,   630] loss: 0.093
[75,   660] loss: 0.072
[75,   690] loss: 0.083
[75,   720] loss: 0.080
[75,   750] loss: 0.086
[75,   780] loss: 0.045
[75,   810] loss: 0.098
[75,   840] loss: 0.078
[75,   870] loss: 0.049
[75,   900] loss: 0.046
[75,   930] loss: 0.106
[75,   960] loss: 0.096
[75,   990] loss: 0.064
[75,  1020] loss: 0.077
[75,  1050] loss: 0.051
[75,  1080] loss: 0.099
[75,  1110] loss: 0.109
[75,  1140] loss: 0.069
[75,  1170] loss: 0.039
[75,  1200] loss: 0.088
[75,  1230] loss: 0.082
```

```
[75, 1260] loss: 0.085
[75, 1290] loss: 0.080
[75, 1320] loss: 0.054
[75, 1350] loss: 0.102
[75, 1380] loss: 0.065
[75, 1410] loss: 0.101
[75, 1440] loss: 0.091
[75, 1470] loss: 0.069
[75, 1500] loss: 0.070
[75, 1530] loss: 0.070
[75, 1560] loss: 0.054
[75, 1590] loss: 0.080
[75, 1620] loss: 0.067
[75, 1650] loss: 0.062
[75, 1680] loss: 0.063
[75, 1710] loss: 0.075
[75, 1740] loss: 0.076
[75, 1770] loss: 0.093
[75, 1800] loss: 0.102
[75, 1830] loss: 0.073
[75, 1860] loss: 0.054
[75, 1890] loss: 0.112
[75, 1920] loss: 0.095
[75, 1950] loss: 0.105
[75, 1980] loss: 0.076
[75, 2010] loss: 0.071
[75, 2040] loss: 0.116
[75, 2070] loss: 0.084
[75, 2100] loss: 0.098
[75, 2130] loss: 0.079
[75, 2160] loss: 0.063
[75, 2190] loss: 0.061
eval intermediate_models/epoch74_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.981711
epoch 74 accuracy on train set is: 0.9817111959287532
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991549

epoch 74 accuracy on test set is: 0.9915485278080698

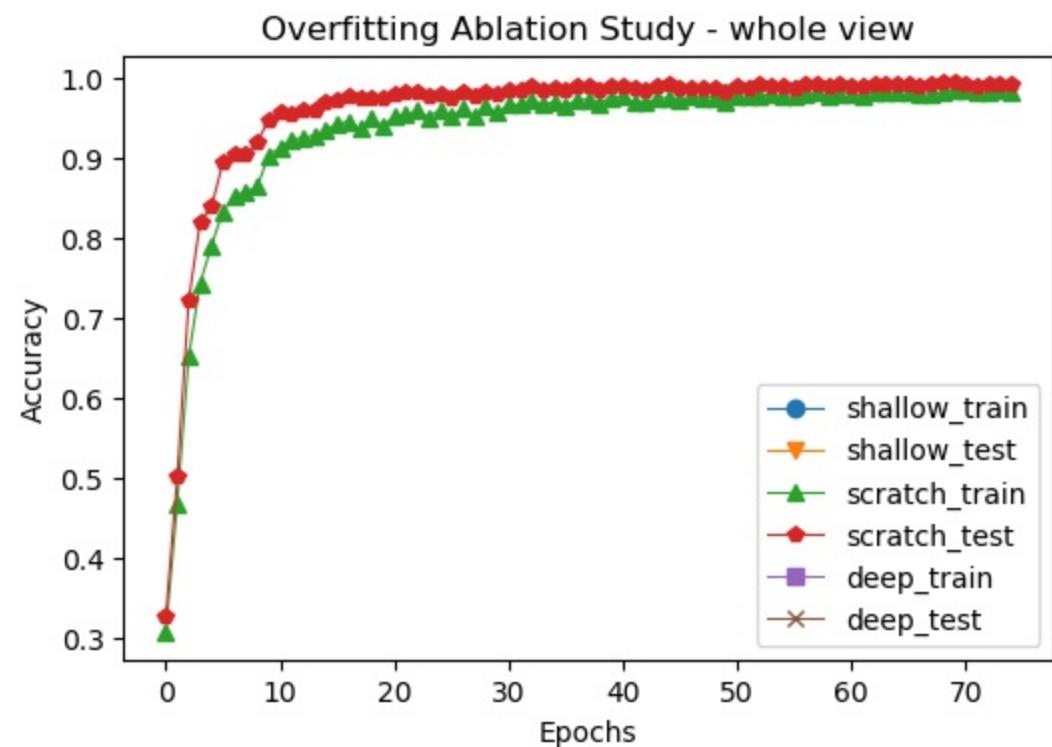
evaluating on test set takes 16 seconds.

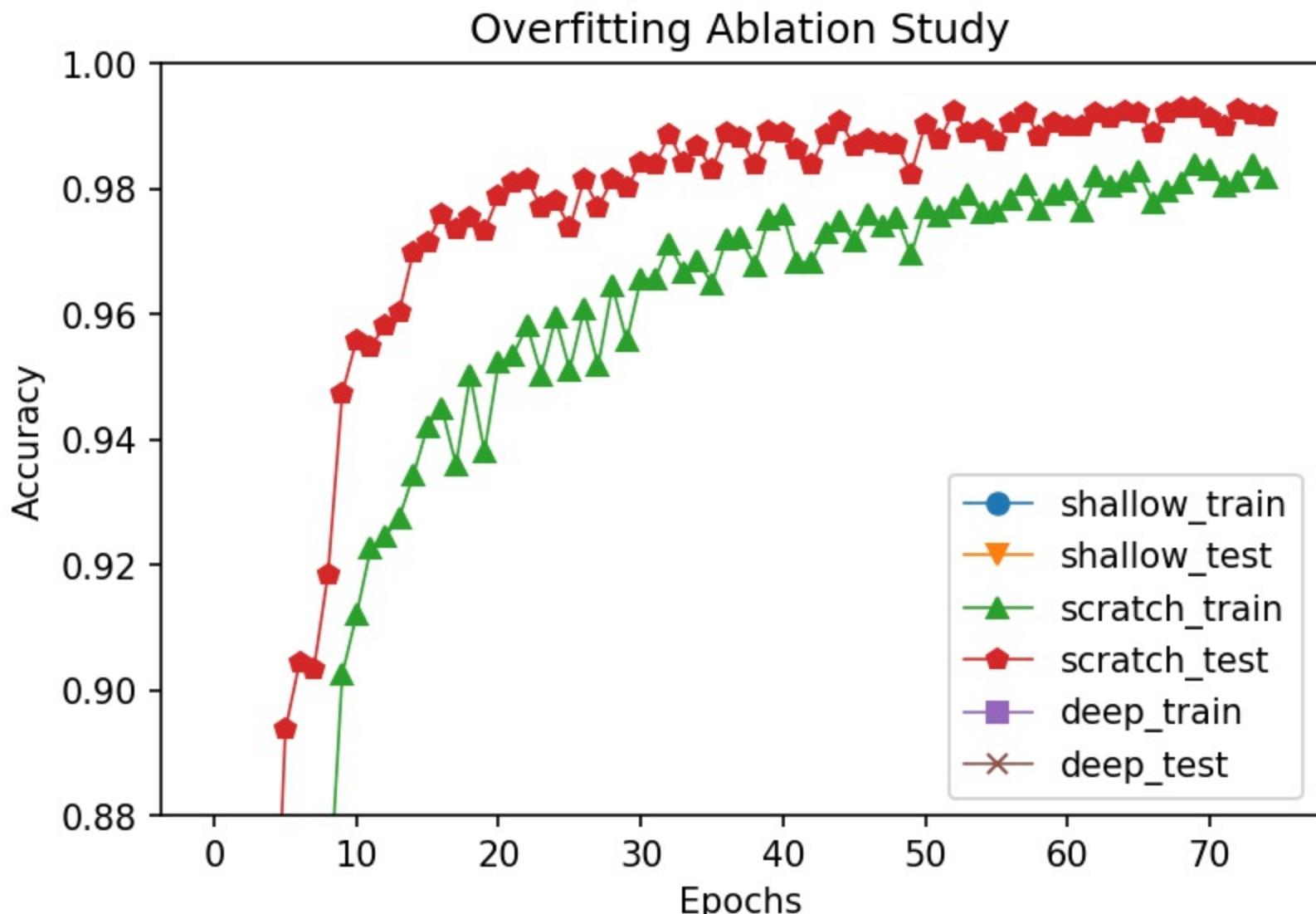
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 74 takes 315 seconds.

[76, 30] loss: 0.037

[76, 60] loss: 0.051

```
[76,    90] loss: 0.063
[76,   120] loss: 0.080
[76,   150] loss: 0.064
[76,   180] loss: 0.046
[76,   210] loss: 0.067
[76,   240] loss: 0.072
[76,   270] loss: 0.059
[76,   300] loss: 0.050
[76,   330] loss: 0.064
[76,   360] loss: 0.064
[76,   390] loss: 0.089
[76,   420] loss: 0.072
[76,   450] loss: 0.074
[76,   480] loss: 0.071
[76,   510] loss: 0.108
[76,   540] loss: 0.112
[76,   570] loss: 0.077
[76,   600] loss: 0.089
[76,   630] loss: 0.067
[76,   660] loss: 0.072
[76,   690] loss: 0.090
[76,   720] loss: 0.088
[76,   750] loss: 0.101
[76,   780] loss: 0.089
[76,   810] loss: 0.055
[76,   840] loss: 0.049
[76,   870] loss: 0.057
[76,   900] loss: 0.093
[76,   930] loss: 0.050
[76,   960] loss: 0.092
[76,   990] loss: 0.054
[76,  1020] loss: 0.047
[76,  1050] loss: 0.082
[76,  1080] loss: 0.073
[76,  1110] loss: 0.073
[76,  1140] loss: 0.078
[76,  1170] loss: 0.099
[76,  1200] loss: 0.079
[76,  1230] loss: 0.081
```

```
[76, 1260] loss: 0.039
[76, 1290] loss: 0.065
[76, 1320] loss: 0.082
[76, 1350] loss: 0.083
[76, 1380] loss: 0.088
[76, 1410] loss: 0.109
[76, 1440] loss: 0.074
[76, 1470] loss: 0.064
[76, 1500] loss: 0.038
[76, 1530] loss: 0.066
[76, 1560] loss: 0.070
[76, 1590] loss: 0.069
[76, 1620] loss: 0.094
[76, 1650] loss: 0.072
[76, 1680] loss: 0.100
[76, 1710] loss: 0.078
[76, 1740] loss: 0.068
[76, 1770] loss: 0.052
[76, 1800] loss: 0.079
[76, 1830] loss: 0.065
[76, 1860] loss: 0.094
[76, 1890] loss: 0.099
[76, 1920] loss: 0.058
[76, 1950] loss: 0.112
[76, 1980] loss: 0.095
[76, 2010] loss: 0.055
[76, 2040] loss: 0.125
[76, 2070] loss: 0.089
[76, 2100] loss: 0.094
[76, 2130] loss: 0.079
[76, 2160] loss: 0.119
[76, 2190] loss: 0.059
eval intermediate_models/epoch75_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.981416
epoch 75 accuracy on train set is: 0.981415848782261
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991276

epoch 75 accuracy on test set is: 0.9912758996728462

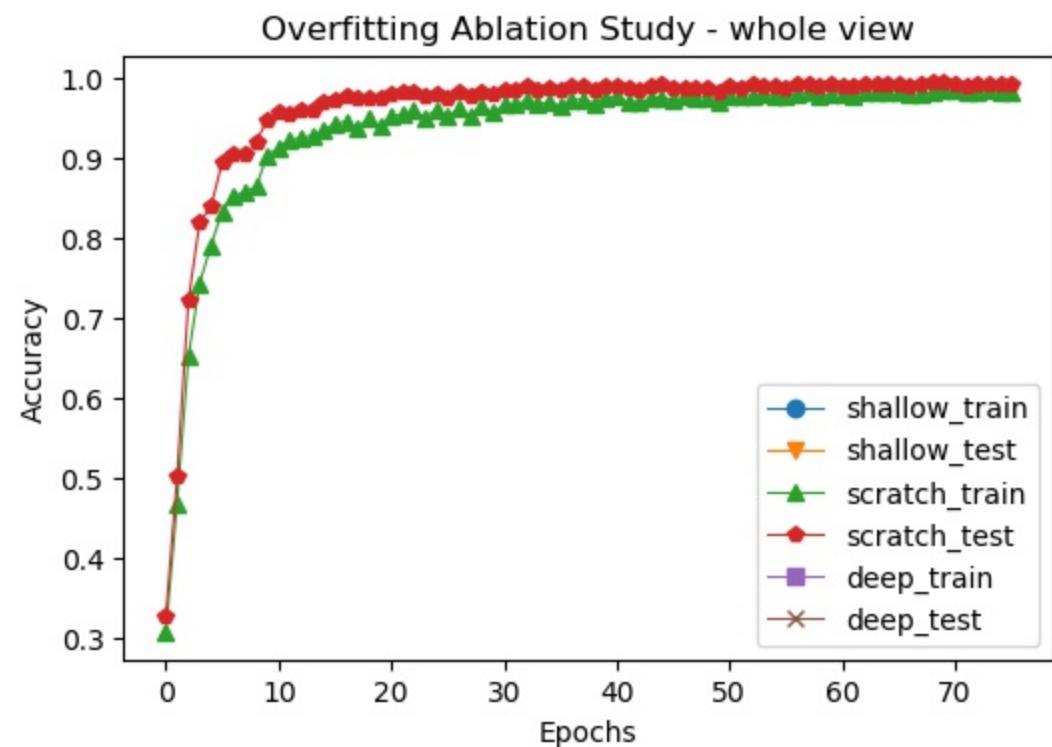
evaluating on test set takes 16 seconds.

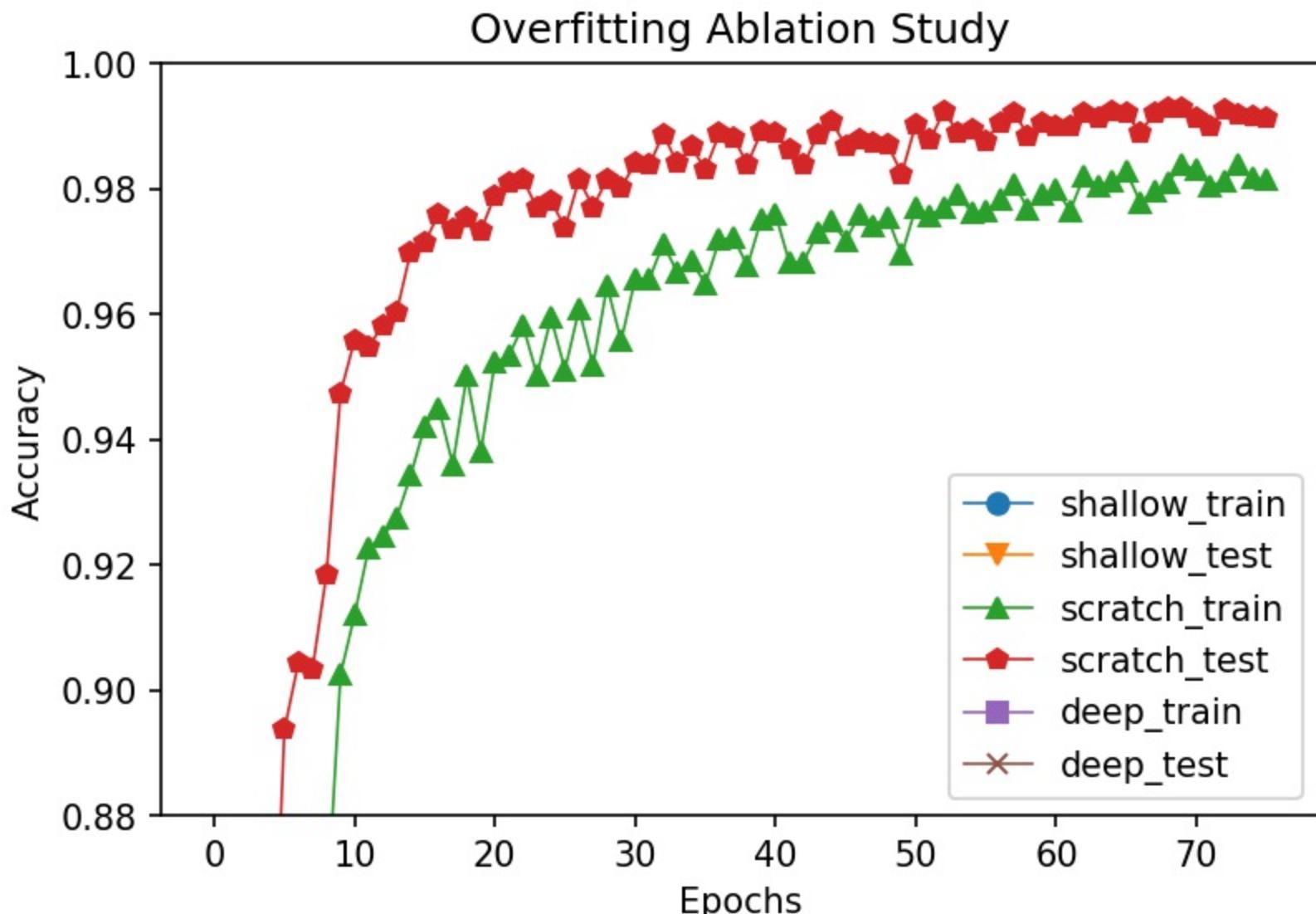
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 75 takes 315 seconds.

[77, 30] loss: 0.062

[77, 60] loss: 0.093

```
[77,    90] loss: 0.075
[77,   120] loss: 0.090
[77,   150] loss: 0.080
[77,   180] loss: 0.077
[77,   210] loss: 0.075
[77,   240] loss: 0.096
[77,   270] loss: 0.069
[77,   300] loss: 0.094
[77,   330] loss: 0.058
[77,   360] loss: 0.065
[77,   390] loss: 0.068
[77,   420] loss: 0.099
[77,   450] loss: 0.100
[77,   480] loss: 0.108
[77,   510] loss: 0.065
[77,   540] loss: 0.060
[77,   570] loss: 0.062
[77,   600] loss: 0.104
[77,   630] loss: 0.052
[77,   660] loss: 0.039
[77,   690] loss: 0.078
[77,   720] loss: 0.102
[77,   750] loss: 0.085
[77,   780] loss: 0.080
[77,   810] loss: 0.057
[77,   840] loss: 0.055
[77,   870] loss: 0.050
[77,   900] loss: 0.062
[77,   930] loss: 0.067
[77,   960] loss: 0.052
[77,   990] loss: 0.065
[77,  1020] loss: 0.052
[77,  1050] loss: 0.047
[77,  1080] loss: 0.052
[77,  1110] loss: 0.115
[77,  1140] loss: 0.064
[77,  1170] loss: 0.058
[77,  1200] loss: 0.082
[77,  1230] loss: 0.080
```

```
[77, 1260] loss: 0.038
[77, 1290] loss: 0.073
[77, 1320] loss: 0.066
[77, 1350] loss: 0.064
[77, 1380] loss: 0.089
[77, 1410] loss: 0.061
[77, 1440] loss: 0.068
[77, 1470] loss: 0.115
[77, 1500] loss: 0.091
[77, 1530] loss: 0.067
[77, 1560] loss: 0.118
[77, 1590] loss: 0.074
[77, 1620] loss: 0.062
[77, 1650] loss: 0.044
[77, 1680] loss: 0.076
[77, 1710] loss: 0.077
[77, 1740] loss: 0.062
[77, 1770] loss: 0.058
[77, 1800] loss: 0.087
[77, 1830] loss: 0.067
[77, 1860] loss: 0.094
[77, 1890] loss: 0.084
[77, 1920] loss: 0.058
[77, 1950] loss: 0.096
[77, 1980] loss: 0.106
[77, 2010] loss: 0.076
[77, 2040] loss: 0.088
[77, 2070] loss: 0.090
[77, 2100] loss: 0.062
[77, 2130] loss: 0.072
[77, 2160] loss: 0.060
[77, 2190] loss: 0.070
eval intermediate_models/epoch76_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.984642
epoch 76 accuracy on train set is: 0.9846419483824064
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991367

epoch 76 accuracy on test set is: 0.9913667757179208

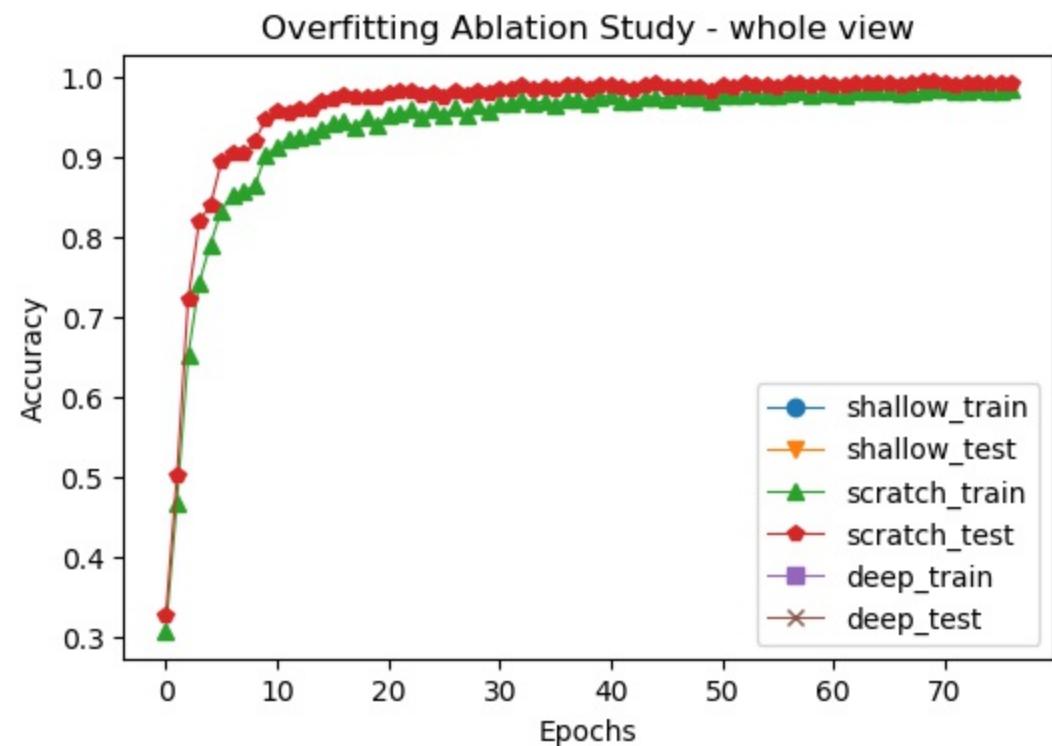
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

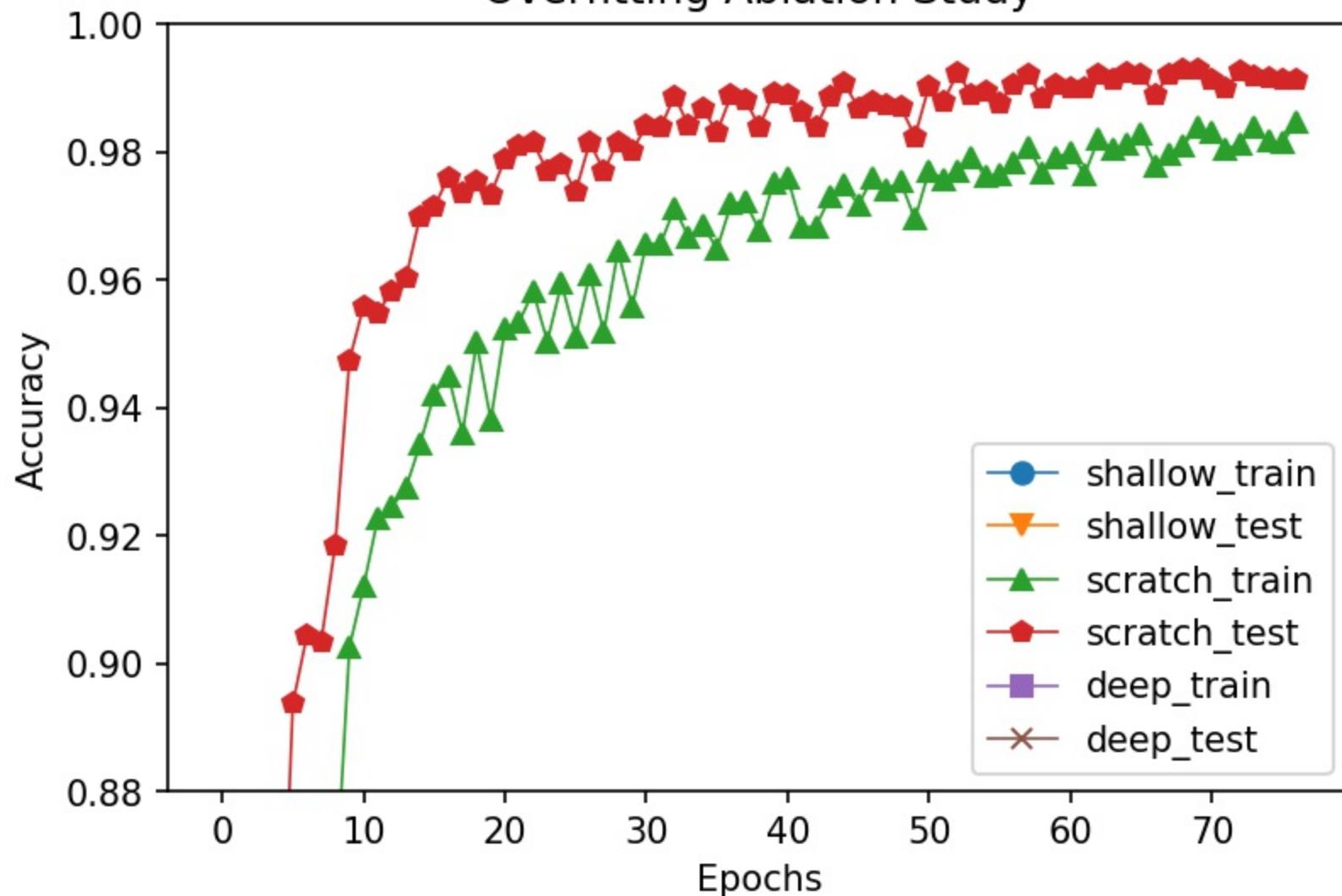
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 76 takes 316 seconds.

[78, 30] loss: 0.096

[78, 60] loss: 0.046

```
[78,    90] loss: 0.060
[78,   120] loss: 0.043
[78,   150] loss: 0.084
[78,   180] loss: 0.044
[78,   210] loss: 0.049
[78,   240] loss: 0.064
[78,   270] loss: 0.066
[78,   300] loss: 0.056
[78,   330] loss: 0.090
[78,   360] loss: 0.074
[78,   390] loss: 0.067
[78,   420] loss: 0.039
[78,   450] loss: 0.088
[78,   480] loss: 0.038
[78,   510] loss: 0.071
[78,   540] loss: 0.063
[78,   570] loss: 0.076
[78,   600] loss: 0.075
[78,   630] loss: 0.087
[78,   660] loss: 0.067
[78,   690] loss: 0.076
[78,   720] loss: 0.069
[78,   750] loss: 0.074
[78,   780] loss: 0.074
[78,   810] loss: 0.113
[78,   840] loss: 0.064
[78,   870] loss: 0.095
[78,   900] loss: 0.063
[78,   930] loss: 0.094
[78,   960] loss: 0.065
[78,   990] loss: 0.082
[78,  1020] loss: 0.061
[78,  1050] loss: 0.081
[78,  1080] loss: 0.090
[78,  1110] loss: 0.076
[78,  1140] loss: 0.056
[78,  1170] loss: 0.082
[78,  1200] loss: 0.051
[78,  1230] loss: 0.062
```

```
[78, 1260] loss: 0.095
[78, 1290] loss: 0.054
[78, 1320] loss: 0.062
[78, 1350] loss: 0.057
[78, 1380] loss: 0.049
[78, 1410] loss: 0.082
[78, 1440] loss: 0.062
[78, 1470] loss: 0.043
[78, 1500] loss: 0.075
[78, 1530] loss: 0.087
[78, 1560] loss: 0.089
[78, 1590] loss: 0.105
[78, 1620] loss: 0.091
[78, 1650] loss: 0.063
[78, 1680] loss: 0.078
[78, 1710] loss: 0.072
[78, 1740] loss: 0.095
[78, 1770] loss: 0.100
[78, 1800] loss: 0.080
[78, 1830] loss: 0.080
[78, 1860] loss: 0.077
[78, 1890] loss: 0.094
[78, 1920] loss: 0.083
[78, 1950] loss: 0.058
[78, 1980] loss: 0.078
[78, 2010] loss: 0.074
[78, 2040] loss: 0.108
[78, 2070] loss: 0.084
[78, 2100] loss: 0.095
[78, 2130] loss: 0.059
[78, 2160] loss: 0.058
[78, 2190] loss: 0.061
eval intermediate_models/epoch77_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.984642
epoch 77 accuracy on train set is: 0.9846419483824064
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	NaN	NaN	NaN

78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993184

epoch 77 accuracy on test set is: 0.9931842966194111

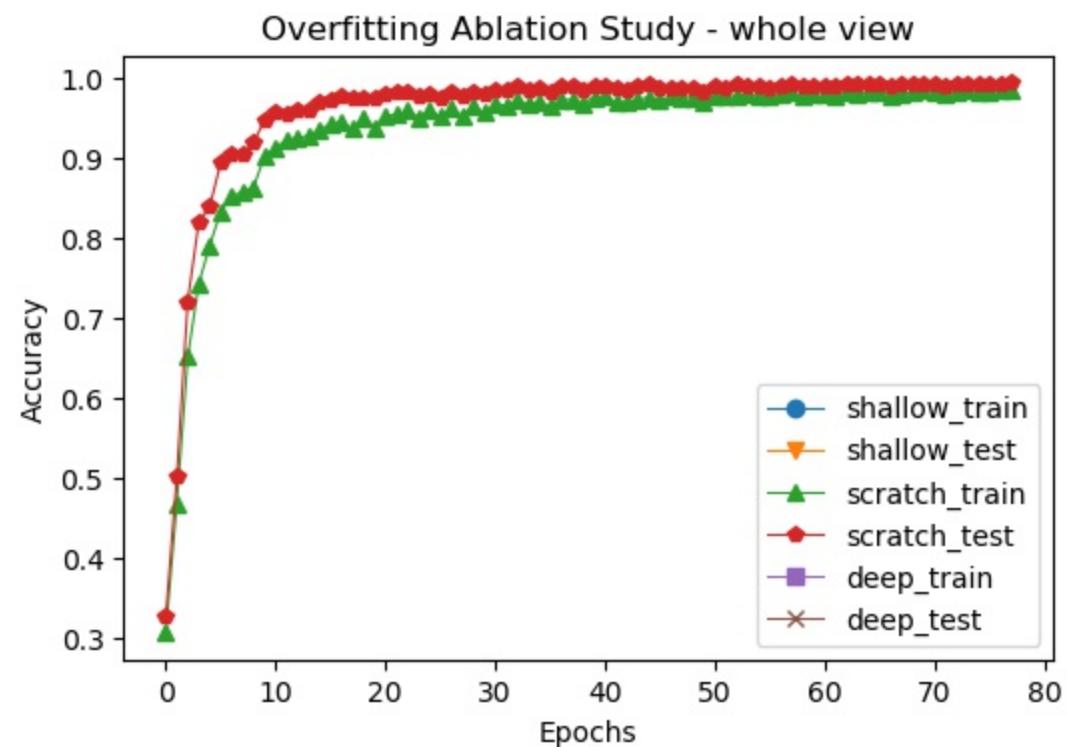
evaluating on test set takes 17 seconds.

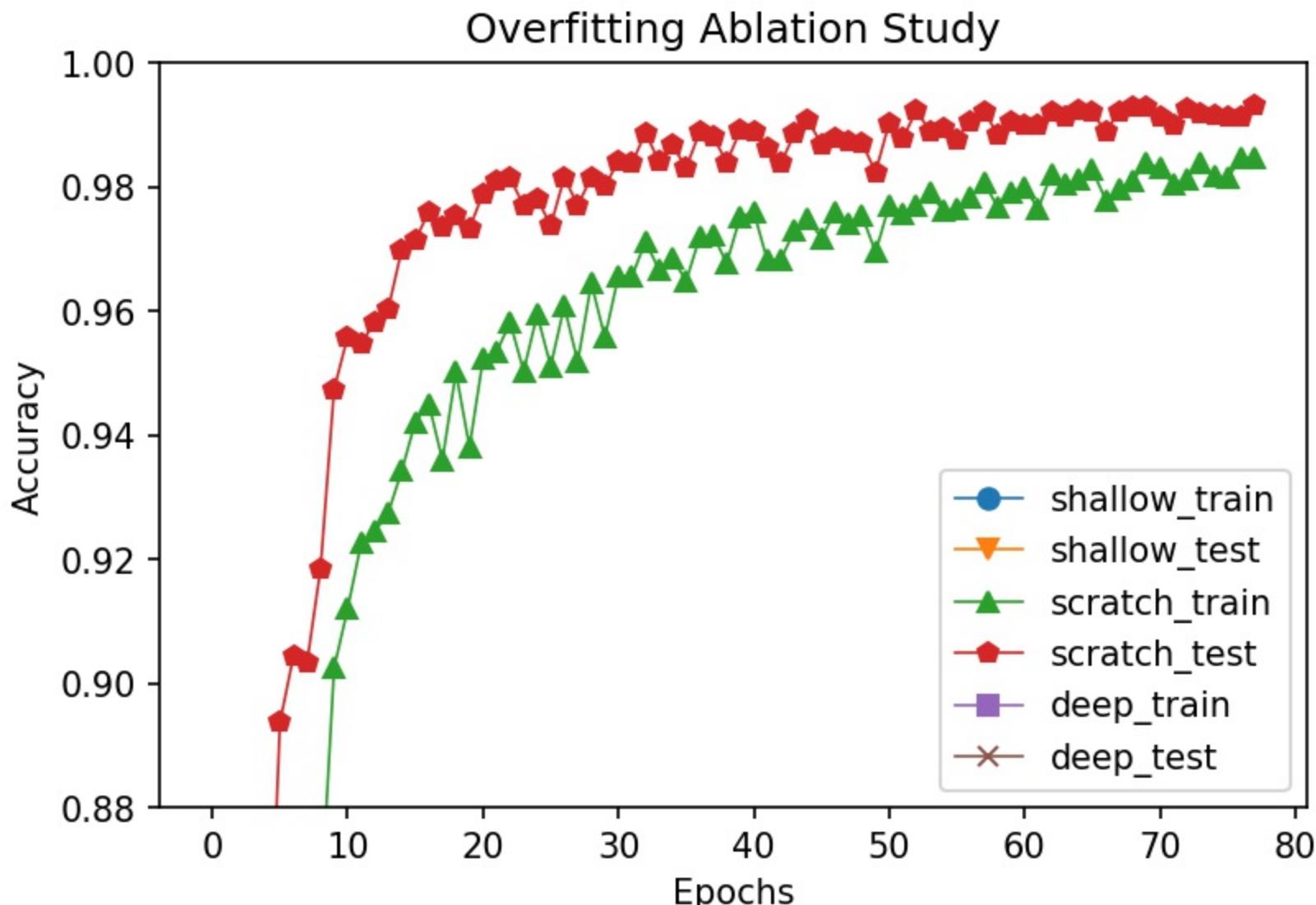
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 77 takes 317 seconds.

[79, 30] loss: 0.108

[79, 60] loss: 0.081

```
[79,    90] loss: 0.073
[79,   120] loss: 0.092
[79,   150] loss: 0.064
[79,   180] loss: 0.086
[79,   210] loss: 0.095
[79,   240] loss: 0.079
[79,   270] loss: 0.089
[79,   300] loss: 0.071
[79,   330] loss: 0.091
[79,   360] loss: 0.075
[79,   390] loss: 0.083
[79,   420] loss: 0.056
[79,   450] loss: 0.062
[79,   480] loss: 0.048
[79,   510] loss: 0.083
[79,   540] loss: 0.088
[79,   570] loss: 0.067
[79,   600] loss: 0.114
[79,   630] loss: 0.107
[79,   660] loss: 0.110
[79,   690] loss: 0.083
[79,   720] loss: 0.034
[79,   750] loss: 0.086
[79,   780] loss: 0.045
[79,   810] loss: 0.080
[79,   840] loss: 0.064
[79,   870] loss: 0.059
[79,   900] loss: 0.038
[79,   930] loss: 0.086
[79,   960] loss: 0.066
[79,   990] loss: 0.065
[79,  1020] loss: 0.077
[79,  1050] loss: 0.106
[79,  1080] loss: 0.077
[79,  1110] loss: 0.069
[79,  1140] loss: 0.047
[79,  1170] loss: 0.065
[79,  1200] loss: 0.053
[79,  1230] loss: 0.066
```

```
[79, 1260] loss: 0.046
[79, 1290] loss: 0.045
[79, 1320] loss: 0.063
[79, 1350] loss: 0.105
[79, 1380] loss: 0.073
[79, 1410] loss: 0.072
[79, 1440] loss: 0.056
[79, 1470] loss: 0.048
[79, 1500] loss: 0.063
[79, 1530] loss: 0.036
[79, 1560] loss: 0.064
[79, 1590] loss: 0.067
[79, 1620] loss: 0.087
[79, 1650] loss: 0.031
[79, 1680] loss: 0.063
[79, 1710] loss: 0.101
[79, 1740] loss: 0.086
[79, 1770] loss: 0.080
[79, 1800] loss: 0.062
[79, 1830] loss: 0.056
[79, 1860] loss: 0.070
[79, 1890] loss: 0.044
[79, 1920] loss: 0.073
[79, 1950] loss: 0.099
[79, 1980] loss: 0.085
[79, 2010] loss: 0.067
[79, 2040] loss: 0.068
[79, 2070] loss: 0.038
[79, 2100] loss: 0.050
[79, 2130] loss: 0.052
[79, 2160] loss: 0.051
[79, 2190] loss: 0.125
eval intermediate_models/epoch78_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.976986
epoch 78 accuracy on train set is: 0.9769856415848782
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988640

epoch 78 accuracy on test set is: 0.9886404943656852

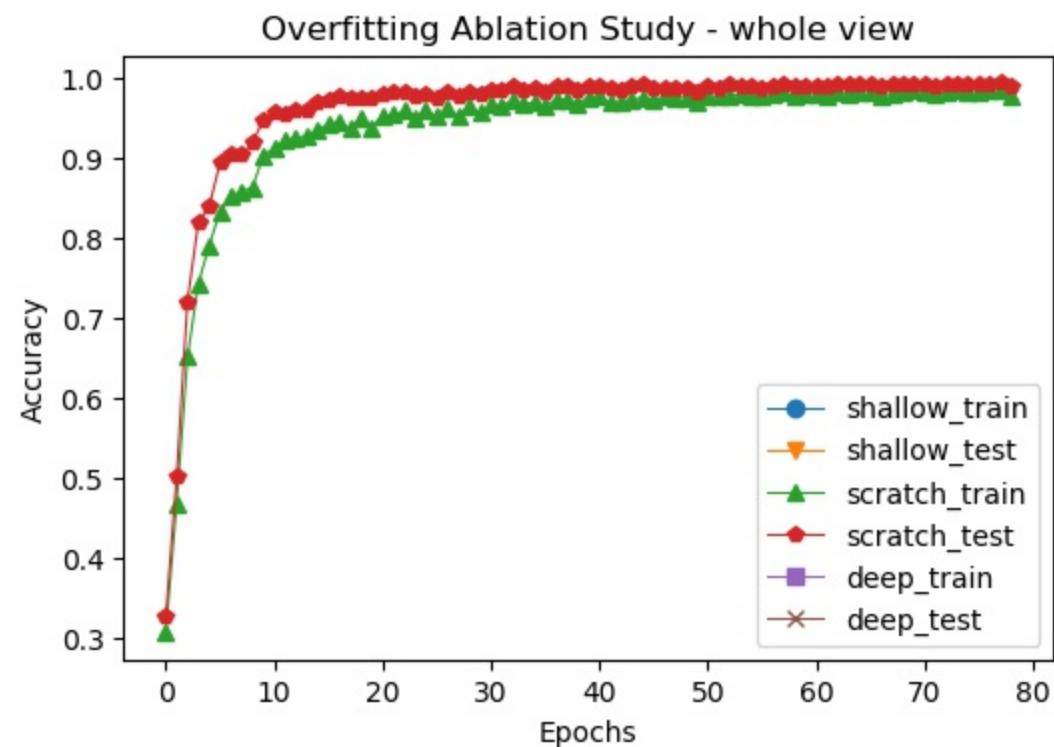
evaluating on test set takes 16 seconds.

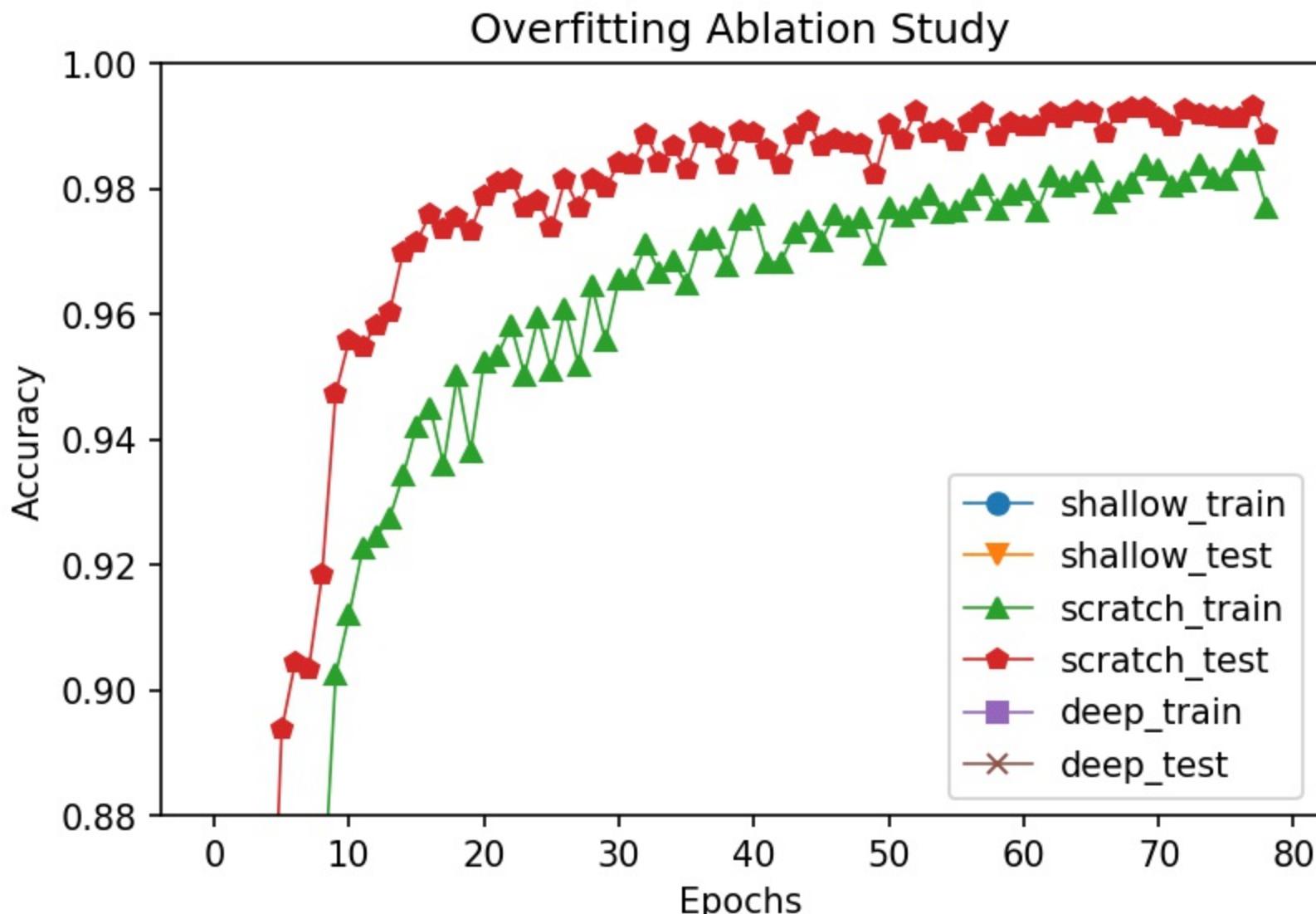
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 78 takes 318 seconds.

[80, 30] loss: 0.068

[80, 60] loss: 0.062

```
[80,    90] loss: 0.082
[80,   120] loss: 0.087
[80,   150] loss: 0.107
[80,   180] loss: 0.059
[80,   210] loss: 0.076
[80,   240] loss: 0.093
[80,   270] loss: 0.076
[80,   300] loss: 0.054
[80,   330] loss: 0.029
[80,   360] loss: 0.082
[80,   390] loss: 0.071
[80,   420] loss: 0.057
[80,   450] loss: 0.057
[80,   480] loss: 0.073
[80,   510] loss: 0.054
[80,   540] loss: 0.077
[80,   570] loss: 0.081
[80,   600] loss: 0.069
[80,   630] loss: 0.065
[80,   660] loss: 0.068
[80,   690] loss: 0.053
[80,   720] loss: 0.088
[80,   750] loss: 0.054
[80,   780] loss: 0.039
[80,   810] loss: 0.066
[80,   840] loss: 0.106
[80,   870] loss: 0.117
[80,   900] loss: 0.054
[80,   930] loss: 0.067
[80,   960] loss: 0.056
[80,   990] loss: 0.048
[80,  1020] loss: 0.058
[80,  1050] loss: 0.055
[80,  1080] loss: 0.044
[80,  1110] loss: 0.056
[80,  1140] loss: 0.057
[80,  1170] loss: 0.086
[80,  1200] loss: 0.072
[80,  1230] loss: 0.096
```

```
[80, 1260] loss: 0.077
[80, 1290] loss: 0.089
[80, 1320] loss: 0.098
[80, 1350] loss: 0.088
[80, 1380] loss: 0.049
[80, 1410] loss: 0.066
[80, 1440] loss: 0.064
[80, 1470] loss: 0.062
[80, 1500] loss: 0.043
[80, 1530] loss: 0.084
[80, 1560] loss: 0.028
[80, 1590] loss: 0.092
[80, 1620] loss: 0.104
[80, 1650] loss: 0.095
[80, 1680] loss: 0.087
[80, 1710] loss: 0.085
[80, 1740] loss: 0.088
[80, 1770] loss: 0.055
[80, 1800] loss: 0.071
[80, 1830] loss: 0.047
[80, 1860] loss: 0.057
[80, 1890] loss: 0.048
[80, 1920] loss: 0.062
[80, 1950] loss: 0.069
[80, 1980] loss: 0.053
[80, 2010] loss: 0.080
[80, 2040] loss: 0.064
[80, 2070] loss: 0.084
[80, 2100] loss: 0.075
[80, 2130] loss: 0.041
[80, 2160] loss: 0.087
[80, 2190] loss: 0.037
evf intermediate_models/epoch79_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.983120
epoch 79 accuracy on train set is: 0.9831197746274082
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	NaN	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993184

epoch 79 accuracy on test set is: 0.9931842966194111

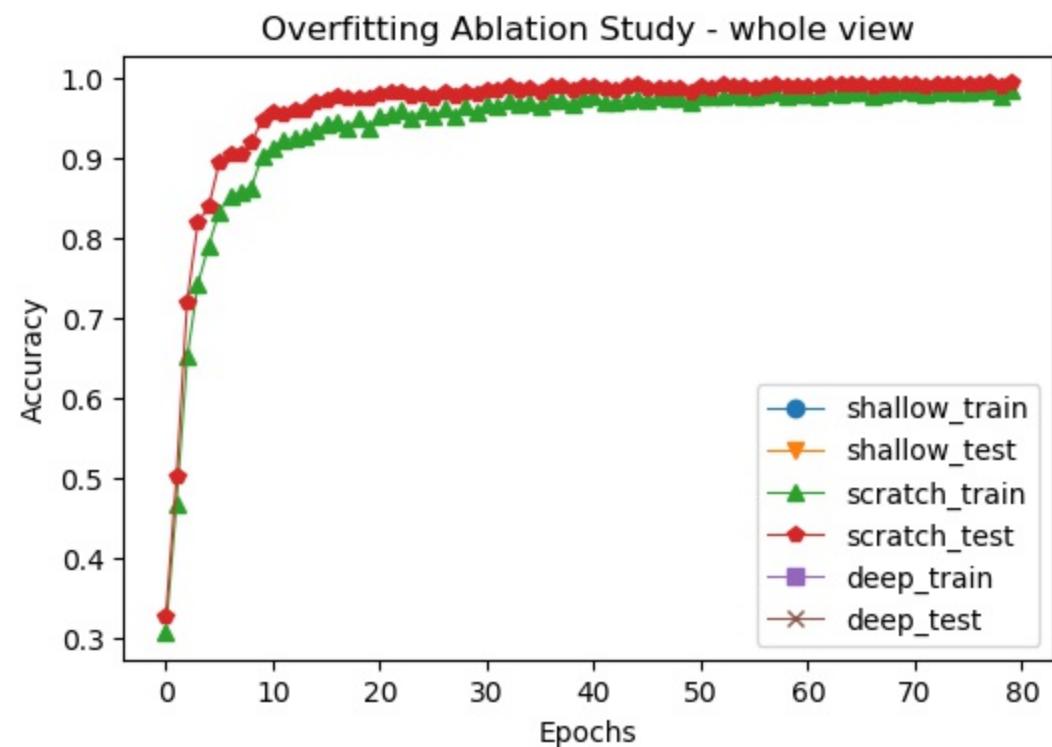
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

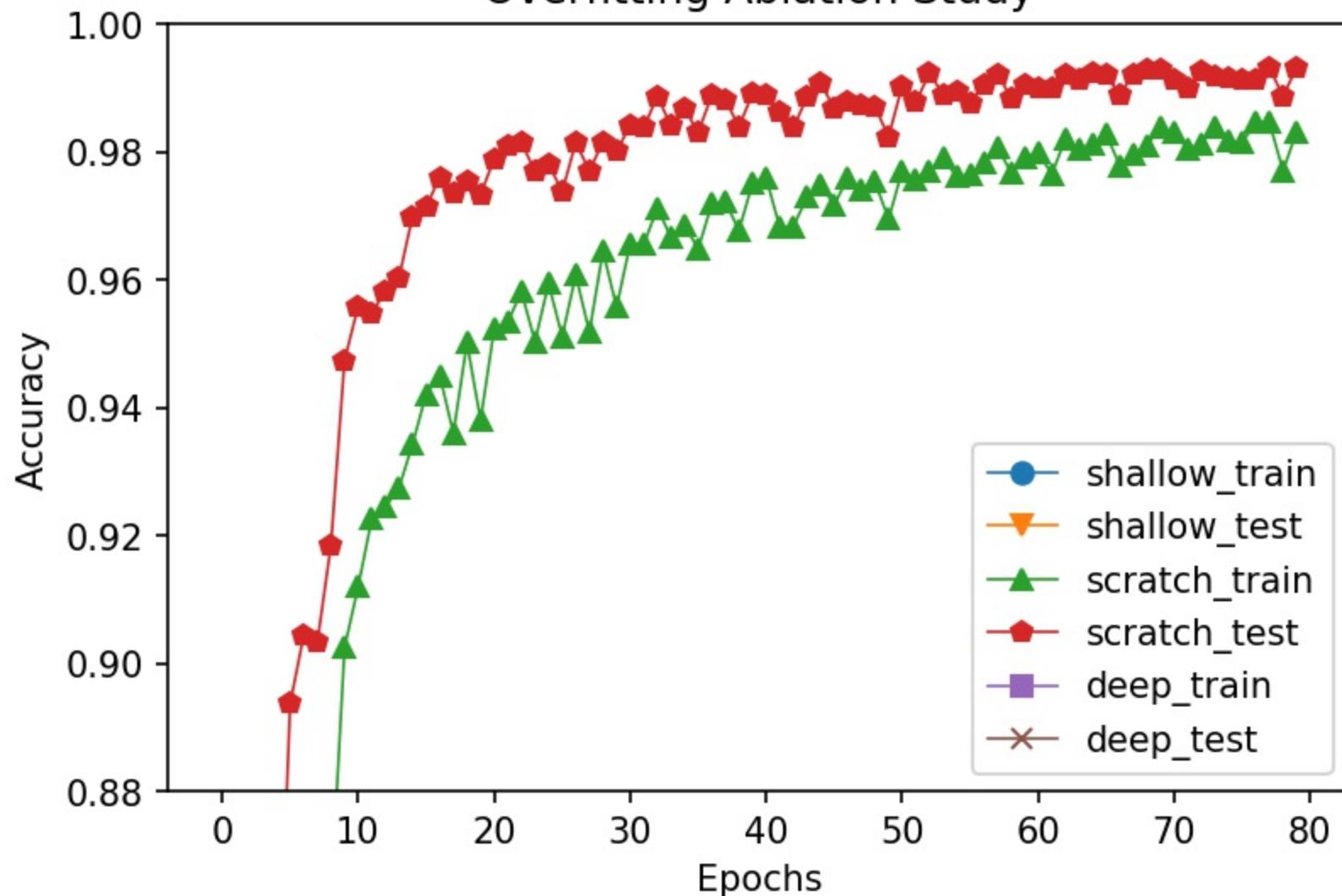
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	NaN	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 79 takes 316 seconds.

[81, 30] loss: 0.075

[81, 60] loss: 0.057

```
[81,    90] loss: 0.088
[81,   120] loss: 0.072
[81,   150] loss: 0.072
[81,   180] loss: 0.129
[81,   210] loss: 0.115
[81,   240] loss: 0.050
[81,   270] loss: 0.075
[81,   300] loss: 0.074
[81,   330] loss: 0.058
[81,   360] loss: 0.105
[81,   390] loss: 0.086
[81,   420] loss: 0.032
[81,   450] loss: 0.084
[81,   480] loss: 0.062
[81,   510] loss: 0.041
[81,   540] loss: 0.082
[81,   570] loss: 0.054
[81,   600] loss: 0.081
[81,   630] loss: 0.056
[81,   660] loss: 0.048
[81,   690] loss: 0.080
[81,   720] loss: 0.091
[81,   750] loss: 0.089
[81,   780] loss: 0.075
[81,   810] loss: 0.070
[81,   840] loss: 0.092
[81,   870] loss: 0.087
[81,   900] loss: 0.062
[81,   930] loss: 0.044
[81,   960] loss: 0.082
[81,   990] loss: 0.053
[81,  1020] loss: 0.060
[81,  1050] loss: 0.043
[81,  1080] loss: 0.054
[81,  1110] loss: 0.044
[81,  1140] loss: 0.045
[81,  1170] loss: 0.054
[81,  1200] loss: 0.051
[81,  1230] loss: 0.065
```

```
[81, 1260] loss: 0.090
[81, 1290] loss: 0.122
[81, 1320] loss: 0.112
[81, 1350] loss: 0.089
[81, 1380] loss: 0.066
[81, 1410] loss: 0.090
[81, 1440] loss: 0.087
[81, 1470] loss: 0.080
[81, 1500] loss: 0.100
[81, 1530] loss: 0.059
[81, 1560] loss: 0.088
[81, 1590] loss: 0.071
[81, 1620] loss: 0.076
[81, 1650] loss: 0.089
[81, 1680] loss: 0.080
[81, 1710] loss: 0.066
[81, 1740] loss: 0.044
[81, 1770] loss: 0.029
[81, 1800] loss: 0.067
[81, 1830] loss: 0.044
[81, 1860] loss: 0.078
[81, 1890] loss: 0.109
[81, 1920] loss: 0.104
[81, 1950] loss: 0.043
[81, 1980] loss: 0.063
[81, 2010] loss: 0.080
[81, 2040] loss: 0.083
[81, 2070] loss: 0.099
[81, 2100] loss: 0.045
[81, 2130] loss: 0.074
[81, 2160] loss: 0.084
[81, 2190] loss: 0.063
eval intermediate_models/epoch80_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.982529
epoch 80 accuracy on train set is: 0.9825290803344239
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	NaN	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991821

epoch 80 accuracy on test set is: 0.9918211559432933

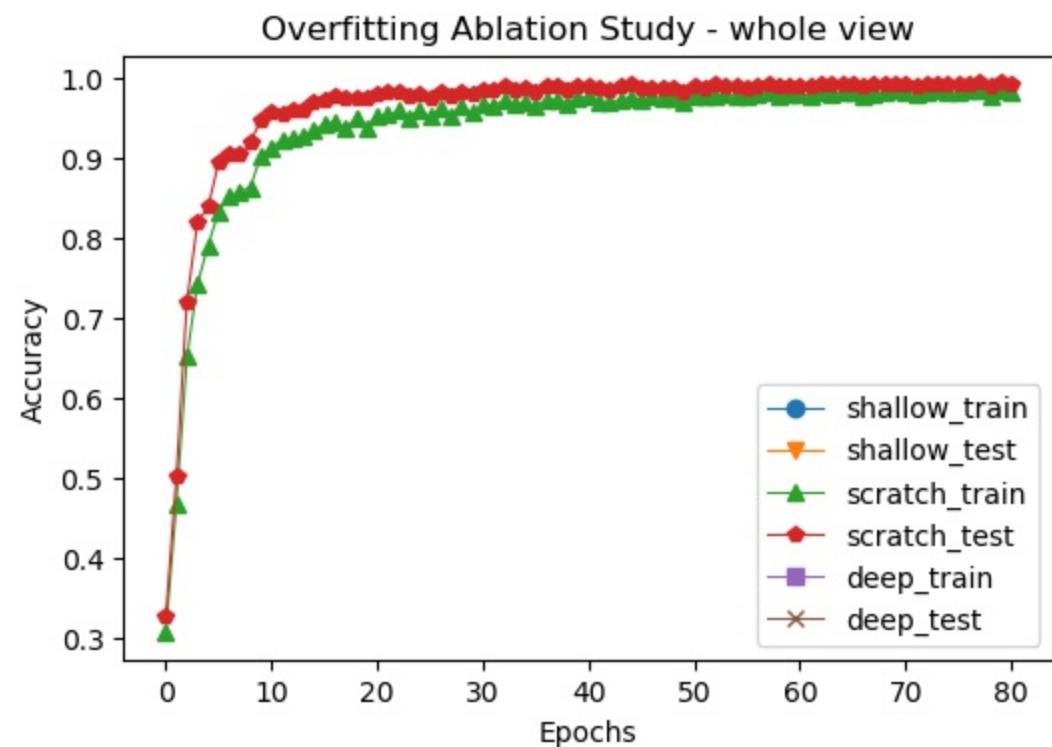
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

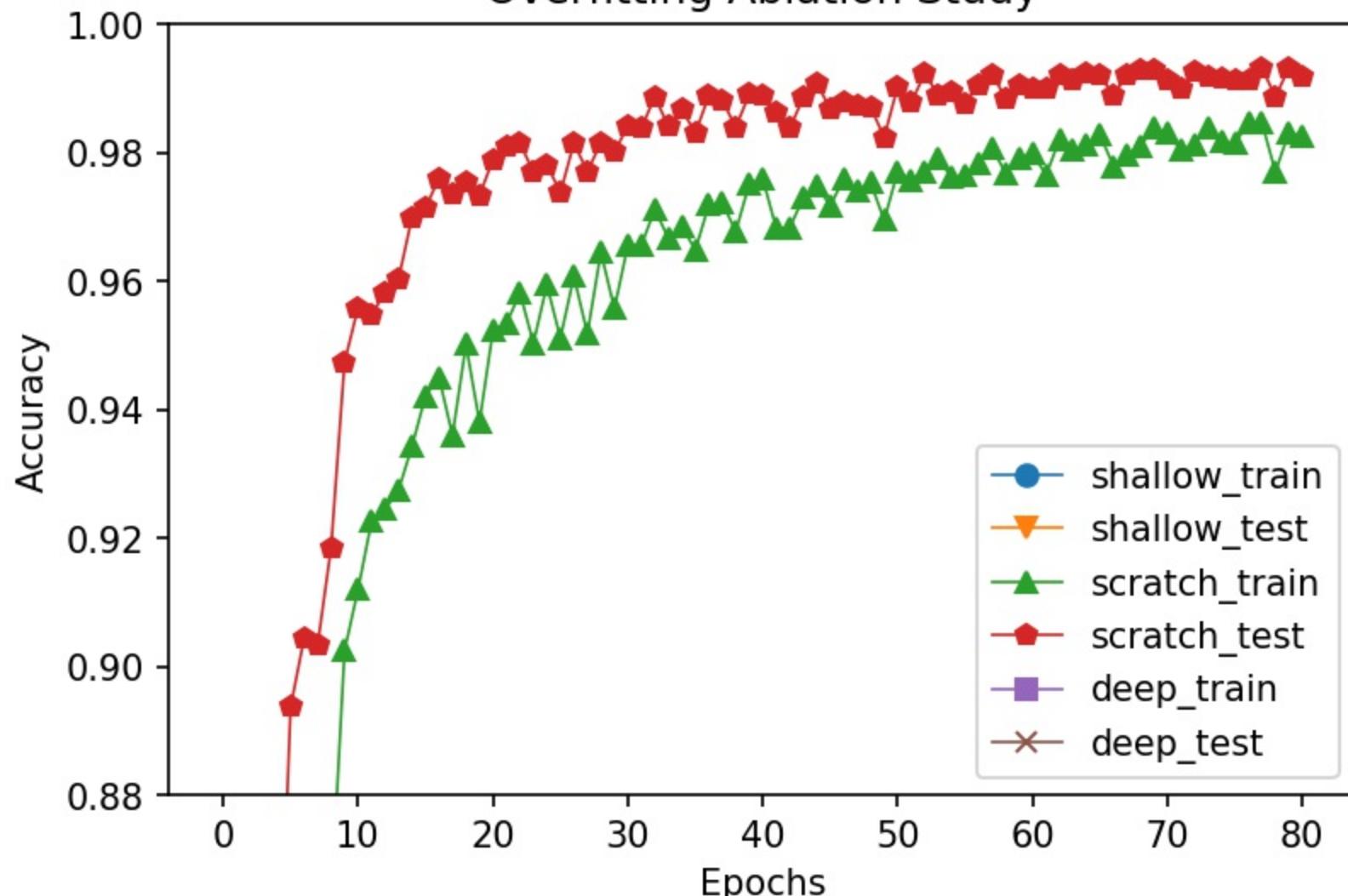
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	NaN	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 80 takes 315 seconds.

[82, 30] loss: 0.062

[82, 60] loss: 0.076

```
[82,    90] loss: 0.083
[82,   120] loss: 0.079
[82,   150] loss: 0.045
[82,   180] loss: 0.060
[82,   210] loss: 0.069
[82,   240] loss: 0.088
[82,   270] loss: 0.045
[82,   300] loss: 0.091
[82,   330] loss: 0.110
[82,   360] loss: 0.064
[82,   390] loss: 0.043
[82,   420] loss: 0.055
[82,   450] loss: 0.092
[82,   480] loss: 0.084
[82,   510] loss: 0.051
[82,   540] loss: 0.053
[82,   570] loss: 0.051
[82,   600] loss: 0.071
[82,   630] loss: 0.071
[82,   660] loss: 0.051
[82,   690] loss: 0.069
[82,   720] loss: 0.072
[82,   750] loss: 0.114
[82,   780] loss: 0.086
[82,   810] loss: 0.080
[82,   840] loss: 0.082
[82,   870] loss: 0.070
[82,   900] loss: 0.123
[82,   930] loss: 0.072
[82,   960] loss: 0.067
[82,   990] loss: 0.093
[82,  1020] loss: 0.065
[82,  1050] loss: 0.092
[82,  1080] loss: 0.055
[82,  1110] loss: 0.073
[82,  1140] loss: 0.059
[82,  1170] loss: 0.058
[82,  1200] loss: 0.069
[82,  1230] loss: 0.064
```

```
[82, 1260] loss: 0.052
[82, 1290] loss: 0.070
[82, 1320] loss: 0.105
[82, 1350] loss: 0.067
[82, 1380] loss: 0.044
[82, 1410] loss: 0.063
[82, 1440] loss: 0.081
[82, 1470] loss: 0.097
[82, 1500] loss: 0.039
[82, 1530] loss: 0.049
[82, 1560] loss: 0.065
[82, 1590] loss: 0.066
[82, 1620] loss: 0.071
[82, 1650] loss: 0.075
[82, 1680] loss: 0.067
[82, 1710] loss: 0.091
[82, 1740] loss: 0.066
[82, 1770] loss: 0.049
[82, 1800] loss: 0.089
[82, 1830] loss: 0.104
[82, 1860] loss: 0.113
[82, 1890] loss: 0.074
[82, 1920] loss: 0.052
[82, 1950] loss: 0.040
[82, 1980] loss: 0.096
[82, 2010] loss: 0.070
[82, 2040] loss: 0.034
[82, 2070] loss: 0.068
[82, 2100] loss: 0.064
[82, 2130] loss: 0.073
[82, 2160] loss: 0.057
[82, 2190] loss: 0.061
eval intermediate_models/epoch81_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.982711
epoch 81 accuracy on train set is: 0.9827108324245729
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	NaN	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992730

epoch 81 accuracy on test set is: 0.9927299163940385

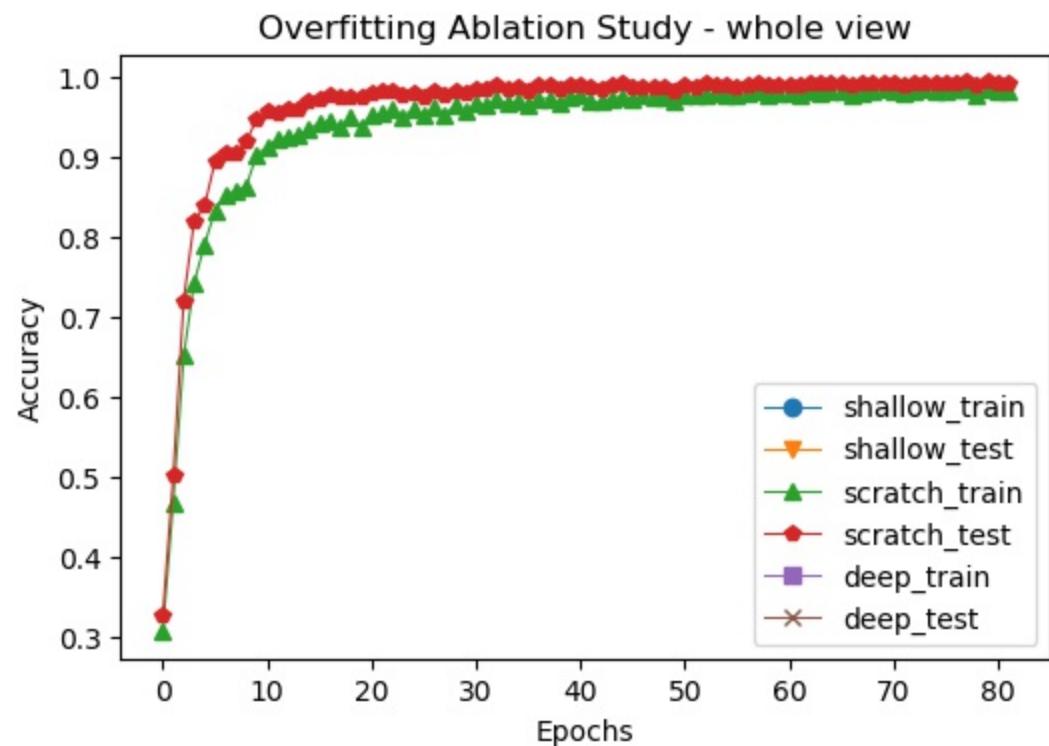
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

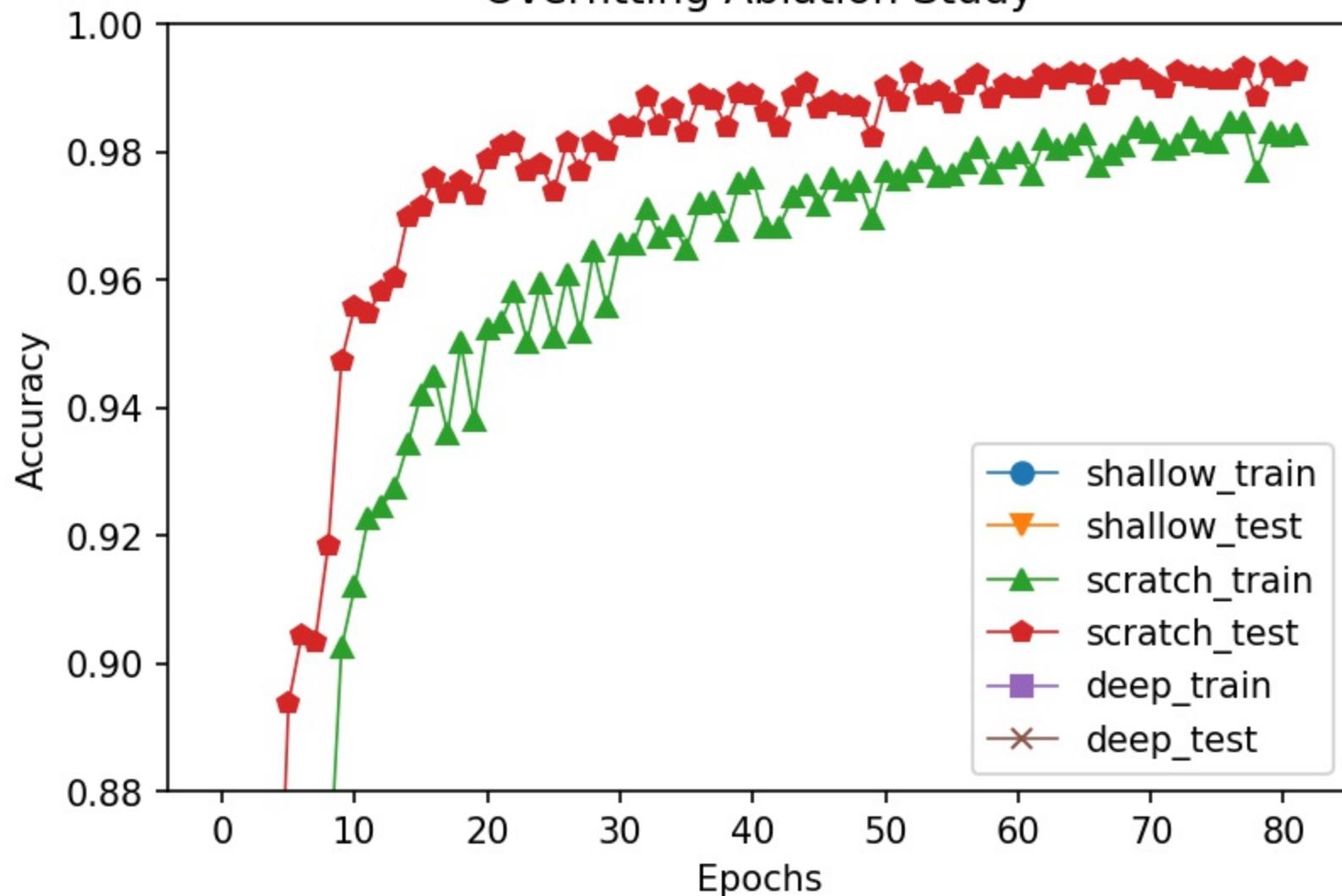
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	NaN	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 81 takes 315 seconds.

[83, 30] loss: 0.084

[83, 60] loss: 0.060

```
[83,    90] loss: 0.072
[83,   120] loss: 0.094
[83,   150] loss: 0.061
[83,   180] loss: 0.061
[83,   210] loss: 0.060
[83,   240] loss: 0.069
[83,   270] loss: 0.073
[83,   300] loss: 0.064
[83,   330] loss: 0.090
[83,   360] loss: 0.048
[83,   390] loss: 0.073
[83,   420] loss: 0.058
[83,   450] loss: 0.100
[83,   480] loss: 0.070
[83,   510] loss: 0.097
[83,   540] loss: 0.097
[83,   570] loss: 0.078
[83,   600] loss: 0.086
[83,   630] loss: 0.054
[83,   660] loss: 0.061
[83,   690] loss: 0.090
[83,   720] loss: 0.054
[83,   750] loss: 0.071
[83,   780] loss: 0.070
[83,   810] loss: 0.087
[83,   840] loss: 0.084
[83,   870] loss: 0.078
[83,   900] loss: 0.091
[83,   930] loss: 0.082
[83,   960] loss: 0.052
[83,   990] loss: 0.042
[83,  1020] loss: 0.058
[83,  1050] loss: 0.079
[83,  1080] loss: 0.093
[83,  1110] loss: 0.054
[83,  1140] loss: 0.069
[83,  1170] loss: 0.072
[83,  1200] loss: 0.022
[83,  1230] loss: 0.078
```

```
[83, 1260] loss: 0.047
[83, 1290] loss: 0.101
[83, 1320] loss: 0.048
[83, 1350] loss: 0.072
[83, 1380] loss: 0.052
[83, 1410] loss: 0.076
[83, 1440] loss: 0.076
[83, 1470] loss: 0.088
[83, 1500] loss: 0.049
[83, 1530] loss: 0.059
[83, 1560] loss: 0.089
[83, 1590] loss: 0.043
[83, 1620] loss: 0.050
[83, 1650] loss: 0.081
[83, 1680] loss: 0.057
[83, 1710] loss: 0.059
[83, 1740] loss: 0.081
[83, 1770] loss: 0.044
[83, 1800] loss: 0.050
[83, 1830] loss: 0.070
[83, 1860] loss: 0.068
[83, 1890] loss: 0.075
[83, 1920] loss: 0.079
[83, 1950] loss: 0.060
[83, 1980] loss: 0.062
[83, 2010] loss: 0.082
[83, 2040] loss: 0.055
[83, 2070] loss: 0.054
[83, 2100] loss: 0.102
[83, 2130] loss: 0.049
[83, 2160] loss: 0.065
[83, 2190] loss: 0.081
eval intermediate_models/epoch82_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.983279
epoch 82 accuracy on train set is: 0.9832788077062886
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	NaN	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.994184

epoch 82 accuracy on test set is: 0.9941839331152308

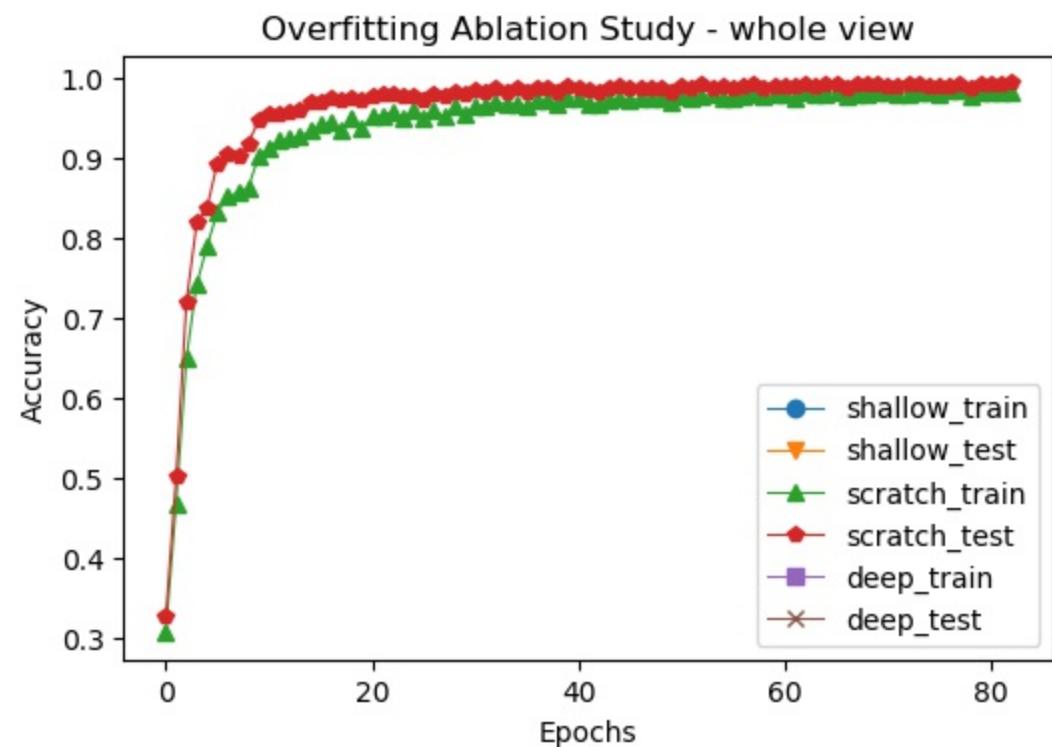
evaluating on test set takes 16 seconds.

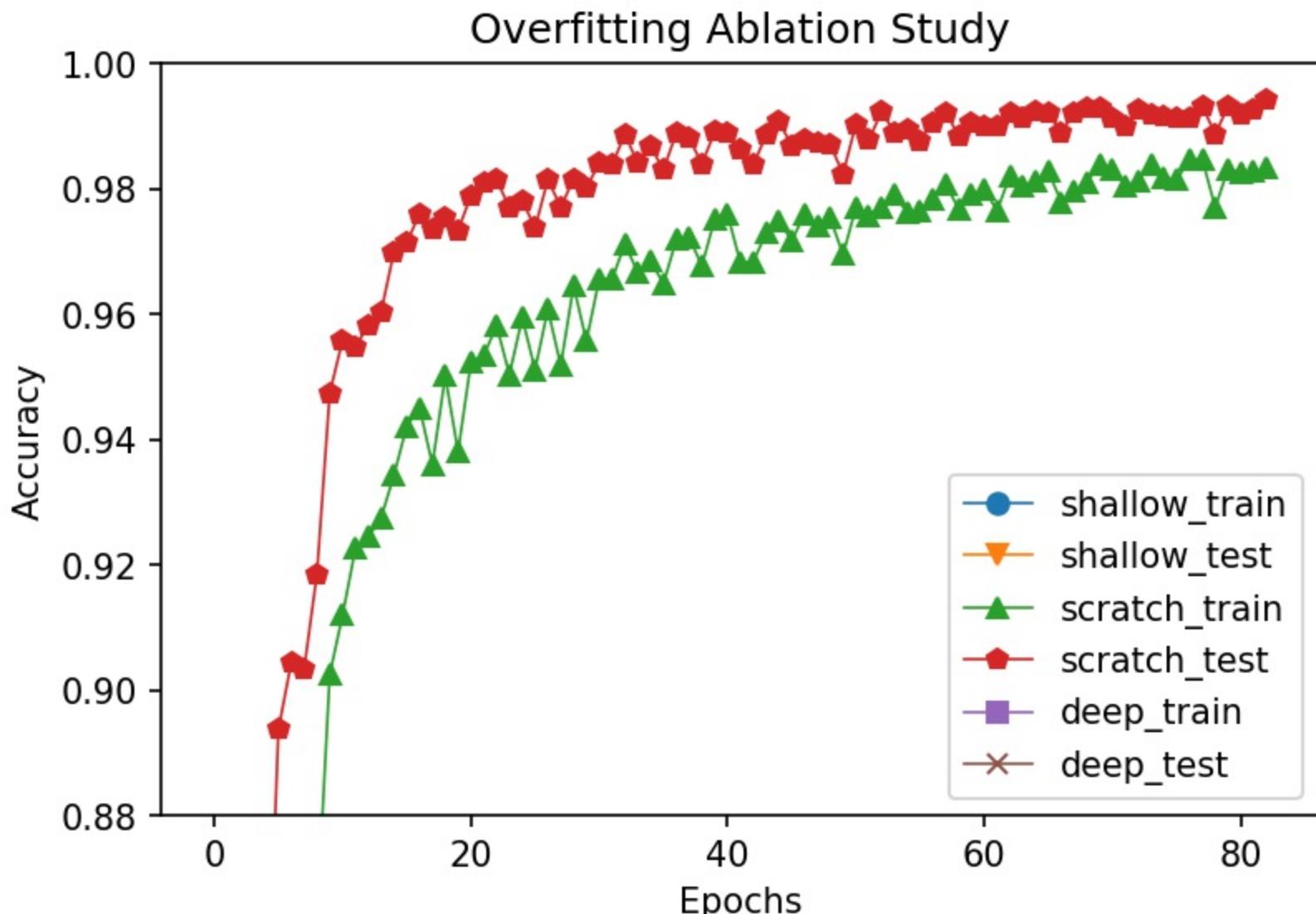
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	NaN	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 82 takes 315 seconds.

[84, 30] loss: 0.049

[84, 60] loss: 0.057

```
[84,    90] loss: 0.051
[84,   120] loss: 0.057
[84,   150] loss: 0.069
[84,   180] loss: 0.054
[84,   210] loss: 0.094
[84,   240] loss: 0.040
[84,   270] loss: 0.072
[84,   300] loss: 0.071
[84,   330] loss: 0.080
[84,   360] loss: 0.098
[84,   390] loss: 0.066
[84,   420] loss: 0.082
[84,   450] loss: 0.121
[84,   480] loss: 0.122
[84,   510] loss: 0.063
[84,   540] loss: 0.069
[84,   570] loss: 0.091
[84,   600] loss: 0.063
[84,   630] loss: 0.057
[84,   660] loss: 0.084
[84,   690] loss: 0.066
[84,   720] loss: 0.049
[84,   750] loss: 0.056
[84,   780] loss: 0.059
[84,   810] loss: 0.069
[84,   840] loss: 0.050
[84,   870] loss: 0.066
[84,   900] loss: 0.059
[84,   930] loss: 0.087
[84,   960] loss: 0.044
[84,   990] loss: 0.065
[84,  1020] loss: 0.063
[84,  1050] loss: 0.088
[84,  1080] loss: 0.063
[84,  1110] loss: 0.094
[84,  1140] loss: 0.045
[84,  1170] loss: 0.127
[84,  1200] loss: 0.101
[84,  1230] loss: 0.067
```

```
[84, 1260] loss: 0.062
[84, 1290] loss: 0.038
[84, 1320] loss: 0.042
[84, 1350] loss: 0.075
[84, 1380] loss: 0.092
[84, 1410] loss: 0.051
[84, 1440] loss: 0.075
[84, 1470] loss: 0.078
[84, 1500] loss: 0.051
[84, 1530] loss: 0.069
[84, 1560] loss: 0.043
[84, 1590] loss: 0.067
[84, 1620] loss: 0.060
[84, 1650] loss: 0.062
[84, 1680] loss: 0.070
[84, 1710] loss: 0.074
[84, 1740] loss: 0.084
[84, 1770] loss: 0.095
[84, 1800] loss: 0.069
[84, 1830] loss: 0.064
[84, 1860] loss: 0.037
[84, 1890] loss: 0.054
[84, 1920] loss: 0.063
[84, 1950] loss: 0.056
[84, 1980] loss: 0.107
[84, 2010] loss: 0.068
[84, 2040] loss: 0.071
[84, 2070] loss: 0.077
[84, 2100] loss: 0.082
[84, 2130] loss: 0.087
[84, 2160] loss: 0.097
[84, 2190] loss: 0.052
eval intermediate_models/epoch83_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.985074
epoch 83 accuracy on train set is: 0.9850736095965104
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	NaN	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991730

epoch 83 accuracy on test set is: 0.9917302798982188

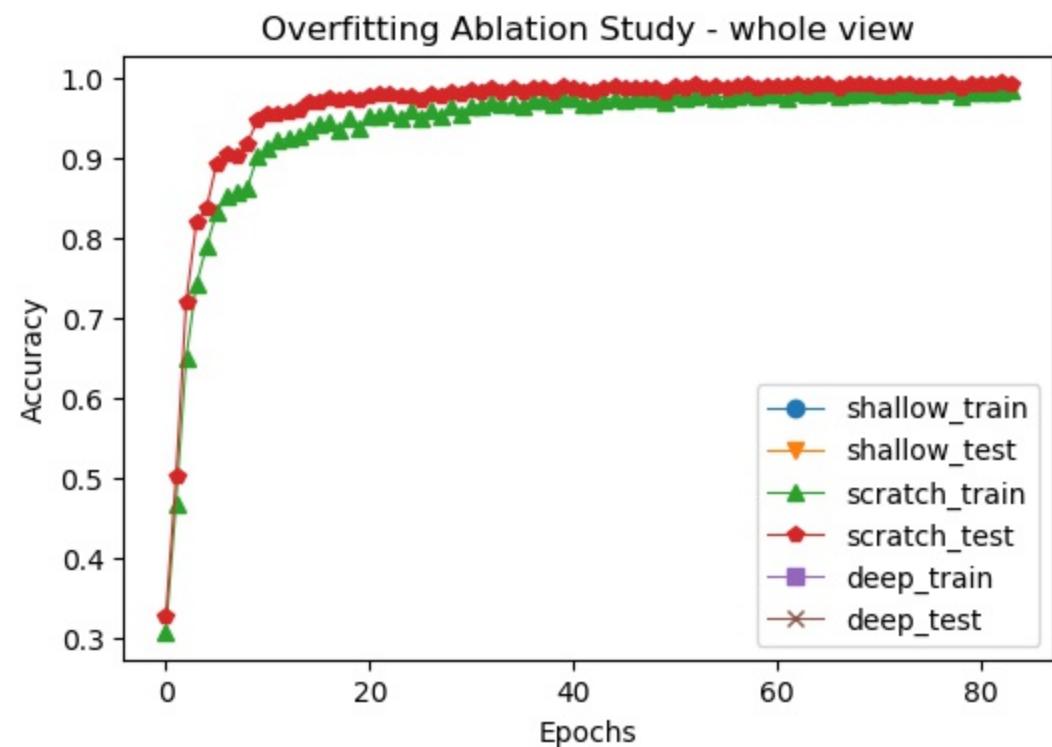
evaluating on test set takes 16 seconds.

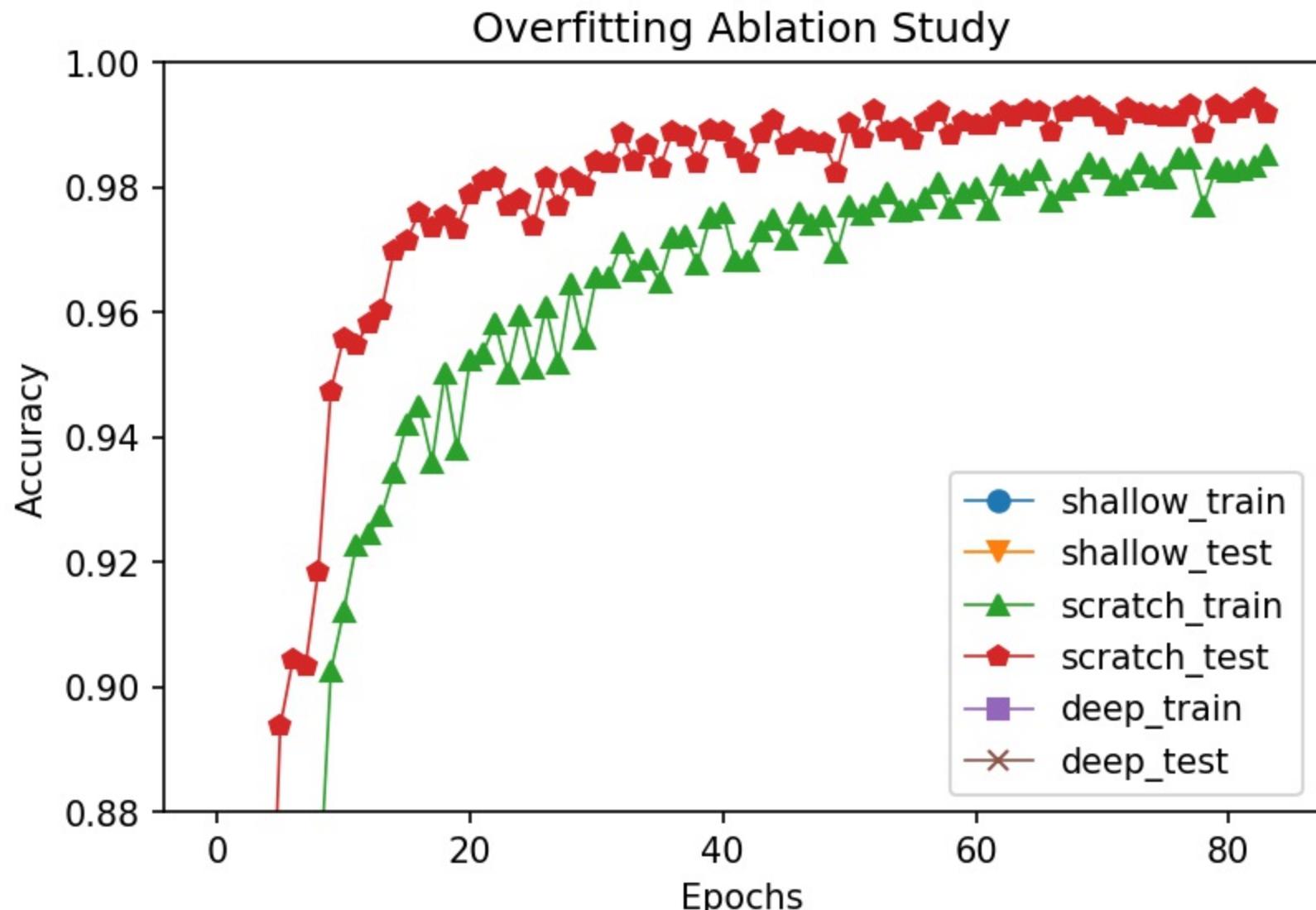
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	NaN	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 83 takes 315 seconds.

[85, 30] loss: 0.057

[85, 60] loss: 0.089

```
[85,    90] loss: 0.075
[85,   120] loss: 0.059
[85,   150] loss: 0.081
[85,   180] loss: 0.055
[85,   210] loss: 0.060
[85,   240] loss: 0.069
[85,   270] loss: 0.065
[85,   300] loss: 0.096
[85,   330] loss: 0.053
[85,   360] loss: 0.092
[85,   390] loss: 0.083
[85,   420] loss: 0.075
[85,   450] loss: 0.062
[85,   480] loss: 0.089
[85,   510] loss: 0.079
[85,   540] loss: 0.051
[85,   570] loss: 0.069
[85,   600] loss: 0.051
[85,   630] loss: 0.064
[85,   660] loss: 0.049
[85,   690] loss: 0.077
[85,   720] loss: 0.073
[85,   750] loss: 0.078
[85,   780] loss: 0.083
[85,   810] loss: 0.101
[85,   840] loss: 0.064
[85,   870] loss: 0.091
[85,   900] loss: 0.074
[85,   930] loss: 0.068
[85,   960] loss: 0.117
[85,   990] loss: 0.047
[85,  1020] loss: 0.102
[85,  1050] loss: 0.067
[85,  1080] loss: 0.050
[85,  1110] loss: 0.069
[85,  1140] loss: 0.054
[85,  1170] loss: 0.074
[85,  1200] loss: 0.076
[85,  1230] loss: 0.072
```

```
[85, 1260] loss: 0.048
[85, 1290] loss: 0.064
[85, 1320] loss: 0.051
[85, 1350] loss: 0.078
[85, 1380] loss: 0.061
[85, 1410] loss: 0.048
[85, 1440] loss: 0.088
[85, 1470] loss: 0.046
[85, 1500] loss: 0.045
[85, 1530] loss: 0.039
[85, 1560] loss: 0.076
[85, 1590] loss: 0.058
[85, 1620] loss: 0.051
[85, 1650] loss: 0.103
[85, 1680] loss: 0.067
[85, 1710] loss: 0.071
[85, 1740] loss: 0.057
[85, 1770] loss: 0.069
[85, 1800] loss: 0.070
[85, 1830] loss: 0.078
[85, 1860] loss: 0.037
[85, 1890] loss: 0.060
[85, 1920] loss: 0.048
[85, 1950] loss: 0.039
[85, 1980] loss: 0.029
[85, 2010] loss: 0.063
[85, 2040] loss: 0.066
[85, 2070] loss: 0.056
[85, 2100] loss: 0.099
[85, 2130] loss: 0.065
[85, 2160] loss: 0.069
[85, 2190] loss: 0.079
evf intermediate_models/epoch84_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.983415
epoch 84 accuracy on train set is: 0.9834151217739004
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	NaN	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992185

epoch 84 accuracy on test set is: 0.9921846601235914

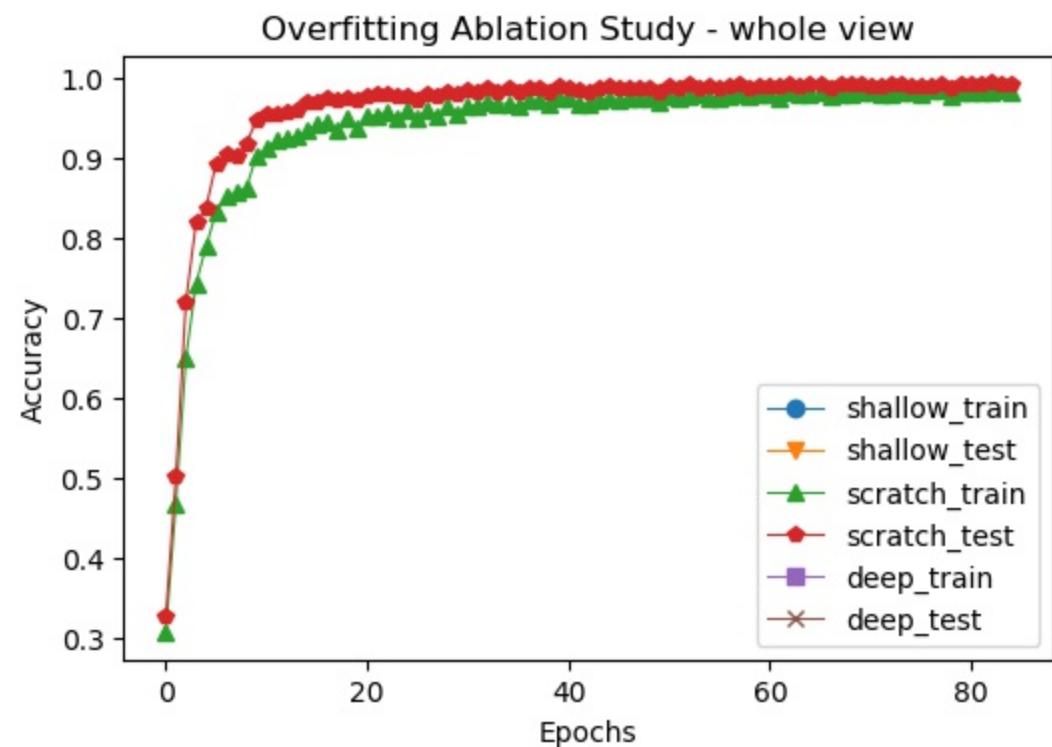
evaluating on test set takes 16 seconds.

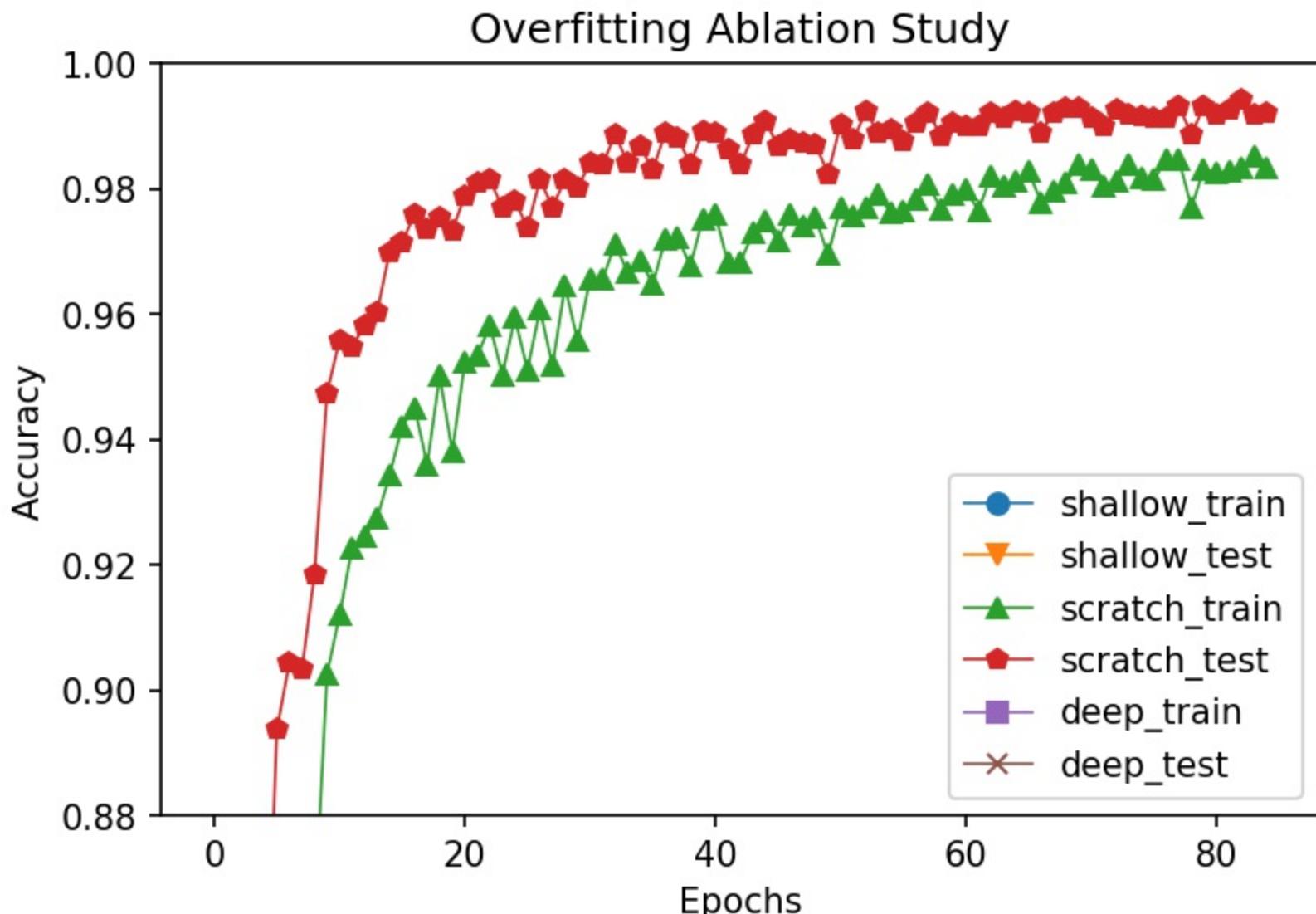
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	NaN	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 84 takes 315 seconds.

[86, 30] loss: 0.083

[86, 60] loss: 0.084

```
[86,    90] loss: 0.059
[86,   120] loss: 0.114
[86,   150] loss: 0.102
[86,   180] loss: 0.044
[86,   210] loss: 0.052
[86,   240] loss: 0.056
[86,   270] loss: 0.036
[86,   300] loss: 0.072
[86,   330] loss: 0.064
[86,   360] loss: 0.079
[86,   390] loss: 0.096
[86,   420] loss: 0.052
[86,   450] loss: 0.070
[86,   480] loss: 0.046
[86,   510] loss: 0.093
[86,   540] loss: 0.062
[86,   570] loss: 0.060
[86,   600] loss: 0.047
[86,   630] loss: 0.048
[86,   660] loss: 0.065
[86,   690] loss: 0.070
[86,   720] loss: 0.063
[86,   750] loss: 0.069
[86,   780] loss: 0.048
[86,   810] loss: 0.053
[86,   840] loss: 0.065
[86,   870] loss: 0.054
[86,   900] loss: 0.075
[86,   930] loss: 0.080
[86,   960] loss: 0.074
[86,   990] loss: 0.061
[86,  1020] loss: 0.095
[86,  1050] loss: 0.084
[86,  1080] loss: 0.074
[86,  1110] loss: 0.066
[86,  1140] loss: 0.074
[86,  1170] loss: 0.066
[86,  1200] loss: 0.039
[86,  1230] loss: 0.065
```

```
[86, 1260] loss: 0.038
[86, 1290] loss: 0.039
[86, 1320] loss: 0.053
[86, 1350] loss: 0.088
[86, 1380] loss: 0.055
[86, 1410] loss: 0.073
[86, 1440] loss: 0.069
[86, 1470] loss: 0.066
[86, 1500] loss: 0.048
[86, 1530] loss: 0.037
[86, 1560] loss: 0.052
[86, 1590] loss: 0.052
[86, 1620] loss: 0.071
[86, 1650] loss: 0.069
[86, 1680] loss: 0.064
[86, 1710] loss: 0.068
[86, 1740] loss: 0.064
[86, 1770] loss: 0.068
[86, 1800] loss: 0.103
[86, 1830] loss: 0.078
[86, 1860] loss: 0.081
[86, 1890] loss: 0.084
[86, 1920] loss: 0.095
[86, 1950] loss: 0.067
[86, 1980] loss: 0.095
[86, 2010] loss: 0.070
[86, 2040] loss: 0.070
[86, 2070] loss: 0.076
[86, 2100] loss: 0.065
[86, 2130] loss: 0.057
[86, 2160] loss: 0.039
[86, 2190] loss: 0.081
eval intermediate_models/epoch85_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.983279
epoch 85 accuracy on train set is: 0.9832788077062886
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	NaN	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990822

epoch 85 accuracy on test set is: 0.9908215194474737

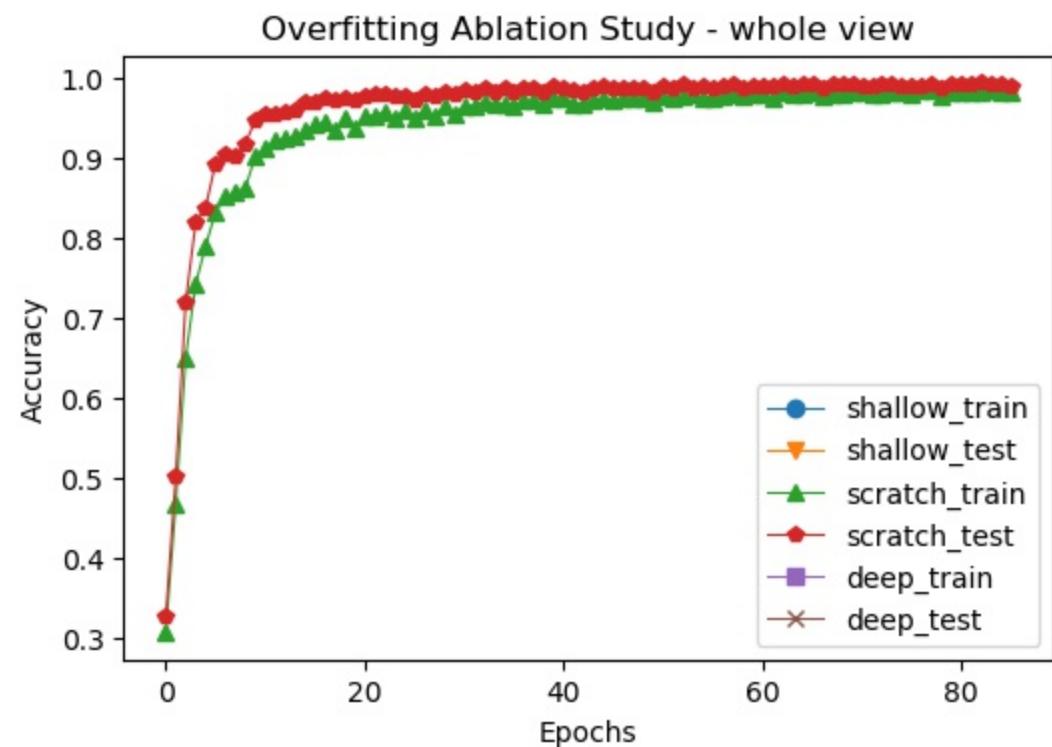
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

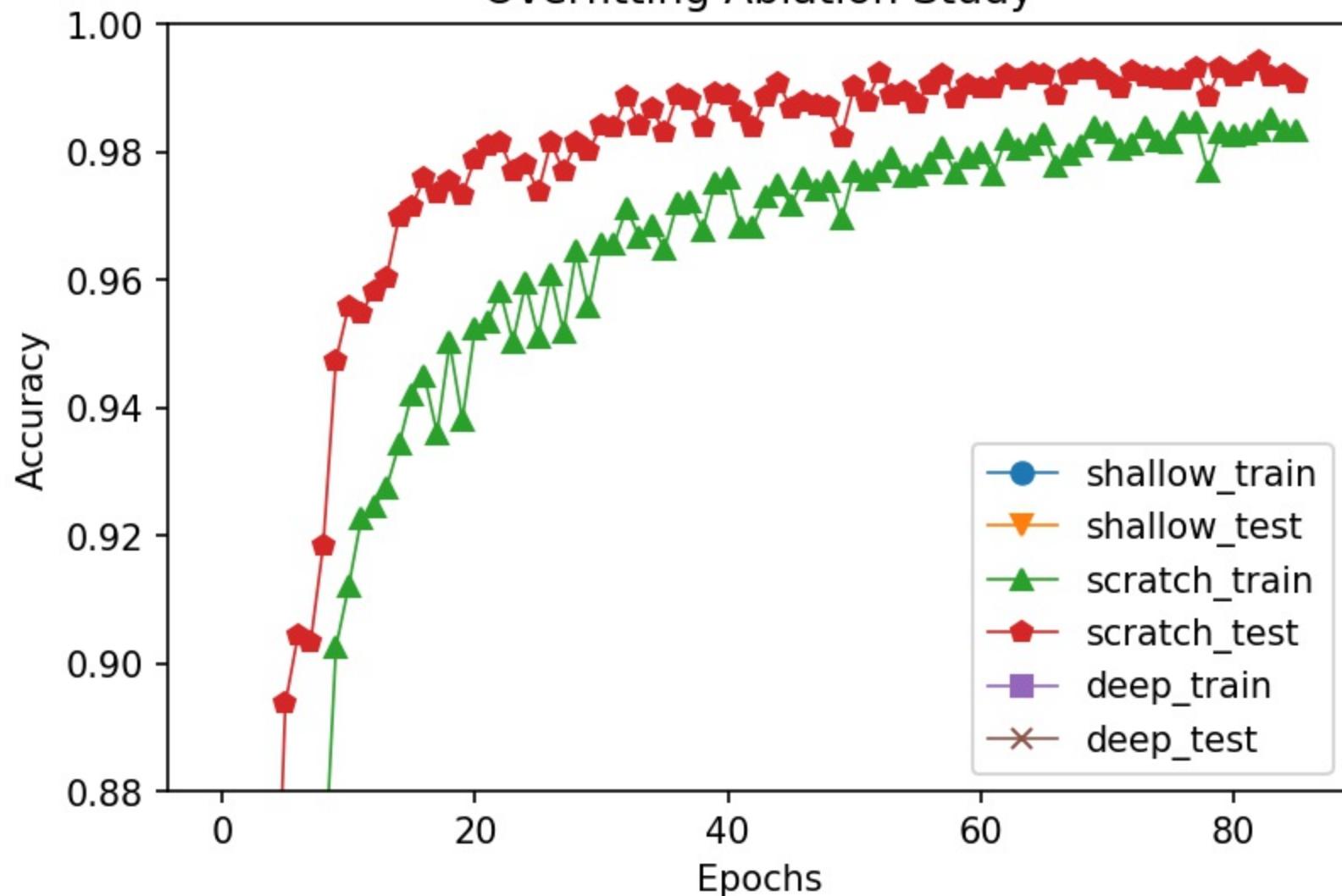
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	NaN	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 85 takes 315 seconds.

[87, 30] loss: 0.041

[87, 60] loss: 0.039

```
[87,    90] loss: 0.100
[87,   120] loss: 0.068
[87,   150] loss: 0.089
[87,   180] loss: 0.071
[87,   210] loss: 0.080
[87,   240] loss: 0.068
[87,   270] loss: 0.065
[87,   300] loss: 0.062
[87,   330] loss: 0.054
[87,   360] loss: 0.042
[87,   390] loss: 0.043
[87,   420] loss: 0.052
[87,   450] loss: 0.037
[87,   480] loss: 0.041
[87,   510] loss: 0.031
[87,   540] loss: 0.046
[87,   570] loss: 0.104
[87,   600] loss: 0.072
[87,   630] loss: 0.049
[87,   660] loss: 0.054
[87,   690] loss: 0.098
[87,   720] loss: 0.068
[87,   750] loss: 0.082
[87,   780] loss: 0.065
[87,   810] loss: 0.044
[87,   840] loss: 0.045
[87,   870] loss: 0.047
[87,   900] loss: 0.066
[87,   930] loss: 0.054
[87,   960] loss: 0.067
[87,   990] loss: 0.098
[87,  1020] loss: 0.048
[87,  1050] loss: 0.073
[87,  1080] loss: 0.061
[87,  1110] loss: 0.074
[87,  1140] loss: 0.044
[87,  1170] loss: 0.045
[87,  1200] loss: 0.079
[87,  1230] loss: 0.053
```

```
[87, 1260] loss: 0.069
[87, 1290] loss: 0.041
[87, 1320] loss: 0.113
[87, 1350] loss: 0.062
[87, 1380] loss: 0.042
[87, 1410] loss: 0.073
[87, 1440] loss: 0.063
[87, 1470] loss: 0.061
[87, 1500] loss: 0.085
[87, 1530] loss: 0.060
[87, 1560] loss: 0.048
[87, 1590] loss: 0.051
[87, 1620] loss: 0.055
[87, 1650] loss: 0.057
[87, 1680] loss: 0.043
[87, 1710] loss: 0.065
[87, 1740] loss: 0.048
[87, 1770] loss: 0.062
[87, 1800] loss: 0.063
[87, 1830] loss: 0.049
[87, 1860] loss: 0.089
[87, 1890] loss: 0.082
[87, 1920] loss: 0.078
[87, 1950] loss: 0.047
[87, 1980] loss: 0.055
[87, 2010] loss: 0.049
[87, 2040] loss: 0.060
[87, 2070] loss: 0.063
[87, 2100] loss: 0.067
[87, 2130] loss: 0.114
[87, 2160] loss: 0.080
[87, 2190] loss: 0.065
eval intermediate_models/epoch86_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.984301
epoch 86 accuracy on train set is: 0.9843011632133769
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	NaN	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991458

epoch 86 accuracy on test set is: 0.9914576517629953

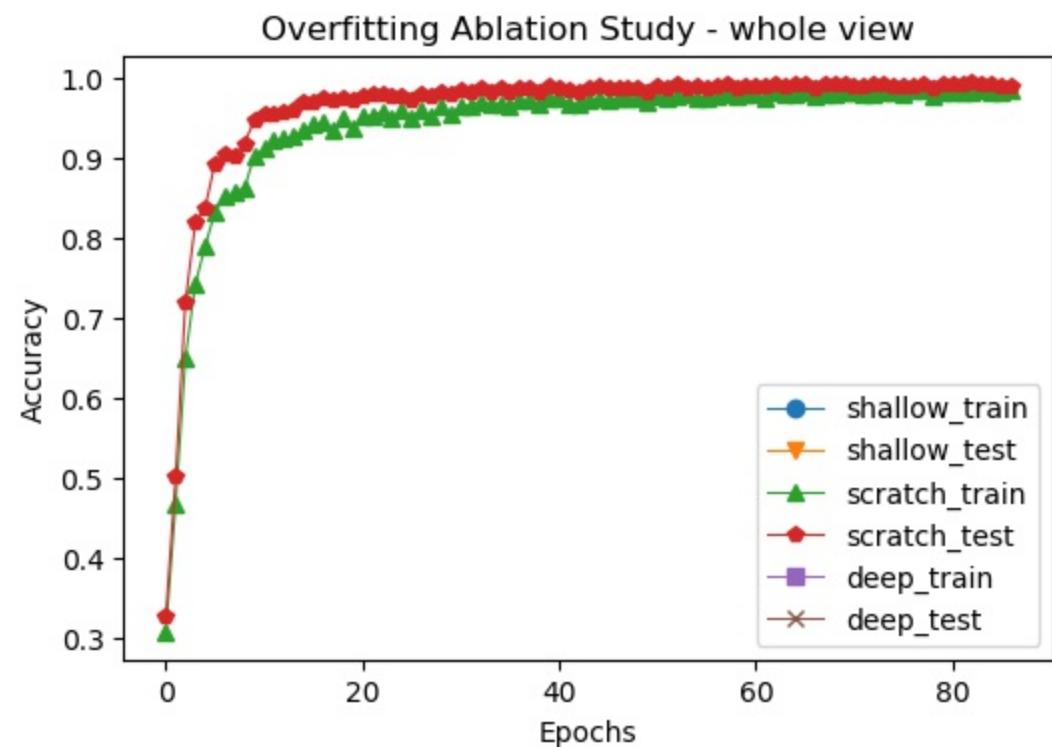
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

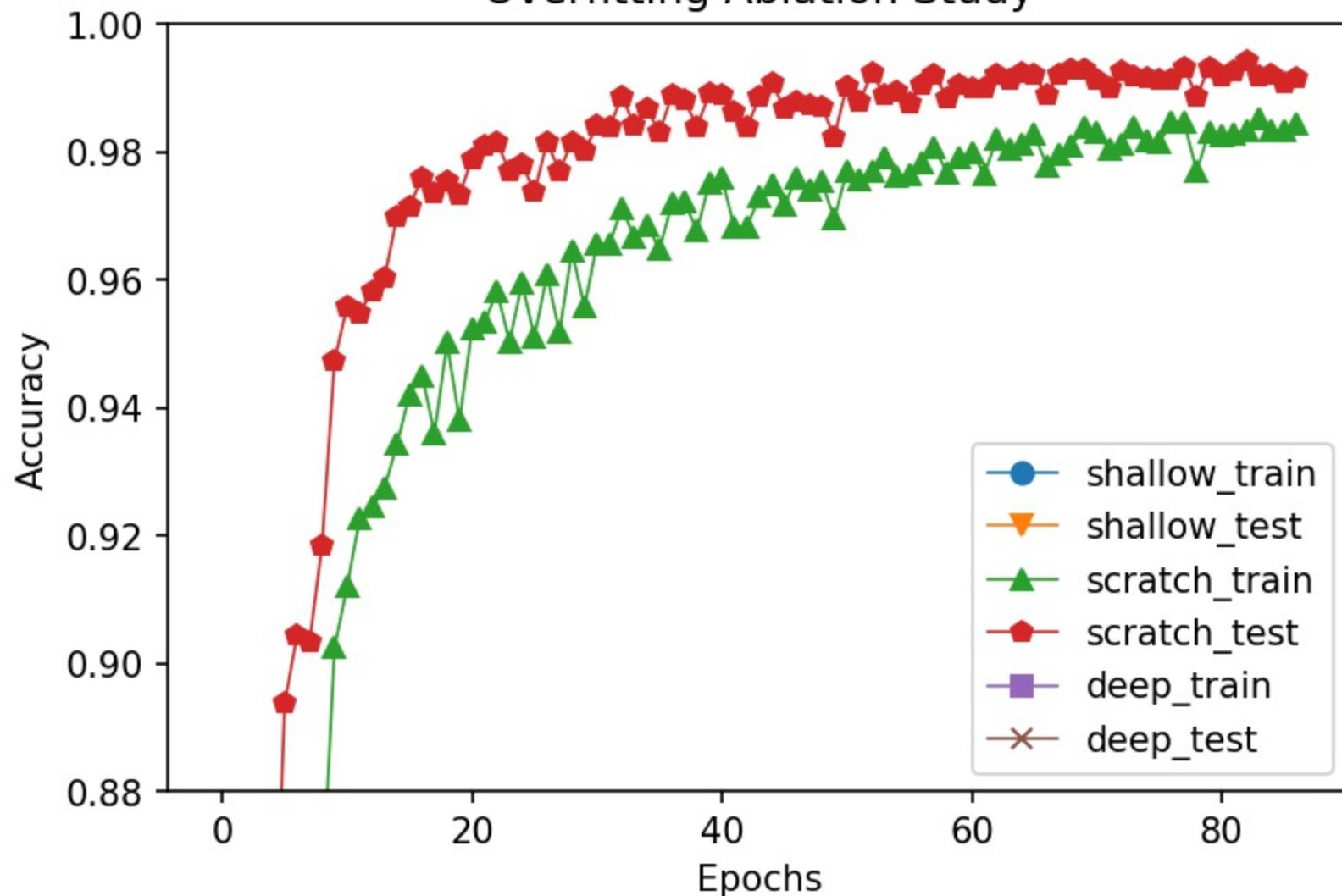
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	NaN	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 86 takes 315 seconds.

[88, 30] loss: 0.078

[88, 60] loss: 0.046

```
[88,    90] loss: 0.064
[88,   120] loss: 0.078
[88,   150] loss: 0.093
[88,   180] loss: 0.052
[88,   210] loss: 0.052
[88,   240] loss: 0.079
[88,   270] loss: 0.065
[88,   300] loss: 0.020
[88,   330] loss: 0.058
[88,   360] loss: 0.051
[88,   390] loss: 0.058
[88,   420] loss: 0.060
[88,   450] loss: 0.063
[88,   480] loss: 0.087
[88,   510] loss: 0.052
[88,   540] loss: 0.052
[88,   570] loss: 0.046
[88,   600] loss: 0.077
[88,   630] loss: 0.068
[88,   660] loss: 0.094
[88,   690] loss: 0.068
[88,   720] loss: 0.048
[88,   750] loss: 0.071
[88,   780] loss: 0.069
[88,   810] loss: 0.054
[88,   840] loss: 0.057
[88,   870] loss: 0.047
[88,   900] loss: 0.035
[88,   930] loss: 0.052
[88,   960] loss: 0.075
[88,   990] loss: 0.044
[88,  1020] loss: 0.061
[88,  1050] loss: 0.068
[88,  1080] loss: 0.061
[88,  1110] loss: 0.081
[88,  1140] loss: 0.104
[88,  1170] loss: 0.131
[88,  1200] loss: 0.087
[88,  1230] loss: 0.107
```

```
[88, 1260] loss: 0.076
[88, 1290] loss: 0.064
[88, 1320] loss: 0.043
[88, 1350] loss: 0.062
[88, 1380] loss: 0.079
[88, 1410] loss: 0.073
[88, 1440] loss: 0.044
[88, 1470] loss: 0.101
[88, 1500] loss: 0.068
[88, 1530] loss: 0.055
[88, 1560] loss: 0.047
[88, 1590] loss: 0.060
[88, 1620] loss: 0.059
[88, 1650] loss: 0.044
[88, 1680] loss: 0.067
[88, 1710] loss: 0.051
[88, 1740] loss: 0.058
[88, 1770] loss: 0.047
[88, 1800] loss: 0.054
[88, 1830] loss: 0.078
[88, 1860] loss: 0.088
[88, 1890] loss: 0.060
[88, 1920] loss: 0.052
[88, 1950] loss: 0.048
[88, 1980] loss: 0.057
[88, 2010] loss: 0.059
[88, 2040] loss: 0.066
[88, 2070] loss: 0.097
[88, 2100] loss: 0.053
[88, 2130] loss: 0.084
[88, 2160] loss: 0.045
[88, 2190] loss: 0.069
evf intermediate_models/epoch87_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.981779
epoch 87 accuracy on train set is: 0.981779352962559
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	NaN	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.991730

epoch 87 accuracy on test set is: 0.9917302798982188

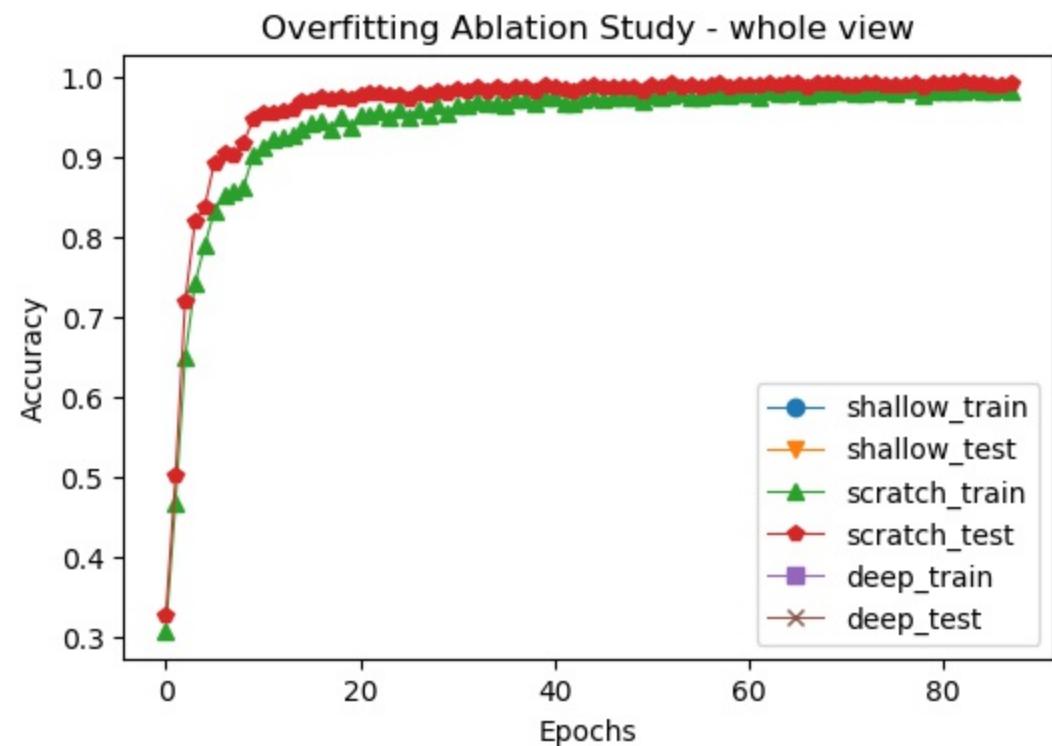
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

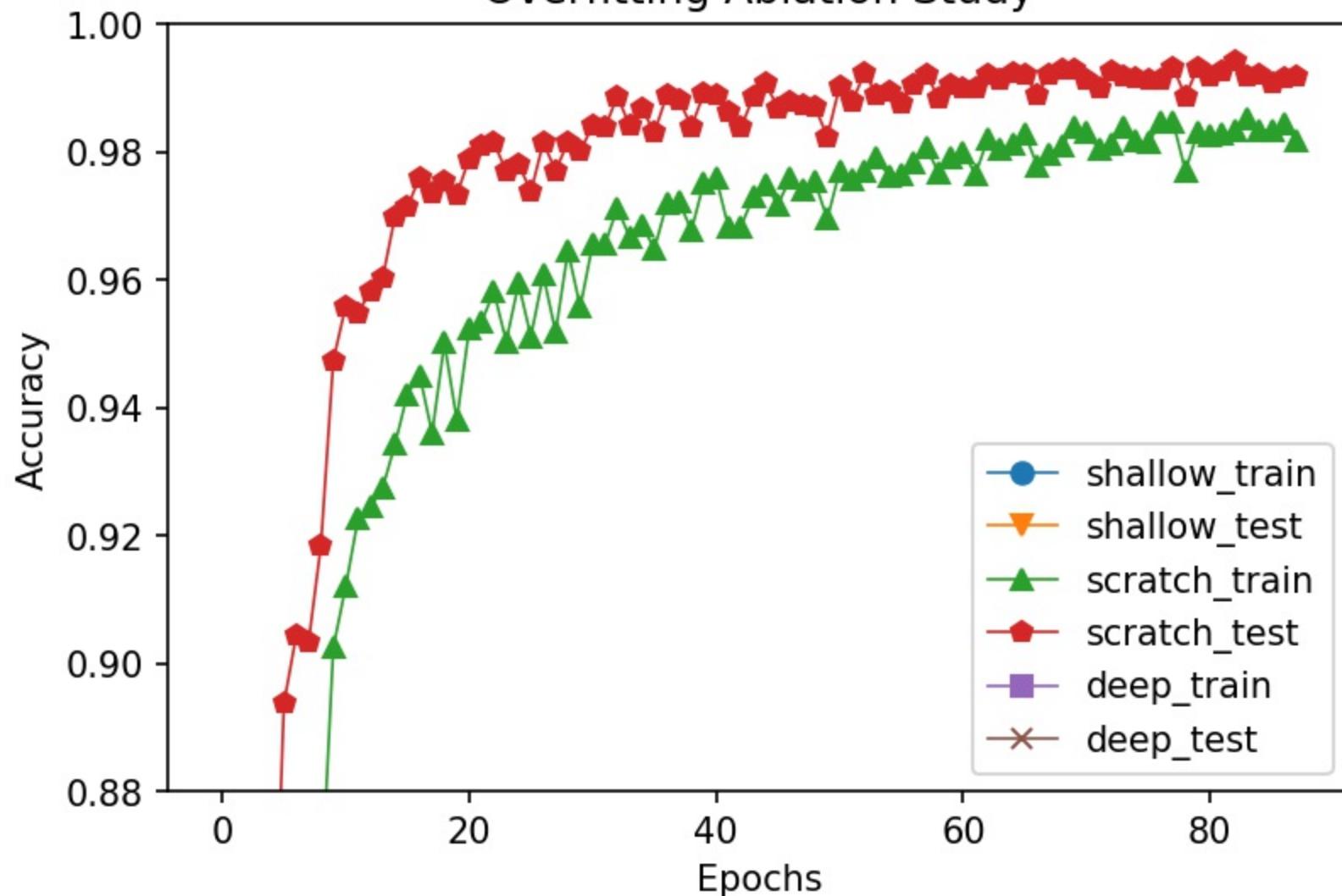
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	NaN	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 87 takes 315 seconds.

[89, 30] loss: 0.072

[89, 60] loss: 0.062

```
[89,    90] loss: 0.059
[89,   120] loss: 0.061
[89,   150] loss: 0.059
[89,   180] loss: 0.041
[89,   210] loss: 0.062
[89,   240] loss: 0.057
[89,   270] loss: 0.088
[89,   300] loss: 0.048
[89,   330] loss: 0.046
[89,   360] loss: 0.055
[89,   390] loss: 0.056
[89,   420] loss: 0.092
[89,   450] loss: 0.092
[89,   480] loss: 0.072
[89,   510] loss: 0.093
[89,   540] loss: 0.088
[89,   570] loss: 0.068
[89,   600] loss: 0.056
[89,   630] loss: 0.085
[89,   660] loss: 0.078
[89,   690] loss: 0.050
[89,   720] loss: 0.051
[89,   750] loss: 0.065
[89,   780] loss: 0.045
[89,   810] loss: 0.077
[89,   840] loss: 0.074
[89,   870] loss: 0.060
[89,   900] loss: 0.076
[89,   930] loss: 0.074
[89,   960] loss: 0.089
[89,   990] loss: 0.079
[89,  1020] loss: 0.096
[89,  1050] loss: 0.060
[89,  1080] loss: 0.048
[89,  1110] loss: 0.079
[89,  1140] loss: 0.054
[89,  1170] loss: 0.078
[89,  1200] loss: 0.070
[89,  1230] loss: 0.076
```

```
[89, 1260] loss: 0.071
[89, 1290] loss: 0.096
[89, 1320] loss: 0.062
[89, 1350] loss: 0.046
[89, 1380] loss: 0.029
[89, 1410] loss: 0.034
[89, 1440] loss: 0.075
[89, 1470] loss: 0.058
[89, 1500] loss: 0.047
[89, 1530] loss: 0.052
[89, 1560] loss: 0.069
[89, 1590] loss: 0.058
[89, 1620] loss: 0.082
[89, 1650] loss: 0.046
[89, 1680] loss: 0.059
[89, 1710] loss: 0.065
[89, 1740] loss: 0.095
[89, 1770] loss: 0.081
[89, 1800] loss: 0.095
[89, 1830] loss: 0.028
[89, 1860] loss: 0.056
[89, 1890] loss: 0.058
[89, 1920] loss: 0.077
[89, 1950] loss: 0.112
[89, 1980] loss: 0.070
[89, 2010] loss: 0.073
[89, 2040] loss: 0.097
[89, 2070] loss: 0.081
[89, 2100] loss: 0.079
[89, 2130] loss: 0.053
[89, 2160] loss: 0.087
[89, 2190] loss: 0.073
eval intermediate_models/epoch88_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.983688
epoch 88 accuracy on train set is: 0.983687749909124
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	NaN	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993275

epoch 88 accuracy on test set is: 0.9932751726644856

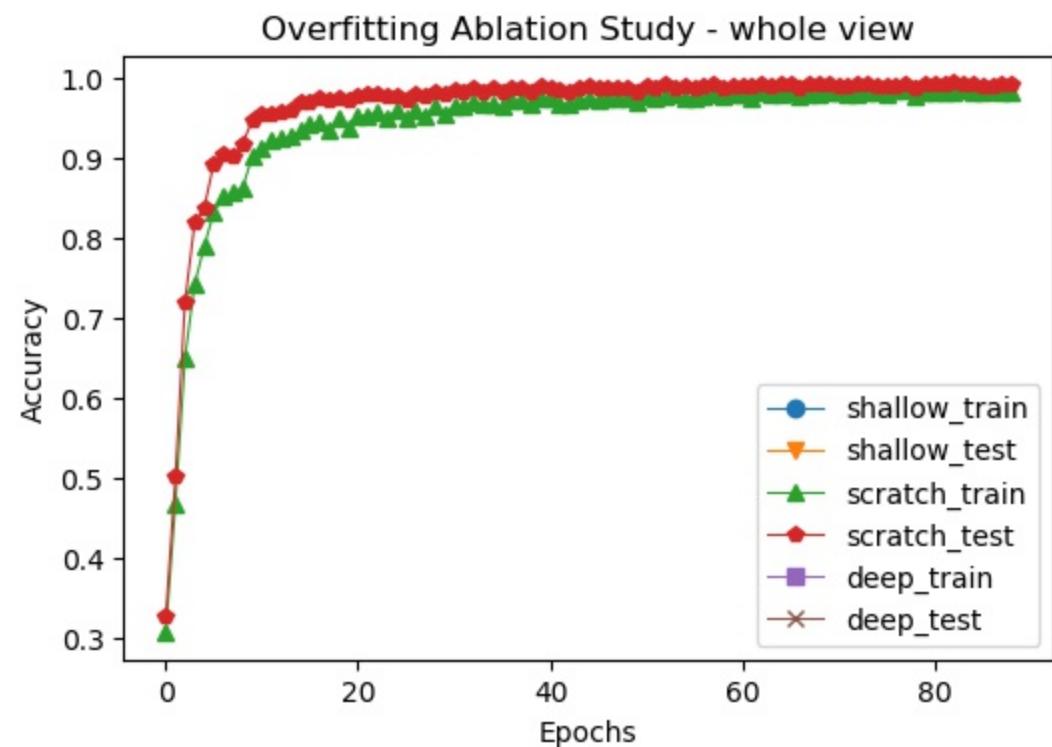
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

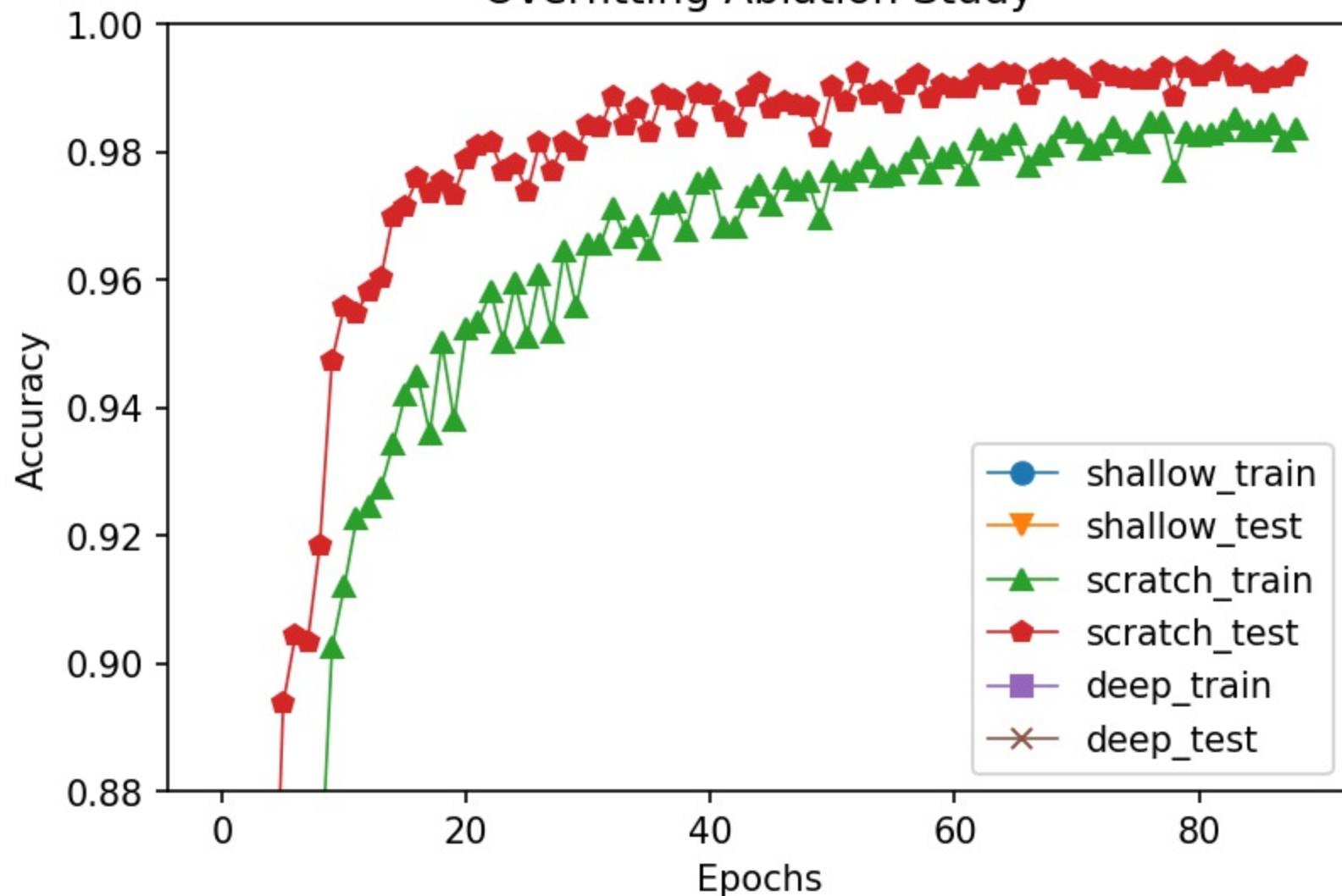
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	NaN	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 88 takes 315 seconds.

[90, 30] loss: 0.044

[90, 60] loss: 0.065

```
[90,    90] loss: 0.048
[90,   120] loss: 0.092
[90,   150] loss: 0.103
[90,   180] loss: 0.086
[90,   210] loss: 0.067
[90,   240] loss: 0.059
[90,   270] loss: 0.049
[90,   300] loss: 0.065
[90,   330] loss: 0.082
[90,   360] loss: 0.072
[90,   390] loss: 0.054
[90,   420] loss: 0.062
[90,   450] loss: 0.064
[90,   480] loss: 0.079
[90,   510] loss: 0.102
[90,   540] loss: 0.097
[90,   570] loss: 0.051
[90,   600] loss: 0.062
[90,   630] loss: 0.067
[90,   660] loss: 0.061
[90,   690] loss: 0.050
[90,   720] loss: 0.055
[90,   750] loss: 0.057
[90,   780] loss: 0.053
[90,   810] loss: 0.057
[90,   840] loss: 0.081
[90,   870] loss: 0.070
[90,   900] loss: 0.048
[90,   930] loss: 0.072
[90,   960] loss: 0.062
[90,   990] loss: 0.067
[90,  1020] loss: 0.043
[90,  1050] loss: 0.088
[90,  1080] loss: 0.078
[90,  1110] loss: 0.061
[90,  1140] loss: 0.054
[90,  1170] loss: 0.040
[90,  1200] loss: 0.064
[90,  1230] loss: 0.056
```

```
[90, 1260] loss: 0.042
[90, 1290] loss: 0.070
[90, 1320] loss: 0.084
[90, 1350] loss: 0.051
[90, 1380] loss: 0.078
[90, 1410] loss: 0.074
[90, 1440] loss: 0.040
[90, 1470] loss: 0.059
[90, 1500] loss: 0.064
[90, 1530] loss: 0.061
[90, 1560] loss: 0.083
[90, 1590] loss: 0.058
[90, 1620] loss: 0.063
[90, 1650] loss: 0.087
[90, 1680] loss: 0.080
[90, 1710] loss: 0.045
[90, 1740] loss: 0.061
[90, 1770] loss: 0.068
[90, 1800] loss: 0.099
[90, 1830] loss: 0.078
[90, 1860] loss: 0.079
[90, 1890] loss: 0.048
[90, 1920] loss: 0.051
[90, 1950] loss: 0.066
[90, 1980] loss: 0.044
[90, 2010] loss: 0.044
[90, 2040] loss: 0.048
[90, 2070] loss: 0.067
[90, 2100] loss: 0.062
[90, 2130] loss: 0.039
[90, 2160] loss: 0.078
[90, 2190] loss: 0.072
eval intermediate_models/epoch89_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.984824
epoch 89 accuracy on train set is: 0.9848237004725554
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	NaN	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993275

epoch 89 accuracy on test set is: 0.9932751726644856

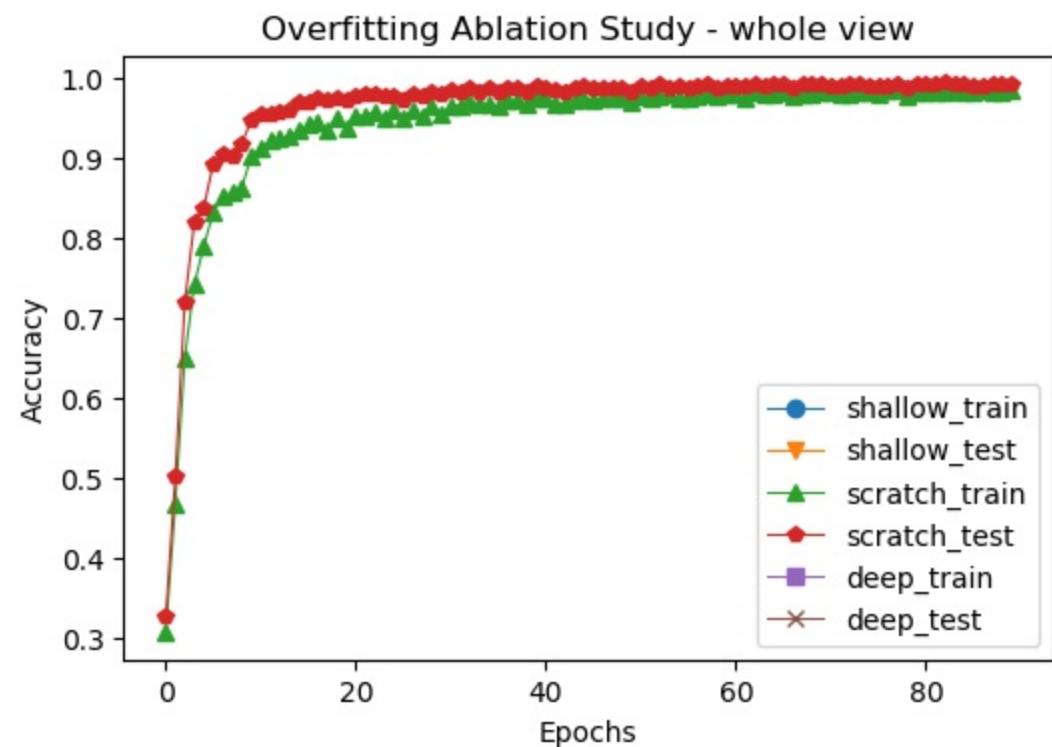
evaluating on test set takes 16 seconds.

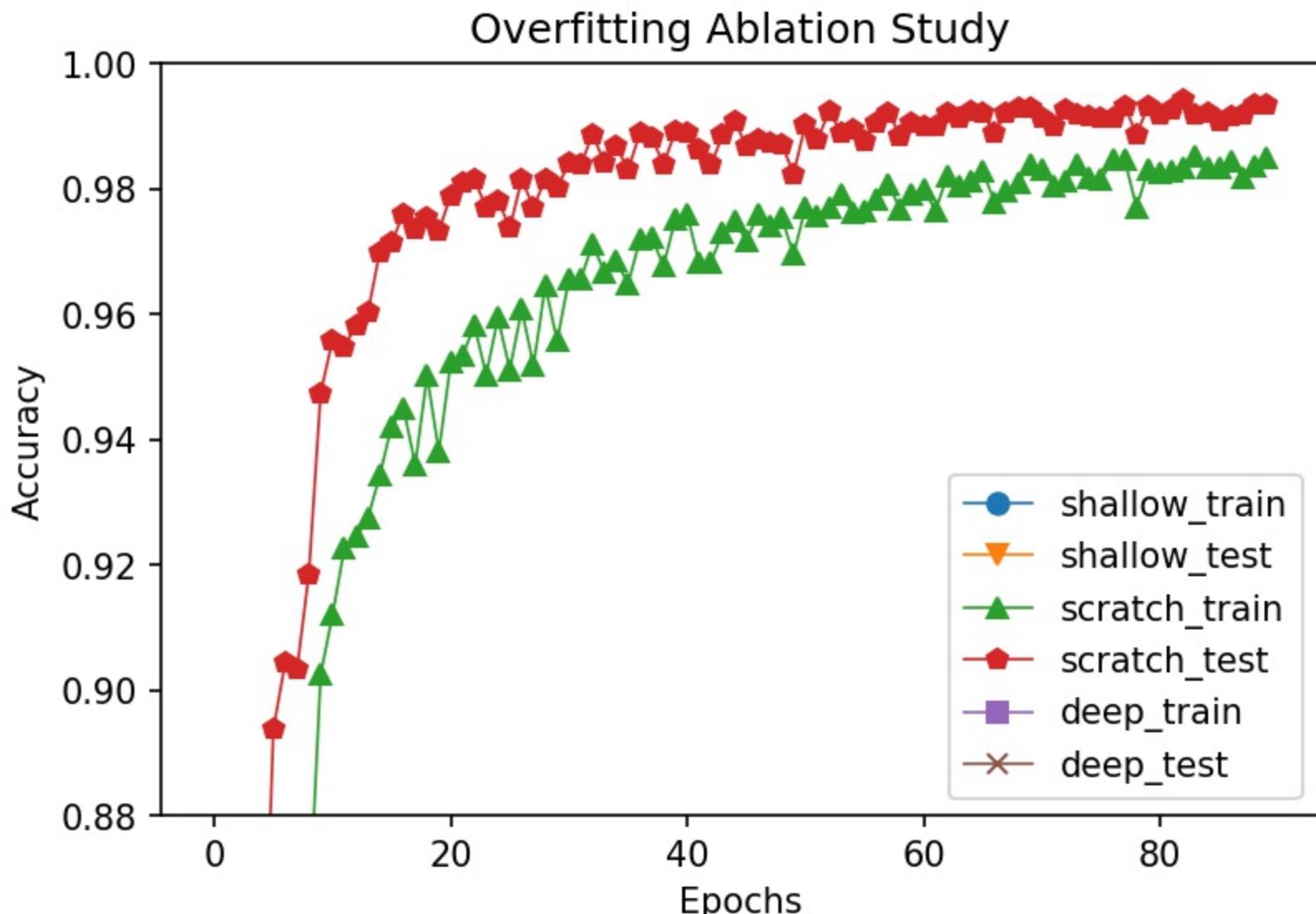
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	NaN	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 89 takes 319 seconds.

[91, 30] loss: 0.066

[91, 60] loss: 0.052

```
[91,    90] loss: 0.055
[91,   120] loss: 0.055
[91,   150] loss: 0.105
[91,   180] loss: 0.069
[91,   210] loss: 0.080
[91,   240] loss: 0.088
[91,   270] loss: 0.065
[91,   300] loss: 0.050
[91,   330] loss: 0.033
[91,   360] loss: 0.073
[91,   390] loss: 0.065
[91,   420] loss: 0.067
[91,   450] loss: 0.070
[91,   480] loss: 0.086
[91,   510] loss: 0.074
[91,   540] loss: 0.078
[91,   570] loss: 0.046
[91,   600] loss: 0.073
[91,   630] loss: 0.057
[91,   660] loss: 0.041
[91,   690] loss: 0.043
[91,   720] loss: 0.063
[91,   750] loss: 0.070
[91,   780] loss: 0.079
[91,   810] loss: 0.063
[91,   840] loss: 0.043
[91,   870] loss: 0.075
[91,   900] loss: 0.055
[91,   930] loss: 0.046
[91,   960] loss: 0.068
[91,   990] loss: 0.078
[91,  1020] loss: 0.057
[91,  1050] loss: 0.046
[91,  1080] loss: 0.096
[91,  1110] loss: 0.072
[91,  1140] loss: 0.061
[91,  1170] loss: 0.048
[91,  1200] loss: 0.074
[91,  1230] loss: 0.079
```

```
[91, 1260] loss: 0.098
[91, 1290] loss: 0.076
[91, 1320] loss: 0.110
[91, 1350] loss: 0.063
[91, 1380] loss: 0.048
[91, 1410] loss: 0.065
[91, 1440] loss: 0.087
[91, 1470] loss: 0.089
[91, 1500] loss: 0.051
[91, 1530] loss: 0.049
[91, 1560] loss: 0.084
[91, 1590] loss: 0.085
[91, 1620] loss: 0.040
[91, 1650] loss: 0.066
[91, 1680] loss: 0.047
[91, 1710] loss: 0.059
[91, 1740] loss: 0.040
[91, 1770] loss: 0.044
[91, 1800] loss: 0.083
[91, 1830] loss: 0.060
[91, 1860] loss: 0.071
[91, 1890] loss: 0.074
[91, 1920] loss: 0.054
[91, 1950] loss: 0.051
[91, 1980] loss: 0.092
[91, 2010] loss: 0.061
[91, 2040] loss: 0.058
[91, 2070] loss: 0.073
[91, 2100] loss: 0.061
[91, 2130] loss: 0.061
[91, 2160] loss: 0.052
[91, 2190] loss: 0.071
eval intermediate_models/epoch90_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.986619
epoch 90 accuracy on train set is: 0.9866185023627771
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	NaN	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992639

epoch 90 accuracy on test set is: 0.992639040348964

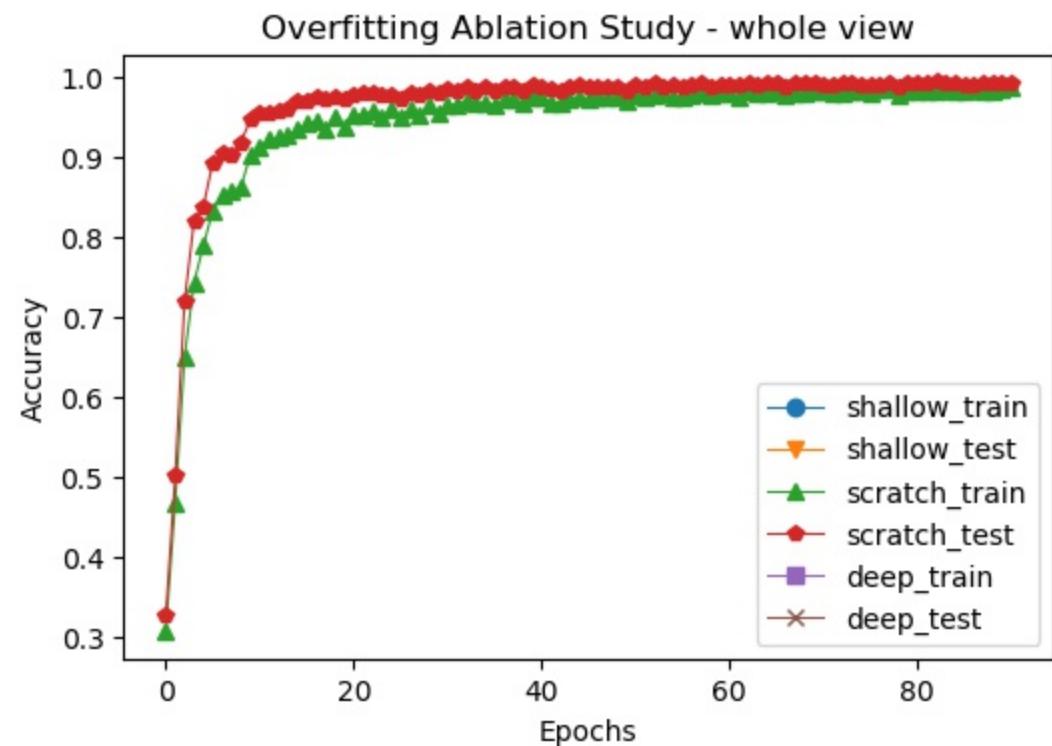
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

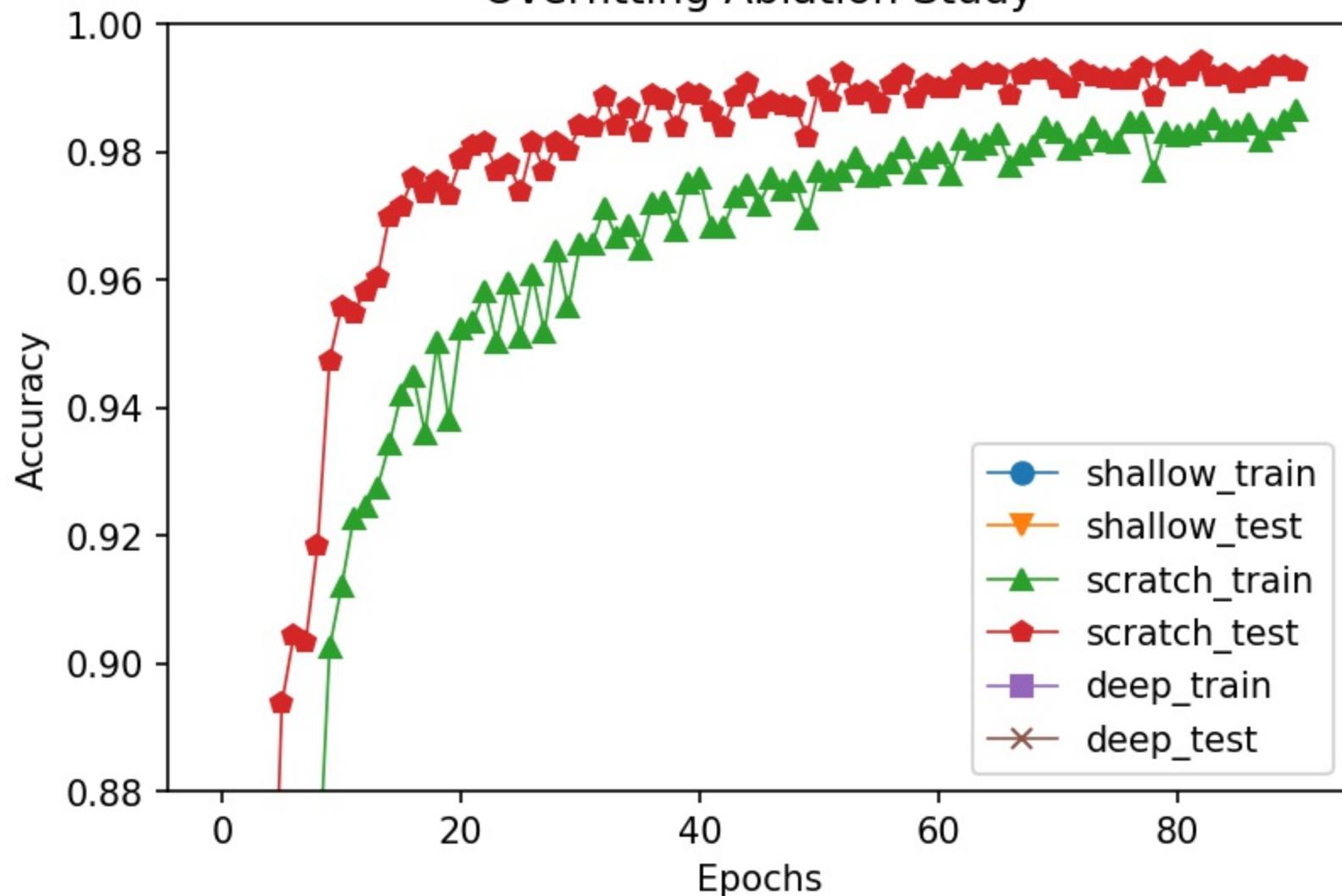
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	NaN	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 90 takes 316 seconds.

[92, 30] loss: 0.038

[92, 60] loss: 0.072

```
[92,    90] loss: 0.031
[92,   120] loss: 0.093
[92,   150] loss: 0.053
[92,   180] loss: 0.072
[92,   210] loss: 0.076
[92,   240] loss: 0.076
[92,   270] loss: 0.058
[92,   300] loss: 0.047
[92,   330] loss: 0.058
[92,   360] loss: 0.081
[92,   390] loss: 0.075
[92,   420] loss: 0.062
[92,   450] loss: 0.066
[92,   480] loss: 0.097
[92,   510] loss: 0.097
[92,   540] loss: 0.068
[92,   570] loss: 0.073
[92,   600] loss: 0.091
[92,   630] loss: 0.105
[92,   660] loss: 0.045
[92,   690] loss: 0.086
[92,   720] loss: 0.078
[92,   750] loss: 0.058
[92,   780] loss: 0.085
[92,   810] loss: 0.050
[92,   840] loss: 0.059
[92,   870] loss: 0.068
[92,   900] loss: 0.043
[92,   930] loss: 0.085
[92,   960] loss: 0.059
[92,   990] loss: 0.085
[92,  1020] loss: 0.072
[92,  1050] loss: 0.057
[92,  1080] loss: 0.062
[92,  1110] loss: 0.062
[92,  1140] loss: 0.072
[92,  1170] loss: 0.048
[92,  1200] loss: 0.054
[92,  1230] loss: 0.102
```

```
[92, 1260] loss: 0.065
[92, 1290] loss: 0.056
[92, 1320] loss: 0.073
[92, 1350] loss: 0.098
[92, 1380] loss: 0.092
[92, 1410] loss: 0.052
[92, 1440] loss: 0.057
[92, 1470] loss: 0.063
[92, 1500] loss: 0.034
[92, 1530] loss: 0.060
[92, 1560] loss: 0.092
[92, 1590] loss: 0.072
[92, 1620] loss: 0.079
[92, 1650] loss: 0.053
[92, 1680] loss: 0.058
[92, 1710] loss: 0.053
[92, 1740] loss: 0.075
[92, 1770] loss: 0.082
[92, 1800] loss: 0.052
[92, 1830] loss: 0.085
[92, 1860] loss: 0.047
[92, 1890] loss: 0.070
[92, 1920] loss: 0.068
[92, 1950] loss: 0.060
[92, 1980] loss: 0.052
[92, 2010] loss: 0.057
[92, 2040] loss: 0.065
[92, 2070] loss: 0.035
[92, 2100] loss: 0.064
[92, 2130] loss: 0.065
[92, 2160] loss: 0.088
[92, 2190] loss: 0.043
eval intermediate_models/epoch91_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.985119
epoch 91 accuracy on train set is: 0.9851190476190477
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	NaN	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992730

epoch 91 accuracy on test set is: 0.9927299163940385

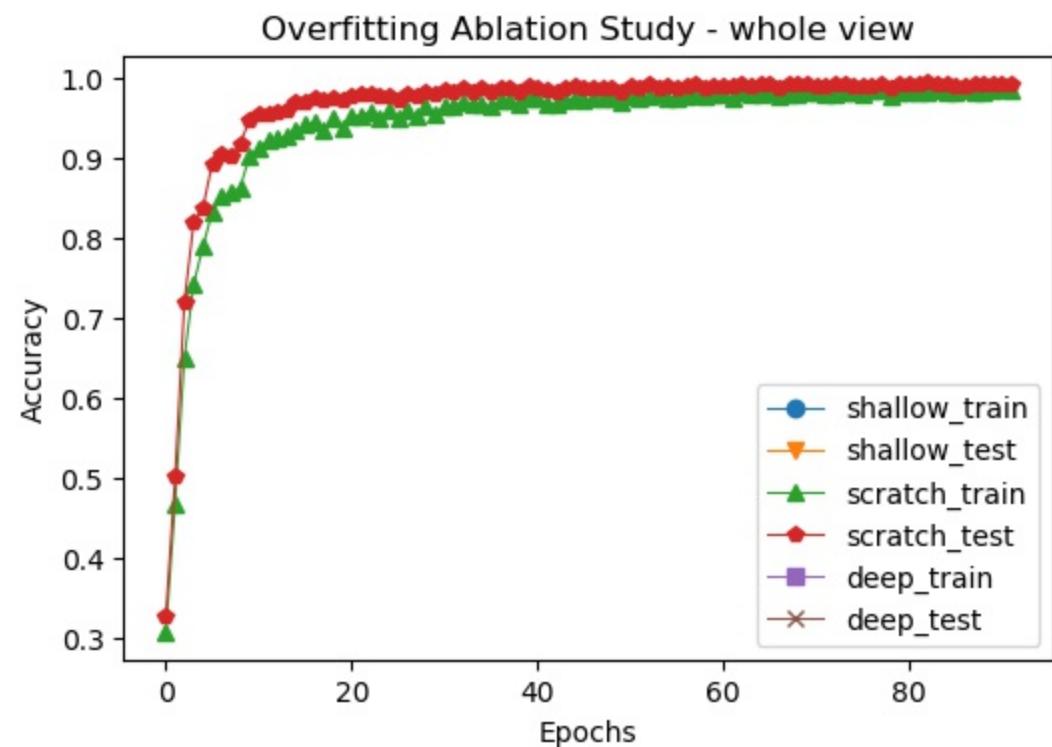
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

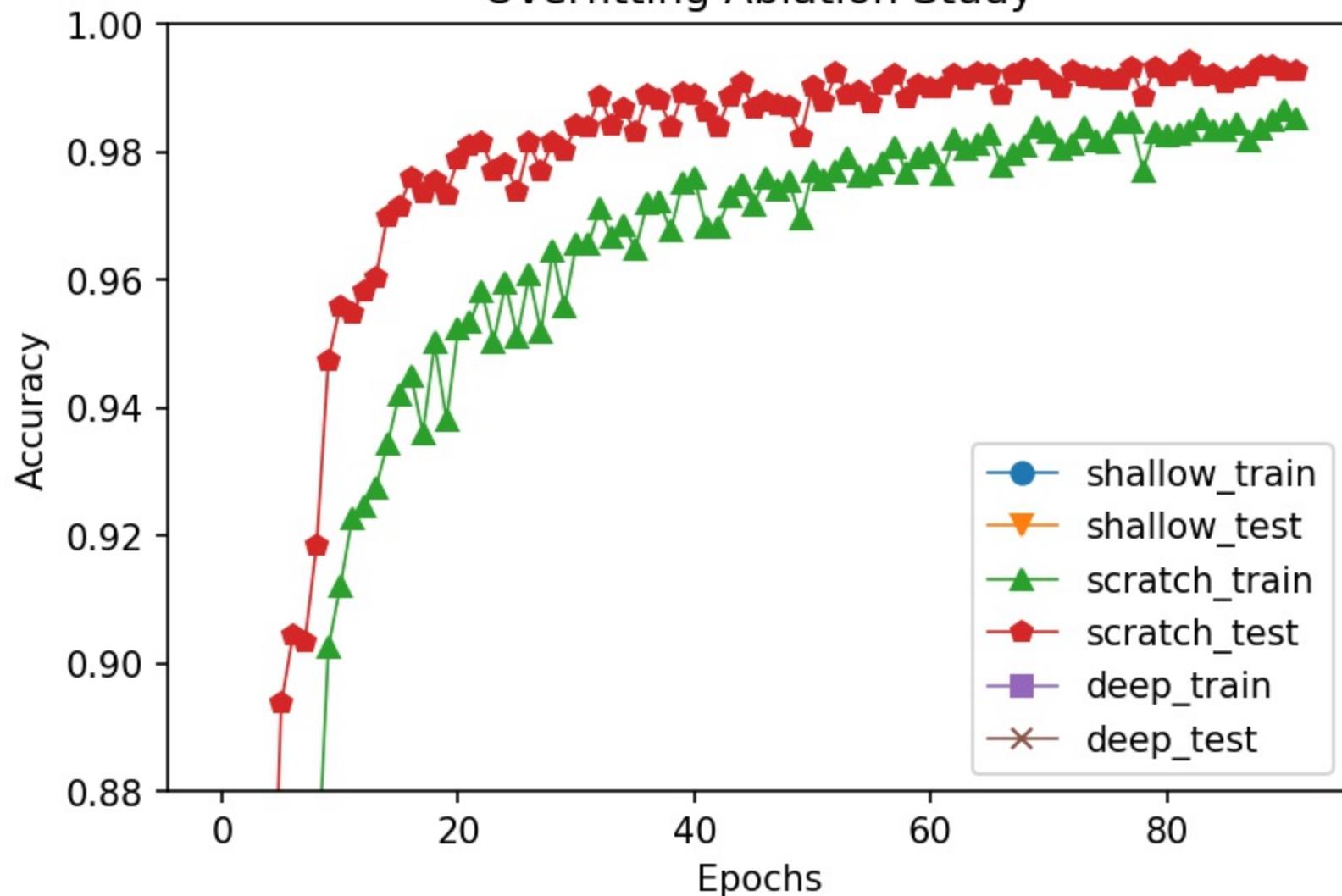
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	NaN	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 91 takes 317 seconds.

[93, 30] loss: 0.063

[93, 60] loss: 0.064

```
[93,    90] loss: 0.053
[93,   120] loss: 0.050
[93,   150] loss: 0.066
[93,   180] loss: 0.077
[93,   210] loss: 0.058
[93,   240] loss: 0.043
[93,   270] loss: 0.060
[93,   300] loss: 0.076
[93,   330] loss: 0.039
[93,   360] loss: 0.042
[93,   390] loss: 0.058
[93,   420] loss: 0.070
[93,   450] loss: 0.079
[93,   480] loss: 0.059
[93,   510] loss: 0.087
[93,   540] loss: 0.068
[93,   570] loss: 0.066
[93,   600] loss: 0.100
[93,   630] loss: 0.069
[93,   660] loss: 0.049
[93,   690] loss: 0.035
[93,   720] loss: 0.064
[93,   750] loss: 0.052
[93,   780] loss: 0.051
[93,   810] loss: 0.084
[93,   840] loss: 0.051
[93,   870] loss: 0.075
[93,   900] loss: 0.052
[93,   930] loss: 0.112
[93,   960] loss: 0.072
[93,   990] loss: 0.116
[93,  1020] loss: 0.060
[93,  1050] loss: 0.086
[93,  1080] loss: 0.038
[93,  1110] loss: 0.057
[93,  1140] loss: 0.067
[93,  1170] loss: 0.042
[93,  1200] loss: 0.033
[93,  1230] loss: 0.044
```

```
[93, 1260] loss: 0.037
[93, 1290] loss: 0.049
[93, 1320] loss: 0.080
[93, 1350] loss: 0.097
[93, 1380] loss: 0.068
[93, 1410] loss: 0.080
[93, 1440] loss: 0.073
[93, 1470] loss: 0.077
[93, 1500] loss: 0.065
[93, 1530] loss: 0.080
[93, 1560] loss: 0.046
[93, 1590] loss: 0.069
[93, 1620] loss: 0.059
[93, 1650] loss: 0.038
[93, 1680] loss: 0.064
[93, 1710] loss: 0.077
[93, 1740] loss: 0.051
[93, 1770] loss: 0.057
[93, 1800] loss: 0.063
[93, 1830] loss: 0.081
[93, 1860] loss: 0.105
[93, 1890] loss: 0.087
[93, 1920] loss: 0.082
[93, 1950] loss: 0.048
[93, 1980] loss: 0.071
[93, 2010] loss: 0.072
[93, 2040] loss: 0.055
[93, 2070] loss: 0.056
[93, 2100] loss: 0.071
[93, 2130] loss: 0.069
[93, 2160] loss: 0.070
[93, 2190] loss: 0.063
eval intermediate_models/epoch92_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.986005
epoch 92 accuracy on train set is: 0.9860050890585241
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	NaN	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993003

epoch 92 accuracy on test set is: 0.9930025445292621

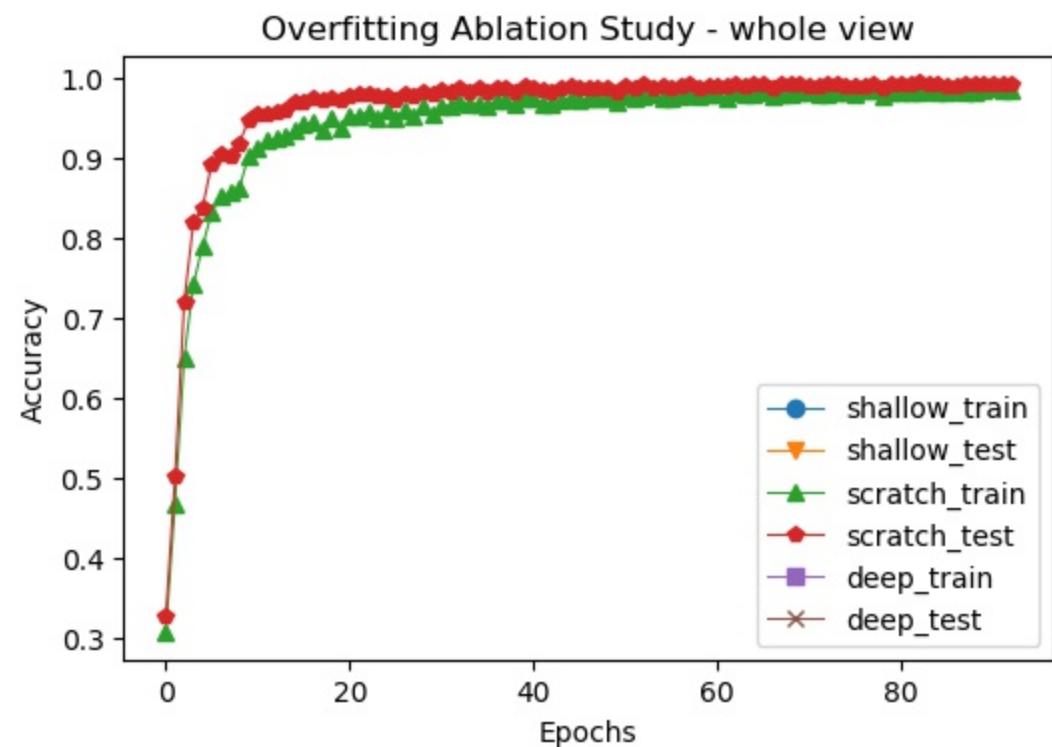
evaluating on test set takes 16 seconds.

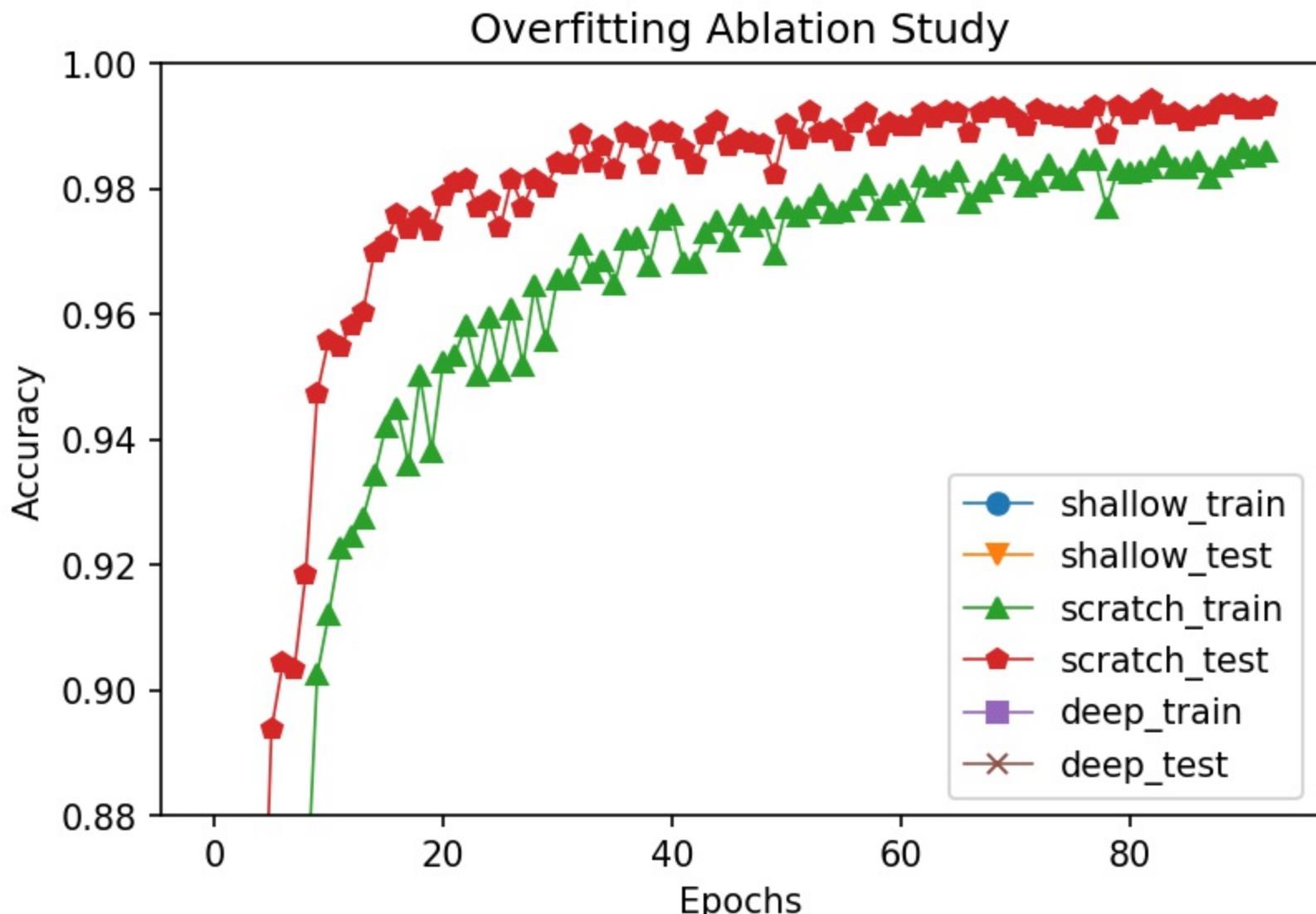
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	NaN	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 92 takes 316 seconds.

[94, 30] loss: 0.084

[94, 60] loss: 0.056

```
[94,    90] loss: 0.053
[94,   120] loss: 0.069
[94,   150] loss: 0.049
[94,   180] loss: 0.077
[94,   210] loss: 0.054
[94,   240] loss: 0.058
[94,   270] loss: 0.043
[94,   300] loss: 0.093
[94,   330] loss: 0.058
[94,   360] loss: 0.072
[94,   390] loss: 0.051
[94,   420] loss: 0.045
[94,   450] loss: 0.076
[94,   480] loss: 0.033
[94,   510] loss: 0.060
[94,   540] loss: 0.071
[94,   570] loss: 0.055
[94,   600] loss: 0.071
[94,   630] loss: 0.063
[94,   660] loss: 0.042
[94,   690] loss: 0.067
[94,   720] loss: 0.094
[94,   750] loss: 0.074
[94,   780] loss: 0.043
[94,   810] loss: 0.092
[94,   840] loss: 0.047
[94,   870] loss: 0.077
[94,   900] loss: 0.054
[94,   930] loss: 0.050
[94,   960] loss: 0.068
[94,   990] loss: 0.054
[94,  1020] loss: 0.054
[94,  1050] loss: 0.067
[94,  1080] loss: 0.073
[94,  1110] loss: 0.076
[94,  1140] loss: 0.069
[94,  1170] loss: 0.043
[94,  1200] loss: 0.077
[94,  1230] loss: 0.062
```

```
[94, 1260] loss: 0.063
[94, 1290] loss: 0.060
[94, 1320] loss: 0.048
[94, 1350] loss: 0.046
[94, 1380] loss: 0.036
[94, 1410] loss: 0.082
[94, 1440] loss: 0.063
[94, 1470] loss: 0.065
[94, 1500] loss: 0.088
[94, 1530] loss: 0.090
[94, 1560] loss: 0.074
[94, 1590] loss: 0.064
[94, 1620] loss: 0.097
[94, 1650] loss: 0.083
[94, 1680] loss: 0.041
[94, 1710] loss: 0.065
[94, 1740] loss: 0.049
[94, 1770] loss: 0.034
[94, 1800] loss: 0.063
[94, 1830] loss: 0.061
[94, 1860] loss: 0.049
[94, 1890] loss: 0.059
[94, 1920] loss: 0.084
[94, 1950] loss: 0.073
[94, 1980] loss: 0.091
[94, 2010] loss: 0.072
[94, 2040] loss: 0.076
[94, 2070] loss: 0.072
[94, 2100] loss: 0.092
[94, 2130] loss: 0.073
[94, 2160] loss: 0.077
[94, 2190] loss: 0.042
eval intermediate_models/epoch93_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.985392
epoch 93 accuracy on train set is: 0.9853916757542712
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	NaN	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992276

epoch 93 accuracy on test set is: 0.9922755361686659

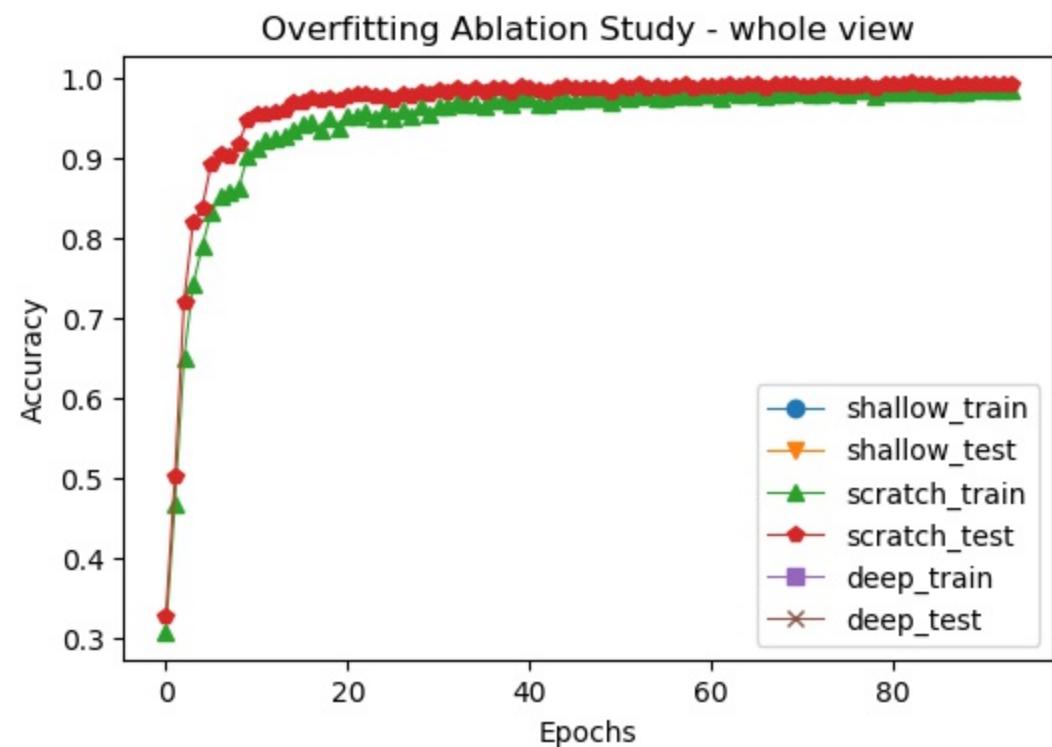
evaluating on test set takes 16 seconds.

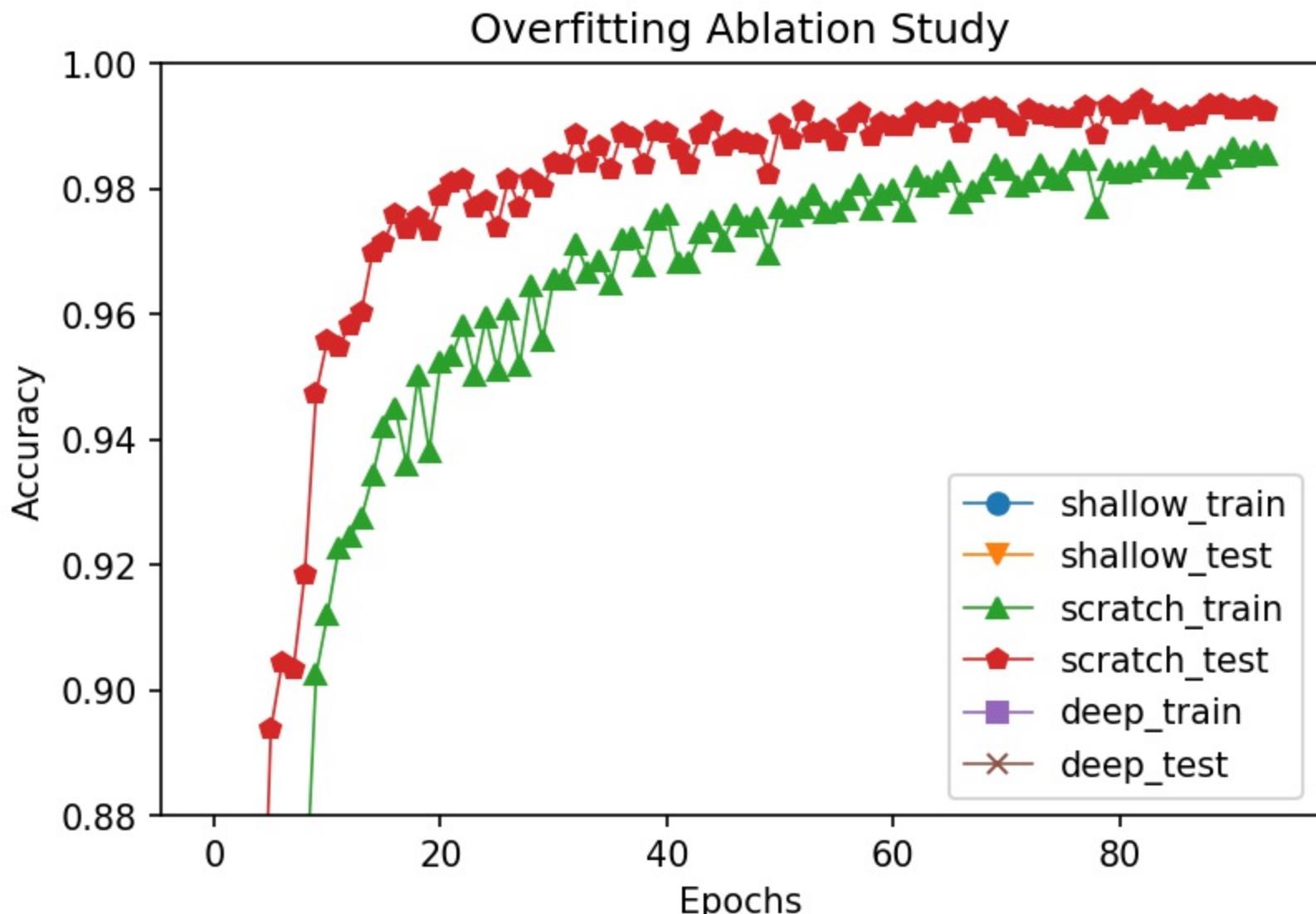
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	NaN	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 93 takes 315 seconds.

[95, 30] loss: 0.080

[95, 60] loss: 0.049

```
[95,    90] loss: 0.065
[95,   120] loss: 0.071
[95,   150] loss: 0.090
[95,   180] loss: 0.052
[95,   210] loss: 0.053
[95,   240] loss: 0.045
[95,   270] loss: 0.098
[95,   300] loss: 0.052
[95,   330] loss: 0.072
[95,   360] loss: 0.066
[95,   390] loss: 0.054
[95,   420] loss: 0.055
[95,   450] loss: 0.036
[95,   480] loss: 0.030
[95,   510] loss: 0.066
[95,   540] loss: 0.052
[95,   570] loss: 0.095
[95,   600] loss: 0.048
[95,   630] loss: 0.068
[95,   660] loss: 0.042
[95,   690] loss: 0.043
[95,   720] loss: 0.060
[95,   750] loss: 0.051
[95,   780] loss: 0.063
[95,   810] loss: 0.048
[95,   840] loss: 0.068
[95,   870] loss: 0.069
[95,   900] loss: 0.054
[95,   930] loss: 0.039
[95,   960] loss: 0.092
[95,   990] loss: 0.051
[95,  1020] loss: 0.052
[95,  1050] loss: 0.057
[95,  1080] loss: 0.090
[95,  1110] loss: 0.064
[95,  1140] loss: 0.055
[95,  1170] loss: 0.089
[95,  1200] loss: 0.080
[95,  1230] loss: 0.050
```

```
[95, 1260] loss: 0.065
[95, 1290] loss: 0.048
[95, 1320] loss: 0.063
[95, 1350] loss: 0.075
[95, 1380] loss: 0.053
[95, 1410] loss: 0.094
[95, 1440] loss: 0.028
[95, 1470] loss: 0.081
[95, 1500] loss: 0.063
[95, 1530] loss: 0.056
[95, 1560] loss: 0.055
[95, 1590] loss: 0.071
[95, 1620] loss: 0.051
[95, 1650] loss: 0.066
[95, 1680] loss: 0.044
[95, 1710] loss: 0.048
[95, 1740] loss: 0.063
[95, 1770] loss: 0.070
[95, 1800] loss: 0.057
[95, 1830] loss: 0.055
[95, 1860] loss: 0.068
[95, 1890] loss: 0.059
[95, 1920] loss: 0.061
[95, 1950] loss: 0.053
[95, 1980] loss: 0.035
[95, 2010] loss: 0.083
[95, 2040] loss: 0.047
[95, 2070] loss: 0.066
[95, 2100] loss: 0.072
[95, 2130] loss: 0.076
[95, 2160] loss: 0.088
[95, 2190] loss: 0.047
eval intermediate_models/epoch94_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.984392
epoch 94 accuracy on train set is: 0.9843920392584514
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	NaN	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992185

epoch 94 accuracy on test set is: 0.9921846601235914

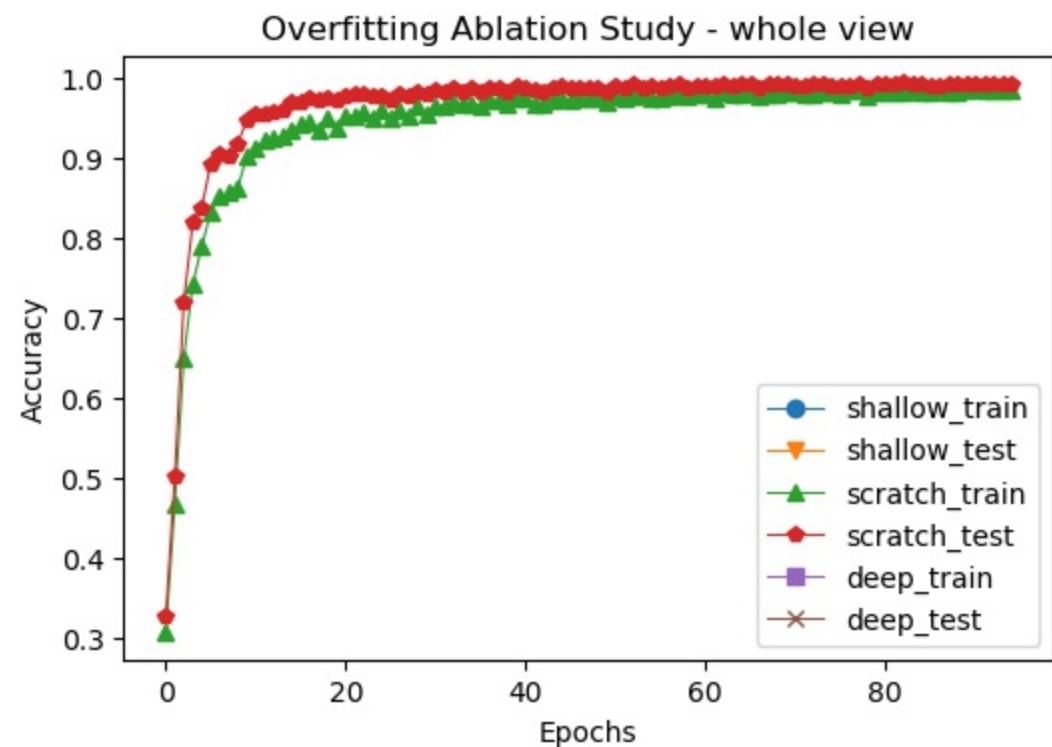
evaluating on test set takes 16 seconds.

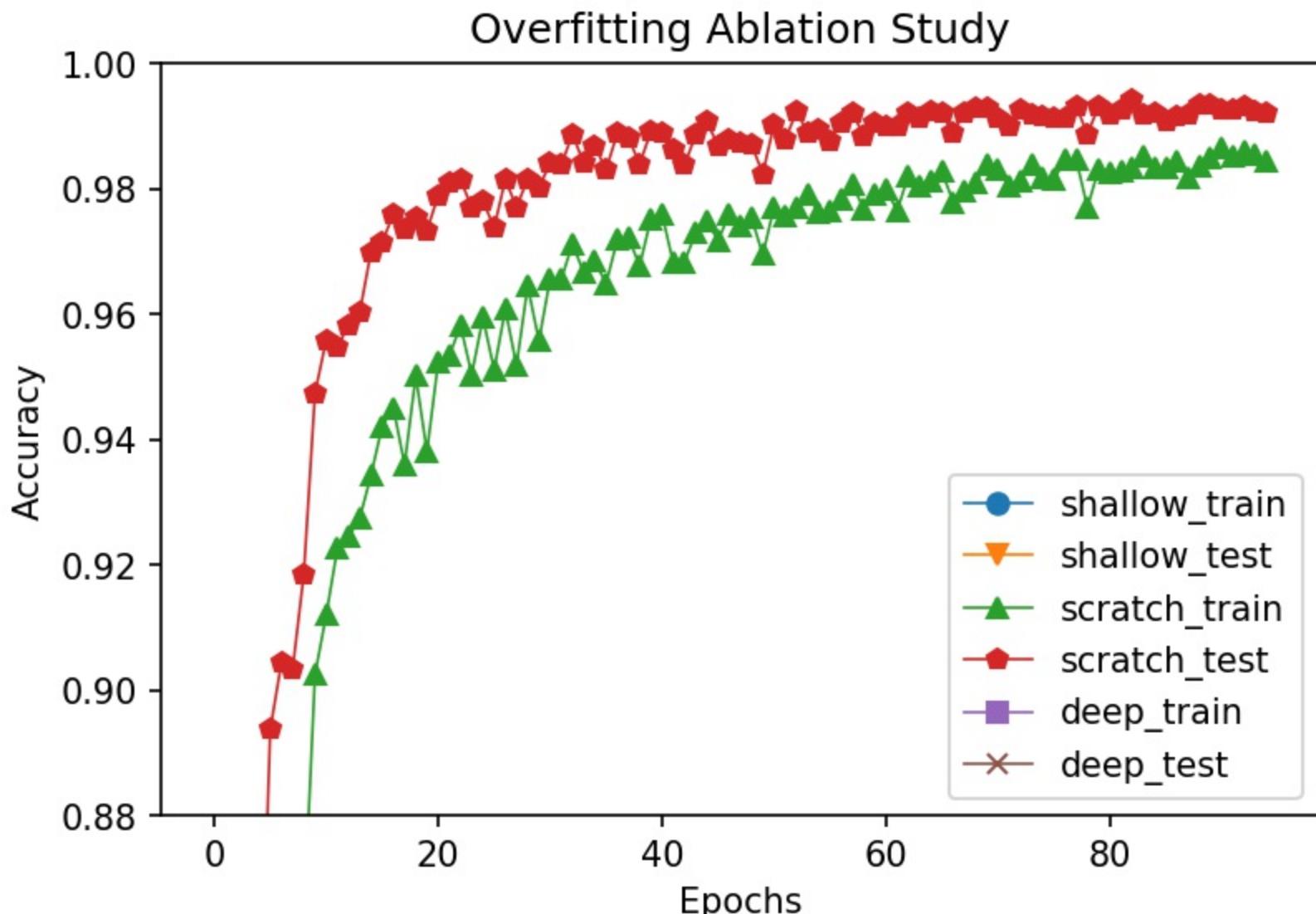
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	NaN	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 94 takes 315 seconds.

[96, 30] loss: 0.030

[96, 60] loss: 0.053

```
[96,    90] loss: 0.087
[96,   120] loss: 0.052
[96,   150] loss: 0.061
[96,   180] loss: 0.053
[96,   210] loss: 0.079
[96,   240] loss: 0.035
[96,   270] loss: 0.065
[96,   300] loss: 0.056
[96,   330] loss: 0.034
[96,   360] loss: 0.063
[96,   390] loss: 0.070
[96,   420] loss: 0.062
[96,   450] loss: 0.031
[96,   480] loss: 0.060
[96,   510] loss: 0.039
[96,   540] loss: 0.090
[96,   570] loss: 0.042
[96,   600] loss: 0.060
[96,   630] loss: 0.049
[96,   660] loss: 0.065
[96,   690] loss: 0.060
[96,   720] loss: 0.076
[96,   750] loss: 0.063
[96,   780] loss: 0.072
[96,   810] loss: 0.052
[96,   840] loss: 0.082
[96,   870] loss: 0.060
[96,   900] loss: 0.048
[96,   930] loss: 0.049
[96,   960] loss: 0.049
[96,   990] loss: 0.082
[96,  1020] loss: 0.062
[96,  1050] loss: 0.052
[96,  1080] loss: 0.066
[96,  1110] loss: 0.084
[96,  1140] loss: 0.069
[96,  1170] loss: 0.057
[96,  1200] loss: 0.055
[96,  1230] loss: 0.062
```

```
[96, 1260] loss: 0.050
[96, 1290] loss: 0.066
[96, 1320] loss: 0.054
[96, 1350] loss: 0.057
[96, 1380] loss: 0.079
[96, 1410] loss: 0.069
[96, 1440] loss: 0.045
[96, 1470] loss: 0.052
[96, 1500] loss: 0.052
[96, 1530] loss: 0.074
[96, 1560] loss: 0.081
[96, 1590] loss: 0.074
[96, 1620] loss: 0.069
[96, 1650] loss: 0.088
[96, 1680] loss: 0.048
[96, 1710] loss: 0.048
[96, 1740] loss: 0.048
[96, 1770] loss: 0.069
[96, 1800] loss: 0.049
[96, 1830] loss: 0.066
[96, 1860] loss: 0.055
[96, 1890] loss: 0.051
[96, 1920] loss: 0.064
[96, 1950] loss: 0.047
[96, 1980] loss: 0.045
[96, 2010] loss: 0.072
[96, 2040] loss: 0.061
[96, 2070] loss: 0.058
[96, 2100] loss: 0.058
[96, 2130] loss: 0.087
[96, 2160] loss: 0.053
[96, 2190] loss: 0.049
eval intermediate_models/epoch95_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.983756
epoch 95 accuracy on train set is: 0.9837559069429298
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	NaN	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992457

epoch 95 accuracy on test set is: 0.992457288258815

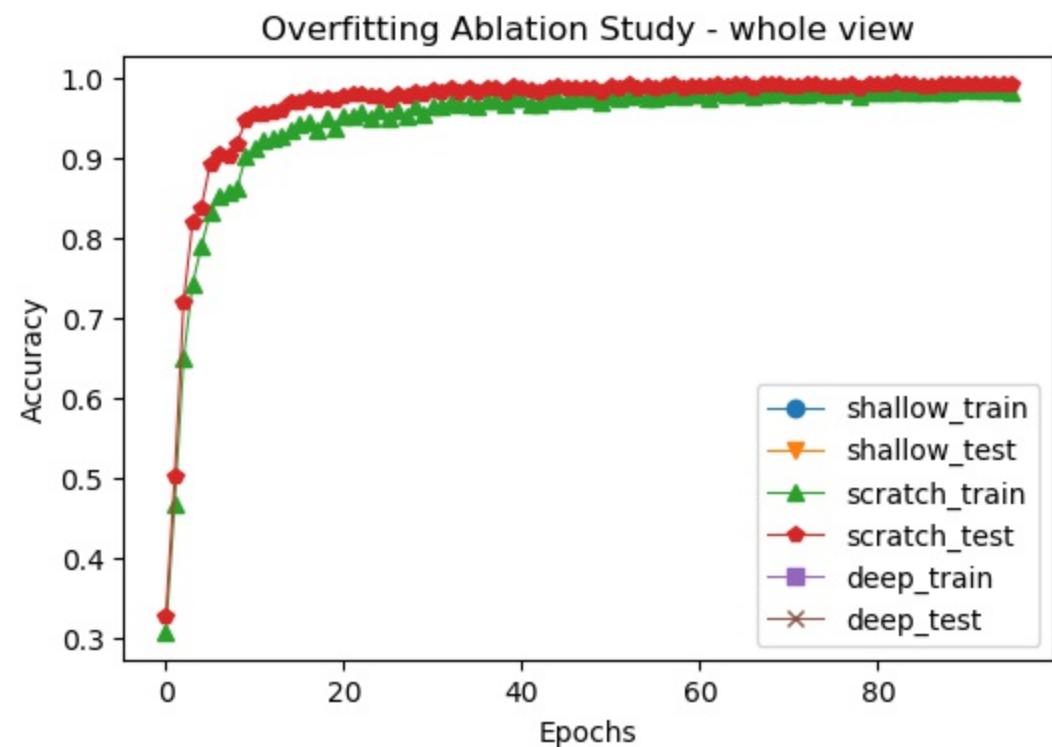
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

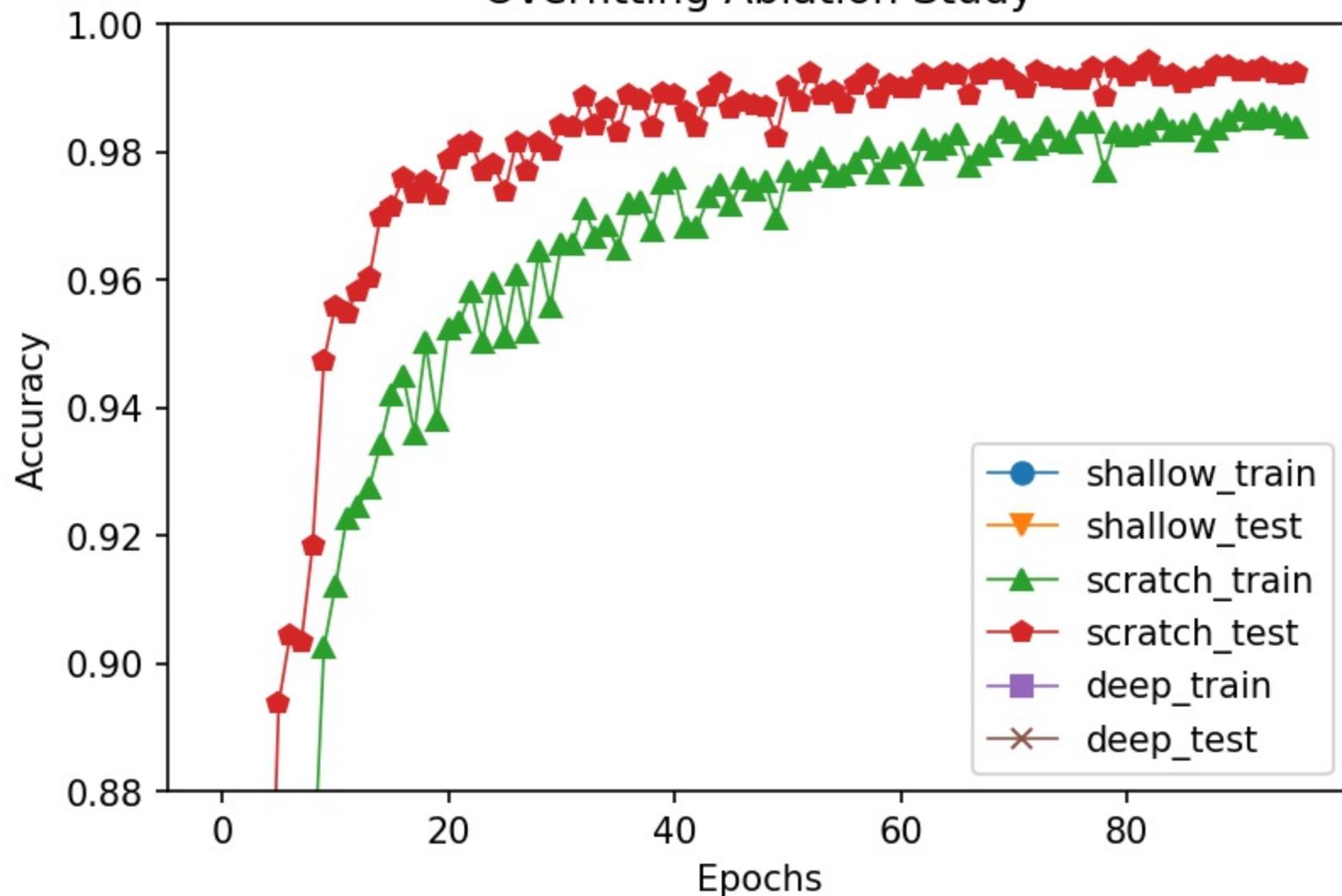
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	NaN	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 95 takes 315 seconds.

[97, 30] loss: 0.056

[97, 60] loss: 0.064

```
[97,    90] loss: 0.093
[97,   120] loss: 0.078
[97,   150] loss: 0.068
[97,   180] loss: 0.079
[97,   210] loss: 0.086
[97,   240] loss: 0.030
[97,   270] loss: 0.045
[97,   300] loss: 0.062
[97,   330] loss: 0.040
[97,   360] loss: 0.051
[97,   390] loss: 0.067
[97,   420] loss: 0.054
[97,   450] loss: 0.036
[97,   480] loss: 0.081
[97,   510] loss: 0.093
[97,   540] loss: 0.050
[97,   570] loss: 0.026
[97,   600] loss: 0.038
[97,   630] loss: 0.060
[97,   660] loss: 0.046
[97,   690] loss: 0.045
[97,   720] loss: 0.047
[97,   750] loss: 0.063
[97,   780] loss: 0.070
[97,   810] loss: 0.090
[97,   840] loss: 0.087
[97,   870] loss: 0.057
[97,   900] loss: 0.053
[97,   930] loss: 0.059
[97,   960] loss: 0.051
[97,   990] loss: 0.050
[97,  1020] loss: 0.070
[97,  1050] loss: 0.078
[97,  1080] loss: 0.064
[97,  1110] loss: 0.091
[97,  1140] loss: 0.054
[97,  1170] loss: 0.043
[97,  1200] loss: 0.056
[97,  1230] loss: 0.056
```

```
[97, 1260] loss: 0.043
[97, 1290] loss: 0.086
[97, 1320] loss: 0.083
[97, 1350] loss: 0.074
[97, 1380] loss: 0.034
[97, 1410] loss: 0.051
[97, 1440] loss: 0.046
[97, 1470] loss: 0.031
[97, 1500] loss: 0.080
[97, 1530] loss: 0.047
[97, 1560] loss: 0.092
[97, 1590] loss: 0.033
[97, 1620] loss: 0.054
[97, 1650] loss: 0.049
[97, 1680] loss: 0.039
[97, 1710] loss: 0.060
[97, 1740] loss: 0.057
[97, 1770] loss: 0.069
[97, 1800] loss: 0.040
[97, 1830] loss: 0.066
[97, 1860] loss: 0.083
[97, 1890] loss: 0.056
[97, 1920] loss: 0.088
[97, 1950] loss: 0.073
[97, 1980] loss: 0.079
[97, 2010] loss: 0.100
[97, 2040] loss: 0.047
[97, 2070] loss: 0.084
[97, 2100] loss: 0.060
[97, 2130] loss: 0.082
[97, 2160] loss: 0.049
[97, 2190] loss: 0.071
eval intermediate_models/epoch96_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.984324
epoch 96 accuracy on train set is: 0.9843238822246456
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	NaN	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990822

epoch 96 accuracy on test set is: 0.9908215194474737

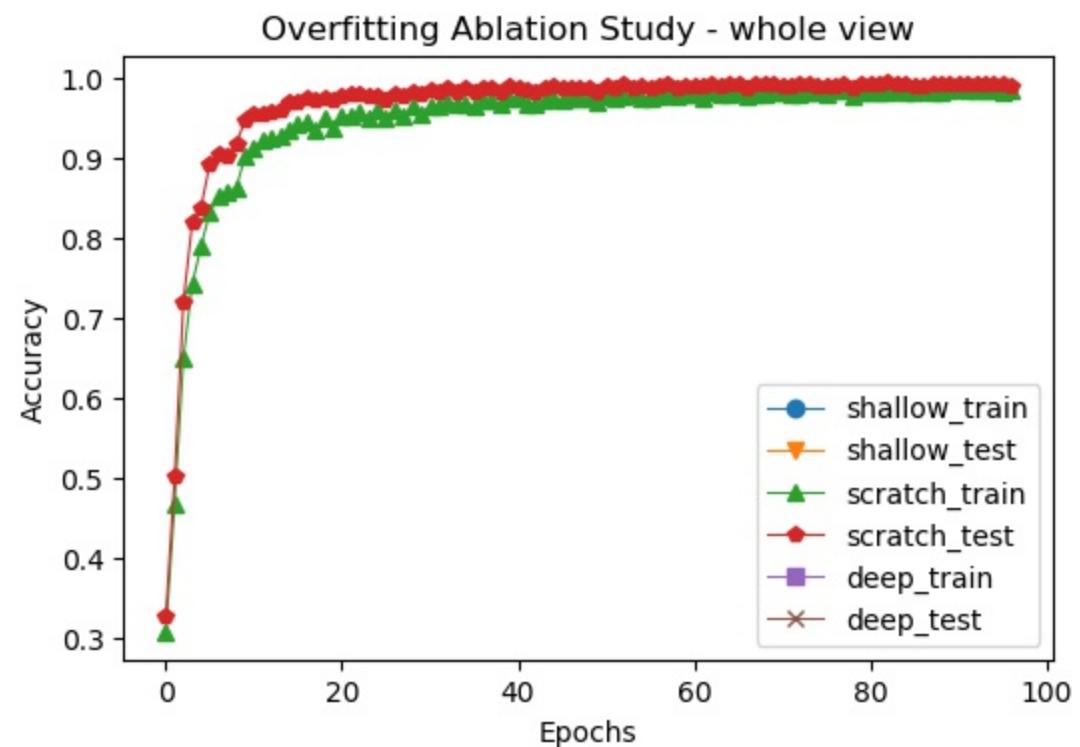
evaluating on test set takes 16 seconds.

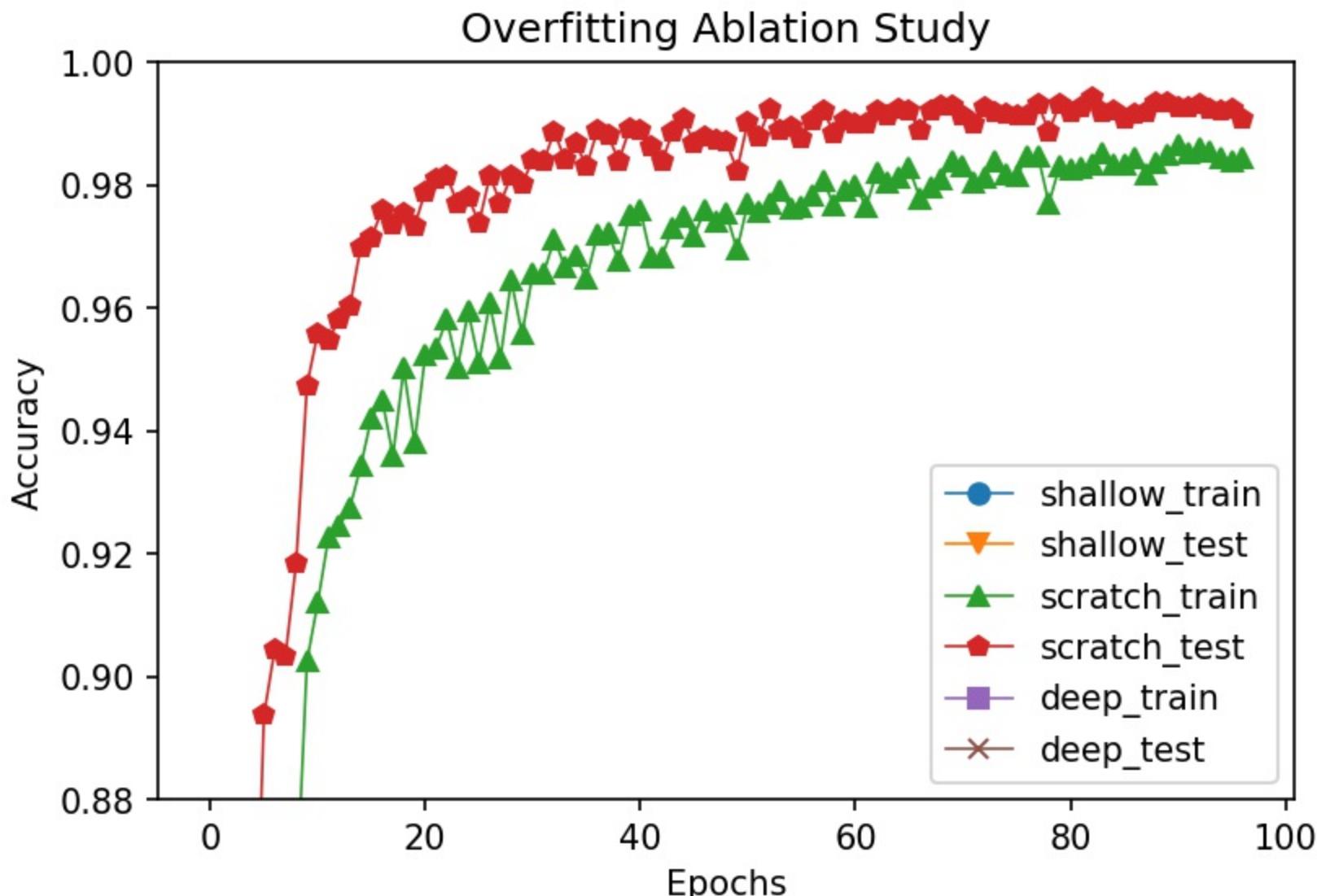
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	NaN	NaN	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 96 takes 315 seconds.

[98, 30] loss: 0.060

[98, 60] loss: 0.081

```
[98,    90] loss: 0.050
[98,   120] loss: 0.064
[98,   150] loss: 0.083
[98,   180] loss: 0.059
[98,   210] loss: 0.048
[98,   240] loss: 0.054
[98,   270] loss: 0.033
[98,   300] loss: 0.068
[98,   330] loss: 0.089
[98,   360] loss: 0.146
[98,   390] loss: 0.057
[98,   420] loss: 0.067
[98,   450] loss: 0.093
[98,   480] loss: 0.092
[98,   510] loss: 0.077
[98,   540] loss: 0.041
[98,   570] loss: 0.075
[98,   600] loss: 0.041
[98,   630] loss: 0.060
[98,   660] loss: 0.051
[98,   690] loss: 0.065
[98,   720] loss: 0.057
[98,   750] loss: 0.066
[98,   780] loss: 0.037
[98,   810] loss: 0.060
[98,   840] loss: 0.067
[98,   870] loss: 0.047
[98,   900] loss: 0.060
[98,   930] loss: 0.053
[98,   960] loss: 0.035
[98,   990] loss: 0.055
[98,  1020] loss: 0.053
[98,  1050] loss: 0.063
[98,  1080] loss: 0.057
[98,  1110] loss: 0.048
[98,  1140] loss: 0.068
[98,  1170] loss: 0.067
[98,  1200] loss: 0.083
[98,  1230] loss: 0.099
```

```
[98, 1260] loss: 0.100
[98, 1290] loss: 0.065
[98, 1320] loss: 0.082
[98, 1350] loss: 0.061
[98, 1380] loss: 0.036
[98, 1410] loss: 0.059
[98, 1440] loss: 0.068
[98, 1470] loss: 0.038
[98, 1500] loss: 0.094
[98, 1530] loss: 0.063
[98, 1560] loss: 0.048
[98, 1590] loss: 0.049
[98, 1620] loss: 0.063
[98, 1650] loss: 0.073
[98, 1680] loss: 0.044
[98, 1710] loss: 0.067
[98, 1740] loss: 0.075
[98, 1770] loss: 0.054
[98, 1800] loss: 0.054
[98, 1830] loss: 0.046
[98, 1860] loss: 0.096
[98, 1890] loss: 0.045
[98, 1920] loss: 0.074
[98, 1950] loss: 0.042
[98, 1980] loss: 0.067
[98, 2010] loss: 0.053
[98, 2040] loss: 0.041
[98, 2070] loss: 0.058
[98, 2100] loss: 0.078
[98, 2130] loss: 0.048
[98, 2160] loss: 0.040
[98, 2190] loss: 0.039
eval intermediate_models/epoch97_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.987050
epoch 97 accuracy on train set is: 0.9870501635768811
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	NaN	NaN	NaN
98	NaN	NaN		NaN	NaN	NaN
99	NaN	NaN		NaN	NaN	NaN
100	NaN	NaN		NaN	NaN	NaN
101	NaN	NaN		NaN	NaN	NaN
102	NaN	NaN		NaN	NaN	NaN
103	NaN	NaN		NaN	NaN	NaN
104	NaN	NaN		NaN	NaN	NaN
105	NaN	NaN		NaN	NaN	NaN
106	NaN	NaN		NaN	NaN	NaN
107	NaN	NaN		NaN	NaN	NaN
108	NaN	NaN		NaN	NaN	NaN
109	NaN	NaN		NaN	NaN	NaN
110	NaN	NaN		NaN	NaN	NaN
111	NaN	NaN		NaN	NaN	NaN
112	NaN	NaN		NaN	NaN	NaN
113	NaN	NaN		NaN	NaN	NaN
114	NaN	NaN		NaN	NaN	NaN
115	NaN	NaN		NaN	NaN	NaN
116	NaN	NaN		NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.994366

epoch 97 accuracy on test set is: 0.9943656852053798

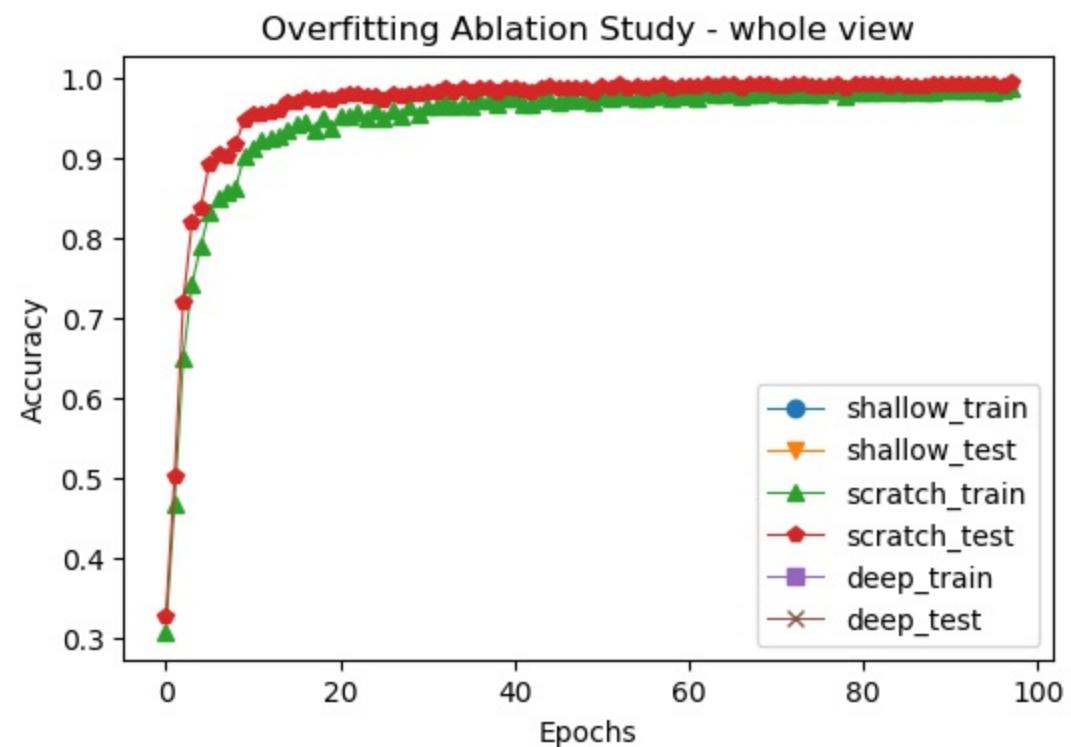
evaluating on test set takes 16 seconds.

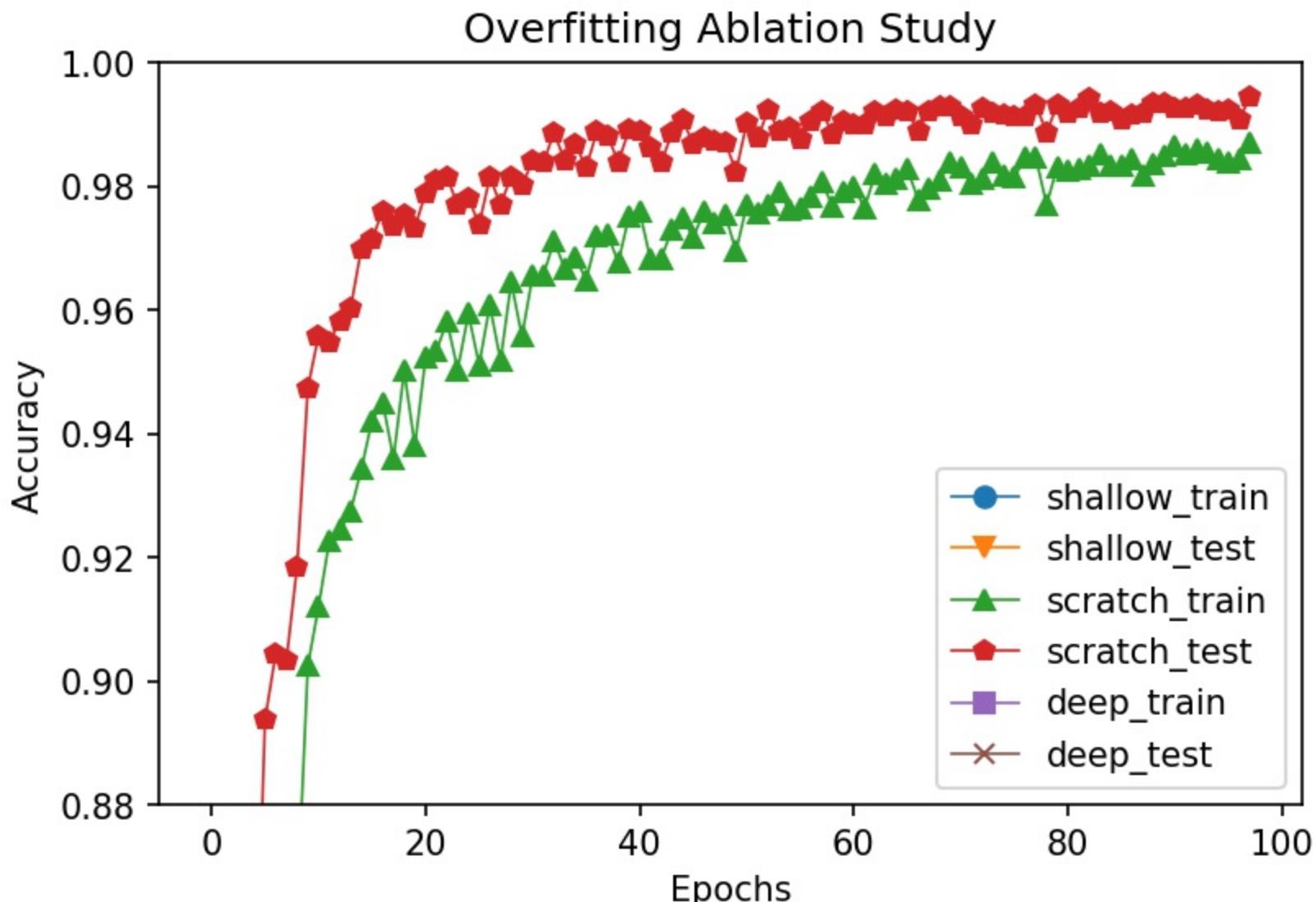
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	NaN	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 97 takes 315 seconds.

[99, 30] loss: 0.056

[99, 60] loss: 0.068

```
[99,    90] loss: 0.090
[99,   120] loss: 0.060
[99,   150] loss: 0.082
[99,   180] loss: 0.050
[99,   210] loss: 0.062
[99,   240] loss: 0.077
[99,   270] loss: 0.083
[99,   300] loss: 0.057
[99,   330] loss: 0.115
[99,   360] loss: 0.072
[99,   390] loss: 0.050
[99,   420] loss: 0.063
[99,   450] loss: 0.025
[99,   480] loss: 0.066
[99,   510] loss: 0.061
[99,   540] loss: 0.043
[99,   570] loss: 0.047
[99,   600] loss: 0.068
[99,   630] loss: 0.052
[99,   660] loss: 0.047
[99,   690] loss: 0.073
[99,   720] loss: 0.055
[99,   750] loss: 0.079
[99,   780] loss: 0.059
[99,   810] loss: 0.081
[99,   840] loss: 0.037
[99,   870] loss: 0.049
[99,   900] loss: 0.066
[99,   930] loss: 0.068
[99,   960] loss: 0.080
[99,   990] loss: 0.058
[99,  1020] loss: 0.055
[99,  1050] loss: 0.053
[99,  1080] loss: 0.044
[99,  1110] loss: 0.078
[99,  1140] loss: 0.091
[99,  1170] loss: 0.050
[99,  1200] loss: 0.038
[99,  1230] loss: 0.055
```

```
[99, 1260] loss: 0.069
[99, 1290] loss: 0.051
[99, 1320] loss: 0.049
[99, 1350] loss: 0.064
[99, 1380] loss: 0.070
[99, 1410] loss: 0.096
[99, 1440] loss: 0.070
[99, 1470] loss: 0.082
[99, 1500] loss: 0.061
[99, 1530] loss: 0.087
[99, 1560] loss: 0.077
[99, 1590] loss: 0.072
[99, 1620] loss: 0.057
[99, 1650] loss: 0.048
[99, 1680] loss: 0.081
[99, 1710] loss: 0.079
[99, 1740] loss: 0.056
[99, 1770] loss: 0.054
[99, 1800] loss: 0.052
[99, 1830] loss: 0.047
[99, 1860] loss: 0.053
[99, 1890] loss: 0.094
[99, 1920] loss: 0.075
[99, 1950] loss: 0.062
[99, 1980] loss: 0.063
[99, 2010] loss: 0.060
[99, 2040] loss: 0.083
[99, 2070] loss: 0.062
[99, 2100] loss: 0.073
[99, 2130] loss: 0.071
[99, 2160] loss: 0.041
[99, 2190] loss: 0.098
eval intermediate_models/epoch98_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.984233
epoch 98 accuracy on train set is: 0.9842330061795711
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	NaN	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992639

epoch 98 accuracy on test set is: 0.992639040348964

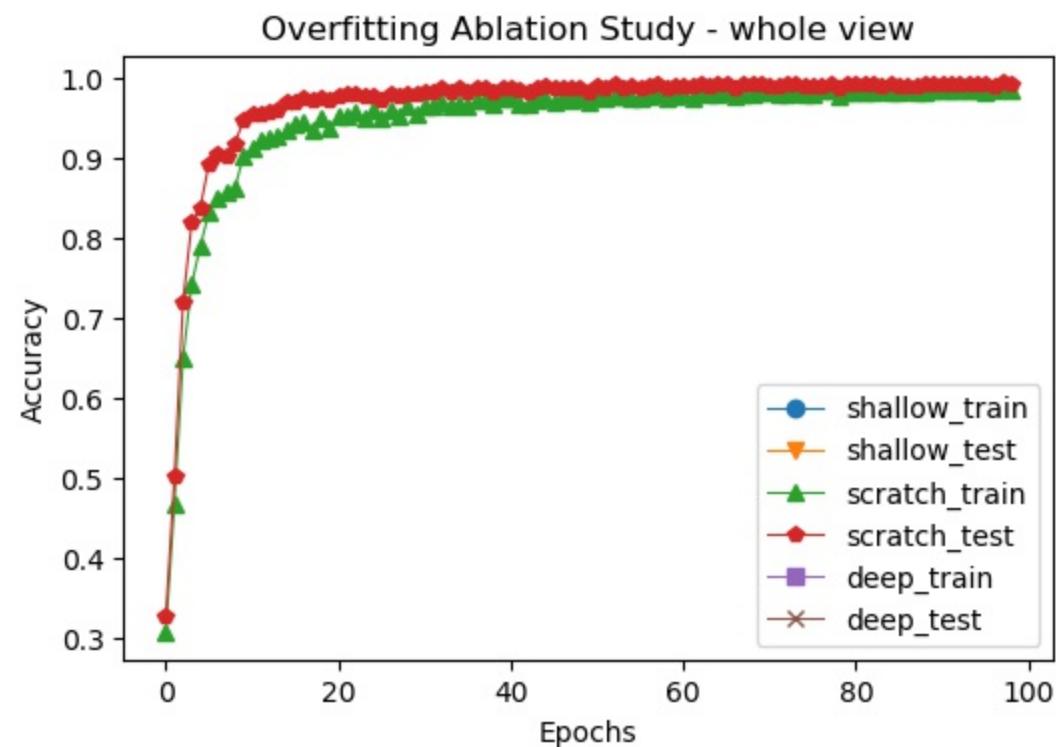
evaluating on test set takes 16 seconds.

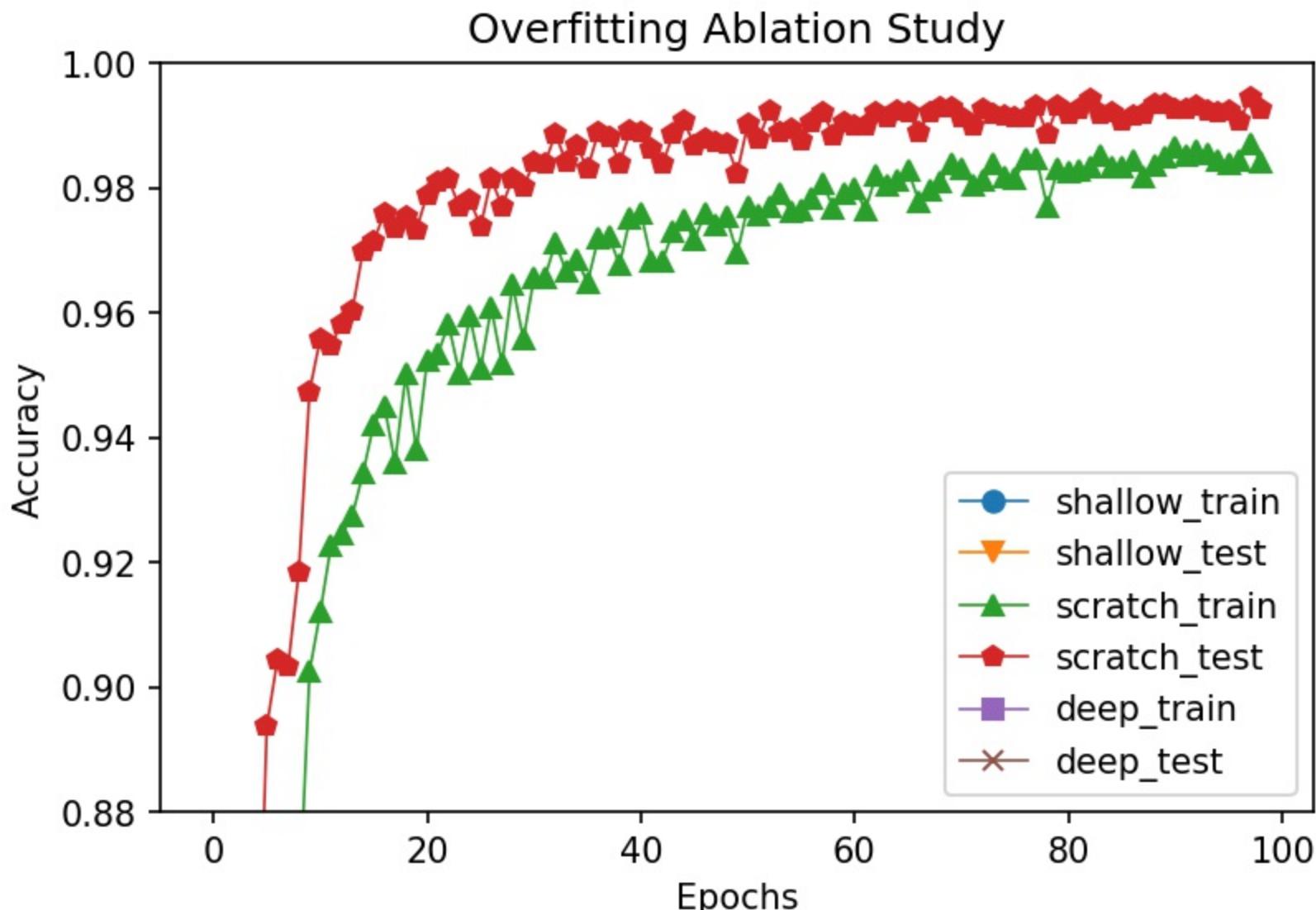
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	NaN	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 98 takes 315 seconds.

[100, 30] loss: 0.042

[100, 60] loss: 0.050

```
[100,    90] loss: 0.052
[100,   120] loss: 0.062
[100,   150] loss: 0.057
[100,   180] loss: 0.061
[100,   210] loss: 0.061
[100,   240] loss: 0.131
[100,   270] loss: 0.064
[100,   300] loss: 0.091
[100,   330] loss: 0.045
[100,   360] loss: 0.038
[100,   390] loss: 0.044
[100,   420] loss: 0.040
[100,   450] loss: 0.075
[100,   480] loss: 0.051
[100,   510] loss: 0.064
[100,   540] loss: 0.054
[100,   570] loss: 0.039
[100,   600] loss: 0.055
[100,   630] loss: 0.071
[100,   660] loss: 0.077
[100,   690] loss: 0.069
[100,   720] loss: 0.085
[100,   750] loss: 0.069
[100,   780] loss: 0.040
[100,   810] loss: 0.048
[100,   840] loss: 0.030
[100,   870] loss: 0.042
[100,   900] loss: 0.045
[100,   930] loss: 0.061
[100,   960] loss: 0.034
[100,   990] loss: 0.054
[100,  1020] loss: 0.078
[100,  1050] loss: 0.027
[100,  1080] loss: 0.049
[100,  1110] loss: 0.048
[100,  1140] loss: 0.038
[100,  1170] loss: 0.079
[100,  1200] loss: 0.064
[100,  1230] loss: 0.062
```

```
[100, 1260] loss: 0.050
[100, 1290] loss: 0.021
[100, 1320] loss: 0.050
[100, 1350] loss: 0.038
[100, 1380] loss: 0.059
[100, 1410] loss: 0.052
[100, 1440] loss: 0.076
[100, 1470] loss: 0.067
[100, 1500] loss: 0.065
[100, 1530] loss: 0.049
[100, 1560] loss: 0.048
[100, 1590] loss: 0.047
[100, 1620] loss: 0.084
[100, 1650] loss: 0.066
[100, 1680] loss: 0.077
[100, 1710] loss: 0.091
[100, 1740] loss: 0.093
[100, 1770] loss: 0.074
[100, 1800] loss: 0.092
[100, 1830] loss: 0.037
[100, 1860] loss: 0.061
[100, 1890] loss: 0.069
[100, 1920] loss: 0.050
[100, 1950] loss: 0.084
[100, 1980] loss: 0.038
[100, 2010] loss: 0.036
[100, 2040] loss: 0.076
[100, 2070] loss: 0.063
[100, 2100] loss: 0.039
[100, 2130] loss: 0.028
[100, 2160] loss: 0.042
[100, 2190] loss: 0.079
eval intermediate_models/epoch99_scratch.pt saved.

evaluating on training set...
Accuracy on test images: 0.986596
epoch 99 accuracy on train set is: 0.9865957833515085
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	NaN	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.994184

epoch 99 accuracy on test set is: 0.9941839331152308

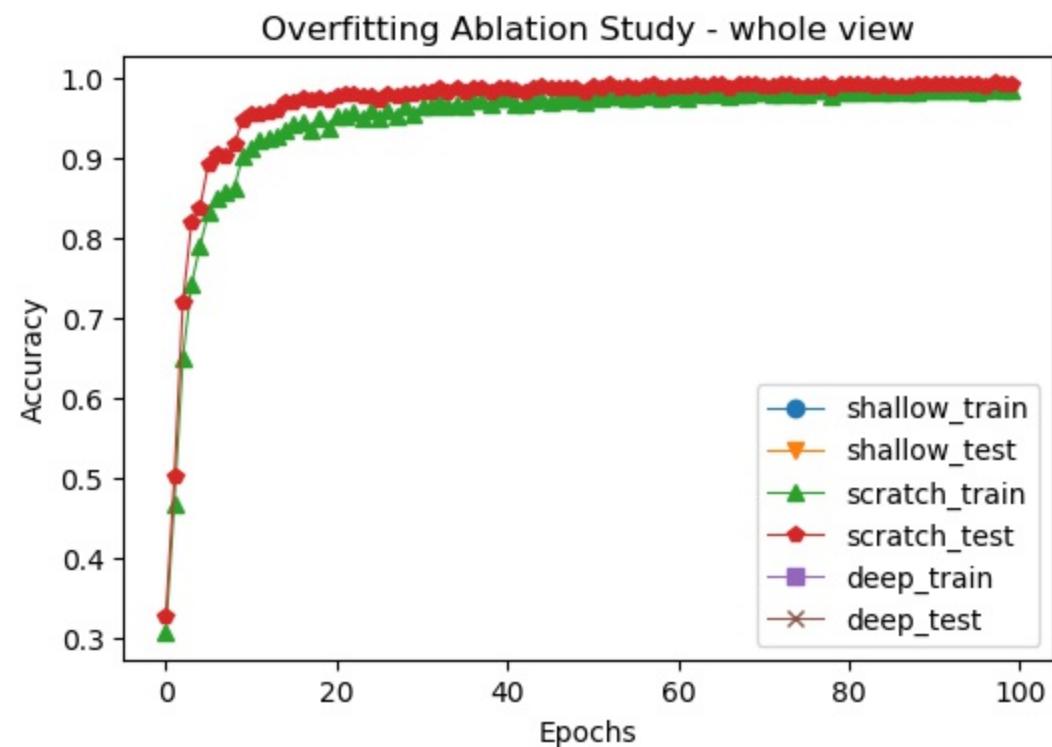
evaluating on test set takes 16 seconds.

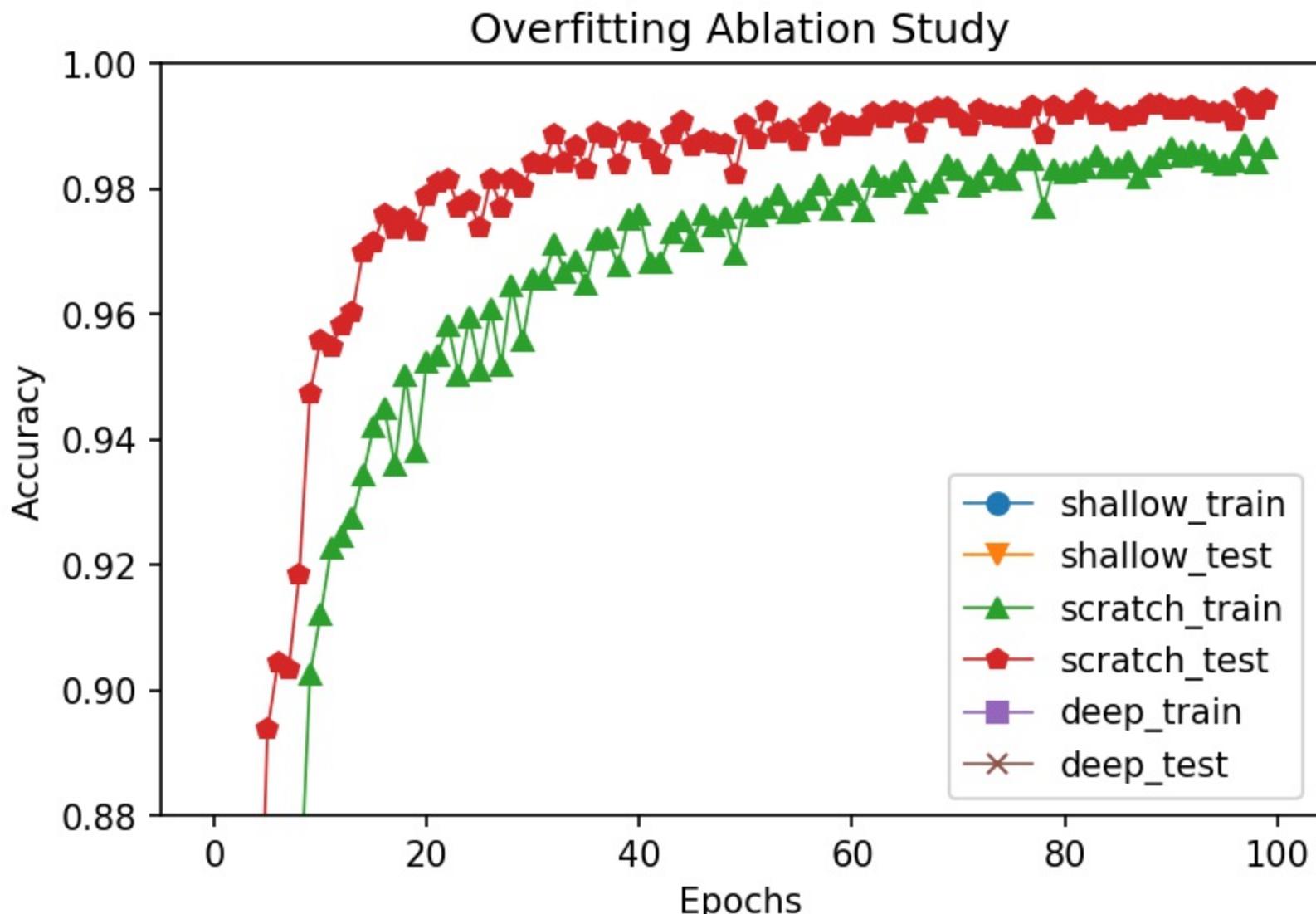
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	NaN	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 99 takes 315 seconds.

[101, 30] loss: 0.062

[101, 60] loss: 0.083

```
[101,    90] loss: 0.073
[101,   120] loss: 0.039
[101,   150] loss: 0.077
[101,   180] loss: 0.098
[101,   210] loss: 0.047
[101,   240] loss: 0.062
[101,   270] loss: 0.057
[101,   300] loss: 0.071
[101,   330] loss: 0.065
[101,   360] loss: 0.058
[101,   390] loss: 0.069
[101,   420] loss: 0.039
[101,   450] loss: 0.113
[101,   480] loss: 0.066
[101,   510] loss: 0.039
[101,   540] loss: 0.070
[101,   570] loss: 0.070
[101,   600] loss: 0.063
[101,   630] loss: 0.042
[101,   660] loss: 0.052
[101,   690] loss: 0.062
[101,   720] loss: 0.071
[101,   750] loss: 0.076
[101,   780] loss: 0.048
[101,   810] loss: 0.054
[101,   840] loss: 0.051
[101,   870] loss: 0.055
[101,   900] loss: 0.041
[101,   930] loss: 0.068
[101,   960] loss: 0.074
[101,   990] loss: 0.065
[101,  1020] loss: 0.052
[101,  1050] loss: 0.054
[101,  1080] loss: 0.050
[101,  1110] loss: 0.053
[101,  1140] loss: 0.084
[101,  1170] loss: 0.045
[101,  1200] loss: 0.049
[101,  1230] loss: 0.039
```

```
[101, 1260] loss: 0.050
[101, 1290] loss: 0.053
[101, 1320] loss: 0.052
[101, 1350] loss: 0.056
[101, 1380] loss: 0.045
[101, 1410] loss: 0.068
[101, 1440] loss: 0.055
[101, 1470] loss: 0.046
[101, 1500] loss: 0.078
[101, 1530] loss: 0.062
[101, 1560] loss: 0.108
[101, 1590] loss: 0.077
[101, 1620] loss: 0.079
[101, 1650] loss: 0.074
[101, 1680] loss: 0.065
[101, 1710] loss: 0.027
[101, 1740] loss: 0.071
[101, 1770] loss: 0.045
[101, 1800] loss: 0.034
[101, 1830] loss: 0.070
[101, 1860] loss: 0.066
[101, 1890] loss: 0.075
[101, 1920] loss: 0.037
[101, 1950] loss: 0.056
[101, 1980] loss: 0.078
[101, 2010] loss: 0.048
[101, 2040] loss: 0.062
[101, 2070] loss: 0.063
[101, 2100] loss: 0.081
[101, 2130] loss: 0.067
[101, 2160] loss: 0.070
[101, 2190] loss: 0.054
eval intermediate_models/epoch100_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.985414
epoch 100 accuracy on train set is: 0.9854143947655398
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	NaN	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990731

epoch 100 accuracy on test set is: 0.9907306434023991

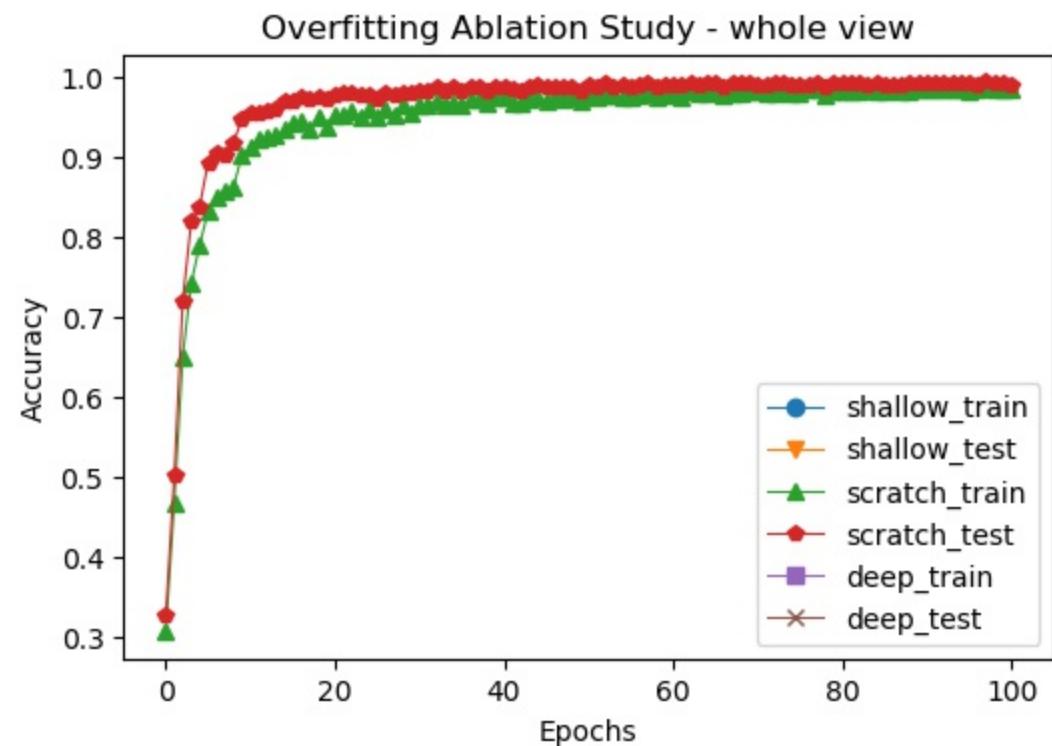
evaluating on test set takes 16 seconds.

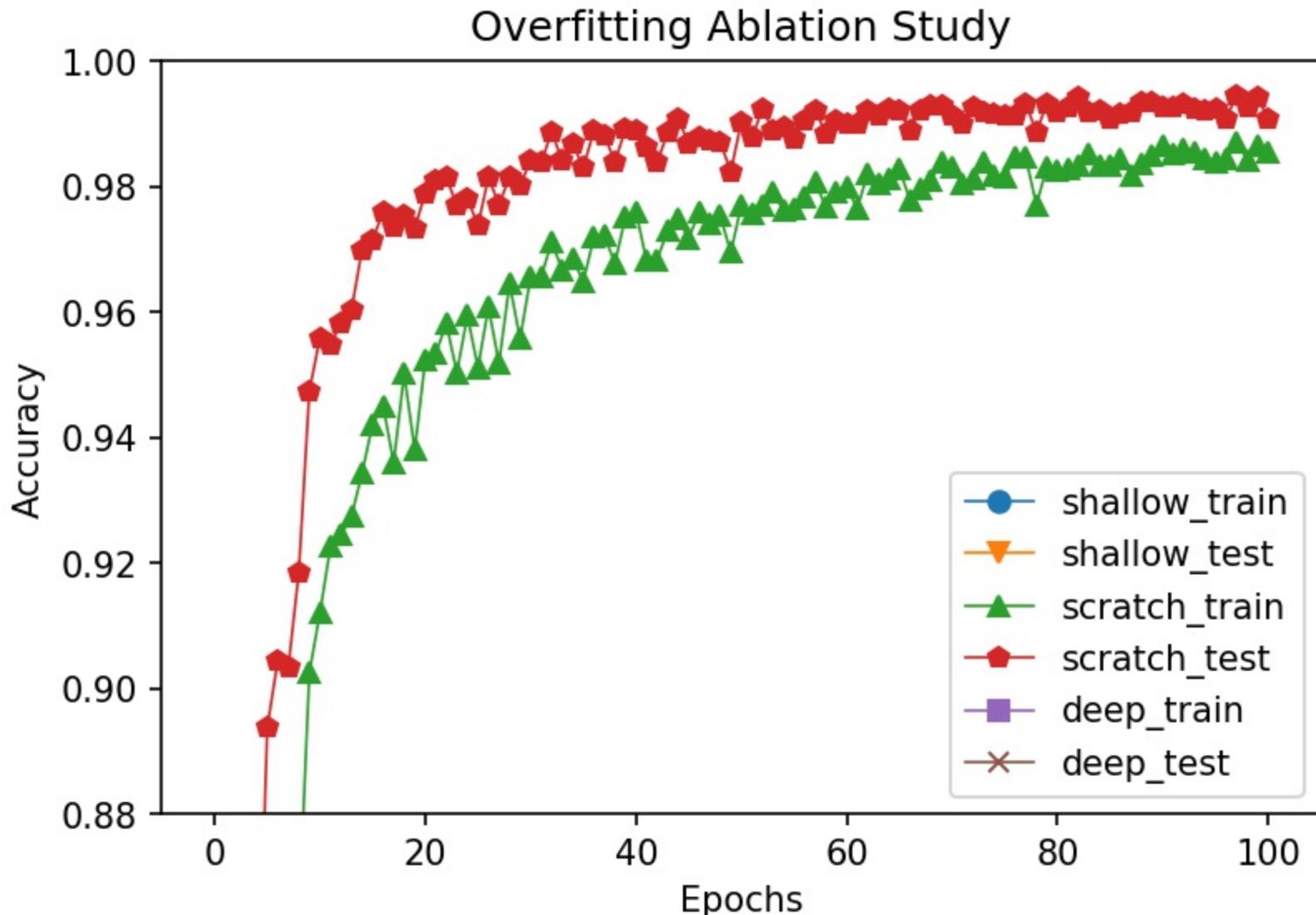
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	NaN	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 100 takes 315 seconds.  
[102,    30] loss: 0.082  
[102,    60] loss: 0.059
```

1014

```
[102,    90] loss: 0.104
[102,   120] loss: 0.067
[102,   150] loss: 0.084
[102,   180] loss: 0.030
[102,   210] loss: 0.044
[102,   240] loss: 0.053
[102,   270] loss: 0.070
[102,   300] loss: 0.068
[102,   330] loss: 0.069
[102,   360] loss: 0.048
[102,   390] loss: 0.086
[102,   420] loss: 0.051
[102,   450] loss: 0.040
[102,   480] loss: 0.043
[102,   510] loss: 0.034
[102,   540] loss: 0.063
[102,   570] loss: 0.075
[102,   600] loss: 0.085
[102,   630] loss: 0.045
[102,   660] loss: 0.044
[102,   690] loss: 0.050
[102,   720] loss: 0.051
[102,   750] loss: 0.046
[102,   780] loss: 0.060
[102,   810] loss: 0.096
[102,   840] loss: 0.061
[102,   870] loss: 0.064
[102,   900] loss: 0.059
[102,   930] loss: 0.072
[102,   960] loss: 0.053
[102,   990] loss: 0.048
[102,  1020] loss: 0.061
[102,  1050] loss: 0.075
[102,  1080] loss: 0.061
[102,  1110] loss: 0.062
[102,  1140] loss: 0.073
[102,  1170] loss: 0.038
[102,  1200] loss: 0.034
[102,  1230] loss: 0.071
```

```
[102, 1260] loss: 0.042
[102, 1290] loss: 0.066
[102, 1320] loss: 0.058
[102, 1350] loss: 0.067
[102, 1380] loss: 0.044
[102, 1410] loss: 0.057
[102, 1440] loss: 0.058
[102, 1470] loss: 0.080
[102, 1500] loss: 0.021
[102, 1530] loss: 0.047
[102, 1560] loss: 0.057
[102, 1590] loss: 0.079
[102, 1620] loss: 0.058
[102, 1650] loss: 0.066
[102, 1680] loss: 0.061
[102, 1710] loss: 0.072
[102, 1740] loss: 0.039
[102, 1770] loss: 0.039
[102, 1800] loss: 0.063
[102, 1830] loss: 0.052
[102, 1860] loss: 0.096
[102, 1890] loss: 0.048
[102, 1920] loss: 0.048
[102, 1950] loss: 0.065
[102, 1980] loss: 0.067
[102, 2010] loss: 0.047
[102, 2040] loss: 0.069
[102, 2070] loss: 0.050
[102, 2100] loss: 0.062
[102, 2130] loss: 0.075
[102, 2160] loss: 0.082
[102, 2190] loss: 0.049
eval intermediate_models/epoch101_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.985414
epoch 101 accuracy on train set is: 0.9854143947655398
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	NaN	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993366

epoch 101 accuracy on test set is: 0.9933660487095601

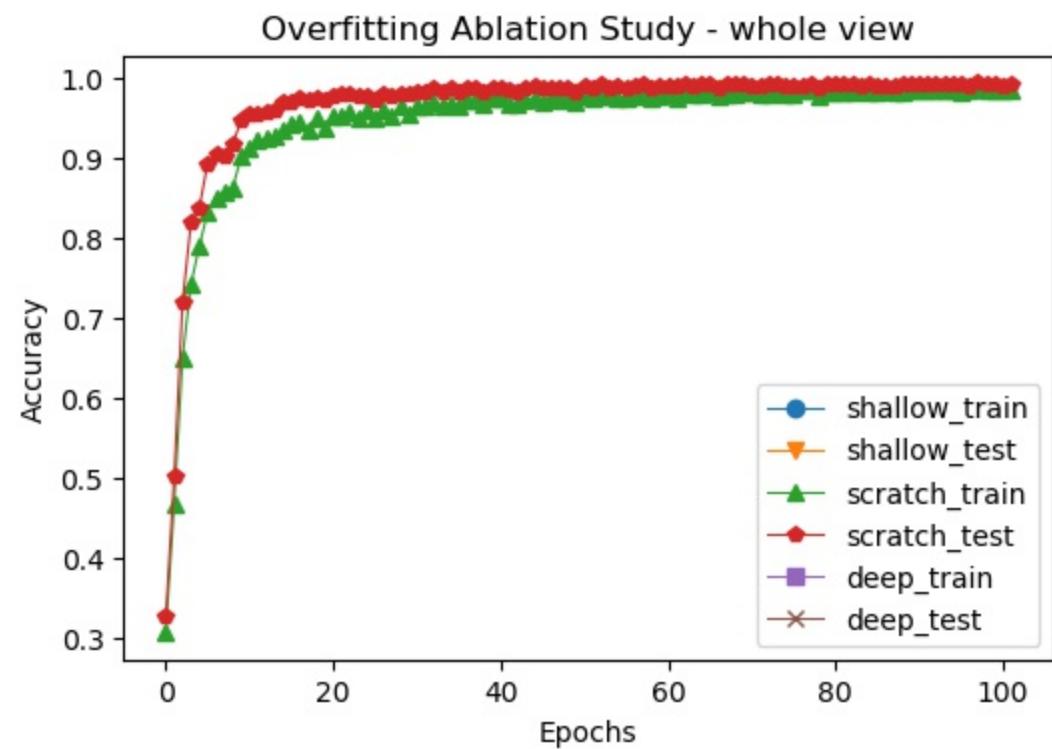
evaluating on test set takes 16 seconds.

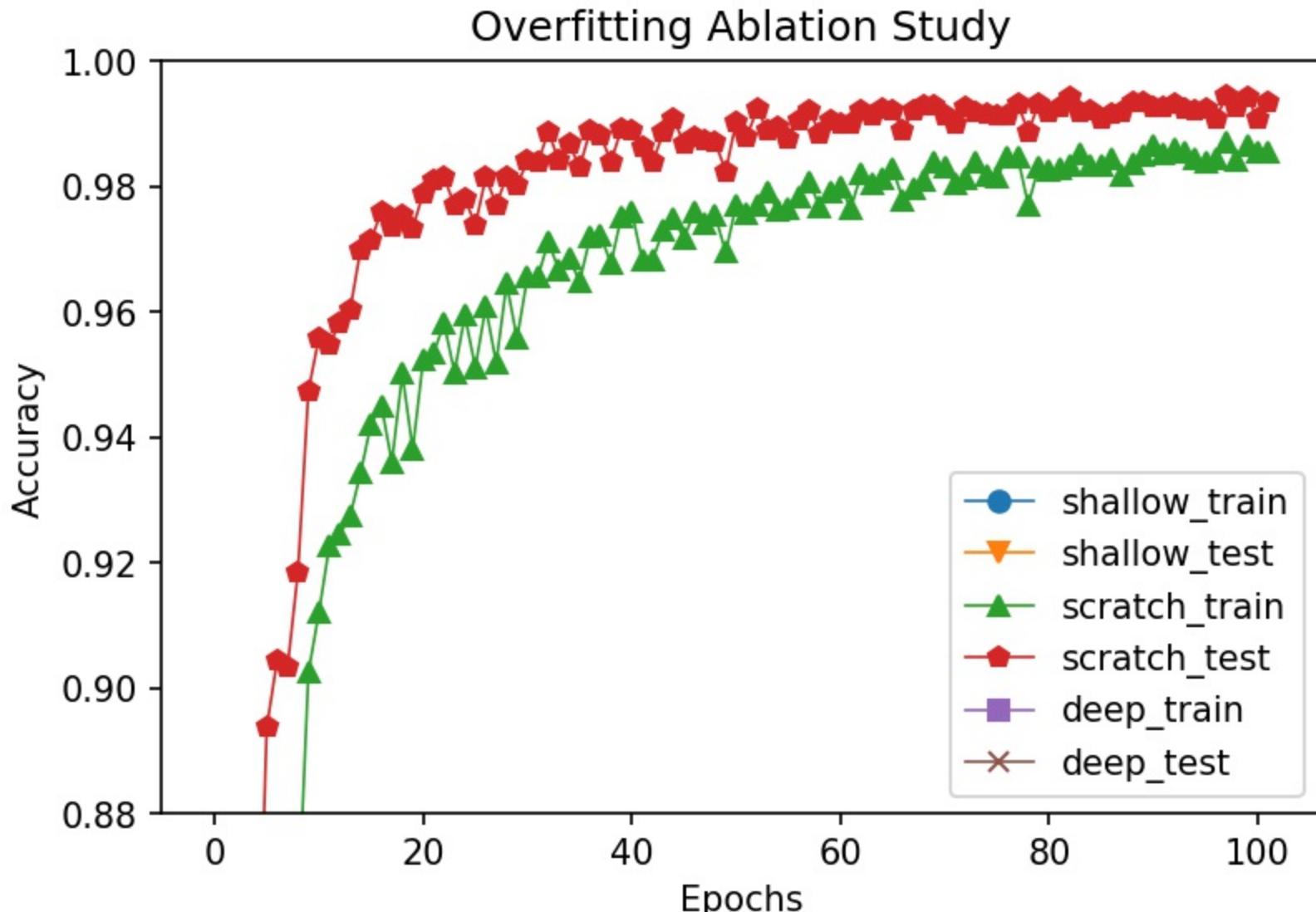
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	NaN	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 101 takes 315 seconds.

[103, 30] loss: 0.080

[103, 60] loss: 0.071

```
[103,    90] loss: 0.030
[103,   120] loss: 0.068
[103,   150] loss: 0.032
[103,   180] loss: 0.074
[103,   210] loss: 0.056
[103,   240] loss: 0.031
[103,   270] loss: 0.075
[103,   300] loss: 0.067
[103,   330] loss: 0.061
[103,   360] loss: 0.052
[103,   390] loss: 0.101
[103,   420] loss: 0.059
[103,   450] loss: 0.071
[103,   480] loss: 0.059
[103,   510] loss: 0.067
[103,   540] loss: 0.051
[103,   570] loss: 0.068
[103,   600] loss: 0.040
[103,   630] loss: 0.072
[103,   660] loss: 0.059
[103,   690] loss: 0.081
[103,   720] loss: 0.055
[103,   750] loss: 0.078
[103,   780] loss: 0.060
[103,   810] loss: 0.087
[103,   840] loss: 0.057
[103,   870] loss: 0.063
[103,   900] loss: 0.040
[103,   930] loss: 0.043
[103,   960] loss: 0.062
[103,   990] loss: 0.050
[103,  1020] loss: 0.068
[103,  1050] loss: 0.077
[103,  1080] loss: 0.057
[103,  1110] loss: 0.047
[103,  1140] loss: 0.094
[103,  1170] loss: 0.069
[103,  1200] loss: 0.056
[103,  1230] loss: 0.052
```

```
[103, 1260] loss: 0.074
[103, 1290] loss: 0.052
[103, 1320] loss: 0.055
[103, 1350] loss: 0.061
[103, 1380] loss: 0.061
[103, 1410] loss: 0.117
[103, 1440] loss: 0.066
[103, 1470] loss: 0.089
[103, 1500] loss: 0.046
[103, 1530] loss: 0.058
[103, 1560] loss: 0.050
[103, 1590] loss: 0.043
[103, 1620] loss: 0.057
[103, 1650] loss: 0.088
[103, 1680] loss: 0.054
[103, 1710] loss: 0.066
[103, 1740] loss: 0.068
[103, 1770] loss: 0.063
[103, 1800] loss: 0.106
[103, 1830] loss: 0.067
[103, 1860] loss: 0.048
[103, 1890] loss: 0.046
[103, 1920] loss: 0.057
[103, 1950] loss: 0.052
[103, 1980] loss: 0.097
[103, 2010] loss: 0.050
[103, 2040] loss: 0.061
[103, 2070] loss: 0.077
[103, 2100] loss: 0.058
[103, 2130] loss: 0.068
[103, 2160] loss: 0.070
[103, 2190] loss: 0.089
eval intermediate_models/epoch102_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.983415
epoch 102 accuracy on train set is: 0.9834151217739004
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	NaN	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992639

epoch 102 accuracy on test set is: 0.992639040348964

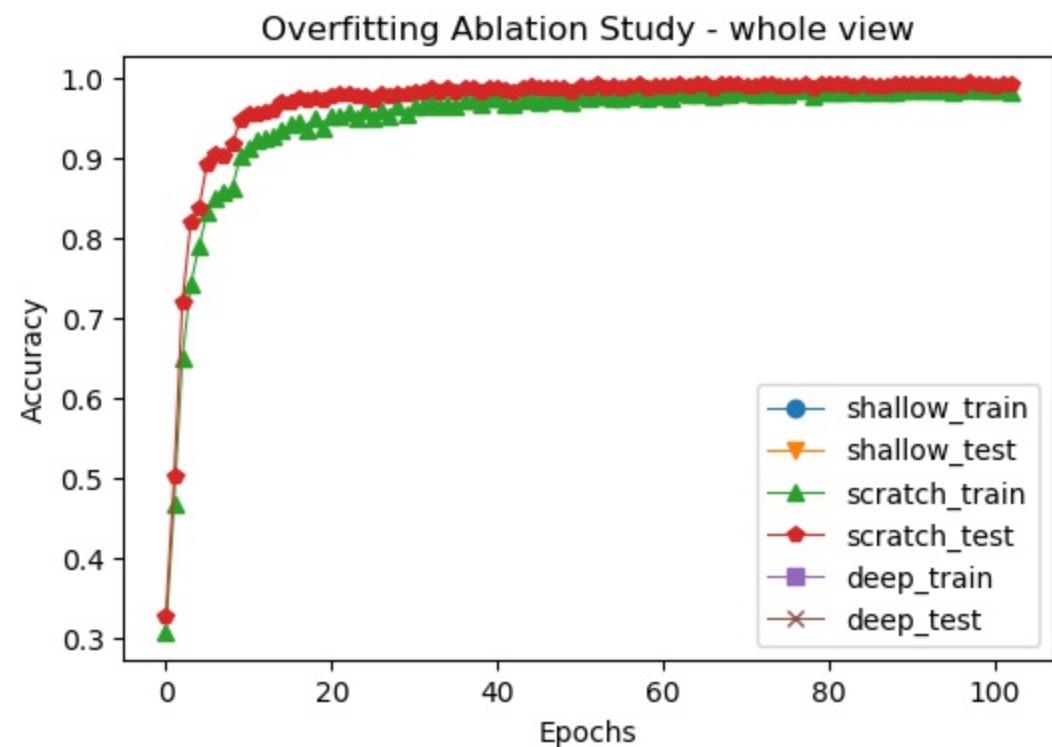
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

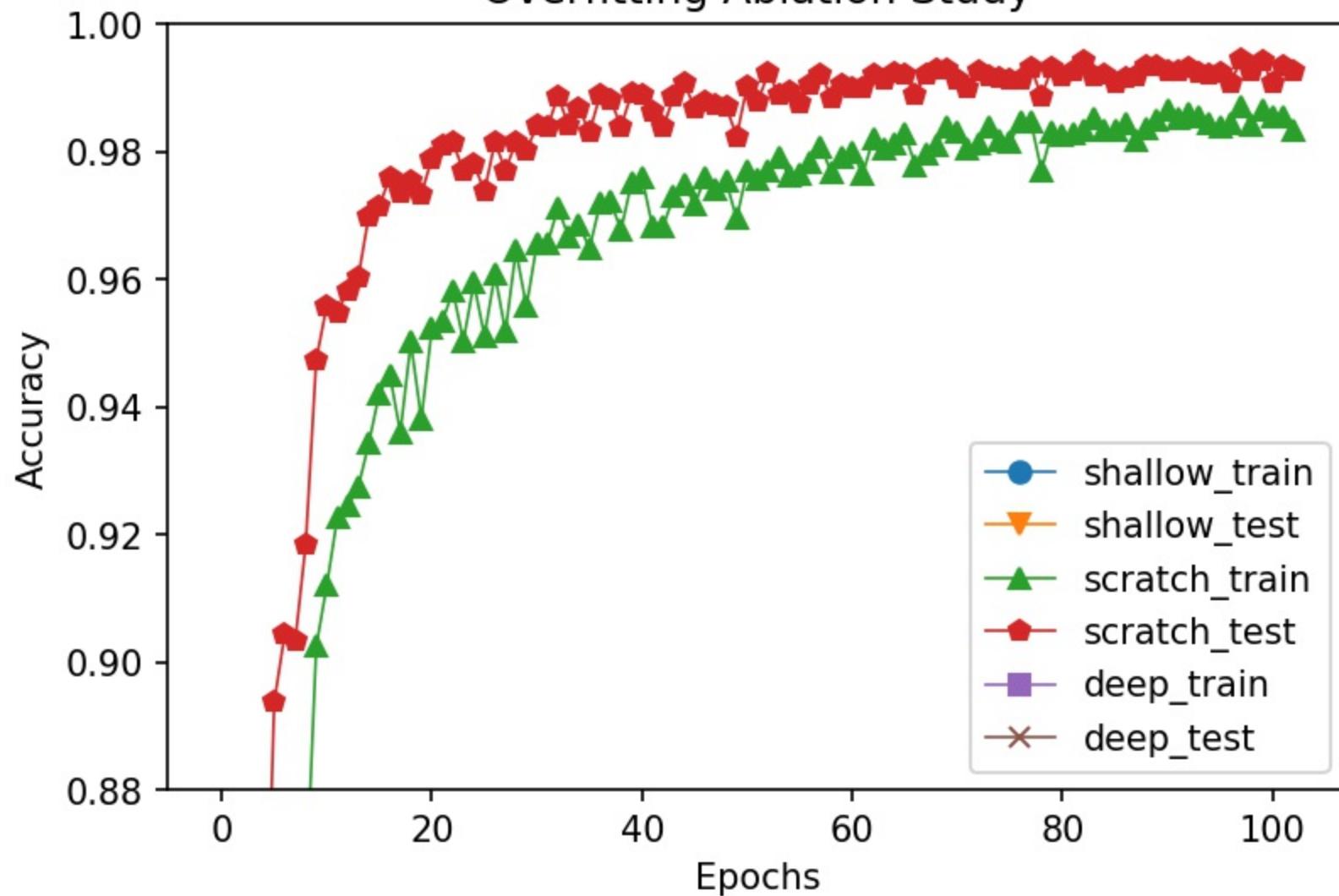
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	NaN	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 102 takes 315 seconds.

[104, 30] loss: 0.051

[104, 60] loss: 0.039

```
[104,    90] loss: 0.058
[104,   120] loss: 0.034
[104,   150] loss: 0.060
[104,   180] loss: 0.057
[104,   210] loss: 0.048
[104,   240] loss: 0.031
[104,   270] loss: 0.051
[104,   300] loss: 0.064
[104,   330] loss: 0.074
[104,   360] loss: 0.095
[104,   390] loss: 0.059
[104,   420] loss: 0.057
[104,   450] loss: 0.047
[104,   480] loss: 0.051
[104,   510] loss: 0.035
[104,   540] loss: 0.045
[104,   570] loss: 0.034
[104,   600] loss: 0.060
[104,   630] loss: 0.050
[104,   660] loss: 0.059
[104,   690] loss: 0.053
[104,   720] loss: 0.045
[104,   750] loss: 0.057
[104,   780] loss: 0.062
[104,   810] loss: 0.045
[104,   840] loss: 0.064
[104,   870] loss: 0.061
[104,   900] loss: 0.061
[104,   930] loss: 0.055
[104,   960] loss: 0.078
[104,   990] loss: 0.056
[104,  1020] loss: 0.045
[104,  1050] loss: 0.043
[104,  1080] loss: 0.073
[104,  1110] loss: 0.057
[104,  1140] loss: 0.089
[104,  1170] loss: 0.055
[104,  1200] loss: 0.065
[104,  1230] loss: 0.041
```

```
[104, 1260] loss: 0.068
[104, 1290] loss: 0.077
[104, 1320] loss: 0.076
[104, 1350] loss: 0.068
[104, 1380] loss: 0.030
[104, 1410] loss: 0.069
[104, 1440] loss: 0.072
[104, 1470] loss: 0.032
[104, 1500] loss: 0.069
[104, 1530] loss: 0.050
[104, 1560] loss: 0.038
[104, 1590] loss: 0.051
[104, 1620] loss: 0.049
[104, 1650] loss: 0.048
[104, 1680] loss: 0.050
[104, 1710] loss: 0.065
[104, 1740] loss: 0.049
[104, 1770] loss: 0.093
[104, 1800] loss: 0.091
[104, 1830] loss: 0.054
[104, 1860] loss: 0.054
[104, 1890] loss: 0.112
[104, 1920] loss: 0.065
[104, 1950] loss: 0.071
[104, 1980] loss: 0.061
[104, 2010] loss: 0.066
[104, 2040] loss: 0.036
[104, 2070] loss: 0.061
[104, 2100] loss: 0.066
[104, 2130] loss: 0.047
[104, 2160] loss: 0.039
[104, 2190] loss: 0.062
evFit intermediate_models/epoch103_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.985982
epoch 103 accuracy on train set is: 0.9859823700472555
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	NaN	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992003

epoch 103 accuracy on test set is: 0.9920029080334424

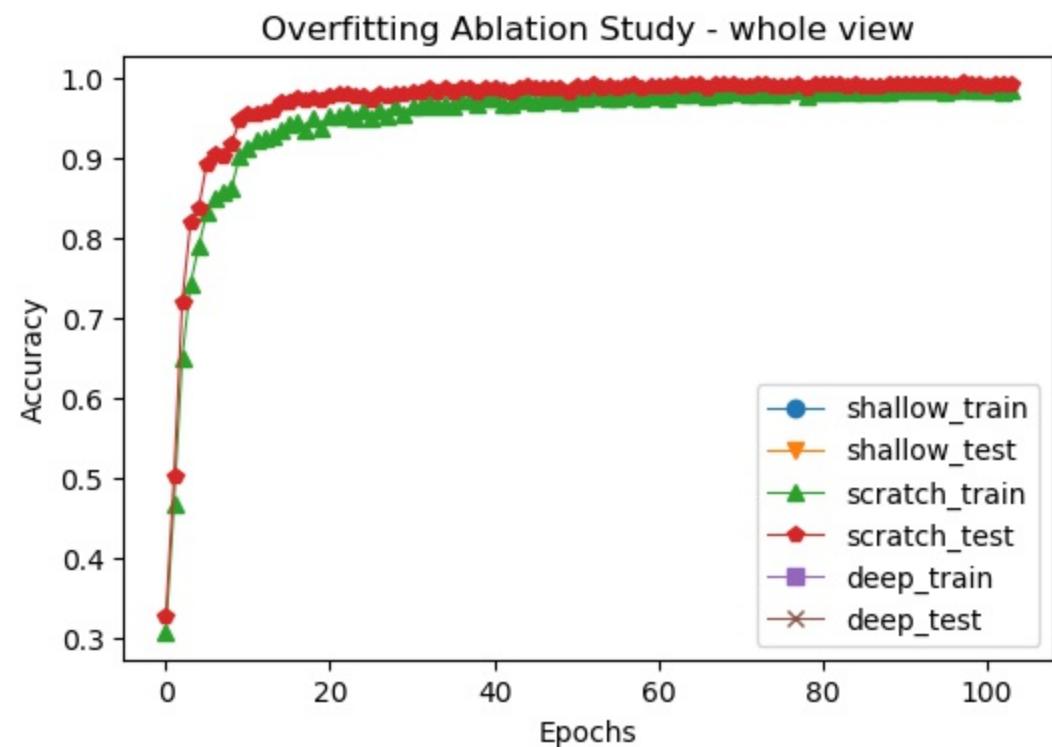
evaluating on test set takes 16 seconds.

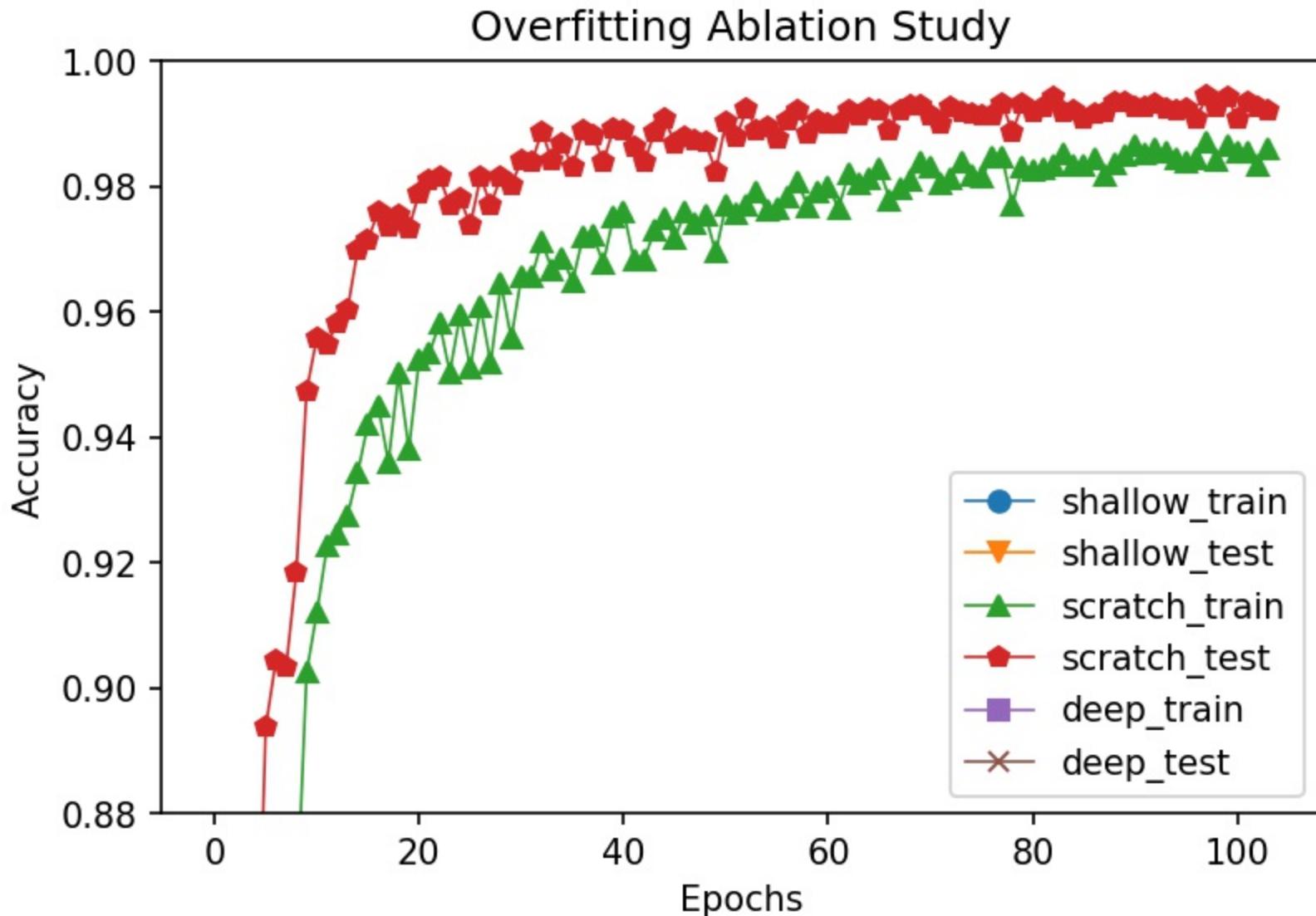
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	NaN	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 103 takes 315 seconds.  
[105,     30] loss: 0.053  
[105,     60] loss: 0.039
```

```
[105,    90] loss: 0.041
[105,   120] loss: 0.053
[105,   150] loss: 0.075
[105,   180] loss: 0.068
[105,   210] loss: 0.075
[105,   240] loss: 0.065
[105,   270] loss: 0.070
[105,   300] loss: 0.074
[105,   330] loss: 0.055
[105,   360] loss: 0.045
[105,   390] loss: 0.042
[105,   420] loss: 0.076
[105,   450] loss: 0.064
[105,   480] loss: 0.037
[105,   510] loss: 0.046
[105,   540] loss: 0.061
[105,   570] loss: 0.076
[105,   600] loss: 0.046
[105,   630] loss: 0.058
[105,   660] loss: 0.068
[105,   690] loss: 0.053
[105,   720] loss: 0.035
[105,   750] loss: 0.075
[105,   780] loss: 0.046
[105,   810] loss: 0.060
[105,   840] loss: 0.055
[105,   870] loss: 0.051
[105,   900] loss: 0.036
[105,   930] loss: 0.067
[105,   960] loss: 0.030
[105,   990] loss: 0.079
[105,  1020] loss: 0.061
[105,  1050] loss: 0.061
[105,  1080] loss: 0.059
[105,  1110] loss: 0.095
[105,  1140] loss: 0.061
[105,  1170] loss: 0.068
[105,  1200] loss: 0.079
[105,  1230] loss: 0.054
```

```
[105, 1260] loss: 0.051
[105, 1290] loss: 0.075
[105, 1320] loss: 0.067
[105, 1350] loss: 0.080
[105, 1380] loss: 0.073
[105, 1410] loss: 0.049
[105, 1440] loss: 0.050
[105, 1470] loss: 0.050
[105, 1500] loss: 0.065
[105, 1530] loss: 0.083
[105, 1560] loss: 0.058
[105, 1590] loss: 0.043
[105, 1620] loss: 0.045
[105, 1650] loss: 0.087
[105, 1680] loss: 0.030
[105, 1710] loss: 0.048
[105, 1740] loss: 0.060
[105, 1770] loss: 0.046
[105, 1800] loss: 0.032
[105, 1830] loss: 0.043
[105, 1860] loss: 0.045
[105, 1890] loss: 0.052
[105, 1920] loss: 0.073
[105, 1950] loss: 0.061
[105, 1980] loss: 0.071
[105, 2010] loss: 0.084
[105, 2040] loss: 0.065
[105, 2070] loss: 0.086
[105, 2100] loss: 0.075
[105, 2130] loss: 0.074
[105, 2160] loss: 0.061
[105, 2190] loss: 0.058
eval intermediate_models/epoch104_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.982120
epoch 104 accuracy on train set is: 0.9821201381315885
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	NaN	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.988186

epoch 104 accuracy on test set is: 0.9881861141403127

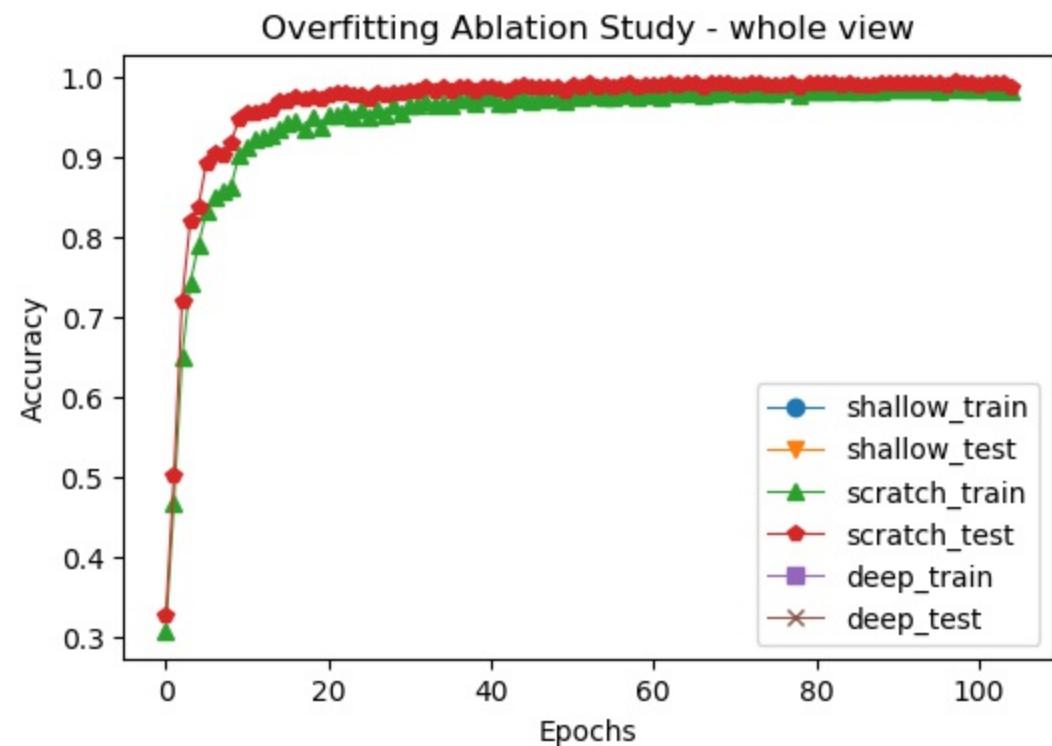
evaluating on test set takes 16 seconds.

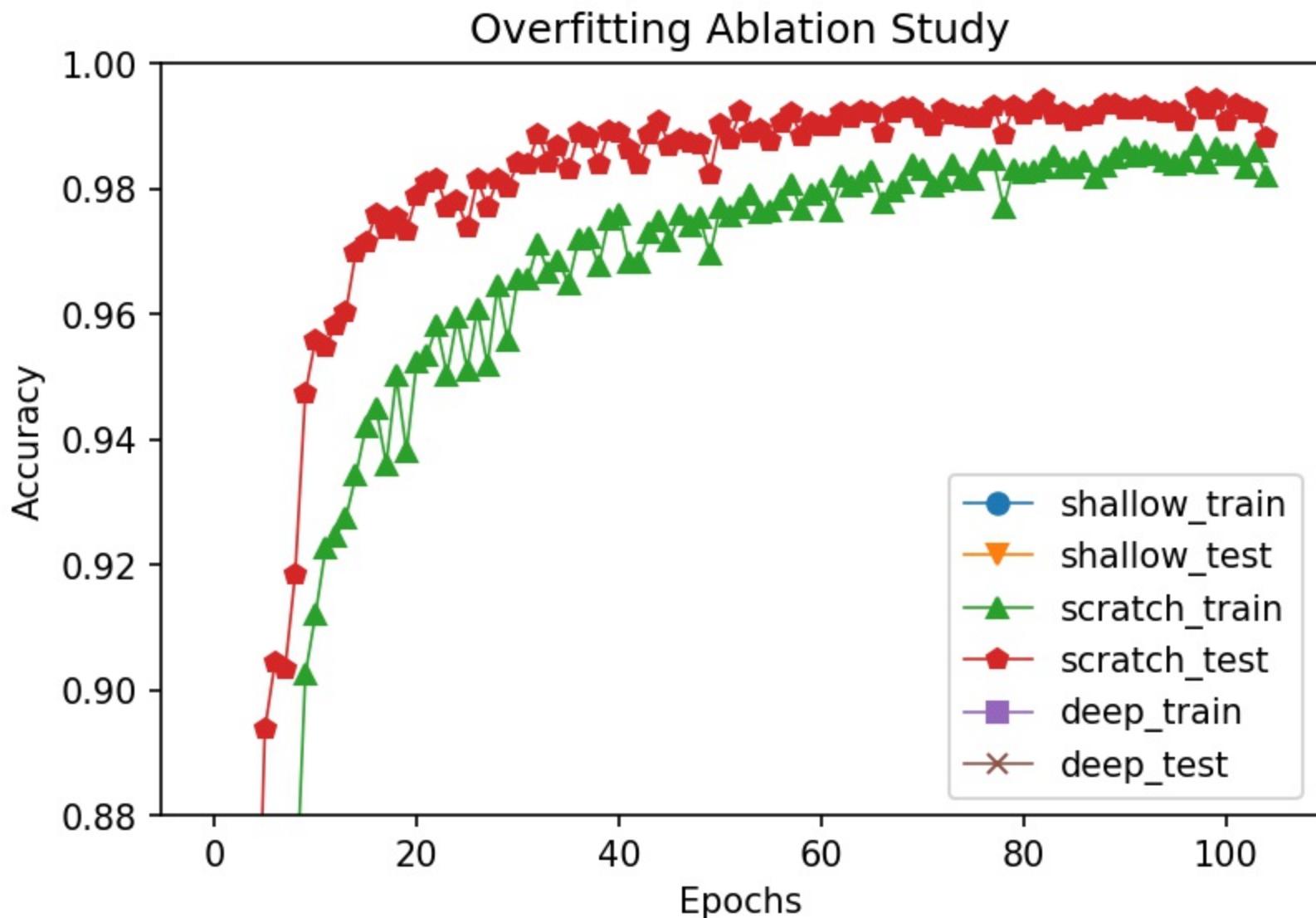
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	NaN	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 104 takes 315 seconds.

[106, 30] loss: 0.058

[106, 60] loss: 0.083

```
[106,    90] loss: 0.057
[106,   120] loss: 0.070
[106,   150] loss: 0.066
[106,   180] loss: 0.064
[106,   210] loss: 0.081
[106,   240] loss: 0.119
[106,   270] loss: 0.077
[106,   300] loss: 0.062
[106,   330] loss: 0.049
[106,   360] loss: 0.047
[106,   390] loss: 0.069
[106,   420] loss: 0.035
[106,   450] loss: 0.059
[106,   480] loss: 0.062
[106,   510] loss: 0.054
[106,   540] loss: 0.080
[106,   570] loss: 0.080
[106,   600] loss: 0.030
[106,   630] loss: 0.055
[106,   660] loss: 0.037
[106,   690] loss: 0.032
[106,   720] loss: 0.049
[106,   750] loss: 0.057
[106,   780] loss: 0.037
[106,   810] loss: 0.048
[106,   840] loss: 0.040
[106,   870] loss: 0.065
[106,   900] loss: 0.065
[106,   930] loss: 0.094
[106,   960] loss: 0.076
[106,   990] loss: 0.061
[106,  1020] loss: 0.052
[106,  1050] loss: 0.077
[106,  1080] loss: 0.060
[106,  1110] loss: 0.077
[106,  1140] loss: 0.062
[106,  1170] loss: 0.053
[106,  1200] loss: 0.030
[106,  1230] loss: 0.047
```

```
[106, 1260] loss: 0.074
[106, 1290] loss: 0.047
[106, 1320] loss: 0.042
[106, 1350] loss: 0.060
[106, 1380] loss: 0.077
[106, 1410] loss: 0.063
[106, 1440] loss: 0.042
[106, 1470] loss: 0.057
[106, 1500] loss: 0.038
[106, 1530] loss: 0.057
[106, 1560] loss: 0.070
[106, 1590] loss: 0.047
[106, 1620] loss: 0.048
[106, 1650] loss: 0.026
[106, 1680] loss: 0.049
[106, 1710] loss: 0.037
[106, 1740] loss: 0.031
[106, 1770] loss: 0.062
[106, 1800] loss: 0.077
[106, 1830] loss: 0.052
[106, 1860] loss: 0.054
[106, 1890] loss: 0.057
[106, 1920] loss: 0.059
[106, 1950] loss: 0.073
[106, 1980] loss: 0.040
[106, 2010] loss: 0.044
[106, 2040] loss: 0.047
[106, 2070] loss: 0.060
[106, 2100] loss: 0.067
[106, 2130] loss: 0.049
[106, 2160] loss: 0.067
[106, 2190] loss: 0.055
eval intermediate_models/epoch105_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.987686
epoch 105 accuracy on train set is: 0.9876862958924028
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	NaN	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992366

epoch 105 accuracy on test set is: 0.9923664122137404

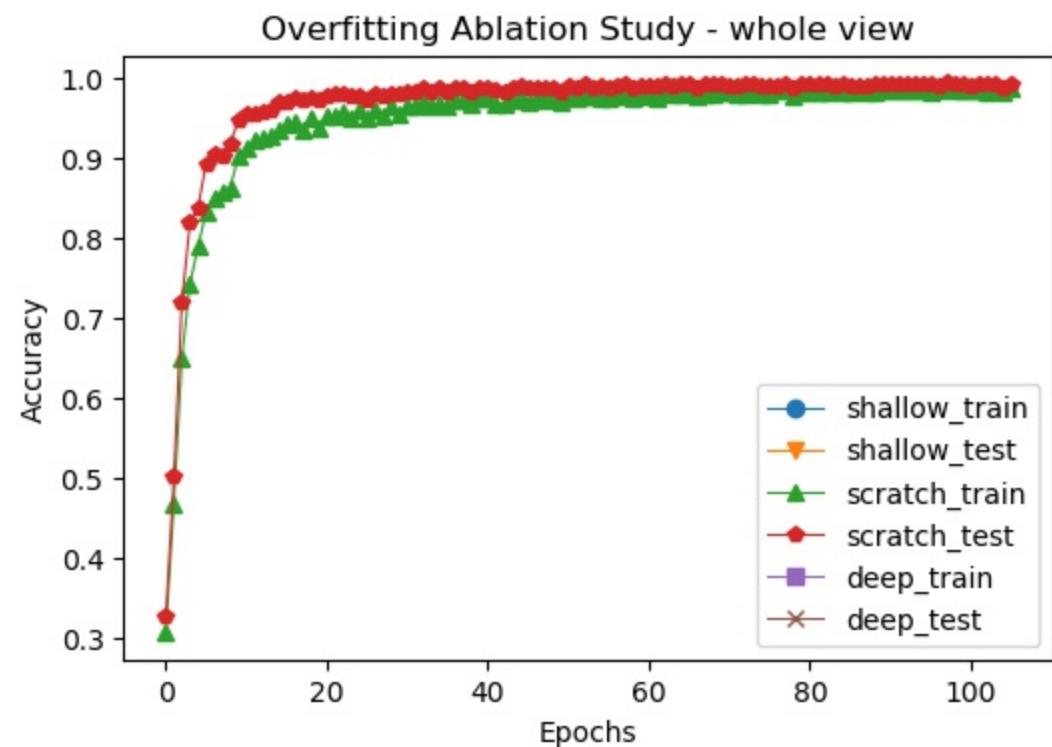
evaluating on test set takes 16 seconds.

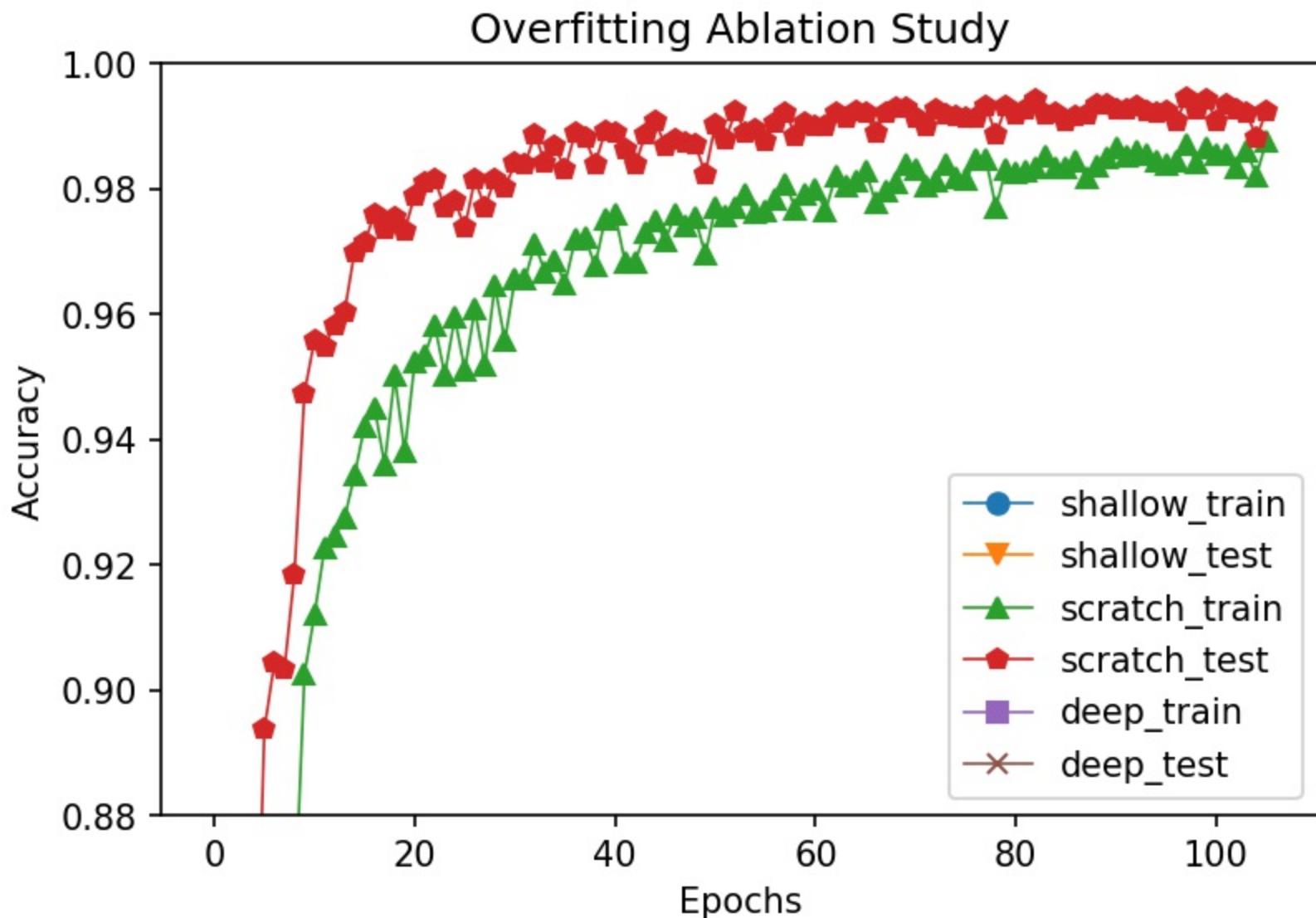
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	NaN	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 105 takes 315 seconds.  
[107,     30] loss: 0.046  
[107,     60] loss: 0.065
```

```
[107,    90] loss: 0.057
[107,   120] loss: 0.036
[107,   150] loss: 0.040
[107,   180] loss: 0.055
[107,   210] loss: 0.054
[107,   240] loss: 0.050
[107,   270] loss: 0.033
[107,   300] loss: 0.064
[107,   330] loss: 0.052
[107,   360] loss: 0.065
[107,   390] loss: 0.078
[107,   420] loss: 0.065
[107,   450] loss: 0.049
[107,   480] loss: 0.074
[107,   510] loss: 0.033
[107,   540] loss: 0.051
[107,   570] loss: 0.022
[107,   600] loss: 0.033
[107,   630] loss: 0.065
[107,   660] loss: 0.045
[107,   690] loss: 0.070
[107,   720] loss: 0.058
[107,   750] loss: 0.042
[107,   780] loss: 0.034
[107,   810] loss: 0.096
[107,   840] loss: 0.041
[107,   870] loss: 0.041
[107,   900] loss: 0.095
[107,   930] loss: 0.053
[107,   960] loss: 0.056
[107,   990] loss: 0.036
[107,  1020] loss: 0.054
[107,  1050] loss: 0.067
[107,  1080] loss: 0.035
[107,  1110] loss: 0.063
[107,  1140] loss: 0.073
[107,  1170] loss: 0.065
[107,  1200] loss: 0.050
[107,  1230] loss: 0.067
```

```
[107, 1260] loss: 0.067
[107, 1290] loss: 0.072
[107, 1320] loss: 0.050
[107, 1350] loss: 0.021
[107, 1380] loss: 0.069
[107, 1410] loss: 0.076
[107, 1440] loss: 0.096
[107, 1470] loss: 0.066
[107, 1500] loss: 0.044
[107, 1530] loss: 0.055
[107, 1560] loss: 0.064
[107, 1590] loss: 0.057
[107, 1620] loss: 0.088
[107, 1650] loss: 0.059
[107, 1680] loss: 0.047
[107, 1710] loss: 0.050
[107, 1740] loss: 0.068
[107, 1770] loss: 0.054
[107, 1800] loss: 0.081
[107, 1830] loss: 0.090
[107, 1860] loss: 0.097
[107, 1890] loss: 0.066
[107, 1920] loss: 0.081
[107, 1950] loss: 0.037
[107, 1980] loss: 0.055
[107, 2010] loss: 0.057
[107, 2040] loss: 0.037
[107, 2070] loss: 0.066
[107, 2100] loss: 0.056
[107, 2130] loss: 0.053
[107, 2160] loss: 0.048
[107, 2190] loss: 0.047
eval intermediate_models/epoch106_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.986573
epoch 106 accuracy on train set is: 0.98657306434024
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	NaN	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993730

epoch 106 accuracy on test set is: 0.9937295528898582

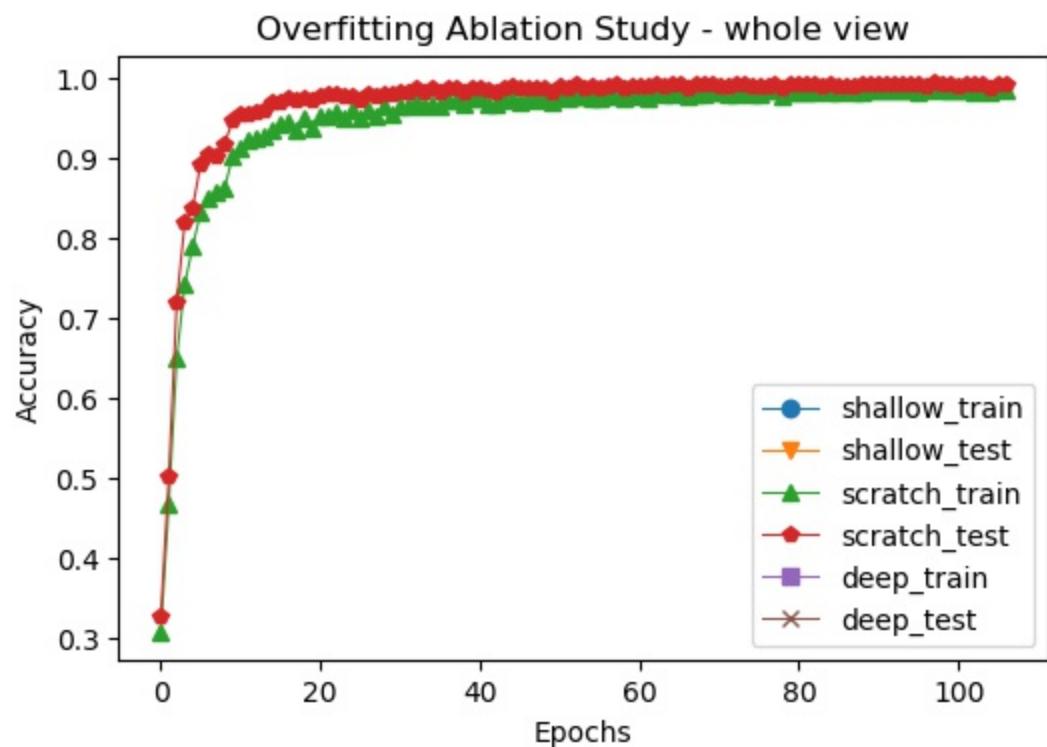
evaluating on test set takes 16 seconds.

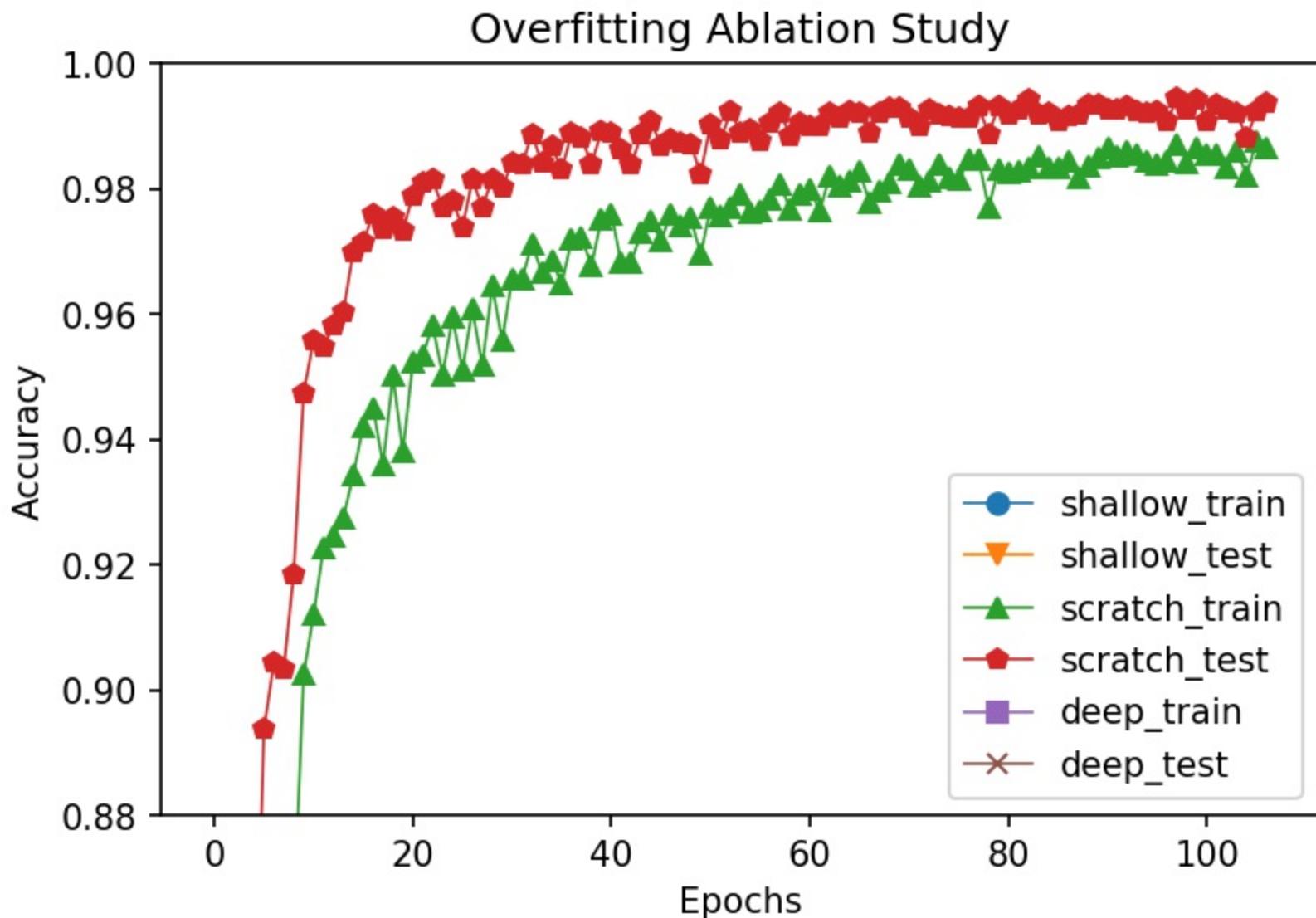
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	NaN	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 106 takes 315 seconds.  
[108,     30] loss: 0.029  
[108,     60] loss: 0.065
```

1074

```
[108,    90] loss: 0.060
[108,   120] loss: 0.055
[108,   150] loss: 0.074
[108,   180] loss: 0.066
[108,   210] loss: 0.046
[108,   240] loss: 0.052
[108,   270] loss: 0.055
[108,   300] loss: 0.086
[108,   330] loss: 0.032
[108,   360] loss: 0.045
[108,   390] loss: 0.068
[108,   420] loss: 0.048
[108,   450] loss: 0.064
[108,   480] loss: 0.080
[108,   510] loss: 0.073
[108,   540] loss: 0.058
[108,   570] loss: 0.079
[108,   600] loss: 0.053
[108,   630] loss: 0.053
[108,   660] loss: 0.053
[108,   690] loss: 0.079
[108,   720] loss: 0.027
[108,   750] loss: 0.062
[108,   780] loss: 0.041
[108,   810] loss: 0.057
[108,   840] loss: 0.050
[108,   870] loss: 0.072
[108,   900] loss: 0.046
[108,   930] loss: 0.053
[108,   960] loss: 0.061
[108,   990] loss: 0.067
[108,  1020] loss: 0.043
[108,  1050] loss: 0.062
[108,  1080] loss: 0.104
[108,  1110] loss: 0.054
[108,  1140] loss: 0.056
[108,  1170] loss: 0.074
[108,  1200] loss: 0.049
[108,  1230] loss: 0.038
```

```
[108, 1260] loss: 0.063
[108, 1290] loss: 0.067
[108, 1320] loss: 0.062
[108, 1350] loss: 0.056
[108, 1380] loss: 0.049
[108, 1410] loss: 0.058
[108, 1440] loss: 0.049
[108, 1470] loss: 0.058
[108, 1500] loss: 0.064
[108, 1530] loss: 0.029
[108, 1560] loss: 0.059
[108, 1590] loss: 0.045
[108, 1620] loss: 0.046
[108, 1650] loss: 0.064
[108, 1680] loss: 0.076
[108, 1710] loss: 0.068
[108, 1740] loss: 0.069
[108, 1770] loss: 0.065
[108, 1800] loss: 0.052
[108, 1830] loss: 0.069
[108, 1860] loss: 0.036
[108, 1890] loss: 0.054
[108, 1920] loss: 0.051
[108, 1950] loss: 0.049
[108, 1980] loss: 0.076
[108, 2010] loss: 0.055
[108, 2040] loss: 0.065
[108, 2070] loss: 0.054
[108, 2100] loss: 0.056
[108, 2130] loss: 0.062
[108, 2160] loss: 0.063
[108, 2190] loss: 0.054
eval intermediate_models/epoch107_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.987368
epoch 107 accuracy on train set is: 0.987368229734642
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	NaN	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993093

epoch 107 accuracy on test set is: 0.9930934205743366

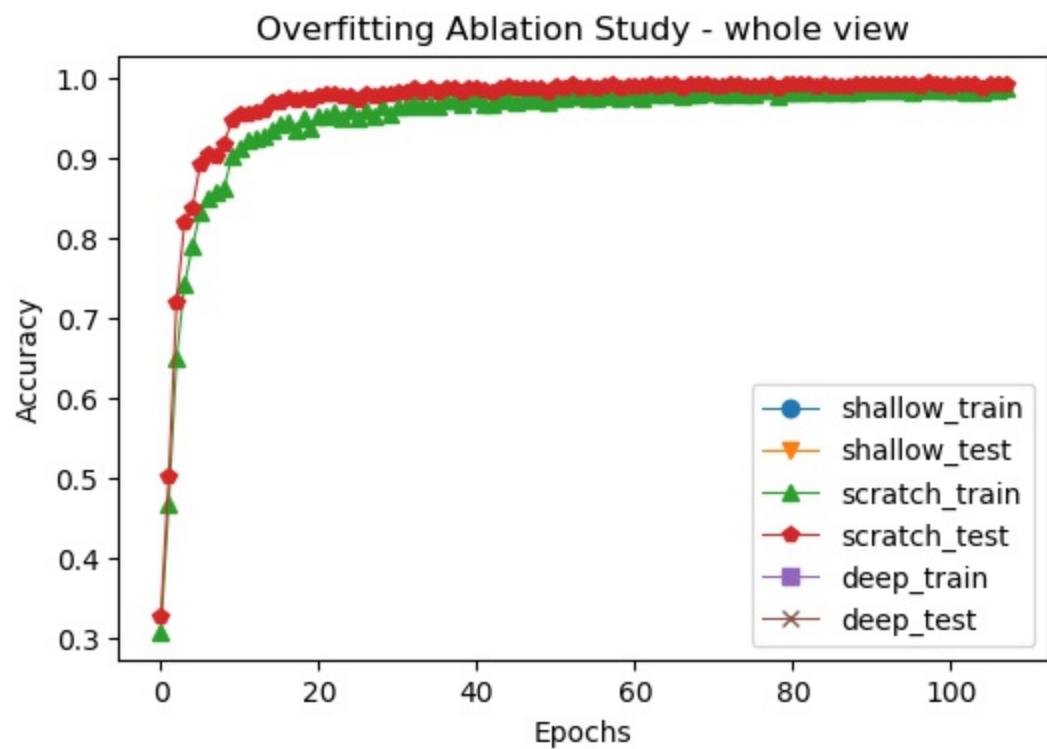
evaluating on test set takes 16 seconds.

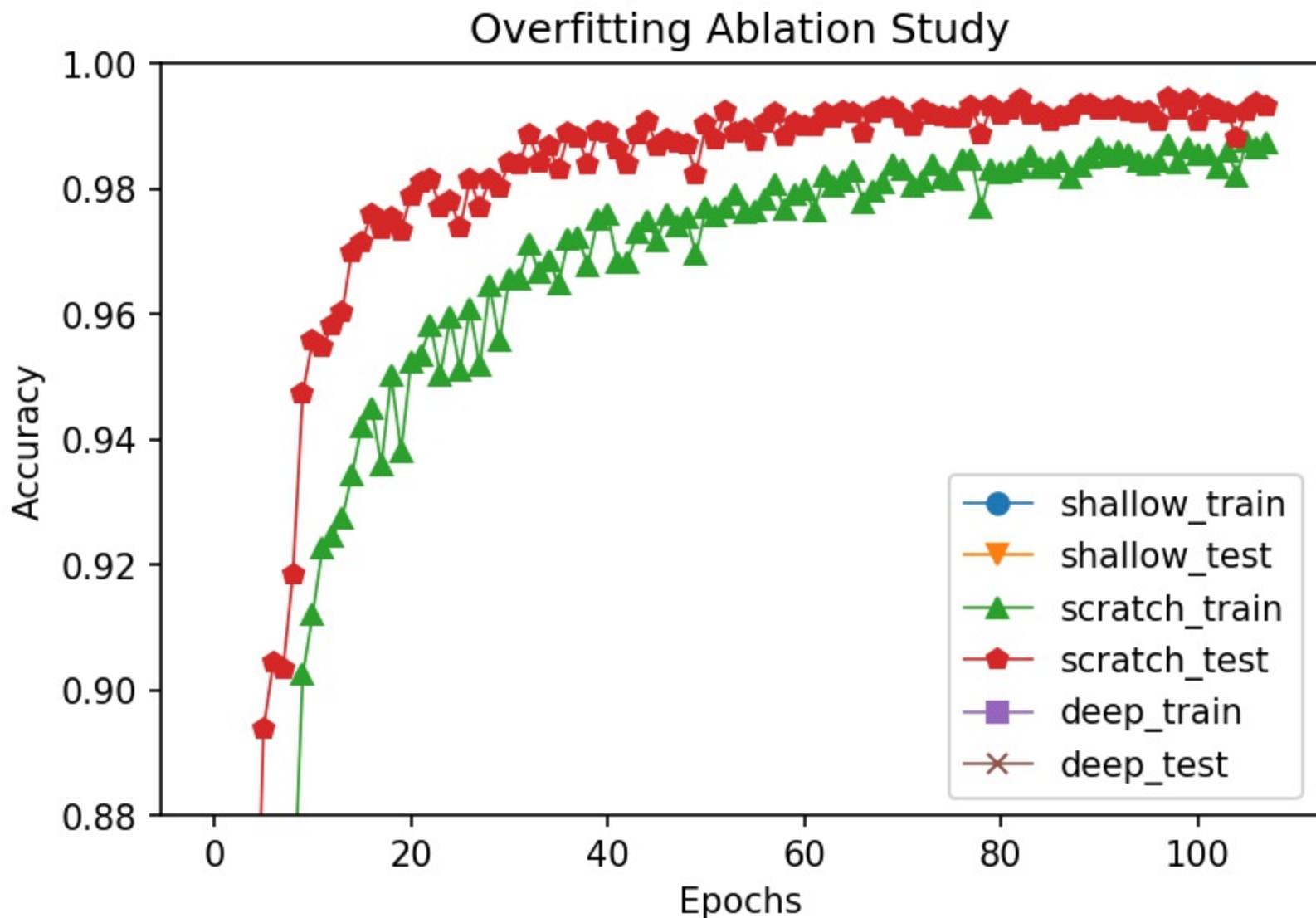
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	NaN	NaN	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 107 takes 314 seconds.  
[109,    30] loss: 0.088  
[109,    60] loss: 0.079
```

1084

```
[109,    90] loss: 0.112
[109,   120] loss: 0.053
[109,   150] loss: 0.066
[109,   180] loss: 0.044
[109,   210] loss: 0.049
[109,   240] loss: 0.045
[109,   270] loss: 0.045
[109,   300] loss: 0.086
[109,   330] loss: 0.075
[109,   360] loss: 0.034
[109,   390] loss: 0.053
[109,   420] loss: 0.078
[109,   450] loss: 0.076
[109,   480] loss: 0.063
[109,   510] loss: 0.050
[109,   540] loss: 0.037
[109,   570] loss: 0.073
[109,   600] loss: 0.039
[109,   630] loss: 0.045
[109,   660] loss: 0.078
[109,   690] loss: 0.046
[109,   720] loss: 0.061
[109,   750] loss: 0.035
[109,   780] loss: 0.079
[109,   810] loss: 0.059
[109,   840] loss: 0.063
[109,   870] loss: 0.041
[109,   900] loss: 0.054
[109,   930] loss: 0.035
[109,   960] loss: 0.060
[109,   990] loss: 0.039
[109,  1020] loss: 0.050
[109,  1050] loss: 0.047
[109,  1080] loss: 0.029
[109,  1110] loss: 0.026
[109,  1140] loss: 0.084
[109,  1170] loss: 0.076
[109,  1200] loss: 0.042
[109,  1230] loss: 0.084
```

```
[109, 1260] loss: 0.034
[109, 1290] loss: 0.050
[109, 1320] loss: 0.061
[109, 1350] loss: 0.053
[109, 1380] loss: 0.079
[109, 1410] loss: 0.072
[109, 1440] loss: 0.085
[109, 1470] loss: 0.039
[109, 1500] loss: 0.029
[109, 1530] loss: 0.044
[109, 1560] loss: 0.039
[109, 1590] loss: 0.080
[109, 1620] loss: 0.069
[109, 1650] loss: 0.034
[109, 1680] loss: 0.046
[109, 1710] loss: 0.054
[109, 1740] loss: 0.065
[109, 1770] loss: 0.026
[109, 1800] loss: 0.076
[109, 1830] loss: 0.053
[109, 1860] loss: 0.063
[109, 1890] loss: 0.057
[109, 1920] loss: 0.081
[109, 1950] loss: 0.058
[109, 1980] loss: 0.046
[109, 2010] loss: 0.065
[109, 2040] loss: 0.046
[109, 2070] loss: 0.040
[109, 2100] loss: 0.061
[109, 2130] loss: 0.067
[109, 2160] loss: 0.049
[109, 2190] loss: 0.082
eval intermediate_models/epoch108_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.985187
epoch 108 accuracy on train set is: 0.9851872046528535
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	NaN	NaN	NaN
109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992821

epoch 108 accuracy on test set is: 0.992820792439113

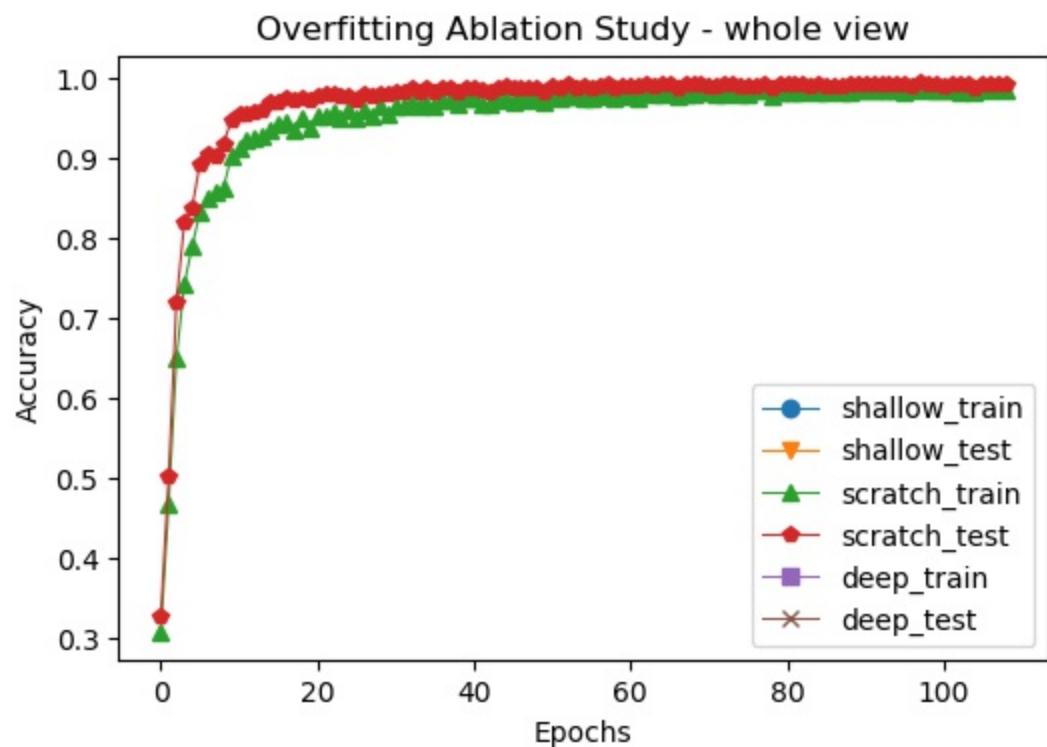
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

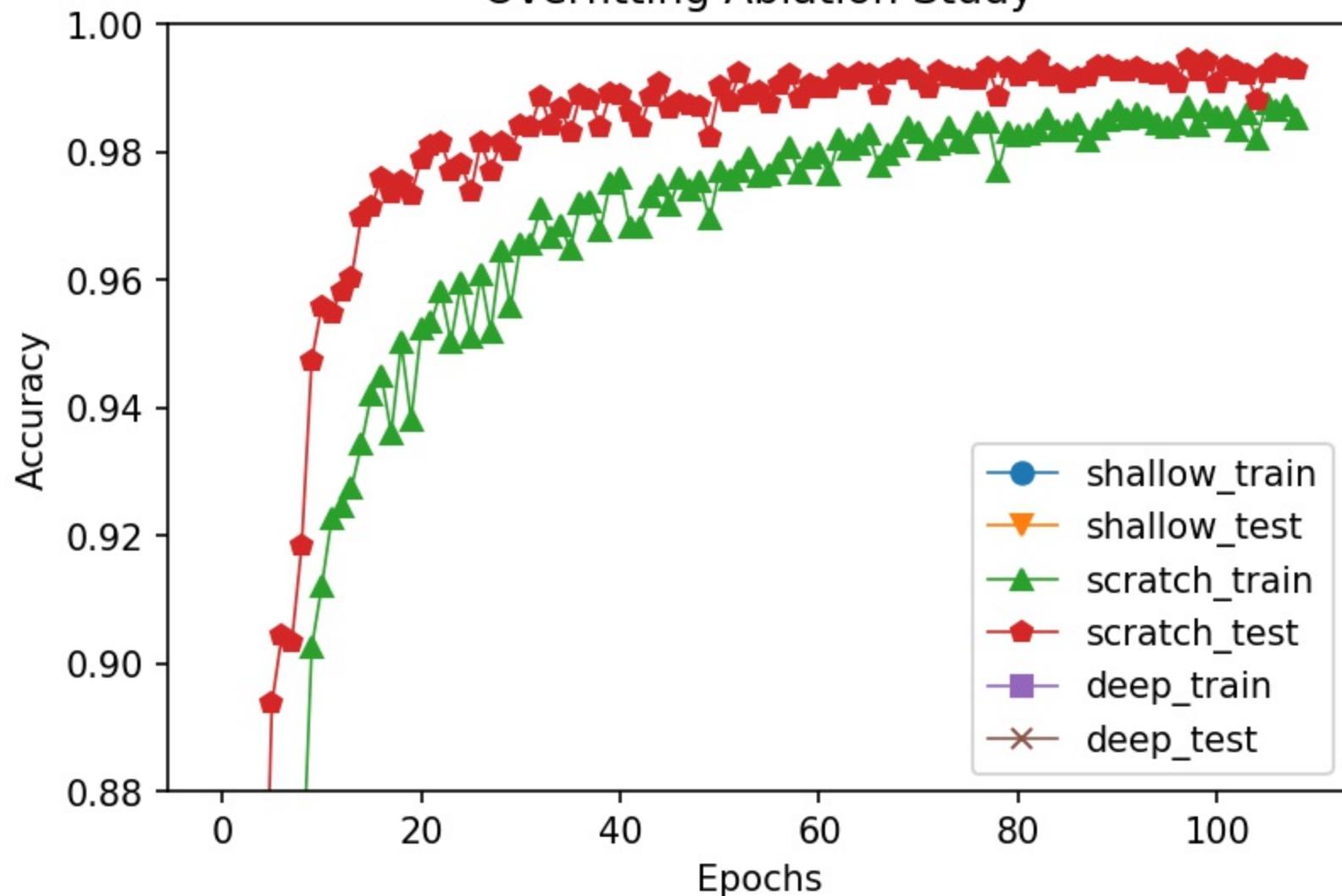
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	NaN	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



epoch 108 takes 314 seconds.  
[110, 30] loss: 0.078  
[110, 60] loss: 0.053

```
[110,    90] loss: 0.043
[110,   120] loss: 0.085
[110,   150] loss: 0.077
[110,   180] loss: 0.081
[110,   210] loss: 0.074
[110,   240] loss: 0.079
[110,   270] loss: 0.053
[110,   300] loss: 0.052
[110,   330] loss: 0.042
[110,   360] loss: 0.084
[110,   390] loss: 0.050
[110,   420] loss: 0.068
[110,   450] loss: 0.026
[110,   480] loss: 0.073
[110,   510] loss: 0.062
[110,   540] loss: 0.091
[110,   570] loss: 0.064
[110,   600] loss: 0.051
[110,   630] loss: 0.060
[110,   660] loss: 0.041
[110,   690] loss: 0.083
[110,   720] loss: 0.048
[110,   750] loss: 0.069
[110,   780] loss: 0.056
[110,   810] loss: 0.044
[110,   840] loss: 0.051
[110,   870] loss: 0.055
[110,   900] loss: 0.063
[110,   930] loss: 0.073
[110,   960] loss: 0.073
[110,   990] loss: 0.060
[110,  1020] loss: 0.063
[110,  1050] loss: 0.056
[110,  1080] loss: 0.031
[110,  1110] loss: 0.049
[110,  1140] loss: 0.066
[110,  1170] loss: 0.085
[110,  1200] loss: 0.032
[110,  1230] loss: 0.039
```

```
[110, 1260] loss: 0.070
[110, 1290] loss: 0.044
[110, 1320] loss: 0.054
[110, 1350] loss: 0.061
[110, 1380] loss: 0.045
[110, 1410] loss: 0.053
[110, 1440] loss: 0.053
[110, 1470] loss: 0.030
[110, 1500] loss: 0.038
[110, 1530] loss: 0.065
[110, 1560] loss: 0.057
[110, 1590] loss: 0.062
[110, 1620] loss: 0.034
[110, 1650] loss: 0.059
[110, 1680] loss: 0.077
[110, 1710] loss: 0.053
[110, 1740] loss: 0.036
[110, 1770] loss: 0.054
[110, 1800] loss: 0.040
[110, 1830] loss: 0.064
[110, 1860] loss: 0.047
[110, 1890] loss: 0.074
[110, 1920] loss: 0.072
[110, 1950] loss: 0.054
[110, 1980] loss: 0.064
[110, 2010] loss: 0.054
[110, 2040] loss: 0.036
[110, 2070] loss: 0.040
[110, 2100] loss: 0.049
[110, 2130] loss: 0.065
[110, 2160] loss: 0.056
[110, 2190] loss: 0.049
eval intermediate_models/epoch109_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.985119
epoch 109 accuracy on train set is: 0.9851190476190477
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	NaN	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.990458

epoch 109 accuracy on test set is: 0.9904580152671756

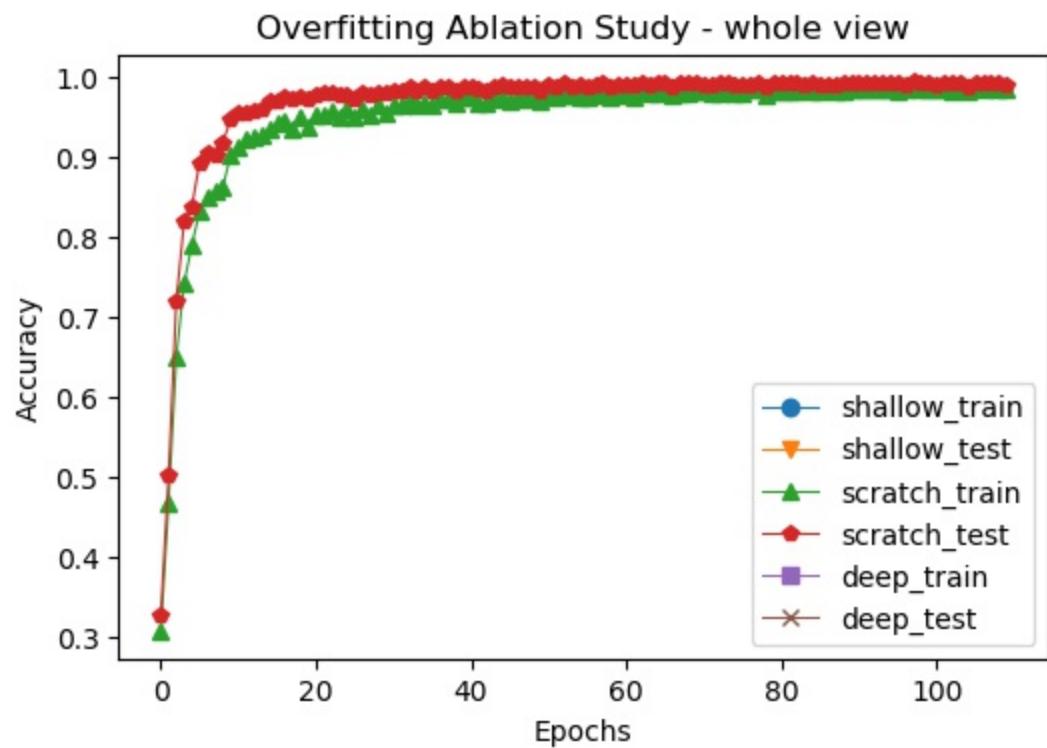
evaluating on test set takes 16 seconds.

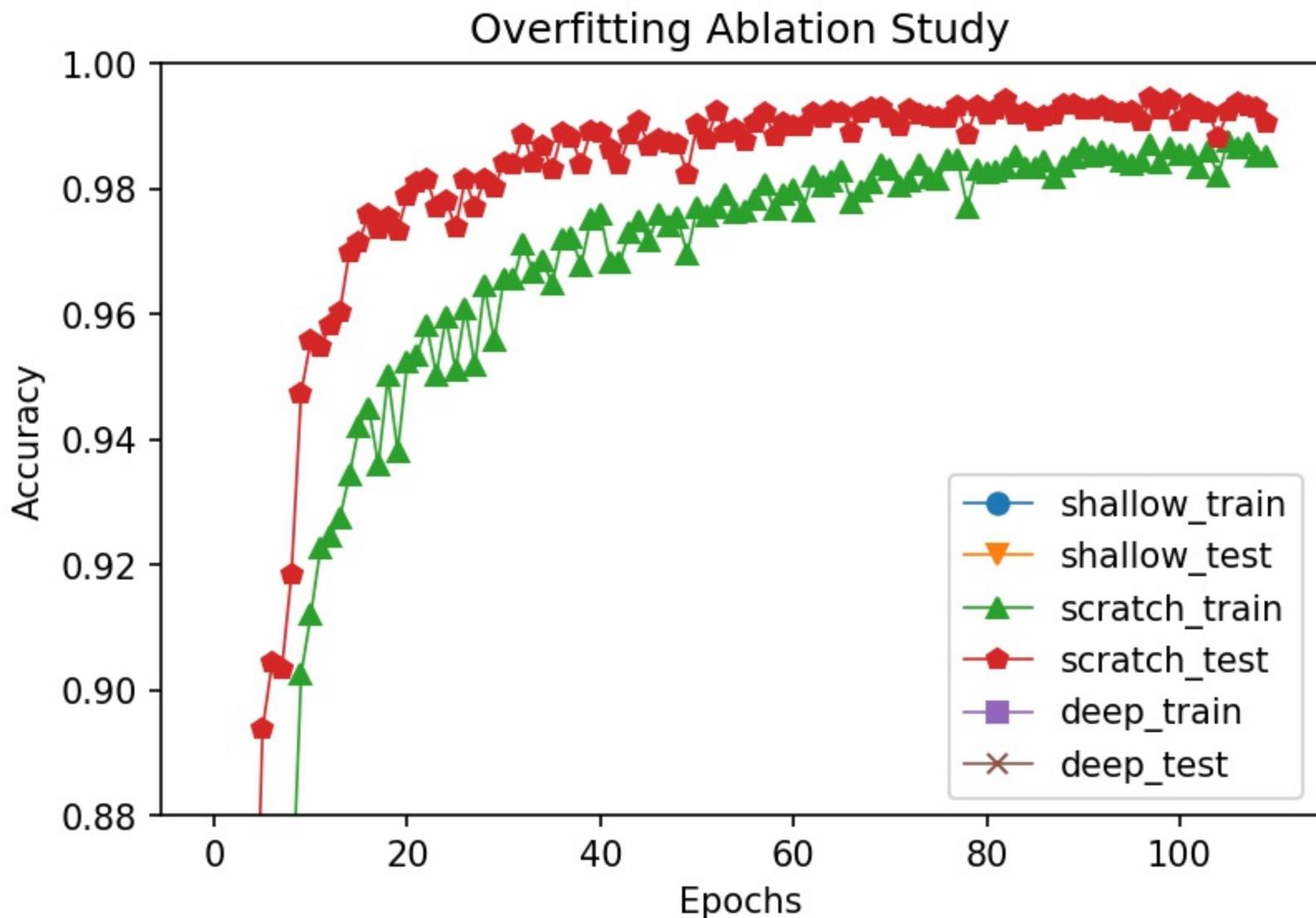
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	NaN	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 109 takes 314 seconds.  
[111,    30] loss: 0.074  
[111,    60] loss: 0.043
```

```
[111,    90] loss: 0.067
[111,   120] loss: 0.061
[111,   150] loss: 0.080
[111,   180] loss: 0.067
[111,   210] loss: 0.046
[111,   240] loss: 0.070
[111,   270] loss: 0.047
[111,   300] loss: 0.058
[111,   330] loss: 0.044
[111,   360] loss: 0.067
[111,   390] loss: 0.063
[111,   420] loss: 0.064
[111,   450] loss: 0.056
[111,   480] loss: 0.029
[111,   510] loss: 0.066
[111,   540] loss: 0.058
[111,   570] loss: 0.071
[111,   600] loss: 0.064
[111,   630] loss: 0.051
[111,   660] loss: 0.054
[111,   690] loss: 0.058
[111,   720] loss: 0.035
[111,   750] loss: 0.049
[111,   780] loss: 0.053
[111,   810] loss: 0.049
[111,   840] loss: 0.037
[111,   870] loss: 0.046
[111,   900] loss: 0.080
[111,   930] loss: 0.068
[111,   960] loss: 0.057
[111,   990] loss: 0.051
[111,  1020] loss: 0.075
[111,  1050] loss: 0.044
[111,  1080] loss: 0.050
[111,  1110] loss: 0.050
[111,  1140] loss: 0.065
[111,  1170] loss: 0.044
[111,  1200] loss: 0.043
[111,  1230] loss: 0.045
```

```
[111, 1260] loss: 0.071
[111, 1290] loss: 0.053
[111, 1320] loss: 0.064
[111, 1350] loss: 0.057
[111, 1380] loss: 0.079
[111, 1410] loss: 0.056
[111, 1440] loss: 0.046
[111, 1470] loss: 0.048
[111, 1500] loss: 0.057
[111, 1530] loss: 0.060
[111, 1560] loss: 0.047
[111, 1590] loss: 0.055
[111, 1620] loss: 0.067
[111, 1650] loss: 0.073
[111, 1680] loss: 0.072
[111, 1710] loss: 0.060
[111, 1740] loss: 0.037
[111, 1770] loss: 0.054
[111, 1800] loss: 0.056
[111, 1830] loss: 0.045
[111, 1860] loss: 0.056
[111, 1890] loss: 0.052
[111, 1920] loss: 0.047
[111, 1950] loss: 0.076
[111, 1980] loss: 0.062
[111, 2010] loss: 0.048
[111, 2040] loss: 0.051
[111, 2070] loss: 0.067
[111, 2100] loss: 0.073
[111, 2130] loss: 0.065
[111, 2160] loss: 0.040
[111, 2190] loss: 0.069
eval intermediate_models/epoch110_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.988141
epoch 110 accuracy on train set is: 0.9881406761177753
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	NaN	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.995093

epoch 110 accuracy on test set is: 0.995092693565976

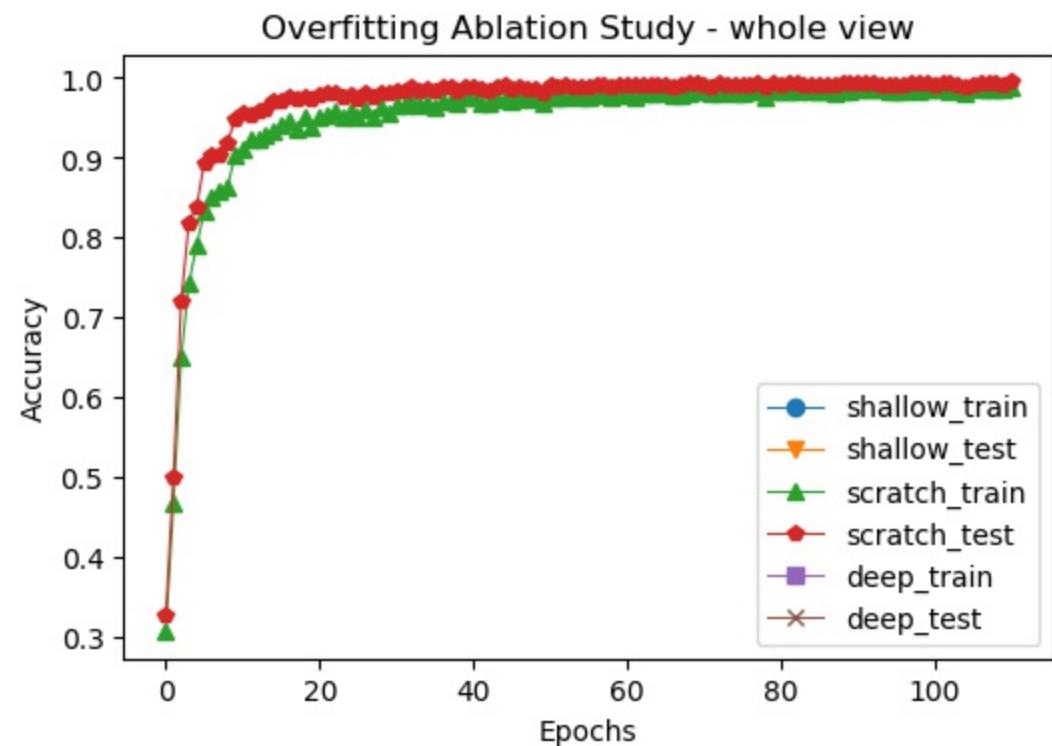
evaluating on test set takes 16 seconds.

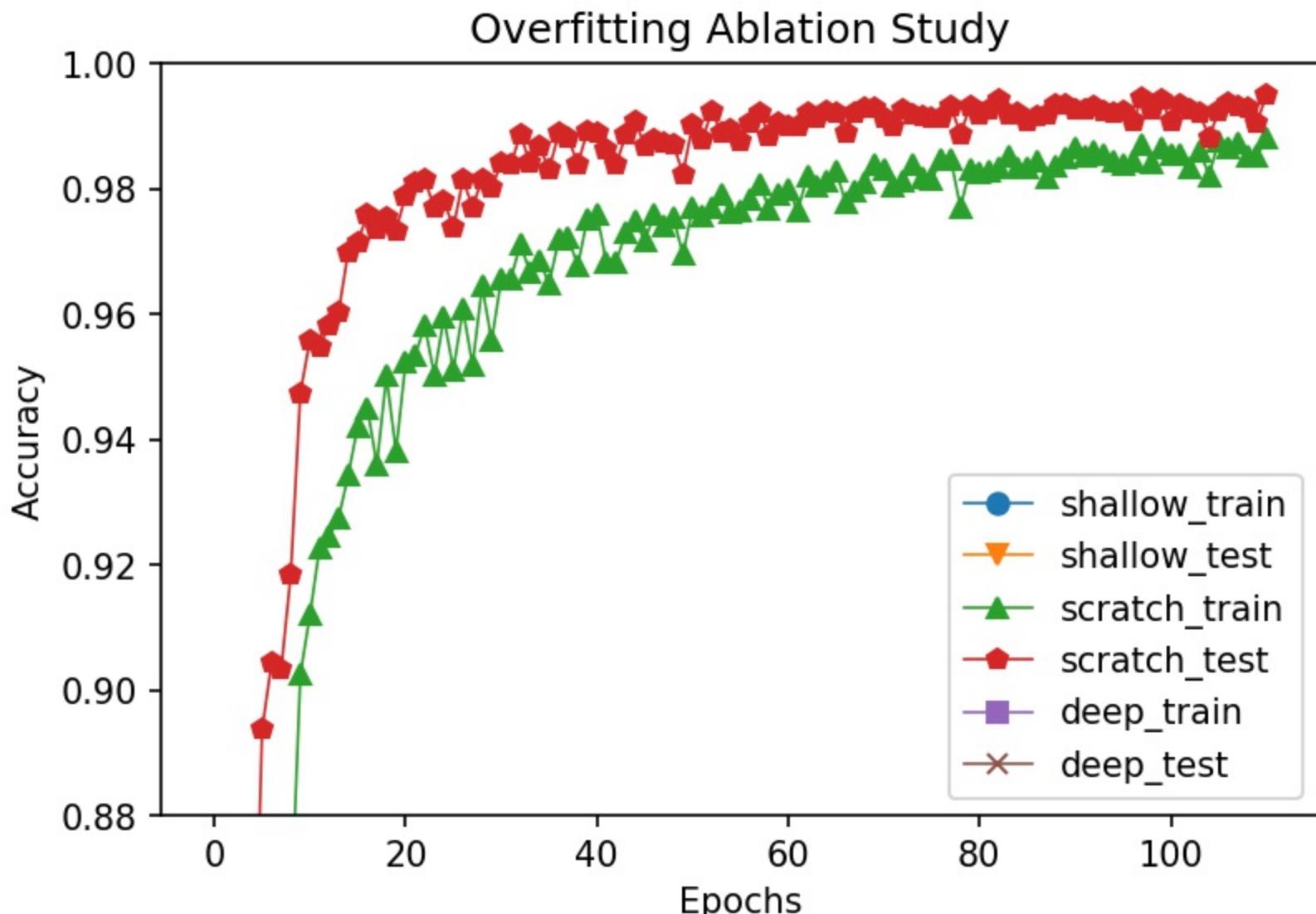
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	NaN	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 110 takes 315 seconds.

[112, 30] loss: 0.046

[112, 60] loss: 0.079

```
[112,    90] loss: 0.044
[112,   120] loss: 0.059
[112,   150] loss: 0.063
[112,   180] loss: 0.115
[112,   210] loss: 0.045
[112,   240] loss: 0.074
[112,   270] loss: 0.054
[112,   300] loss: 0.066
[112,   330] loss: 0.060
[112,   360] loss: 0.068
[112,   390] loss: 0.043
[112,   420] loss: 0.055
[112,   450] loss: 0.056
[112,   480] loss: 0.082
[112,   510] loss: 0.061
[112,   540] loss: 0.043
[112,   570] loss: 0.065
[112,   600] loss: 0.100
[112,   630] loss: 0.071
[112,   660] loss: 0.047
[112,   690] loss: 0.047
[112,   720] loss: 0.076
[112,   750] loss: 0.057
[112,   780] loss: 0.050
[112,   810] loss: 0.037
[112,   840] loss: 0.060
[112,   870] loss: 0.065
[112,   900] loss: 0.060
[112,   930] loss: 0.059
[112,   960] loss: 0.055
[112,   990] loss: 0.041
[112,  1020] loss: 0.066
[112,  1050] loss: 0.076
[112,  1080] loss: 0.087
[112,  1110] loss: 0.040
[112,  1140] loss: 0.034
[112,  1170] loss: 0.066
[112,  1200] loss: 0.047
[112,  1230] loss: 0.075
```

```
[112, 1260] loss: 0.064
[112, 1290] loss: 0.070
[112, 1320] loss: 0.057
[112, 1350] loss: 0.062
[112, 1380] loss: 0.053
[112, 1410] loss: 0.031
[112, 1440] loss: 0.067
[112, 1470] loss: 0.021
[112, 1500] loss: 0.042
[112, 1530] loss: 0.065
[112, 1560] loss: 0.044
[112, 1590] loss: 0.036
[112, 1620] loss: 0.085
[112, 1650] loss: 0.080
[112, 1680] loss: 0.070
[112, 1710] loss: 0.071
[112, 1740] loss: 0.062
[112, 1770] loss: 0.068
[112, 1800] loss: 0.073
[112, 1830] loss: 0.032
[112, 1860] loss: 0.046
[112, 1890] loss: 0.074
[112, 1920] loss: 0.077
[112, 1950] loss: 0.037
[112, 1980] loss: 0.051
[112, 2010] loss: 0.068
[112, 2040] loss: 0.062
[112, 2070] loss: 0.055
[112, 2100] loss: 0.043
[112, 2130] loss: 0.046
[112, 2160] loss: 0.068
[112, 2190] loss: 0.044
eval intermediate_models/epoch111_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.988731
epoch 111 accuracy on train set is: 0.9887313704107598
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	NaN	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.994457

epoch 111 accuracy on test set is: 0.9944565612504543

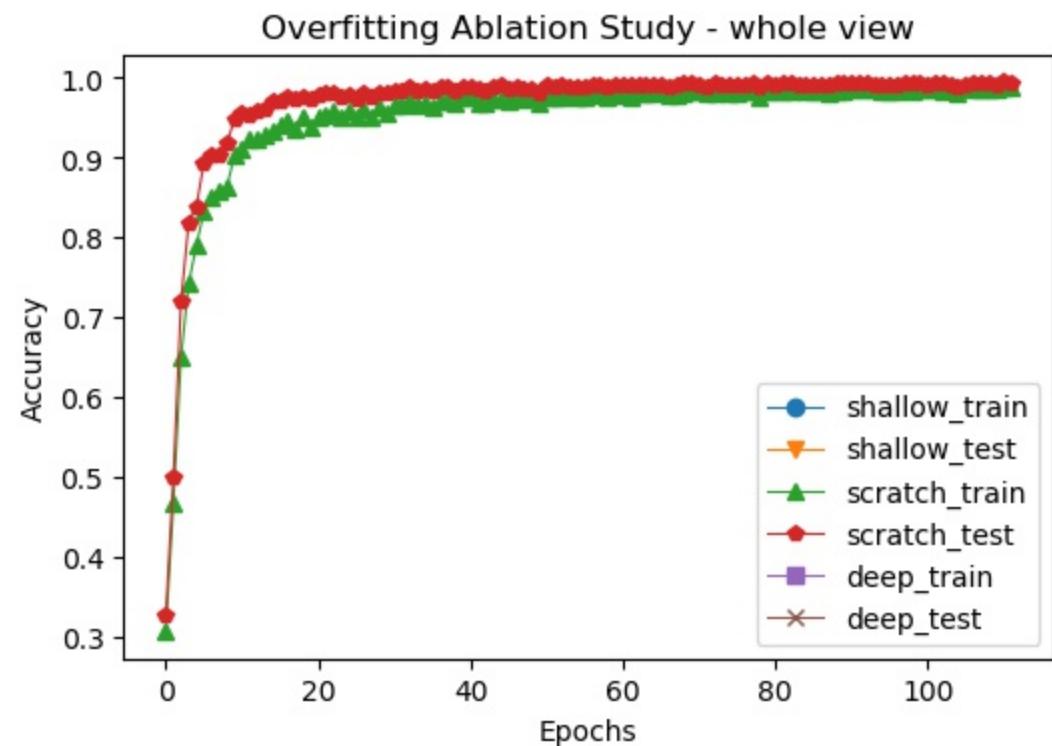
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

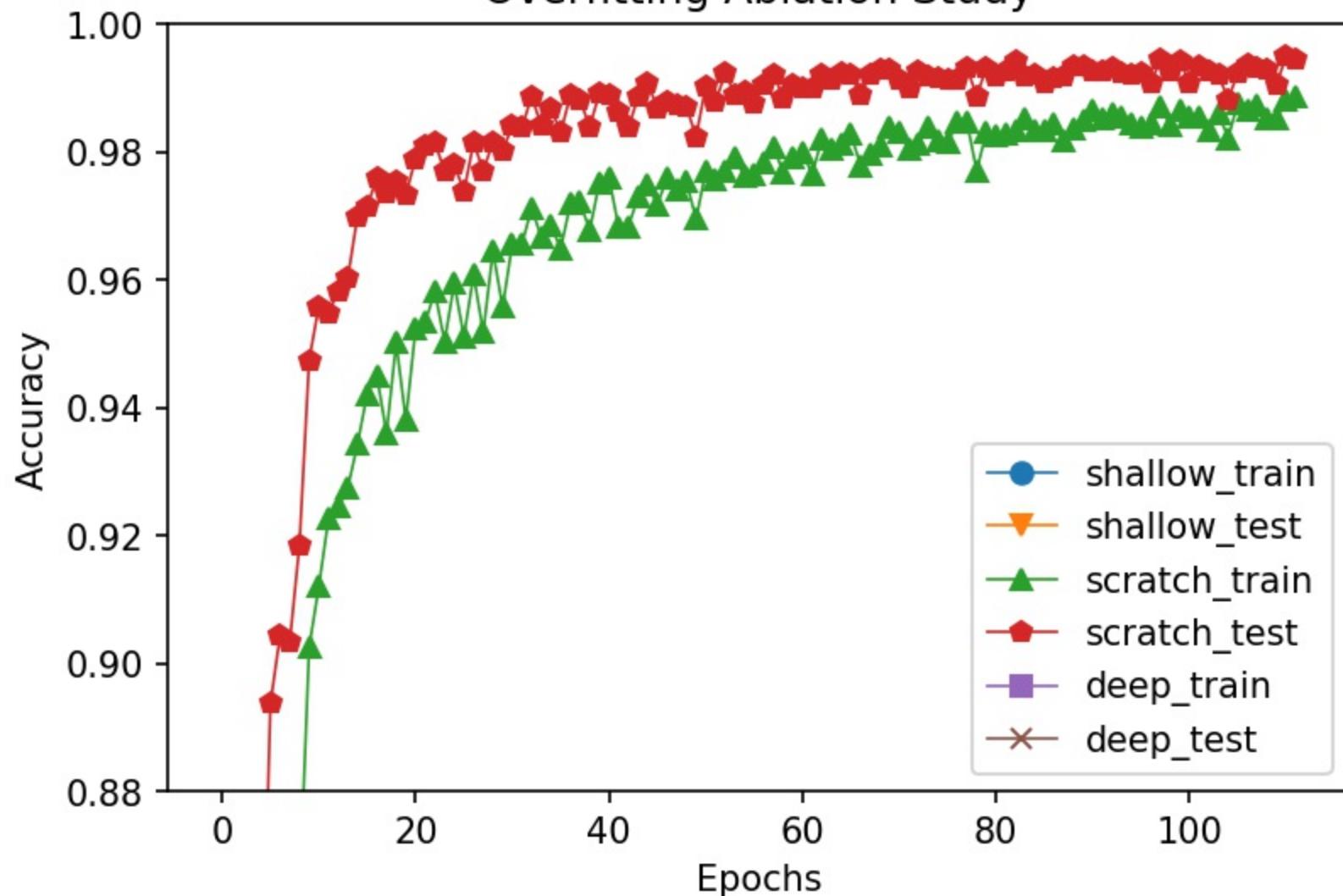
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	NaN	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



```
epoch 111 takes 315 seconds.  
[113,    30] loss: 0.051  
[113,    60] loss: 0.028
```

```
[113,    90] loss: 0.057
[113,   120] loss: 0.063
[113,   150] loss: 0.083
[113,   180] loss: 0.074
[113,   210] loss: 0.050
[113,   240] loss: 0.054
[113,   270] loss: 0.040
[113,   300] loss: 0.036
[113,   330] loss: 0.046
[113,   360] loss: 0.034
[113,   390] loss: 0.060
[113,   420] loss: 0.055
[113,   450] loss: 0.071
[113,   480] loss: 0.048
[113,   510] loss: 0.066
[113,   540] loss: 0.059
[113,   570] loss: 0.088
[113,   600] loss: 0.066
[113,   630] loss: 0.059
[113,   660] loss: 0.036
[113,   690] loss: 0.039
[113,   720] loss: 0.026
[113,   750] loss: 0.043
[113,   780] loss: 0.034
[113,   810] loss: 0.062
[113,   840] loss: 0.039
[113,   870] loss: 0.042
[113,   900] loss: 0.048
[113,   930] loss: 0.064
[113,   960] loss: 0.036
[113,   990] loss: 0.075
[113,  1020] loss: 0.043
[113,  1050] loss: 0.046
[113,  1080] loss: 0.043
[113,  1110] loss: 0.043
[113,  1140] loss: 0.102
[113,  1170] loss: 0.047
[113,  1200] loss: 0.062
[113,  1230] loss: 0.061
```

```
[113, 1260] loss: 0.069
[113, 1290] loss: 0.052
[113, 1320] loss: 0.031
[113, 1350] loss: 0.056
[113, 1380] loss: 0.052
[113, 1410] loss: 0.032
[113, 1440] loss: 0.029
[113, 1470] loss: 0.048
[113, 1500] loss: 0.056
[113, 1530] loss: 0.048
[113, 1560] loss: 0.057
[113, 1590] loss: 0.039
[113, 1620] loss: 0.045
[113, 1650] loss: 0.064
[113, 1680] loss: 0.038
[113, 1710] loss: 0.080
[113, 1740] loss: 0.060
[113, 1770] loss: 0.058
[113, 1800] loss: 0.059
[113, 1830] loss: 0.048
[113, 1860] loss: 0.109
[113, 1890] loss: 0.057
[113, 1920] loss: 0.068
[113, 1950] loss: 0.070
[113, 1980] loss: 0.045
[113, 2010] loss: 0.028
[113, 2040] loss: 0.066
[113, 2070] loss: 0.089
[113, 2100] loss: 0.049
[113, 2130] loss: 0.074
[113, 2160] loss: 0.044
[113, 2190] loss: 0.059
eval intermediate_models/epoch112_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.986187
epoch 112 accuracy on train set is: 0.9861868411486732
evaluating on training set takes 63 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	NaN	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.994457

epoch 112 accuracy on test set is: 0.9944565612504543

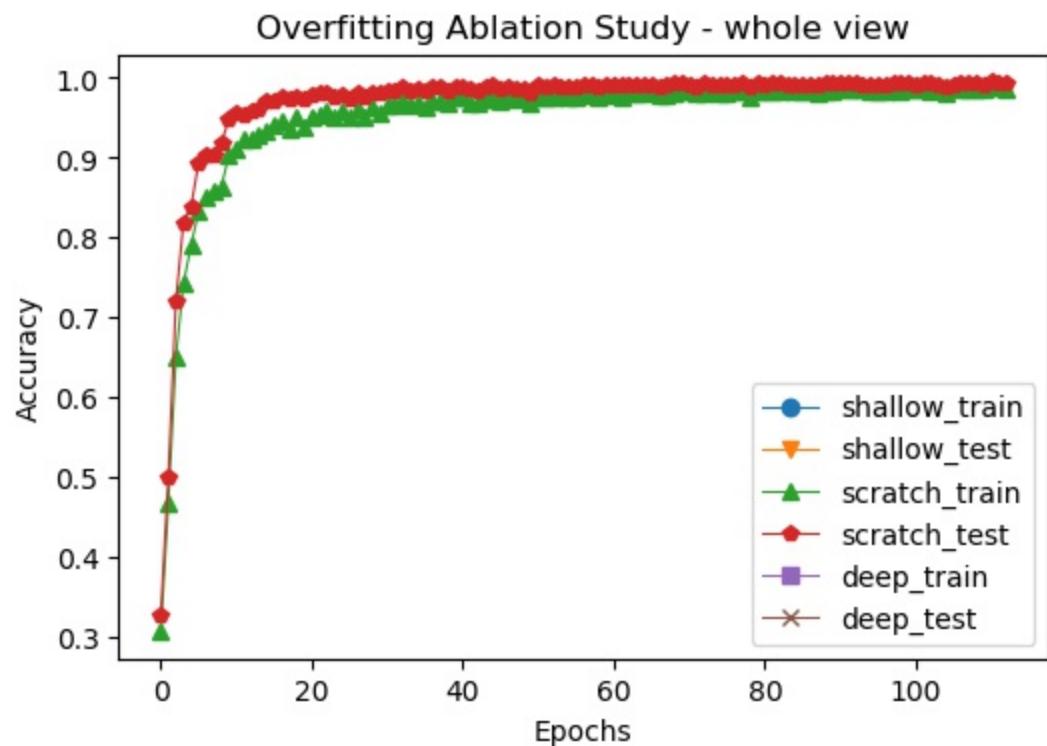
evaluating on test set takes 16 seconds.

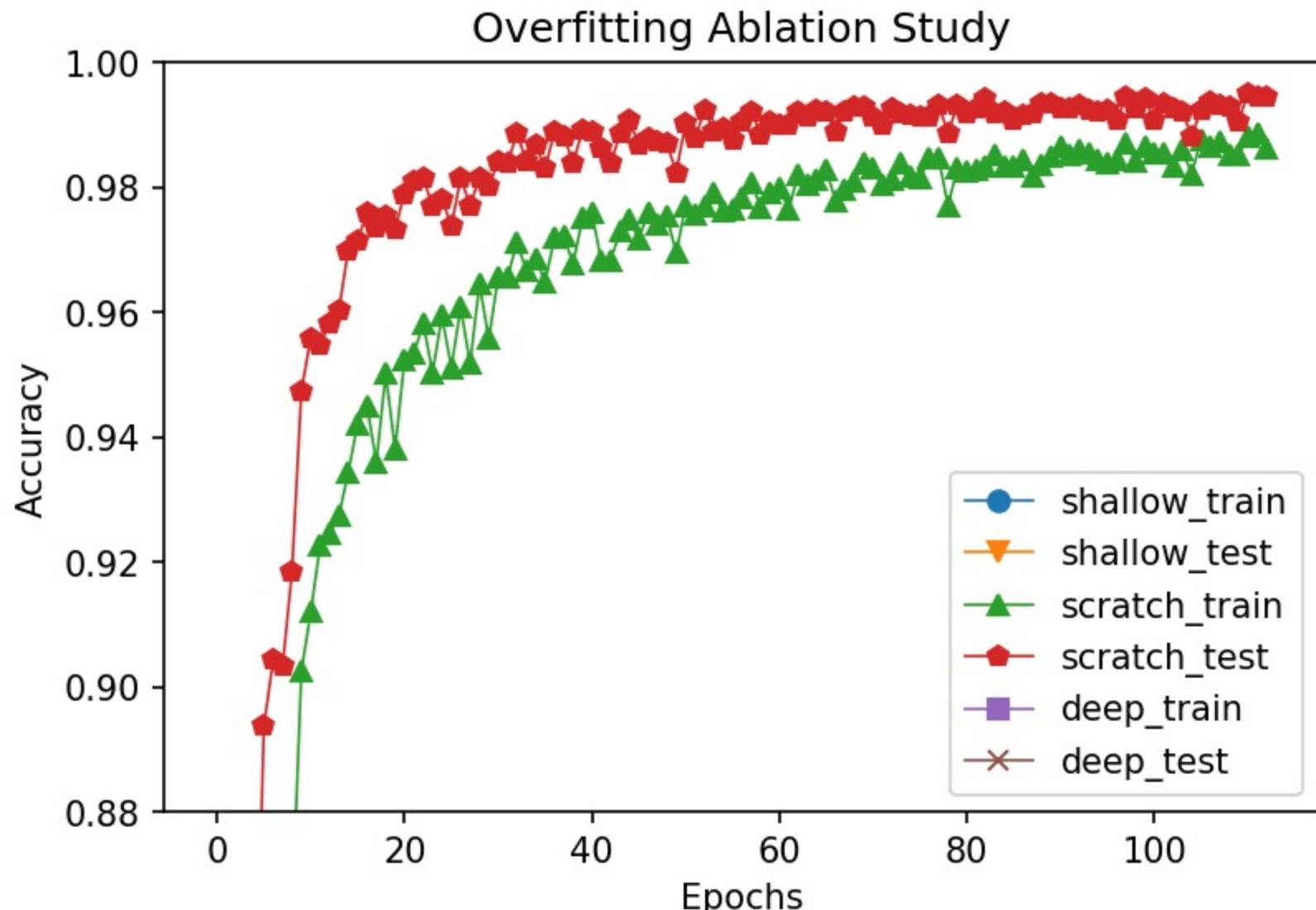
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	NaN	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





```
[114,    90] loss: 0.062
[114,   120] loss: 0.059
[114,   150] loss: 0.070
[114,   180] loss: 0.084
[114,   210] loss: 0.056
[114,   240] loss: 0.046
[114,   270] loss: 0.033
[114,   300] loss: 0.023
[114,   330] loss: 0.043
[114,   360] loss: 0.024
[114,   390] loss: 0.037
[114,   420] loss: 0.039
[114,   450] loss: 0.030
[114,   480] loss: 0.053
[114,   510] loss: 0.052
[114,   540] loss: 0.056
[114,   570] loss: 0.070
[114,   600] loss: 0.067
[114,   630] loss: 0.074
[114,   660] loss: 0.054
[114,   690] loss: 0.040
[114,   720] loss: 0.081
[114,   750] loss: 0.077
[114,   780] loss: 0.060
[114,   810] loss: 0.053
[114,   840] loss: 0.054
[114,   870] loss: 0.086
[114,   900] loss: 0.040
[114,   930] loss: 0.044
[114,   960] loss: 0.054
[114,   990] loss: 0.075
[114,  1020] loss: 0.060
[114,  1050] loss: 0.061
[114,  1080] loss: 0.062
[114,  1110] loss: 0.049
[114,  1140] loss: 0.085
[114,  1170] loss: 0.031
[114,  1200] loss: 0.068
[114,  1230] loss: 0.045
```

```
[114, 1260] loss: 0.052
[114, 1290] loss: 0.047
[114, 1320] loss: 0.034
[114, 1350] loss: 0.052
[114, 1380] loss: 0.050
[114, 1410] loss: 0.035
[114, 1440] loss: 0.055
[114, 1470] loss: 0.045
[114, 1500] loss: 0.055
[114, 1530] loss: 0.047
[114, 1560] loss: 0.062
[114, 1590] loss: 0.067
[114, 1620] loss: 0.070
[114, 1650] loss: 0.042
[114, 1680] loss: 0.084
[114, 1710] loss: 0.052
[114, 1740] loss: 0.070
[114, 1770] loss: 0.044
[114, 1800] loss: 0.045
[114, 1830] loss: 0.042
[114, 1860] loss: 0.053
[114, 1890] loss: 0.070
[114, 1920] loss: 0.070
[114, 1950] loss: 0.060
[114, 1980] loss: 0.064
[114, 2010] loss: 0.037
[114, 2040] loss: 0.058
[114, 2070] loss: 0.063
[114, 2100] loss: 0.069
[114, 2130] loss: 0.073
[114, 2160] loss: 0.065
[114, 2190] loss: 0.064
eval intermediate_models/epoch113_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.987209
epoch 113 accuracy on train set is: 0.9872091966557616
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	NaN	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993911

epoch 113 accuracy on test set is: 0.9939113049800072

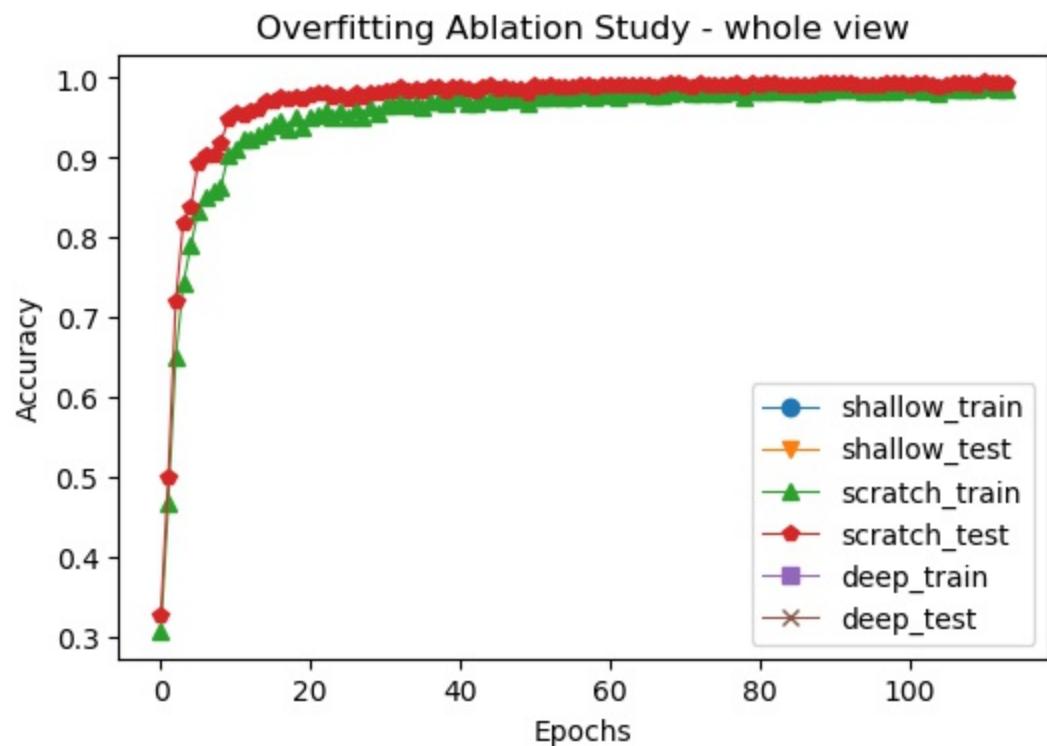
evaluating on test set takes 16 seconds.

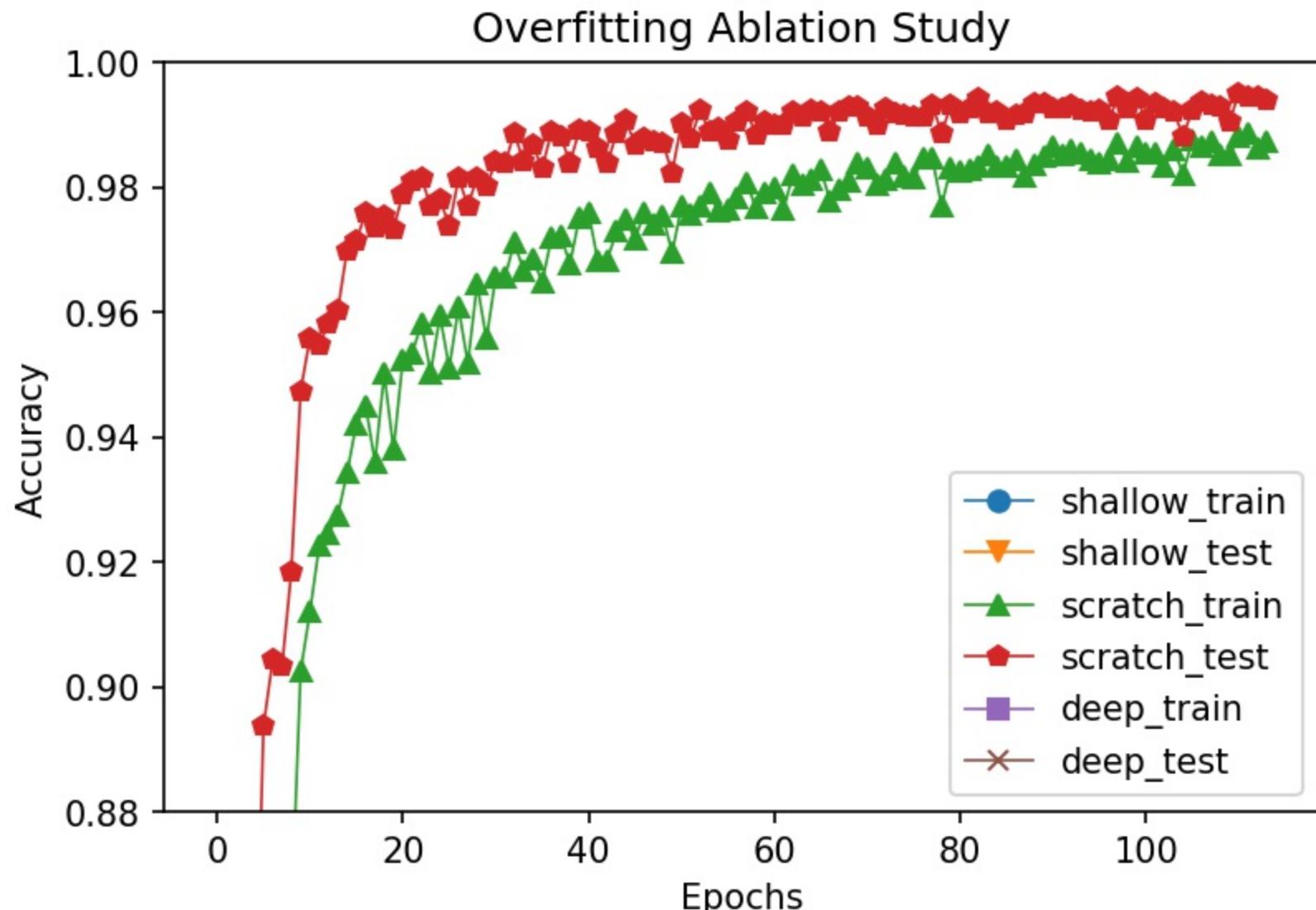
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	NaN	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 113 takes 314 seconds.

[115, 30] loss: 0.050

[115, 60] loss: 0.033

```
[115,    90] loss: 0.059
[115,   120] loss: 0.066
[115,   150] loss: 0.038
[115,   180] loss: 0.055
[115,   210] loss: 0.056
[115,   240] loss: 0.063
[115,   270] loss: 0.052
[115,   300] loss: 0.059
[115,   330] loss: 0.049
[115,   360] loss: 0.046
[115,   390] loss: 0.072
[115,   420] loss: 0.058
[115,   450] loss: 0.056
[115,   480] loss: 0.061
[115,   510] loss: 0.080
[115,   540] loss: 0.070
[115,   570] loss: 0.066
[115,   600] loss: 0.059
[115,   630] loss: 0.047
[115,   660] loss: 0.054
[115,   690] loss: 0.055
[115,   720] loss: 0.050
[115,   750] loss: 0.076
[115,   780] loss: 0.078
[115,   810] loss: 0.043
[115,   840] loss: 0.065
[115,   870] loss: 0.040
[115,   900] loss: 0.096
[115,   930] loss: 0.043
[115,   960] loss: 0.046
[115,   990] loss: 0.039
[115,  1020] loss: 0.061
[115,  1050] loss: 0.061
[115,  1080] loss: 0.069
[115,  1110] loss: 0.051
[115,  1140] loss: 0.041
[115,  1170] loss: 0.068
[115,  1200] loss: 0.044
[115,  1230] loss: 0.033
```

```
[115, 1260] loss: 0.056
[115, 1290] loss: 0.042
[115, 1320] loss: 0.041
[115, 1350] loss: 0.039
[115, 1380] loss: 0.052
[115, 1410] loss: 0.049
[115, 1440] loss: 0.040
[115, 1470] loss: 0.085
[115, 1500] loss: 0.034
[115, 1530] loss: 0.032
[115, 1560] loss: 0.097
[115, 1590] loss: 0.049
[115, 1620] loss: 0.047
[115, 1650] loss: 0.039
[115, 1680] loss: 0.037
[115, 1710] loss: 0.036
[115, 1740] loss: 0.046
[115, 1770] loss: 0.064
[115, 1800] loss: 0.093
[115, 1830] loss: 0.050
[115, 1860] loss: 0.024
[115, 1890] loss: 0.060
[115, 1920] loss: 0.044
[115, 1950] loss: 0.044
[115, 1980] loss: 0.055
[115, 2010] loss: 0.039
[115, 2040] loss: 0.057
[115, 2070] loss: 0.049
[115, 2100] loss: 0.049
[115, 2130] loss: 0.046
[115, 2160] loss: 0.026
[115, 2190] loss: 0.056
eval intermediate_models/epoch114_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.987573
epoch 114 accuracy on train set is: 0.9875727008360596
evaluating on training set takes 63 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	NaN	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993548

epoch 114 accuracy on test set is: 0.9935478007997092

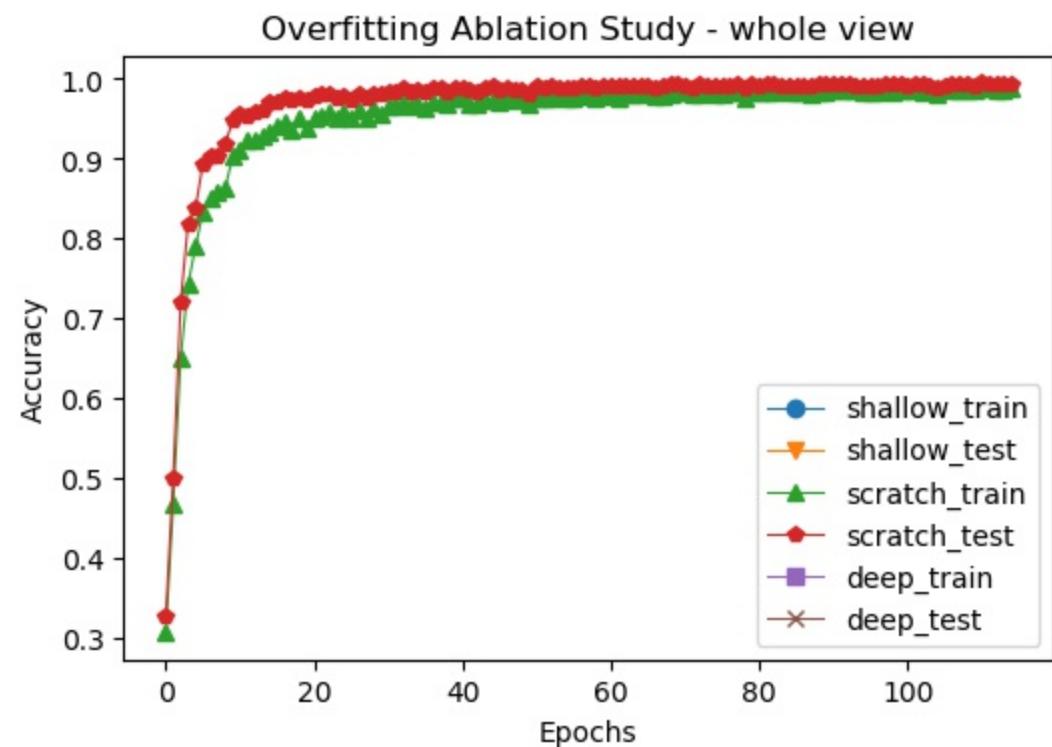
evaluating on test set takes 16 seconds.

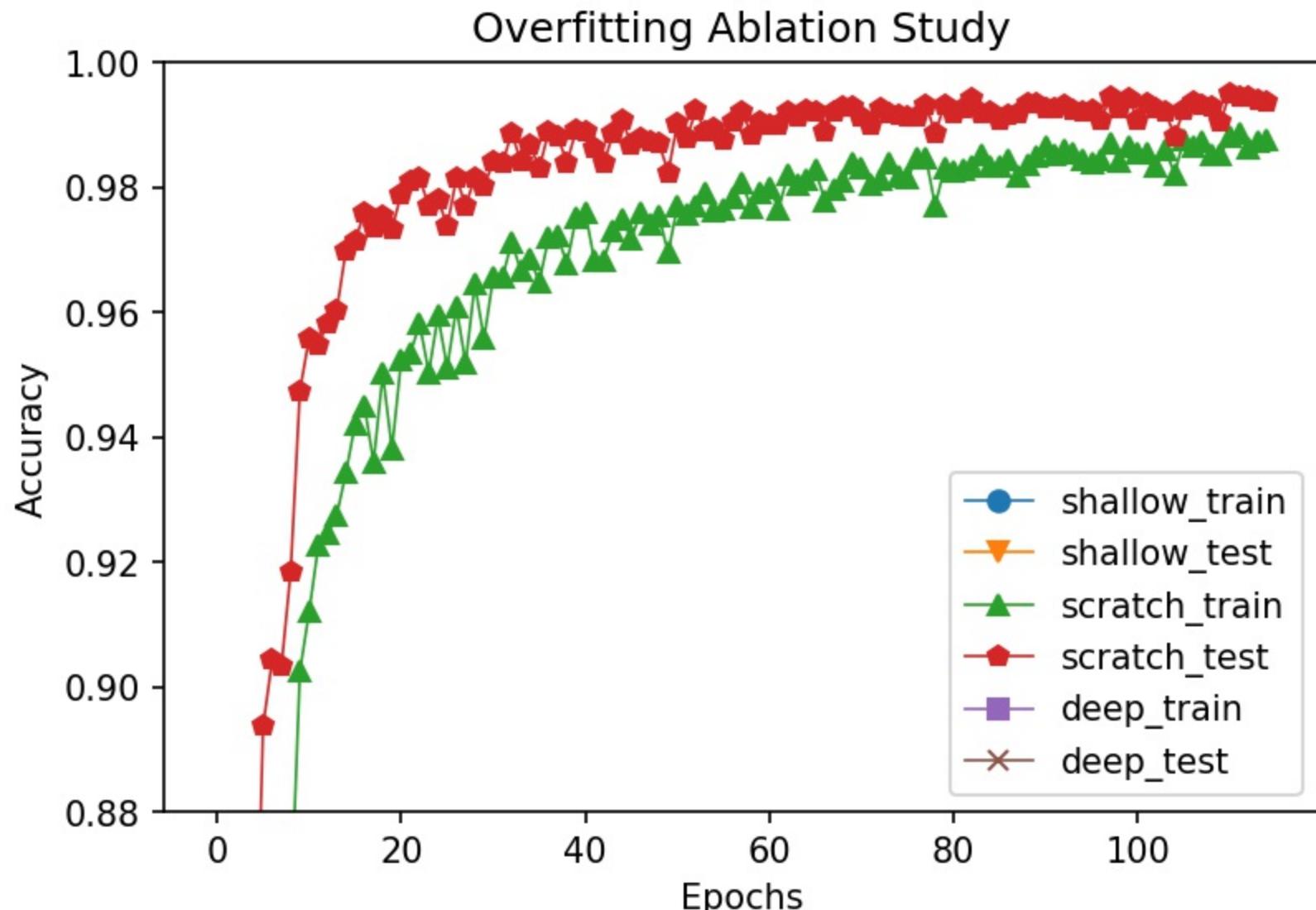
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	NaN	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 114 takes 315 seconds.

[116, 30] loss: 0.045

[116, 60] loss: 0.040

```
[116,    90] loss: 0.044
[116,   120] loss: 0.089
[116,   150] loss: 0.055
[116,   180] loss: 0.020
[116,   210] loss: 0.049
[116,   240] loss: 0.054
[116,   270] loss: 0.022
[116,   300] loss: 0.056
[116,   330] loss: 0.053
[116,   360] loss: 0.025
[116,   390] loss: 0.068
[116,   420] loss: 0.057
[116,   450] loss: 0.040
[116,   480] loss: 0.052
[116,   510] loss: 0.052
[116,   540] loss: 0.047
[116,   570] loss: 0.024
[116,   600] loss: 0.039
[116,   630] loss: 0.061
[116,   660] loss: 0.055
[116,   690] loss: 0.022
[116,   720] loss: 0.018
[116,   750] loss: 0.065
[116,   780] loss: 0.049
[116,   810] loss: 0.086
[116,   840] loss: 0.053
[116,   870] loss: 0.039
[116,   900] loss: 0.067
[116,   930] loss: 0.048
[116,   960] loss: 0.080
[116,   990] loss: 0.052
[116,  1020] loss: 0.059
[116,  1050] loss: 0.031
[116,  1080] loss: 0.039
[116,  1110] loss: 0.052
[116,  1140] loss: 0.068
[116,  1170] loss: 0.062
[116,  1200] loss: 0.047
[116,  1230] loss: 0.065
```

```
[116, 1260] loss: 0.025
[116, 1290] loss: 0.036
[116, 1320] loss: 0.037
[116, 1350] loss: 0.082
[116, 1380] loss: 0.076
[116, 1410] loss: 0.035
[116, 1440] loss: 0.043
[116, 1470] loss: 0.048
[116, 1500] loss: 0.060
[116, 1530] loss: 0.050
[116, 1560] loss: 0.056
[116, 1590] loss: 0.068
[116, 1620] loss: 0.059
[116, 1650] loss: 0.056
[116, 1680] loss: 0.056
[116, 1710] loss: 0.034
[116, 1740] loss: 0.057
[116, 1770] loss: 0.049
[116, 1800] loss: 0.041
[116, 1830] loss: 0.047
[116, 1860] loss: 0.081
[116, 1890] loss: 0.059
[116, 1920] loss: 0.051
[116, 1950] loss: 0.066
[116, 1980] loss: 0.069
[116, 2010] loss: 0.037
[116, 2040] loss: 0.074
[116, 2070] loss: 0.071
[116, 2100] loss: 0.074
[116, 2130] loss: 0.052
[116, 2160] loss: 0.090
[116, 2190] loss: 0.064
eval intermediate_models/epoch115_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.986005
epoch 115 accuracy on train set is: 0.9860050890585241
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	NaN	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993820

epoch 115 accuracy on test set is: 0.9938204289349327

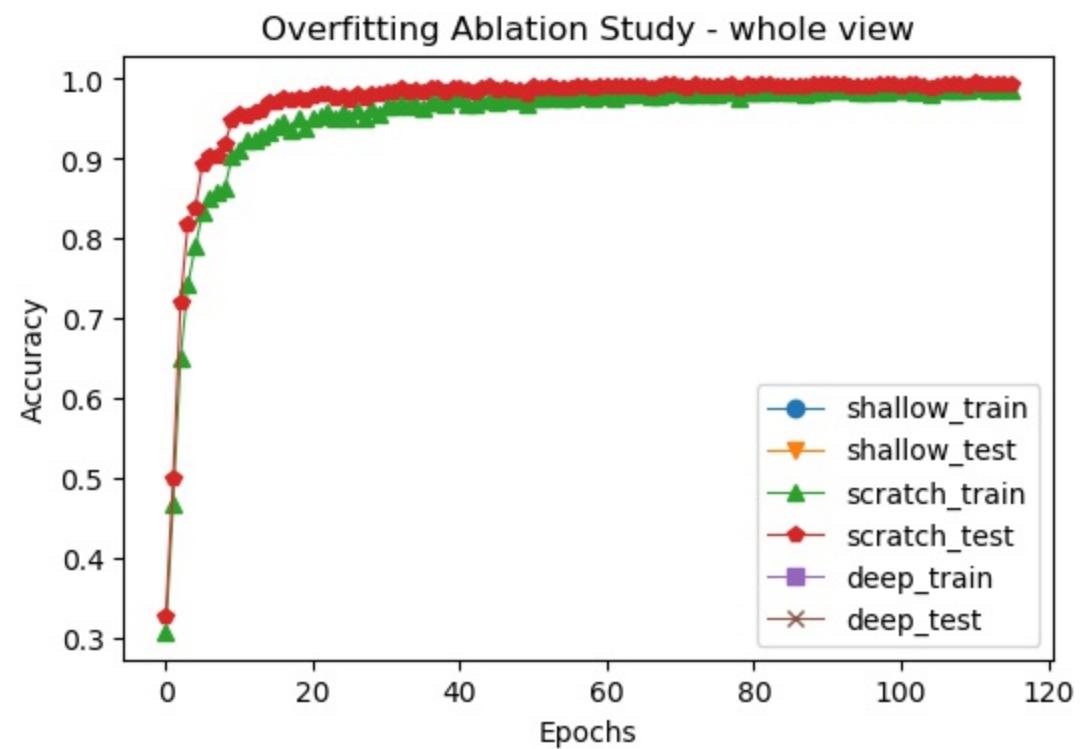
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

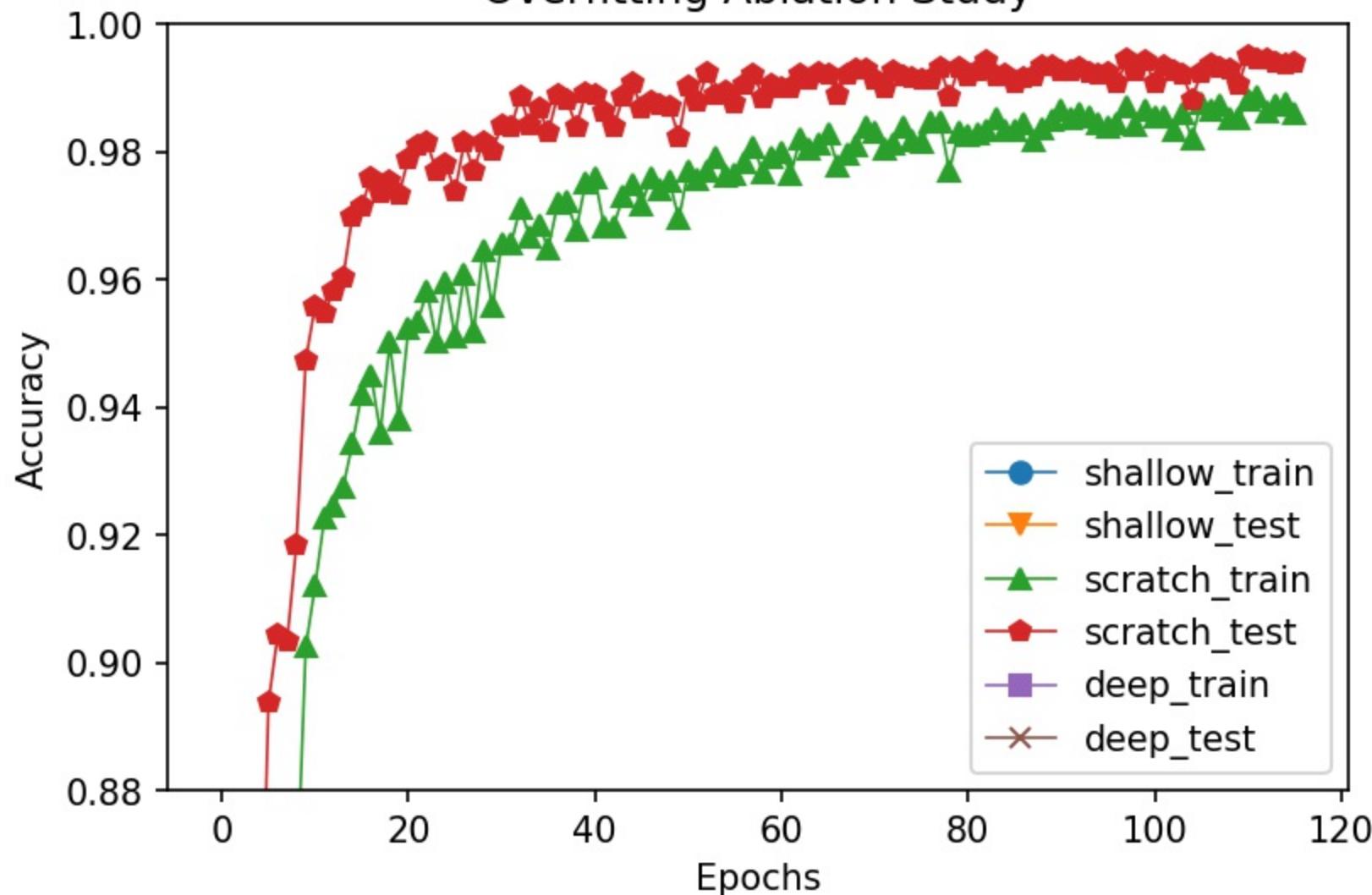
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	NaN	NaN	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



```
epoch 115 takes 315 seconds.  
[117,    30] loss: 0.073  
[117,    60] loss: 0.056
```

```
[117,    90] loss: 0.028
[117,   120] loss: 0.064
[117,   150] loss: 0.071
[117,   180] loss: 0.046
[117,   210] loss: 0.043
[117,   240] loss: 0.084
[117,   270] loss: 0.048
[117,   300] loss: 0.057
[117,   330] loss: 0.037
[117,   360] loss: 0.040
[117,   390] loss: 0.047
[117,   420] loss: 0.052
[117,   450] loss: 0.075
[117,   480] loss: 0.048
[117,   510] loss: 0.043
[117,   540] loss: 0.036
[117,   570] loss: 0.045
[117,   600] loss: 0.061
[117,   630] loss: 0.039
[117,   660] loss: 0.061
[117,   690] loss: 0.066
[117,   720] loss: 0.061
[117,   750] loss: 0.042
[117,   780] loss: 0.051
[117,   810] loss: 0.051
[117,   840] loss: 0.041
[117,   870] loss: 0.040
[117,   900] loss: 0.044
[117,   930] loss: 0.035
[117,   960] loss: 0.059
[117,   990] loss: 0.056
[117,  1020] loss: 0.055
[117,  1050] loss: 0.089
[117,  1080] loss: 0.031
[117,  1110] loss: 0.071
[117,  1140] loss: 0.053
[117,  1170] loss: 0.040
[117,  1200] loss: 0.084
[117,  1230] loss: 0.097
```

```
[117, 1260] loss: 0.059
[117, 1290] loss: 0.043
[117, 1320] loss: 0.035
[117, 1350] loss: 0.042
[117, 1380] loss: 0.031
[117, 1410] loss: 0.045
[117, 1440] loss: 0.071
[117, 1470] loss: 0.058
[117, 1500] loss: 0.047
[117, 1530] loss: 0.069
[117, 1560] loss: 0.051
[117, 1590] loss: 0.049
[117, 1620] loss: 0.038
[117, 1650] loss: 0.061
[117, 1680] loss: 0.041
[117, 1710] loss: 0.070
[117, 1740] loss: 0.053
[117, 1770] loss: 0.070
[117, 1800] loss: 0.077
[117, 1830] loss: 0.062
[117, 1860] loss: 0.057
[117, 1890] loss: 0.057
[117, 1920] loss: 0.042
[117, 1950] loss: 0.047
[117, 1980] loss: 0.034
[117, 2010] loss: 0.063
[117, 2040] loss: 0.059
[117, 2070] loss: 0.059
[117, 2100] loss: 0.052
[117, 2130] loss: 0.040
[117, 2160] loss: 0.071
[117, 2190] loss: 0.057
eval intermediate_models/epoch116_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.986891
epoch 116 accuracy on train set is: 0.9868911304980007
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	NaN	NaN	NaN

117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.994638

epoch 116 accuracy on test set is: 0.9946383133406034

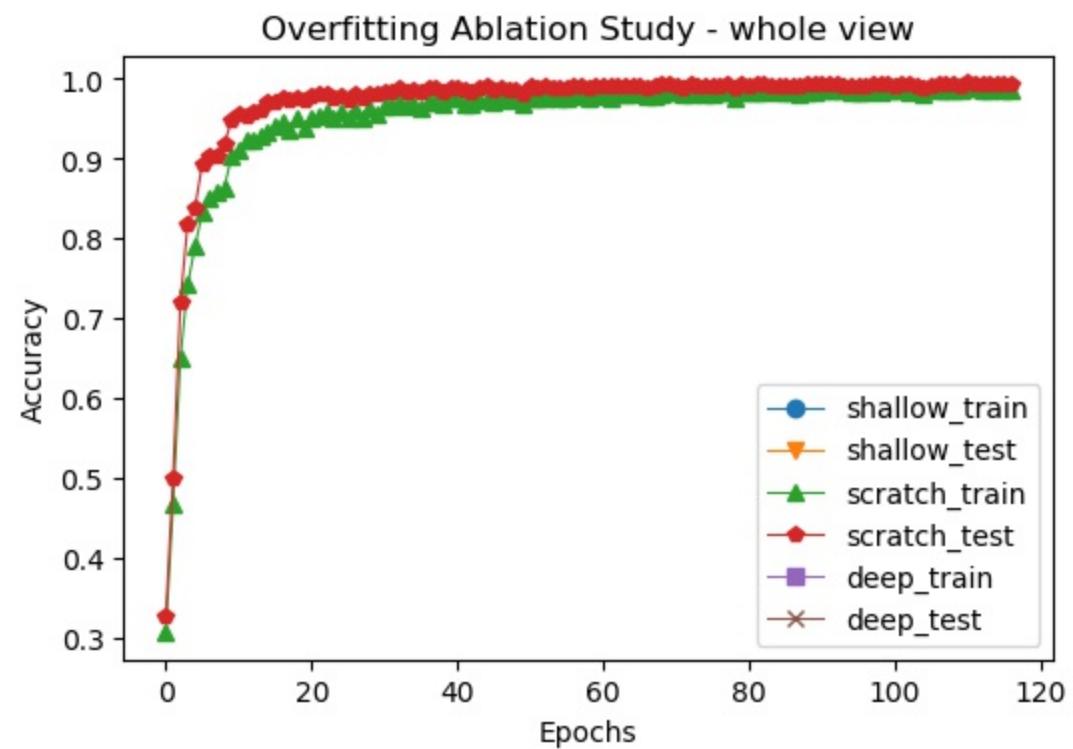
evaluating on test set takes 16 seconds.

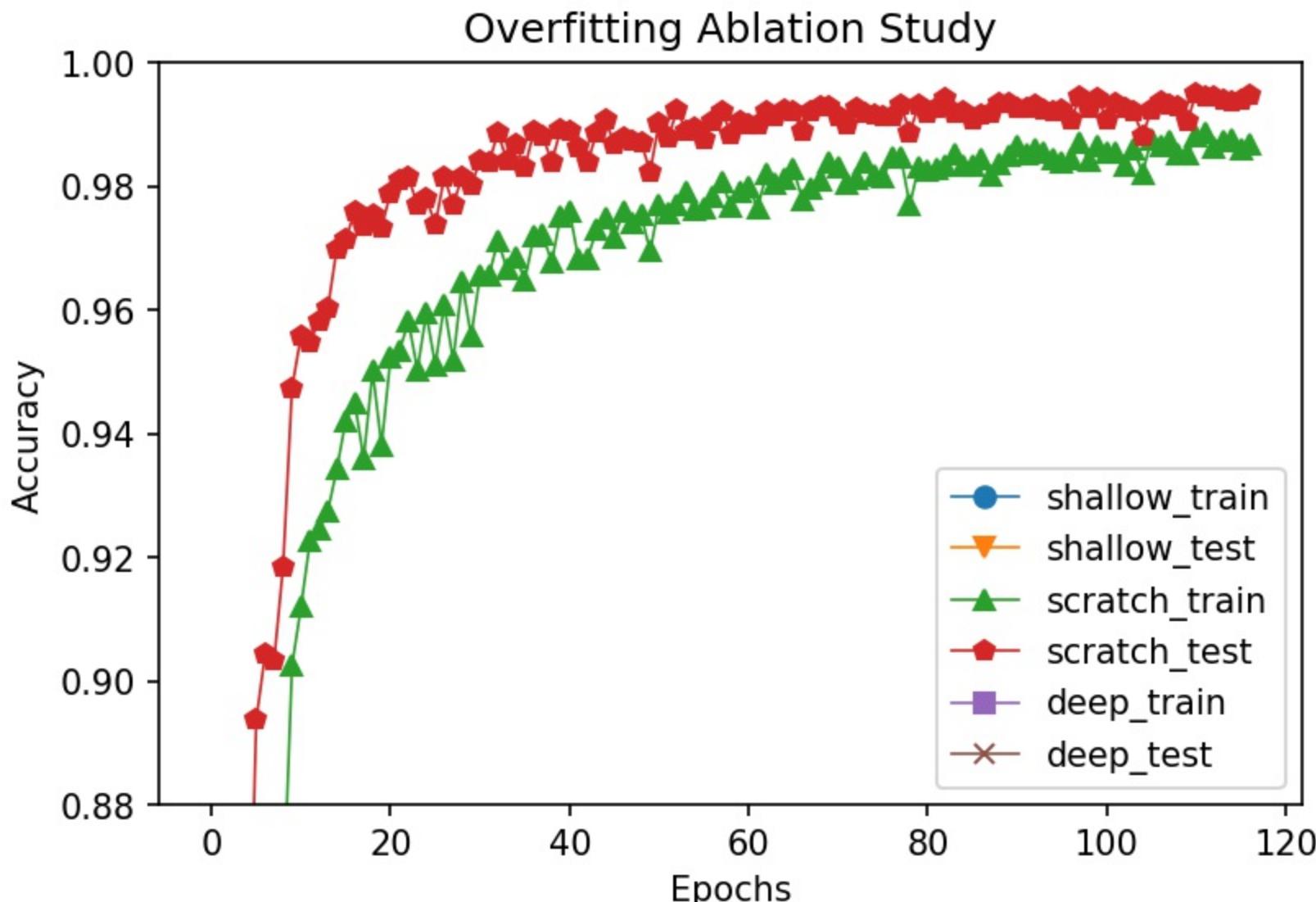
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	0.994638	NaN	NaN
117	NaN	NaN	NaN	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN





epoch 116 takes 316 seconds.

[118, 30] loss: 0.090

[118, 60] loss: 0.061

```
[118,    90] loss: 0.038
[118,   120] loss: 0.046
[118,   150] loss: 0.052
[118,   180] loss: 0.043
[118,   210] loss: 0.038
[118,   240] loss: 0.058
[118,   270] loss: 0.049
[118,   300] loss: 0.060
[118,   330] loss: 0.072
[118,   360] loss: 0.042
[118,   390] loss: 0.043
[118,   420] loss: 0.082
[118,   450] loss: 0.046
[118,   480] loss: 0.063
[118,   510] loss: 0.049
[118,   540] loss: 0.067
[118,   570] loss: 0.049
[118,   600] loss: 0.072
[118,   630] loss: 0.048
[118,   660] loss: 0.048
[118,   690] loss: 0.046
[118,   720] loss: 0.035
[118,   750] loss: 0.086
[118,   780] loss: 0.075
[118,   810] loss: 0.055
[118,   840] loss: 0.033
[118,   870] loss: 0.068
[118,   900] loss: 0.053
[118,   930] loss: 0.051
[118,   960] loss: 0.080
[118,   990] loss: 0.062
[118,  1020] loss: 0.066
[118,  1050] loss: 0.060
[118,  1080] loss: 0.056
[118,  1110] loss: 0.061
[118,  1140] loss: 0.041
[118,  1170] loss: 0.037
[118,  1200] loss: 0.051
[118,  1230] loss: 0.037
```

```
[118, 1260] loss: 0.047
[118, 1290] loss: 0.056
[118, 1320] loss: 0.048
[118, 1350] loss: 0.025
[118, 1380] loss: 0.044
[118, 1410] loss: 0.081
[118, 1440] loss: 0.053
[118, 1470] loss: 0.081
[118, 1500] loss: 0.047
[118, 1530] loss: 0.039
[118, 1560] loss: 0.075
[118, 1590] loss: 0.051
[118, 1620] loss: 0.027
[118, 1650] loss: 0.056
[118, 1680] loss: 0.063
[118, 1710] loss: 0.044
[118, 1740] loss: 0.041
[118, 1770] loss: 0.067
[118, 1800] loss: 0.046
[118, 1830] loss: 0.040
[118, 1860] loss: 0.074
[118, 1890] loss: 0.034
[118, 1920] loss: 0.046
[118, 1950] loss: 0.042
[118, 1980] loss: 0.039
[118, 2010] loss: 0.069
[118, 2040] loss: 0.043
[118, 2070] loss: 0.051
[118, 2100] loss: 0.032
[118, 2130] loss: 0.050
[118, 2160] loss: 0.038
[118, 2190] loss: 0.048
eval intermediate_models/epoch117_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.985710
epoch 117 accuracy on train set is: 0.985709741912032
evaluating on training set takes 65 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	0.994638	NaN	NaN

117	NaN	NaN	0.98571	NaN	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993820

epoch 117 accuracy on test set is: 0.9938204289349327

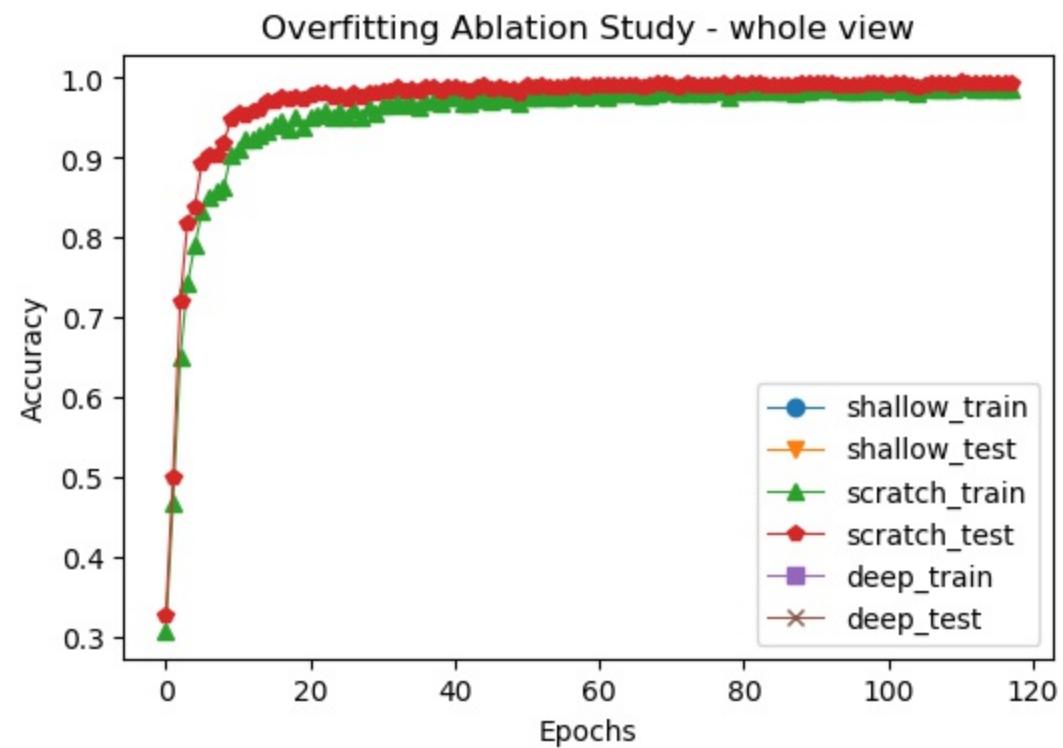
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

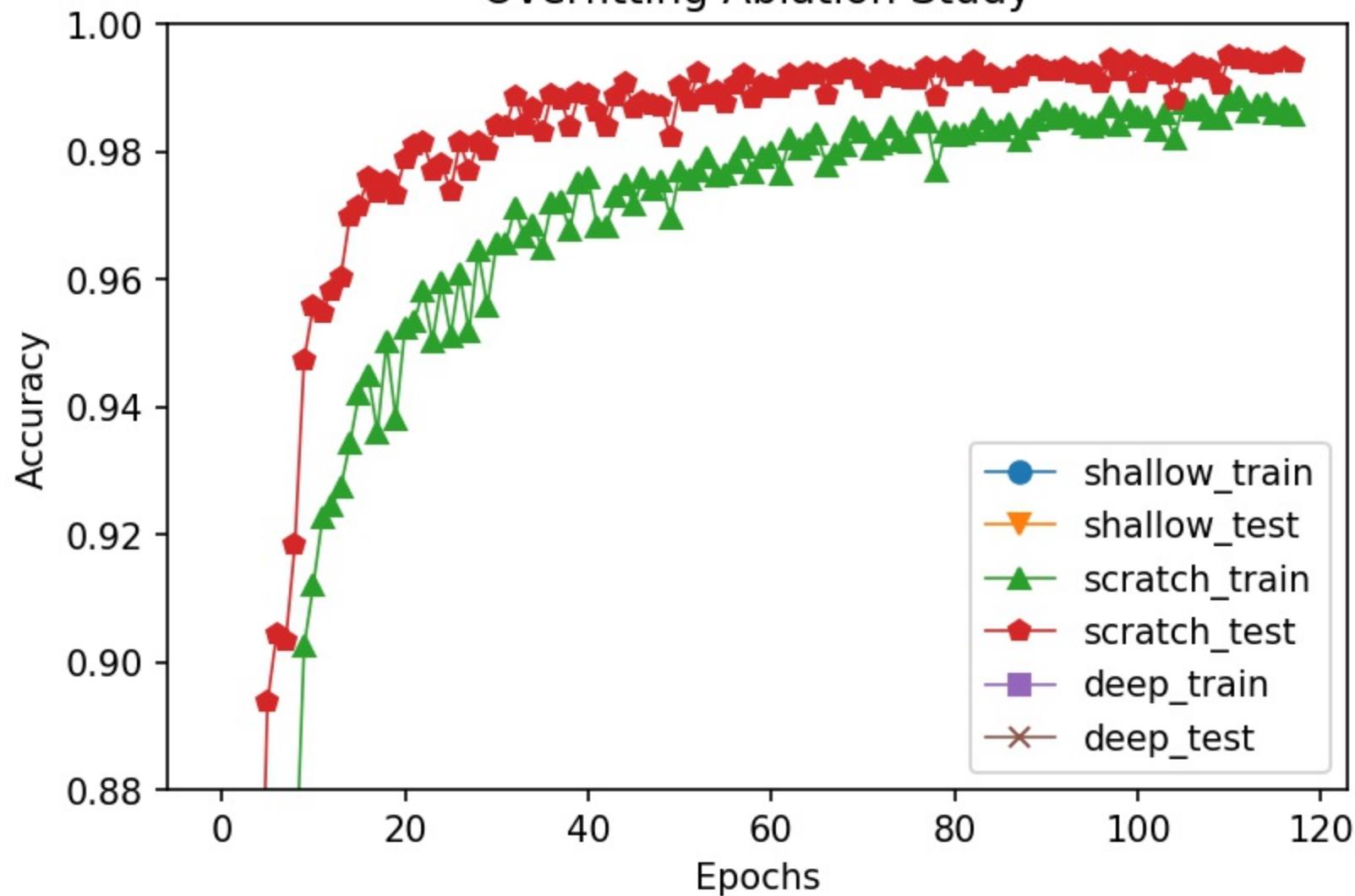
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	0.994638	NaN	NaN
117	NaN	NaN	0.98571	0.99382	NaN	NaN
118	NaN	NaN	NaN	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



```
epoch 117 takes 316 seconds.  
[119,      30] loss: 0.039  
[119,      60] loss: 0.039
```

```
[119,    90] loss: 0.102
[119,   120] loss: 0.087
[119,   150] loss: 0.041
[119,   180] loss: 0.044
[119,   210] loss: 0.063
[119,   240] loss: 0.033
[119,   270] loss: 0.045
[119,   300] loss: 0.041
[119,   330] loss: 0.036
[119,   360] loss: 0.030
[119,   390] loss: 0.041
[119,   420] loss: 0.061
[119,   450] loss: 0.047
[119,   480] loss: 0.045
[119,   510] loss: 0.060
[119,   540] loss: 0.027
[119,   570] loss: 0.069
[119,   600] loss: 0.019
[119,   630] loss: 0.039
[119,   660] loss: 0.063
[119,   690] loss: 0.044
[119,   720] loss: 0.061
[119,   750] loss: 0.068
[119,   780] loss: 0.061
[119,   810] loss: 0.057
[119,   840] loss: 0.079
[119,   870] loss: 0.051
[119,   900] loss: 0.052
[119,   930] loss: 0.082
[119,   960] loss: 0.046
[119,   990] loss: 0.055
[119,  1020] loss: 0.036
[119,  1050] loss: 0.040
[119,  1080] loss: 0.049
[119,  1110] loss: 0.078
[119,  1140] loss: 0.074
[119,  1170] loss: 0.040
[119,  1200] loss: 0.053
[119,  1230] loss: 0.075
```

```
[119, 1260] loss: 0.071
[119, 1290] loss: 0.055
[119, 1320] loss: 0.037
[119, 1350] loss: 0.064
[119, 1380] loss: 0.056
[119, 1410] loss: 0.090
[119, 1440] loss: 0.028
[119, 1470] loss: 0.063
[119, 1500] loss: 0.049
[119, 1530] loss: 0.037
[119, 1560] loss: 0.058
[119, 1590] loss: 0.053
[119, 1620] loss: 0.039
[119, 1650] loss: 0.055
[119, 1680] loss: 0.079
[119, 1710] loss: 0.069
[119, 1740] loss: 0.043
[119, 1770] loss: 0.065
[119, 1800] loss: 0.050
[119, 1830] loss: 0.052
[119, 1860] loss: 0.060
[119, 1890] loss: 0.073
[119, 1920] loss: 0.040
[119, 1950] loss: 0.053
[119, 1980] loss: 0.081
[119, 2010] loss: 0.050
[119, 2040] loss: 0.055
[119, 2070] loss: 0.053
[119, 2100] loss: 0.040
[119, 2130] loss: 0.036
[119, 2160] loss: 0.091
[119, 2190] loss: 0.046
eval intermediate_models/epoch118_scratch.pt saved.
evaluating on training set...
Accuracy on test images: 0.986346
epoch 118 accuracy on train set is: 0.9863458742275536
evaluating on training set takes 64 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	0.994638	NaN	NaN

117	NaN	NaN	0.98571	0.99382	NaN	NaN
118	NaN	NaN	0.986346	NaN	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993639

epoch 118 accuracy on test set is: 0.9936386768447837

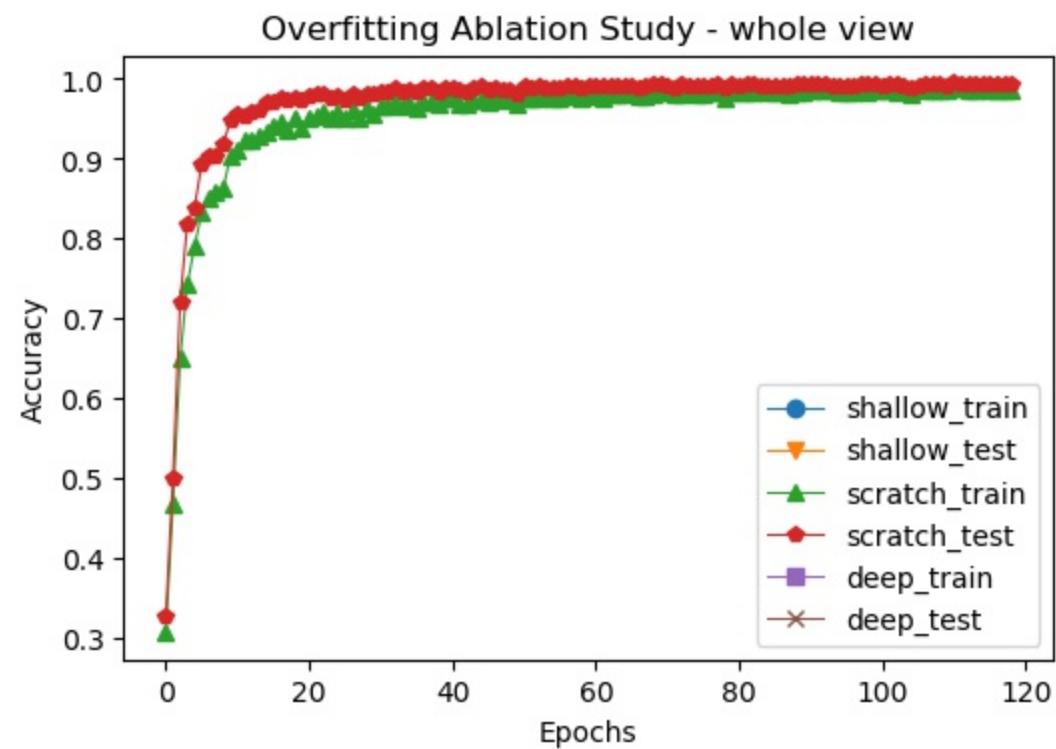
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

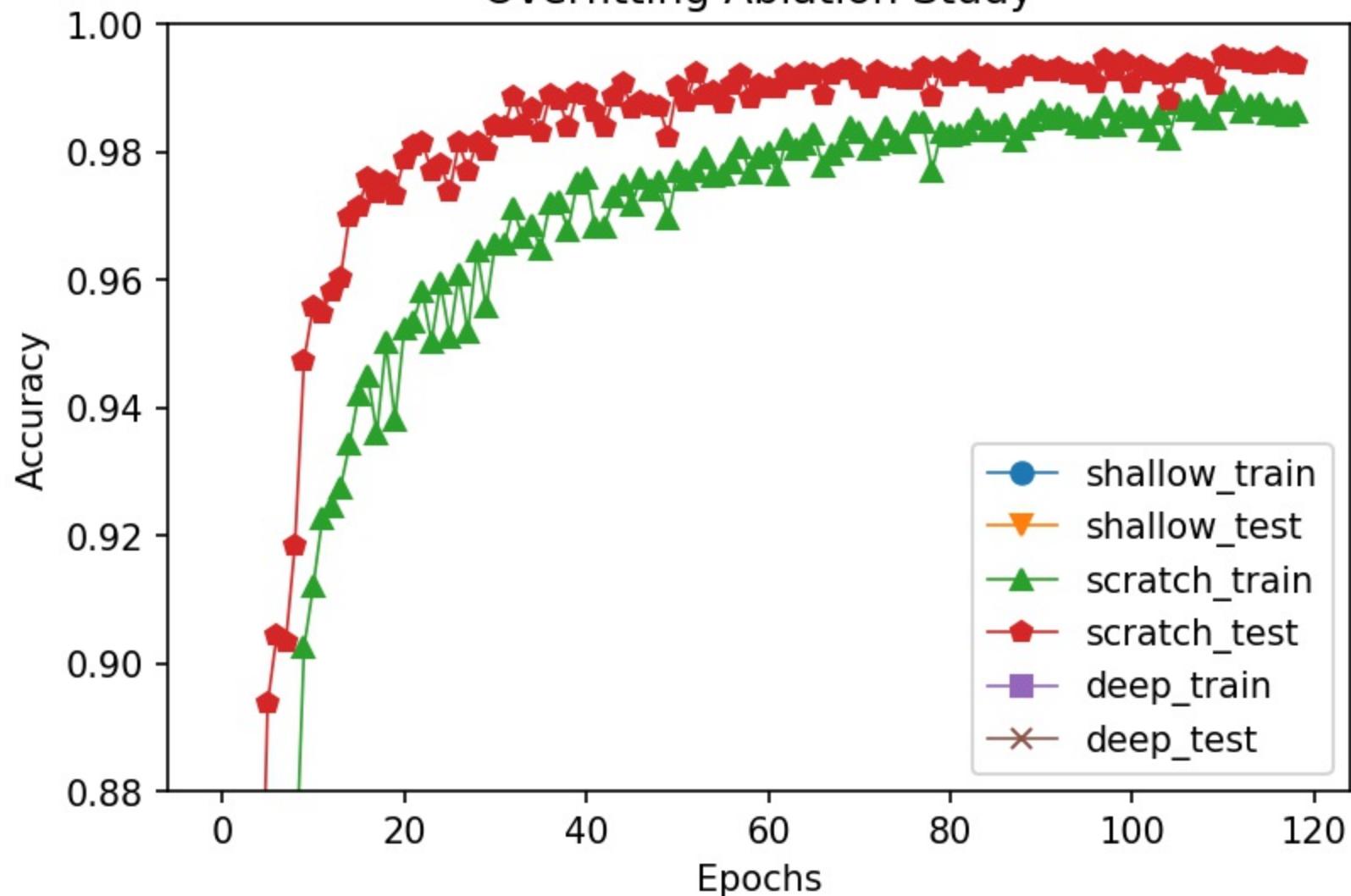
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	0.994638	NaN	NaN
117	NaN	NaN	0.98571	0.99382	NaN	NaN
118	NaN	NaN	0.986346	0.993639	NaN	NaN
119	NaN	NaN	NaN	NaN	NaN	NaN



## Overfitting Ablation Study



```
epoch 118 takes 317 seconds.  
[120,      30] loss: 0.061  
[120,      60] loss: 0.056
```

```
[120,    90] loss: 0.067
[120,   120] loss: 0.059
[120,   150] loss: 0.090
[120,   180] loss: 0.043
[120,   210] loss: 0.036
[120,   240] loss: 0.041
[120,   270] loss: 0.038
[120,   300] loss: 0.046
[120,   330] loss: 0.046
[120,   360] loss: 0.100
[120,   390] loss: 0.039
[120,   420] loss: 0.055
[120,   450] loss: 0.067
[120,   480] loss: 0.064
[120,   510] loss: 0.052
[120,   540] loss: 0.040
[120,   570] loss: 0.050
[120,   600] loss: 0.058
[120,   630] loss: 0.065
[120,   660] loss: 0.053
[120,   690] loss: 0.052
[120,   720] loss: 0.025
[120,   750] loss: 0.056
[120,   780] loss: 0.060
[120,   810] loss: 0.066
[120,   840] loss: 0.096
[120,   870] loss: 0.070
[120,   900] loss: 0.068
[120,   930] loss: 0.045
[120,   960] loss: 0.074
[120,   990] loss: 0.082
[120,  1020] loss: 0.082
[120,  1050] loss: 0.038
[120,  1080] loss: 0.056
[120,  1110] loss: 0.077
[120,  1140] loss: 0.048
[120,  1170] loss: 0.041
[120,  1200] loss: 0.067
[120,  1230] loss: 0.055
```

```
[120, 1260] loss: 0.067
[120, 1290] loss: 0.062
[120, 1320] loss: 0.048
[120, 1350] loss: 0.037
[120, 1380] loss: 0.039
[120, 1410] loss: 0.070
[120, 1440] loss: 0.032
[120, 1470] loss: 0.067
[120, 1500] loss: 0.053
[120, 1530] loss: 0.042
[120, 1560] loss: 0.065
[120, 1590] loss: 0.042
[120, 1620] loss: 0.039
[120, 1650] loss: 0.059
[120, 1680] loss: 0.049
[120, 1710] loss: 0.048
[120, 1740] loss: 0.061
[120, 1770] loss: 0.070
[120, 1800] loss: 0.049
[120, 1830] loss: 0.043
[120, 1860] loss: 0.057
[120, 1890] loss: 0.046
[120, 1920] loss: 0.035
[120, 1950] loss: 0.051
[120, 1980] loss: 0.043
[120, 2010] loss: 0.063
[120, 2040] loss: 0.052
[120, 2070] loss: 0.063
[120, 2100] loss: 0.083
[120, 2130] loss: 0.053
[120, 2160] loss: 0.043
[120, 2190] loss: 0.017
eval intermediate_models/epoch119_scratch.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.988322
epoch 119 accuracy on train set is: 0.9883224282079244
evaluating on training set takes 63 seconds.
    shallow_train shallow_test scratch_train scratch_test deep_train deep_test
```

0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN

39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN
70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN

78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN
109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	0.994638	NaN	NaN

117	NaN	NaN	0.98571	0.99382	NaN	NaN
118	NaN	NaN	0.986346	0.993639	NaN	NaN
119	NaN	NaN	0.988322	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.994820

epoch 119 accuracy on test set is: 0.9948200654307524

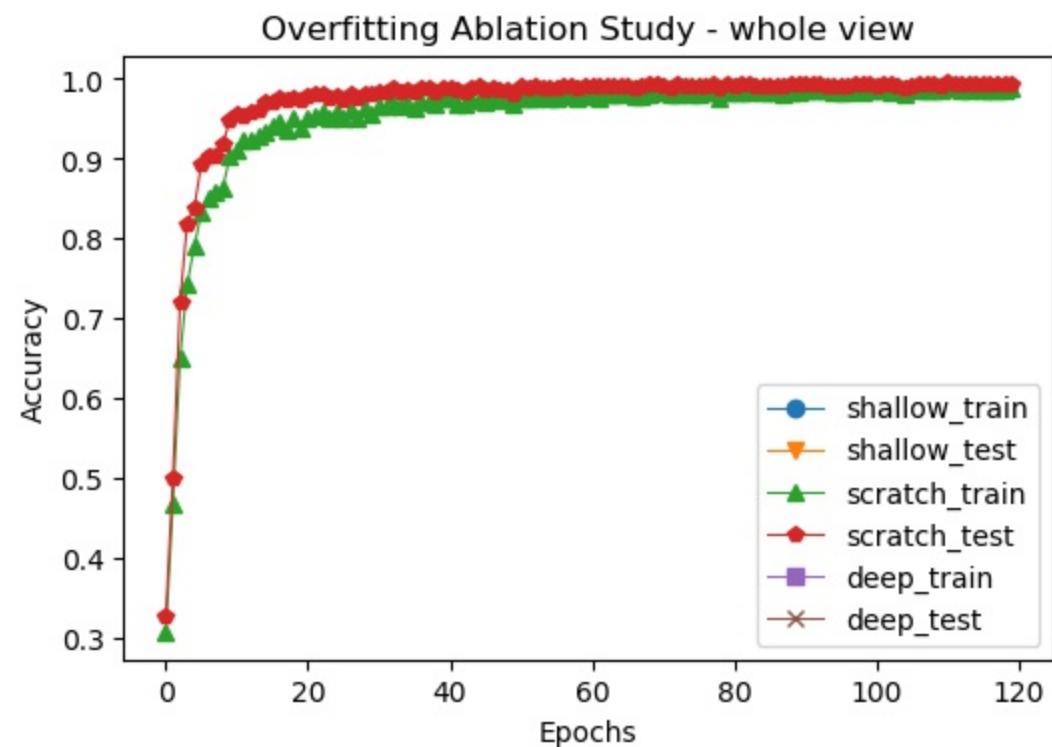
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
0	NaN	NaN	0.307297	0.32679	NaN	NaN
1	NaN	NaN	0.467444	0.501363	NaN	NaN
2	NaN	NaN	0.6509	0.720738	NaN	NaN
3	NaN	NaN	0.74248	0.819066	NaN	NaN
4	NaN	NaN	0.790008	0.838332	NaN	NaN
5	NaN	NaN	0.832652	0.893675	NaN	NaN
6	NaN	NaN	0.851531	0.904398	NaN	NaN
7	NaN	NaN	0.857256	0.903308	NaN	NaN
8	NaN	NaN	0.86305	0.918484	NaN	NaN
9	NaN	NaN	0.902467	0.947383	NaN	NaN
10	NaN	NaN	0.911964	0.955834	NaN	NaN
11	NaN	NaN	0.922574	0.954653	NaN	NaN
12	NaN	NaN	0.924414	0.958288	NaN	NaN
13	NaN	NaN	0.927435	0.960287	NaN	NaN
14	NaN	NaN	0.934251	0.96992	NaN	NaN
15	NaN	NaN	0.942112	0.971374	NaN	NaN
16	NaN	NaN	0.94502	0.976009	NaN	NaN
17	NaN	NaN	0.935955	0.973555	NaN	NaN
18	NaN	NaN	0.950109	0.975282	NaN	NaN
19	NaN	NaN	0.938023	0.973282	NaN	NaN
20	NaN	NaN	0.952245	0.978917	NaN	NaN
21	NaN	NaN	0.953335	0.980825	NaN	NaN
22	NaN	NaN	0.958152	0.981461	NaN	NaN
23	NaN	NaN	0.950314	0.976917	NaN	NaN
24	NaN	NaN	0.959378	0.978099	NaN	NaN
25	NaN	NaN	0.951086	0.973737	NaN	NaN
26	NaN	NaN	0.960742	0.981461	NaN	NaN
27	NaN	NaN	0.951768	0.977008	NaN	NaN
28	NaN	NaN	0.96449	0.981552	NaN	NaN
29	NaN	NaN	0.955675	0.980098	NaN	NaN
30	NaN	NaN	0.965581	0.984097	NaN	NaN

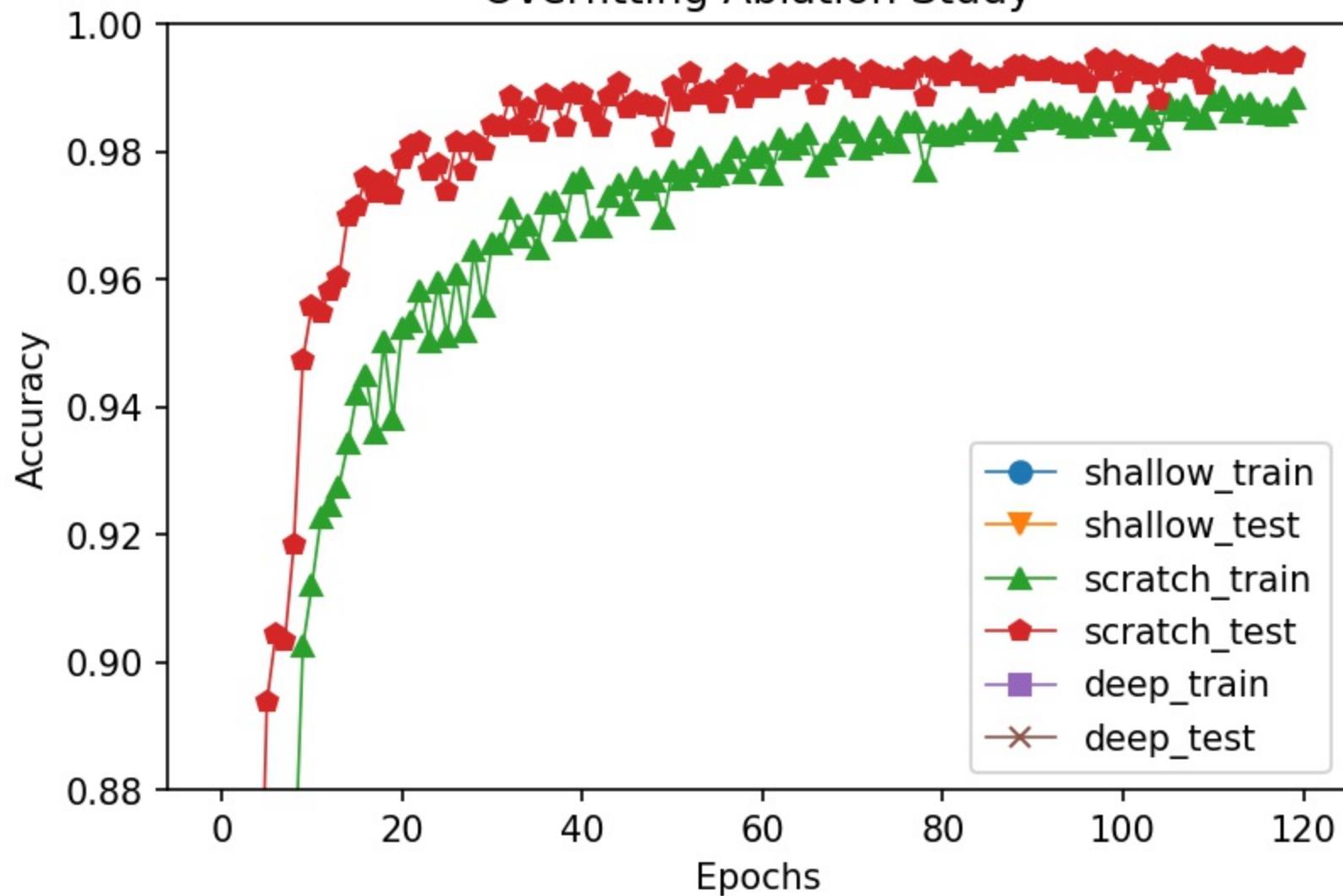
31	NaN	NaN	0.965535	0.983733	NaN	NaN
32	NaN	NaN	0.971192	0.988731	NaN	NaN
33	NaN	NaN	0.966603	0.984006	NaN	NaN
34	NaN	NaN	0.968534	0.986732	NaN	NaN
35	NaN	NaN	0.964786	0.983006	NaN	NaN
36	NaN	NaN	0.971828	0.988822	NaN	NaN
37	NaN	NaN	0.972192	0.988095	NaN	NaN
38	NaN	NaN	0.967671	0.983915	NaN	NaN
39	NaN	NaN	0.975168	0.989095	NaN	NaN
40	NaN	NaN	0.975941	0.988822	NaN	NaN
41	NaN	NaN	0.968125	0.986278	NaN	NaN
42	NaN	NaN	0.968307	0.983733	NaN	NaN
43	NaN	NaN	0.972896	0.98855	NaN	NaN
44	NaN	NaN	0.974895	0.990731	NaN	NaN
45	NaN	NaN	0.971669	0.986823	NaN	NaN
46	NaN	NaN	0.975941	0.987732	NaN	NaN
47	NaN	NaN	0.974123	0.987186	NaN	NaN
48	NaN	NaN	0.975304	0.987005	NaN	NaN
49	NaN	NaN	0.969443	0.98237	NaN	NaN
50	NaN	NaN	0.977008	0.990276	NaN	NaN
51	NaN	NaN	0.975782	0.987732	NaN	NaN
52	NaN	NaN	0.977031	0.992276	NaN	NaN
53	NaN	NaN	0.979053	0.989004	NaN	NaN
54	NaN	NaN	0.9761	0.989549	NaN	NaN
55	NaN	NaN	0.976509	0.987641	NaN	NaN
56	NaN	NaN	0.97819	0.990549	NaN	NaN
57	NaN	NaN	0.980802	0.992185	NaN	NaN
58	NaN	NaN	0.976713	0.988459	NaN	NaN
59	NaN	NaN	0.979099	0.990458	NaN	NaN
60	NaN	NaN	0.979939	0.990004	NaN	NaN
61	NaN	NaN	0.97644	0.990004	NaN	NaN
62	NaN	NaN	0.982075	0.992003	NaN	NaN
63	NaN	NaN	0.980348	0.991367	NaN	NaN
64	NaN	NaN	0.981325	0.992457	NaN	NaN
65	NaN	NaN	0.982734	0.992003	NaN	NaN
66	NaN	NaN	0.977849	0.988822	NaN	NaN
67	NaN	NaN	0.979621	0.992185	NaN	NaN
68	NaN	NaN	0.981007	0.992912	NaN	NaN
69	NaN	NaN	0.983915	0.992821	NaN	NaN

70	NaN	NaN	0.982938	0.991185	NaN	NaN
71	NaN	NaN	0.980371	0.989913	NaN	NaN
72	NaN	NaN	0.981302	0.992639	NaN	NaN
73	NaN	NaN	0.983938	0.991821	NaN	NaN
74	NaN	NaN	0.981711	0.991549	NaN	NaN
75	NaN	NaN	0.981416	0.991276	NaN	NaN
76	NaN	NaN	0.984642	0.991367	NaN	NaN
77	NaN	NaN	0.984642	0.993184	NaN	NaN
78	NaN	NaN	0.976986	0.98864	NaN	NaN
79	NaN	NaN	0.98312	0.993184	NaN	NaN
80	NaN	NaN	0.982529	0.991821	NaN	NaN
81	NaN	NaN	0.982711	0.99273	NaN	NaN
82	NaN	NaN	0.983279	0.994184	NaN	NaN
83	NaN	NaN	0.985074	0.99173	NaN	NaN
84	NaN	NaN	0.983415	0.992185	NaN	NaN
85	NaN	NaN	0.983279	0.990822	NaN	NaN
86	NaN	NaN	0.984301	0.991458	NaN	NaN
87	NaN	NaN	0.981779	0.99173	NaN	NaN
88	NaN	NaN	0.983688	0.993275	NaN	NaN
89	NaN	NaN	0.984824	0.993275	NaN	NaN
90	NaN	NaN	0.986619	0.992639	NaN	NaN
91	NaN	NaN	0.985119	0.99273	NaN	NaN
92	NaN	NaN	0.986005	0.993003	NaN	NaN
93	NaN	NaN	0.985392	0.992276	NaN	NaN
94	NaN	NaN	0.984392	0.992185	NaN	NaN
95	NaN	NaN	0.983756	0.992457	NaN	NaN
96	NaN	NaN	0.984324	0.990822	NaN	NaN
97	NaN	NaN	0.98705	0.994366	NaN	NaN
98	NaN	NaN	0.984233	0.992639	NaN	NaN
99	NaN	NaN	0.986596	0.994184	NaN	NaN
100	NaN	NaN	0.985414	0.990731	NaN	NaN
101	NaN	NaN	0.985414	0.993366	NaN	NaN
102	NaN	NaN	0.983415	0.992639	NaN	NaN
103	NaN	NaN	0.985982	0.992003	NaN	NaN
104	NaN	NaN	0.98212	0.988186	NaN	NaN
105	NaN	NaN	0.987686	0.992366	NaN	NaN
106	NaN	NaN	0.986573	0.99373	NaN	NaN
107	NaN	NaN	0.987368	0.993093	NaN	NaN
108	NaN	NaN	0.985187	0.992821	NaN	NaN

109	NaN	NaN	0.985119	0.990458	NaN	NaN
110	NaN	NaN	0.988141	0.995093	NaN	NaN
111	NaN	NaN	0.988731	0.994457	NaN	NaN
112	NaN	NaN	0.986187	0.994457	NaN	NaN
113	NaN	NaN	0.987209	0.993911	NaN	NaN
114	NaN	NaN	0.987573	0.993548	NaN	NaN
115	NaN	NaN	0.986005	0.99382	NaN	NaN
116	NaN	NaN	0.986891	0.994638	NaN	NaN
117	NaN	NaN	0.98571	0.99382	NaN	NaN
118	NaN	NaN	0.986346	0.993639	NaN	NaN
119	NaN	NaN	0.988322	0.99482	NaN	NaN



## Overfitting Ablation Study



epoch 119 takes 315 seconds.

Finished Training

Evaluating...

FINISHED

In [2]: