

```

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_continued_train.py'      = '/home/
administrator/Documents/OU/Deep_learning/project/DeepLearning_PlantDiseases/
Scripts'

shallow_train shallow_test scratch_train scratch_test deep_train deep_test
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

Directory ovft_intermediate_models already exists
RETRAINING deep

Targeting alexnet with 39 classes
-----
loading ovft_intermediate_models/epoch39_deep.pt to continue training...
Resizing input images to max of (224, 224)
Transferring models to GPU(s)

```

```
Training...
/home/administrator/anaconda3/envs/torch_nvidia_390/lib/python3.7/site-packages/
torchvision/transforms/transforms.py:562: UserWarning: The use of the [REDACTED]
transforms.RandomSizedCrop transform is deprecated, please use [REDACTED]
transforms.RandomResizedCrop instead.
    warnings.warn("The use of the transforms.RandomSizedCrop transform is
deprecated, " + [REDACTED]
/home/administrator/anaconda3/envs/torch_nvidia_390/lib/python3.7/site-packages/
torchvision/transforms/transforms.py:187: UserWarning: The use of the [REDACTED]
transforms.Scale transform is deprecated, please use transforms.Resize instead.
    warnings.warn("The use of the transforms.Scale transform is deprecated, " +
[41,    30] loss: 0.047
[41,    60] loss: 0.026
[41,    90] loss: 0.049
[41,   120] loss: 0.043
[41,   150] loss: 0.066
[41,   180] loss: 0.057
[41,   210] loss: 0.053
[41,   240] loss: 0.073
[41,   270] loss: 0.038
[41,   300] loss: 0.037
[41,   330] loss: 0.072
[41,   360] loss: 0.053
[41,   390] loss: 0.066
[41,   420] loss: 0.071
[41,   450] loss: 0.055
[41,   480] loss: 0.076
[41,   510] loss: 0.069
[41,   540] loss: 0.071
[41,   570] loss: 0.057
[41,   600] loss: 0.073
[41,   630] loss: 0.055
[41,   660] loss: 0.044
[41,   690] loss: 0.034
[41,   720] loss: 0.074
[41,   750] loss: 0.062
[41,   780] loss: 0.067
[41,   810] loss: 0.048
[41,   840] loss: 0.047
[41,   870] loss: 0.055
[41,   900] loss: 0.060
[41,   930] loss: 0.037
[41,   960] loss: 0.049
[41,   990] loss: 0.038
[41,  1020] loss: 0.069
[41,  1050] loss: 0.044
[41,  1080] loss: 0.050
[41,  1110] loss: 0.063
[41,  1140] loss: 0.052
[41,  1170] loss: 0.054
[41,  1200] loss: 0.067
[41,  1230] loss: 0.044
[41,  1260] loss: 0.060
[41,  1290] loss: 0.029
[41,  1320] loss: 0.062
[41,  1350] loss: 0.066
[41,  1380] loss: 0.086
[41,  1410] loss: 0.062
[41,  1440] loss: 0.046
```

```

[41, 1470] loss: 0.064
[41, 1500] loss: 0.071
[41, 1530] loss: 0.056
[41, 1560] loss: 0.057
[41, 1590] loss: 0.052
[41, 1620] loss: 0.055
[41, 1650] loss: 0.049
[41, 1680] loss: 0.046
[41, 1710] loss: 0.068
[41, 1740] loss: 0.061
[41, 1770] loss: 0.048
[41, 1800] loss: 0.050
[41, 1830] loss: 0.047
[41, 1860] loss: 0.062
[41, 1890] loss: 0.062
[41, 1920] loss: 0.039
[41, 1950] loss: 0.052
[41, 1980] loss: 0.081
[41, 2010] loss: 0.051
[41, 2040] loss: 0.047
[41, 2070] loss: 0.050
[41, 2100] loss: 0.053
[41, 2130] loss: 0.066
[41, 2160] loss: 0.070
[41, 2190] loss: 0.060

```

`ovft_intermediate_models/epoch40_deep.pt saved.`

evaluating on training set...

Accuracy on test images: 0.987027

epoch 40 accuracy on train set is: 0.9870274445656125

evaluating on training set takes 64 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	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

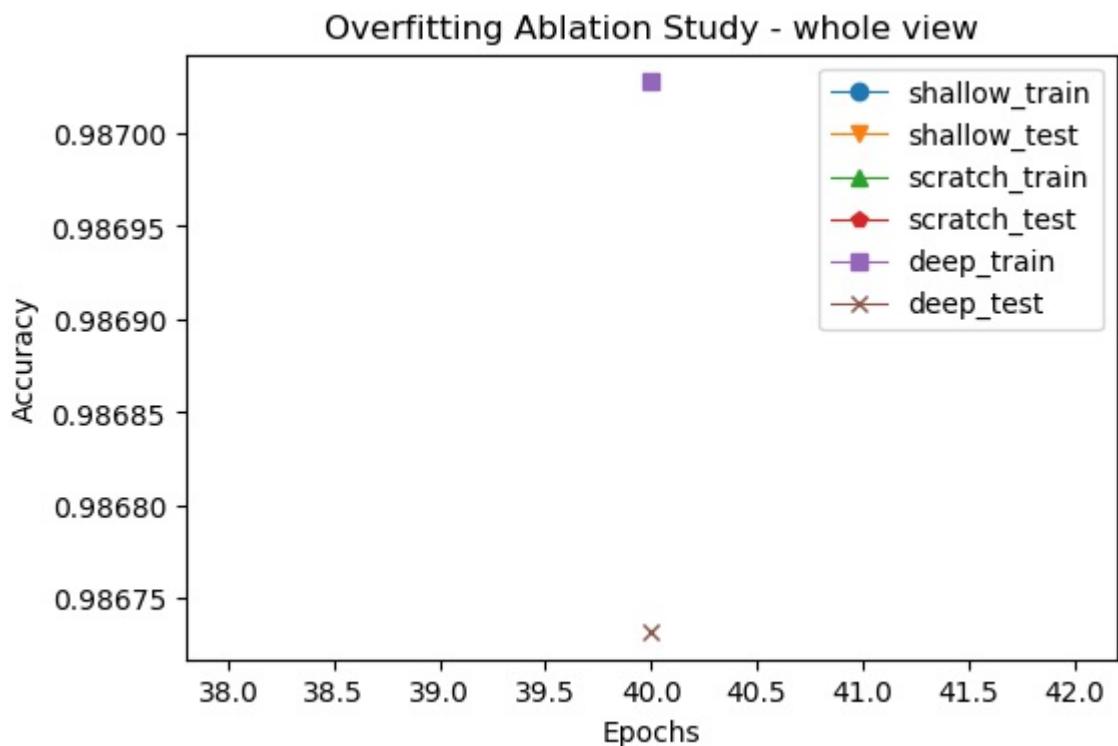
evaluating on test set...

Accuracy on test images: 0.986732

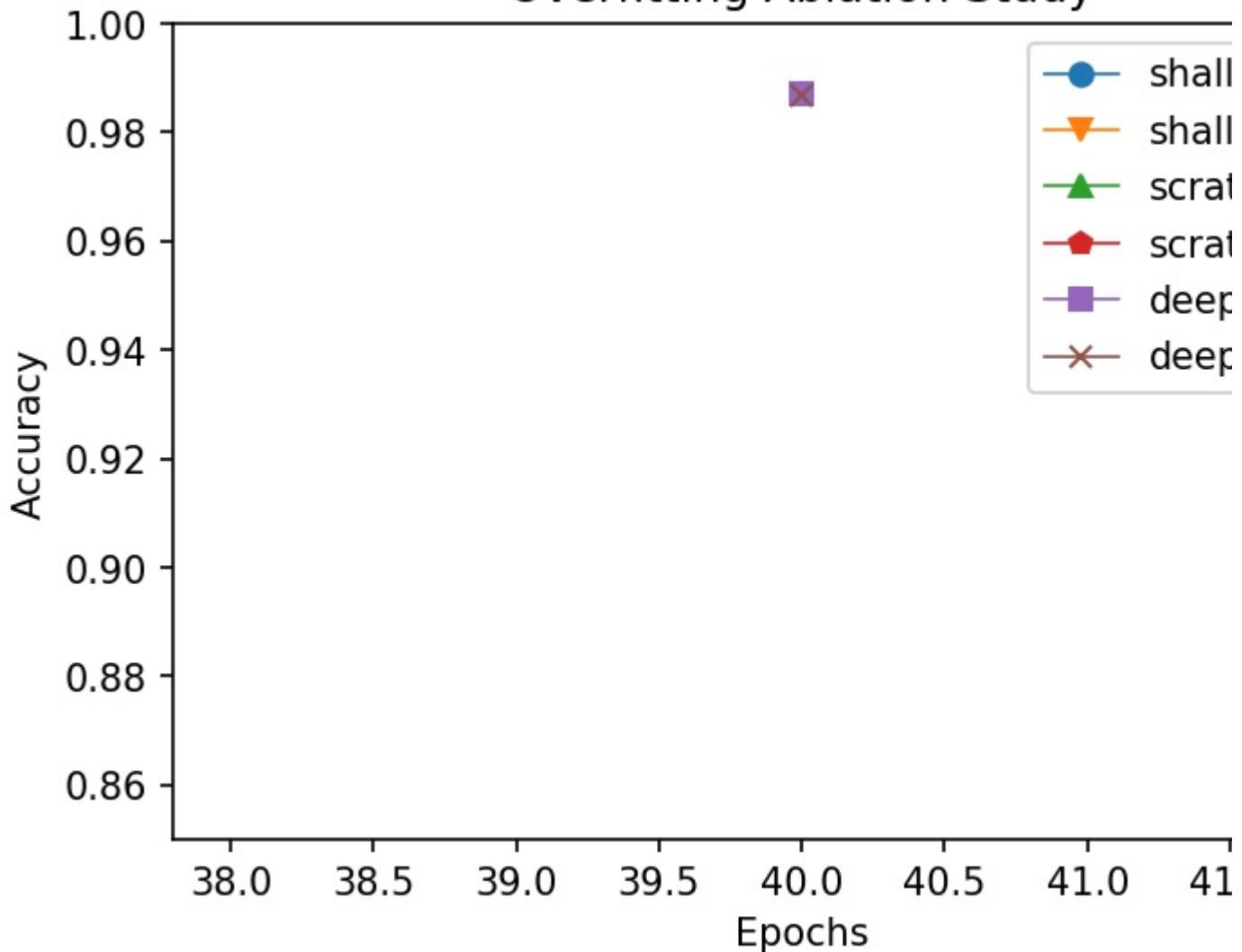
epoch 40 accuracy on test set is: 0.9867320974191203

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
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



Overfitting Ablation Study



```
epoch 40 takes 324 seconds.  
[42,    30] loss: 0.066  
[42,    60] loss: 0.039  
[42,   90] loss: 0.050  
[42,  120] loss: 0.053  
[42,  150] loss: 0.067  
[42,  180] loss: 0.080  
[42,  210] loss: 0.063  
[42,  240] loss: 0.073  
[42,  270] loss: 0.064  
[42,  300] loss: 0.047  
[42,  330] loss: 0.080  
[42,  360] loss: 0.034  
[42,  390] loss: 0.028  
[42,  420] loss: 0.037  
[42,  450] loss: 0.068  
[42,  480] loss: 0.035  
[42,  510] loss: 0.051  
[42,  540] loss: 0.063  
[42,  570] loss: 0.029  
[42,  600] loss: 0.074  
[42,  630] loss: 0.036  
[42,  660] loss: 0.039
```

```

[42, 690] loss: 0.052
[42, 720] loss: 0.036
[42, 750] loss: 0.048
[42, 780] loss: 0.060
[42, 810] loss: 0.041
[42, 840] loss: 0.058
[42, 870] loss: 0.052
[42, 900] loss: 0.094
[42, 930] loss: 0.062
[42, 960] loss: 0.035
[42, 990] loss: 0.048
[42, 1020] loss: 0.078
[42, 1050] loss: 0.061
[42, 1080] loss: 0.076
[42, 1110] loss: 0.067
[42, 1140] loss: 0.061
[42, 1170] loss: 0.055
[42, 1200] loss: 0.027
[42, 1230] loss: 0.045
[42, 1260] loss: 0.083
[42, 1290] loss: 0.052
[42, 1320] loss: 0.042
[42, 1350] loss: 0.061
[42, 1380] loss: 0.020
[42, 1410] loss: 0.047
[42, 1440] loss: 0.047
[42, 1470] loss: 0.034
[42, 1500] loss: 0.070
[42, 1530] loss: 0.053
[42, 1560] loss: 0.098
[42, 1590] loss: 0.060
[42, 1620] loss: 0.055
[42, 1650] loss: 0.044
[42, 1680] loss: 0.056
[42, 1710] loss: 0.048
[42, 1740] loss: 0.070
[42, 1770] loss: 0.050
[42, 1800] loss: 0.059
[42, 1830] loss: 0.058
[42, 1860] loss: 0.045
[42, 1890] loss: 0.066
[42, 1920] loss: 0.057
[42, 1950] loss: 0.050
[42, 1980] loss: 0.106
[42, 2010] loss: 0.077
[42, 2040] loss: 0.082
[42, 2070] loss: 0.052
[42, 2100] loss: 0.040
[42, 2130] loss: 0.032
[42, 2160] loss: 0.045
[42, 2190] loss: 0.042
ovft_intermediate_models/epoch41_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.987709
epoch 41 accuracy on train set is: 0.9877090149036714
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	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

evaluating on test set...

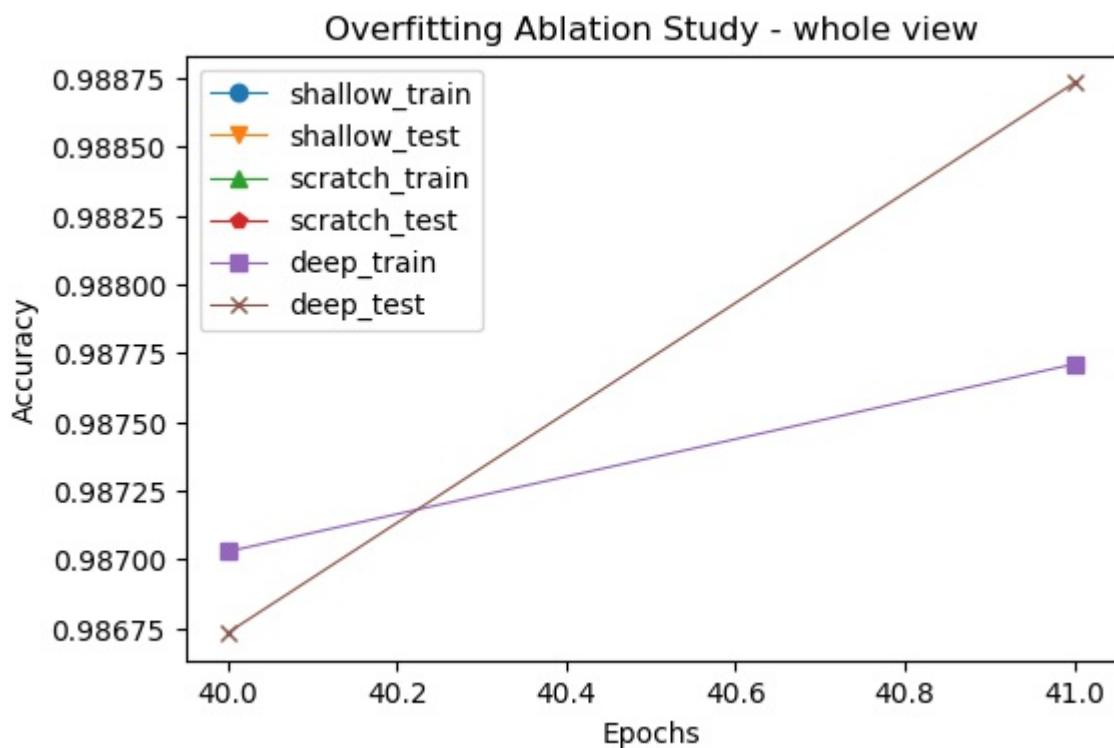
Accuracy on test images: 0.988731

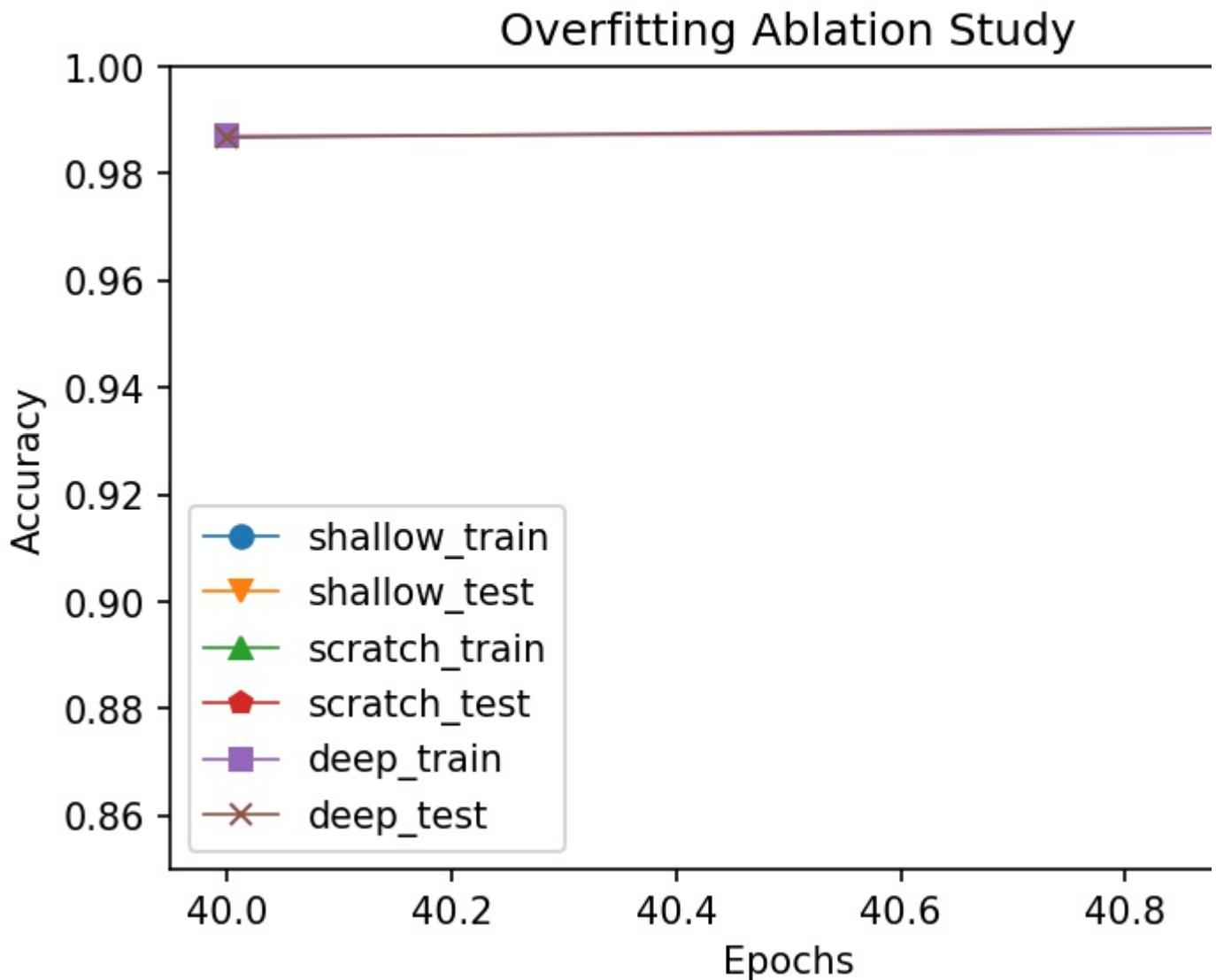
epoch 41 accuracy on test set is: 0.9887313704107598

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
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





```
epoch 41 takes 316 seconds.  
[43,    30] loss: 0.094  
[43,    60] loss: 0.047  
[43,    90] loss: 0.043  
[43,   120] loss: 0.061  
[43,   150] loss: 0.050  
[43,   180] loss: 0.059  
[43,   210] loss: 0.029  
[43,   240] loss: 0.074  
[43,   270] loss: 0.044  
[43,   300] loss: 0.070  
[43,   330] loss: 0.032  
[43,   360] loss: 0.042  
[43,   390] loss: 0.039  
[43,   420] loss: 0.065  
[43,   450] loss: 0.067  
[43,   480] loss: 0.033  
[43,   510] loss: 0.047  
[43,   540] loss: 0.027  
[43,   570] loss: 0.046  
[43,   600] loss: 0.031  
[43,   630] loss: 0.064  
[43,   660] loss: 0.068
```

```

[43, 690] loss: 0.045
[43, 720] loss: 0.056
[43, 750] loss: 0.037
[43, 780] loss: 0.061
[43, 810] loss: 0.057
[43, 840] loss: 0.037
[43, 870] loss: 0.076
[43, 900] loss: 0.036
[43, 930] loss: 0.067
[43, 960] loss: 0.067
[43, 990] loss: 0.067
[43, 1020] loss: 0.045
[43, 1050] loss: 0.034
[43, 1080] loss: 0.046
[43, 1110] loss: 0.055
[43, 1140] loss: 0.057
[43, 1170] loss: 0.045
[43, 1200] loss: 0.036
[43, 1230] loss: 0.083
[43, 1260] loss: 0.060
[43, 1290] loss: 0.055
[43, 1320] loss: 0.044
[43, 1350] loss: 0.040
[43, 1380] loss: 0.072
[43, 1410] loss: 0.058
[43, 1440] loss: 0.044
[43, 1470] loss: 0.087
[43, 1500] loss: 0.055
[43, 1530] loss: 0.059
[43, 1560] loss: 0.055
[43, 1590] loss: 0.039
[43, 1620] loss: 0.067
[43, 1650] loss: 0.079
[43, 1680] loss: 0.064
[43, 1710] loss: 0.039
[43, 1740] loss: 0.030
[43, 1770] loss: 0.035
[43, 1800] loss: 0.055
[43, 1830] loss: 0.047
[43, 1860] loss: 0.052
[43, 1890] loss: 0.051
[43, 1920] loss: 0.050
[43, 1950] loss: 0.038
[43, 1980] loss: 0.077
[43, 2010] loss: 0.109
[43, 2040] loss: 0.043
[43, 2070] loss: 0.073
[43, 2100] loss: 0.053
[43, 2130] loss: 0.045
[43, 2160] loss: 0.067
[43, 2190] loss: 0.098
ovft_intermediate_models/epoch42_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.984551
epoch 42 accuracy on train set is: 0.9845510723373319
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	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

evaluating on test set...

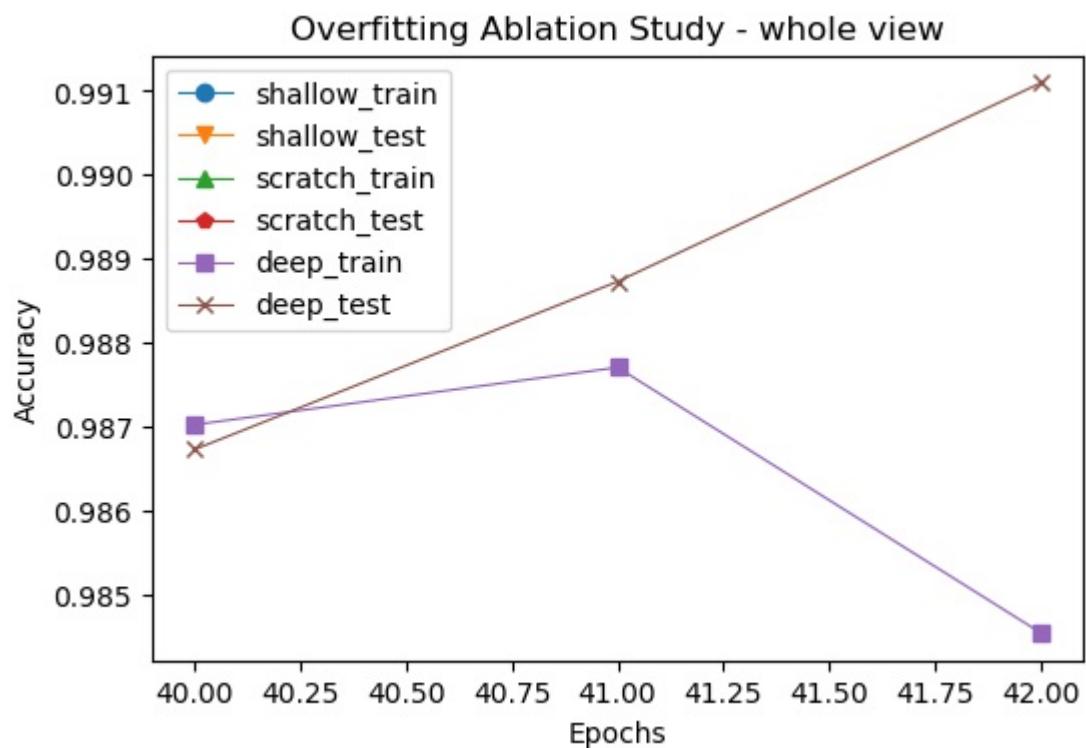
Accuracy on test images: 0.991094

epoch 42 accuracy on test set is: 0.9910941475826972

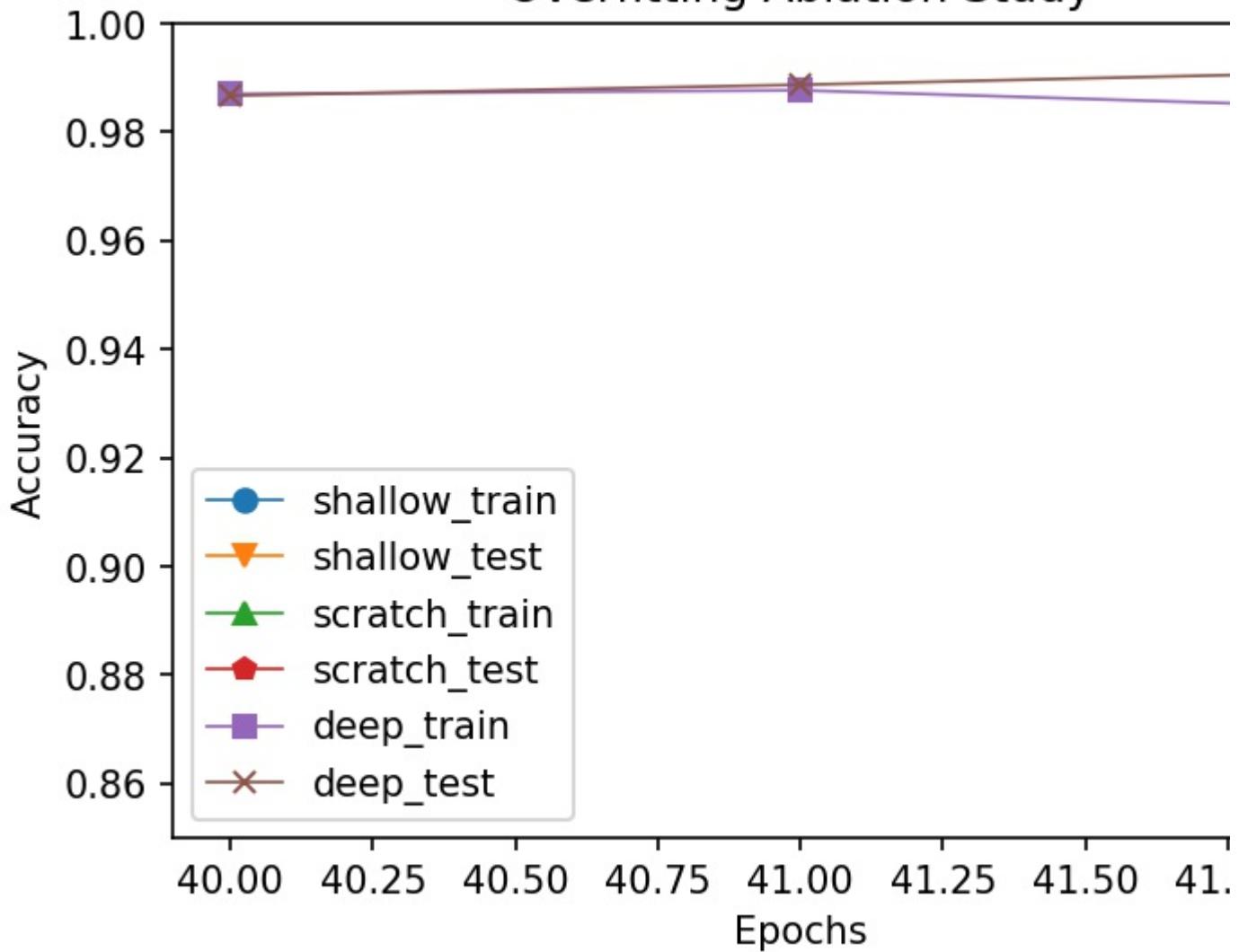
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
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



Overfitting Ablation Study



```
epoch 42 takes 315 seconds.  
[44,     30] loss: 0.066  
[44,     60] loss: 0.068  
[44,    90] loss: 0.079  
[44,   120] loss: 0.062  
[44,   150] loss: 0.034  
[44,   180] loss: 0.044  
[44,   210] loss: 0.049  
[44,   240] loss: 0.048  
[44,   270] loss: 0.041  
[44,   300] loss: 0.066  
[44,   330] loss: 0.039  
[44,   360] loss: 0.050  
[44,   390] loss: 0.038  
[44,   420] loss: 0.053  
[44,   450] loss: 0.023  
[44,   480] loss: 0.051  
[44,   510] loss: 0.052  
[44,   540] loss: 0.041  
[44,   570] loss: 0.073  
[44,   600] loss: 0.028  
[44,   630] loss: 0.024  
[44,   660] loss: 0.048
```

```

[44, 690] loss: 0.095
[44, 720] loss: 0.044
[44, 750] loss: 0.043
[44, 780] loss: 0.071
[44, 810] loss: 0.066
[44, 840] loss: 0.048
[44, 870] loss: 0.045
[44, 900] loss: 0.076
[44, 930] loss: 0.054
[44, 960] loss: 0.050
[44, 990] loss: 0.081
[44, 1020] loss: 0.061
[44, 1050] loss: 0.069
[44, 1080] loss: 0.046
[44, 1110] loss: 0.047
[44, 1140] loss: 0.059
[44, 1170] loss: 0.060
[44, 1200] loss: 0.048
[44, 1230] loss: 0.037
[44, 1260] loss: 0.068
[44, 1290] loss: 0.031
[44, 1320] loss: 0.062
[44, 1350] loss: 0.085
[44, 1380] loss: 0.036
[44, 1410] loss: 0.033
[44, 1440] loss: 0.033
[44, 1470] loss: 0.060
[44, 1500] loss: 0.035
[44, 1530] loss: 0.041
[44, 1560] loss: 0.061
[44, 1590] loss: 0.040
[44, 1620] loss: 0.078
[44, 1650] loss: 0.052
[44, 1680] loss: 0.037
[44, 1710] loss: 0.036
[44, 1740] loss: 0.039
[44, 1770] loss: 0.066
[44, 1800] loss: 0.112
[44, 1830] loss: 0.071
[44, 1860] loss: 0.035
[44, 1890] loss: 0.049
[44, 1920] loss: 0.021
[44, 1950] loss: 0.060
[44, 1980] loss: 0.077
[44, 2010] loss: 0.073
[44, 2040] loss: 0.056
[44, 2070] loss: 0.043
[44, 2100] loss: 0.030
[44, 2130] loss: 0.086
[44, 2160] loss: 0.033
[44, 2190] loss: 0.066
ovft_intermediate_models/epoch43_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.986255
epoch 43 accuracy on train set is: 0.9862549981824791
evaluating on training set takes 63 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	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

evaluating on test set...

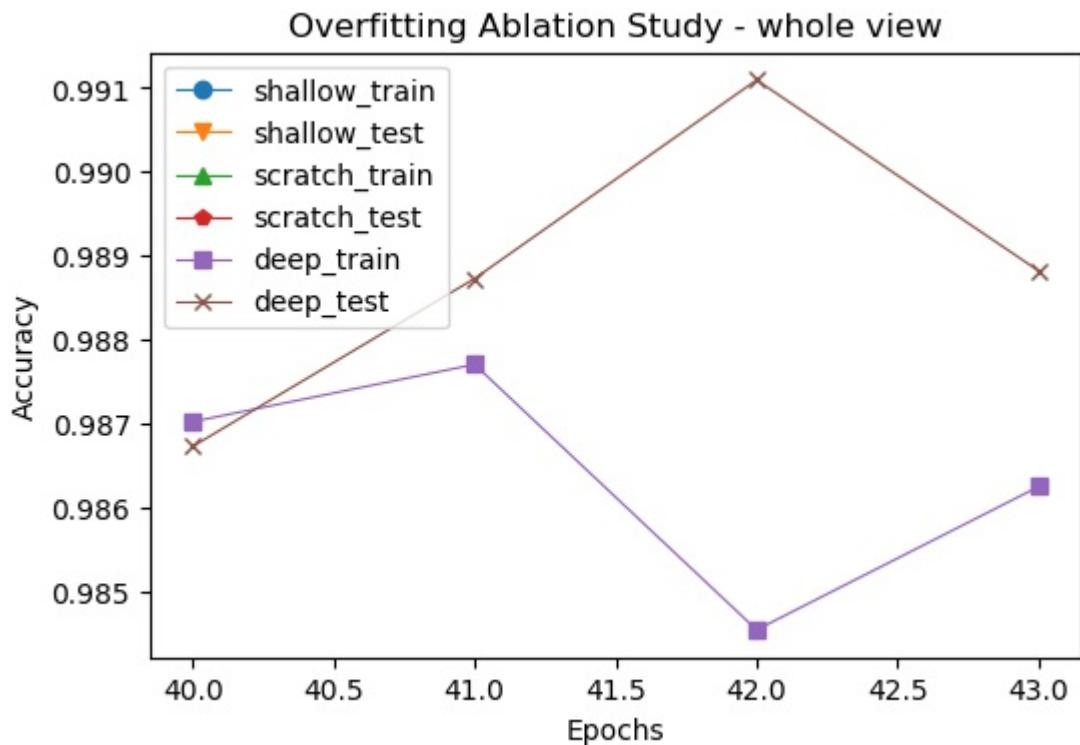
Accuracy on test images: 0.988822

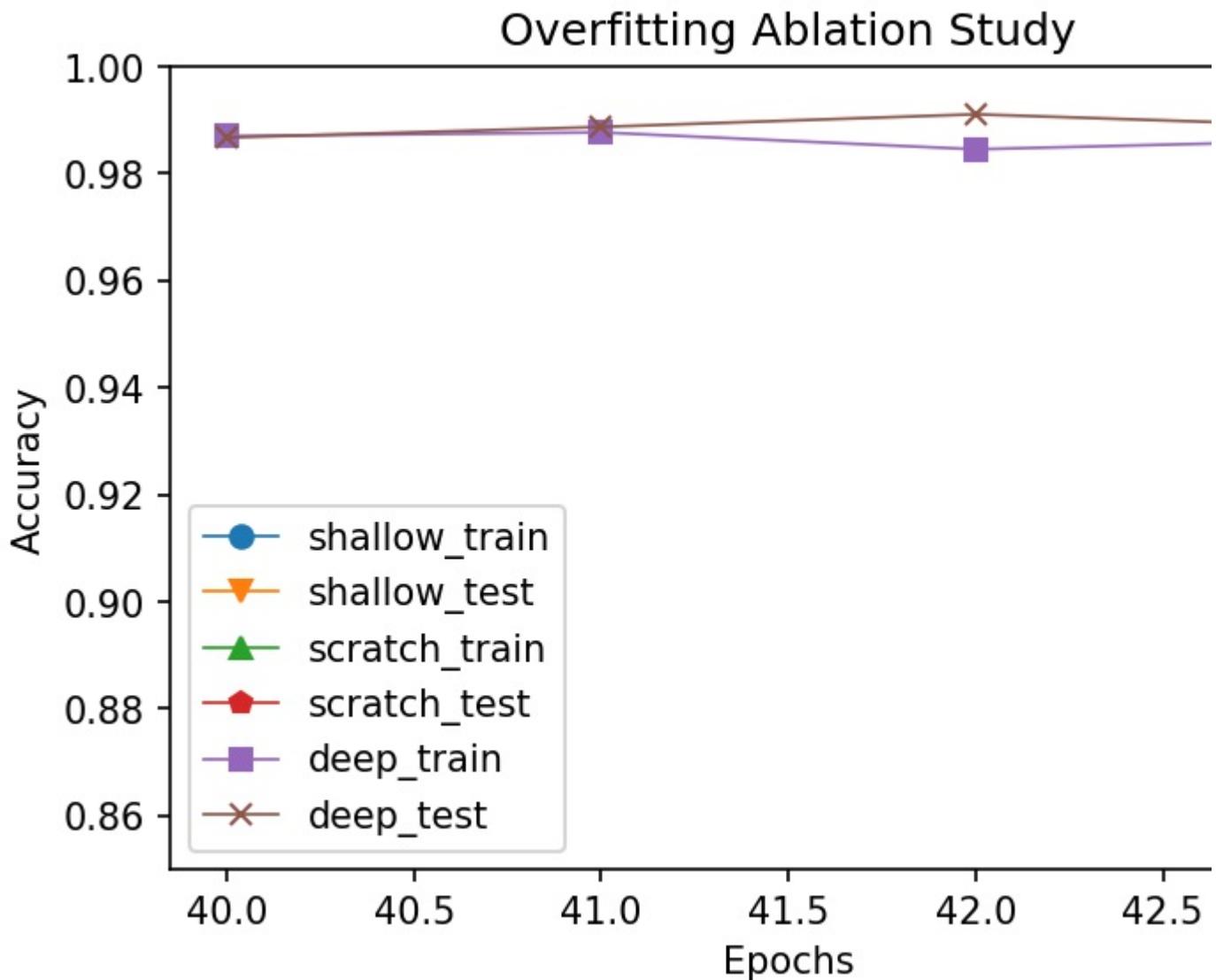
epoch 43 accuracy on test set is: 0.9888222464558343

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
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





```
epoch 43 takes 315 seconds.  
[45,    30] loss: 0.065  
[45,    60] loss: 0.062  
[45,   90] loss: 0.070  
[45,  120] loss: 0.055  
[45,  150] loss: 0.043  
[45,  180] loss: 0.043  
[45,  210] loss: 0.047  
[45,  240] loss: 0.081  
[45,  270] loss: 0.064  
[45,  300] loss: 0.069  
[45,  330] loss: 0.050  
[45,  360] loss: 0.059  
[45,  390] loss: 0.052  
[45,  420] loss: 0.055  
[45,  450] loss: 0.034  
[45,  480] loss: 0.061  
[45,  510] loss: 0.040  
[45,  540] loss: 0.060  
[45,  570] loss: 0.037  
[45,  600] loss: 0.047  
[45,  630] loss: 0.050  
[45,  660] loss: 0.055
```

```

[45, 690] loss: 0.045
[45, 720] loss: 0.045
[45, 750] loss: 0.051
[45, 780] loss: 0.040
[45, 810] loss: 0.036
[45, 840] loss: 0.042
[45, 870] loss: 0.047
[45, 900] loss: 0.049
[45, 930] loss: 0.034
[45, 960] loss: 0.055
[45, 990] loss: 0.036
[45, 1020] loss: 0.073
[45, 1050] loss: 0.025
[45, 1080] loss: 0.074
[45, 1110] loss: 0.035
[45, 1140] loss: 0.055
[45, 1170] loss: 0.063
[45, 1200] loss: 0.081
[45, 1230] loss: 0.061
[45, 1260] loss: 0.077
[45, 1290] loss: 0.074
[45, 1320] loss: 0.052
[45, 1350] loss: 0.054
[45, 1380] loss: 0.056
[45, 1410] loss: 0.059
[45, 1440] loss: 0.040
[45, 1470] loss: 0.067
[45, 1500] loss: 0.080
[45, 1530] loss: 0.038
[45, 1560] loss: 0.044
[45, 1590] loss: 0.057
[45, 1620] loss: 0.058
[45, 1650] loss: 0.048
[45, 1680] loss: 0.020
[45, 1710] loss: 0.084
[45, 1740] loss: 0.051
[45, 1770] loss: 0.047
[45, 1800] loss: 0.053
[45, 1830] loss: 0.059
[45, 1860] loss: 0.042
[45, 1890] loss: 0.032
[45, 1920] loss: 0.030
[45, 1950] loss: 0.052
[45, 1980] loss: 0.047
[45, 2010] loss: 0.043
[45, 2040] loss: 0.081
[45, 2070] loss: 0.074
[45, 2100] loss: 0.060
[45, 2130] loss: 0.076
[45, 2160] loss: 0.044
[45, 2190] loss: 0.047
ovft_intermediate_models/epoch44_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.986619
epoch 44 accuracy on train set is: 0.9866185023627771
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	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

evaluating on test set...

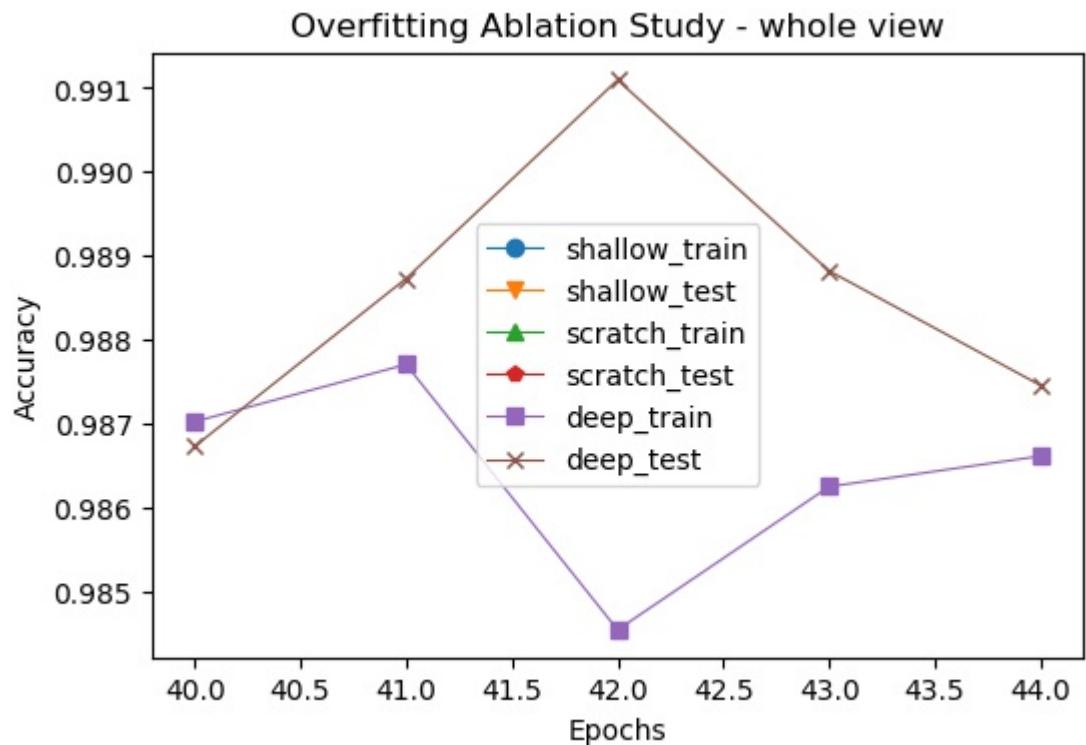
Accuracy on test images: 0.987459

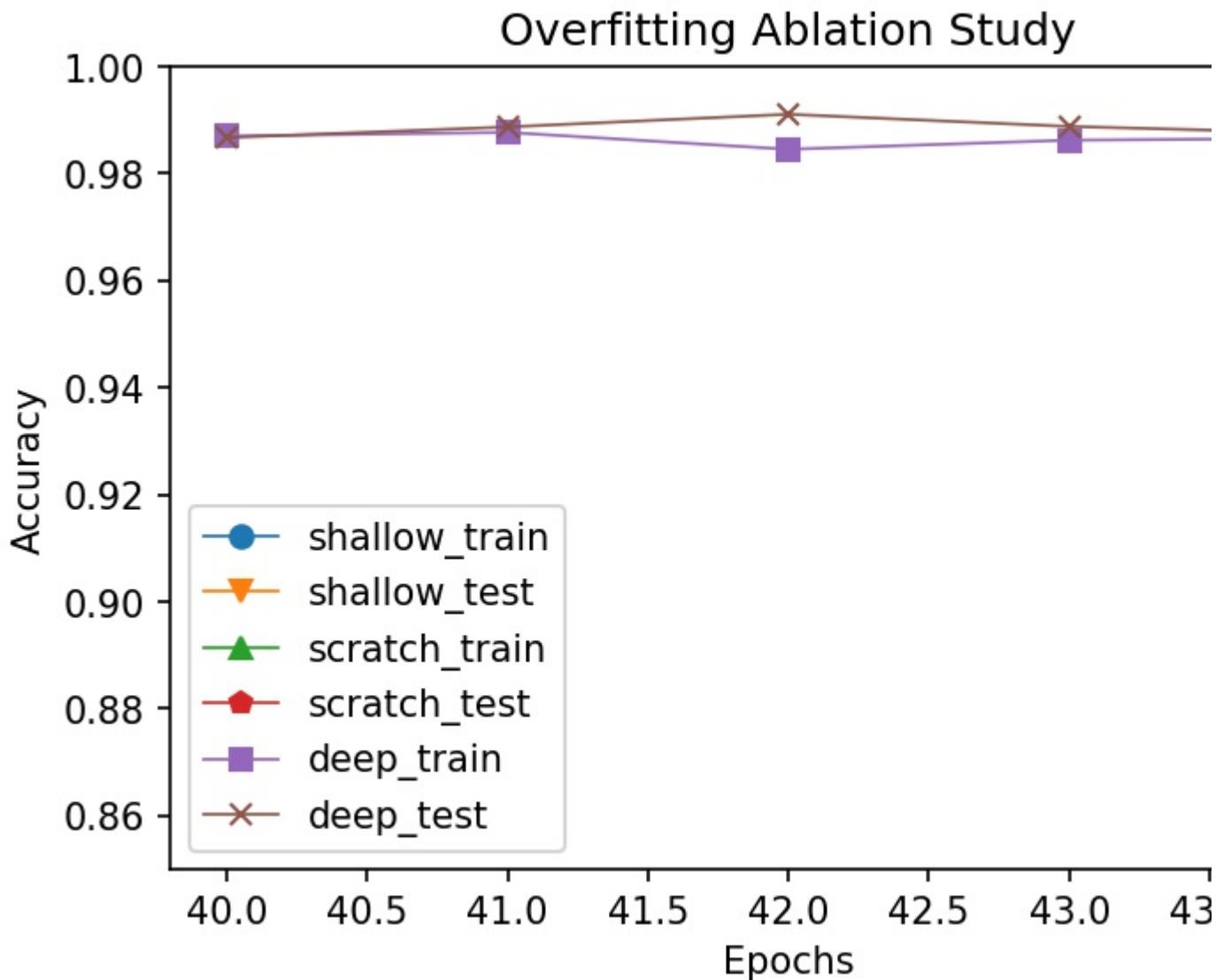
epoch 44 accuracy on test set is: 0.9874591057797165

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
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





```

epoch 44 takes 316 seconds.
[46,    30] loss: 0.060
[46,    60] loss: 0.086
[46,    90] loss: 0.027
[46,   120] loss: 0.077
[46,   150] loss: 0.056
[46,   180] loss: 0.076
[46,   210] loss: 0.043
[46,   240] loss: 0.045
[46,   270] loss: 0.044
[46,   300] loss: 0.050
[46,   330] loss: 0.045
[46,   360] loss: 0.045
[46,   390] loss: 0.058
[46,   420] loss: 0.072
[46,   450] loss: 0.060
[46,   480] loss: 0.052
[46,   510] loss: 0.045
[46,   540] loss: 0.037
[46,   570] loss: 0.068
[46,   600] loss: 0.035
[46,   630] loss: 0.063
[46,   660] loss: 0.073

```

```

[46, 690] loss: 0.076
[46, 720] loss: 0.034
[46, 750] loss: 0.063
[46, 780] loss: 0.053
[46, 810] loss: 0.065
[46, 840] loss: 0.061
[46, 870] loss: 0.029
[46, 900] loss: 0.048
[46, 930] loss: 0.091
[46, 960] loss: 0.036
[46, 990] loss: 0.059
[46, 1020] loss: 0.031
[46, 1050] loss: 0.037
[46, 1080] loss: 0.058
[46, 1110] loss: 0.059
[46, 1140] loss: 0.048
[46, 1170] loss: 0.085
[46, 1200] loss: 0.061
[46, 1230] loss: 0.074
[46, 1260] loss: 0.048
[46, 1290] loss: 0.048
[46, 1320] loss: 0.054
[46, 1350] loss: 0.088
[46, 1380] loss: 0.072
[46, 1410] loss: 0.038
[46, 1440] loss: 0.108
[46, 1470] loss: 0.051
[46, 1500] loss: 0.078
[46, 1530] loss: 0.043
[46, 1560] loss: 0.068
[46, 1590] loss: 0.046
[46, 1620] loss: 0.060
[46, 1650] loss: 0.037
[46, 1680] loss: 0.056
[46, 1710] loss: 0.046
[46, 1740] loss: 0.036
[46, 1770] loss: 0.036
[46, 1800] loss: 0.037
[46, 1830] loss: 0.057
[46, 1860] loss: 0.059
[46, 1890] loss: 0.079
[46, 1920] loss: 0.063
[46, 1950] loss: 0.059
[46, 1980] loss: 0.043
[46, 2010] loss: 0.046
[46, 2040] loss: 0.050
[46, 2070] loss: 0.059
[46, 2100] loss: 0.077
[46, 2130] loss: 0.062
[46, 2160] loss: 0.065
[46, 2190] loss: 0.045
ovft_intermediate_models/epoch45_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.987255
epoch 45 accuracy on train set is: 0.9872546346782988
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	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

evaluating on test set...

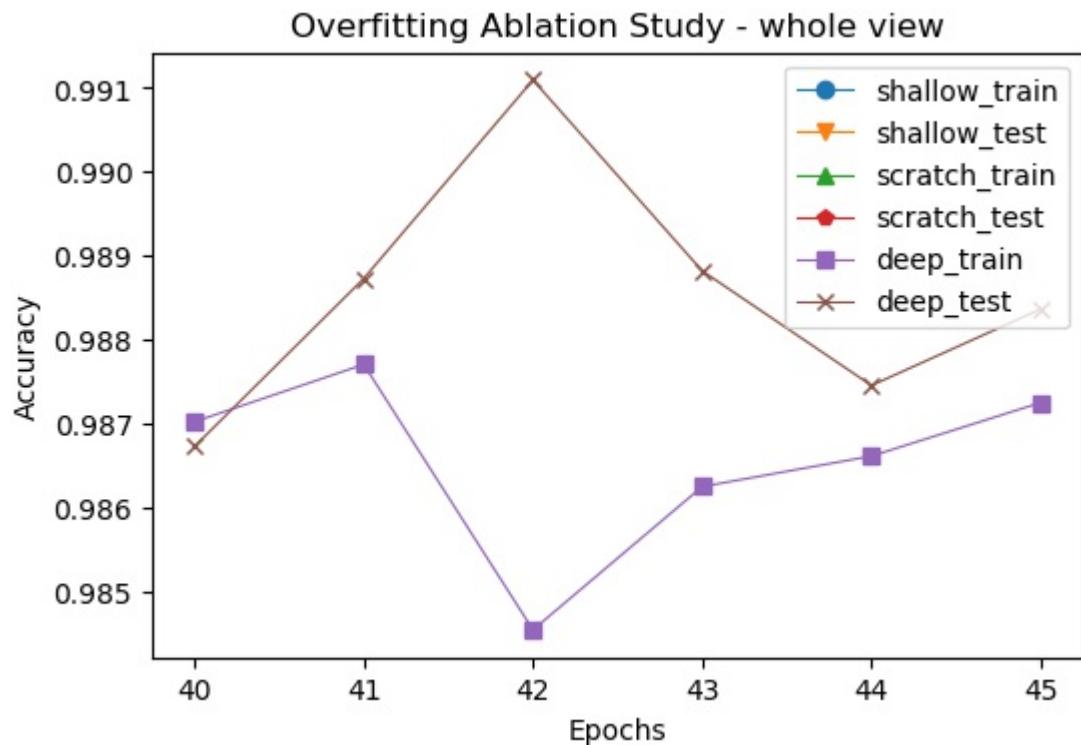
Accuracy on test images: 0.988368

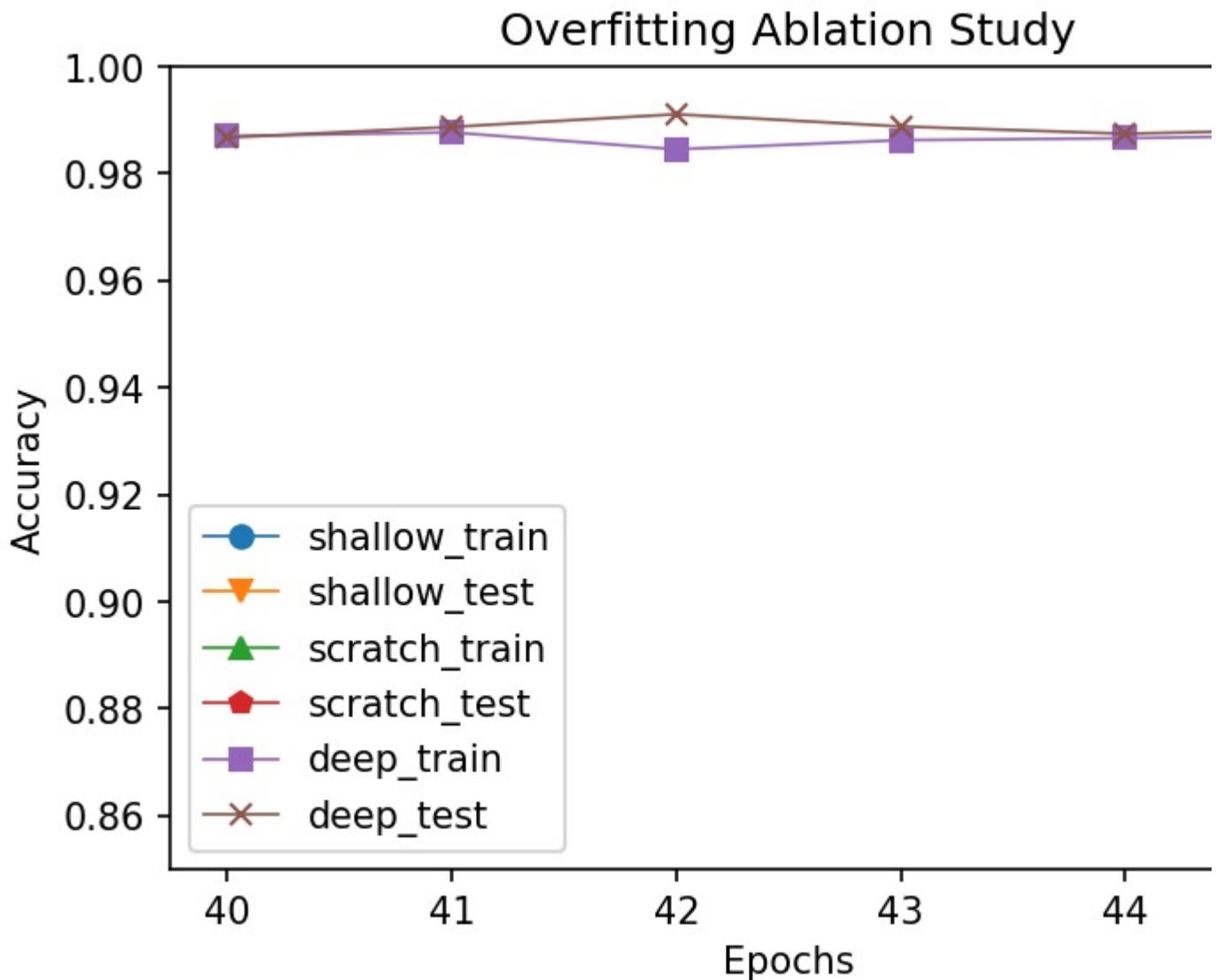
epoch 45 accuracy on test set is: 0.9883678662304617

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
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





```

epoch 45 takes 316 seconds.
[47,    30] loss: 0.051
[47,    60] loss: 0.050
[47,    90] loss: 0.055
[47,   120] loss: 0.062
[47,   150] loss: 0.068
[47,   180] loss: 0.047
[47,   210] loss: 0.059
[47,   240] loss: 0.024
[47,   270] loss: 0.069
[47,   300] loss: 0.037
[47,   330] loss: 0.025
[47,   360] loss: 0.042
[47,   390] loss: 0.042
[47,   420] loss: 0.062
[47,   450] loss: 0.053
[47,   480] loss: 0.062
[47,   510] loss: 0.038
[47,   540] loss: 0.036
[47,   570] loss: 0.041
[47,   600] loss: 0.028
[47,   630] loss: 0.069
[47,   660] loss: 0.063

```

```

[47, 690] loss: 0.063
[47, 720] loss: 0.066
[47, 750] loss: 0.040
[47, 780] loss: 0.044
[47, 810] loss: 0.032
[47, 840] loss: 0.051
[47, 870] loss: 0.050
[47, 900] loss: 0.068
[47, 930] loss: 0.074
[47, 960] loss: 0.049
[47, 990] loss: 0.053
[47, 1020] loss: 0.049
[47, 1050] loss: 0.046
[47, 1080] loss: 0.061
[47, 1110] loss: 0.063
[47, 1140] loss: 0.041
[47, 1170] loss: 0.023
[47, 1200] loss: 0.045
[47, 1230] loss: 0.061
[47, 1260] loss: 0.043
[47, 1290] loss: 0.065
[47, 1320] loss: 0.046
[47, 1350] loss: 0.068
[47, 1380] loss: 0.057
[47, 1410] loss: 0.048
[47, 1440] loss: 0.036
[47, 1470] loss: 0.071
[47, 1500] loss: 0.070
[47, 1530] loss: 0.054
[47, 1560] loss: 0.042
[47, 1590] loss: 0.037
[47, 1620] loss: 0.070
[47, 1650] loss: 0.052
[47, 1680] loss: 0.045
[47, 1710] loss: 0.065
[47, 1740] loss: 0.072
[47, 1770] loss: 0.059
[47, 1800] loss: 0.034
[47, 1830] loss: 0.049
[47, 1860] loss: 0.024
[47, 1890] loss: 0.028
[47, 1920] loss: 0.025
[47, 1950] loss: 0.069
[47, 1980] loss: 0.077
[47, 2010] loss: 0.096
[47, 2040] loss: 0.032
[47, 2070] loss: 0.060
[47, 2100] loss: 0.047
[47, 2130] loss: 0.055
[47, 2160] loss: 0.078
[47, 2190] loss: 0.045
ovft_intermediate_models/epoch46_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.988709
epoch 46 accuracy on train set is: 0.988708651399491
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	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

evaluating on test set...

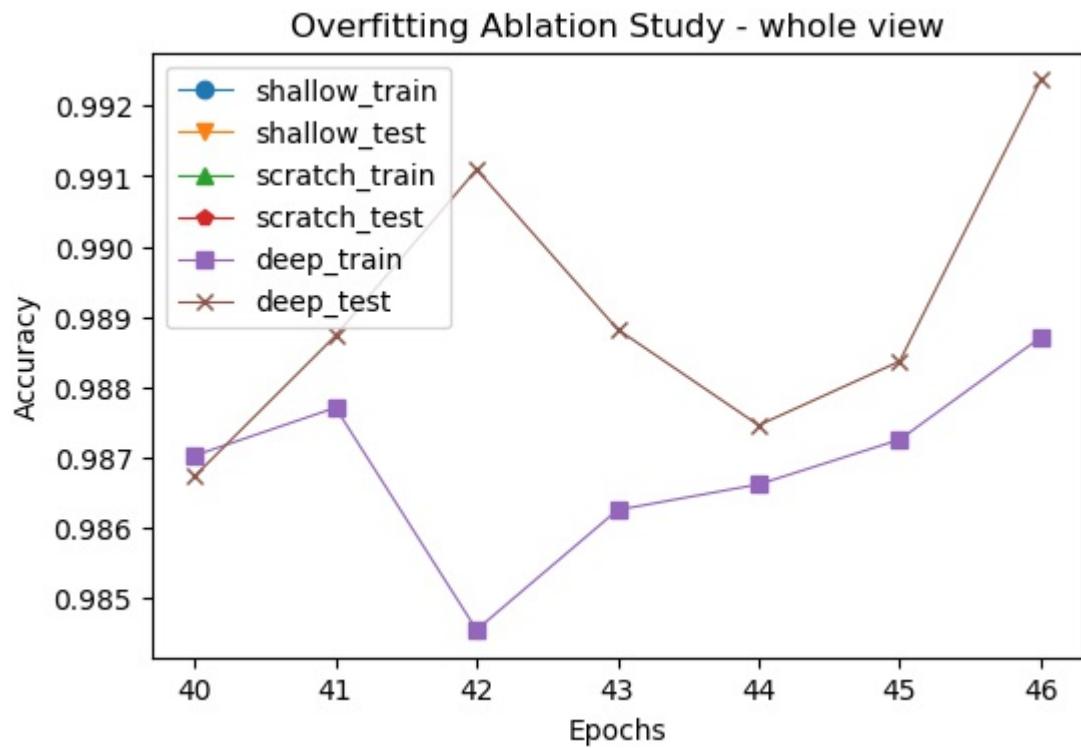
Accuracy on test images: 0.992366

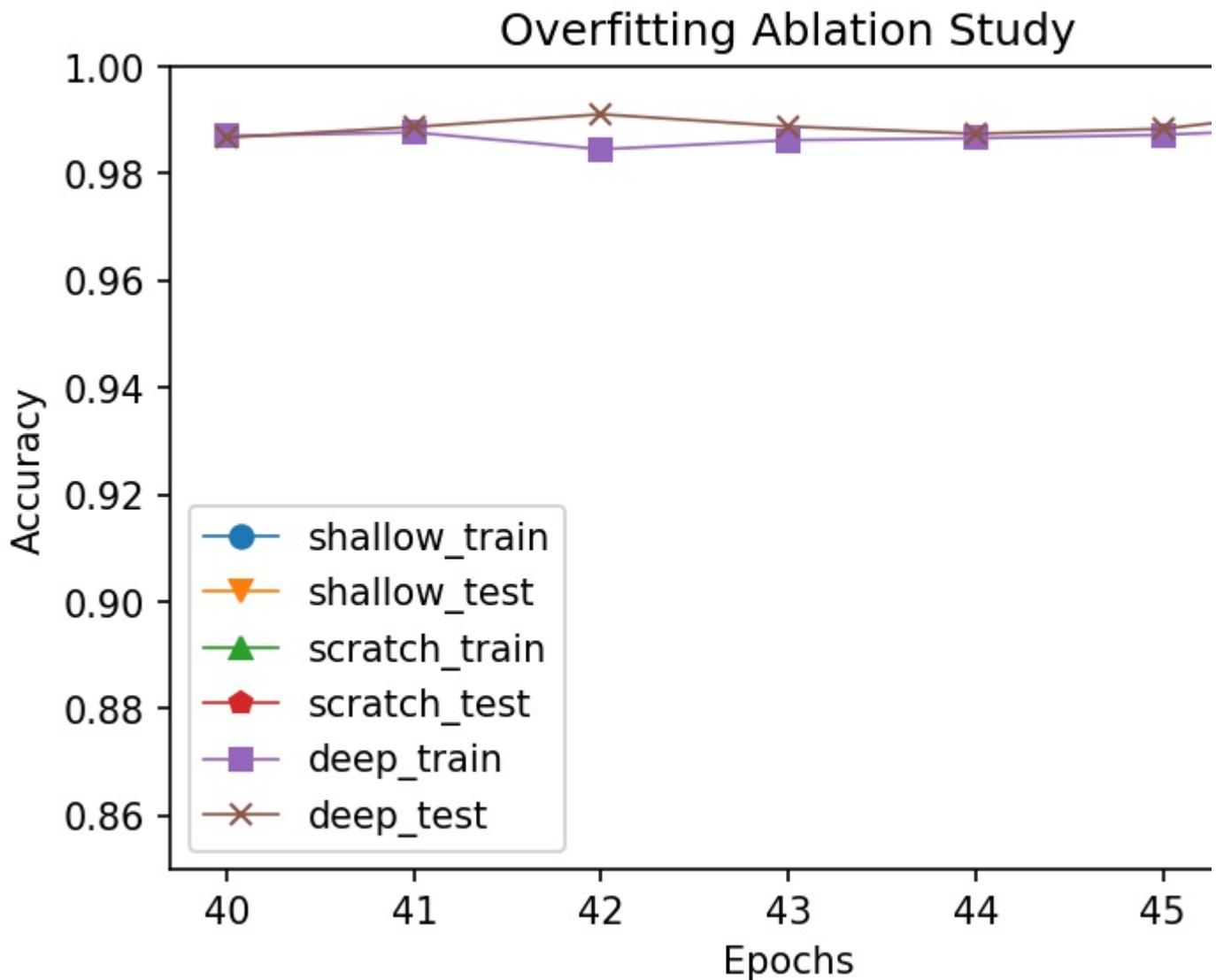
epoch 46 accuracy on test set is: 0.9923664122137404

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
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





```
epoch 46 takes 316 seconds.  
[48,    30] loss: 0.058  
[48,    60] loss: 0.045  
[48,   90] loss: 0.049  
[48,  120] loss: 0.045  
[48,  150] loss: 0.065  
[48,  180] loss: 0.074  
[48,  210] loss: 0.031  
[48,  240] loss: 0.037  
[48,  270] loss: 0.062  
[48,  300] loss: 0.080  
[48,  330] loss: 0.057  
[48,  360] loss: 0.034  
[48,  390] loss: 0.047  
[48,  420] loss: 0.063  
[48,  450] loss: 0.048  
[48,  480] loss: 0.043  
[48,  510] loss: 0.082  
[48,  540] loss: 0.077  
[48,  570] loss: 0.038  
[48,  600] loss: 0.043  
[48,  630] loss: 0.035  
[48,  660] loss: 0.053
```

```

[48, 690] loss: 0.060
[48, 720] loss: 0.072
[48, 750] loss: 0.049
[48, 780] loss: 0.045
[48, 810] loss: 0.053
[48, 840] loss: 0.049
[48, 870] loss: 0.043
[48, 900] loss: 0.071
[48, 930] loss: 0.027
[48, 960] loss: 0.064
[48, 990] loss: 0.037
[48, 1020] loss: 0.073
[48, 1050] loss: 0.067
[48, 1080] loss: 0.059
[48, 1110] loss: 0.052
[48, 1140] loss: 0.036
[48, 1170] loss: 0.047
[48, 1200] loss: 0.028
[48, 1230] loss: 0.077
[48, 1260] loss: 0.060
[48, 1290] loss: 0.054
[48, 1320] loss: 0.045
[48, 1350] loss: 0.056
[48, 1380] loss: 0.036
[48, 1410] loss: 0.044
[48, 1440] loss: 0.070
[48, 1470] loss: 0.048
[48, 1500] loss: 0.070
[48, 1530] loss: 0.076
[48, 1560] loss: 0.079
[48, 1590] loss: 0.045
[48, 1620] loss: 0.070
[48, 1650] loss: 0.060
[48, 1680] loss: 0.027
[48, 1710] loss: 0.035
[48, 1740] loss: 0.047
[48, 1770] loss: 0.020
[48, 1800] loss: 0.038
[48, 1830] loss: 0.042
[48, 1860] loss: 0.039
[48, 1890] loss: 0.023
[48, 1920] loss: 0.076
[48, 1950] loss: 0.058
[48, 1980] loss: 0.066
[48, 2010] loss: 0.046
[48, 2040] loss: 0.038
[48, 2070] loss: 0.064
[48, 2100] loss: 0.067
[48, 2130] loss: 0.071
[48, 2160] loss: 0.062
[48, 2190] loss: 0.053
ovft_intermediate_models/epoch47_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.988277
epoch 47 accuracy on train set is: 0.9882769901853872
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	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

evaluating on test set...

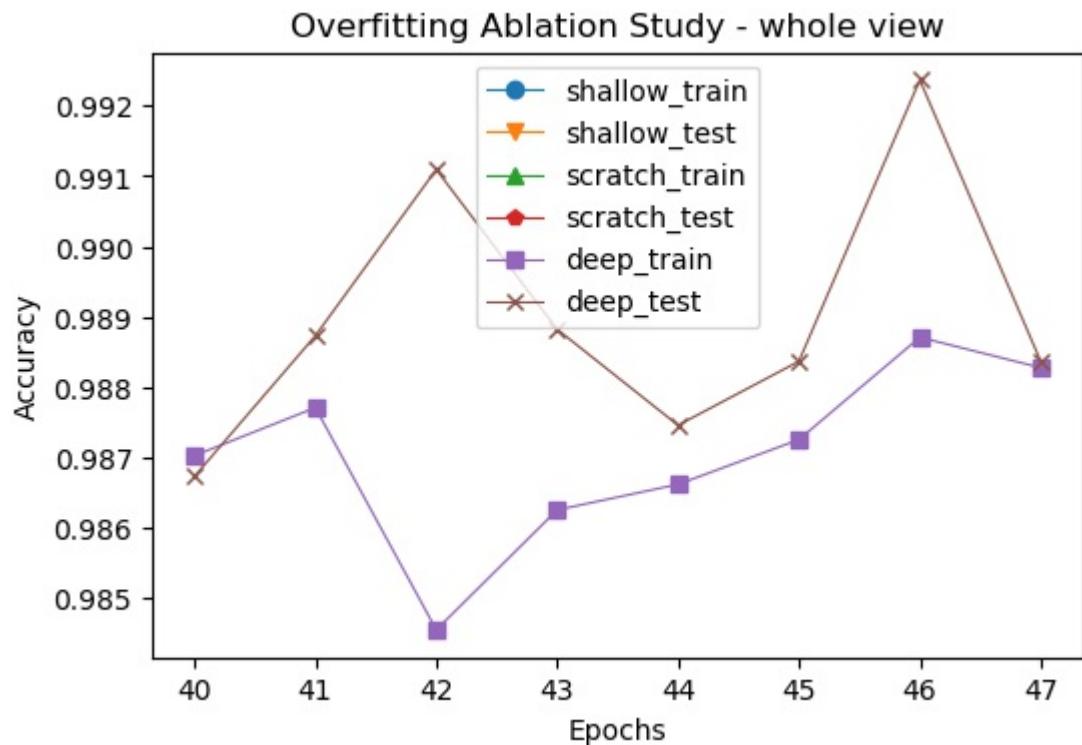
Accuracy on test images: 0.988368

epoch 47 accuracy on test set is: 0.9883678662304617

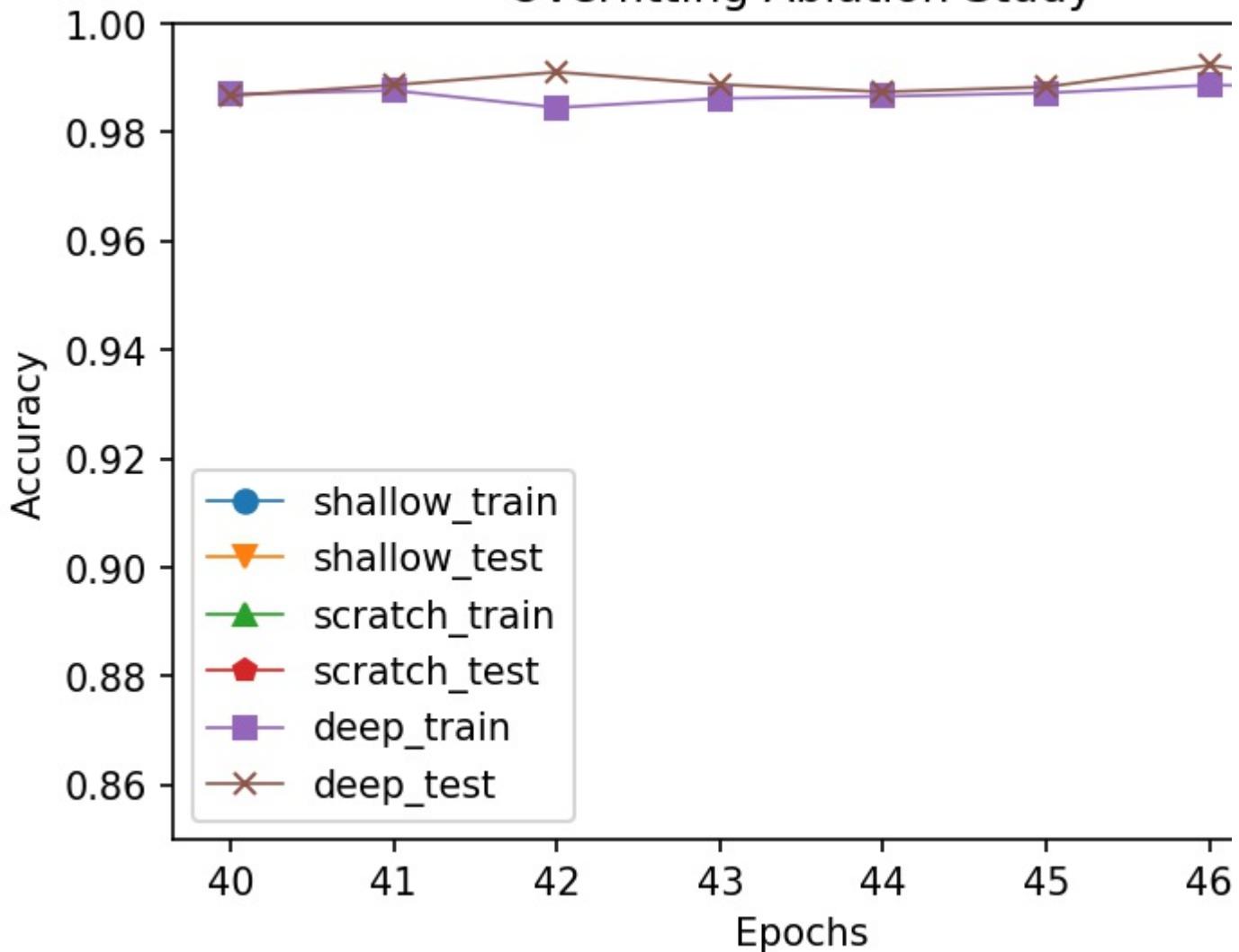
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
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



Overfitting Ablation Study



```
epoch 47 takes 316 seconds.  
[49,    30] loss: 0.042  
[49,    60] loss: 0.023  
[49,   90] loss: 0.046  
[49,  120] loss: 0.066  
[49,  150] loss: 0.036  
[49,  180] loss: 0.060  
[49,  210] loss: 0.049  
[49,  240] loss: 0.040  
[49,  270] loss: 0.037  
[49,  300] loss: 0.047  
[49,  330] loss: 0.060  
[49,  360] loss: 0.060  
[49,  390] loss: 0.039  
[49,  420] loss: 0.046  
[49,  450] loss: 0.042  
[49,  480] loss: 0.051  
[49,  510] loss: 0.045  
[49,  540] loss: 0.070  
[49,  570] loss: 0.040  
[49,  600] loss: 0.044  
[49,  630] loss: 0.036  
[49,  660] loss: 0.056
```

```

[49,   690] loss: 0.035
[49,   720] loss: 0.052
[49,   750] loss: 0.041
[49,   780] loss: 0.080
[49,   810] loss: 0.041
[49,   840] loss: 0.044
[49,   870] loss: 0.048
[49,   900] loss: 0.049
[49,   930] loss: 0.062
[49,   960] loss: 0.050
[49,   990] loss: 0.048
[49,  1020] loss: 0.063
[49,  1050] loss: 0.028
[49,  1080] loss: 0.041
[49,  1110] loss: 0.029
[49,  1140] loss: 0.036
[49,  1170] loss: 0.047
[49,  1200] loss: 0.053
[49,  1230] loss: 0.045
[49,  1260] loss: 0.039
[49,  1290] loss: 0.025
[49,  1320] loss: 0.040
[49,  1350] loss: 0.042
[49,  1380] loss: 0.045
[49,  1410] loss: 0.063
[49,  1440] loss: 0.049
[49,  1470] loss: 0.052
[49,  1500] loss: 0.117
[49,  1530] loss: 0.061
[49,  1560] loss: 0.074
[49,  1590] loss: 0.076
[49,  1620] loss: 0.044
[49,  1650] loss: 0.042
[49,  1680] loss: 0.044
[49,  1710] loss: 0.054
[49,  1740] loss: 0.047
[49,  1770] loss: 0.046
[49,  1800] loss: 0.062
[49,  1830] loss: 0.025
[49,  1860] loss: 0.040
[49,  1890] loss: 0.057
[49,  1920] loss: 0.064
[49,  1950] loss: 0.038
[49,  1980] loss: 0.038
[49,  2010] loss: 0.052
[49,  2040] loss: 0.024
[49,  2070] loss: 0.028
[49,  2100] loss: 0.049
[49,  2130] loss: 0.046
[49,  2160] loss: 0.069
[49,  2190] loss: 0.034
ovft_intermediate_models/epoch48_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989095
epoch 48 accuracy on train set is: 0.9890948745910578
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN    0.987027  0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	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

evaluating on test set...

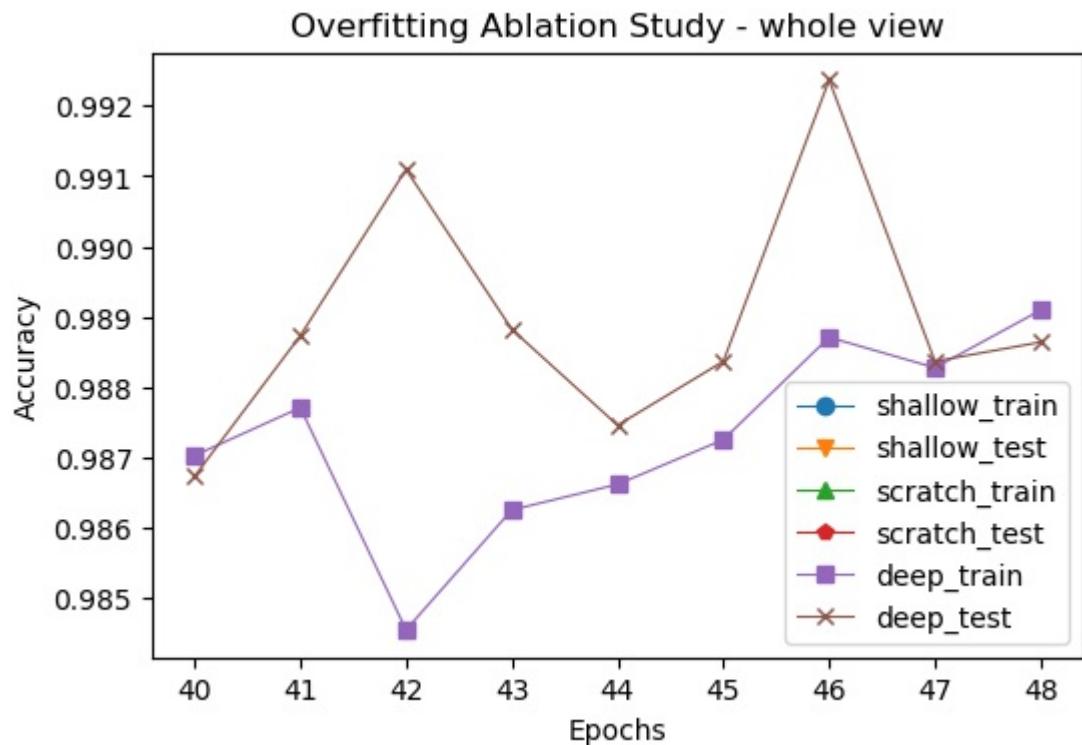
Accuracy on test images: 0.988640

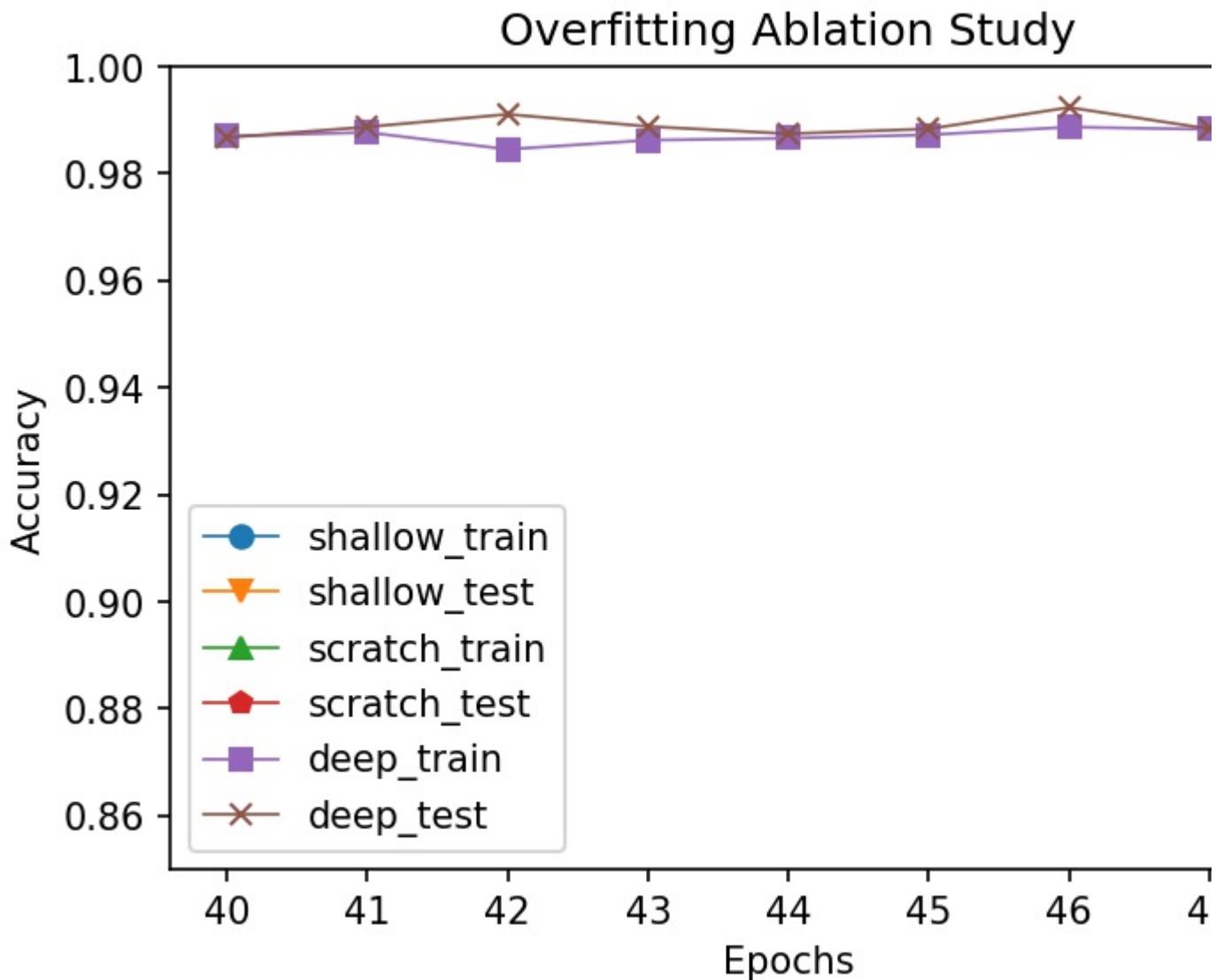
epoch 48 accuracy on test set is: 0.9886404943656852

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
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





```

epoch 48 takes 316 seconds.
[50,    30] loss: 0.073
[50,    60] loss: 0.064
[50,   90] loss: 0.053
[50,  120] loss: 0.039
[50,  150] loss: 0.040
[50,  180] loss: 0.020
[50,  210] loss: 0.056
[50,  240] loss: 0.032
[50,  270] loss: 0.044
[50,  300] loss: 0.045
[50,  330] loss: 0.033
[50,  360] loss: 0.059
[50,  390] loss: 0.034
[50,  420] loss: 0.048
[50,  450] loss: 0.047
[50,  480] loss: 0.046
[50,  510] loss: 0.066
[50,  540] loss: 0.058
[50,  570] loss: 0.056
[50,  600] loss: 0.053
[50,  630] loss: 0.061
[50,  660] loss: 0.070

```

```

[50, 690] loss: 0.071
[50, 720] loss: 0.054
[50, 750] loss: 0.038
[50, 780] loss: 0.062
[50, 810] loss: 0.074
[50, 840] loss: 0.068
[50, 870] loss: 0.070
[50, 900] loss: 0.091
[50, 930] loss: 0.071
[50, 960] loss: 0.061
[50, 990] loss: 0.041
[50, 1020] loss: 0.048
[50, 1050] loss: 0.049
[50, 1080] loss: 0.034
[50, 1110] loss: 0.069
[50, 1140] loss: 0.059
[50, 1170] loss: 0.047
[50, 1200] loss: 0.051
[50, 1230] loss: 0.045
[50, 1260] loss: 0.045
[50, 1290] loss: 0.041
[50, 1320] loss: 0.051
[50, 1350] loss: 0.028
[50, 1380] loss: 0.034
[50, 1410] loss: 0.037
[50, 1440] loss: 0.062
[50, 1470] loss: 0.055
[50, 1500] loss: 0.051
[50, 1530] loss: 0.071
[50, 1560] loss: 0.041
[50, 1590] loss: 0.038
[50, 1620] loss: 0.077
[50, 1650] loss: 0.064
[50, 1680] loss: 0.065
[50, 1710] loss: 0.080
[50, 1740] loss: 0.058
[50, 1770] loss: 0.065
[50, 1800] loss: 0.035
[50, 1830] loss: 0.052
[50, 1860] loss: 0.085
[50, 1890] loss: 0.082
[50, 1920] loss: 0.044
[50, 1950] loss: 0.037
[50, 1980] loss: 0.041
[50, 2010] loss: 0.053
[50, 2040] loss: 0.069
[50, 2070] loss: 0.045
[50, 2100] loss: 0.030
[50, 2130] loss: 0.087
[50, 2160] loss: 0.051
[50, 2190] loss: 0.037
ovft_intermediate_models/epoch49_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.988595
epoch 49 accuracy on train set is: 0.9885950563431479
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	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

evaluating on test set...

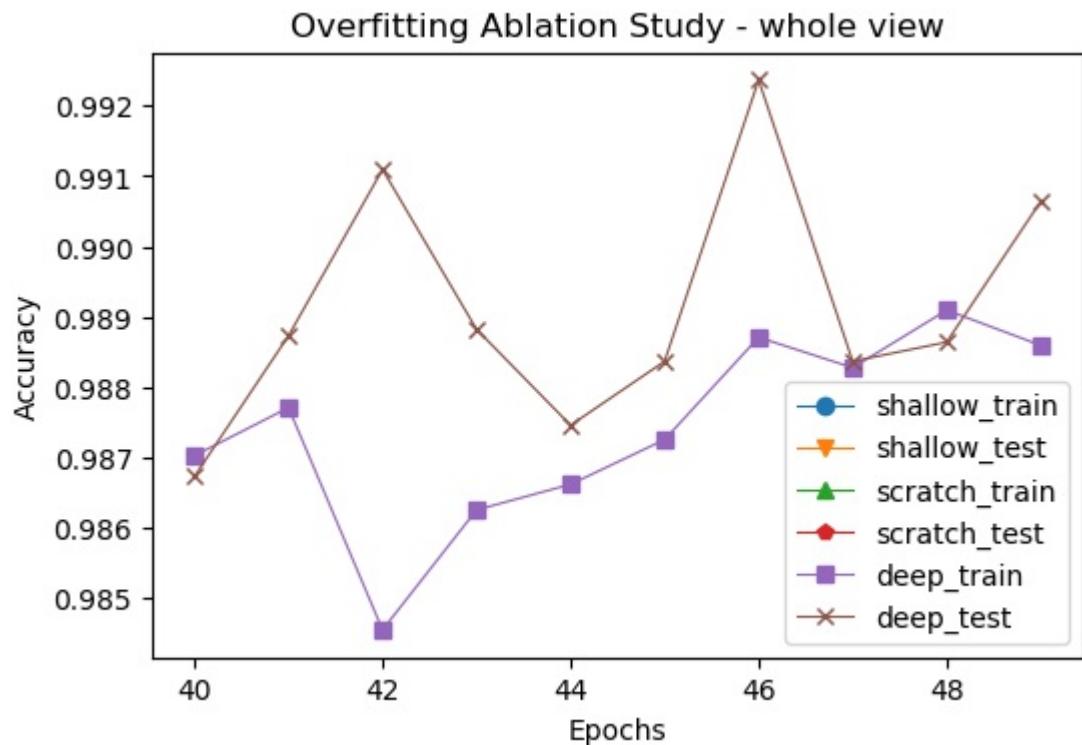
Accuracy on test images: 0.990640

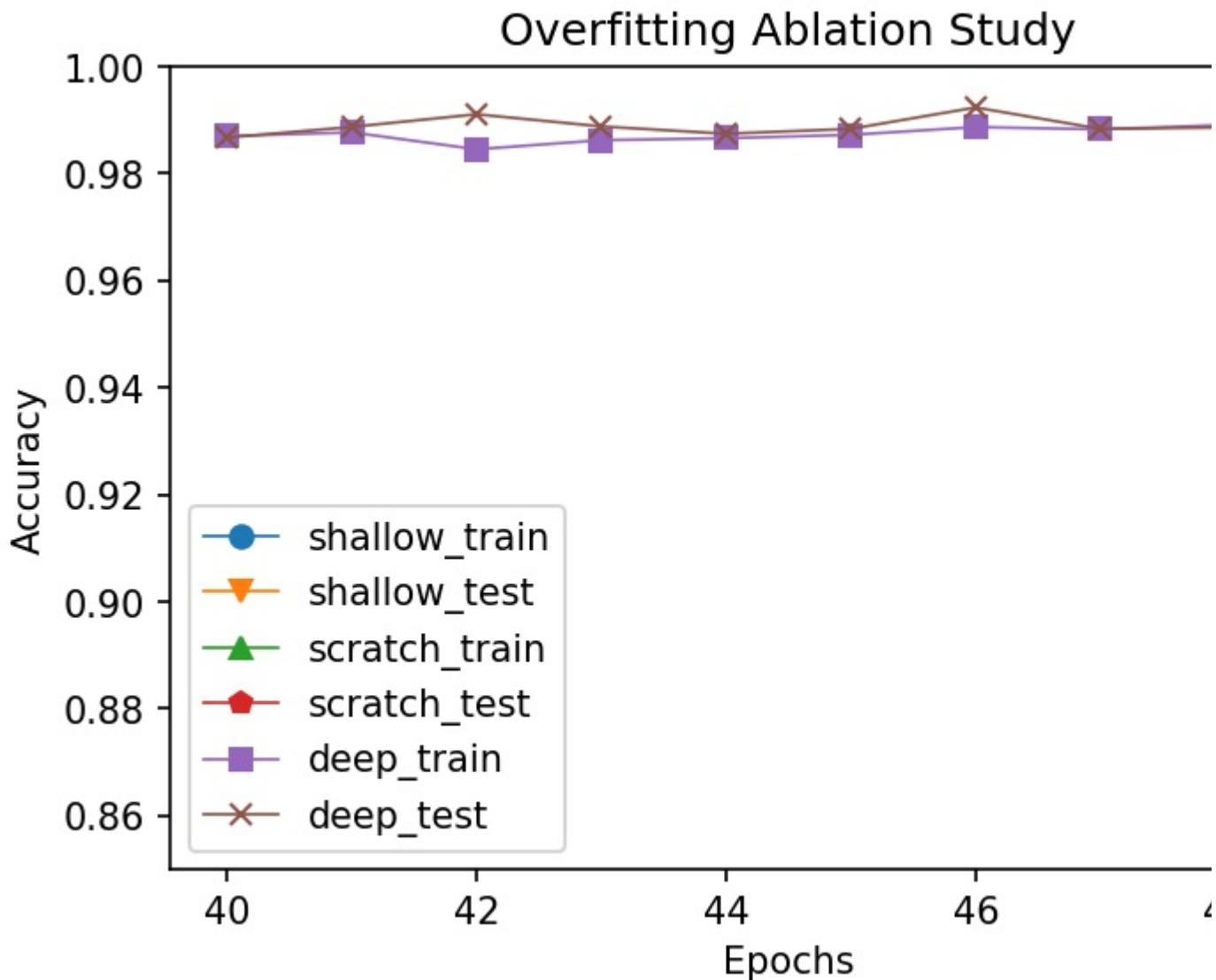
epoch 49 accuracy on test set is: 0.9906397673573246

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
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





```
epoch 49 takes 316 seconds.  
[51,    30] loss: 0.066  
[51,    60] loss: 0.071  
[51,    90] loss: 0.058  
[51,   120] loss: 0.058  
[51,   150] loss: 0.054  
[51,   180] loss: 0.048  
[51,   210] loss: 0.028  
[51,   240] loss: 0.037  
[51,   270] loss: 0.054  
[51,   300] loss: 0.067  
[51,   330] loss: 0.063  
[51,   360] loss: 0.050  
[51,   390] loss: 0.060  
[51,   420] loss: 0.064  
[51,   450] loss: 0.034  
[51,   480] loss: 0.060  
[51,   510] loss: 0.054  
[51,   540] loss: 0.078  
[51,   570] loss: 0.052  
[51,   600] loss: 0.054  
[51,   630] loss: 0.053  
[51,   660] loss: 0.041
```

```

[51, 690] loss: 0.042
[51, 720] loss: 0.040
[51, 750] loss: 0.056
[51, 780] loss: 0.084
[51, 810] loss: 0.057
[51, 840] loss: 0.051
[51, 870] loss: 0.052
[51, 900] loss: 0.070
[51, 930] loss: 0.072
[51, 960] loss: 0.047
[51, 990] loss: 0.053
[51, 1020] loss: 0.040
[51, 1050] loss: 0.077
[51, 1080] loss: 0.060
[51, 1110] loss: 0.064
[51, 1140] loss: 0.054
[51, 1170] loss: 0.034
[51, 1200] loss: 0.042
[51, 1230] loss: 0.045
[51, 1260] loss: 0.051
[51, 1290] loss: 0.034
[51, 1320] loss: 0.039
[51, 1350] loss: 0.052
[51, 1380] loss: 0.050
[51, 1410] loss: 0.030
[51, 1440] loss: 0.022
[51, 1470] loss: 0.054
[51, 1500] loss: 0.052
[51, 1530] loss: 0.046
[51, 1560] loss: 0.060
[51, 1590] loss: 0.073
[51, 1620] loss: 0.065
[51, 1650] loss: 0.068
[51, 1680] loss: 0.072
[51, 1710] loss: 0.041
[51, 1740] loss: 0.047
[51, 1770] loss: 0.038
[51, 1800] loss: 0.086
[51, 1830] loss: 0.034
[51, 1860] loss: 0.049
[51, 1890] loss: 0.040
[51, 1920] loss: 0.049
[51, 1950] loss: 0.058
[51, 1980] loss: 0.043
[51, 2010] loss: 0.049
[51, 2040] loss: 0.026
[51, 2070] loss: 0.029
[51, 2100] loss: 0.049
[51, 2130] loss: 0.060
[51, 2160] loss: 0.045
[51, 2190] loss: 0.038
ovft_intermediate_models/epoch50_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.988913
epoch 50 accuracy on train set is: 0.9889131225009088
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	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

evaluating on test set...

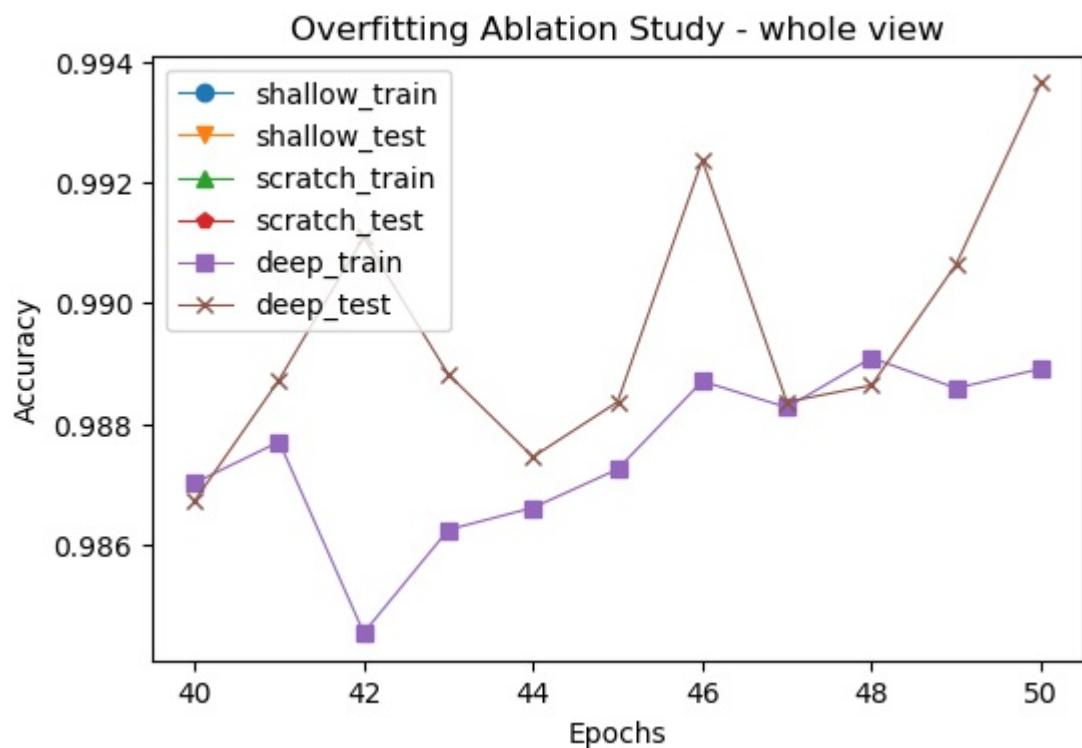
Accuracy on test images: 0.993639

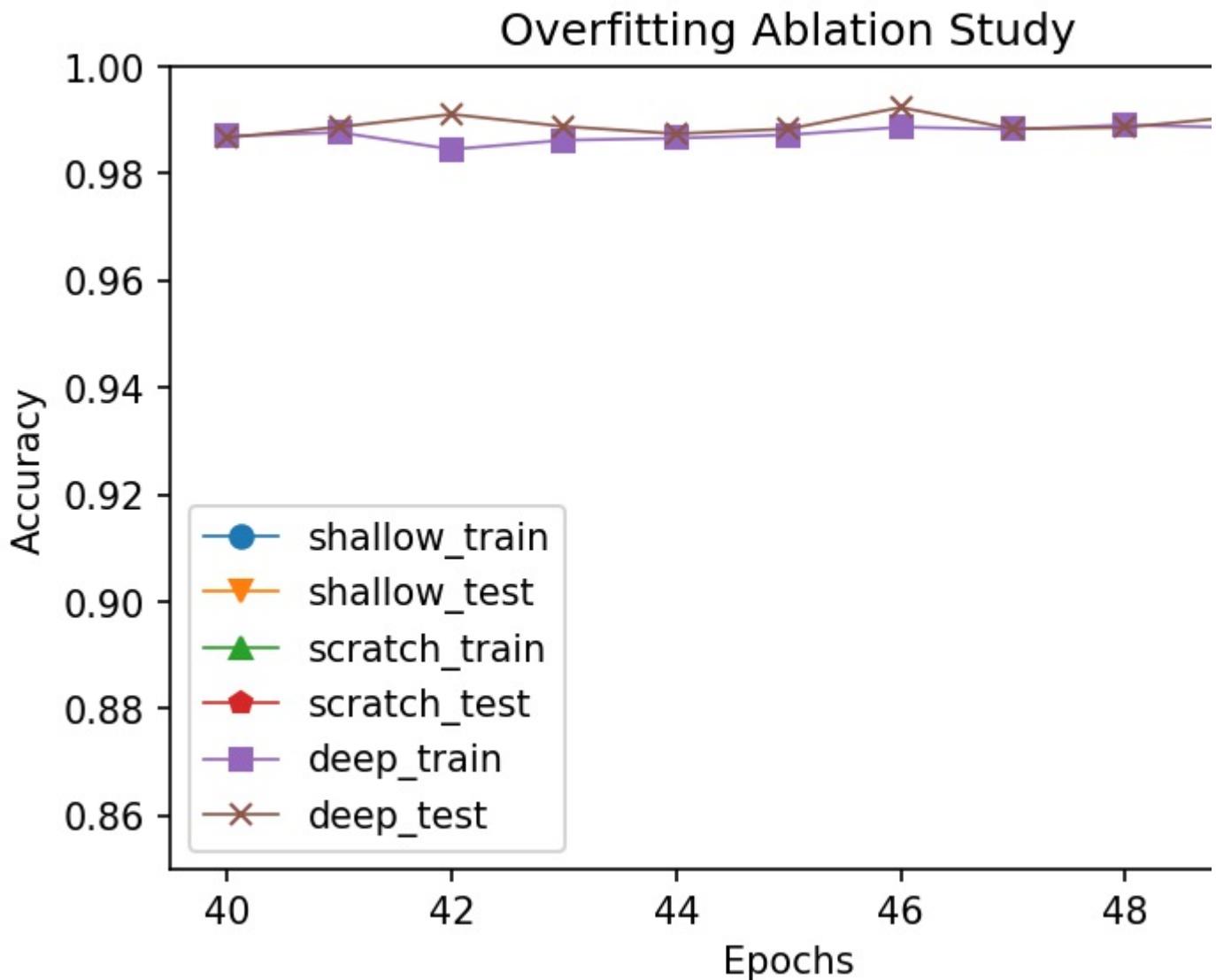
epoch 50 accuracy on test set is: 0.9936386768447837

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
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





```
epoch 50 takes 316 seconds.  
[52,    30] loss: 0.071  
[52,    60] loss: 0.039  
[52,    90] loss: 0.047  
[52,   120] loss: 0.040  
[52,   150] loss: 0.031  
[52,   180] loss: 0.044  
[52,   210] loss: 0.056  
[52,   240] loss: 0.060  
[52,   270] loss: 0.035  
[52,   300] loss: 0.052  
[52,   330] loss: 0.040  
[52,   360] loss: 0.042  
[52,   390] loss: 0.060  
[52,   420] loss: 0.036  
[52,   450] loss: 0.044  
[52,   480] loss: 0.042  
[52,   510] loss: 0.053  
[52,   540] loss: 0.053  
[52,   570] loss: 0.069  
[52,   600] loss: 0.075  
[52,   630] loss: 0.051  
[52,   660] loss: 0.067
```

```

[52, 690] loss: 0.060
[52, 720] loss: 0.028
[52, 750] loss: 0.041
[52, 780] loss: 0.032
[52, 810] loss: 0.058
[52, 840] loss: 0.034
[52, 870] loss: 0.063
[52, 900] loss: 0.073
[52, 930] loss: 0.031
[52, 960] loss: 0.072
[52, 990] loss: 0.030
[52, 1020] loss: 0.049
[52, 1050] loss: 0.036
[52, 1080] loss: 0.016
[52, 1110] loss: 0.091
[52, 1140] loss: 0.061
[52, 1170] loss: 0.070
[52, 1200] loss: 0.046
[52, 1230] loss: 0.061
[52, 1260] loss: 0.025
[52, 1290] loss: 0.043
[52, 1320] loss: 0.030
[52, 1350] loss: 0.048
[52, 1380] loss: 0.047
[52, 1410] loss: 0.072
[52, 1440] loss: 0.077
[52, 1470] loss: 0.036
[52, 1500] loss: 0.076
[52, 1530] loss: 0.051
[52, 1560] loss: 0.054
[52, 1590] loss: 0.061
[52, 1620] loss: 0.057
[52, 1650] loss: 0.046
[52, 1680] loss: 0.048
[52, 1710] loss: 0.051
[52, 1740] loss: 0.079
[52, 1770] loss: 0.033
[52, 1800] loss: 0.071
[52, 1830] loss: 0.067
[52, 1860] loss: 0.083
[52, 1890] loss: 0.062
[52, 1920] loss: 0.057
[52, 1950] loss: 0.060
[52, 1980] loss: 0.042
[52, 2010] loss: 0.057
[52, 2040] loss: 0.089
[52, 2070] loss: 0.046
[52, 2100] loss: 0.028
[52, 2130] loss: 0.022
[52, 2160] loss: 0.026
[52, 2190] loss: 0.032
ovft_intermediate_models/epoch51_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989527
epoch 51 accuracy on train set is: 0.9895265358051618
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	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

evaluating on test set...

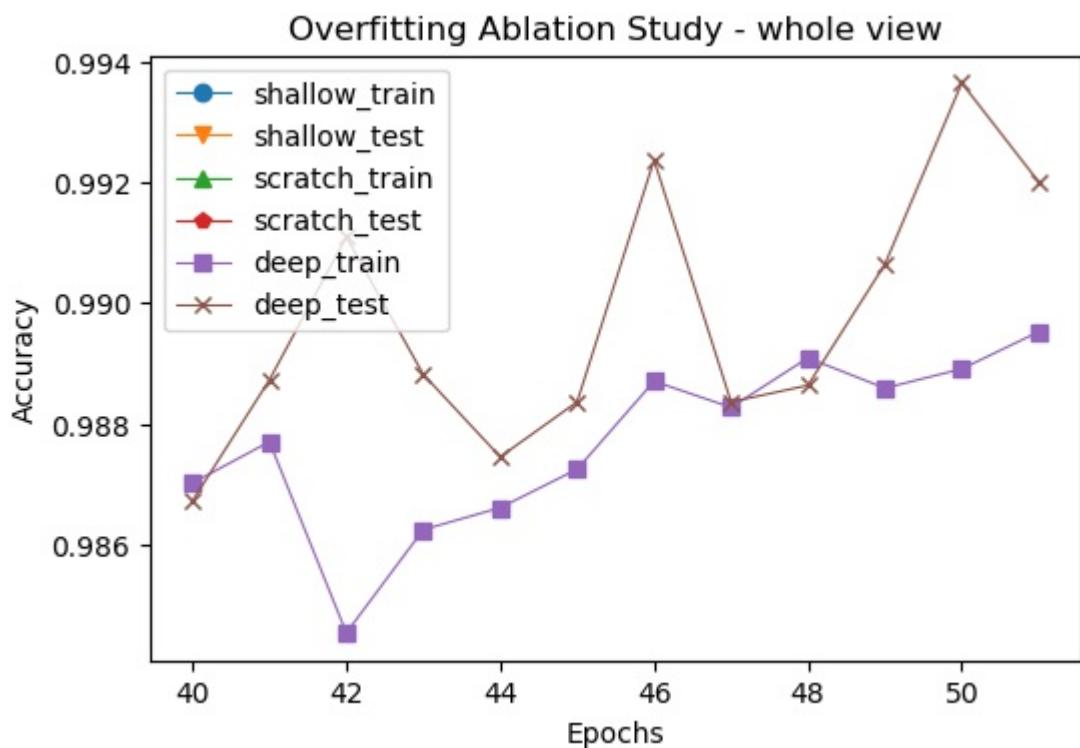
Accuracy on test images: 0.992003

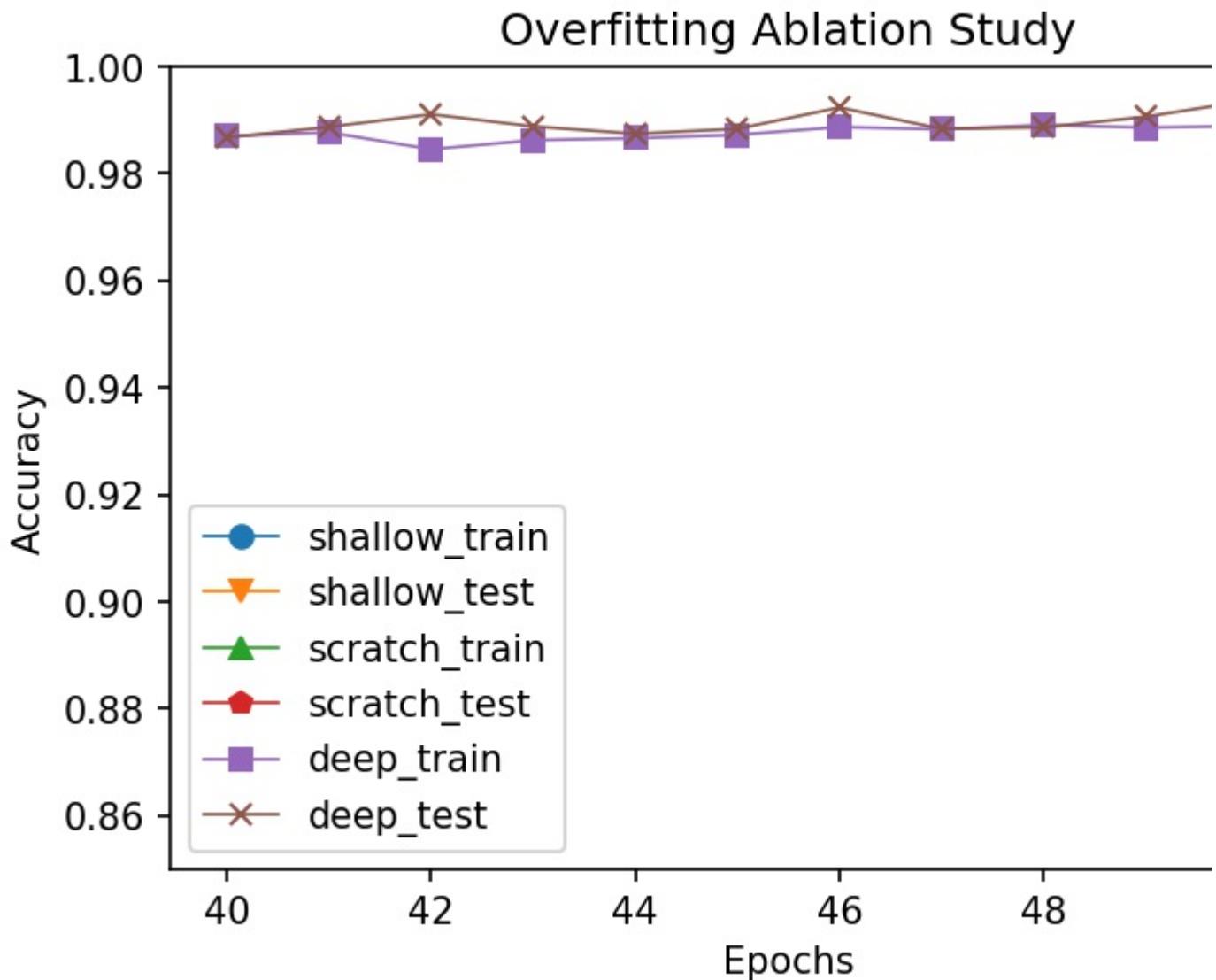
epoch 51 accuracy on test set is: 0.9920029080334424

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
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





```
epoch 51 takes 316 seconds.  
[53,    30] loss: 0.044  
[53,    60] loss: 0.060  
[53,    90] loss: 0.042  
[53,   120] loss: 0.029  
[53,   150] loss: 0.058  
[53,   180] loss: 0.026  
[53,   210] loss: 0.048  
[53,   240] loss: 0.071  
[53,   270] loss: 0.037  
[53,   300] loss: 0.055  
[53,   330] loss: 0.056  
[53,   360] loss: 0.049  
[53,   390] loss: 0.039  
[53,   420] loss: 0.042  
[53,   450] loss: 0.068  
[53,   480] loss: 0.049  
[53,   510] loss: 0.044  
[53,   540] loss: 0.071  
[53,   570] loss: 0.057  
[53,   600] loss: 0.082  
[53,   630] loss: 0.068  
[53,   660] loss: 0.049
```

```

[53, 690] loss: 0.028
[53, 720] loss: 0.051
[53, 750] loss: 0.040
[53, 780] loss: 0.039
[53, 810] loss: 0.058
[53, 840] loss: 0.023
[53, 870] loss: 0.070
[53, 900] loss: 0.070
[53, 930] loss: 0.054
[53, 960] loss: 0.060
[53, 990] loss: 0.049
[53, 1020] loss: 0.031
[53, 1050] loss: 0.061
[53, 1080] loss: 0.051
[53, 1110] loss: 0.047
[53, 1140] loss: 0.054
[53, 1170] loss: 0.031
[53, 1200] loss: 0.038
[53, 1230] loss: 0.068
[53, 1260] loss: 0.103
[53, 1290] loss: 0.049
[53, 1320] loss: 0.047
[53, 1350] loss: 0.049
[53, 1380] loss: 0.052
[53, 1410] loss: 0.043
[53, 1440] loss: 0.066
[53, 1470] loss: 0.036
[53, 1500] loss: 0.046
[53, 1530] loss: 0.041
[53, 1560] loss: 0.070
[53, 1590] loss: 0.052
[53, 1620] loss: 0.041
[53, 1650] loss: 0.052
[53, 1680] loss: 0.041
[53, 1710] loss: 0.041
[53, 1740] loss: 0.068
[53, 1770] loss: 0.051
[53, 1800] loss: 0.049
[53, 1830] loss: 0.032
[53, 1860] loss: 0.042
[53, 1890] loss: 0.043
[53, 1920] loss: 0.037
[53, 1950] loss: 0.031
[53, 1980] loss: 0.069
[53, 2010] loss: 0.047
[53, 2040] loss: 0.055
[53, 2070] loss: 0.083
[53, 2100] loss: 0.033
[53, 2130] loss: 0.056
[53, 2160] loss: 0.039
[53, 2190] loss: 0.060
ovft_intermediate_models/epoch52_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.987118
epoch 52 accuracy on train set is: 0.987118320610687
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	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

evaluating on test set...

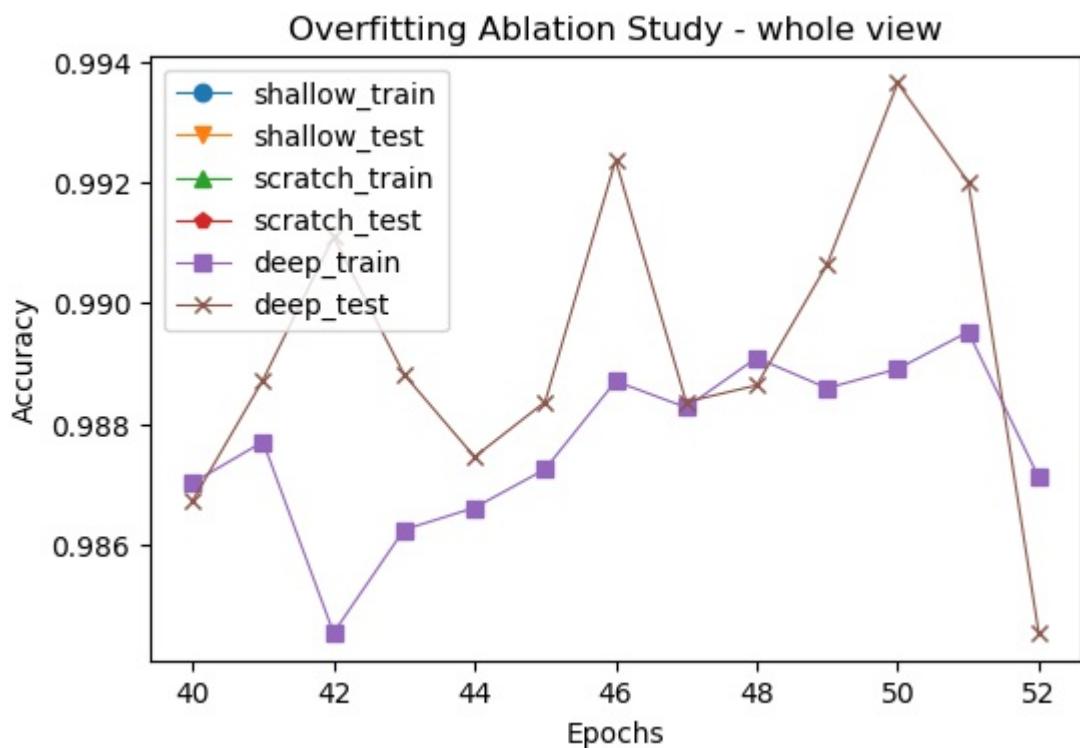
Accuracy on test images: 0.984551

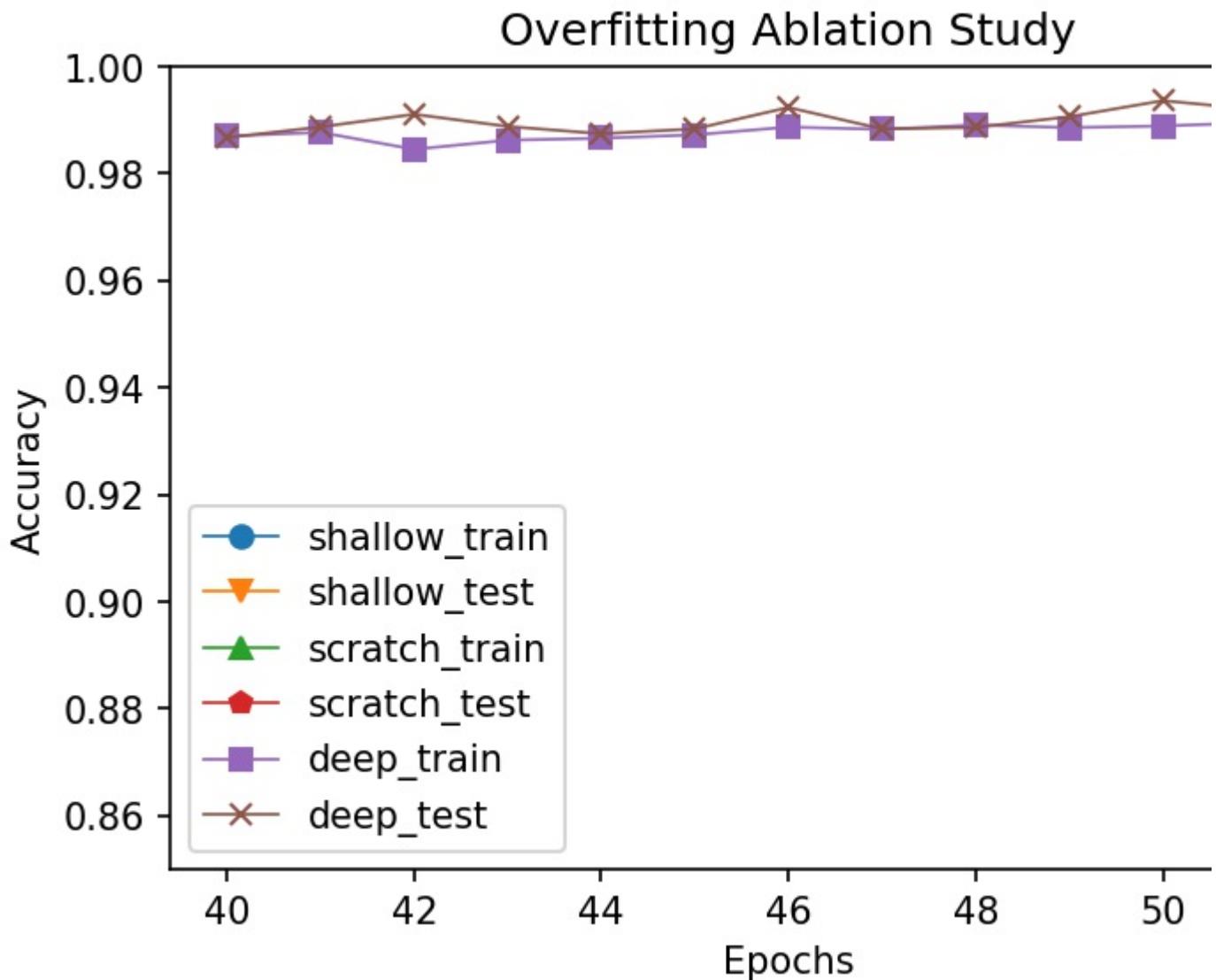
epoch 52 accuracy on test set is: 0.9845510723373319

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
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





```
epoch 52 takes 316 seconds.  
[54,    30] loss: 0.050  
[54,    60] loss: 0.085  
[54,    90] loss: 0.066  
[54,   120] loss: 0.022  
[54,   150] loss: 0.062  
[54,   180] loss: 0.060  
[54,   210] loss: 0.053  
[54,   240] loss: 0.086  
[54,   270] loss: 0.042  
[54,   300] loss: 0.040  
[54,   330] loss: 0.035  
[54,   360] loss: 0.026  
[54,   390] loss: 0.023  
[54,   420] loss: 0.025  
[54,   450] loss: 0.092  
[54,   480] loss: 0.025  
[54,   510] loss: 0.041  
[54,   540] loss: 0.023  
[54,   570] loss: 0.064  
[54,   600] loss: 0.046  
[54,   630] loss: 0.027  
[54,   660] loss: 0.071
```

```

[54, 690] loss: 0.064
[54, 720] loss: 0.036
[54, 750] loss: 0.071
[54, 780] loss: 0.031
[54, 810] loss: 0.060
[54, 840] loss: 0.062
[54, 870] loss: 0.075
[54, 900] loss: 0.036
[54, 930] loss: 0.035
[54, 960] loss: 0.076
[54, 990] loss: 0.052
[54, 1020] loss: 0.051
[54, 1050] loss: 0.030
[54, 1080] loss: 0.041
[54, 1110] loss: 0.051
[54, 1140] loss: 0.044
[54, 1170] loss: 0.030
[54, 1200] loss: 0.049
[54, 1230] loss: 0.058
[54, 1260] loss: 0.068
[54, 1290] loss: 0.043
[54, 1320] loss: 0.048
[54, 1350] loss: 0.057
[54, 1380] loss: 0.041
[54, 1410] loss: 0.053
[54, 1440] loss: 0.050
[54, 1470] loss: 0.034
[54, 1500] loss: 0.049
[54, 1530] loss: 0.054
[54, 1560] loss: 0.052
[54, 1590] loss: 0.040
[54, 1620] loss: 0.048
[54, 1650] loss: 0.060
[54, 1680] loss: 0.043
[54, 1710] loss: 0.036
[54, 1740] loss: 0.083
[54, 1770] loss: 0.022
[54, 1800] loss: 0.052
[54, 1830] loss: 0.041
[54, 1860] loss: 0.054
[54, 1890] loss: 0.040
[54, 1920] loss: 0.039
[54, 1950] loss: 0.036
[54, 1980] loss: 0.049
[54, 2010] loss: 0.047
[54, 2040] loss: 0.041
[54, 2070] loss: 0.048
[54, 2100] loss: 0.036
[54, 2130] loss: 0.034
[54, 2160] loss: 0.059
[54, 2190] loss: 0.046
ovft_intermediate_models/epoch53_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989799
epoch 53 accuracy on train set is: 0.9897991639403854
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	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

evaluating on test set...

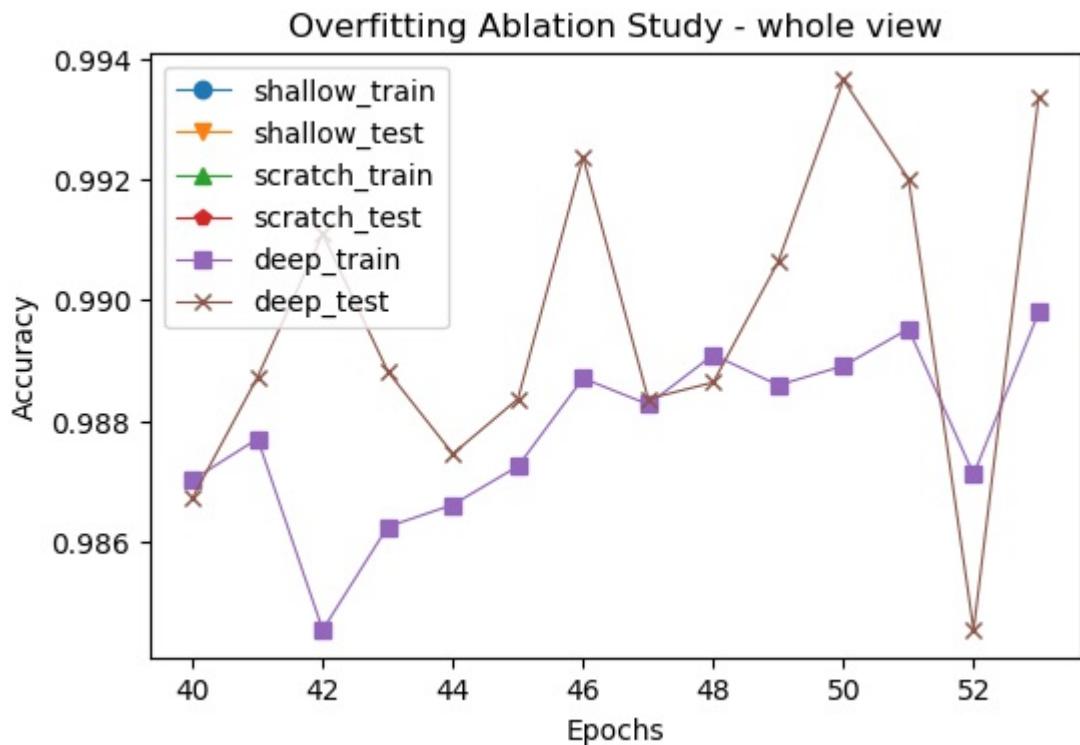
Accuracy on test images: 0.993366

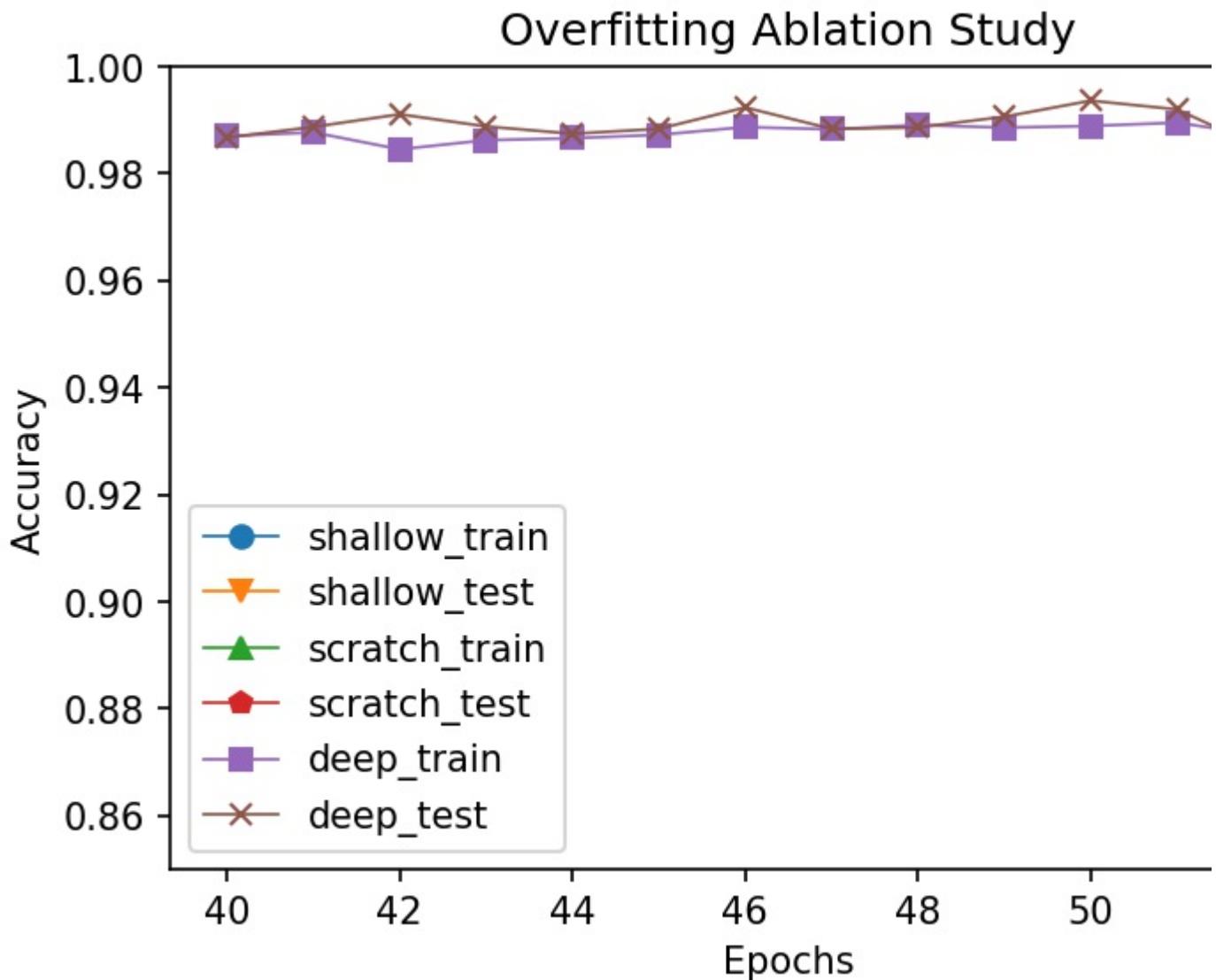
epoch 53 accuracy on test set is: 0.9933660487095601

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
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





```
epoch 53 takes 316 seconds.  
[55,    30] loss: 0.064  
[55,    60] loss: 0.058  
[55,   90] loss: 0.048  
[55,  120] loss: 0.048  
[55,  150] loss: 0.038  
[55,  180] loss: 0.036  
[55,  210] loss: 0.033  
[55,  240] loss: 0.040  
[55,  270] loss: 0.077  
[55,  300] loss: 0.046  
[55,  330] loss: 0.058  
[55,  360] loss: 0.064  
[55,  390] loss: 0.059  
[55,  420] loss: 0.034  
[55,  450] loss: 0.029  
[55,  480] loss: 0.037  
[55,  510] loss: 0.058  
[55,  540] loss: 0.040  
[55,  570] loss: 0.053  
[55,  600] loss: 0.047  
[55,  630] loss: 0.058  
[55,  660] loss: 0.041
```

```

[55, 690] loss: 0.037
[55, 720] loss: 0.047
[55, 750] loss: 0.072
[55, 780] loss: 0.056
[55, 810] loss: 0.090
[55, 840] loss: 0.053
[55, 870] loss: 0.048
[55, 900] loss: 0.045
[55, 930] loss: 0.060
[55, 960] loss: 0.061
[55, 990] loss: 0.045
[55, 1020] loss: 0.039
[55, 1050] loss: 0.015
[55, 1080] loss: 0.052
[55, 1110] loss: 0.043
[55, 1140] loss: 0.048
[55, 1170] loss: 0.041
[55, 1200] loss: 0.093
[55, 1230] loss: 0.036
[55, 1260] loss: 0.046
[55, 1290] loss: 0.063
[55, 1320] loss: 0.080
[55, 1350] loss: 0.052
[55, 1380] loss: 0.051
[55, 1410] loss: 0.034
[55, 1440] loss: 0.063
[55, 1470] loss: 0.055
[55, 1500] loss: 0.034
[55, 1530] loss: 0.069
[55, 1560] loss: 0.059
[55, 1590] loss: 0.042
[55, 1620] loss: 0.048
[55, 1650] loss: 0.035
[55, 1680] loss: 0.020
[55, 1710] loss: 0.079
[55, 1740] loss: 0.065
[55, 1770] loss: 0.065
[55, 1800] loss: 0.043
[55, 1830] loss: 0.046
[55, 1860] loss: 0.063
[55, 1890] loss: 0.086
[55, 1920] loss: 0.048
[55, 1950] loss: 0.067
[55, 1980] loss: 0.050
[55, 2010] loss: 0.027
[55, 2040] loss: 0.050
[55, 2070] loss: 0.038
[55, 2100] loss: 0.047
[55, 2130] loss: 0.029
[55, 2160] loss: 0.047
[55, 2190] loss: 0.077
ovft_intermediate_models/epoch54_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989981
epoch 54 accuracy on train set is: 0.9899809160305344
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	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

evaluating on test set...

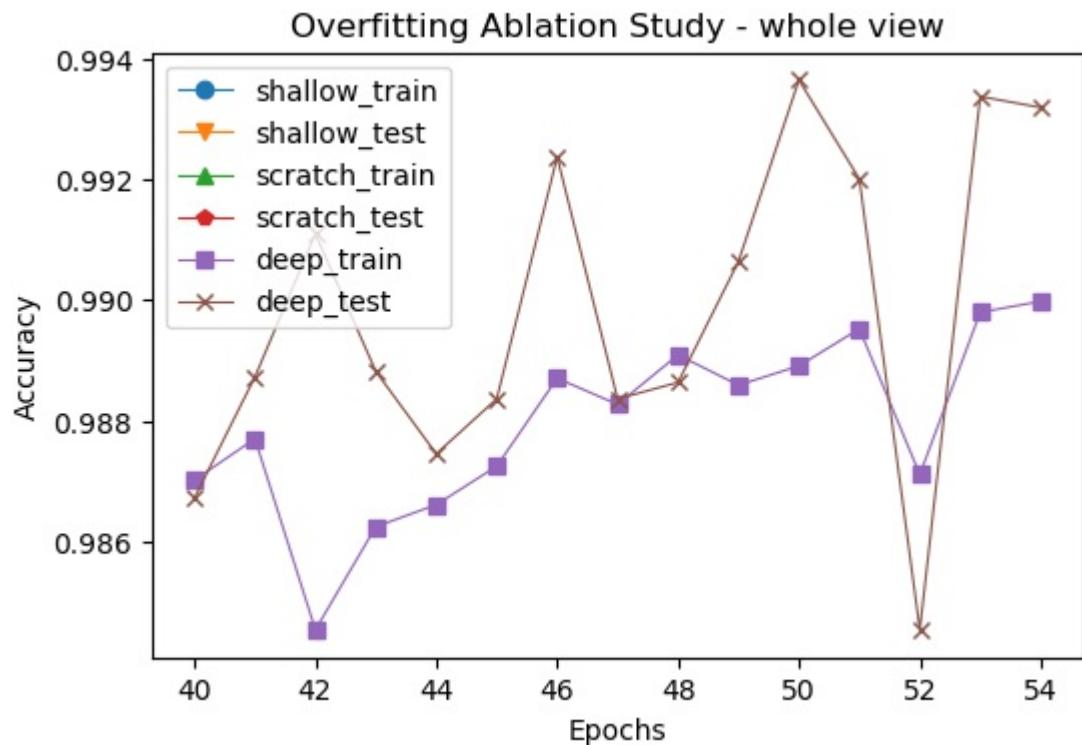
Accuracy on test images: 0.993184

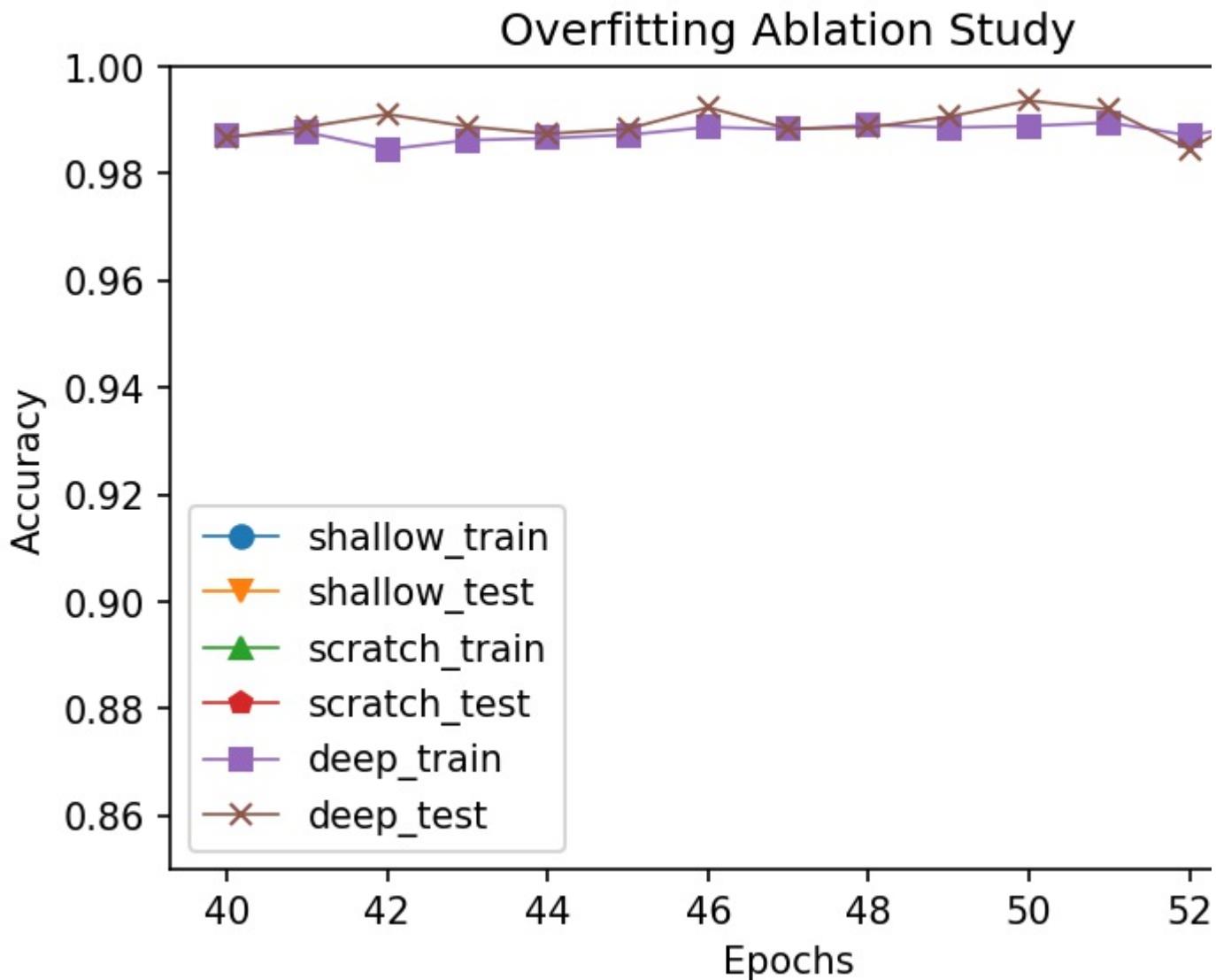
epoch 54 accuracy on test set is: 0.9931842966194111

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

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





```
epoch 54 takes 316 seconds.  
[56,    30] loss: 0.061  
[56,    60] loss: 0.088  
[56,    90] loss: 0.052  
[56,   120] loss: 0.047  
[56,   150] loss: 0.063  
[56,   180] loss: 0.042  
[56,   210] loss: 0.020  
[56,   240] loss: 0.030  
[56,   270] loss: 0.089  
[56,   300] loss: 0.080  
[56,   330] loss: 0.050  
[56,   360] loss: 0.063  
[56,   390] loss: 0.059  
[56,   420] loss: 0.046  
[56,   450] loss: 0.048  
[56,   480] loss: 0.037  
[56,   510] loss: 0.047  
[56,   540] loss: 0.063  
[56,   570] loss: 0.058  
[56,   600] loss: 0.078  
[56,   630] loss: 0.053  
[56,   660] loss: 0.040
```

```

[56, 690] loss: 0.046
[56, 720] loss: 0.042
[56, 750] loss: 0.047
[56, 780] loss: 0.061
[56, 810] loss: 0.054
[56, 840] loss: 0.043
[56, 870] loss: 0.043
[56, 900] loss: 0.041
[56, 930] loss: 0.031
[56, 960] loss: 0.029
[56, 990] loss: 0.040
[56, 1020] loss: 0.049
[56, 1050] loss: 0.068
[56, 1080] loss: 0.045
[56, 1110] loss: 0.054
[56, 1140] loss: 0.079
[56, 1170] loss: 0.057
[56, 1200] loss: 0.053
[56, 1230] loss: 0.042
[56, 1260] loss: 0.024
[56, 1290] loss: 0.030
[56, 1320] loss: 0.053
[56, 1350] loss: 0.063
[56, 1380] loss: 0.036
[56, 1410] loss: 0.071
[56, 1440] loss: 0.045
[56, 1470] loss: 0.048
[56, 1500] loss: 0.058
[56, 1530] loss: 0.052
[56, 1560] loss: 0.029
[56, 1590] loss: 0.053
[56, 1620] loss: 0.023
[56, 1650] loss: 0.053
[56, 1680] loss: 0.043
[56, 1710] loss: 0.055
[56, 1740] loss: 0.058
[56, 1770] loss: 0.064
[56, 1800] loss: 0.033
[56, 1830] loss: 0.059
[56, 1860] loss: 0.066
[56, 1890] loss: 0.062
[56, 1920] loss: 0.047
[56, 1950] loss: 0.063
[56, 1980] loss: 0.044
[56, 2010] loss: 0.058
[56, 2040] loss: 0.062
[56, 2070] loss: 0.047
[56, 2100] loss: 0.075
[56, 2130] loss: 0.042
[56, 2160] loss: 0.031
[56, 2190] loss: 0.056
ovft_intermediate_models/epoch55_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.988822
epoch 55 accuracy on train set is: 0.9888222464558343
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	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

evaluating on test set...

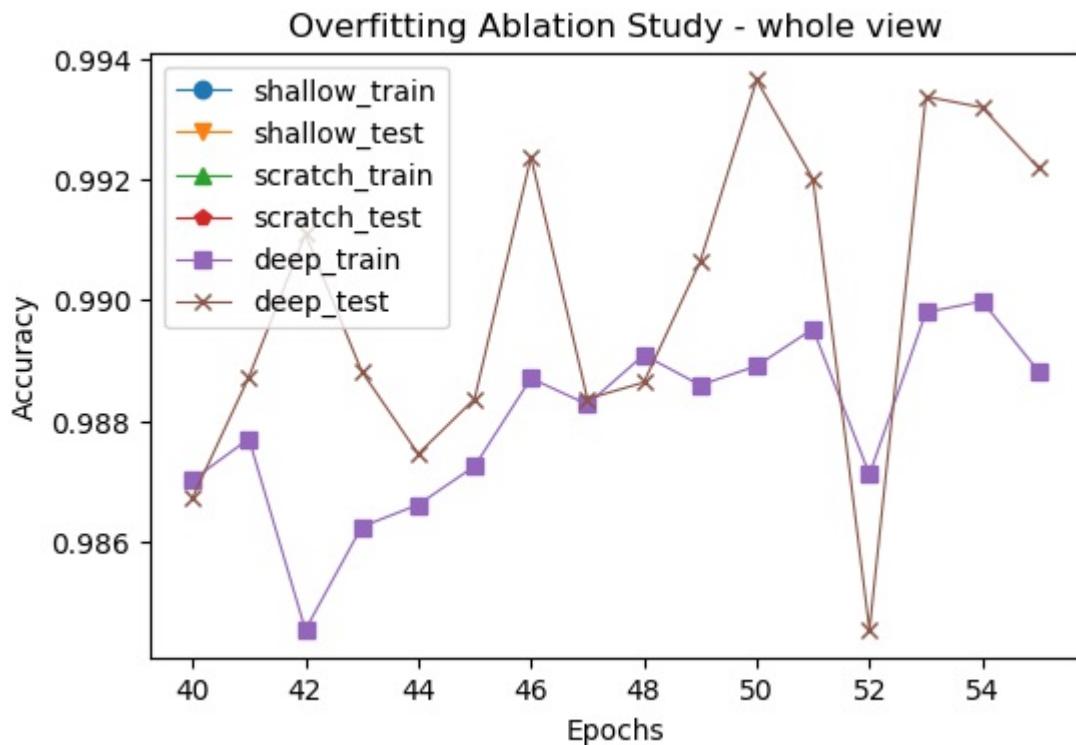
Accuracy on test images: 0.992185

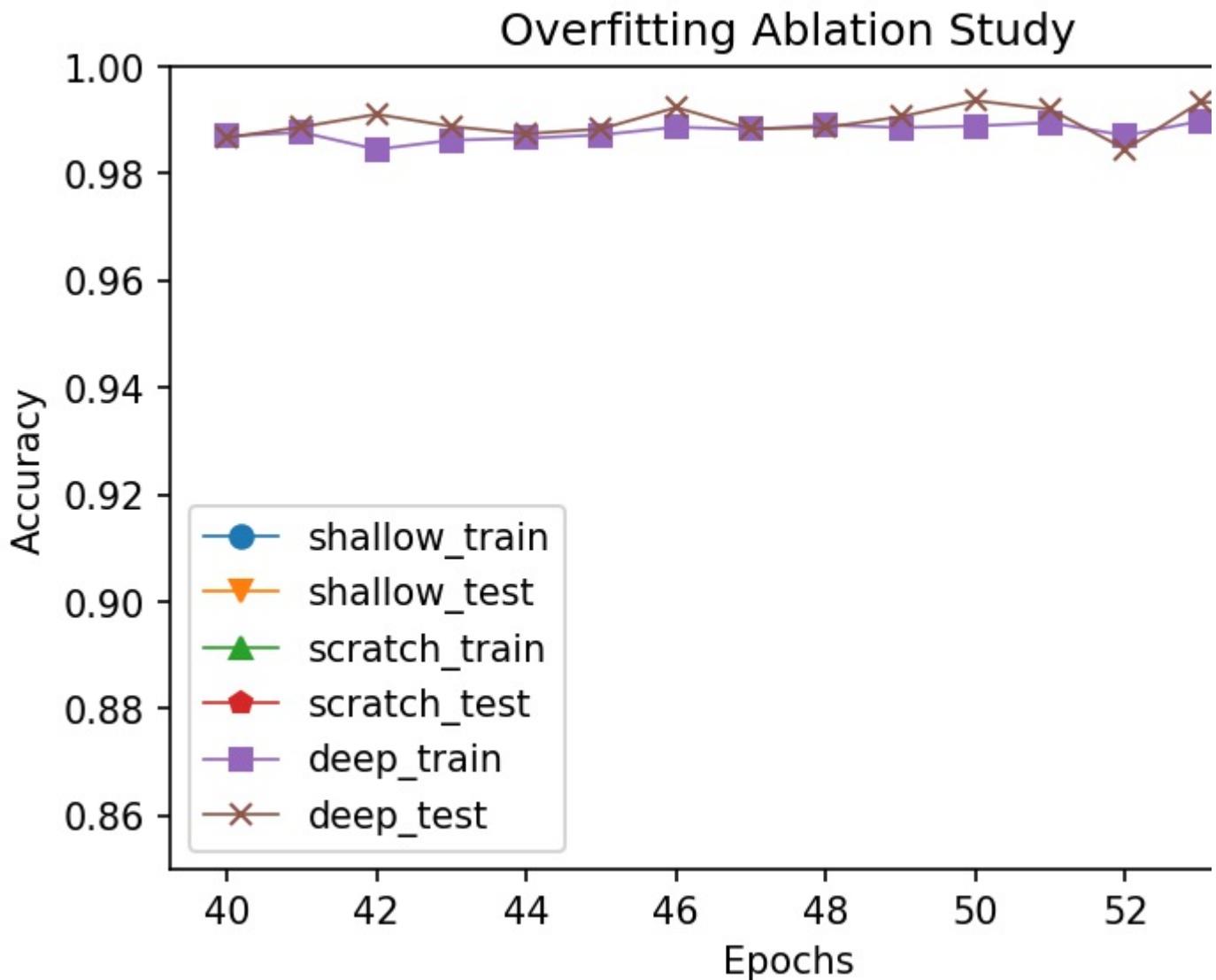
epoch 55 accuracy on test set is: 0.9921846601235914

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
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





```

epoch 55 takes 316 seconds.
[57,    30] loss: 0.049
[57,    60] loss: 0.047
[57,   90] loss: 0.035
[57,  120] loss: 0.021
[57,  150] loss: 0.043
[57,  180] loss: 0.050
[57,  210] loss: 0.021
[57,  240] loss: 0.064
[57,  270] loss: 0.054
[57,  300] loss: 0.033
[57,  330] loss: 0.061
[57,  360] loss: 0.051
[57,  390] loss: 0.056
[57,  420] loss: 0.069
[57,  450] loss: 0.050
[57,  480] loss: 0.067
[57,  510] loss: 0.041
[57,  540] loss: 0.087
[57,  570] loss: 0.038
[57,  600] loss: 0.048
[57,  630] loss: 0.011
[57,  660] loss: 0.033

```

```

[57,   690] loss: 0.067
[57,   720] loss: 0.042
[57,   750] loss: 0.048
[57,   780] loss: 0.068
[57,   810] loss: 0.088
[57,   840] loss: 0.069
[57,   870] loss: 0.046
[57,   900] loss: 0.056
[57,   930] loss: 0.037
[57,   960] loss: 0.078
[57,   990] loss: 0.022
[57,  1020] loss: 0.029
[57,  1050] loss: 0.032
[57,  1080] loss: 0.030
[57,  1110] loss: 0.039
[57,  1140] loss: 0.022
[57,  1170] loss: 0.038
[57,  1200] loss: 0.050
[57,  1230] loss: 0.041
[57,  1260] loss: 0.067
[57,  1290] loss: 0.045
[57,  1320] loss: 0.037
[57,  1350] loss: 0.012
[57,  1380] loss: 0.038
[57,  1410] loss: 0.038
[57,  1440] loss: 0.034
[57,  1470] loss: 0.072
[57,  1500] loss: 0.026
[57,  1530] loss: 0.029
[57,  1560] loss: 0.021
[57,  1590] loss: 0.062
[57,  1620] loss: 0.055
[57,  1650] loss: 0.040
[57,  1680] loss: 0.043
[57,  1710] loss: 0.029
[57,  1740] loss: 0.052
[57,  1770] loss: 0.093
[57,  1800] loss: 0.039
[57,  1830] loss: 0.038
[57,  1860] loss: 0.050
[57,  1890] loss: 0.044
[57,  1920] loss: 0.043
[57,  1950] loss: 0.086
[57,  1980] loss: 0.054
[57,  2010] loss: 0.036
[57,  2040] loss: 0.041
[57,  2070] loss: 0.038
[57,  2100] loss: 0.060
[57,  2130] loss: 0.077
[57,  2160] loss: 0.060
[57,  2190] loss: 0.059
ovft_intermediate_models/epoch56_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.990322
epoch 56 accuracy on train set is: 0.9903217011995638
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027  0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	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

evaluating on test set...

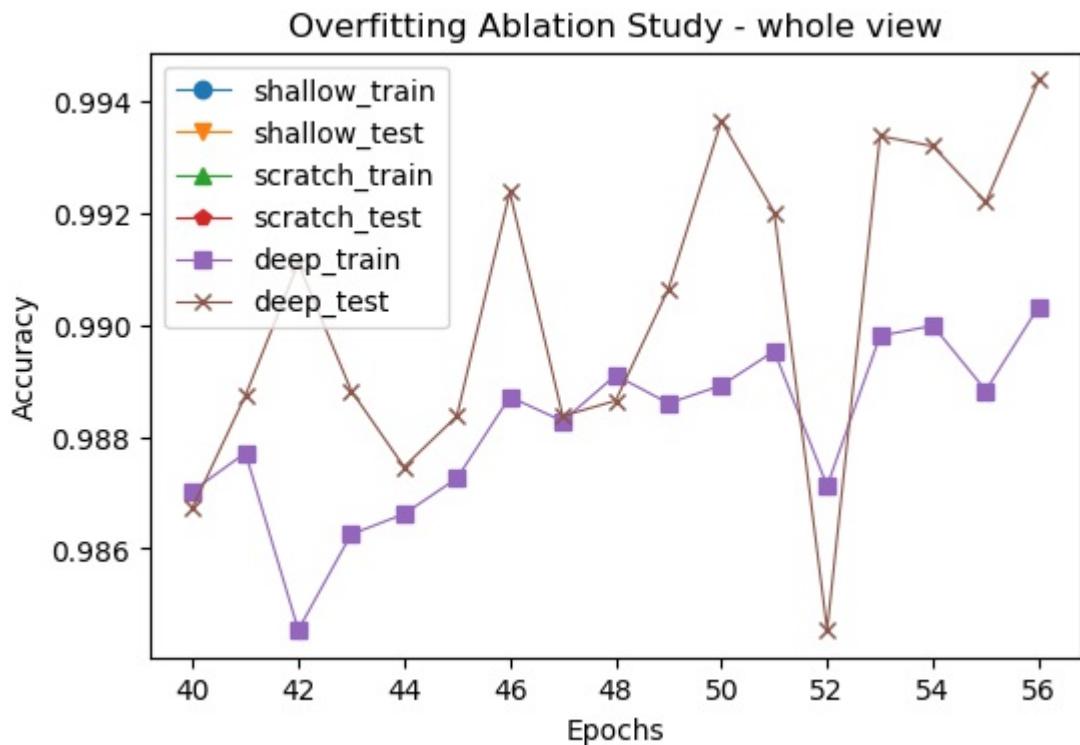
Accuracy on test images: 0.994366

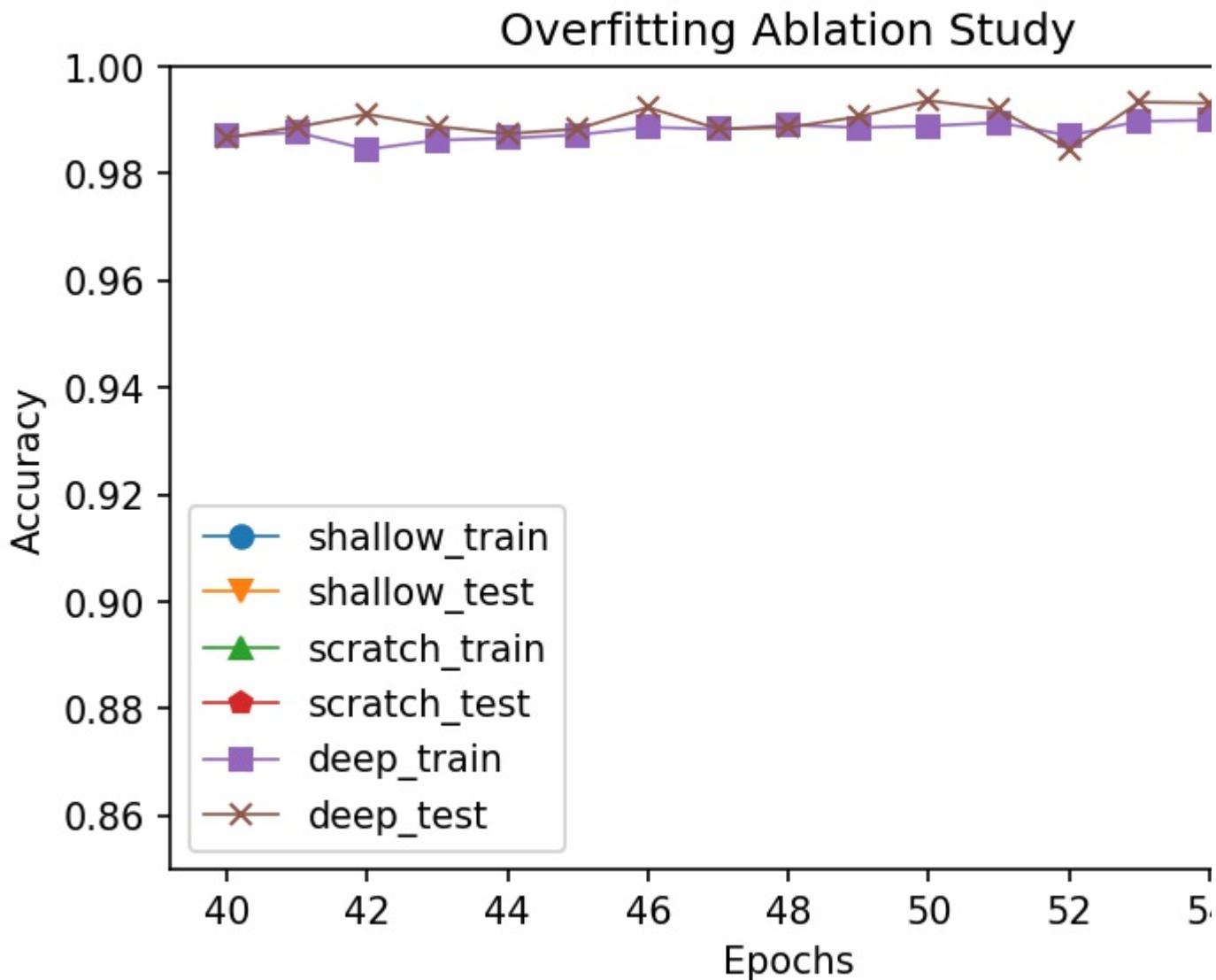
epoch 56 accuracy on test set is: 0.9943656852053798

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
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





```
epoch 56 takes 316 seconds.  
[58,    30] loss: 0.070  
[58,    60] loss: 0.070  
[58,   90] loss: 0.069  
[58,  120] loss: 0.045  
[58,  150] loss: 0.054  
[58,  180] loss: 0.031  
[58,  210] loss: 0.060  
[58,  240] loss: 0.050  
[58,  270] loss: 0.046  
[58,  300] loss: 0.060  
[58,  330] loss: 0.063  
[58,  360] loss: 0.043  
[58,  390] loss: 0.045  
[58,  420] loss: 0.079  
[58,  450] loss: 0.051  
[58,  480] loss: 0.029  
[58,  510] loss: 0.037  
[58,  540] loss: 0.054  
[58,  570] loss: 0.051  
[58,  600] loss: 0.080  
[58,  630] loss: 0.072  
[58,  660] loss: 0.055
```

```

[58, 690] loss: 0.037
[58, 720] loss: 0.049
[58, 750] loss: 0.038
[58, 780] loss: 0.043
[58, 810] loss: 0.047
[58, 840] loss: 0.033
[58, 870] loss: 0.050
[58, 900] loss: 0.038
[58, 930] loss: 0.049
[58, 960] loss: 0.049
[58, 990] loss: 0.051
[58, 1020] loss: 0.036
[58, 1050] loss: 0.038
[58, 1080] loss: 0.033
[58, 1110] loss: 0.062
[58, 1140] loss: 0.040
[58, 1170] loss: 0.040
[58, 1200] loss: 0.079
[58, 1230] loss: 0.054
[58, 1260] loss: 0.037
[58, 1290] loss: 0.071
[58, 1320] loss: 0.047
[58, 1350] loss: 0.045
[58, 1380] loss: 0.032
[58, 1410] loss: 0.076
[58, 1440] loss: 0.052
[58, 1470] loss: 0.069
[58, 1500] loss: 0.042
[58, 1530] loss: 0.052
[58, 1560] loss: 0.051
[58, 1590] loss: 0.037
[58, 1620] loss: 0.055
[58, 1650] loss: 0.033
[58, 1680] loss: 0.026
[58, 1710] loss: 0.036
[58, 1740] loss: 0.054
[58, 1770] loss: 0.072
[58, 1800] loss: 0.036
[58, 1830] loss: 0.054
[58, 1860] loss: 0.064
[58, 1890] loss: 0.041
[58, 1920] loss: 0.042
[58, 1950] loss: 0.075
[58, 1980] loss: 0.024
[58, 2010] loss: 0.045
[58, 2040] loss: 0.046
[58, 2070] loss: 0.053
[58, 2100] loss: 0.023
[58, 2130] loss: 0.042
[58, 2160] loss: 0.050
[58, 2190] loss: 0.044
ovft_intermediate_models/epoch57_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989186
epoch 57 accuracy on train set is: 0.9891857506361323
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	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

evaluating on test set...

Accuracy on test images: 0.994093

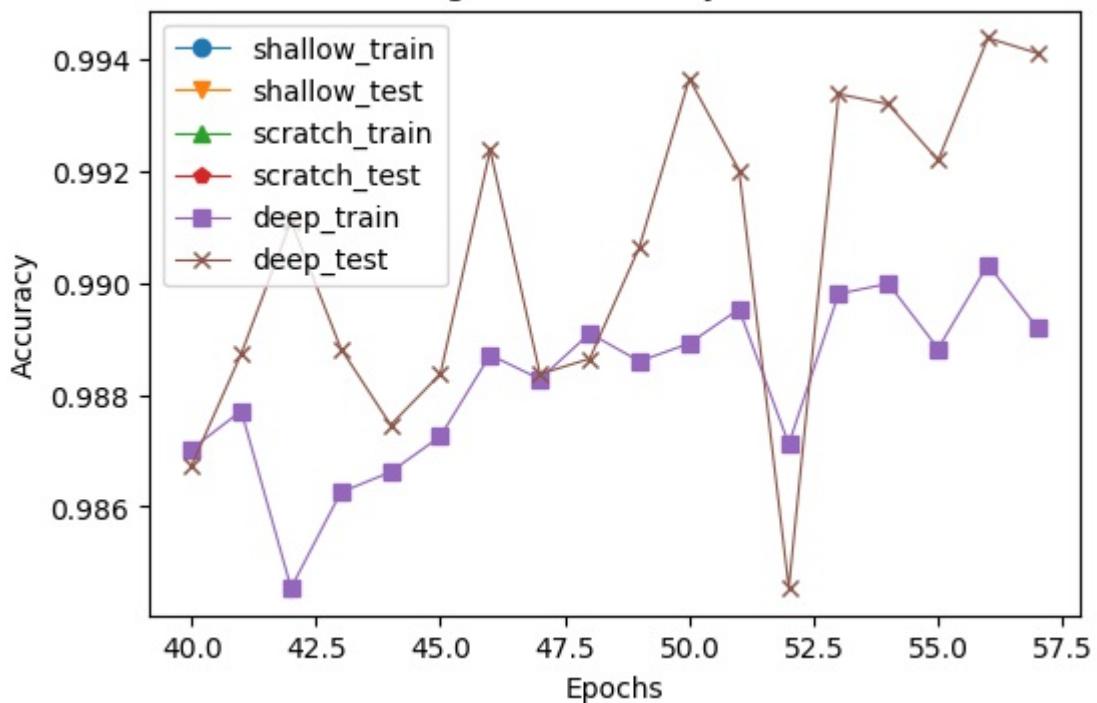
epoch 57 accuracy on test set is: 0.9940930570701563

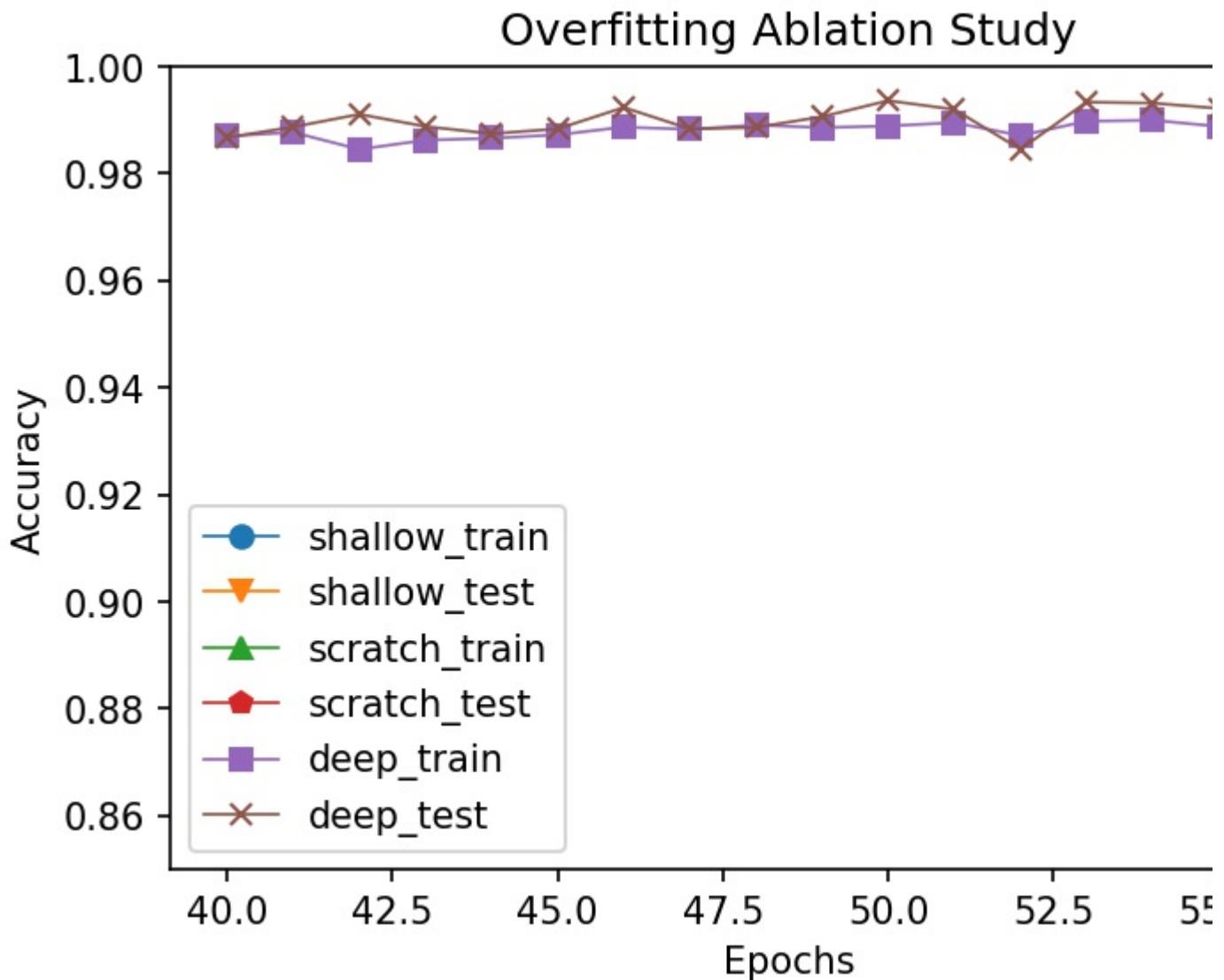
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
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

Overfitting Ablation Study - whole view





```

epoch 57 takes 316 seconds.
[59,    30] loss: 0.024
[59,    60] loss: 0.033
[59,   90] loss: 0.034
[59,  120] loss: 0.035
[59,  150] loss: 0.033
[59,  180] loss: 0.041
[59,  210] loss: 0.059
[59,  240] loss: 0.033
[59,  270] loss: 0.042
[59,  300] loss: 0.038
[59,  330] loss: 0.031
[59,  360] loss: 0.053
[59,  390] loss: 0.075
[59,  420] loss: 0.054
[59,  450] loss: 0.051
[59,  480] loss: 0.063
[59,  510] loss: 0.044
[59,  540] loss: 0.042
[59,  570] loss: 0.027
[59,  600] loss: 0.043
[59,  630] loss: 0.020
[59,  660] loss: 0.072

```

```
[59, 690] loss: 0.080
[59, 720] loss: 0.045
[59, 750] loss: 0.084
[59, 780] loss: 0.060
[59, 810] loss: 0.036
[59, 840] loss: 0.043
[59, 870] loss: 0.043
[59, 900] loss: 0.049
[59, 930] loss: 0.033
[59, 960] loss: 0.035
[59, 990] loss: 0.050
[59, 1020] loss: 0.046
[59, 1050] loss: 0.048
[59, 1080] loss: 0.049
[59, 1110] loss: 0.034
[59, 1140] loss: 0.051
[59, 1170] loss: 0.051
[59, 1200] loss: 0.064
[59, 1230] loss: 0.069
[59, 1260] loss: 0.035
[59, 1290] loss: 0.056
[59, 1320] loss: 0.062
[59, 1350] loss: 0.065
[59, 1380] loss: 0.043
[59, 1410] loss: 0.071
[59, 1440] loss: 0.054
[59, 1470] loss: 0.036
[59, 1500] loss: 0.050
[59, 1530] loss: 0.051
[59, 1560] loss: 0.061
[59, 1590] loss: 0.061
[59, 1620] loss: 0.047
[59, 1650] loss: 0.062
[59, 1680] loss: 0.051
[59, 1710] loss: 0.027
[59, 1740] loss: 0.061
[59, 1770] loss: 0.023
[59, 1800] loss: 0.059
[59, 1830] loss: 0.042
[59, 1860] loss: 0.054
[59, 1890] loss: 0.043
[59, 1920] loss: 0.055
[59, 1950] loss: 0.039
[59, 1980] loss: 0.069
[59, 2010] loss: 0.054
[59, 2040] loss: 0.079
[59, 2070] loss: 0.048
[59, 2100] loss: 0.056
[59, 2130] loss: 0.053
[59, 2160] loss: 0.029
[59, 2190] loss: 0.041
ovft_intermediate_models/epoch58_deep.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.989458
epoch 58 accuracy on train set is: 0.9894583787713559
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732
```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	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

evaluating on test set...

Accuracy on test images: 0.991458

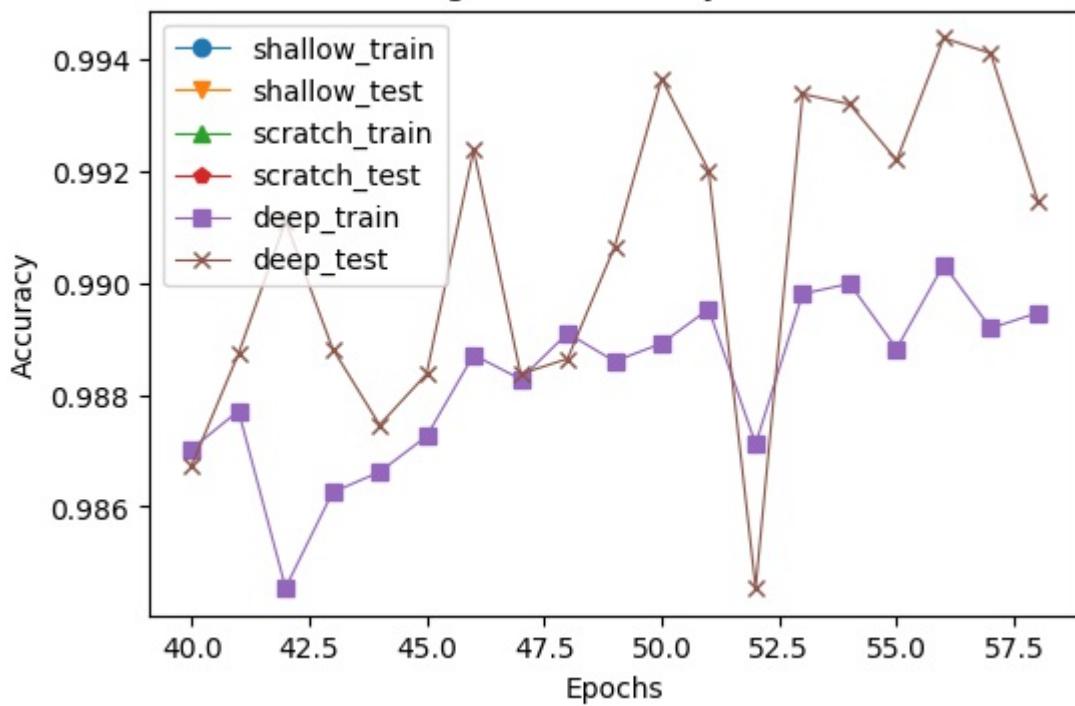
epoch 58 accuracy on test set is: 0.9914576517629953

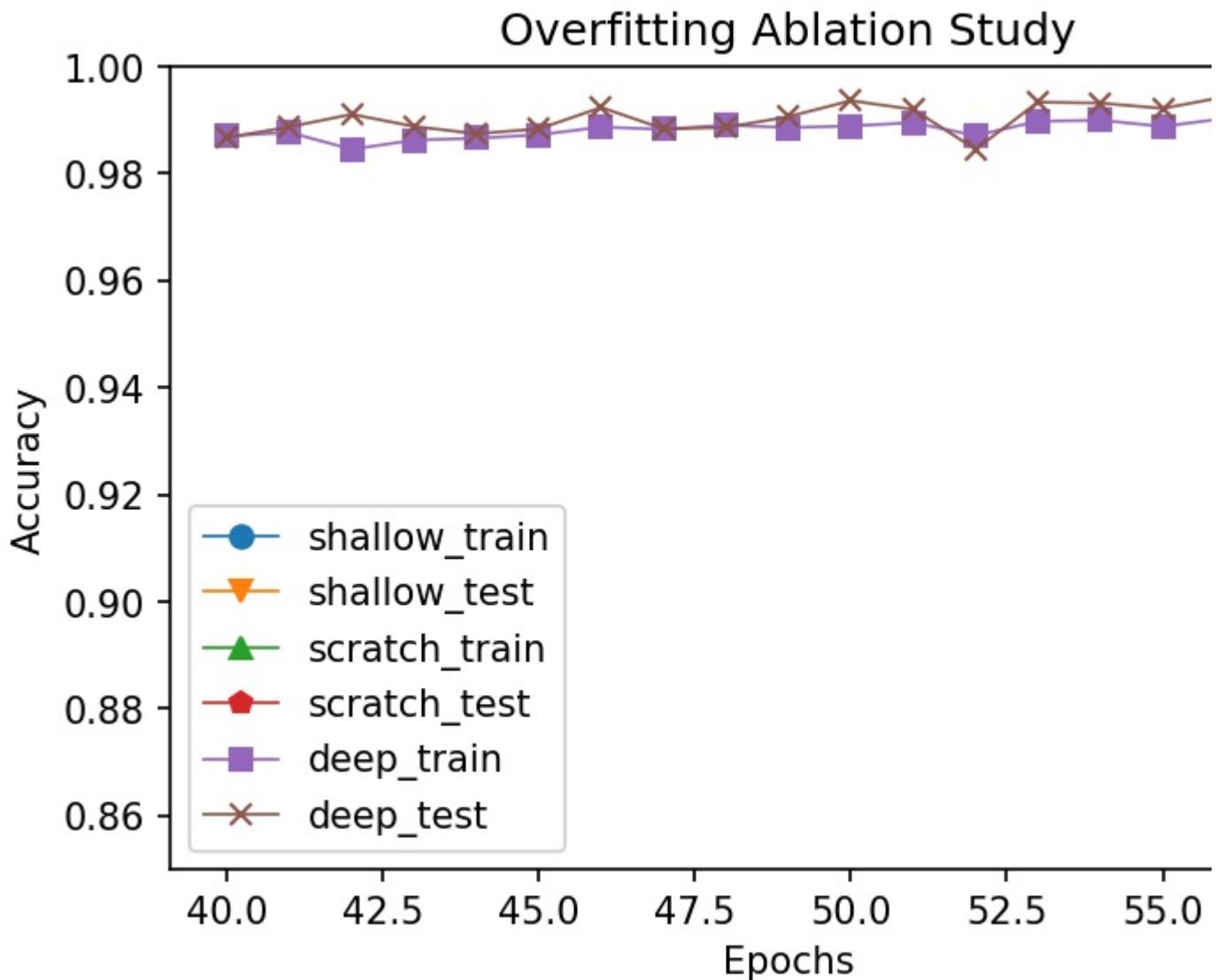
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	NaN	NaN
60	NaN	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view





```

epoch 58 takes 316 seconds.
[60,    30] loss: 0.049
[60,    60] loss: 0.061
[60,    90] loss: 0.052
[60,   120] loss: 0.038
[60,   150] loss: 0.081
[60,   180] loss: 0.068
[60,   210] loss: 0.060
[60,   240] loss: 0.054
[60,   270] loss: 0.039
[60,   300] loss: 0.042
[60,   330] loss: 0.037
[60,   360] loss: 0.026
[60,   390] loss: 0.045
[60,   420] loss: 0.033
[60,   450] loss: 0.041
[60,   480] loss: 0.035
[60,   510] loss: 0.029
[60,   540] loss: 0.037
[60,   570] loss: 0.057
[60,   600] loss: 0.057
[60,   630] loss: 0.038
[60,   660] loss: 0.028

```

```

[60, 690] loss: 0.040
[60, 720] loss: 0.047
[60, 750] loss: 0.063
[60, 780] loss: 0.027
[60, 810] loss: 0.055
[60, 840] loss: 0.027
[60, 870] loss: 0.050
[60, 900] loss: 0.019
[60, 930] loss: 0.029
[60, 960] loss: 0.028
[60, 990] loss: 0.011
[60, 1020] loss: 0.031
[60, 1050] loss: 0.023
[60, 1080] loss: 0.054
[60, 1110] loss: 0.041
[60, 1140] loss: 0.052
[60, 1170] loss: 0.050
[60, 1200] loss: 0.055
[60, 1230] loss: 0.045
[60, 1260] loss: 0.030
[60, 1290] loss: 0.068
[60, 1320] loss: 0.043
[60, 1350] loss: 0.057
[60, 1380] loss: 0.054
[60, 1410] loss: 0.064
[60, 1440] loss: 0.071
[60, 1470] loss: 0.020
[60, 1500] loss: 0.052
[60, 1530] loss: 0.021
[60, 1560] loss: 0.060
[60, 1590] loss: 0.015
[60, 1620] loss: 0.018
[60, 1650] loss: 0.060
[60, 1680] loss: 0.018
[60, 1710] loss: 0.050
[60, 1740] loss: 0.024
[60, 1770] loss: 0.021
[60, 1800] loss: 0.061
[60, 1830] loss: 0.036
[60, 1860] loss: 0.069
[60, 1890] loss: 0.030
[60, 1920] loss: 0.044
[60, 1950] loss: 0.023
[60, 1980] loss: 0.024
[60, 2010] loss: 0.064
[60, 2040] loss: 0.058
[60, 2070] loss: 0.052
[60, 2100] loss: 0.064
[60, 2130] loss: 0.053
[60, 2160] loss: 0.037
[60, 2190] loss: 0.029
ovft_intermediate_models/epoch59_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989663
epoch 59 accuracy on train set is: 0.9896628498727735
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	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

evaluating on test set...

Accuracy on test images: 0.988095

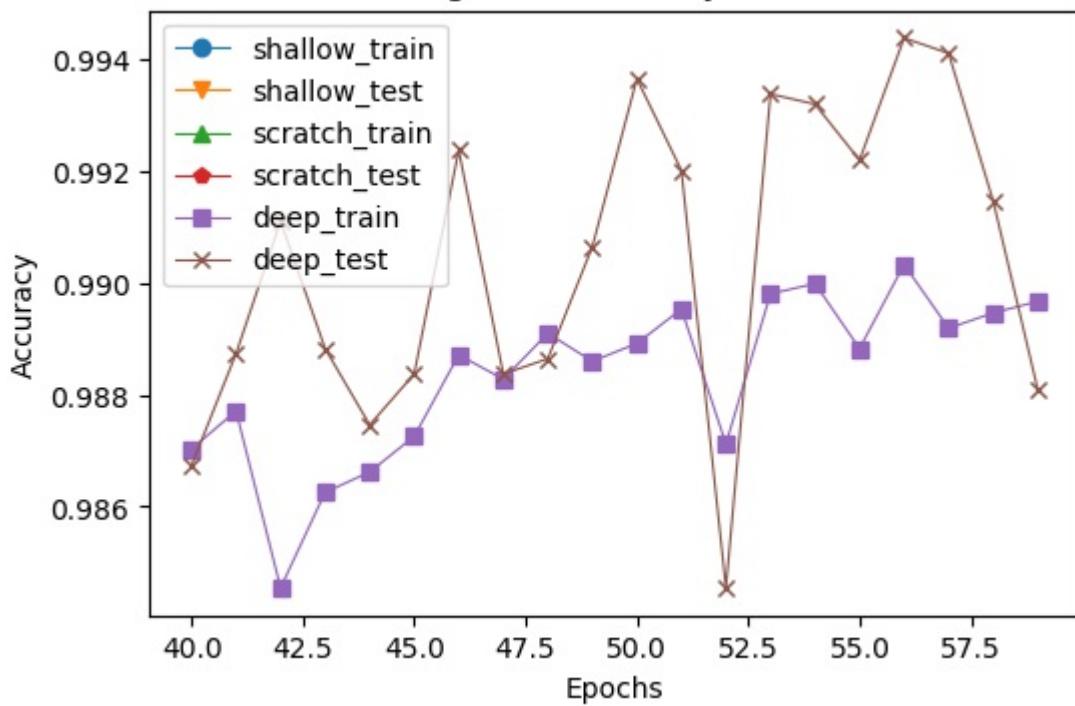
epoch 59 accuracy on test set is: 0.9880952380952381

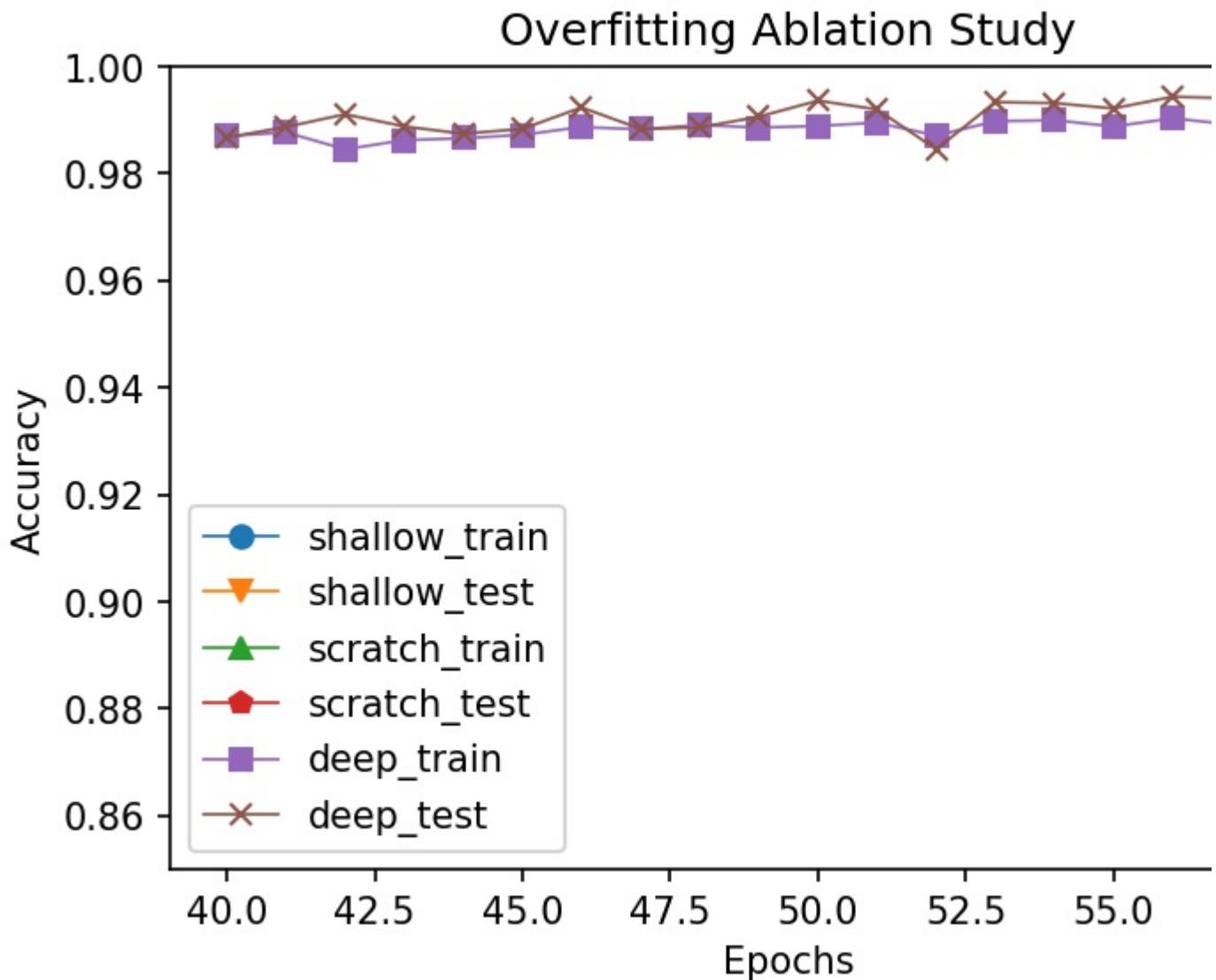
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	NaN	NaN
61	NaN	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view





```

epoch 59 takes 316 seconds.
[61,    30] loss: 0.069
[61,    60] loss: 0.066
[61,    90] loss: 0.080
[61,   120] loss: 0.040
[61,   150] loss: 0.039
[61,   180] loss: 0.059
[61,   210] loss: 0.044
[61,   240] loss: 0.050
[61,   270] loss: 0.047
[61,   300] loss: 0.089
[61,   330] loss: 0.045
[61,   360] loss: 0.110
[61,   390] loss: 0.044
[61,   420] loss: 0.056
[61,   450] loss: 0.035
[61,   480] loss: 0.042
[61,   510] loss: 0.052
[61,   540] loss: 0.033
[61,   570] loss: 0.030
[61,   600] loss: 0.057
[61,   630] loss: 0.058
[61,   660] loss: 0.028

```

```

[61, 690] loss: 0.056
[61, 720] loss: 0.019
[61, 750] loss: 0.038
[61, 780] loss: 0.029
[61, 810] loss: 0.033
[61, 840] loss: 0.057
[61, 870] loss: 0.064
[61, 900] loss: 0.035
[61, 930] loss: 0.029
[61, 960] loss: 0.053
[61, 990] loss: 0.066
[61, 1020] loss: 0.077
[61, 1050] loss: 0.038
[61, 1080] loss: 0.028
[61, 1110] loss: 0.042
[61, 1140] loss: 0.047
[61, 1170] loss: 0.039
[61, 1200] loss: 0.066
[61, 1230] loss: 0.058
[61, 1260] loss: 0.044
[61, 1290] loss: 0.043
[61, 1320] loss: 0.035
[61, 1350] loss: 0.045
[61, 1380] loss: 0.044
[61, 1410] loss: 0.098
[61, 1440] loss: 0.078
[61, 1470] loss: 0.036
[61, 1500] loss: 0.042
[61, 1530] loss: 0.060
[61, 1560] loss: 0.049
[61, 1590] loss: 0.064
[61, 1620] loss: 0.028
[61, 1650] loss: 0.049
[61, 1680] loss: 0.036
[61, 1710] loss: 0.030
[61, 1740] loss: 0.029
[61, 1770] loss: 0.036
[61, 1800] loss: 0.040
[61, 1830] loss: 0.026
[61, 1860] loss: 0.050
[61, 1890] loss: 0.067
[61, 1920] loss: 0.039
[61, 1950] loss: 0.044
[61, 1980] loss: 0.022
[61, 2010] loss: 0.044
[61, 2040] loss: 0.040
[61, 2070] loss: 0.051
[61, 2100] loss: 0.075
[61, 2130] loss: 0.035
[61, 2160] loss: 0.044
[61, 2190] loss: 0.033
ovft_intermediate_models/epoch60_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.987118
epoch 60 accuracy on train set is: 0.987118320610687
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	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

evaluating on test set...

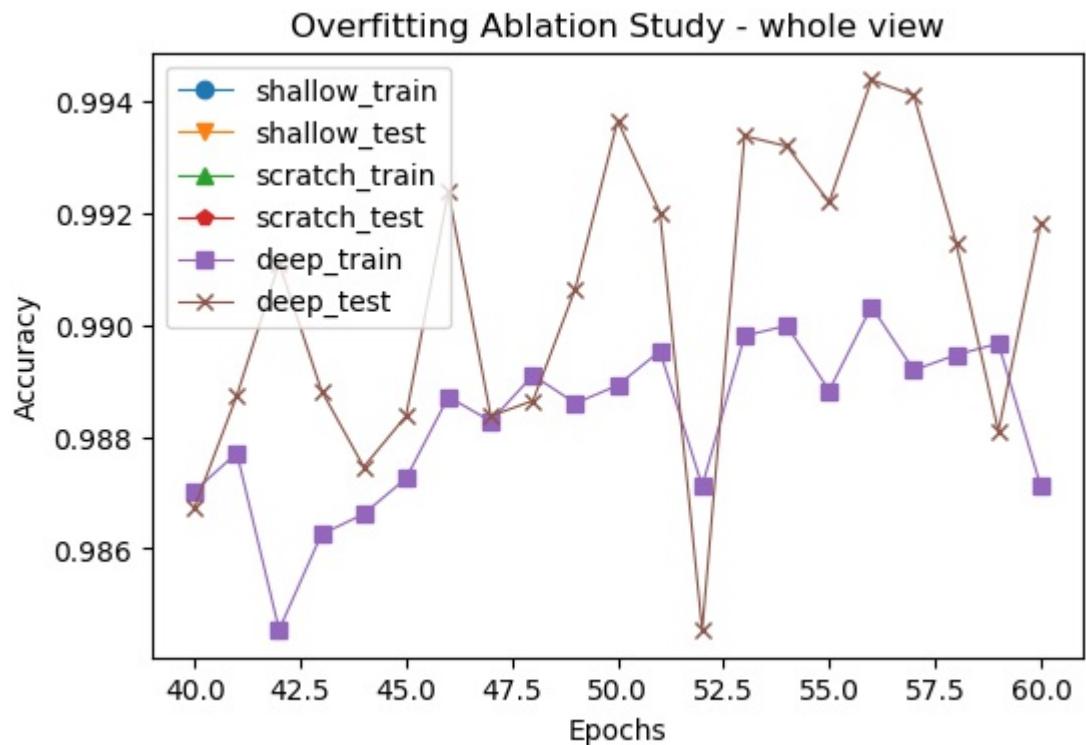
Accuracy on test images: 0.991821

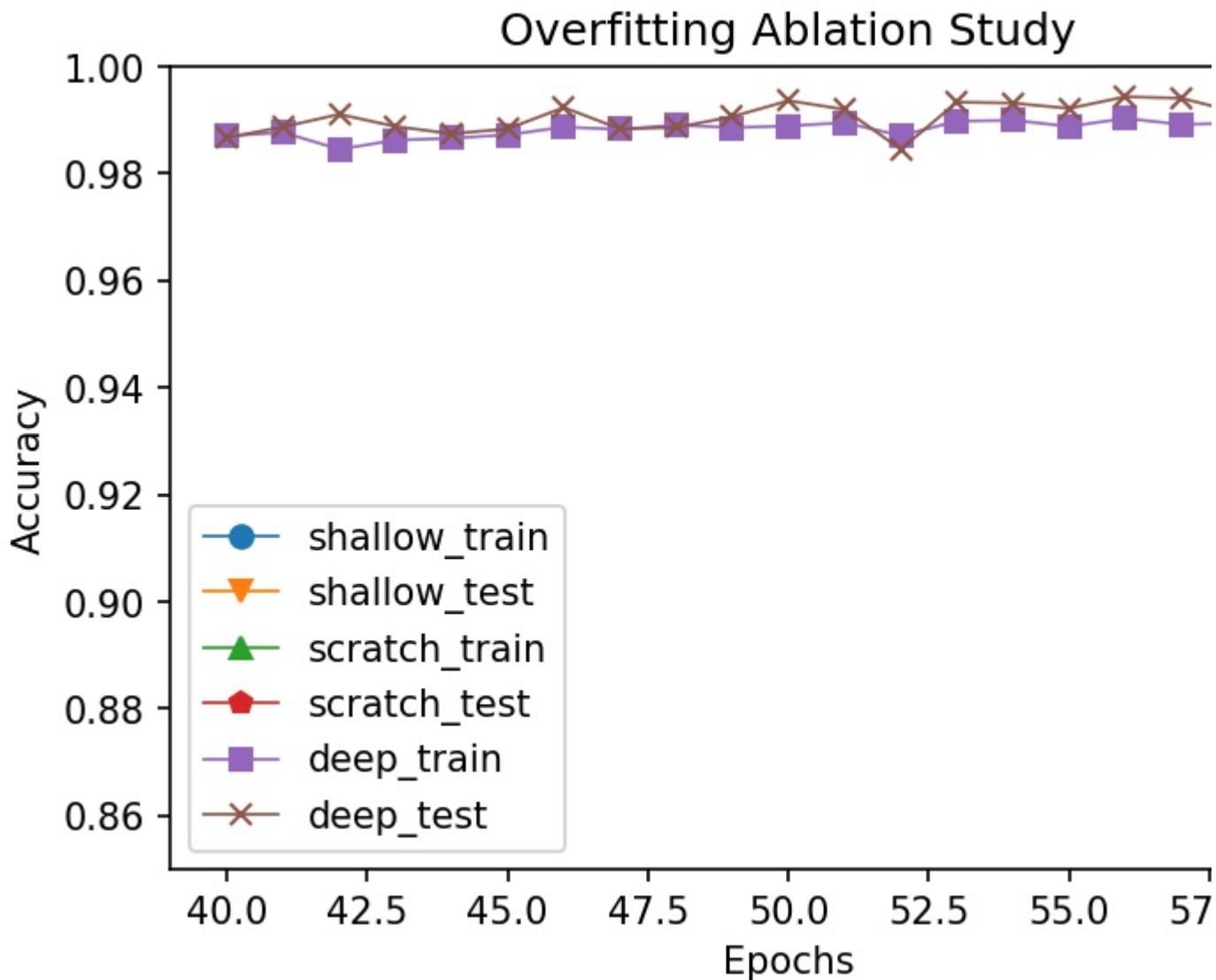
epoch 60 accuracy on test set is: 0.9918211559432933

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	NaN	NaN
62	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```

epoch 60 takes 316 seconds.
[62,    30] loss: 0.069
[62,    60] loss: 0.040
[62,    90] loss: 0.066
[62,   120] loss: 0.050
[62,   150] loss: 0.065
[62,   180] loss: 0.032
[62,   210] loss: 0.050
[62,   240] loss: 0.041
[62,   270] loss: 0.043
[62,   300] loss: 0.032
[62,   330] loss: 0.030
[62,   360] loss: 0.034
[62,   390] loss: 0.057
[62,   420] loss: 0.058
[62,   450] loss: 0.053
[62,   480] loss: 0.058
[62,   510] loss: 0.042
[62,   540] loss: 0.076
[62,   570] loss: 0.053
[62,   600] loss: 0.056
[62,   630] loss: 0.045
[62,   660] loss: 0.052

```

```

[62, 690] loss: 0.040
[62, 720] loss: 0.054
[62, 750] loss: 0.035
[62, 780] loss: 0.043
[62, 810] loss: 0.039
[62, 840] loss: 0.042
[62, 870] loss: 0.045
[62, 900] loss: 0.063
[62, 930] loss: 0.020
[62, 960] loss: 0.057
[62, 990] loss: 0.066
[62, 1020] loss: 0.035
[62, 1050] loss: 0.036
[62, 1080] loss: 0.062
[62, 1110] loss: 0.072
[62, 1140] loss: 0.031
[62, 1170] loss: 0.050
[62, 1200] loss: 0.049
[62, 1230] loss: 0.057
[62, 1260] loss: 0.045
[62, 1290] loss: 0.050
[62, 1320] loss: 0.032
[62, 1350] loss: 0.044
[62, 1380] loss: 0.058
[62, 1410] loss: 0.042
[62, 1440] loss: 0.040
[62, 1470] loss: 0.070
[62, 1500] loss: 0.063
[62, 1530] loss: 0.024
[62, 1560] loss: 0.049
[62, 1590] loss: 0.076
[62, 1620] loss: 0.053
[62, 1650] loss: 0.056
[62, 1680] loss: 0.076
[62, 1710] loss: 0.033
[62, 1740] loss: 0.031
[62, 1770] loss: 0.045
[62, 1800] loss: 0.038
[62, 1830] loss: 0.053
[62, 1860] loss: 0.058
[62, 1890] loss: 0.041
[62, 1920] loss: 0.030
[62, 1950] loss: 0.070
[62, 1980] loss: 0.052
[62, 2010] loss: 0.034
[62, 2040] loss: 0.034
[62, 2070] loss: 0.038
[62, 2100] loss: 0.042
[62, 2130] loss: 0.033
[62, 2160] loss: 0.041
[62, 2190] loss: 0.021
ovft_intermediate_models/epoch61_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.990435
epoch 61 accuracy on train set is: 0.990435296255907
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	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

evaluating on test set...

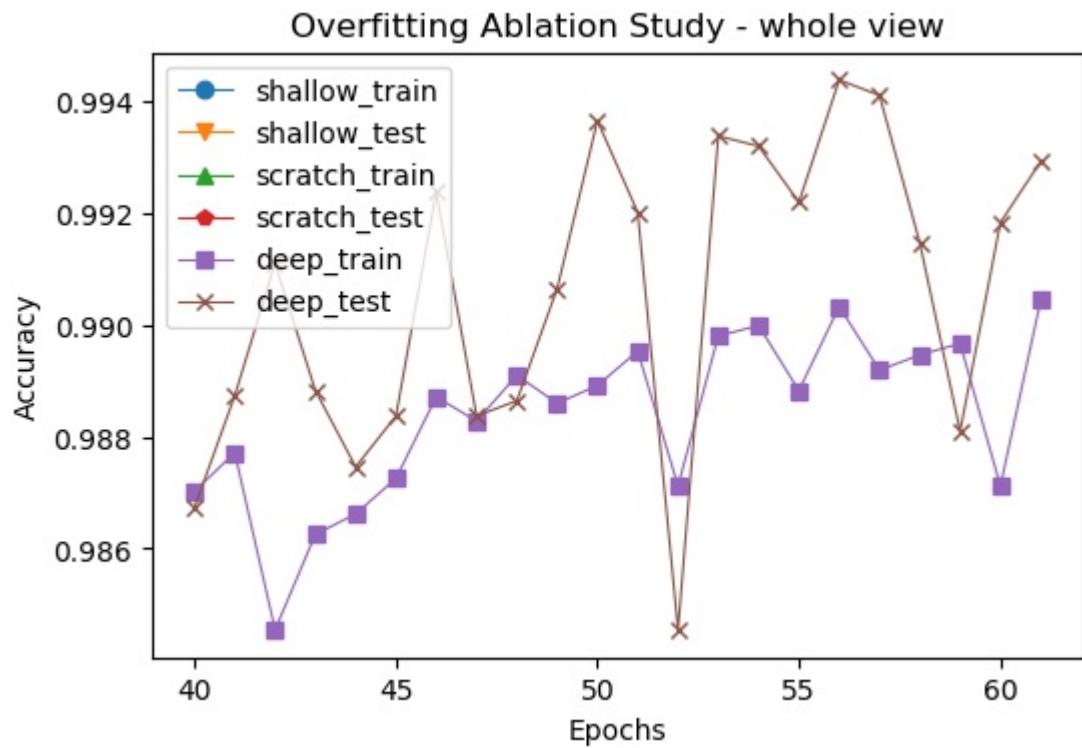
Accuracy on test images: 0.992912

epoch 61 accuracy on test set is: 0.9929116684841875

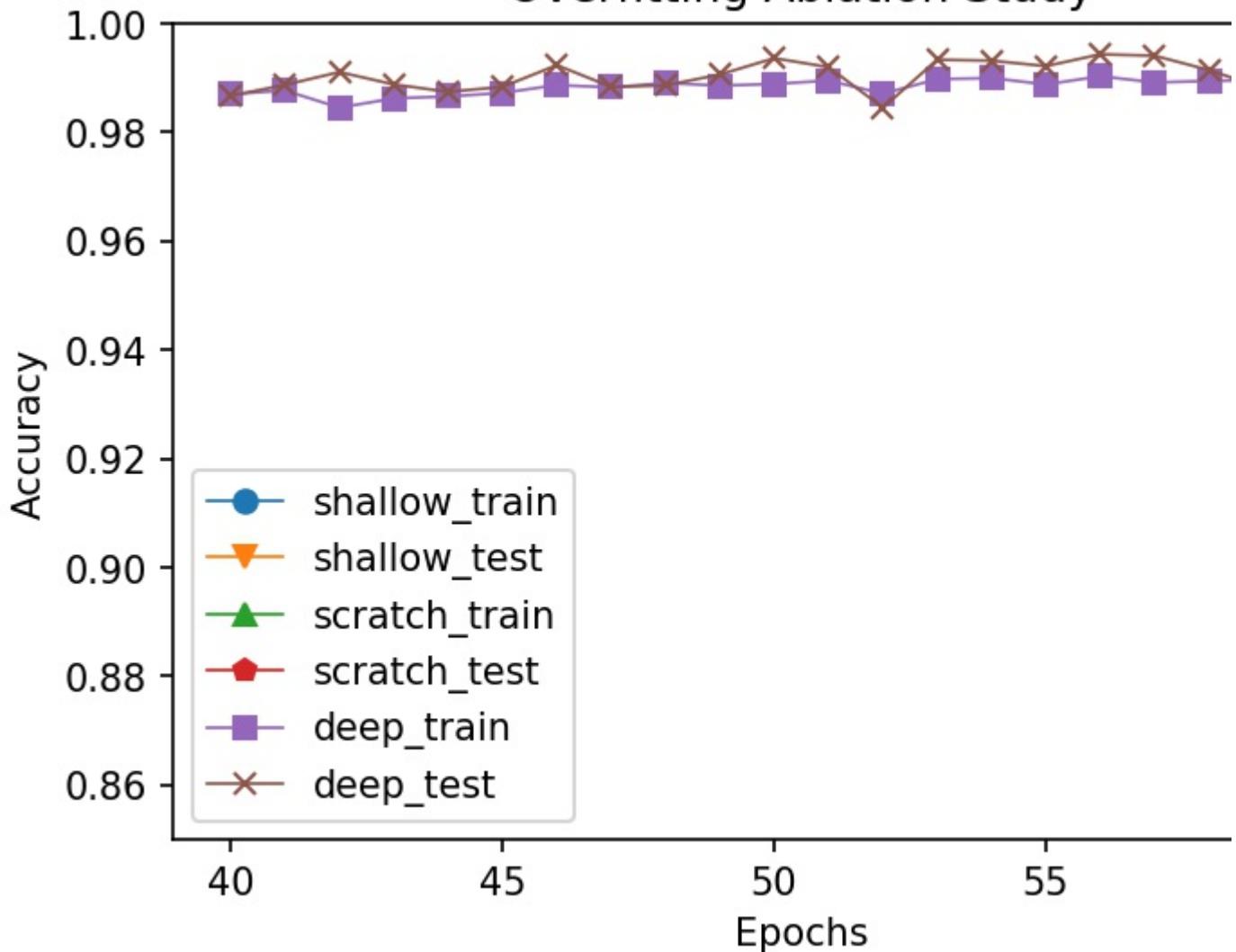
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63	NaN	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN



Overfitting Ablation Study



```
epoch 61 takes 316 seconds.  
[63,    30] loss: 0.055  
[63,    60] loss: 0.088  
[63,   90] loss: 0.024  
[63,  120] loss: 0.039  
[63,  150] loss: 0.052  
[63,  180] loss: 0.059  
[63,  210] loss: 0.064  
[63,  240] loss: 0.029  
[63,  270] loss: 0.021  
[63,  300] loss: 0.030  
[63,  330] loss: 0.043  
[63,  360] loss: 0.037  
[63,  390] loss: 0.062  
[63,  420] loss: 0.027  
[63,  450] loss: 0.019  
[63,  480] loss: 0.041  
[63,  510] loss: 0.036  
[63,  540] loss: 0.048  
[63,  570] loss: 0.045  
[63,  600] loss: 0.057  
[63,  630] loss: 0.046  
[63,  660] loss: 0.035
```

```

[63, 690] loss: 0.035
[63, 720] loss: 0.048
[63, 750] loss: 0.025
[63, 780] loss: 0.046
[63, 810] loss: 0.061
[63, 840] loss: 0.044
[63, 870] loss: 0.048
[63, 900] loss: 0.037
[63, 930] loss: 0.042
[63, 960] loss: 0.036
[63, 990] loss: 0.052
[63, 1020] loss: 0.067
[63, 1050] loss: 0.041
[63, 1080] loss: 0.061
[63, 1110] loss: 0.041
[63, 1140] loss: 0.059
[63, 1170] loss: 0.031
[63, 1200] loss: 0.041
[63, 1230] loss: 0.059
[63, 1260] loss: 0.026
[63, 1290] loss: 0.038
[63, 1320] loss: 0.025
[63, 1350] loss: 0.071
[63, 1380] loss: 0.084
[63, 1410] loss: 0.055
[63, 1440] loss: 0.041
[63, 1470] loss: 0.043
[63, 1500] loss: 0.047
[63, 1530] loss: 0.039
[63, 1560] loss: 0.036
[63, 1590] loss: 0.048
[63, 1620] loss: 0.059
[63, 1650] loss: 0.025
[63, 1680] loss: 0.039
[63, 1710] loss: 0.057
[63, 1740] loss: 0.032
[63, 1770] loss: 0.039
[63, 1800] loss: 0.032
[63, 1830] loss: 0.041
[63, 1860] loss: 0.046
[63, 1890] loss: 0.050
[63, 1920] loss: 0.019
[63, 1950] loss: 0.059
[63, 1980] loss: 0.057
[63, 2010] loss: 0.039
[63, 2040] loss: 0.015
[63, 2070] loss: 0.029
[63, 2100] loss: 0.043
[63, 2130] loss: 0.037
[63, 2160] loss: 0.049
[63, 2190] loss: 0.040
ovft_intermediate_models/epoch62_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989754
epoch 62 accuracy on train set is: 0.989753725917848
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	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

evaluating on test set...

Accuracy on test images: 0.991549

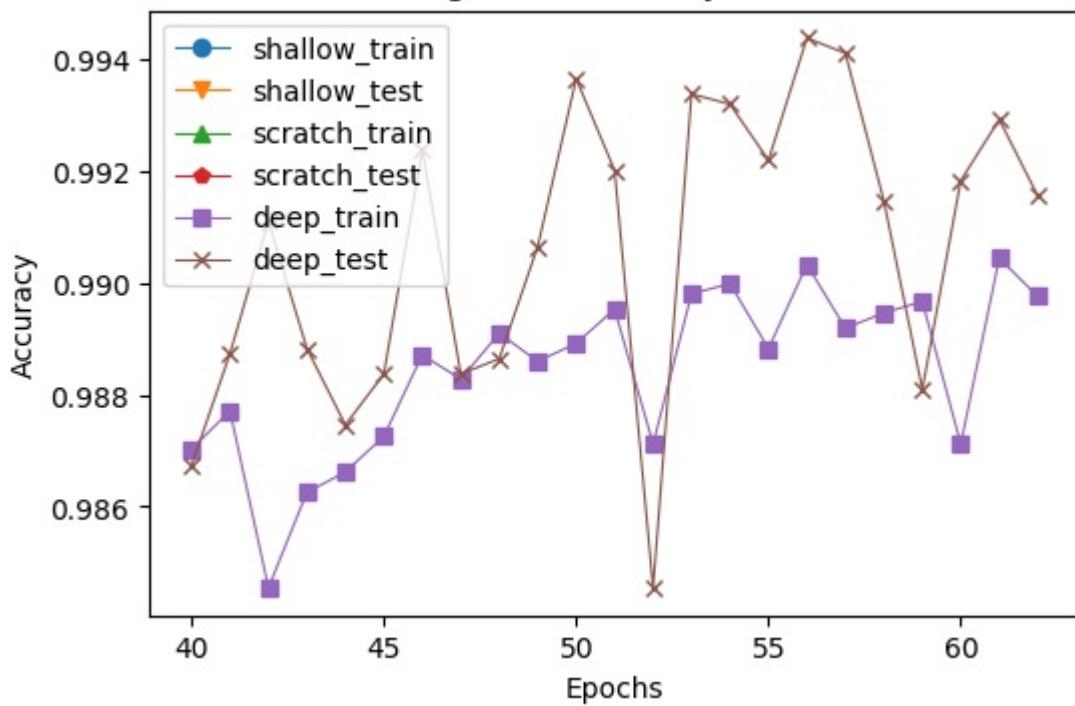
epoch 62 accuracy on test set is: 0.9915485278080698

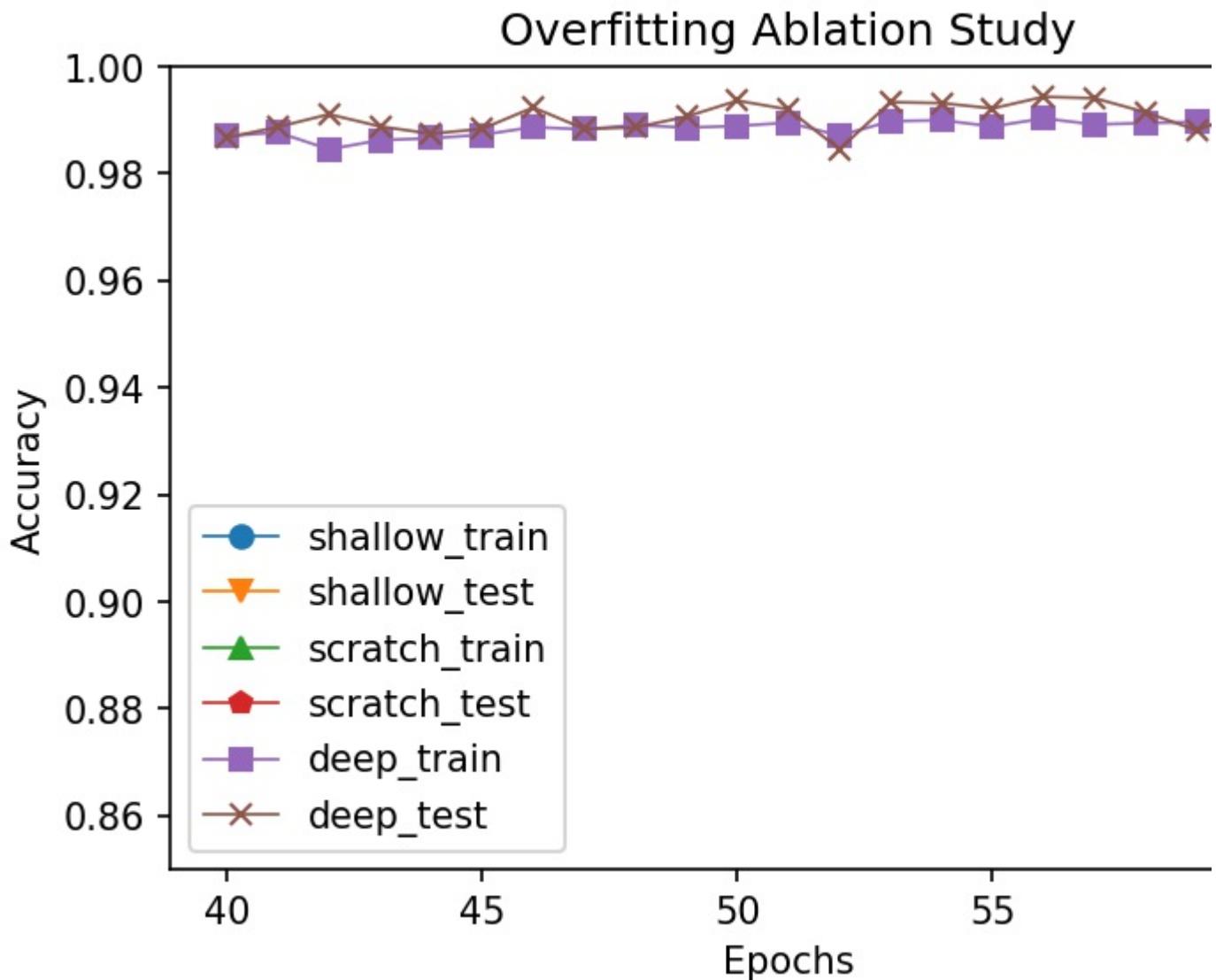
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	NaN	NaN
64	NaN	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view





```
epoch 62 takes 317 seconds.  
[64,     30] loss: 0.034  
[64,     60] loss: 0.038  
[64,    90] loss: 0.047  
[64,   120] loss: 0.061  
[64,   150] loss: 0.059  
[64,   180] loss: 0.038  
[64,   210] loss: 0.045  
[64,   240] loss: 0.053  
[64,   270] loss: 0.058  
[64,   300] loss: 0.043  
[64,   330] loss: 0.053  
[64,   360] loss: 0.030  
[64,   390] loss: 0.045  
[64,   420] loss: 0.039  
[64,   450] loss: 0.025  
[64,   480] loss: 0.052  
[64,   510] loss: 0.037  
[64,   540] loss: 0.040  
[64,   570] loss: 0.082  
[64,   600] loss: 0.053  
[64,   630] loss: 0.028  
[64,   660] loss: 0.031
```

```
[64, 690] loss: 0.039
[64, 720] loss: 0.060
[64, 750] loss: 0.054
[64, 780] loss: 0.046
[64, 810] loss: 0.063
[64, 840] loss: 0.031
[64, 870] loss: 0.055
[64, 900] loss: 0.054
[64, 930] loss: 0.071
[64, 960] loss: 0.033
[64, 990] loss: 0.042
[64, 1020] loss: 0.030
[64, 1050] loss: 0.039
[64, 1080] loss: 0.042
[64, 1110] loss: 0.055
[64, 1140] loss: 0.043
[64, 1170] loss: 0.042
[64, 1200] loss: 0.047
[64, 1230] loss: 0.044
[64, 1260] loss: 0.047
[64, 1290] loss: 0.035
[64, 1320] loss: 0.037
[64, 1350] loss: 0.041
[64, 1380] loss: 0.042
[64, 1410] loss: 0.057
[64, 1440] loss: 0.045
[64, 1470] loss: 0.048
[64, 1500] loss: 0.030
[64, 1530] loss: 0.049
[64, 1560] loss: 0.061
[64, 1590] loss: 0.036
[64, 1620] loss: 0.047
[64, 1650] loss: 0.056
[64, 1680] loss: 0.036
[64, 1710] loss: 0.020
[64, 1740] loss: 0.059
[64, 1770] loss: 0.065
[64, 1800] loss: 0.035
[64, 1830] loss: 0.050
[64, 1860] loss: 0.036
[64, 1890] loss: 0.038
[64, 1920] loss: 0.055
[64, 1950] loss: 0.037
[64, 1980] loss: 0.061
[64, 2010] loss: 0.043
[64, 2040] loss: 0.046
[64, 2070] loss: 0.024
[64, 2100] loss: 0.029
[64, 2130] loss: 0.037
[64, 2160] loss: 0.045
[64, 2190] loss: 0.033
ovft_intermediate_models/epoch63_deep.pt saved.
```

```
evaluating on training set...
Accuracy on test images: 0.990322
epoch 63 accuracy on train set is: 0.9903217011995638
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732
```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	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

evaluating on test set...

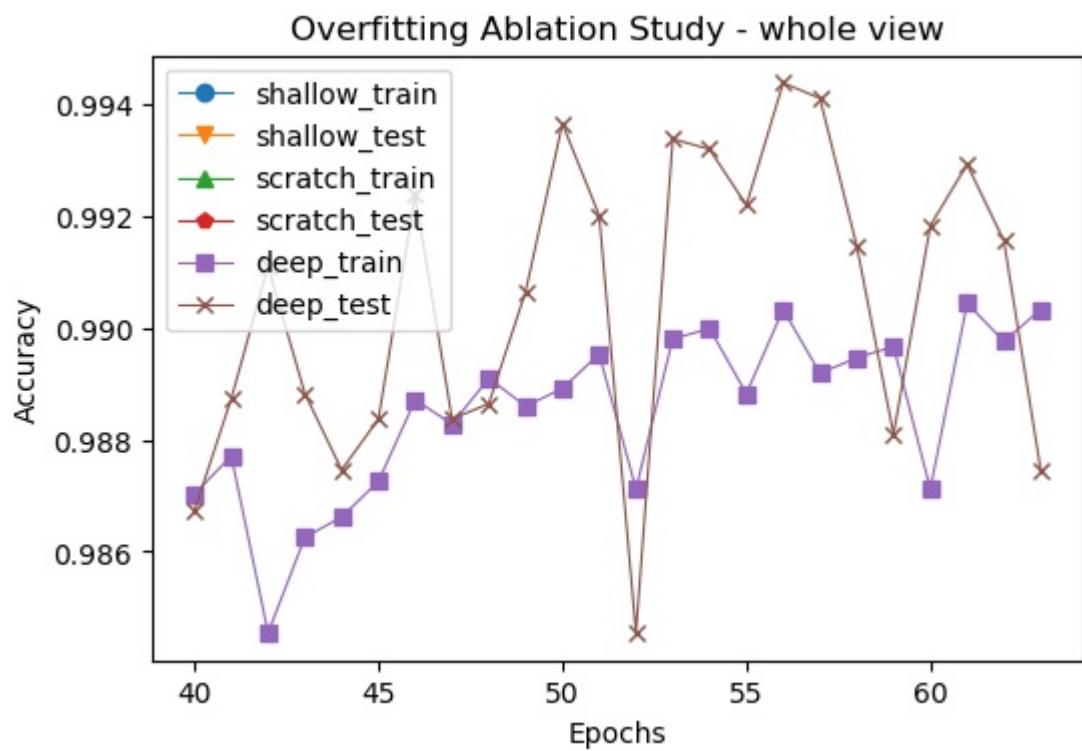
Accuracy on test images: 0.987459

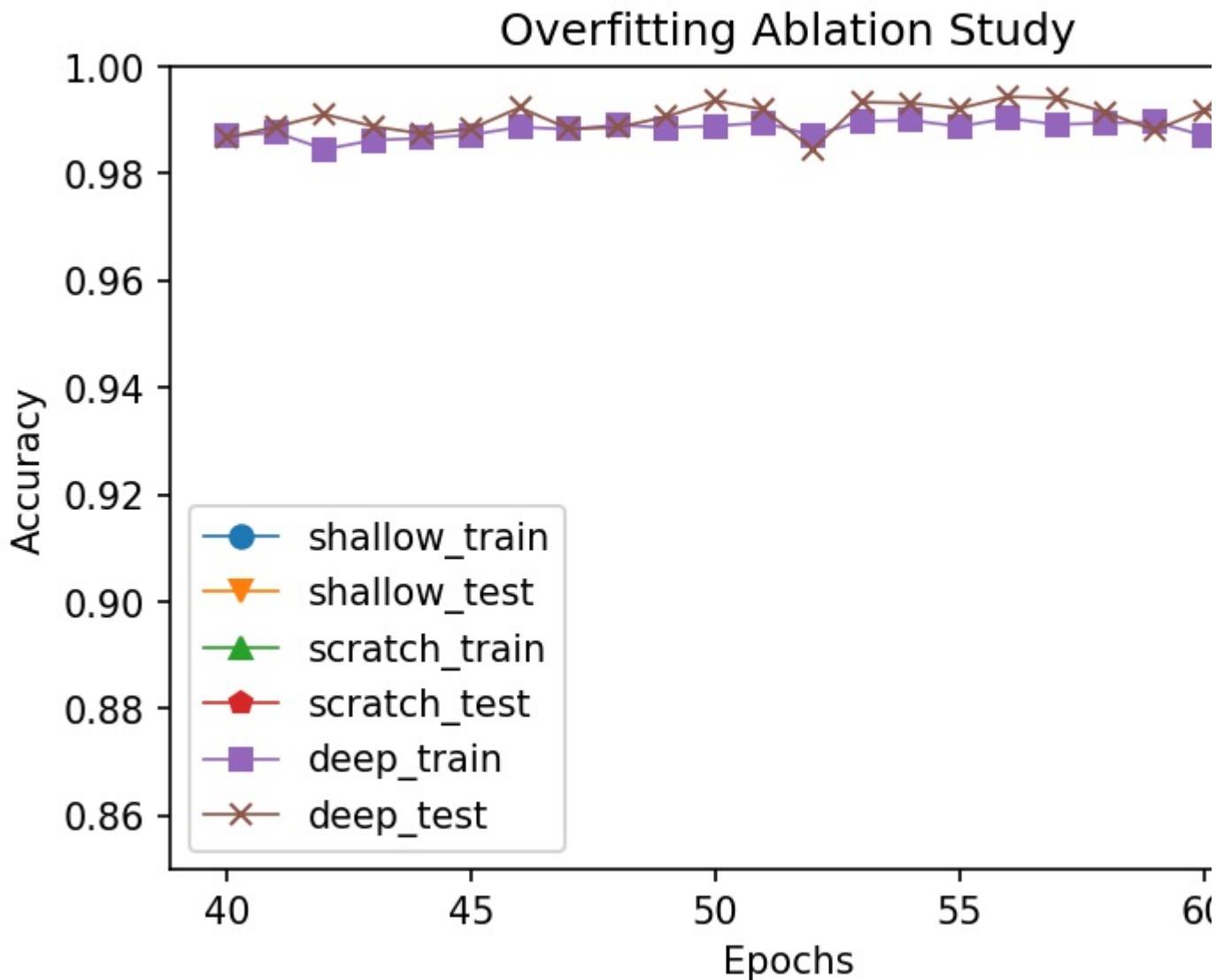
epoch 63 accuracy on test set is: 0.9874591057797165

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	NaN	NaN
65	NaN	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 63 takes 316 seconds.  
[65,    30] loss: 0.029  
[65,    60] loss: 0.049  
[65,    90] loss: 0.045  
[65,   120] loss: 0.057  
[65,   150] loss: 0.056  
[65,   180] loss: 0.060  
[65,   210] loss: 0.033  
[65,   240] loss: 0.054  
[65,   270] loss: 0.027  
[65,   300] loss: 0.039  
[65,   330] loss: 0.033  
[65,   360] loss: 0.027  
[65,   390] loss: 0.038  
[65,   420] loss: 0.058  
[65,   450] loss: 0.043  
[65,   480] loss: 0.032  
[65,   510] loss: 0.102  
[65,   540] loss: 0.030  
[65,   570] loss: 0.053  
[65,   600] loss: 0.050  
[65,   630] loss: 0.040  
[65,   660] loss: 0.048
```

```

[65, 690] loss: 0.046
[65, 720] loss: 0.052
[65, 750] loss: 0.043
[65, 780] loss: 0.043
[65, 810] loss: 0.052
[65, 840] loss: 0.024
[65, 870] loss: 0.018
[65, 900] loss: 0.098
[65, 930] loss: 0.046
[65, 960] loss: 0.067
[65, 990] loss: 0.045
[65, 1020] loss: 0.042
[65, 1050] loss: 0.042
[65, 1080] loss: 0.057
[65, 1110] loss: 0.055
[65, 1140] loss: 0.050
[65, 1170] loss: 0.040
[65, 1200] loss: 0.051
[65, 1230] loss: 0.065
[65, 1260] loss: 0.044
[65, 1290] loss: 0.036
[65, 1320] loss: 0.035
[65, 1350] loss: 0.059
[65, 1380] loss: 0.019
[65, 1410] loss: 0.036
[65, 1440] loss: 0.048
[65, 1470] loss: 0.030
[65, 1500] loss: 0.070
[65, 1530] loss: 0.039
[65, 1560] loss: 0.043
[65, 1590] loss: 0.041
[65, 1620] loss: 0.038
[65, 1650] loss: 0.093
[65, 1680] loss: 0.038
[65, 1710] loss: 0.033
[65, 1740] loss: 0.062
[65, 1770] loss: 0.055
[65, 1800] loss: 0.055
[65, 1830] loss: 0.036
[65, 1860] loss: 0.031
[65, 1890] loss: 0.008
[65, 1920] loss: 0.030
[65, 1950] loss: 0.063
[65, 1980] loss: 0.076
[65, 2010] loss: 0.029
[65, 2040] loss: 0.041
[65, 2070] loss: 0.029
[65, 2100] loss: 0.054
[65, 2130] loss: 0.028
[65, 2160] loss: 0.094
[65, 2190] loss: 0.084
ovft_intermediate_models/epoch64_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.988959
epoch 64 accuracy on train set is: 0.988958560523446
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	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

evaluating on test set...

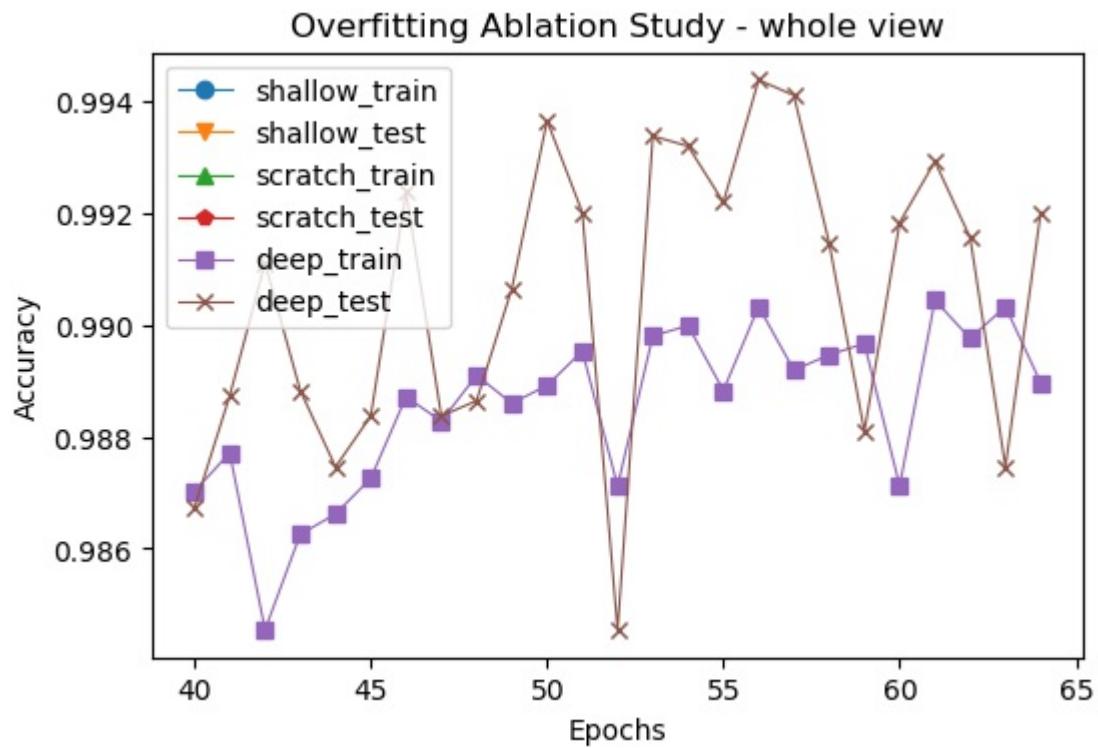
Accuracy on test images: 0.992003

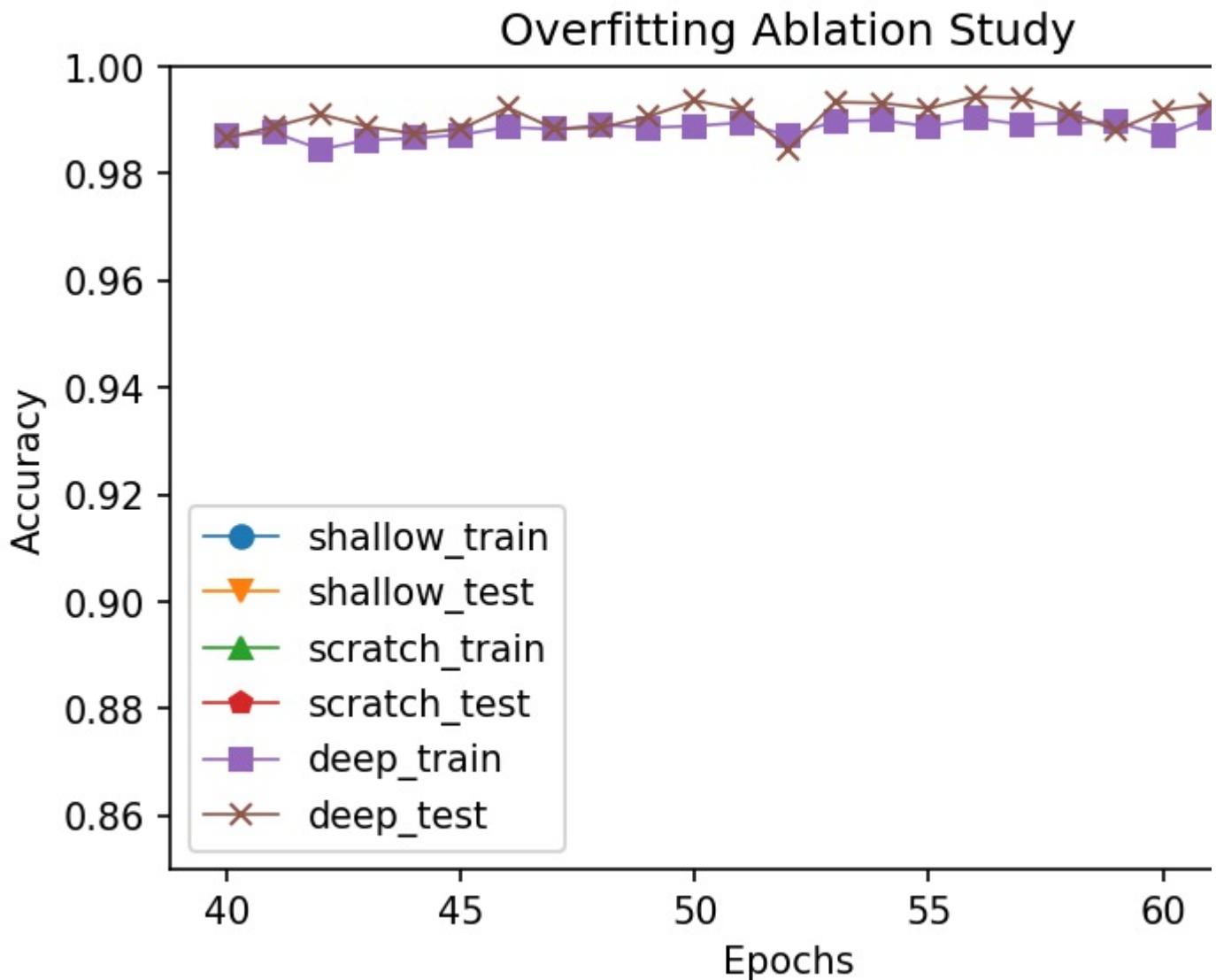
epoch 64 accuracy on test set is: 0.9920029080334424

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	NaN	NaN
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 64 takes 316 seconds.  
[66,    30] loss: 0.048  
[66,    60] loss: 0.039  
[66,    90] loss: 0.030  
[66,   120] loss: 0.039  
[66,   150] loss: 0.032  
[66,   180] loss: 0.025  
[66,   210] loss: 0.030  
[66,   240] loss: 0.064  
[66,   270] loss: 0.045  
[66,   300] loss: 0.046  
[66,   330] loss: 0.044  
[66,   360] loss: 0.049  
[66,   390] loss: 0.047  
[66,   420] loss: 0.049  
[66,   450] loss: 0.049  
[66,   480] loss: 0.035  
[66,   510] loss: 0.041  
[66,   540] loss: 0.028  
[66,   570] loss: 0.054  
[66,   600] loss: 0.056  
[66,   630] loss: 0.057  
[66,   660] loss: 0.038
```

```

[66, 690] loss: 0.062
[66, 720] loss: 0.037
[66, 750] loss: 0.042
[66, 780] loss: 0.052
[66, 810] loss: 0.045
[66, 840] loss: 0.036
[66, 870] loss: 0.033
[66, 900] loss: 0.047
[66, 930] loss: 0.039
[66, 960] loss: 0.036
[66, 990] loss: 0.030
[66, 1020] loss: 0.041
[66, 1050] loss: 0.036
[66, 1080] loss: 0.032
[66, 1110] loss: 0.065
[66, 1140] loss: 0.037
[66, 1170] loss: 0.062
[66, 1200] loss: 0.043
[66, 1230] loss: 0.036
[66, 1260] loss: 0.040
[66, 1290] loss: 0.037
[66, 1320] loss: 0.055
[66, 1350] loss: 0.044
[66, 1380] loss: 0.077
[66, 1410] loss: 0.057
[66, 1440] loss: 0.034
[66, 1470] loss: 0.067
[66, 1500] loss: 0.053
[66, 1530] loss: 0.042
[66, 1560] loss: 0.057
[66, 1590] loss: 0.058
[66, 1620] loss: 0.042
[66, 1650] loss: 0.042
[66, 1680] loss: 0.019
[66, 1710] loss: 0.088
[66, 1740] loss: 0.051
[66, 1770] loss: 0.032
[66, 1800] loss: 0.045
[66, 1830] loss: 0.038
[66, 1860] loss: 0.040
[66, 1890] loss: 0.038
[66, 1920] loss: 0.034
[66, 1950] loss: 0.041
[66, 1980] loss: 0.068
[66, 2010] loss: 0.046
[66, 2040] loss: 0.074
[66, 2070] loss: 0.075
[66, 2100] loss: 0.064
[66, 2130] loss: 0.044
[66, 2160] loss: 0.087
[66, 2190] loss: 0.025
ovft_intermediate_models/epoch65_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.990004
epoch 65 accuracy on train set is: 0.990003635041803
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	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

evaluating on test set...

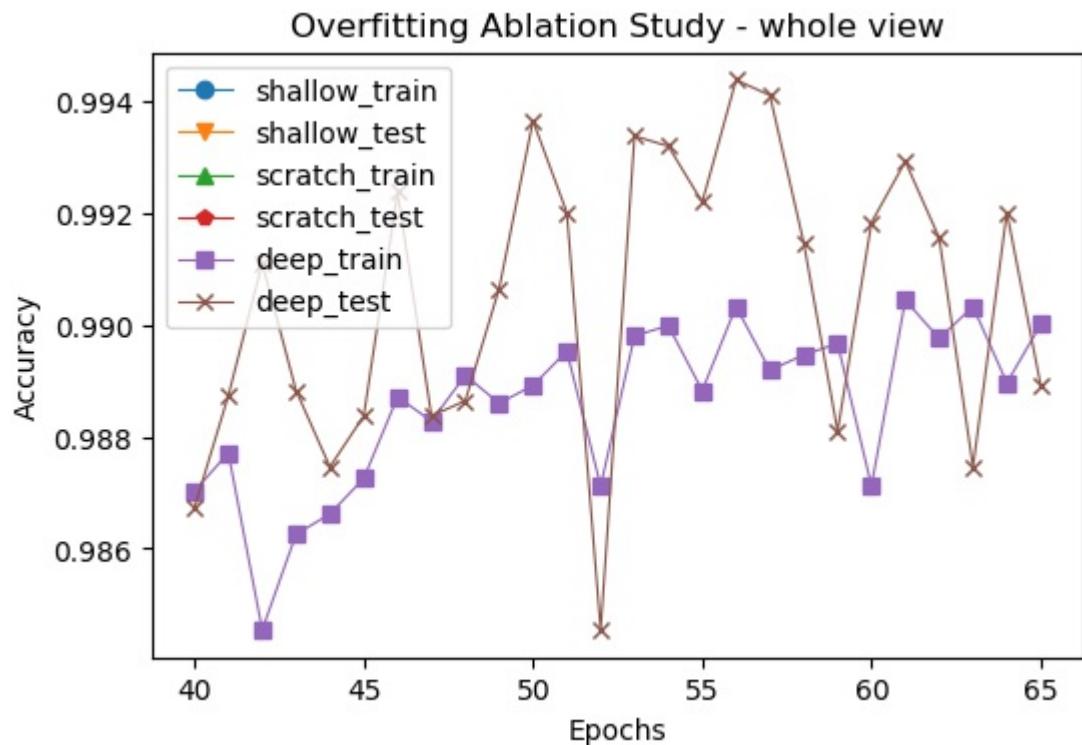
Accuracy on test images: 0.988913

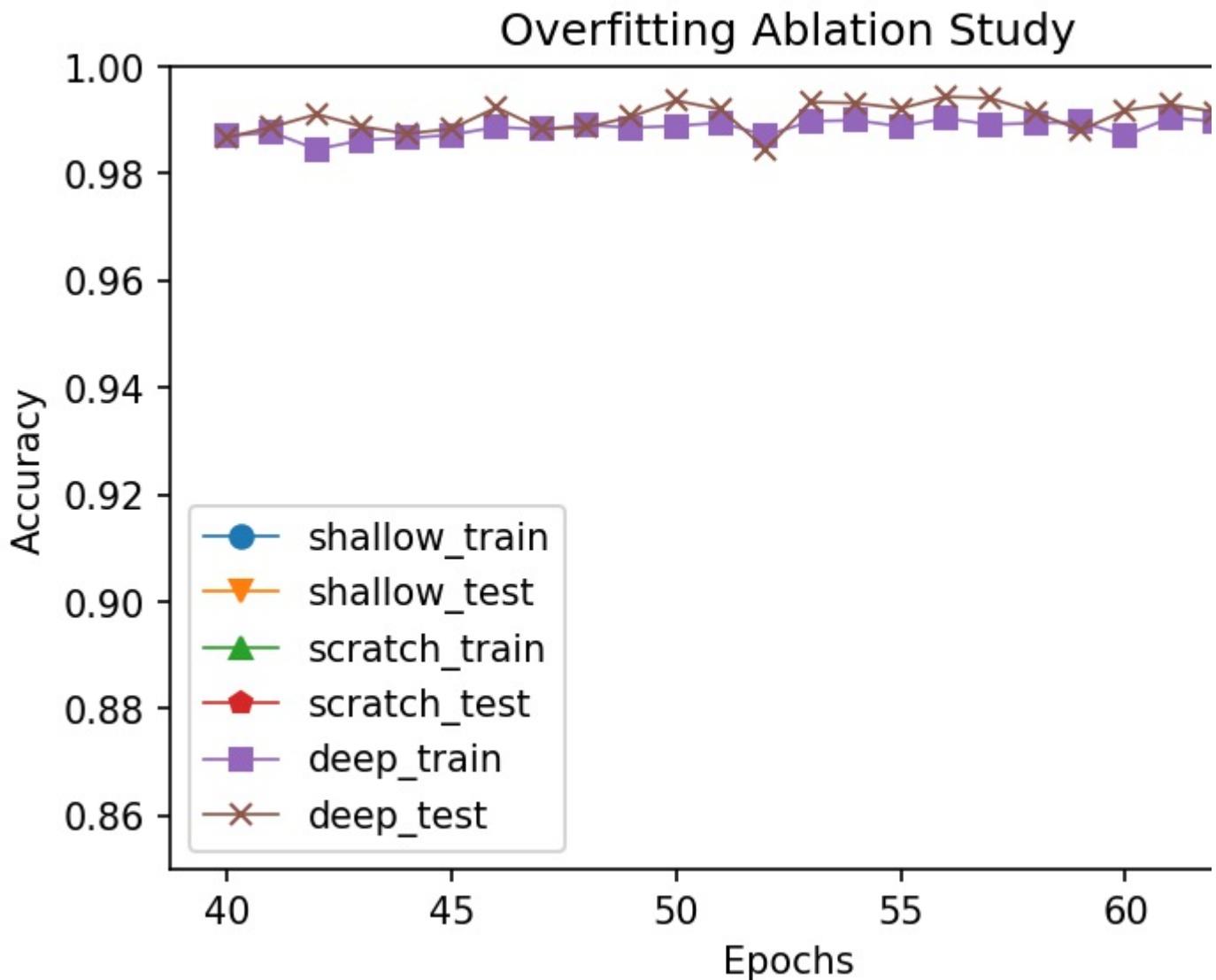
epoch 65 accuracy on test set is: 0.9889131225009088

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	NaN	NaN
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 65 takes 316 seconds.  
[67,    30] loss: 0.058  
[67,    60] loss: 0.050  
[67,    90] loss: 0.033  
[67,   120] loss: 0.036  
[67,   150] loss: 0.058  
[67,   180] loss: 0.064  
[67,   210] loss: 0.042  
[67,   240] loss: 0.031  
[67,   270] loss: 0.055  
[67,   300] loss: 0.045  
[67,   330] loss: 0.031  
[67,   360] loss: 0.047  
[67,   390] loss: 0.044  
[67,   420] loss: 0.044  
[67,   450] loss: 0.039  
[67,   480] loss: 0.058  
[67,   510] loss: 0.075  
[67,   540] loss: 0.052  
[67,   570] loss: 0.040  
[67,   600] loss: 0.067  
[67,   630] loss: 0.076  
[67,   660] loss: 0.051
```

```

[67, 690] loss: 0.038
[67, 720] loss: 0.069
[67, 750] loss: 0.047
[67, 780] loss: 0.050
[67, 810] loss: 0.044
[67, 840] loss: 0.030
[67, 870] loss: 0.046
[67, 900] loss: 0.075
[67, 930] loss: 0.044
[67, 960] loss: 0.025
[67, 990] loss: 0.026
[67, 1020] loss: 0.022
[67, 1050] loss: 0.040
[67, 1080] loss: 0.071
[67, 1110] loss: 0.030
[67, 1140] loss: 0.037
[67, 1170] loss: 0.033
[67, 1200] loss: 0.063
[67, 1230] loss: 0.038
[67, 1260] loss: 0.054
[67, 1290] loss: 0.050
[67, 1320] loss: 0.027
[67, 1350] loss: 0.040
[67, 1380] loss: 0.051
[67, 1410] loss: 0.015
[67, 1440] loss: 0.039
[67, 1470] loss: 0.043
[67, 1500] loss: 0.026
[67, 1530] loss: 0.044
[67, 1560] loss: 0.065
[67, 1590] loss: 0.033
[67, 1620] loss: 0.048
[67, 1650] loss: 0.024
[67, 1680] loss: 0.042
[67, 1710] loss: 0.038
[67, 1740] loss: 0.049
[67, 1770] loss: 0.039
[67, 1800] loss: 0.052
[67, 1830] loss: 0.052
[67, 1860] loss: 0.044
[67, 1890] loss: 0.029
[67, 1920] loss: 0.033
[67, 1950] loss: 0.041
[67, 1980] loss: 0.046
[67, 2010] loss: 0.057
[67, 2040] loss: 0.060
[67, 2070] loss: 0.028
[67, 2100] loss: 0.046
[67, 2130] loss: 0.044
[67, 2160] loss: 0.033
[67, 2190] loss: 0.056
ovft_intermediate_models/epoch66_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.987368
epoch 66 accuracy on train set is: 0.987368229734642
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	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

evaluating on test set...

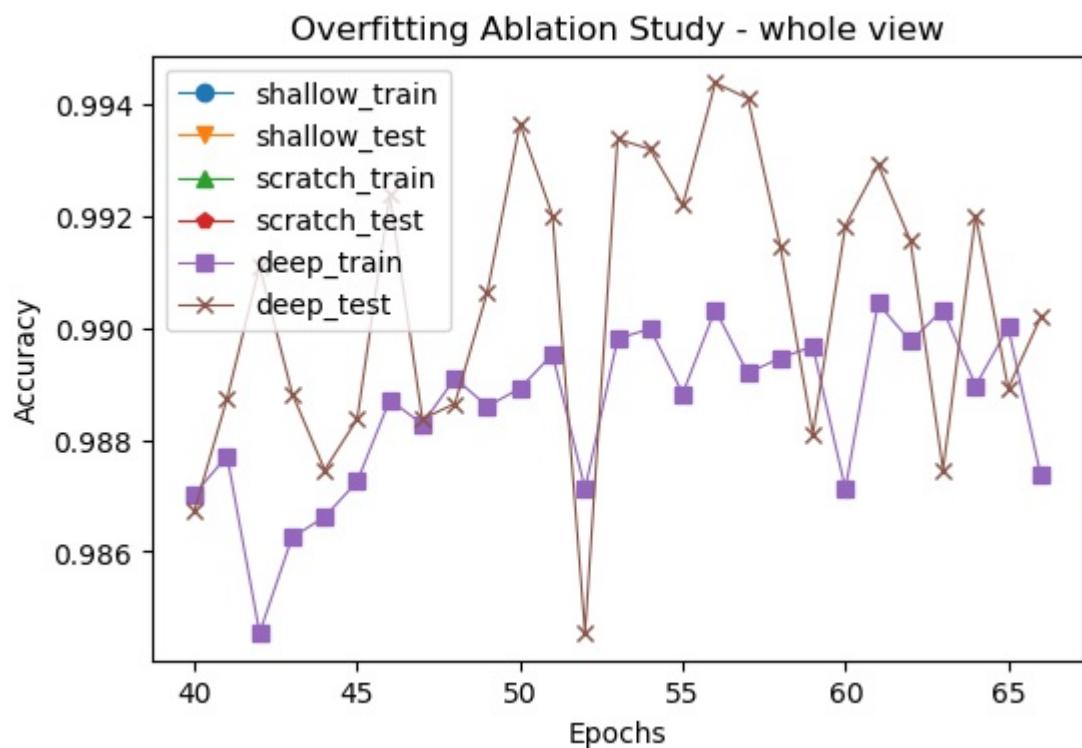
Accuracy on test images: 0.990185

epoch 66 accuracy on test set is: 0.990185387131952

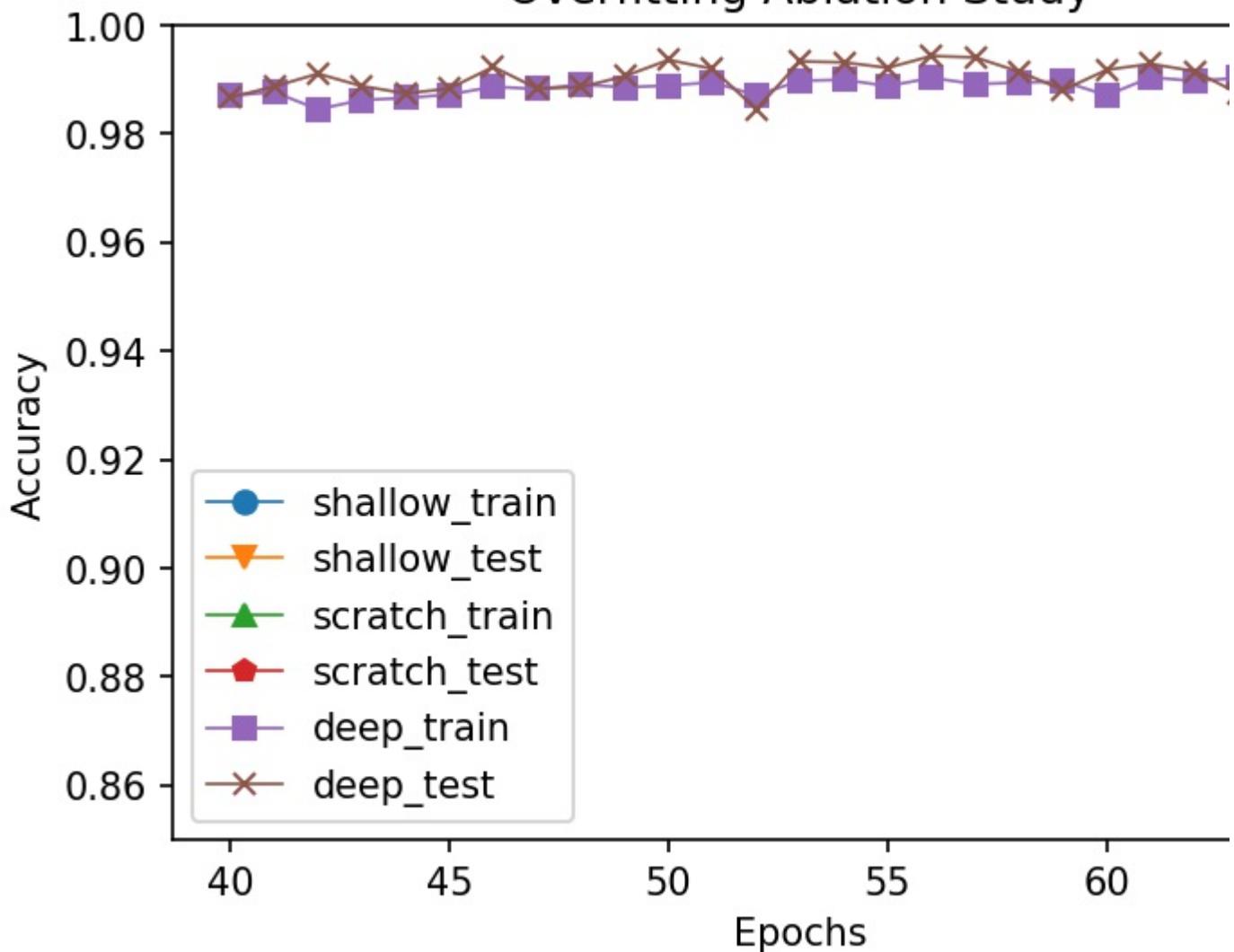
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	NaN	NaN
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN



Overfitting Ablation Study



```
epoch 66 takes 316 seconds.  
[68,    30] loss: 0.033  
[68,    60] loss: 0.023  
[68,   90] loss: 0.056  
[68,  120] loss: 0.031  
[68,  150] loss: 0.043  
[68,  180] loss: 0.040  
[68,  210] loss: 0.034  
[68,  240] loss: 0.042  
[68,  270] loss: 0.032  
[68,  300] loss: 0.036  
[68,  330] loss: 0.032  
[68,  360] loss: 0.066  
[68,  390] loss: 0.051  
[68,  420] loss: 0.059  
[68,  450] loss: 0.046  
[68,  480] loss: 0.054  
[68,  510] loss: 0.042  
[68,  540] loss: 0.037  
[68,  570] loss: 0.040  
[68,  600] loss: 0.046  
[68,  630] loss: 0.017  
[68,  660] loss: 0.083
```

```

[68, 690] loss: 0.038
[68, 720] loss: 0.042
[68, 750] loss: 0.074
[68, 780] loss: 0.042
[68, 810] loss: 0.018
[68, 840] loss: 0.053
[68, 870] loss: 0.052
[68, 900] loss: 0.027
[68, 930] loss: 0.047
[68, 960] loss: 0.044
[68, 990] loss: 0.030
[68, 1020] loss: 0.036
[68, 1050] loss: 0.060
[68, 1080] loss: 0.045
[68, 1110] loss: 0.041
[68, 1140] loss: 0.036
[68, 1170] loss: 0.070
[68, 1200] loss: 0.072
[68, 1230] loss: 0.056
[68, 1260] loss: 0.034
[68, 1290] loss: 0.056
[68, 1320] loss: 0.037
[68, 1350] loss: 0.033
[68, 1380] loss: 0.026
[68, 1410] loss: 0.029
[68, 1440] loss: 0.039
[68, 1470] loss: 0.042
[68, 1500] loss: 0.042
[68, 1530] loss: 0.025
[68, 1560] loss: 0.029
[68, 1590] loss: 0.030
[68, 1620] loss: 0.044
[68, 1650] loss: 0.056
[68, 1680] loss: 0.081
[68, 1710] loss: 0.066
[68, 1740] loss: 0.110
[68, 1770] loss: 0.074
[68, 1800] loss: 0.044
[68, 1830] loss: 0.034
[68, 1860] loss: 0.061
[68, 1890] loss: 0.036
[68, 1920] loss: 0.038
[68, 1950] loss: 0.063
[68, 1980] loss: 0.043
[68, 2010] loss: 0.032
[68, 2040] loss: 0.017
[68, 2070] loss: 0.034
[68, 2100] loss: 0.055
[68, 2130] loss: 0.060
[68, 2160] loss: 0.070
[68, 2190] loss: 0.085
ovft_intermediate_models/epoch67_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989867
epoch 67 accuracy on train set is: 0.9898673209741912
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	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

evaluating on test set...

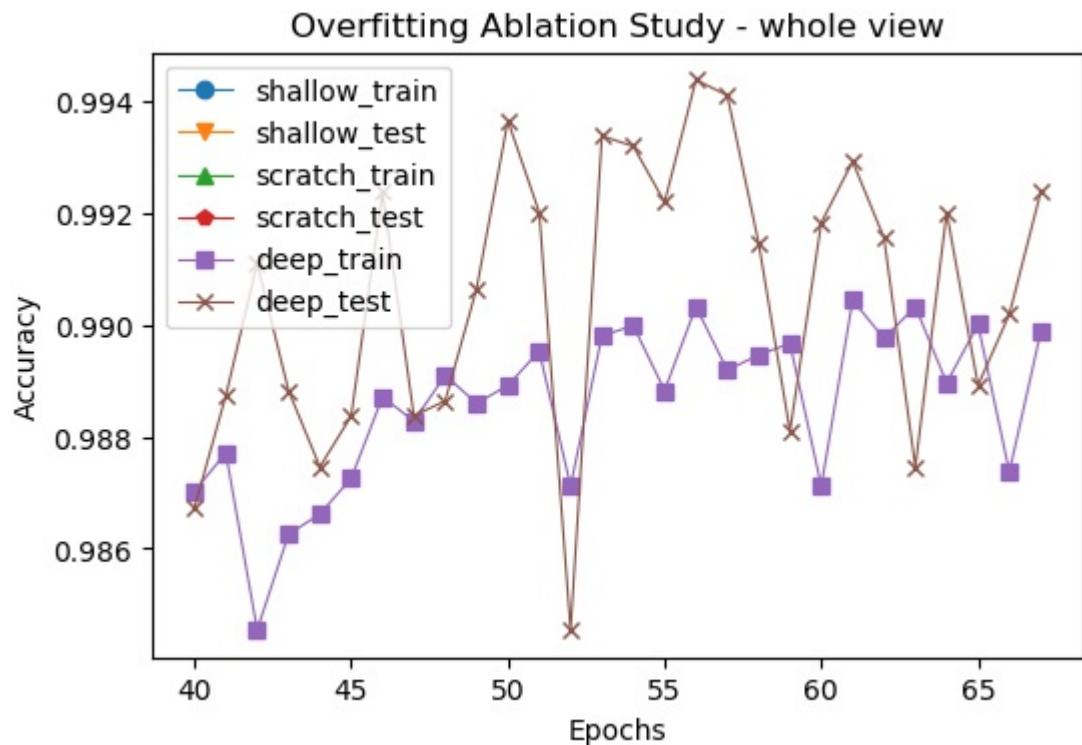
Accuracy on test images: 0.992366

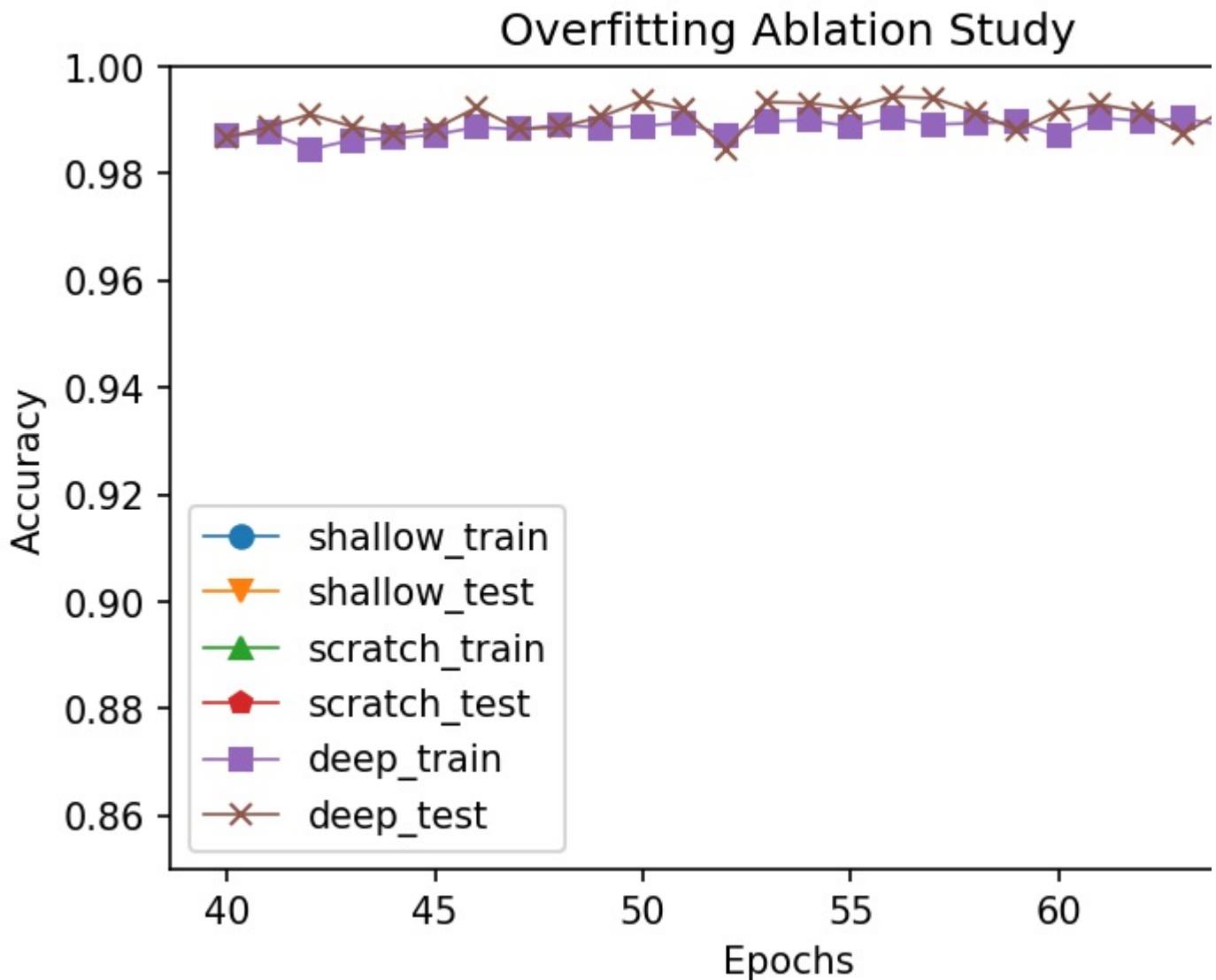
epoch 67 accuracy on test set is: 0.9923664122137404

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	NaN	NaN
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```

epoch 67 takes 316 seconds.
[69,    30] loss: 0.061
[69,    60] loss: 0.046
[69,   90] loss: 0.032
[69,  120] loss: 0.028
[69,  150] loss: 0.037
[69,  180] loss: 0.055
[69,  210] loss: 0.037
[69,  240] loss: 0.069
[69,  270] loss: 0.042
[69,  300] loss: 0.028
[69,  330] loss: 0.026
[69,  360] loss: 0.051
[69,  390] loss: 0.038
[69,  420] loss: 0.034
[69,  450] loss: 0.040
[69,  480] loss: 0.024
[69,  510] loss: 0.043
[69,  540] loss: 0.038
[69,  570] loss: 0.066
[69,  600] loss: 0.057
[69,  630] loss: 0.038
[69,  660] loss: 0.024

```

```

[69, 690] loss: 0.068
[69, 720] loss: 0.039
[69, 750] loss: 0.041
[69, 780] loss: 0.045
[69, 810] loss: 0.051
[69, 840] loss: 0.044
[69, 870] loss: 0.031
[69, 900] loss: 0.055
[69, 930] loss: 0.062
[69, 960] loss: 0.019
[69, 990] loss: 0.040
[69, 1020] loss: 0.044
[69, 1050] loss: 0.025
[69, 1080] loss: 0.079
[69, 1110] loss: 0.035
[69, 1140] loss: 0.049
[69, 1170] loss: 0.042
[69, 1200] loss: 0.052
[69, 1230] loss: 0.051
[69, 1260] loss: 0.048
[69, 1290] loss: 0.070
[69, 1320] loss: 0.024
[69, 1350] loss: 0.050
[69, 1380] loss: 0.052
[69, 1410] loss: 0.051
[69, 1440] loss: 0.051
[69, 1470] loss: 0.078
[69, 1500] loss: 0.030
[69, 1530] loss: 0.026
[69, 1560] loss: 0.039
[69, 1590] loss: 0.029
[69, 1620] loss: 0.040
[69, 1650] loss: 0.034
[69, 1680] loss: 0.041
[69, 1710] loss: 0.057
[69, 1740] loss: 0.056
[69, 1770] loss: 0.033
[69, 1800] loss: 0.064
[69, 1830] loss: 0.052
[69, 1860] loss: 0.068
[69, 1890] loss: 0.064
[69, 1920] loss: 0.015
[69, 1950] loss: 0.042
[69, 1980] loss: 0.041
[69, 2010] loss: 0.048
[69, 2040] loss: 0.072
[69, 2070] loss: 0.039
[69, 2100] loss: 0.050
[69, 2130] loss: 0.039
[69, 2160] loss: 0.027
[69, 2190] loss: 0.022
ovft_intermediate_models/epoch68_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.992185
epoch 68 accuracy on train set is: 0.9921846601235914
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	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

evaluating on test set...

Accuracy on test images: 0.993184

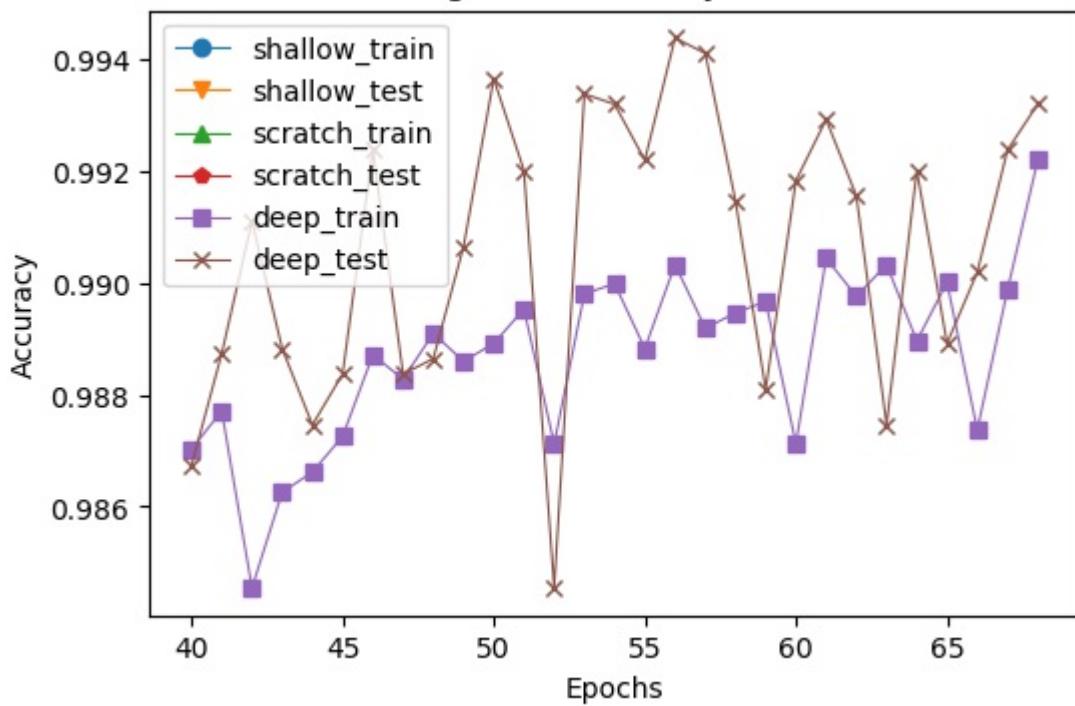
epoch 68 accuracy on test set is: 0.9931842966194111

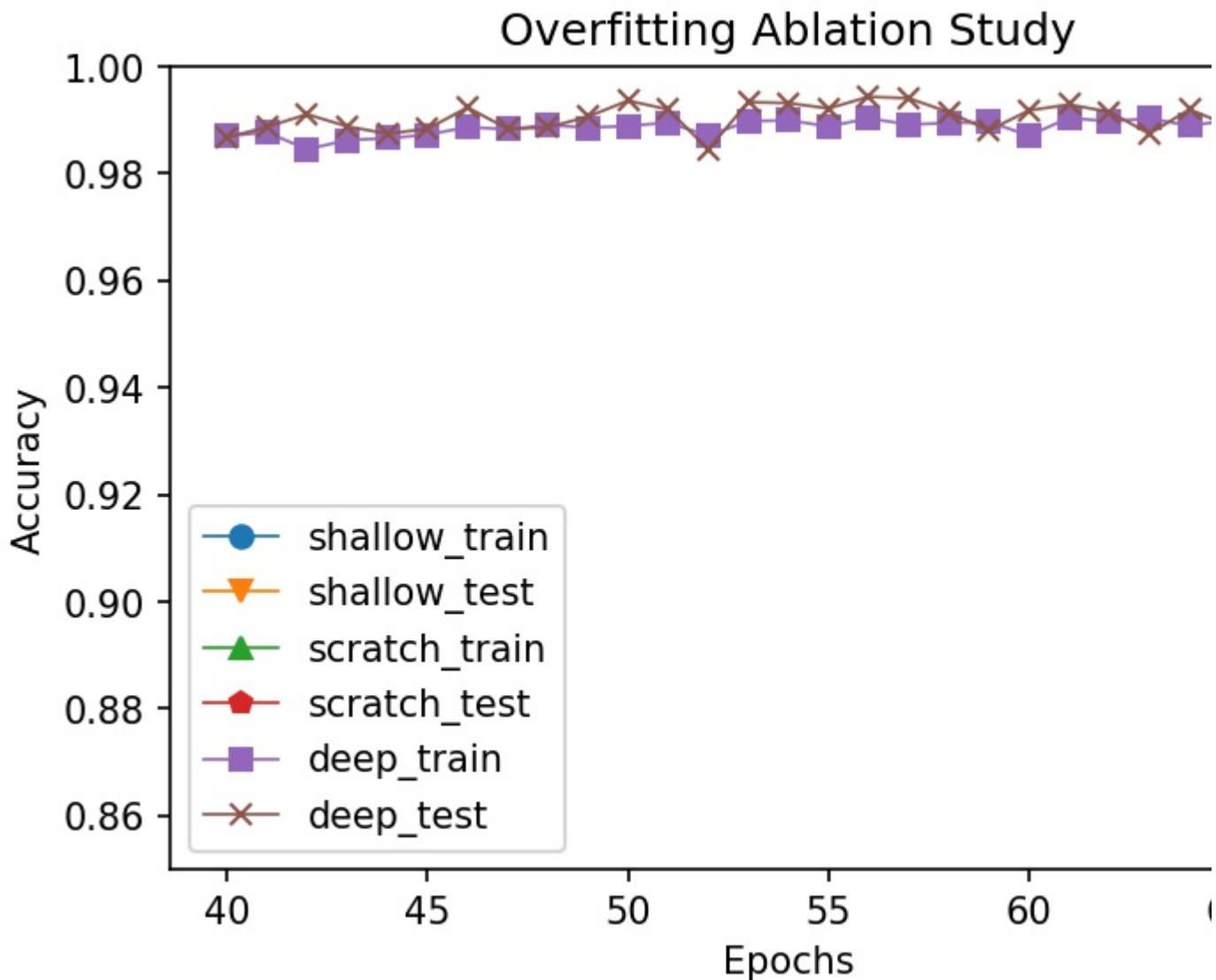
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	NaN	NaN
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view





```
epoch 68 takes 317 seconds.  
[70,    30] loss: 0.028  
[70,    60] loss: 0.040  
[70,   90] loss: 0.032  
[70,  120] loss: 0.057  
[70,  150] loss: 0.026  
[70,  180] loss: 0.043  
[70,  210] loss: 0.028  
[70,  240] loss: 0.040  
[70,  270] loss: 0.047  
[70,  300] loss: 0.026  
[70,  330] loss: 0.051  
[70,  360] loss: 0.039  
[70,  390] loss: 0.065  
[70,  420] loss: 0.038  
[70,  450] loss: 0.043  
[70,  480] loss: 0.077  
[70,  510] loss: 0.045  
[70,  540] loss: 0.039  
[70,  570] loss: 0.041  
[70,  600] loss: 0.049  
[70,  630] loss: 0.042  
[70,  660] loss: 0.031
```

```

[70, 690] loss: 0.034
[70, 720] loss: 0.025
[70, 750] loss: 0.060
[70, 780] loss: 0.042
[70, 810] loss: 0.090
[70, 840] loss: 0.053
[70, 870] loss: 0.028
[70, 900] loss: 0.049
[70, 930] loss: 0.072
[70, 960] loss: 0.046
[70, 990] loss: 0.045
[70, 1020] loss: 0.038
[70, 1050] loss: 0.049
[70, 1080] loss: 0.032
[70, 1110] loss: 0.055
[70, 1140] loss: 0.057
[70, 1170] loss: 0.039
[70, 1200] loss: 0.055
[70, 1230] loss: 0.021
[70, 1260] loss: 0.067
[70, 1290] loss: 0.022
[70, 1320] loss: 0.028
[70, 1350] loss: 0.029
[70, 1380] loss: 0.051
[70, 1410] loss: 0.053
[70, 1440] loss: 0.049
[70, 1470] loss: 0.035
[70, 1500] loss: 0.031
[70, 1530] loss: 0.035
[70, 1560] loss: 0.045
[70, 1590] loss: 0.046
[70, 1620] loss: 0.025
[70, 1650] loss: 0.052
[70, 1680] loss: 0.065
[70, 1710] loss: 0.021
[70, 1740] loss: 0.043
[70, 1770] loss: 0.079
[70, 1800] loss: 0.026
[70, 1830] loss: 0.063
[70, 1860] loss: 0.037
[70, 1890] loss: 0.054
[70, 1920] loss: 0.030
[70, 1950] loss: 0.063
[70, 1980] loss: 0.032
[70, 2010] loss: 0.041
[70, 2040] loss: 0.056
[70, 2070] loss: 0.048
[70, 2100] loss: 0.022
[70, 2130] loss: 0.027
[70, 2160] loss: 0.039
[70, 2190] loss: 0.026
ovft_intermediate_models/epoch69_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.991367
epoch 69 accuracy on train set is: 0.9913667757179208
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	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

evaluating on test set...

Accuracy on test images: 0.993003

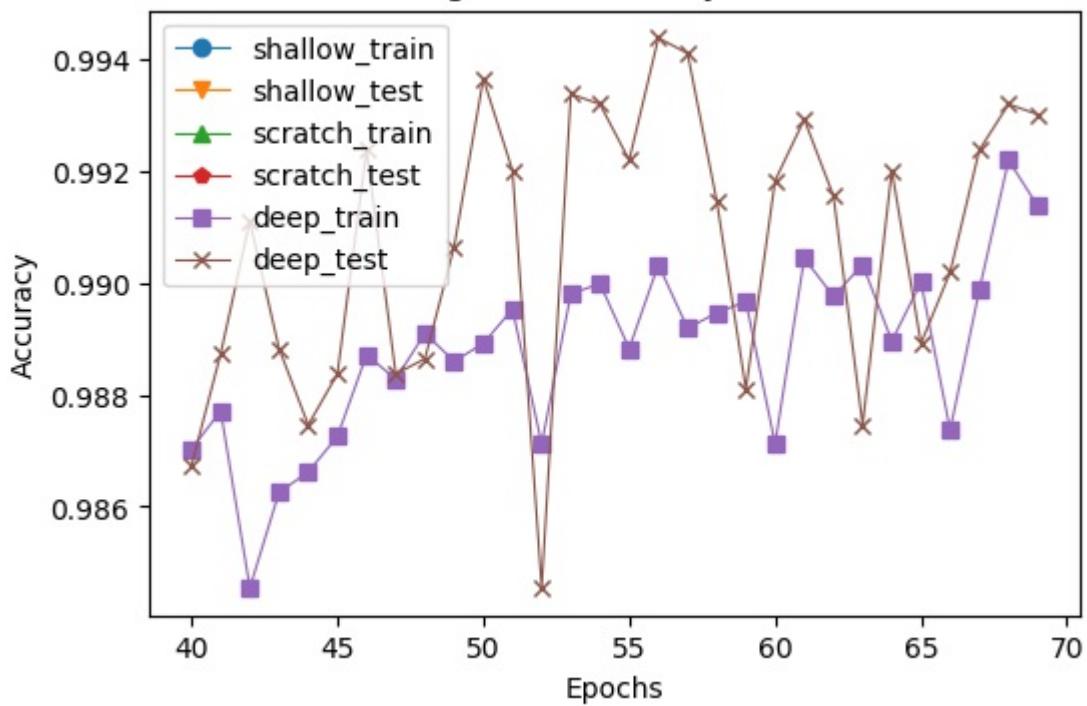
epoch 69 accuracy on test set is: 0.9930025445292621

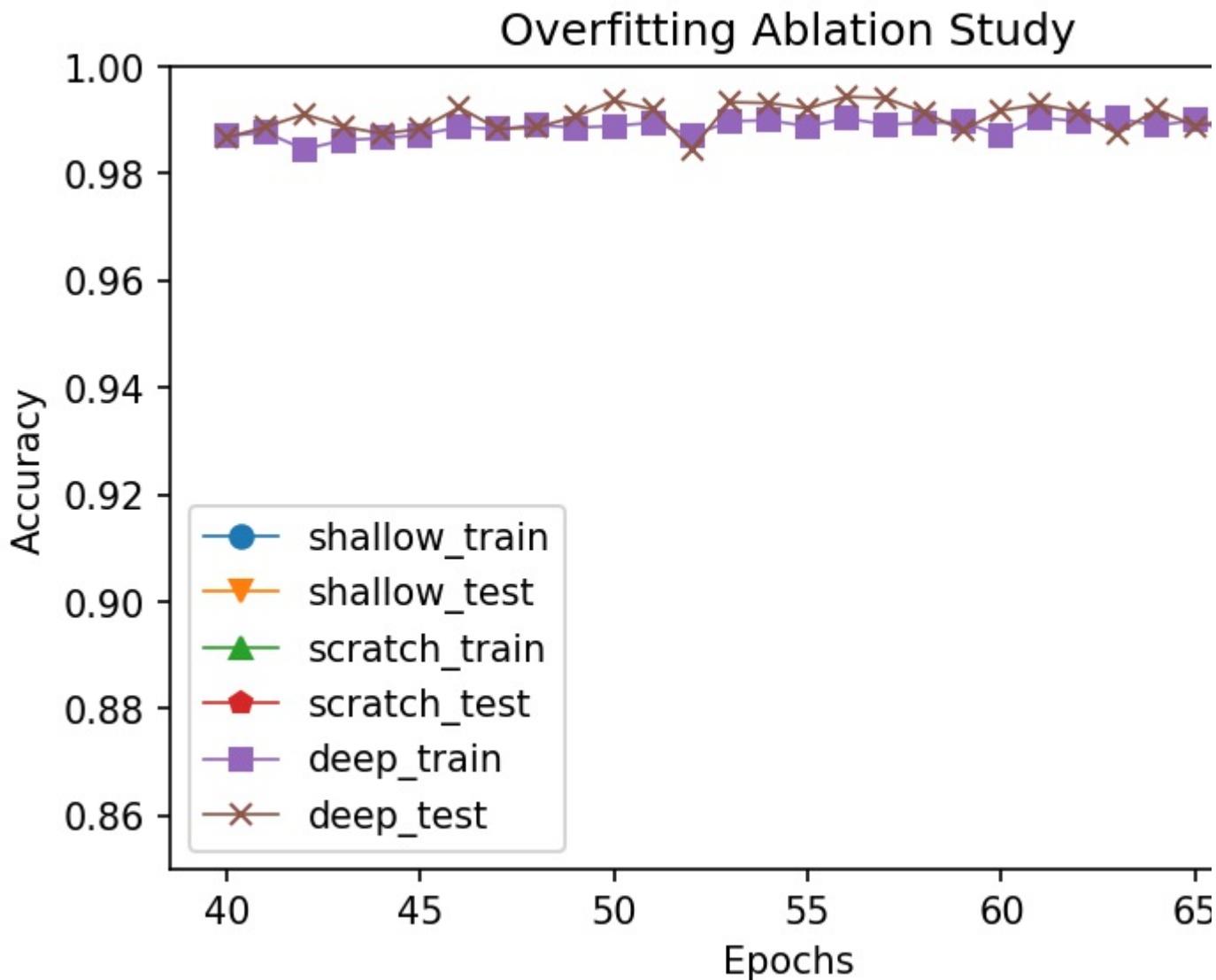
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	NaN	NaN
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view





```
epoch 69 takes 317 seconds.  
[71,    30] loss: 0.066  
[71,    60] loss: 0.045  
[71,    90] loss: 0.055  
[71,   120] loss: 0.058  
[71,   150] loss: 0.077  
[71,   180] loss: 0.050  
[71,   210] loss: 0.044  
[71,   240] loss: 0.027  
[71,   270] loss: 0.043  
[71,   300] loss: 0.044  
[71,   330] loss: 0.039  
[71,   360] loss: 0.058  
[71,   390] loss: 0.041  
[71,   420] loss: 0.055  
[71,   450] loss: 0.062  
[71,   480] loss: 0.062  
[71,   510] loss: 0.049  
[71,   540] loss: 0.056  
[71,   570] loss: 0.024  
[71,   600] loss: 0.036  
[71,   630] loss: 0.050  
[71,   660] loss: 0.061
```

```

[71, 690] loss: 0.036
[71, 720] loss: 0.040
[71, 750] loss: 0.033
[71, 780] loss: 0.047
[71, 810] loss: 0.037
[71, 840] loss: 0.043
[71, 870] loss: 0.049
[71, 900] loss: 0.063
[71, 930] loss: 0.068
[71, 960] loss: 0.056
[71, 990] loss: 0.033
[71, 1020] loss: 0.068
[71, 1050] loss: 0.022
[71, 1080] loss: 0.032
[71, 1110] loss: 0.047
[71, 1140] loss: 0.042
[71, 1170] loss: 0.047
[71, 1200] loss: 0.042
[71, 1230] loss: 0.030
[71, 1260] loss: 0.048
[71, 1290] loss: 0.038
[71, 1320] loss: 0.056
[71, 1350] loss: 0.038
[71, 1380] loss: 0.053
[71, 1410] loss: 0.015
[71, 1440] loss: 0.017
[71, 1470] loss: 0.065
[71, 1500] loss: 0.048
[71, 1530] loss: 0.035
[71, 1560] loss: 0.024
[71, 1590] loss: 0.042
[71, 1620] loss: 0.038
[71, 1650] loss: 0.035
[71, 1680] loss: 0.067
[71, 1710] loss: 0.060
[71, 1740] loss: 0.072
[71, 1770] loss: 0.024
[71, 1800] loss: 0.043
[71, 1830] loss: 0.059
[71, 1860] loss: 0.030
[71, 1890] loss: 0.045
[71, 1920] loss: 0.043
[71, 1950] loss: 0.049
[71, 1980] loss: 0.053
[71, 2010] loss: 0.041
[71, 2040] loss: 0.054
[71, 2070] loss: 0.044
[71, 2100] loss: 0.029
[71, 2130] loss: 0.026
[71, 2160] loss: 0.043
[71, 2190] loss: 0.025
ovft_intermediate_models/epoch70_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989935
epoch 70 accuracy on train set is: 0.9899354780079971
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	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

evaluating on test set...

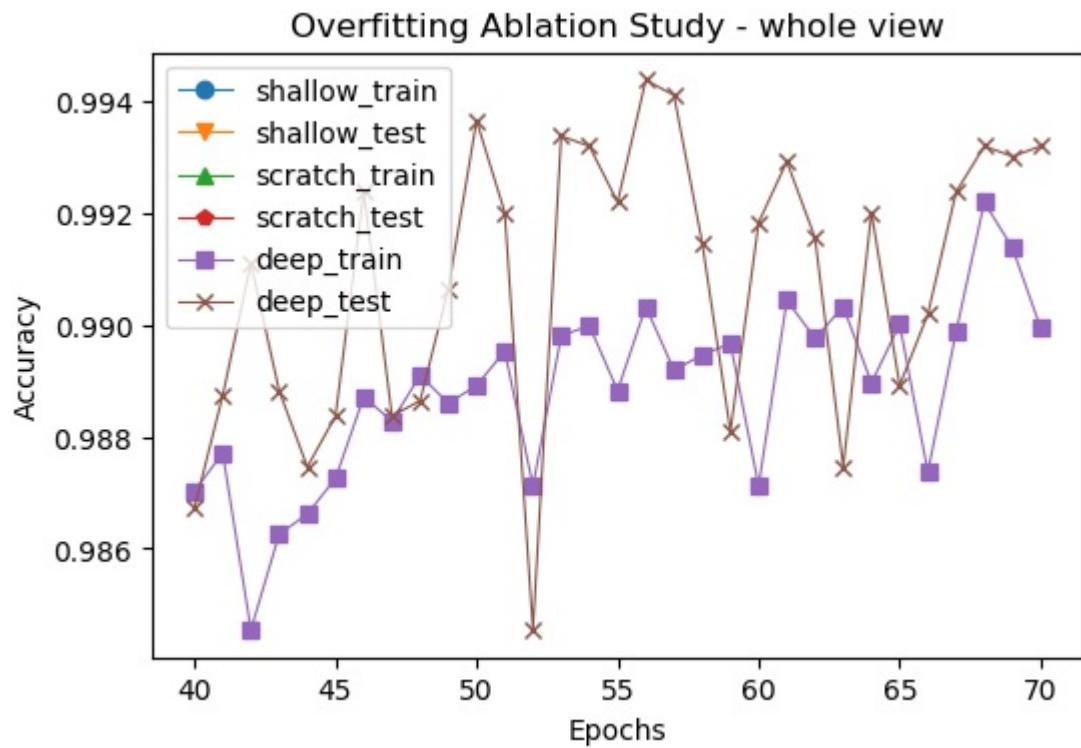
Accuracy on test images: 0.993184

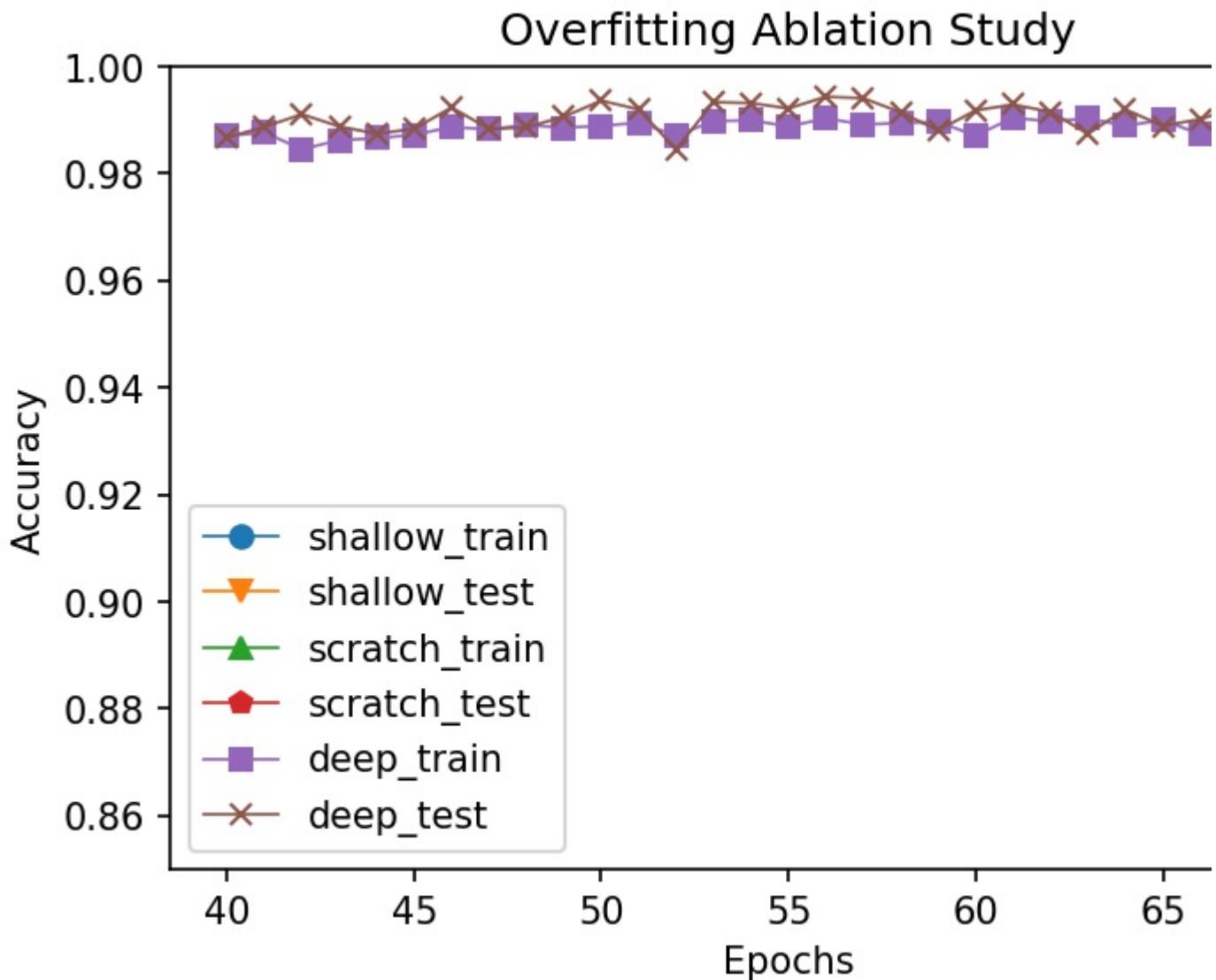
epoch 70 accuracy on test set is: 0.9931842966194111

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	NaN	NaN
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 70 takes 318 seconds.  
[72,    30] loss: 0.075  
[72,    60] loss: 0.037  
[72,    90] loss: 0.031  
[72,   120] loss: 0.049  
[72,   150] loss: 0.070  
[72,   180] loss: 0.049  
[72,   210] loss: 0.042  
[72,   240] loss: 0.023  
[72,   270] loss: 0.056  
[72,   300] loss: 0.038  
[72,   330] loss: 0.042  
[72,   360] loss: 0.036  
[72,   390] loss: 0.058  
[72,   420] loss: 0.054  
[72,   450] loss: 0.062  
[72,   480] loss: 0.030  
[72,   510] loss: 0.024  
[72,   540] loss: 0.041  
[72,   570] loss: 0.047  
[72,   600] loss: 0.059  
[72,   630] loss: 0.048  
[72,   660] loss: 0.031
```

```

[72, 690] loss: 0.042
[72, 720] loss: 0.061
[72, 750] loss: 0.029
[72, 780] loss: 0.037
[72, 810] loss: 0.026
[72, 840] loss: 0.041
[72, 870] loss: 0.039
[72, 900] loss: 0.023
[72, 930] loss: 0.069
[72, 960] loss: 0.049
[72, 990] loss: 0.046
[72, 1020] loss: 0.032
[72, 1050] loss: 0.052
[72, 1080] loss: 0.038
[72, 1110] loss: 0.025
[72, 1140] loss: 0.038
[72, 1170] loss: 0.057
[72, 1200] loss: 0.046
[72, 1230] loss: 0.052
[72, 1260] loss: 0.040
[72, 1290] loss: 0.061
[72, 1320] loss: 0.033
[72, 1350] loss: 0.066
[72, 1380] loss: 0.038
[72, 1410] loss: 0.031
[72, 1440] loss: 0.035
[72, 1470] loss: 0.071
[72, 1500] loss: 0.018
[72, 1530] loss: 0.038
[72, 1560] loss: 0.050
[72, 1590] loss: 0.054
[72, 1620] loss: 0.033
[72, 1650] loss: 0.030
[72, 1680] loss: 0.039
[72, 1710] loss: 0.033
[72, 1740] loss: 0.055
[72, 1770] loss: 0.040
[72, 1800] loss: 0.028
[72, 1830] loss: 0.037
[72, 1860] loss: 0.042
[72, 1890] loss: 0.041
[72, 1920] loss: 0.042
[72, 1950] loss: 0.053
[72, 1980] loss: 0.049
[72, 2010] loss: 0.037
[72, 2040] loss: 0.043
[72, 2070] loss: 0.038
[72, 2100] loss: 0.055
[72, 2130] loss: 0.019
[72, 2160] loss: 0.038
[72, 2190] loss: 0.024
ovft_intermediate_models/epoch71_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.990617
epoch 71 accuracy on train set is: 0.990617048346056
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	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

evaluating on test set...

Accuracy on test images: 0.992730

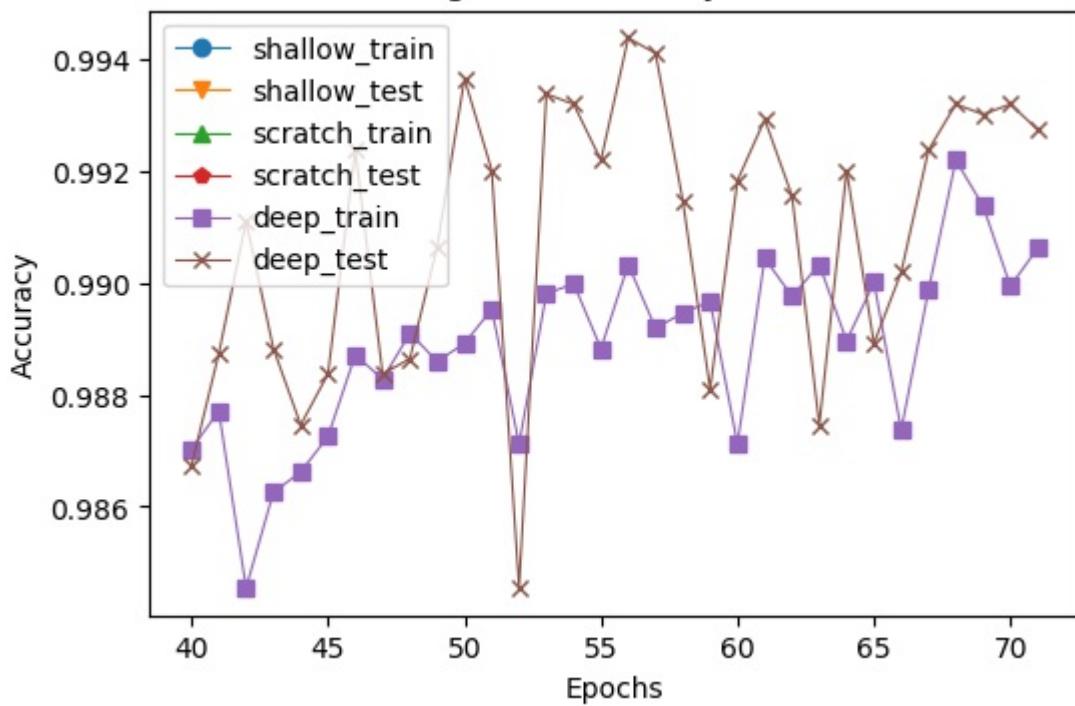
epoch 71 accuracy on test set is: 0.9927299163940385

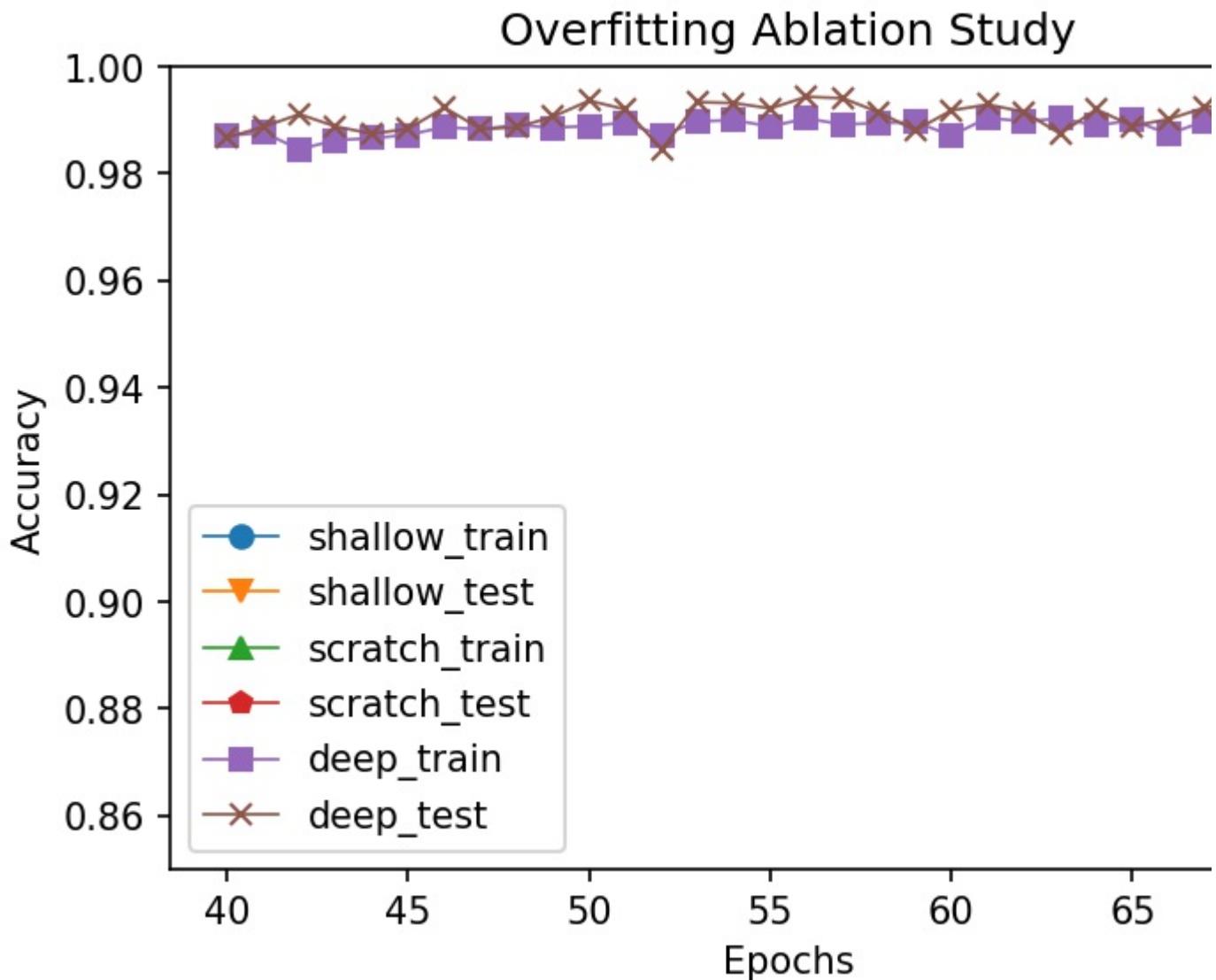
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	NaN	NaN
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view





```
epoch 71 takes 316 seconds.  
[73,    30] loss: 0.046  
[73,    60] loss: 0.076  
[73,    90] loss: 0.042  
[73,   120] loss: 0.076  
[73,   150] loss: 0.034  
[73,   180] loss: 0.036  
[73,   210] loss: 0.049  
[73,   240] loss: 0.032  
[73,   270] loss: 0.052  
[73,   300] loss: 0.030  
[73,   330] loss: 0.028  
[73,   360] loss: 0.036  
[73,   390] loss: 0.033  
[73,   420] loss: 0.061  
[73,   450] loss: 0.052  
[73,   480] loss: 0.027  
[73,   510] loss: 0.036  
[73,   540] loss: 0.044  
[73,   570] loss: 0.028  
[73,   600] loss: 0.084  
[73,   630] loss: 0.035  
[73,   660] loss: 0.053
```

```

[73, 690] loss: 0.062
[73, 720] loss: 0.051
[73, 750] loss: 0.045
[73, 780] loss: 0.039
[73, 810] loss: 0.056
[73, 840] loss: 0.025
[73, 870] loss: 0.034
[73, 900] loss: 0.056
[73, 930] loss: 0.042
[73, 960] loss: 0.034
[73, 990] loss: 0.026
[73, 1020] loss: 0.048
[73, 1050] loss: 0.043
[73, 1080] loss: 0.066
[73, 1110] loss: 0.026
[73, 1140] loss: 0.030
[73, 1170] loss: 0.027
[73, 1200] loss: 0.068
[73, 1230] loss: 0.036
[73, 1260] loss: 0.069
[73, 1290] loss: 0.059
[73, 1320] loss: 0.063
[73, 1350] loss: 0.078
[73, 1380] loss: 0.060
[73, 1410] loss: 0.073
[73, 1440] loss: 0.029
[73, 1470] loss: 0.026
[73, 1500] loss: 0.035
[73, 1530] loss: 0.039
[73, 1560] loss: 0.051
[73, 1590] loss: 0.047
[73, 1620] loss: 0.032
[73, 1650] loss: 0.076
[73, 1680] loss: 0.037
[73, 1710] loss: 0.065
[73, 1740] loss: 0.057
[73, 1770] loss: 0.042
[73, 1800] loss: 0.035
[73, 1830] loss: 0.049
[73, 1860] loss: 0.041
[73, 1890] loss: 0.016
[73, 1920] loss: 0.025
[73, 1950] loss: 0.044
[73, 1980] loss: 0.036
[73, 2010] loss: 0.031
[73, 2040] loss: 0.045
[73, 2070] loss: 0.042
[73, 2100] loss: 0.047
[73, 2130] loss: 0.057
[73, 2160] loss: 0.060
[73, 2190] loss: 0.027
ovft_intermediate_models/epoch72_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.991389
epoch 72 accuracy on train set is: 0.9913894947291894
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	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

evaluating on test set...

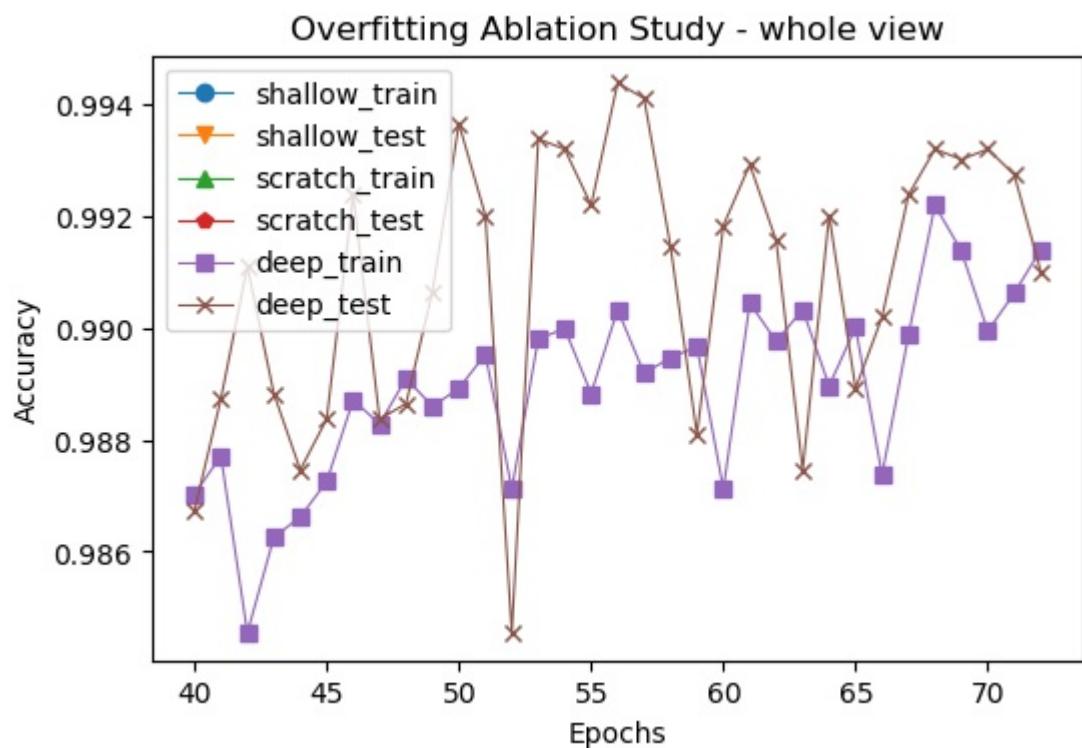
Accuracy on test images: 0.991003

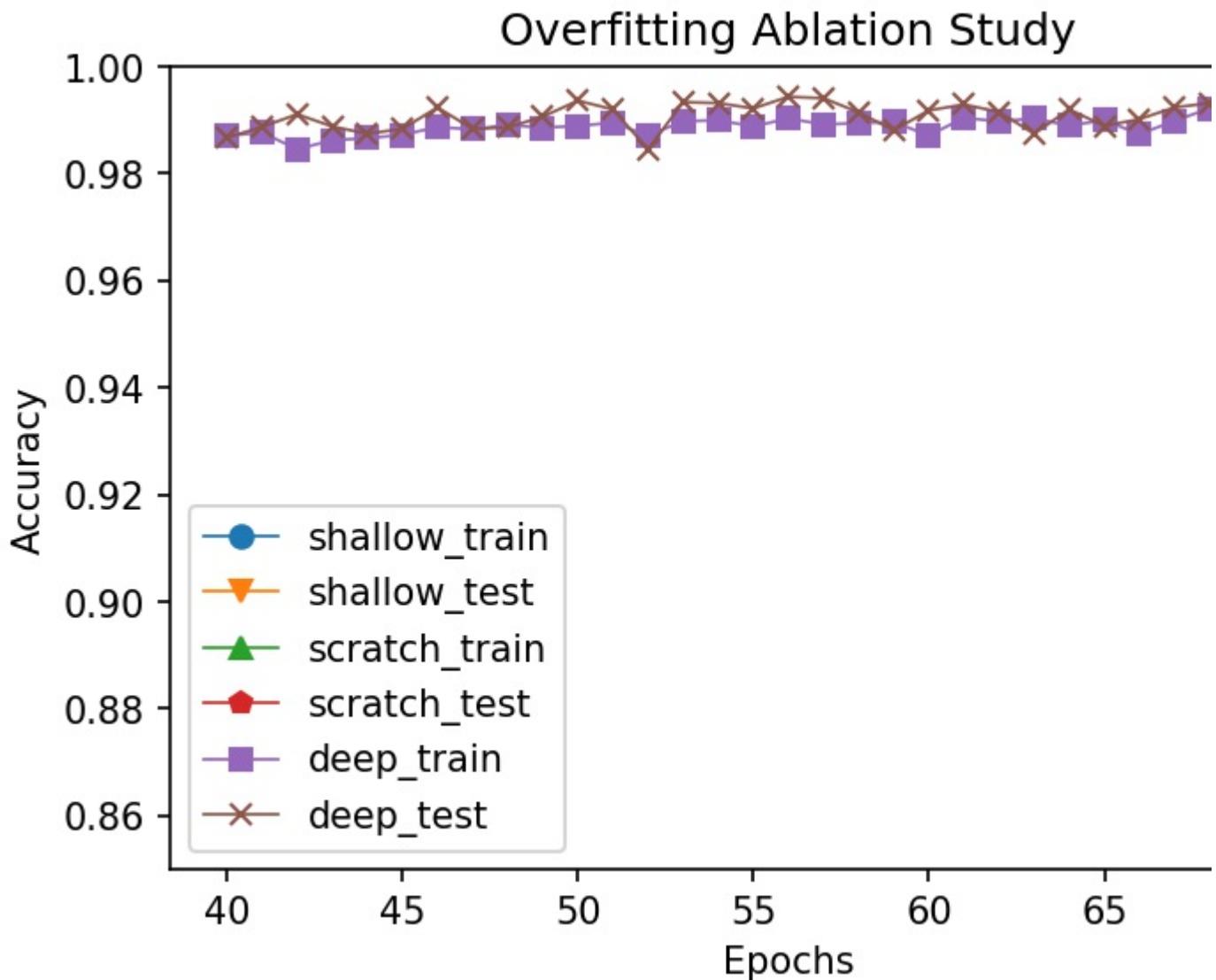
epoch 72 accuracy on test set is: 0.9910032715376227

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	NaN	NaN
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 72 takes 315 seconds.  
[74,    30] loss: 0.037  
[74,    60] loss: 0.057  
[74,    90] loss: 0.044  
[74,   120] loss: 0.024  
[74,   150] loss: 0.022  
[74,   180] loss: 0.036  
[74,   210] loss: 0.011  
[74,   240] loss: 0.052  
[74,   270] loss: 0.056  
[74,   300] loss: 0.050  
[74,   330] loss: 0.043  
[74,   360] loss: 0.052  
[74,   390] loss: 0.062  
[74,   420] loss: 0.050  
[74,   450] loss: 0.047  
[74,   480] loss: 0.021  
[74,   510] loss: 0.031  
[74,   540] loss: 0.073  
[74,   570] loss: 0.060  
[74,   600] loss: 0.056  
[74,   630] loss: 0.053  
[74,   660] loss: 0.035
```

```

[74, 690] loss: 0.043
[74, 720] loss: 0.058
[74, 750] loss: 0.076
[74, 780] loss: 0.053
[74, 810] loss: 0.073
[74, 840] loss: 0.037
[74, 870] loss: 0.043
[74, 900] loss: 0.029
[74, 930] loss: 0.044
[74, 960] loss: 0.030
[74, 990] loss: 0.030
[74, 1020] loss: 0.036
[74, 1050] loss: 0.021
[74, 1080] loss: 0.033
[74, 1110] loss: 0.045
[74, 1140] loss: 0.063
[74, 1170] loss: 0.040
[74, 1200] loss: 0.055
[74, 1230] loss: 0.035
[74, 1260] loss: 0.025
[74, 1290] loss: 0.029
[74, 1320] loss: 0.091
[74, 1350] loss: 0.052
[74, 1380] loss: 0.057
[74, 1410] loss: 0.029
[74, 1440] loss: 0.039
[74, 1470] loss: 0.041
[74, 1500] loss: 0.067
[74, 1530] loss: 0.031
[74, 1560] loss: 0.050
[74, 1590] loss: 0.036
[74, 1620] loss: 0.041
[74, 1650] loss: 0.039
[74, 1680] loss: 0.047
[74, 1710] loss: 0.051
[74, 1740] loss: 0.045
[74, 1770] loss: 0.044
[74, 1800] loss: 0.046
[74, 1830] loss: 0.025
[74, 1860] loss: 0.049
[74, 1890] loss: 0.022
[74, 1920] loss: 0.039
[74, 1950] loss: 0.029
[74, 1980] loss: 0.048
[74, 2010] loss: 0.043
[74, 2040] loss: 0.043
[74, 2070] loss: 0.037
[74, 2100] loss: 0.057
[74, 2130] loss: 0.038
[74, 2160] loss: 0.026
[74, 2190] loss: 0.051
ovft_intermediate_models/epoch73_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.991753
epoch 73 accuracy on train set is: 0.9917529989094874
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	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

evaluating on test set...

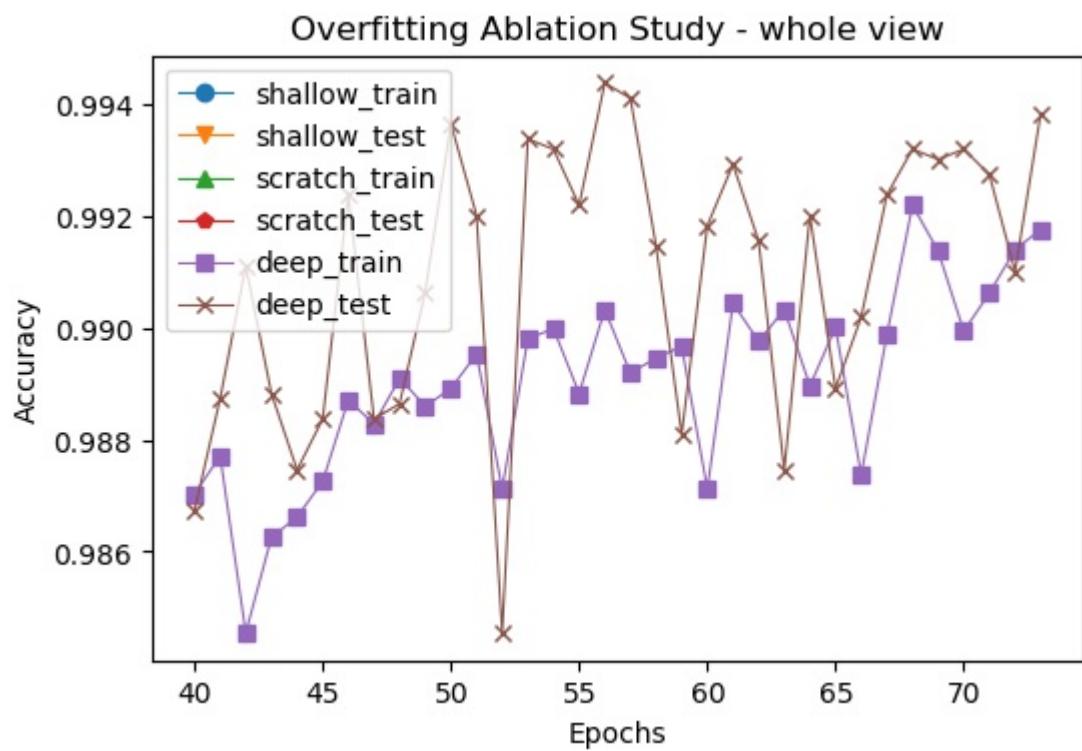
Accuracy on test images: 0.993820

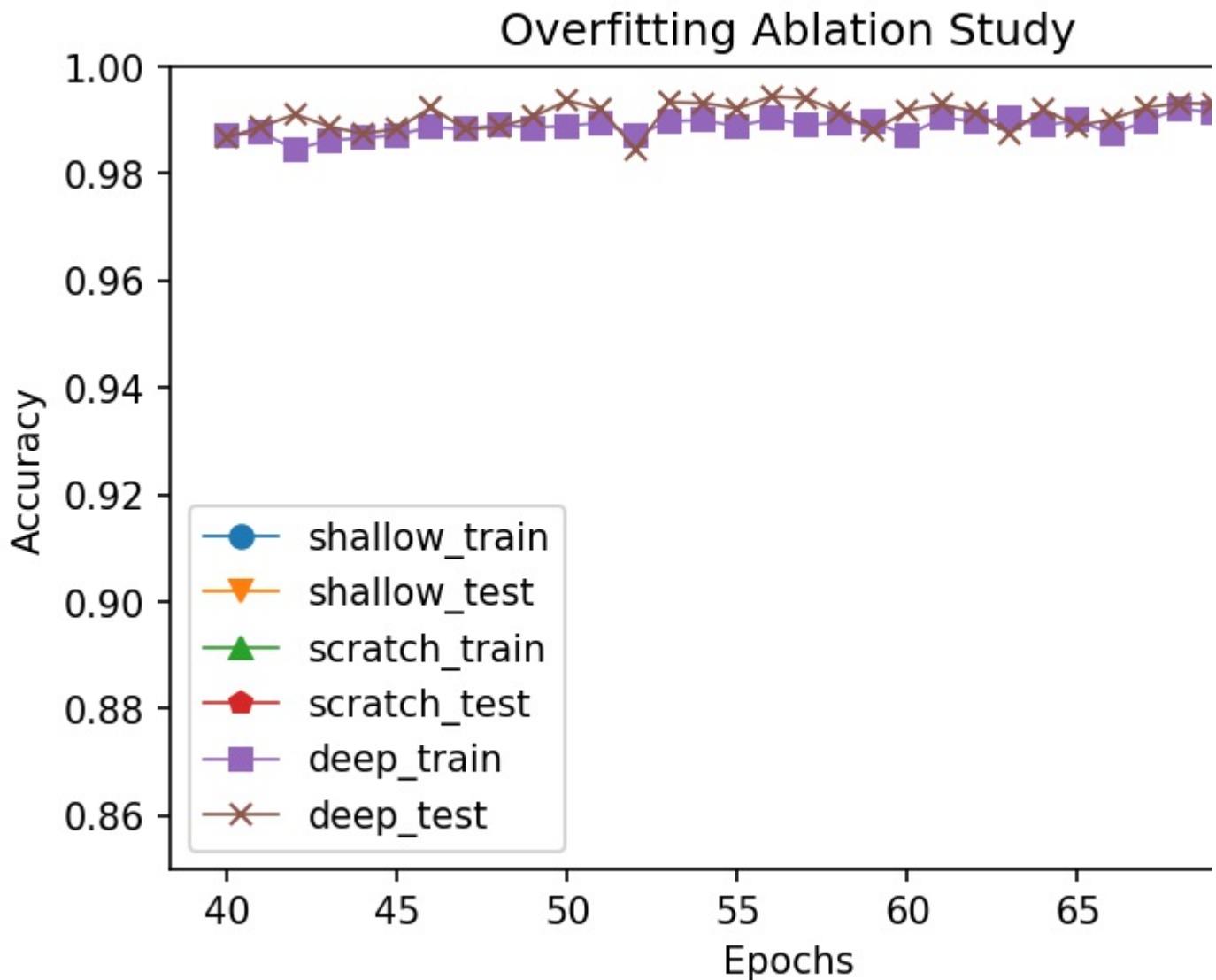
epoch 73 accuracy on test set is: 0.9938204289349327

evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN





```
epoch 73 takes 317 seconds.  
[75,    30] loss: 0.036  
[75,    60] loss: 0.032  
[75,   90] loss: 0.017  
[75,  120] loss: 0.042  
[75,  150] loss: 0.035  
[75,  180] loss: 0.026  
[75,  210] loss: 0.024  
[75,  240] loss: 0.038  
[75,  270] loss: 0.060  
[75,  300] loss: 0.061  
[75,  330] loss: 0.051  
[75,  360] loss: 0.069  
[75,  390] loss: 0.053  
[75,  420] loss: 0.051  
[75,  450] loss: 0.041  
[75,  480] loss: 0.069  
[75,  510] loss: 0.025  
[75,  540] loss: 0.029  
[75,  570] loss: 0.066  
[75,  600] loss: 0.035  
[75,  630] loss: 0.013  
[75,  660] loss: 0.036
```

```

[75, 690] loss: 0.042
[75, 720] loss: 0.022
[75, 750] loss: 0.026
[75, 780] loss: 0.031
[75, 810] loss: 0.033
[75, 840] loss: 0.045
[75, 870] loss: 0.059
[75, 900] loss: 0.039
[75, 930] loss: 0.056
[75, 960] loss: 0.058
[75, 990] loss: 0.045
[75, 1020] loss: 0.043
[75, 1050] loss: 0.052
[75, 1080] loss: 0.033
[75, 1110] loss: 0.048
[75, 1140] loss: 0.060
[75, 1170] loss: 0.024
[75, 1200] loss: 0.070
[75, 1230] loss: 0.038
[75, 1260] loss: 0.042
[75, 1290] loss: 0.043
[75, 1320] loss: 0.038
[75, 1350] loss: 0.048
[75, 1380] loss: 0.030
[75, 1410] loss: 0.027
[75, 1440] loss: 0.048
[75, 1470] loss: 0.037
[75, 1500] loss: 0.034
[75, 1530] loss: 0.032
[75, 1560] loss: 0.037
[75, 1590] loss: 0.040
[75, 1620] loss: 0.039
[75, 1650] loss: 0.021
[75, 1680] loss: 0.021
[75, 1710] loss: 0.036
[75, 1740] loss: 0.048
[75, 1770] loss: 0.067
[75, 1800] loss: 0.064
[75, 1830] loss: 0.048
[75, 1860] loss: 0.040
[75, 1890] loss: 0.059
[75, 1920] loss: 0.045
[75, 1950] loss: 0.056
[75, 1980] loss: 0.031
[75, 2010] loss: 0.042
[75, 2040] loss: 0.040
[75, 2070] loss: 0.048
[75, 2100] loss: 0.032
[75, 2130] loss: 0.023
[75, 2160] loss: 0.042
[75, 2190] loss: 0.036
ovft_intermediate_models/epoch74_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.989731
epoch 74 accuracy on train set is: 0.9897310069065794
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	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

evaluating on test set...

Accuracy on test images: 0.993730

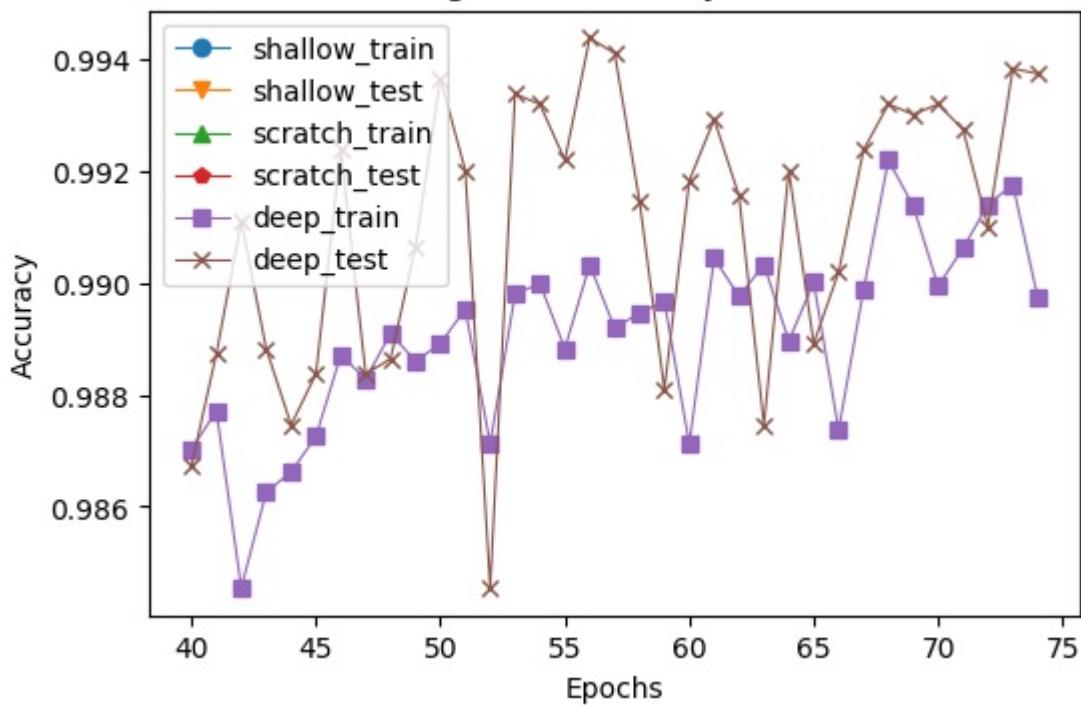
epoch 74 accuracy on test set is: 0.9937295528898582

evaluating on test set takes 16 seconds.

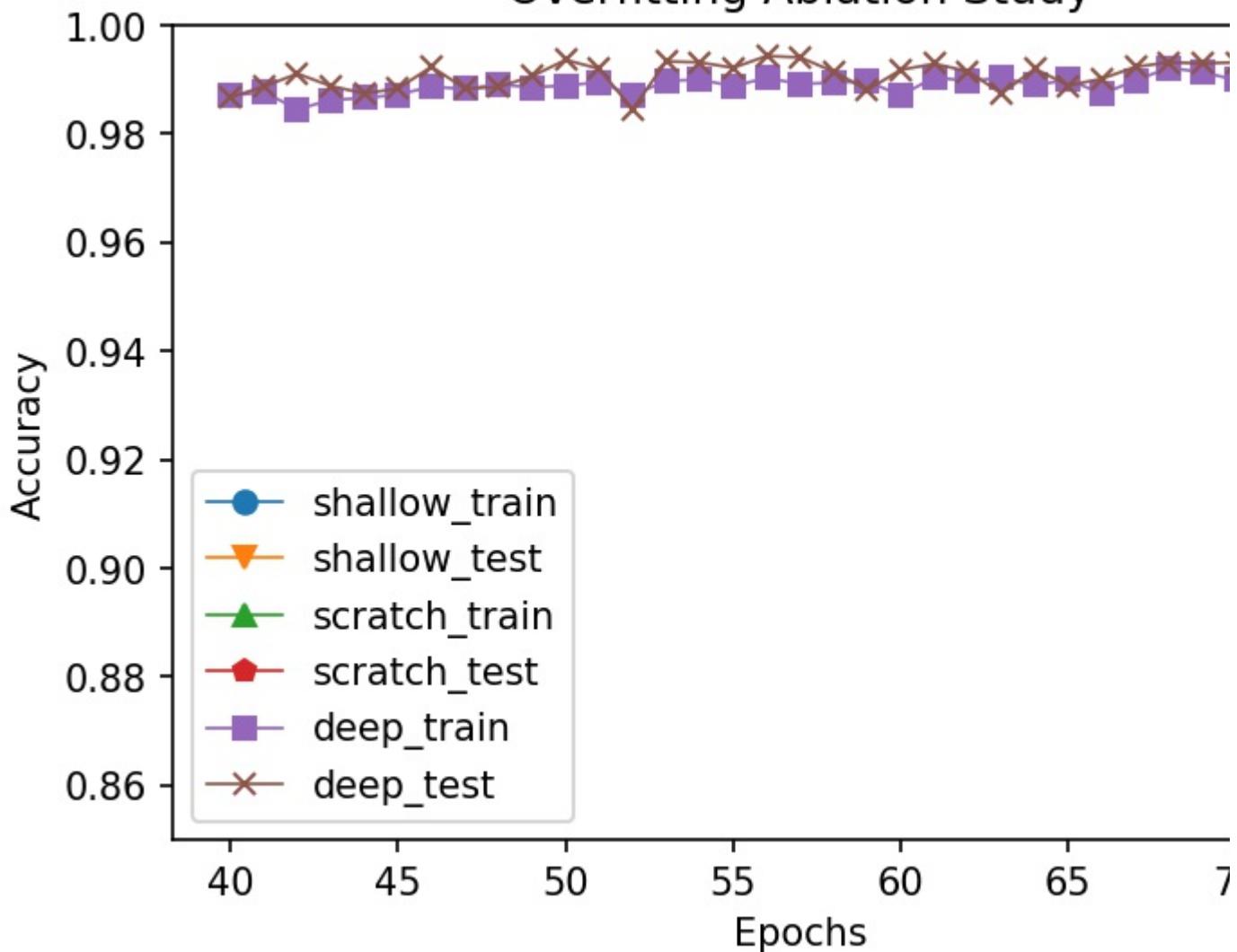
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	NaN	NaN	NaN
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 74 takes 316 seconds.  
[76,    30] loss: 0.039  
[76,    60] loss: 0.050  
[76,   90] loss: 0.034  
[76,  120] loss: 0.033  
[76,  150] loss: 0.042  
[76,  180] loss: 0.031  
[76,  210] loss: 0.054  
[76,  240] loss: 0.045  
[76,  270] loss: 0.047  
[76,  300] loss: 0.024  
[76,  330] loss: 0.018  
[76,  360] loss: 0.040  
[76,  390] loss: 0.045  
[76,  420] loss: 0.059  
[76,  450] loss: 0.067  
[76,  480] loss: 0.039  
[76,  510] loss: 0.034  
[76,  540] loss: 0.040  
[76,  570] loss: 0.043  
[76,  600] loss: 0.025  
[76,  630] loss: 0.018  
[76,  660] loss: 0.047
```

```

[76, 690] loss: 0.054
[76, 720] loss: 0.029
[76, 750] loss: 0.029
[76, 780] loss: 0.051
[76, 810] loss: 0.032
[76, 840] loss: 0.036
[76, 870] loss: 0.060
[76, 900] loss: 0.045
[76, 930] loss: 0.045
[76, 960] loss: 0.021
[76, 990] loss: 0.027
[76, 1020] loss: 0.070
[76, 1050] loss: 0.053
[76, 1080] loss: 0.024
[76, 1110] loss: 0.066
[76, 1140] loss: 0.045
[76, 1170] loss: 0.051
[76, 1200] loss: 0.036
[76, 1230] loss: 0.058
[76, 1260] loss: 0.035
[76, 1290] loss: 0.033
[76, 1320] loss: 0.018
[76, 1350] loss: 0.039
[76, 1380] loss: 0.024
[76, 1410] loss: 0.036
[76, 1440] loss: 0.050
[76, 1470] loss: 0.037
[76, 1500] loss: 0.050
[76, 1530] loss: 0.063
[76, 1560] loss: 0.029
[76, 1590] loss: 0.032
[76, 1620] loss: 0.039
[76, 1650] loss: 0.042
[76, 1680] loss: 0.026
[76, 1710] loss: 0.032
[76, 1740] loss: 0.043
[76, 1770] loss: 0.022
[76, 1800] loss: 0.070
[76, 1830] loss: 0.036
[76, 1860] loss: 0.052
[76, 1890] loss: 0.033
[76, 1920] loss: 0.034
[76, 1950] loss: 0.042
[76, 1980] loss: 0.063
[76, 2010] loss: 0.060
[76, 2040] loss: 0.026
[76, 2070] loss: 0.035
[76, 2100] loss: 0.045
[76, 2130] loss: 0.038
[76, 2160] loss: 0.018
[76, 2190] loss: 0.046
ovft_intermediate_models/epoch75_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.990594
epoch 75 accuracy on train set is: 0.9905943293347873
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	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

evaluating on test set...

Accuracy on test images: 0.993366

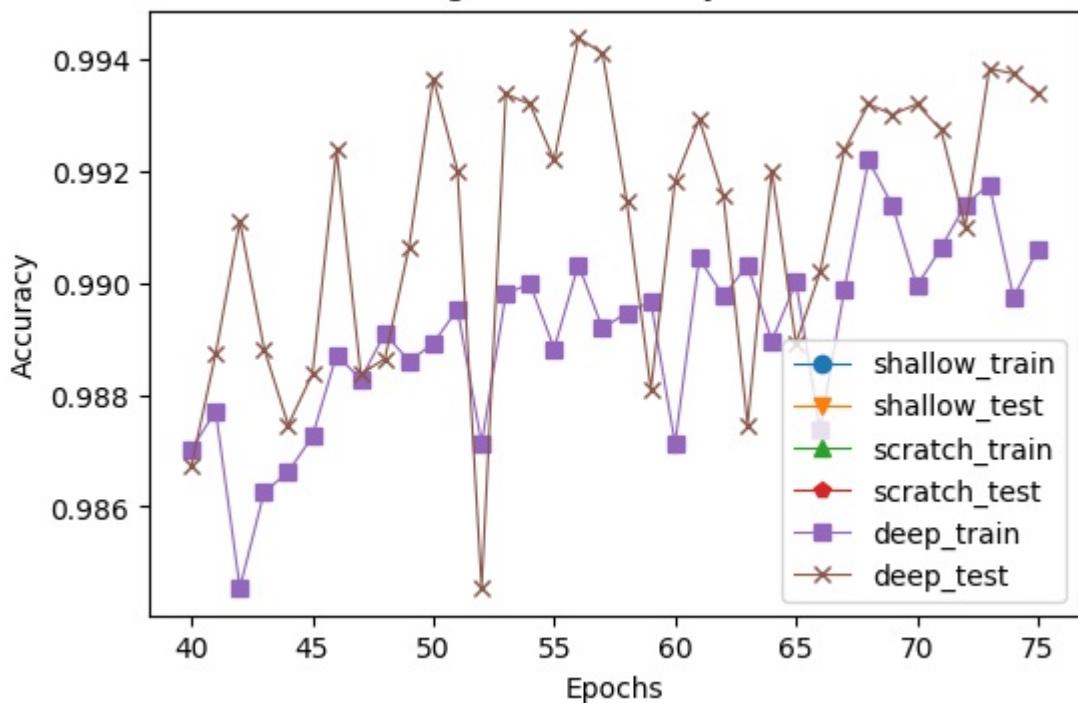
epoch 75 accuracy on test set is: 0.9933660487095601

evaluating on test set takes 16 seconds.

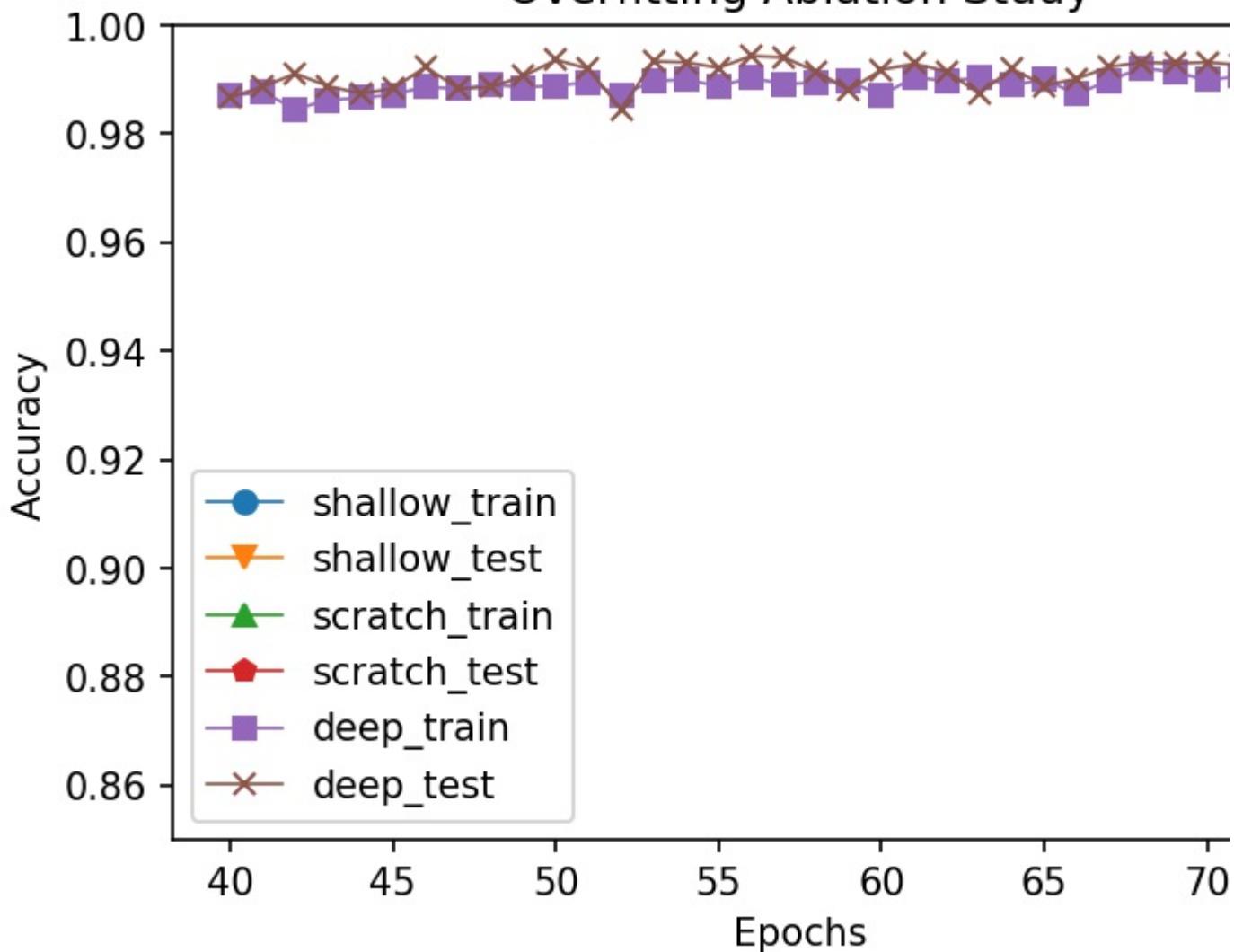
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	NaN	NaN	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 75 takes 319 seconds.  
[77,    30] loss: 0.031  
[77,    60] loss: 0.041  
[77,   90] loss: 0.045  
[77,  120] loss: 0.064  
[77,  150] loss: 0.048  
[77,  180] loss: 0.052  
[77,  210] loss: 0.035  
[77,  240] loss: 0.043  
[77,  270] loss: 0.033  
[77,  300] loss: 0.027  
[77,  330] loss: 0.053  
[77,  360] loss: 0.031  
[77,  390] loss: 0.055  
[77,  420] loss: 0.047  
[77,  450] loss: 0.038  
[77,  480] loss: 0.043  
[77,  510] loss: 0.031  
[77,  540] loss: 0.044  
[77,  570] loss: 0.040  
[77,  600] loss: 0.037  
[77,  630] loss: 0.020  
[77,  660] loss: 0.047
```

```

[77, 690] loss: 0.039
[77, 720] loss: 0.063
[77, 750] loss: 0.036
[77, 780] loss: 0.068
[77, 810] loss: 0.071
[77, 840] loss: 0.046
[77, 870] loss: 0.024
[77, 900] loss: 0.058
[77, 930] loss: 0.040
[77, 960] loss: 0.025
[77, 990] loss: 0.049
[77, 1020] loss: 0.045
[77, 1050] loss: 0.044
[77, 1080] loss: 0.034
[77, 1110] loss: 0.045
[77, 1140] loss: 0.030
[77, 1170] loss: 0.018
[77, 1200] loss: 0.024
[77, 1230] loss: 0.022
[77, 1260] loss: 0.066
[77, 1290] loss: 0.048
[77, 1320] loss: 0.021
[77, 1350] loss: 0.030
[77, 1380] loss: 0.038
[77, 1410] loss: 0.038
[77, 1440] loss: 0.046
[77, 1470] loss: 0.019
[77, 1500] loss: 0.034
[77, 1530] loss: 0.033
[77, 1560] loss: 0.064
[77, 1590] loss: 0.040
[77, 1620] loss: 0.055
[77, 1650] loss: 0.053
[77, 1680] loss: 0.089
[77, 1710] loss: 0.025
[77, 1740] loss: 0.030
[77, 1770] loss: 0.054
[77, 1800] loss: 0.055
[77, 1830] loss: 0.070
[77, 1860] loss: 0.029
[77, 1890] loss: 0.059
[77, 1920] loss: 0.047
[77, 1950] loss: 0.042
[77, 1980] loss: 0.027
[77, 2010] loss: 0.030
[77, 2040] loss: 0.043
[77, 2070] loss: 0.045
[77, 2100] loss: 0.036
[77, 2130] loss: 0.031
[77, 2160] loss: 0.027
[77, 2190] loss: 0.039
ovft_intermediate_models/epoch76_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.992253
epoch 76 accuracy on train set is: 0.9922528171573973
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	NaN
77	NaN	NaN	NaN	NaN	NaN	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993275

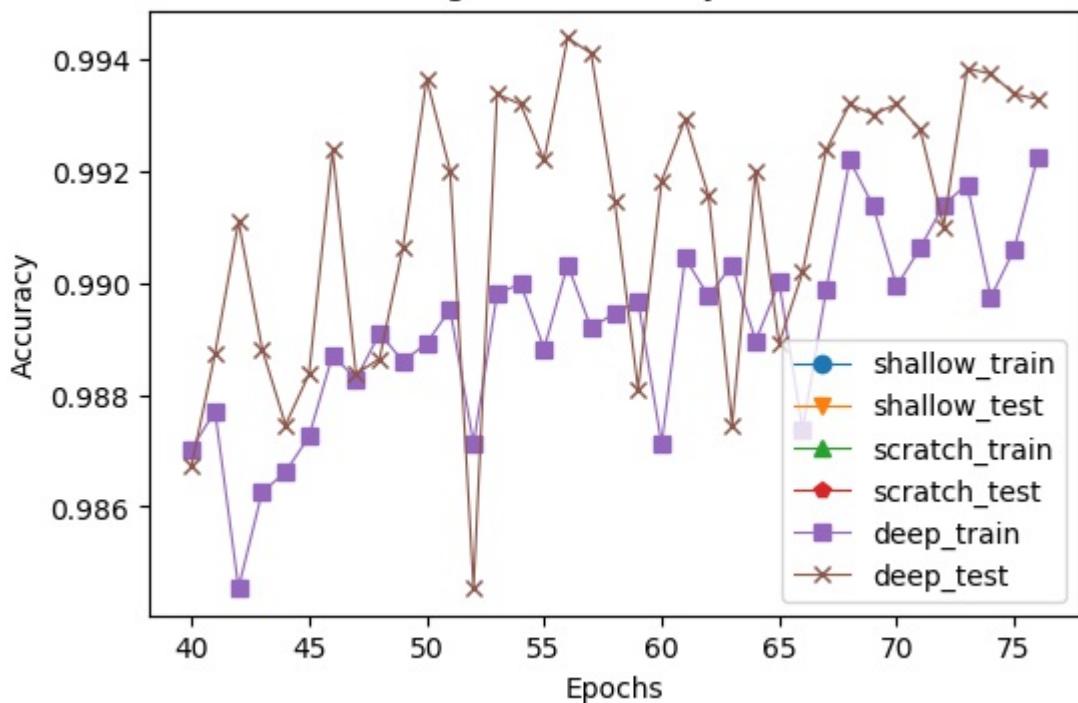
epoch 76 accuracy on test set is: 0.9932751726644856

evaluating on test set takes 16 seconds.

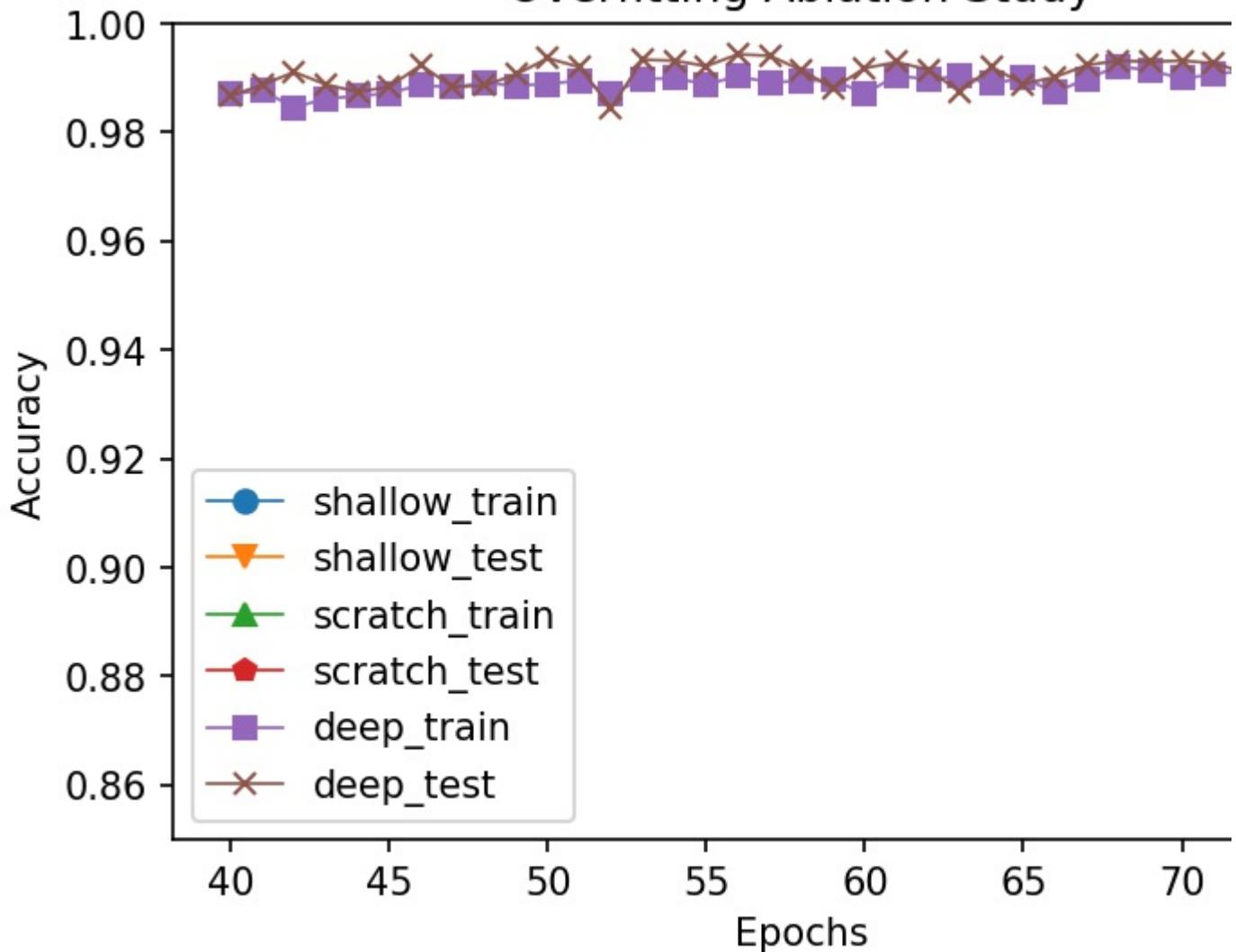
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	Nan	Nan	NaN
78	NaN	NaN	NaN	NaN	Nan	Nan	NaN
79	NaN	NaN	NaN	NaN	Nan	Nan	NaN

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 76 takes 316 seconds.  
[78,    30] loss: 0.029  
[78,    60] loss: 0.045  
[78,   90] loss: 0.030  
[78,  120] loss: 0.041  
[78,  150] loss: 0.036  
[78,  180] loss: 0.023  
[78,  210] loss: 0.024  
[78,  240] loss: 0.049  
[78,  270] loss: 0.049  
[78,  300] loss: 0.051  
[78,  330] loss: 0.037  
[78,  360] loss: 0.049  
[78,  390] loss: 0.059  
[78,  420] loss: 0.025  
[78,  450] loss: 0.032  
[78,  480] loss: 0.022  
[78,  510] loss: 0.029  
[78,  540] loss: 0.052  
[78,  570] loss: 0.079  
[78,  600] loss: 0.075  
[78,  630] loss: 0.039  
[78,  660] loss: 0.048
```

```

[78, 690] loss: 0.045
[78, 720] loss: 0.063
[78, 750] loss: 0.043
[78, 780] loss: 0.038
[78, 810] loss: 0.059
[78, 840] loss: 0.021
[78, 870] loss: 0.067
[78, 900] loss: 0.060
[78, 930] loss: 0.047
[78, 960] loss: 0.049
[78, 990] loss: 0.028
[78, 1020] loss: 0.037
[78, 1050] loss: 0.043
[78, 1080] loss: 0.029
[78, 1110] loss: 0.043
[78, 1140] loss: 0.028
[78, 1170] loss: 0.054
[78, 1200] loss: 0.028
[78, 1230] loss: 0.038
[78, 1260] loss: 0.019
[78, 1290] loss: 0.058
[78, 1320] loss: 0.023
[78, 1350] loss: 0.046
[78, 1380] loss: 0.044
[78, 1410] loss: 0.041
[78, 1440] loss: 0.053
[78, 1470] loss: 0.051
[78, 1500] loss: 0.032
[78, 1530] loss: 0.050
[78, 1560] loss: 0.023
[78, 1590] loss: 0.021
[78, 1620] loss: 0.034
[78, 1650] loss: 0.023
[78, 1680] loss: 0.067
[78, 1710] loss: 0.033
[78, 1740] loss: 0.022
[78, 1770] loss: 0.038
[78, 1800] loss: 0.037
[78, 1830] loss: 0.042
[78, 1860] loss: 0.026
[78, 1890] loss: 0.027
[78, 1920] loss: 0.044
[78, 1950] loss: 0.061
[78, 1980] loss: 0.036
[78, 2010] loss: 0.024
[78, 2040] loss: 0.042
[78, 2070] loss: 0.060
[78, 2100] loss: 0.031
[78, 2130] loss: 0.062
[78, 2160] loss: 0.033
[78, 2190] loss: 0.074
ovft_intermediate_models/epoch77_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.990549
epoch 77 accuracy on train set is: 0.9905488913122501
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	NaN
78	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.992821

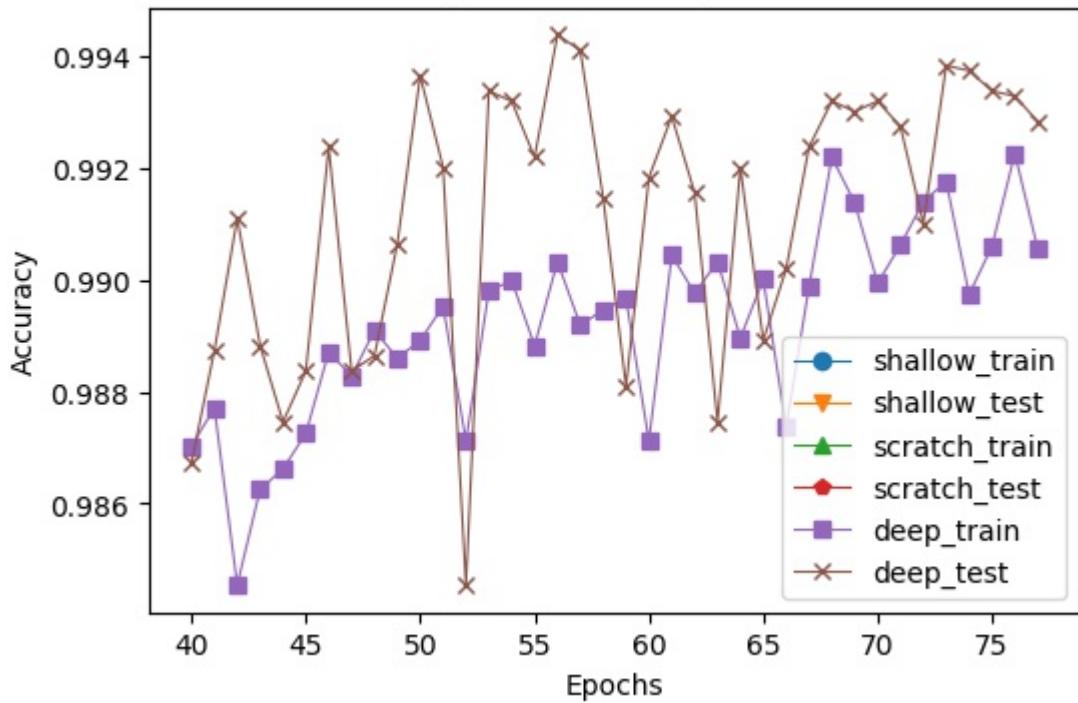
epoch 77 accuracy on test set is: 0.992820792439113

evaluating on test set takes 16 seconds.

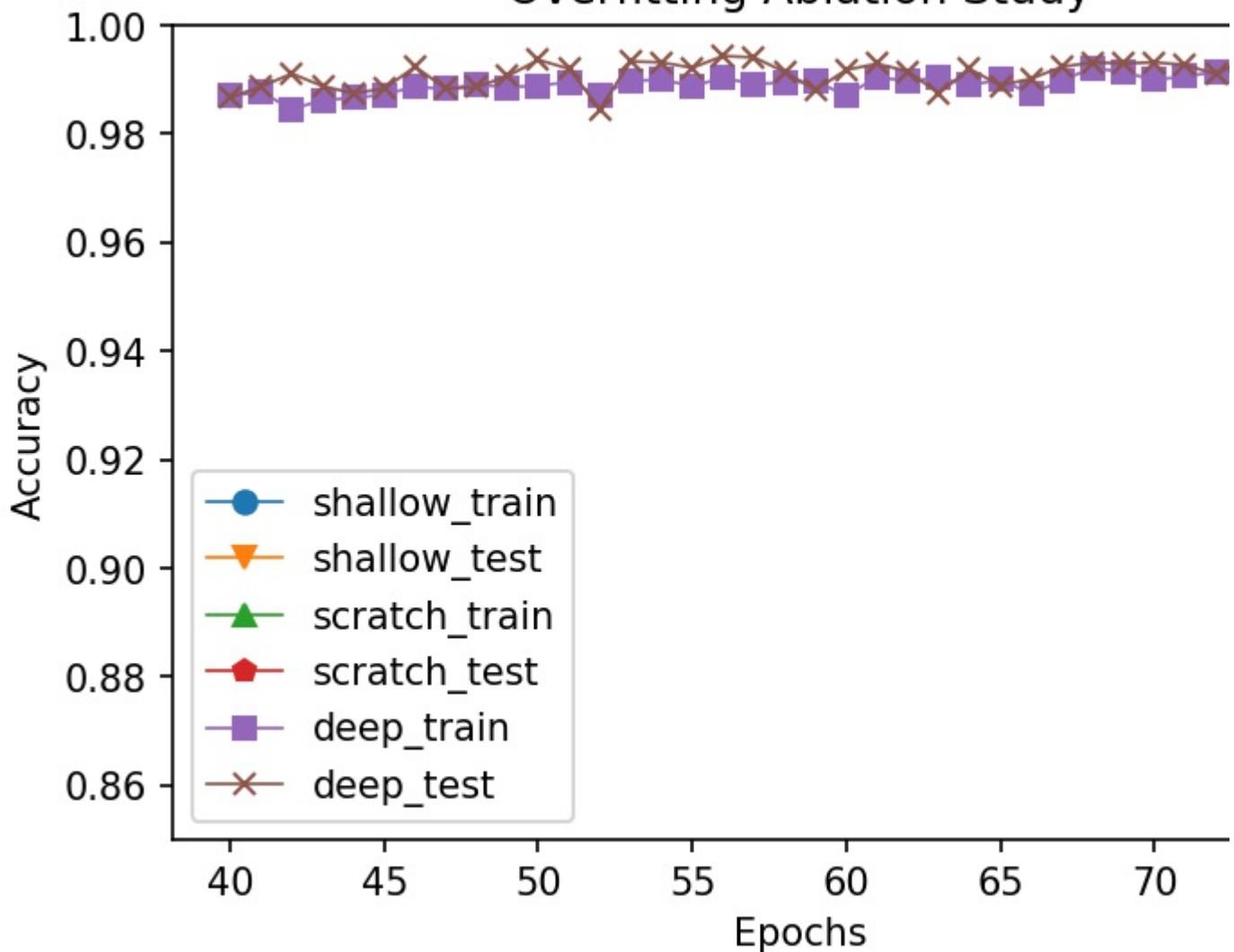
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	NaN	NaN	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 77 takes 316 seconds.  
[79,    30] loss: 0.063  
[79,    60] loss: 0.063  
[79,   90] loss: 0.030  
[79,  120] loss: 0.044  
[79,  150] loss: 0.033  
[79,  180] loss: 0.024  
[79,  210] loss: 0.044  
[79,  240] loss: 0.042  
[79,  270] loss: 0.037  
[79,  300] loss: 0.029  
[79,  330] loss: 0.045  
[79,  360] loss: 0.028  
[79,  390] loss: 0.029  
[79,  420] loss: 0.037  
[79,  450] loss: 0.043  
[79,  480] loss: 0.032  
[79,  510] loss: 0.044  
[79,  540] loss: 0.036  
[79,  570] loss: 0.038  
[79,  600] loss: 0.039  
[79,  630] loss: 0.057  
[79,  660] loss: 0.037
```

```

[79, 690] loss: 0.029
[79, 720] loss: 0.036
[79, 750] loss: 0.052
[79, 780] loss: 0.058
[79, 810] loss: 0.060
[79, 840] loss: 0.024
[79, 870] loss: 0.033
[79, 900] loss: 0.032
[79, 930] loss: 0.039
[79, 960] loss: 0.031
[79, 990] loss: 0.026
[79, 1020] loss: 0.021
[79, 1050] loss: 0.030
[79, 1080] loss: 0.050
[79, 1110] loss: 0.062
[79, 1140] loss: 0.026
[79, 1170] loss: 0.031
[79, 1200] loss: 0.086
[79, 1230] loss: 0.043
[79, 1260] loss: 0.040
[79, 1290] loss: 0.046
[79, 1320] loss: 0.040
[79, 1350] loss: 0.019
[79, 1380] loss: 0.020
[79, 1410] loss: 0.039
[79, 1440] loss: 0.027
[79, 1470] loss: 0.033
[79, 1500] loss: 0.038
[79, 1530] loss: 0.047
[79, 1560] loss: 0.038
[79, 1590] loss: 0.045
[79, 1620] loss: 0.029
[79, 1650] loss: 0.038
[79, 1680] loss: 0.046
[79, 1710] loss: 0.054
[79, 1740] loss: 0.027
[79, 1770] loss: 0.034
[79, 1800] loss: 0.040
[79, 1830] loss: 0.051
[79, 1860] loss: 0.033
[79, 1890] loss: 0.051
[79, 1920] loss: 0.036
[79, 1950] loss: 0.042
[79, 1980] loss: 0.045
[79, 2010] loss: 0.049
[79, 2040] loss: 0.032
[79, 2070] loss: 0.040
[79, 2100] loss: 0.069
[79, 2130] loss: 0.032
[79, 2160] loss: 0.035
[79, 2190] loss: 0.041
ovft_intermediate_models/epoch78_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.991594
epoch 78 accuracy on train set is: 0.9915939658306071
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN          NaN          NaN          NaN    0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	NaN
79	NaN	NaN	NaN	NaN	NaN	NaN

evaluating on test set...

Accuracy on test images: 0.993003

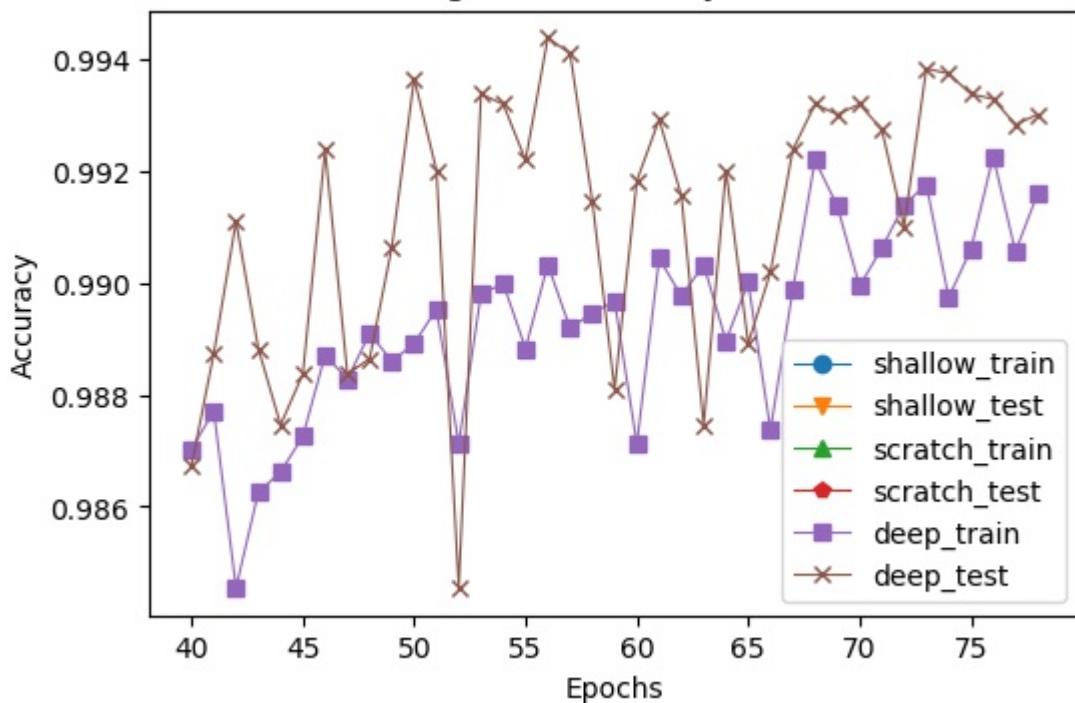
epoch 78 accuracy on test set is: 0.9930025445292621

evaluating on test set takes 16 seconds.

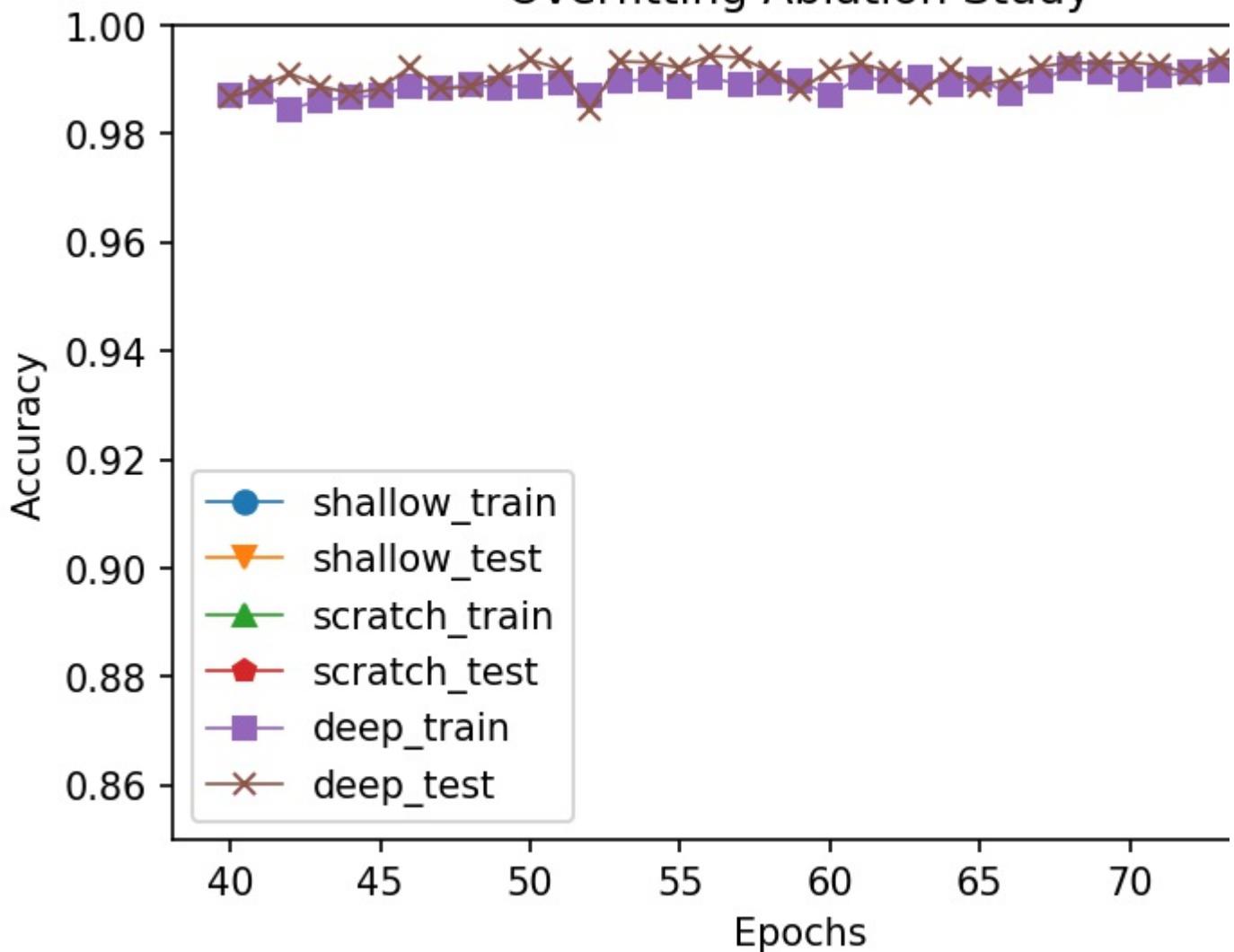
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 78 takes 316 seconds.  
[80,    30] loss: 0.020  
[80,    60] loss: 0.042  
[80,   90] loss: 0.048  
[80,  120] loss: 0.024  
[80,  150] loss: 0.036  
[80,  180] loss: 0.040  
[80,  210] loss: 0.038  
[80,  240] loss: 0.049  
[80,  270] loss: 0.077  
[80,  300] loss: 0.025  
[80,  330] loss: 0.039  
[80,  360] loss: 0.039  
[80,  390] loss: 0.055  
[80,  420] loss: 0.046  
[80,  450] loss: 0.024  
[80,  480] loss: 0.045  
[80,  510] loss: 0.061  
[80,  540] loss: 0.059  
[80,  570] loss: 0.038  
[80,  600] loss: 0.036  
[80,  630] loss: 0.027  
[80,  660] loss: 0.063
```

```

[80, 690] loss: 0.041
[80, 720] loss: 0.035
[80, 750] loss: 0.041
[80, 780] loss: 0.032
[80, 810] loss: 0.040
[80, 840] loss: 0.045
[80, 870] loss: 0.029
[80, 900] loss: 0.035
[80, 930] loss: 0.053
[80, 960] loss: 0.017
[80, 990] loss: 0.045
[80, 1020] loss: 0.020
[80, 1050] loss: 0.042
[80, 1080] loss: 0.043
[80, 1110] loss: 0.036
[80, 1140] loss: 0.084
[80, 1170] loss: 0.031
[80, 1200] loss: 0.043
[80, 1230] loss: 0.051
[80, 1260] loss: 0.039
[80, 1290] loss: 0.033
[80, 1320] loss: 0.034
[80, 1350] loss: 0.022
[80, 1380] loss: 0.037
[80, 1410] loss: 0.067
[80, 1440] loss: 0.042
[80, 1470] loss: 0.058
[80, 1500] loss: 0.047
[80, 1530] loss: 0.040
[80, 1560] loss: 0.051
[80, 1590] loss: 0.039
[80, 1620] loss: 0.021
[80, 1650] loss: 0.071
[80, 1680] loss: 0.039
[80, 1710] loss: 0.038
[80, 1740] loss: 0.044
[80, 1770] loss: 0.036
[80, 1800] loss: 0.029
[80, 1830] loss: 0.040
[80, 1860] loss: 0.037
[80, 1890] loss: 0.077
[80, 1920] loss: 0.052
[80, 1950] loss: 0.080
[80, 1980] loss: 0.052
[80, 2010] loss: 0.060
[80, 2040] loss: 0.054
[80, 2070] loss: 0.034
[80, 2100] loss: 0.037
[80, 2130] loss: 0.034
[80, 2160] loss: 0.057
[80, 2190] loss: 0.035
ovft_intermediate_models/epoch79_deep.pt saved.

evaluating on training set...
Accuracy on test images: 0.990731
epoch 79 accuracy on train set is: 0.9907306434023991
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40           NaN           NaN           NaN           NaN   0.987027   0.986732

```

41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	NaN

evaluating on test set...

Accuracy on test images: 0.993639

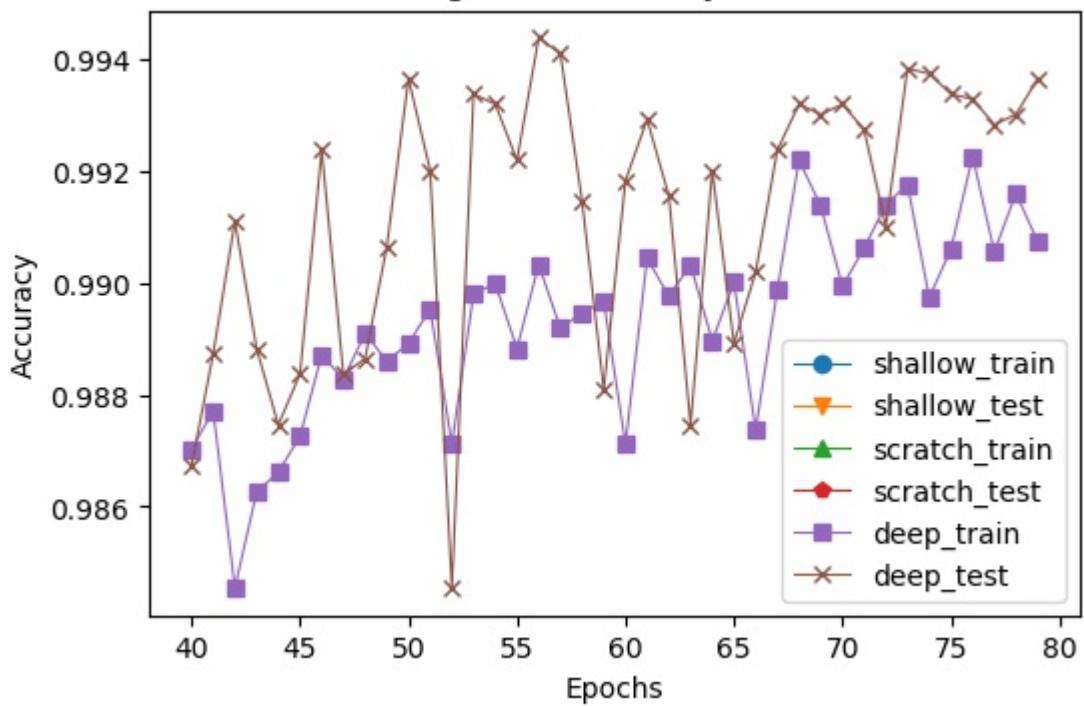
epoch 79 accuracy on test set is: 0.9936386768447837

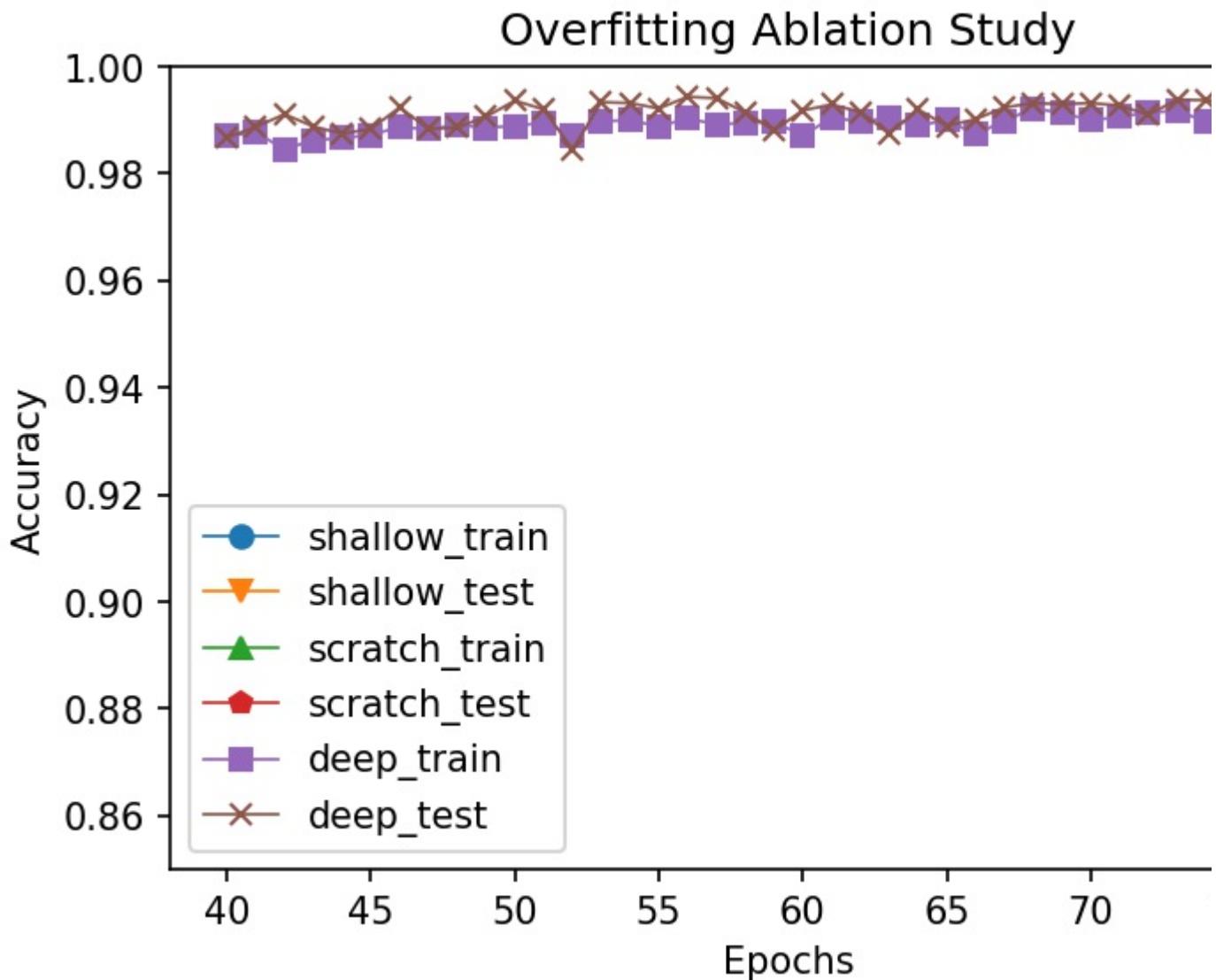
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	NaN	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view





epoch 79 takes 315 seconds.

Finished Training

Evaluating...

RETRAINING

Targeting alexnet with 39 classes

Checking layer size difference between torchvision.models and
torch.utils.model_zoo...

for layer classifier.6.weight torchvision.models has size 39 and model_zoo has
size 1000

for layer classifier.6.bias torchvision.models has size 39 and model_zoo has size
1000

Replacing the following state from initialized alexnet : ['classifier.6.weight',
'classifier.6.bias']

loading ovft_intermediate_models/epoch39_shallow.pt to continue training...

Resizing input images to max of (224, 224)
Transferring models to GPU(s)

```
Retraining...
/home/administrator/anaconda3/envs/torch_nvidia_390/lib/python3.7/site-packages/
torchvision/transforms/transforms.py:562: UserWarning: The use of the [REDACTED]
transforms.RandomSizedCrop transform is deprecated, please use [REDACTED]
transforms.RandomResizedCrop instead.
    warnings.warn("The use of the transforms.RandomSizedCrop transform is
deprecated, " + [REDACTED]
/home/administrator/anaconda3/envs/torch_nvidia_390/lib/python3.7/site-packages/
torchvision/transforms/transforms.py:187: UserWarning: The use of the [REDACTED]
transforms.Scale transform is deprecated, please use transforms.Resize instead.
    warnings.warn("The use of the transforms.Scale transform is deprecated, " +
[41,    30] loss: 0.545
[41,    60] loss: 0.464
[41,    90] loss: 0.414
[41,   120] loss: 0.498
[41,   150] loss: 0.442
[41,   180] loss: 0.434
[41,   210] loss: 0.462
[41,   240] loss: 0.567
[41,   270] loss: 0.387
[41,   300] loss: 0.453
[41,   330] loss: 0.459
[41,   360] loss: 0.547
[41,   390] loss: 0.516
[41,   420] loss: 0.416
[41,   450] loss: 0.387
[41,   480] loss: 0.420
[41,   510] loss: 0.453
[41,   540] loss: 0.339
[41,   570] loss: 0.417
[41,   600] loss: 0.425
[41,   630] loss: 0.501
[41,   660] loss: 0.464
[41,   690] loss: 0.515
[41,   720] loss: 0.521
[41,   750] loss: 0.550
[41,   780] loss: 0.557
[41,   810] loss: 0.520
[41,   840] loss: 0.449
[41,   870] loss: 0.440
[41,   900] loss: 0.554
[41,   930] loss: 0.507
[41,   960] loss: 0.465
[41,   990] loss: 0.505
[41,  1020] loss: 0.463
[41,  1050] loss: 0.551
[41,  1080] loss: 0.549
[41,  1110] loss: 0.510
[41,  1140] loss: 0.511
[41,  1170] loss: 0.510
[41,  1200] loss: 0.423
[41,  1230] loss: 0.447
[41,  1260] loss: 0.531
[41,  1290] loss: 0.450
[41,  1320] loss: 0.464
[41,  1350] loss: 0.475
[41,  1380] loss: 0.457
[41,  1410] loss: 0.494
[41,  1440] loss: 0.446
```

```

[41, 1470] loss: 0.423
[41, 1500] loss: 0.432
[41, 1530] loss: 0.479
[41, 1560] loss: 0.457
[41, 1590] loss: 0.481
[41, 1620] loss: 0.516
[41, 1650] loss: 0.459
[41, 1680] loss: 0.459
[41, 1710] loss: 0.514
[41, 1740] loss: 0.400
[41, 1770] loss: 0.420
[41, 1800] loss: 0.445
[41, 1830] loss: 0.445
[41, 1860] loss: 0.490
[41, 1890] loss: 0.473
[41, 1920] loss: 0.465
[41, 1950] loss: 0.455
[41, 1980] loss: 0.437
[41, 2010] loss: 0.455
[41, 2040] loss: 0.466
[41, 2070] loss: 0.432
[41, 2100] loss: 0.440
[41, 2130] loss: 0.412
[41, 2160] loss: 0.500
[41, 2190] loss: 0.487

```

`ovft_intermediate_models/epoch40_shallow.pt saved`

evaluating on training set...

Accuracy on test images: 0.910987

epoch 40 accuracy on train set is: 0.9109869138495092

evaluating on training set takes 66 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	NaN	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185

67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

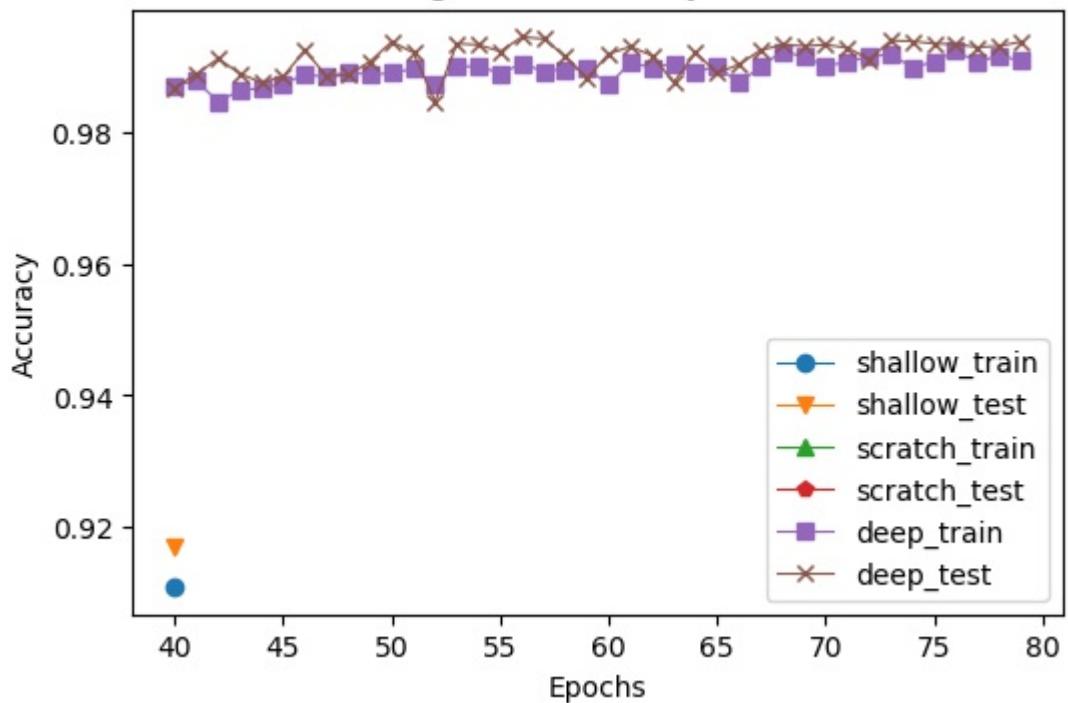
Accuracy on test images: 0.917030

epoch 40 accuracy on test set is: 0.9170301708469647

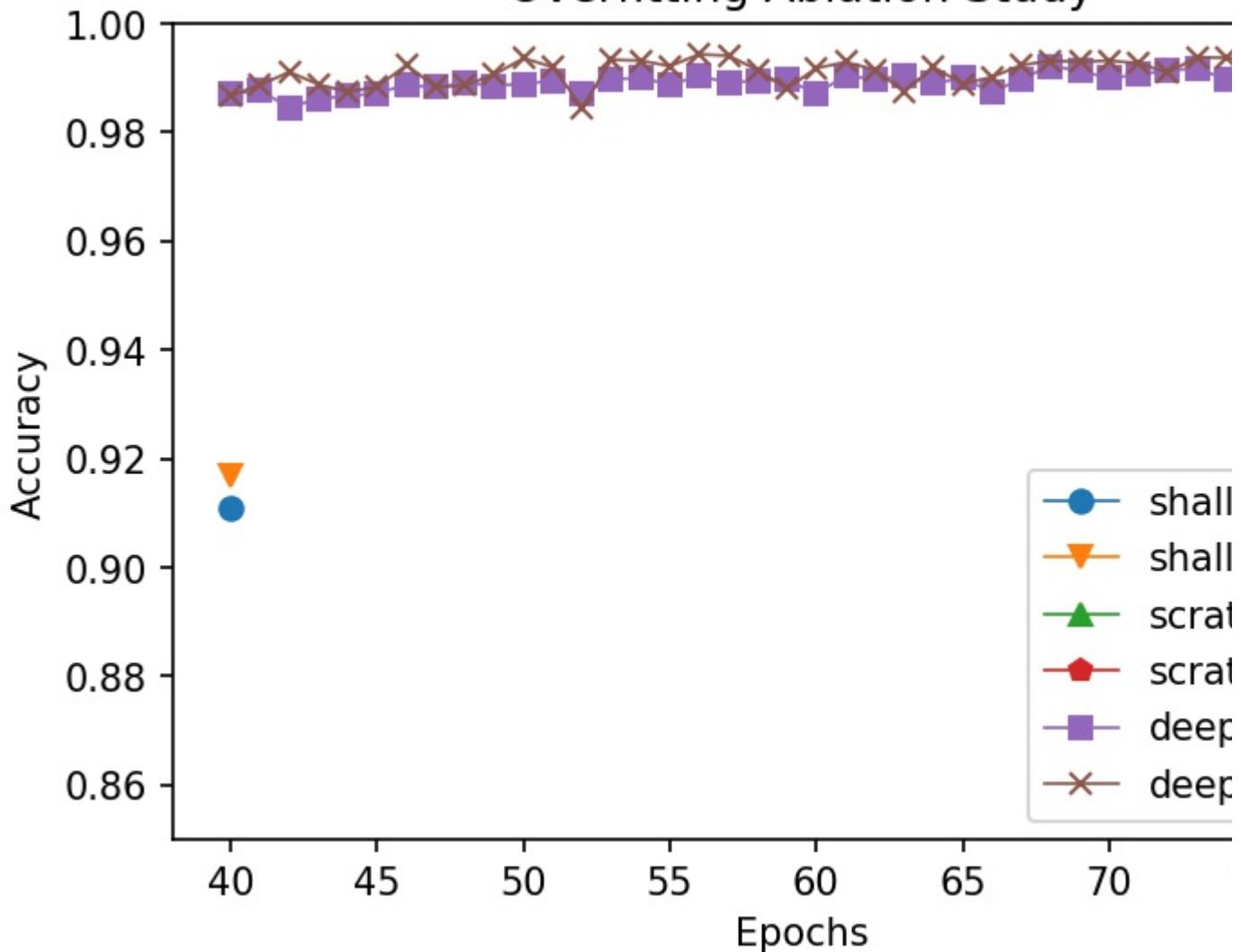
evaluating on test set takes 16 seconds.

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	NaN	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 40 takes 152 seconds.  
[42,    30] loss: 0.433  
[42,    60] loss: 0.428  
[42,   90] loss: 0.428  
[42,  120] loss: 0.432  
[42,  150] loss: 0.476  
[42,  180] loss: 0.454  
[42,  210] loss: 0.440  
[42,  240] loss: 0.524  
[42,  270] loss: 0.463  
[42,  300] loss: 0.537  
[42,  330] loss: 0.578  
[42,  360] loss: 0.398  
[42,  390] loss: 0.408  
[42,  420] loss: 0.393  
[42,  450] loss: 0.564  
[42,  480] loss: 0.434  
[42,  510] loss: 0.476  
[42,  540] loss: 0.481  
[42,  570] loss: 0.438  
[42,  600] loss: 0.450  
[42,  630] loss: 0.499  
[42,  660] loss: 0.543
```

```

[42, 690] loss: 0.433
[42, 720] loss: 0.499
[42, 750] loss: 0.450
[42, 780] loss: 0.515
[42, 810] loss: 0.428
[42, 840] loss: 0.516
[42, 870] loss: 0.495
[42, 900] loss: 0.579
[42, 930] loss: 0.453
[42, 960] loss: 0.508
[42, 990] loss: 0.424
[42, 1020] loss: 0.452
[42, 1050] loss: 0.421
[42, 1080] loss: 0.430
[42, 1110] loss: 0.448
[42, 1140] loss: 0.568
[42, 1170] loss: 0.433
[42, 1200] loss: 0.560
[42, 1230] loss: 0.401
[42, 1260] loss: 0.473
[42, 1290] loss: 0.476
[42, 1320] loss: 0.478
[42, 1350] loss: 0.430
[42, 1380] loss: 0.479
[42, 1410] loss: 0.493
[42, 1440] loss: 0.509
[42, 1470] loss: 0.557
[42, 1500] loss: 0.449
[42, 1530] loss: 0.398
[42, 1560] loss: 0.474
[42, 1590] loss: 0.480
[42, 1620] loss: 0.496
[42, 1650] loss: 0.488
[42, 1680] loss: 0.414
[42, 1710] loss: 0.432
[42, 1740] loss: 0.499
[42, 1770] loss: 0.461
[42, 1800] loss: 0.542
[42, 1830] loss: 0.434
[42, 1860] loss: 0.475
[42, 1890] loss: 0.541
[42, 1920] loss: 0.479
[42, 1950] loss: 0.460
[42, 1980] loss: 0.530
[42, 2010] loss: 0.485
[42, 2040] loss: 0.482
[42, 2070] loss: 0.486
[42, 2100] loss: 0.500
[42, 2130] loss: 0.463
[42, 2160] loss: 0.461
[42, 2190] loss: 0.446
ovft_intermediate_models/epoch41_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.909147
epoch 41 accuracy on train set is: 0.9091466739367503
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	NaN	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.901945

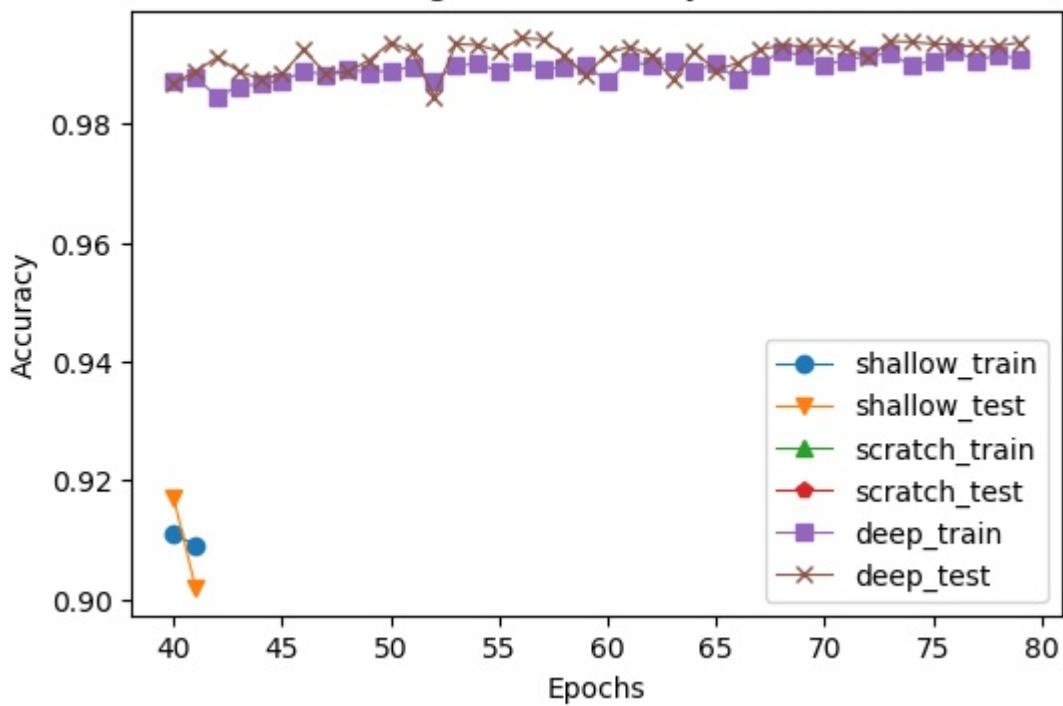
epoch 41 accuracy on test set is: 0.9019447473645947

evaluating on test set takes 16 seconds.

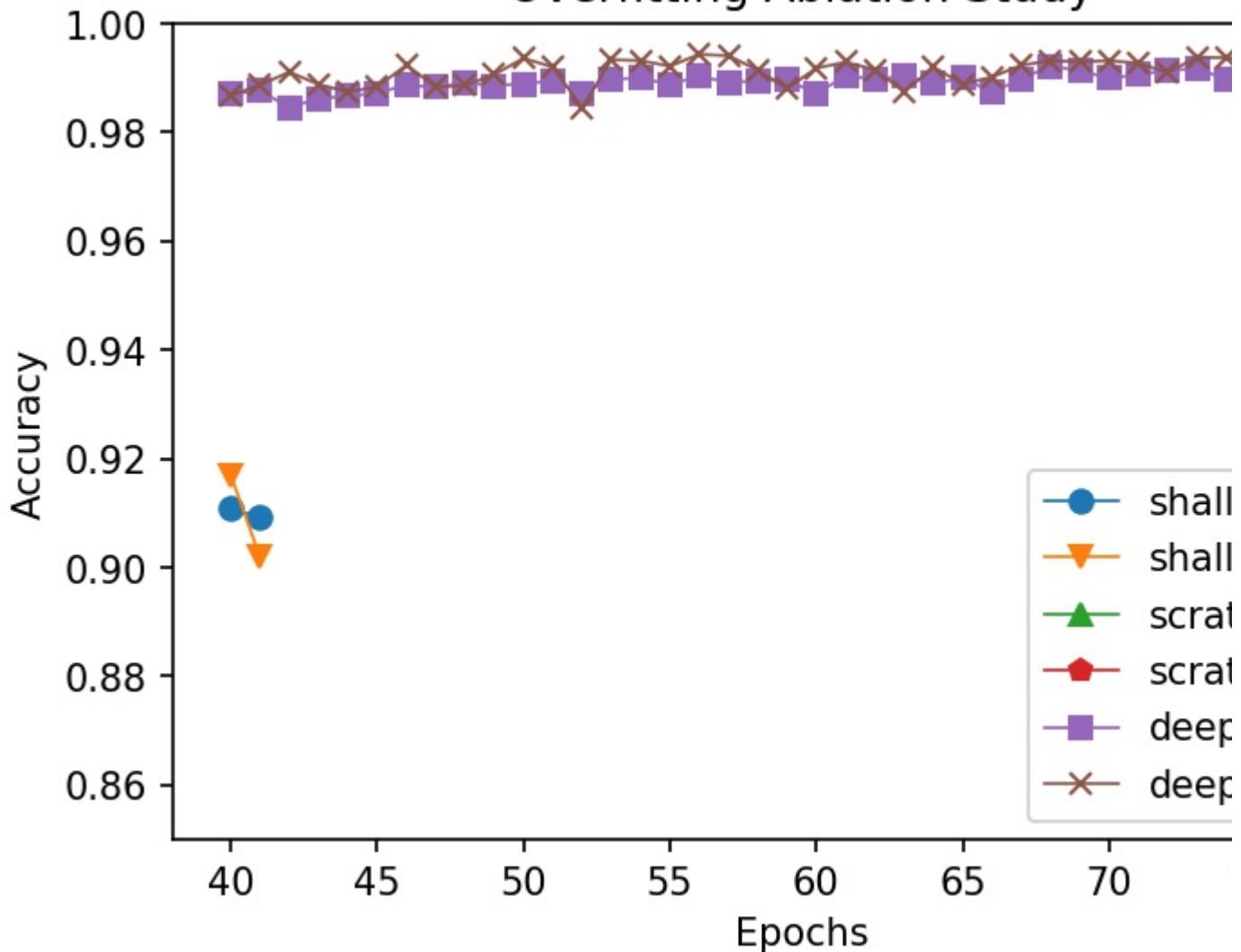
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	NaN	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 41 takes 152 seconds.  
[43,    30] loss: 0.446  
[43,    60] loss: 0.501  
[43,    90] loss: 0.397  
[43,   120] loss: 0.421  
[43,   150] loss: 0.493  
[43,   180] loss: 0.473  
[43,   210] loss: 0.435  
[43,   240] loss: 0.491  
[43,   270] loss: 0.461  
[43,   300] loss: 0.471  
[43,   330] loss: 0.504  
[43,   360] loss: 0.399  
[43,   390] loss: 0.552  
[43,   420] loss: 0.458  
[43,   450] loss: 0.398  
[43,   480] loss: 0.453  
[43,   510] loss: 0.413  
[43,   540] loss: 0.450  
[43,   570] loss: 0.380  
[43,   600] loss: 0.435  
[43,   630] loss: 0.421  
[43,   660] loss: 0.385
```

```

[43, 690] loss: 0.525
[43, 720] loss: 0.509
[43, 750] loss: 0.506
[43, 780] loss: 0.417
[43, 810] loss: 0.462
[43, 840] loss: 0.379
[43, 870] loss: 0.458
[43, 900] loss: 0.481
[43, 930] loss: 0.532
[43, 960] loss: 0.535
[43, 990] loss: 0.439
[43, 1020] loss: 0.521
[43, 1050] loss: 0.588
[43, 1080] loss: 0.521
[43, 1110] loss: 0.463
[43, 1140] loss: 0.417
[43, 1170] loss: 0.360
[43, 1200] loss: 0.402
[43, 1230] loss: 0.446
[43, 1260] loss: 0.469
[43, 1290] loss: 0.488
[43, 1320] loss: 0.386
[43, 1350] loss: 0.408
[43, 1380] loss: 0.539
[43, 1410] loss: 0.480
[43, 1440] loss: 0.485
[43, 1470] loss: 0.472
[43, 1500] loss: 0.486
[43, 1530] loss: 0.487
[43, 1560] loss: 0.455
[43, 1590] loss: 0.554
[43, 1620] loss: 0.541
[43, 1650] loss: 0.428
[43, 1680] loss: 0.533
[43, 1710] loss: 0.453
[43, 1740] loss: 0.437
[43, 1770] loss: 0.500
[43, 1800] loss: 0.482
[43, 1830] loss: 0.423
[43, 1860] loss: 0.444
[43, 1890] loss: 0.413
[43, 1920] loss: 0.478
[43, 1950] loss: 0.452
[43, 1980] loss: 0.470
[43, 2010] loss: 0.525
[43, 2040] loss: 0.534
[43, 2070] loss: 0.479
[43, 2100] loss: 0.499
[43, 2130] loss: 0.410
[43, 2160] loss: 0.489
[43, 2190] loss: 0.464
ovft_intermediate_models/epoch42_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.900945
epoch 42 accuracy on train set is: 0.900945110868775
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	NaN	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.910033

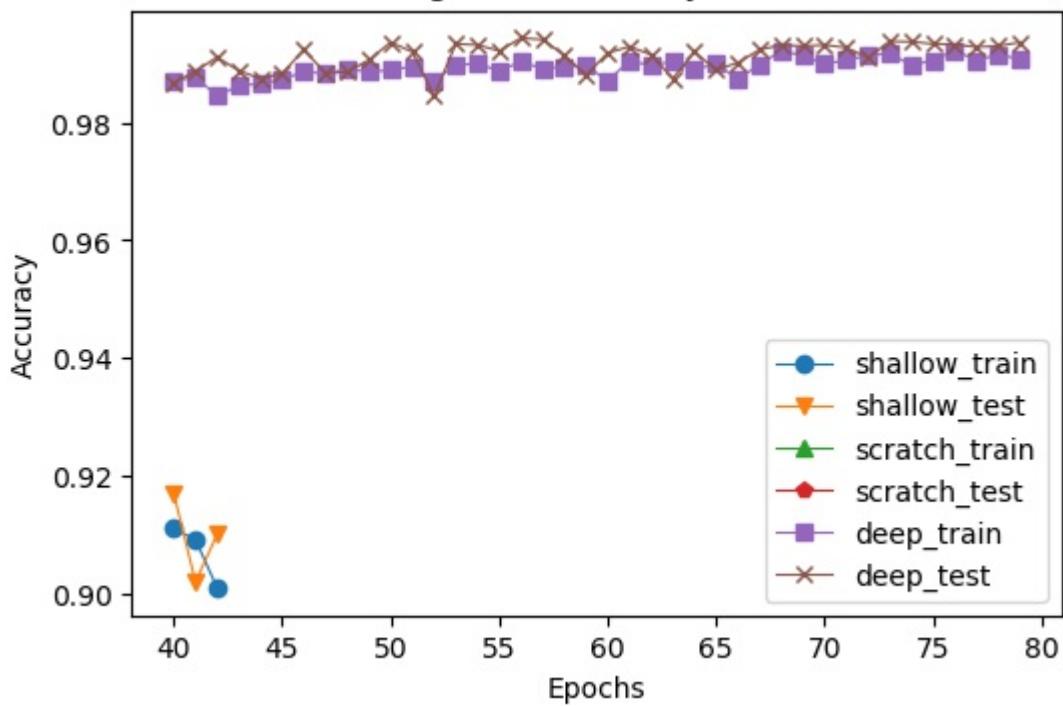
epoch 42 accuracy on test set is: 0.9100327153762269

evaluating on test set takes 16 seconds.

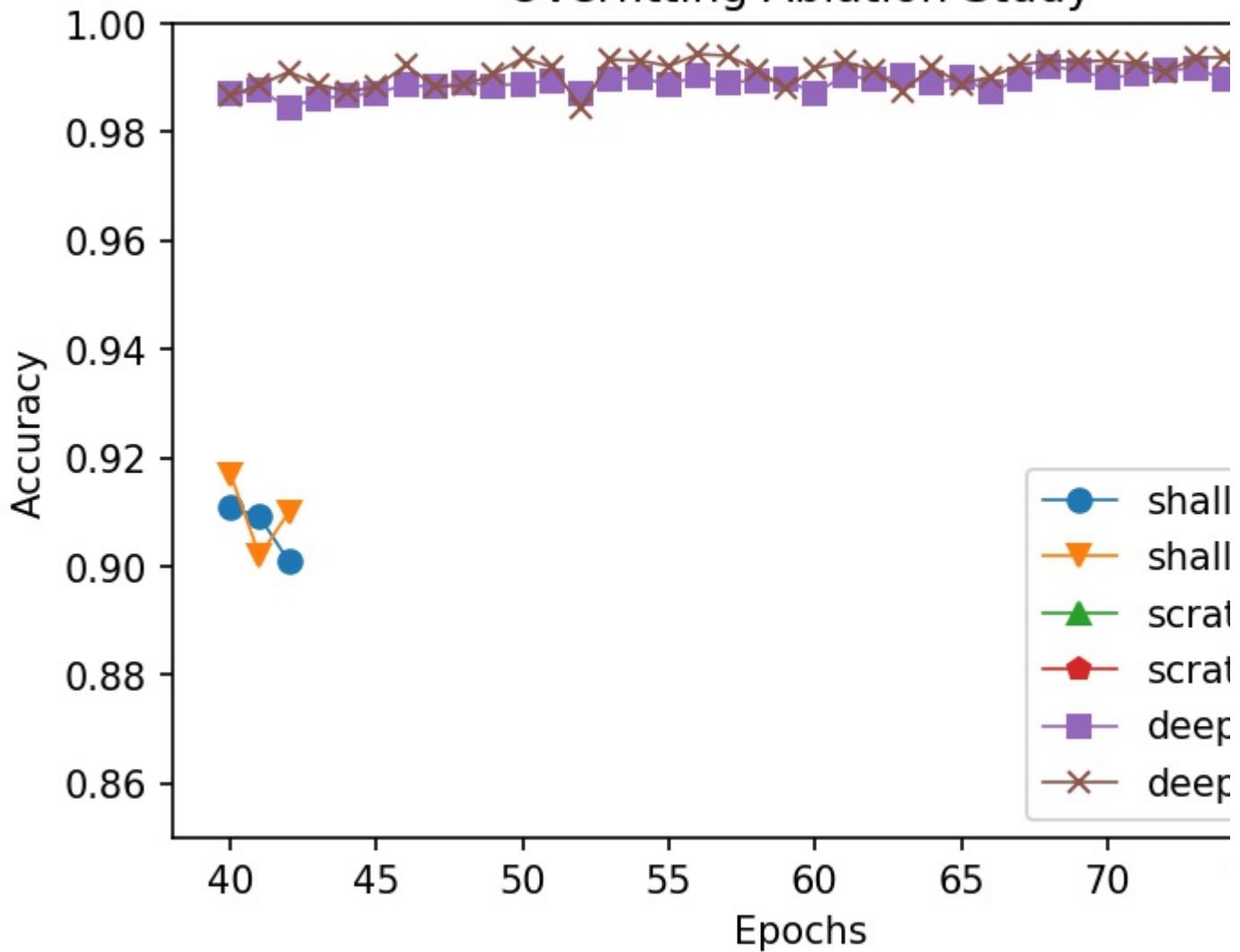
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	NaN	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 42 takes 149 seconds.  
[44,    30] loss: 0.532  
[44,    60] loss: 0.464  
[44,   90] loss: 0.437  
[44,  120] loss: 0.523  
[44,  150] loss: 0.399  
[44,  180] loss: 0.415  
[44,  210] loss: 0.413  
[44,  240] loss: 0.410  
[44,  270] loss: 0.453  
[44,  300] loss: 0.482  
[44,  330] loss: 0.395  
[44,  360] loss: 0.528  
[44,  390] loss: 0.480  
[44,  420] loss: 0.510  
[44,  450] loss: 0.517  
[44,  480] loss: 0.476  
[44,  510] loss: 0.522  
[44,  540] loss: 0.435  
[44,  570] loss: 0.488  
[44,  600] loss: 0.454  
[44,  630] loss: 0.520  
[44,  660] loss: 0.378
```

```

[44, 690] loss: 0.449
[44, 720] loss: 0.503
[44, 750] loss: 0.530
[44, 780] loss: 0.521
[44, 810] loss: 0.531
[44, 840] loss: 0.542
[44, 870] loss: 0.467
[44, 900] loss: 0.564
[44, 930] loss: 0.492
[44, 960] loss: 0.441
[44, 990] loss: 0.471
[44, 1020] loss: 0.391
[44, 1050] loss: 0.543
[44, 1080] loss: 0.486
[44, 1110] loss: 0.445
[44, 1140] loss: 0.455
[44, 1170] loss: 0.459
[44, 1200] loss: 0.406
[44, 1230] loss: 0.504
[44, 1260] loss: 0.450
[44, 1290] loss: 0.379
[44, 1320] loss: 0.517
[44, 1350] loss: 0.490
[44, 1380] loss: 0.388
[44, 1410] loss: 0.432
[44, 1440] loss: 0.396
[44, 1470] loss: 0.420
[44, 1500] loss: 0.469
[44, 1530] loss: 0.448
[44, 1560] loss: 0.412
[44, 1590] loss: 0.488
[44, 1620] loss: 0.484
[44, 1650] loss: 0.428
[44, 1680] loss: 0.458
[44, 1710] loss: 0.459
[44, 1740] loss: 0.478
[44, 1770] loss: 0.427
[44, 1800] loss: 0.406
[44, 1830] loss: 0.467
[44, 1860] loss: 0.584
[44, 1890] loss: 0.548
[44, 1920] loss: 0.391
[44, 1950] loss: 0.527
[44, 1980] loss: 0.428
[44, 2010] loss: 0.519
[44, 2040] loss: 0.432
[44, 2070] loss: 0.451
[44, 2100] loss: 0.461
[44, 2130] loss: 0.413
[44, 2160] loss: 0.528
[44, 2190] loss: 0.501
ovft_intermediate_models/epoch43_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.902763
epoch 43 accuracy on train set is: 0.9027626317702654
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	NaN	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.914940

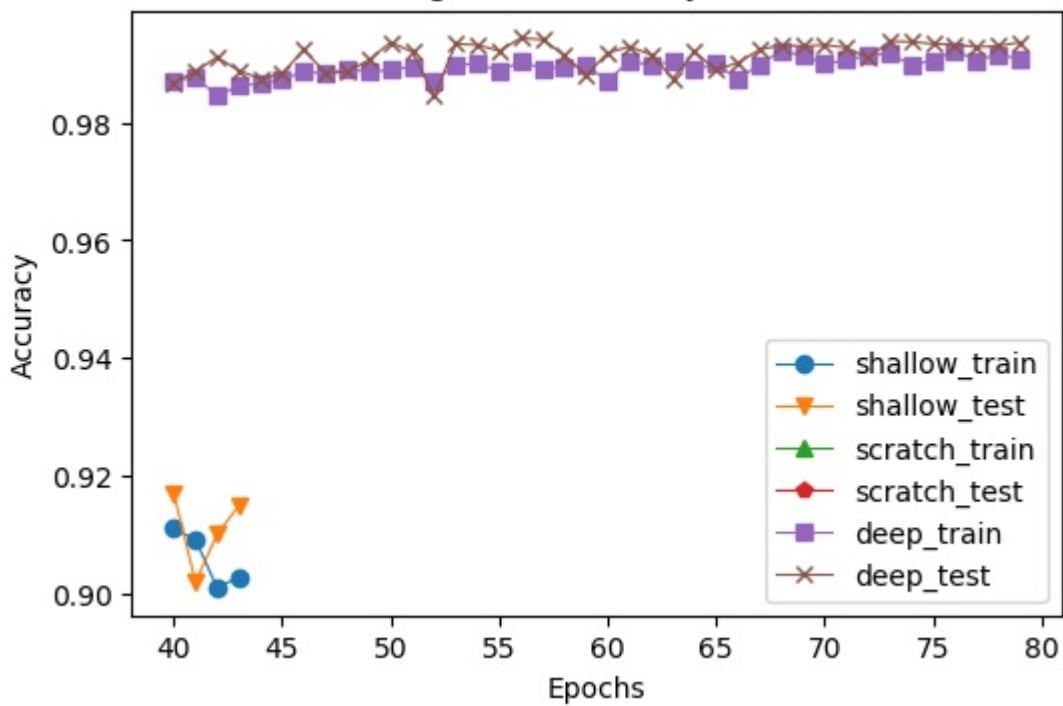
epoch 43 accuracy on test set is: 0.9149400218102508

evaluating on test set takes 16 seconds.

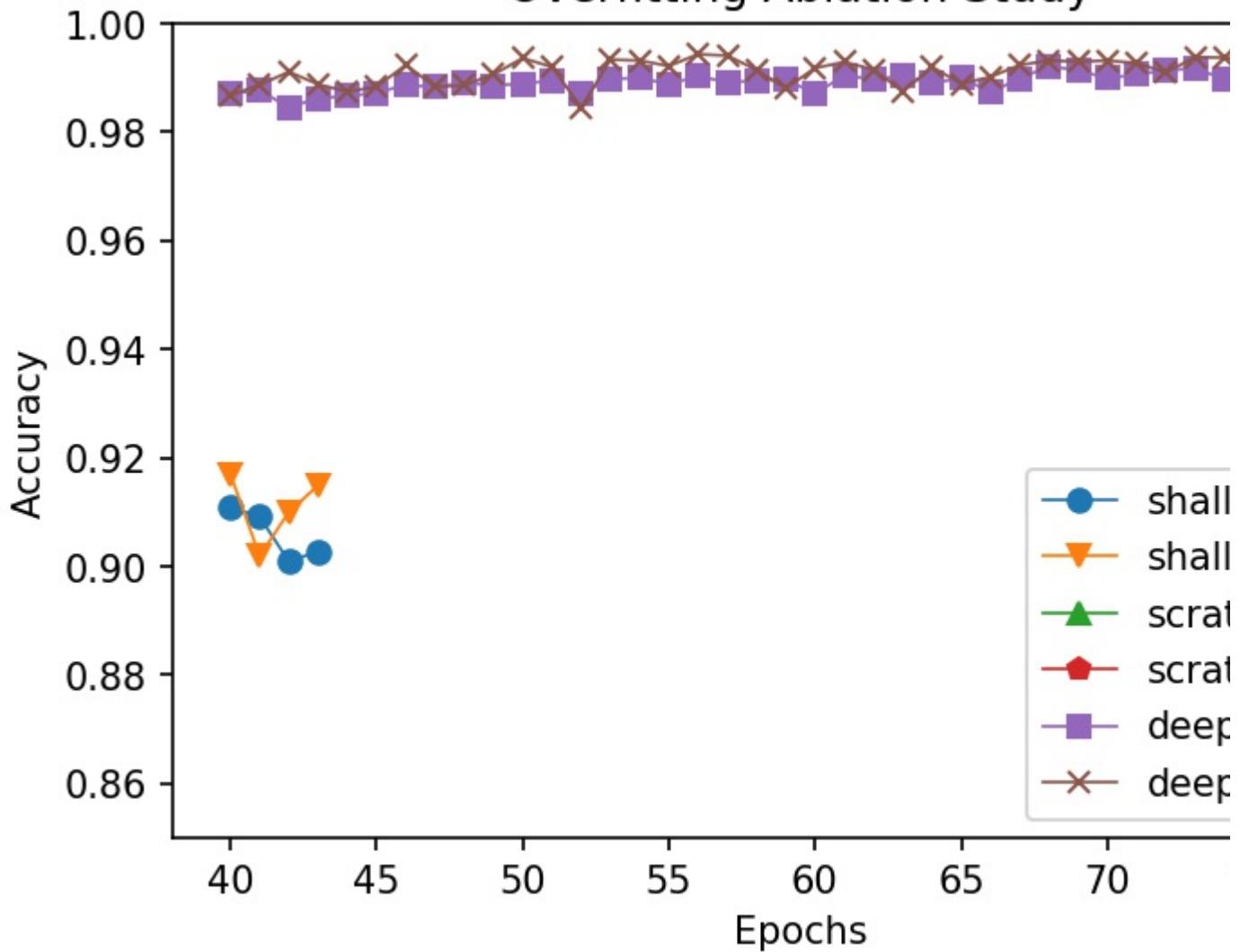
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	NaN	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 43 takes 149 seconds.  
[45,    30] loss: 0.442  
[45,    60] loss: 0.499  
[45,    90] loss: 0.499  
[45,   120] loss: 0.476  
[45,   150] loss: 0.459  
[45,   180] loss: 0.459  
[45,   210] loss: 0.522  
[45,   240] loss: 0.432  
[45,   270] loss: 0.476  
[45,   300] loss: 0.466  
[45,   330] loss: 0.532  
[45,   360] loss: 0.464  
[45,   390] loss: 0.524  
[45,   420] loss: 0.437  
[45,   450] loss: 0.529  
[45,   480] loss: 0.483  
[45,   510] loss: 0.540  
[45,   540] loss: 0.458  
[45,   570] loss: 0.542  
[45,   600] loss: 0.427  
[45,   630] loss: 0.522  
[45,   660] loss: 0.415
```

```

[45, 690] loss: 0.499
[45, 720] loss: 0.460
[45, 750] loss: 0.472
[45, 780] loss: 0.421
[45, 810] loss: 0.464
[45, 840] loss: 0.437
[45, 870] loss: 0.447
[45, 900] loss: 0.442
[45, 930] loss: 0.511
[45, 960] loss: 0.515
[45, 990] loss: 0.411
[45, 1020] loss: 0.505
[45, 1050] loss: 0.416
[45, 1080] loss: 0.423
[45, 1110] loss: 0.454
[45, 1140] loss: 0.465
[45, 1170] loss: 0.466
[45, 1200] loss: 0.475
[45, 1230] loss: 0.491
[45, 1260] loss: 0.434
[45, 1290] loss: 0.452
[45, 1320] loss: 0.479
[45, 1350] loss: 0.402
[45, 1380] loss: 0.547
[45, 1410] loss: 0.445
[45, 1440] loss: 0.470
[45, 1470] loss: 0.476
[45, 1500] loss: 0.537
[45, 1530] loss: 0.518
[45, 1560] loss: 0.434
[45, 1590] loss: 0.359
[45, 1620] loss: 0.483
[45, 1650] loss: 0.435
[45, 1680] loss: 0.438
[45, 1710] loss: 0.500
[45, 1740] loss: 0.445
[45, 1770] loss: 0.440
[45, 1800] loss: 0.426
[45, 1830] loss: 0.483
[45, 1860] loss: 0.496
[45, 1890] loss: 0.370
[45, 1920] loss: 0.418
[45, 1950] loss: 0.479
[45, 1980] loss: 0.509
[45, 2010] loss: 0.405
[45, 2040] loss: 0.389
[45, 2070] loss: 0.410
[45, 2100] loss: 0.504
[45, 2130] loss: 0.486
[45, 2160] loss: 0.499
[45, 2190] loss: 0.440
ovft_intermediate_models/epoch44_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.907556
epoch 44 accuracy on train set is: 0.9075563431479462
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	NaN	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.918121

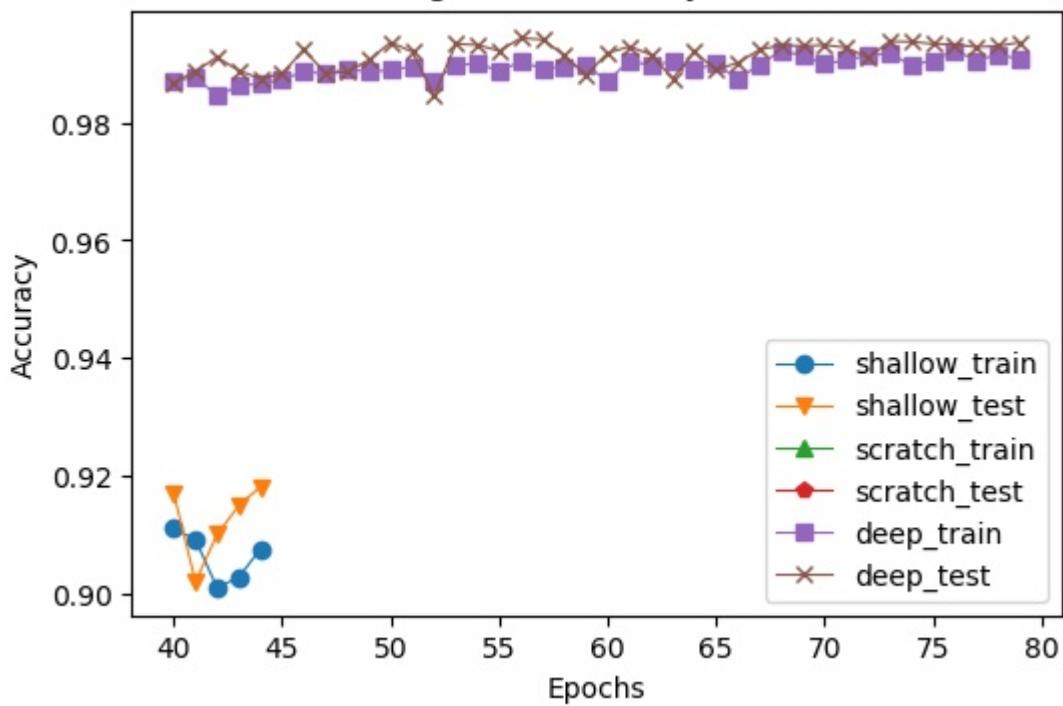
epoch 44 accuracy on test set is: 0.918120683387859

evaluating on test set takes 16 seconds.

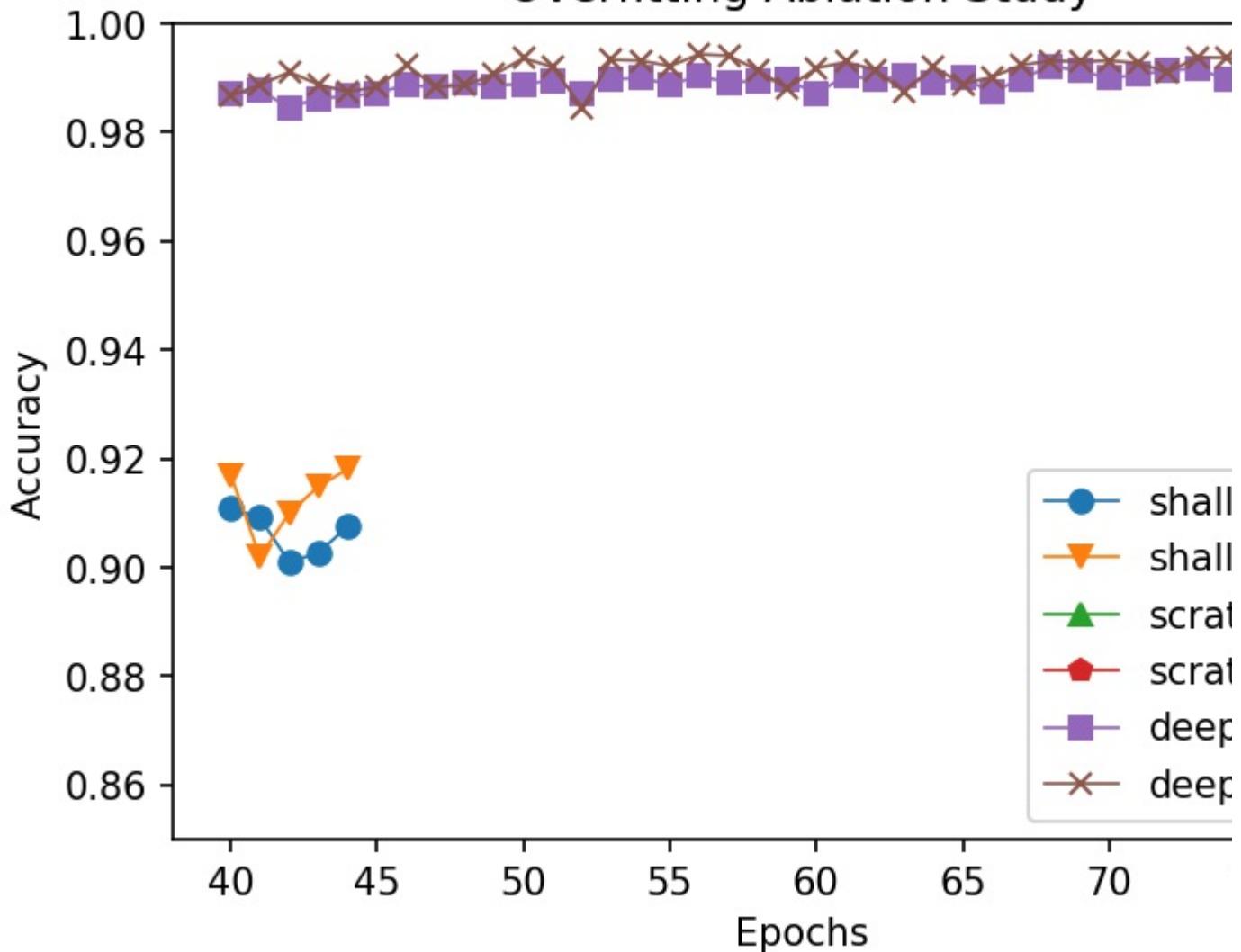
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	NaN	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 44 takes 149 seconds.  
[46,    30] loss: 0.523  
[46,    60] loss: 0.398  
[46,    90] loss: 0.543  
[46,   120] loss: 0.483  
[46,   150] loss: 0.446  
[46,   180] loss: 0.444  
[46,   210] loss: 0.423  
[46,   240] loss: 0.484  
[46,   270] loss: 0.502  
[46,   300] loss: 0.487  
[46,   330] loss: 0.483  
[46,   360] loss: 0.433  
[46,   390] loss: 0.417  
[46,   420] loss: 0.511  
[46,   450] loss: 0.481  
[46,   480] loss: 0.399  
[46,   510] loss: 0.439  
[46,   540] loss: 0.531  
[46,   570] loss: 0.424  
[46,   600] loss: 0.459  
[46,   630] loss: 0.437  
[46,   660] loss: 0.337
```

```

[46, 690] loss: 0.499
[46, 720] loss: 0.491
[46, 750] loss: 0.434
[46, 780] loss: 0.494
[46, 810] loss: 0.498
[46, 840] loss: 0.475
[46, 870] loss: 0.367
[46, 900] loss: 0.510
[46, 930] loss: 0.489
[46, 960] loss: 0.443
[46, 990] loss: 0.435
[46, 1020] loss: 0.388
[46, 1050] loss: 0.458
[46, 1080] loss: 0.361
[46, 1110] loss: 0.469
[46, 1140] loss: 0.418
[46, 1170] loss: 0.490
[46, 1200] loss: 0.436
[46, 1230] loss: 0.475
[46, 1260] loss: 0.583
[46, 1290] loss: 0.453
[46, 1320] loss: 0.443
[46, 1350] loss: 0.457
[46, 1380] loss: 0.515
[46, 1410] loss: 0.442
[46, 1440] loss: 0.409
[46, 1470] loss: 0.414
[46, 1500] loss: 0.453
[46, 1530] loss: 0.420
[46, 1560] loss: 0.536
[46, 1590] loss: 0.469
[46, 1620] loss: 0.425
[46, 1650] loss: 0.450
[46, 1680] loss: 0.519
[46, 1710] loss: 0.374
[46, 1740] loss: 0.503
[46, 1770] loss: 0.468
[46, 1800] loss: 0.493
[46, 1830] loss: 0.478
[46, 1860] loss: 0.449
[46, 1890] loss: 0.458
[46, 1920] loss: 0.490
[46, 1950] loss: 0.513
[46, 1980] loss: 0.350
[46, 2010] loss: 0.481
[46, 2040] loss: 0.533
[46, 2070] loss: 0.426
[46, 2100] loss: 0.506
[46, 2130] loss: 0.440
[46, 2160] loss: 0.512
[46, 2190] loss: 0.415
ovft_intermediate_models/epoch45_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.907511
epoch 45 accuracy on train set is: 0.907510905125409
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	NaN	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.922664

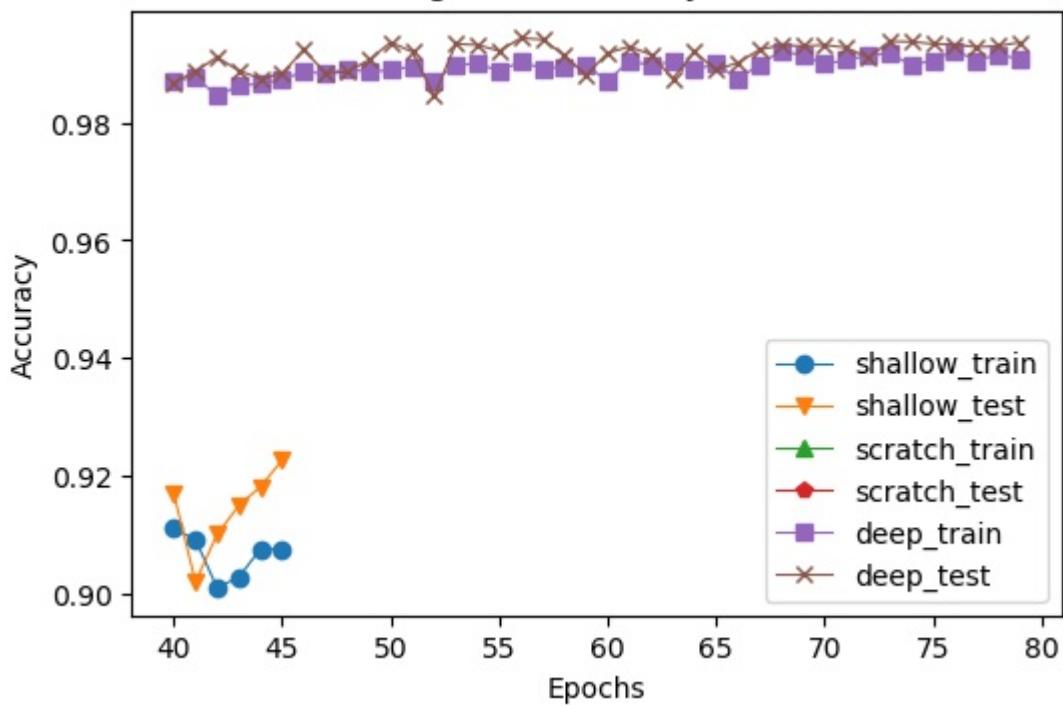
epoch 45 accuracy on test set is: 0.9226644856415849

evaluating on test set takes 16 seconds.

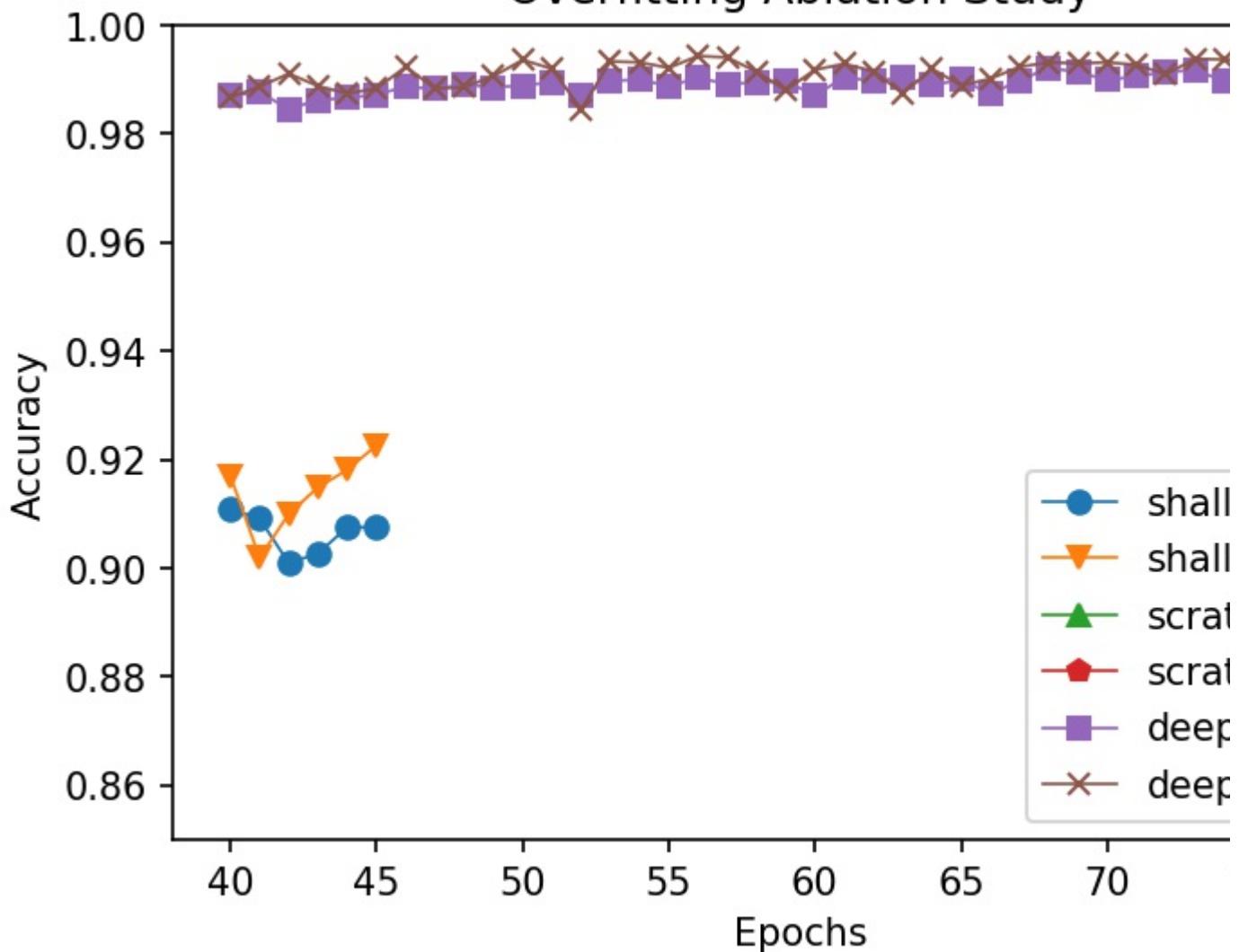
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	NaN	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 45 takes 149 seconds.  
[47,     30] loss: 0.463  
[47,     60] loss: 0.478  
[47,    90] loss: 0.488  
[47,   120] loss: 0.461  
[47,   150] loss: 0.431  
[47,   180] loss: 0.434  
[47,   210] loss: 0.491  
[47,   240] loss: 0.492  
[47,   270] loss: 0.381  
[47,   300] loss: 0.495  
[47,   330] loss: 0.520  
[47,   360] loss: 0.441  
[47,   390] loss: 0.354  
[47,   420] loss: 0.428  
[47,   450] loss: 0.511  
[47,   480] loss: 0.392  
[47,   510] loss: 0.521  
[47,   540] loss: 0.508  
[47,   570] loss: 0.453  
[47,   600] loss: 0.516  
[47,   630] loss: 0.506  
[47,   660] loss: 0.466
```

```

[47, 690] loss: 0.432
[47, 720] loss: 0.395
[47, 750] loss: 0.469
[47, 780] loss: 0.430
[47, 810] loss: 0.459
[47, 840] loss: 0.477
[47, 870] loss: 0.393
[47, 900] loss: 0.348
[47, 930] loss: 0.405
[47, 960] loss: 0.556
[47, 990] loss: 0.471
[47, 1020] loss: 0.495
[47, 1050] loss: 0.497
[47, 1080] loss: 0.551
[47, 1110] loss: 0.490
[47, 1140] loss: 0.497
[47, 1170] loss: 0.513
[47, 1200] loss: 0.416
[47, 1230] loss: 0.456
[47, 1260] loss: 0.450
[47, 1290] loss: 0.501
[47, 1320] loss: 0.444
[47, 1350] loss: 0.497
[47, 1380] loss: 0.436
[47, 1410] loss: 0.451
[47, 1440] loss: 0.468
[47, 1470] loss: 0.473
[47, 1500] loss: 0.555
[47, 1530] loss: 0.467
[47, 1560] loss: 0.482
[47, 1590] loss: 0.489
[47, 1620] loss: 0.501
[47, 1650] loss: 0.464
[47, 1680] loss: 0.499
[47, 1710] loss: 0.470
[47, 1740] loss: 0.404
[47, 1770] loss: 0.412
[47, 1800] loss: 0.507
[47, 1830] loss: 0.479
[47, 1860] loss: 0.425
[47, 1890] loss: 0.475
[47, 1920] loss: 0.509
[47, 1950] loss: 0.496
[47, 1980] loss: 0.443
[47, 2010] loss: 0.481
[47, 2040] loss: 0.439
[47, 2070] loss: 0.504
[47, 2100] loss: 0.463
[47, 2130] loss: 0.457
[47, 2160] loss: 0.492
[47, 2190] loss: 0.502
ovft_intermediate_models/epoch46_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.910533
epoch 46 accuracy on train set is: 0.9105325336241367
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027  0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	NaN	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.933206

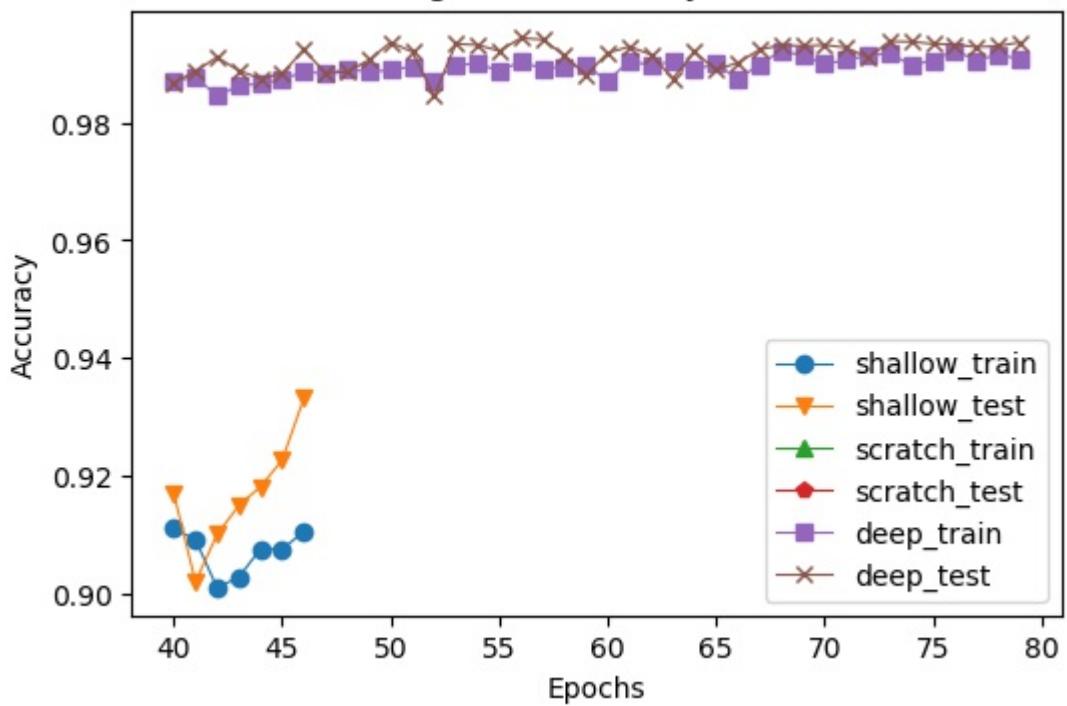
epoch 46 accuracy on test set is: 0.933206106870229

evaluating on test set takes 16 seconds.

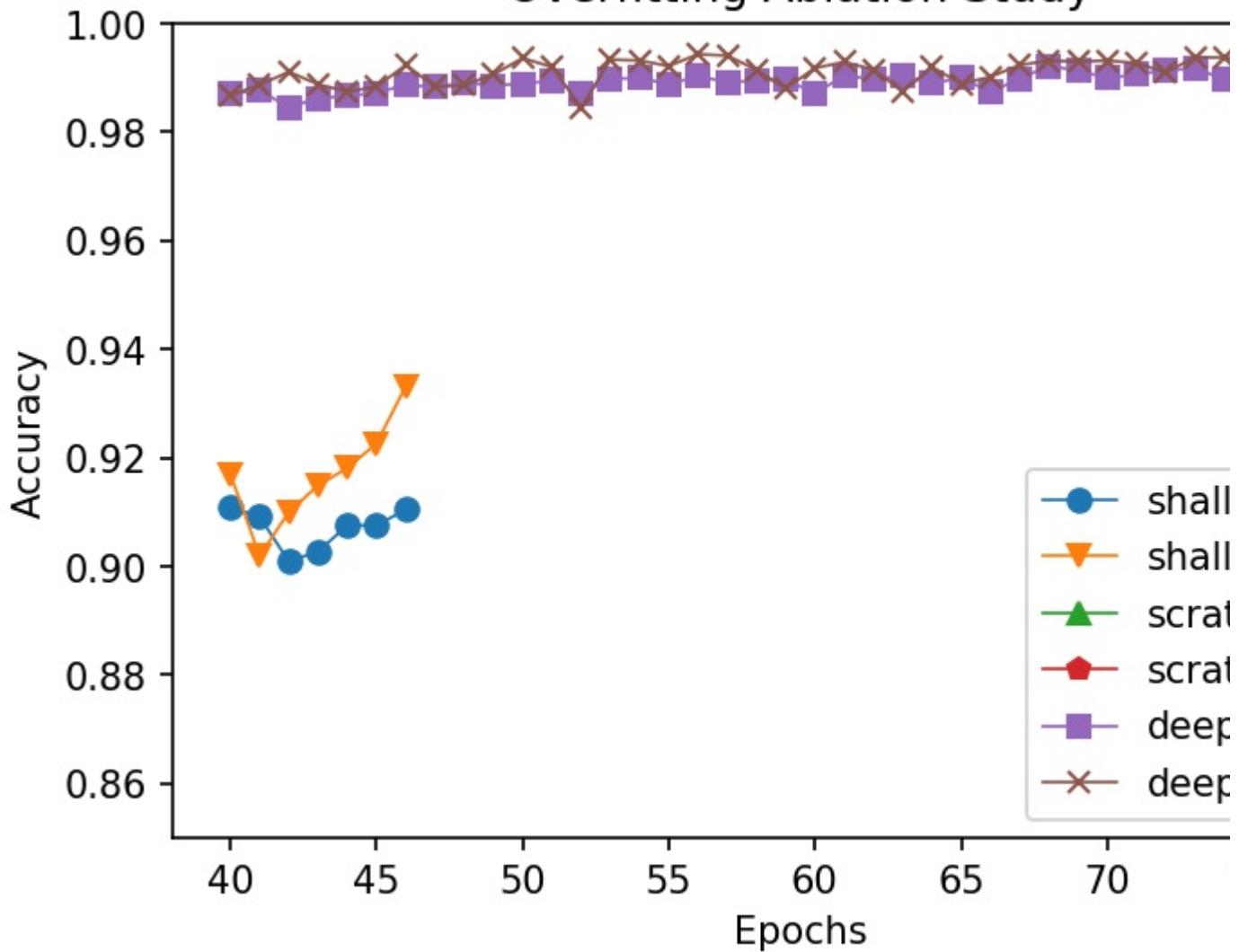
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	NaN	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 46 takes 149 seconds.  
[48,    30] loss: 0.467  
[48,    60] loss: 0.411  
[48,    90] loss: 0.480  
[48,   120] loss: 0.478  
[48,   150] loss: 0.518  
[48,   180] loss: 0.495  
[48,   210] loss: 0.435  
[48,   240] loss: 0.431  
[48,   270] loss: 0.509  
[48,   300] loss: 0.431  
[48,   330] loss: 0.432  
[48,   360] loss: 0.560  
[48,   390] loss: 0.500  
[48,   420] loss: 0.561  
[48,   450] loss: 0.419  
[48,   480] loss: 0.456  
[48,   510] loss: 0.481  
[48,   540] loss: 0.530  
[48,   570] loss: 0.401  
[48,   600] loss: 0.523  
[48,   630] loss: 0.434  
[48,   660] loss: 0.445
```

```

[48,   690] loss: 0.424
[48,   720] loss: 0.581
[48,   750] loss: 0.477
[48,   780] loss: 0.436
[48,   810] loss: 0.479
[48,   840] loss: 0.333
[48,   870] loss: 0.492
[48,   900] loss: 0.416
[48,   930] loss: 0.479
[48,   960] loss: 0.382
[48,   990] loss: 0.509
[48,  1020] loss: 0.453
[48,  1050] loss: 0.450
[48,  1080] loss: 0.418
[48,  1110] loss: 0.432
[48,  1140] loss: 0.546
[48,  1170] loss: 0.509
[48,  1200] loss: 0.451
[48,  1230] loss: 0.545
[48,  1260] loss: 0.532
[48,  1290] loss: 0.470
[48,  1320] loss: 0.423
[48,  1350] loss: 0.398
[48,  1380] loss: 0.441
[48,  1410] loss: 0.484
[48,  1440] loss: 0.393
[48,  1470] loss: 0.470
[48,  1500] loss: 0.507
[48,  1530] loss: 0.445
[48,  1560] loss: 0.517
[48,  1590] loss: 0.479
[48,  1620] loss: 0.505
[48,  1650] loss: 0.492
[48,  1680] loss: 0.426
[48,  1710] loss: 0.461
[48,  1740] loss: 0.473
[48,  1770] loss: 0.481
[48,  1800] loss: 0.457
[48,  1830] loss: 0.454
[48,  1860] loss: 0.405
[48,  1890] loss: 0.470
[48,  1920] loss: 0.501
[48,  1950] loss: 0.475
[48,  1980] loss: 0.478
[48,  2010] loss: 0.533
[48,  2040] loss: 0.390
[48,  2070] loss: 0.488
[48,  2100] loss: 0.414
[48,  2130] loss: 0.460
[48,  2160] loss: 0.466
[48,  2190] loss: 0.434
ovft_intermediate_models/epoch47_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.906466
epoch 47 accuracy on train set is: 0.906465830607052
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	NaN	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.917666

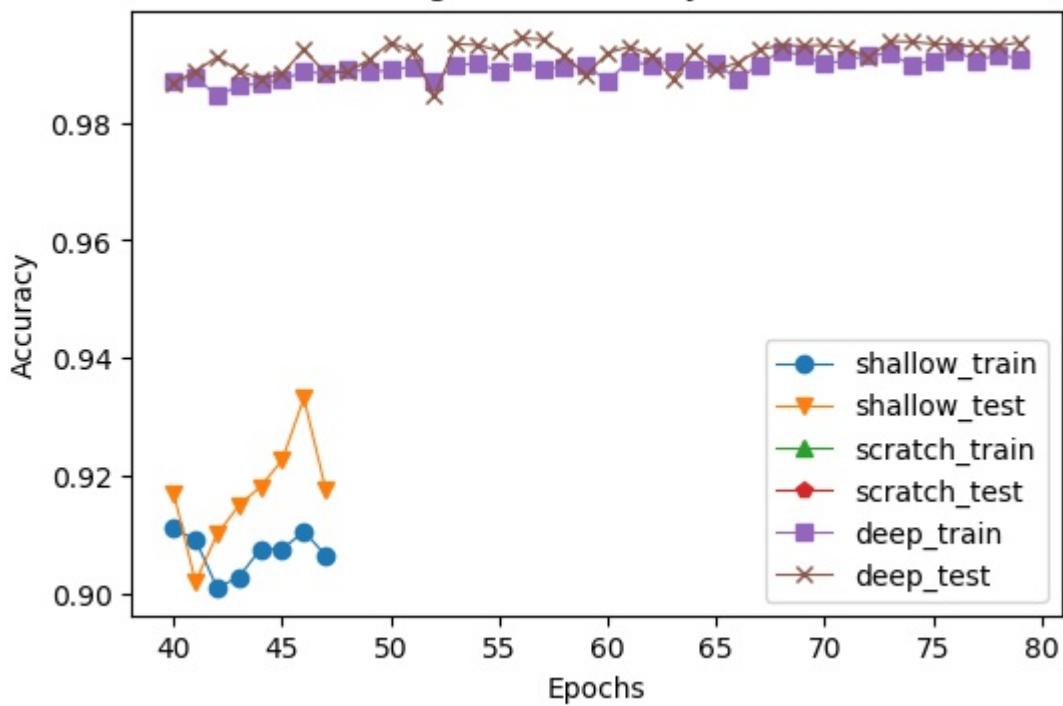
epoch 47 accuracy on test set is: 0.9176663031624863

evaluating on test set takes 16 seconds.

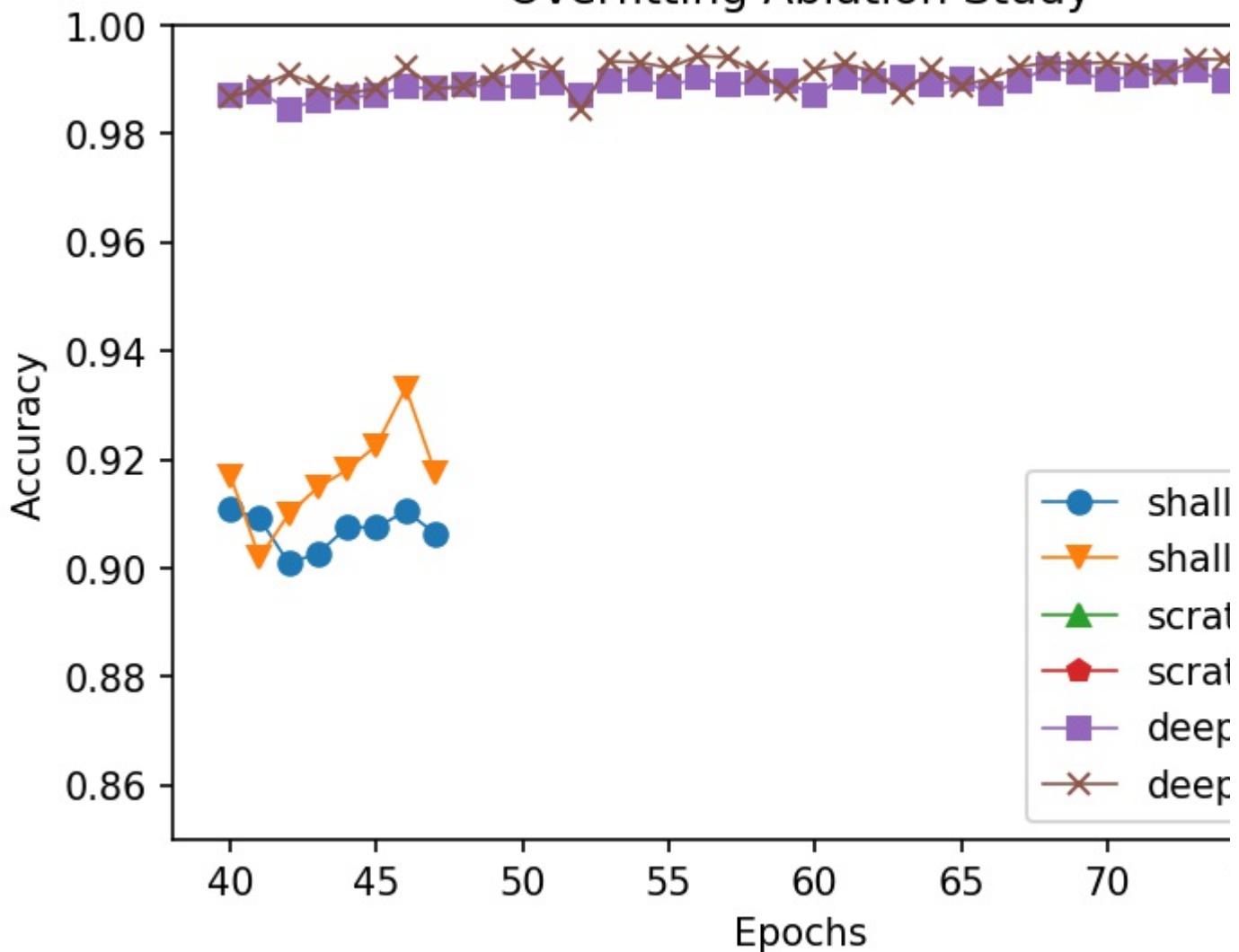
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	NaN	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 47 takes 149 seconds.  
[49,     30] loss: 0.445  
[49,     60] loss: 0.455  
[49,    90] loss: 0.519  
[49,   120] loss: 0.475  
[49,   150] loss: 0.442  
[49,   180] loss: 0.434  
[49,   210] loss: 0.381  
[49,   240] loss: 0.436  
[49,   270] loss: 0.430  
[49,   300] loss: 0.496  
[49,   330] loss: 0.413  
[49,   360] loss: 0.448  
[49,   390] loss: 0.478  
[49,   420] loss: 0.413  
[49,   450] loss: 0.454  
[49,   480] loss: 0.521  
[49,   510] loss: 0.556  
[49,   540] loss: 0.482  
[49,   570] loss: 0.503  
[49,   600] loss: 0.334  
[49,   630] loss: 0.426  
[49,   660] loss: 0.468
```

```

[49, 690] loss: 0.431
[49, 720] loss: 0.512
[49, 750] loss: 0.473
[49, 780] loss: 0.505
[49, 810] loss: 0.460
[49, 840] loss: 0.413
[49, 870] loss: 0.404
[49, 900] loss: 0.432
[49, 930] loss: 0.420
[49, 960] loss: 0.432
[49, 990] loss: 0.454
[49, 1020] loss: 0.455
[49, 1050] loss: 0.466
[49, 1080] loss: 0.441
[49, 1110] loss: 0.374
[49, 1140] loss: 0.489
[49, 1170] loss: 0.539
[49, 1200] loss: 0.486
[49, 1230] loss: 0.432
[49, 1260] loss: 0.432
[49, 1290] loss: 0.430
[49, 1320] loss: 0.484
[49, 1350] loss: 0.417
[49, 1380] loss: 0.460
[49, 1410] loss: 0.461
[49, 1440] loss: 0.438
[49, 1470] loss: 0.420
[49, 1500] loss: 0.483
[49, 1530] loss: 0.443
[49, 1560] loss: 0.424
[49, 1590] loss: 0.403
[49, 1620] loss: 0.496
[49, 1650] loss: 0.454
[49, 1680] loss: 0.446
[49, 1710] loss: 0.452
[49, 1740] loss: 0.502
[49, 1770] loss: 0.411
[49, 1800] loss: 0.499
[49, 1830] loss: 0.461
[49, 1860] loss: 0.407
[49, 1890] loss: 0.387
[49, 1920] loss: 0.465
[49, 1950] loss: 0.472
[49, 1980] loss: 0.532
[49, 2010] loss: 0.455
[49, 2040] loss: 0.463
[49, 2070] loss: 0.409
[49, 2100] loss: 0.460
[49, 2130] loss: 0.450
[49, 2160] loss: 0.611
[49, 2190] loss: 0.384
ovft_intermediate_models/epoch48_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.911055
epoch 48 accuracy on train set is: 0.9110550708833152
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	NaN	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.923210

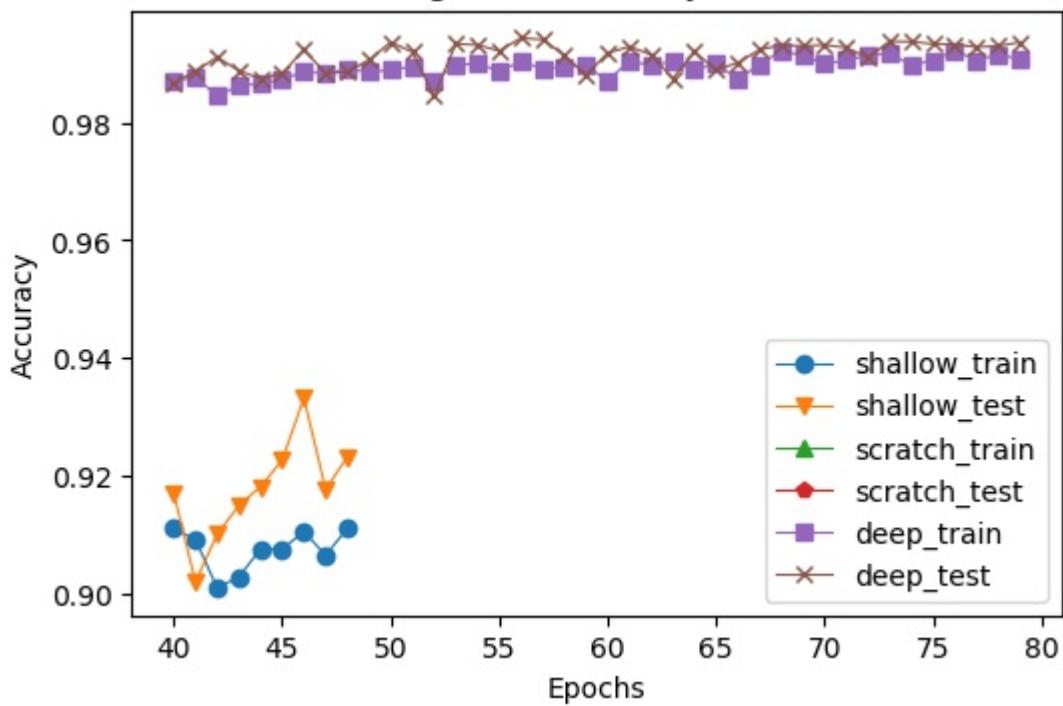
epoch 48 accuracy on test set is: 0.923209741912032

evaluating on test set takes 16 seconds.

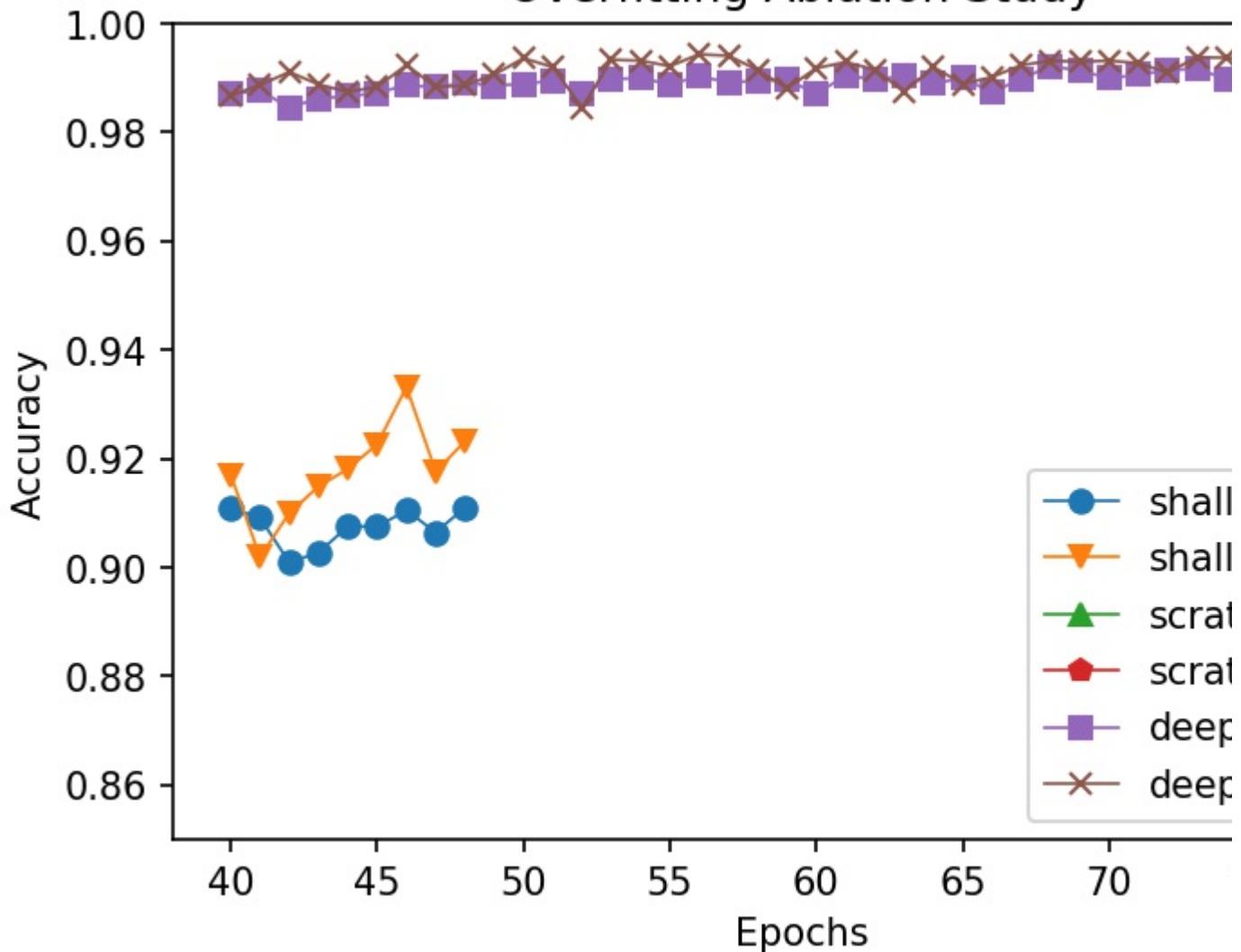
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	NaN	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 48 takes 150 seconds.  
[50,    30] loss: 0.444  
[50,    60] loss: 0.494  
[50,    90] loss: 0.423  
[50,   120] loss: 0.476  
[50,   150] loss: 0.384  
[50,   180] loss: 0.461  
[50,   210] loss: 0.524  
[50,   240] loss: 0.562  
[50,   270] loss: 0.473  
[50,   300] loss: 0.428  
[50,   330] loss: 0.422  
[50,   360] loss: 0.461  
[50,   390] loss: 0.481  
[50,   420] loss: 0.481  
[50,   450] loss: 0.436  
[50,   480] loss: 0.509  
[50,   510] loss: 0.516  
[50,   540] loss: 0.518  
[50,   570] loss: 0.454  
[50,   600] loss: 0.504  
[50,   630] loss: 0.481  
[50,   660] loss: 0.384
```

```

[50, 690] loss: 0.537
[50, 720] loss: 0.422
[50, 750] loss: 0.425
[50, 780] loss: 0.442
[50, 810] loss: 0.468
[50, 840] loss: 0.501
[50, 870] loss: 0.448
[50, 900] loss: 0.466
[50, 930] loss: 0.496
[50, 960] loss: 0.584
[50, 990] loss: 0.426
[50, 1020] loss: 0.522
[50, 1050] loss: 0.505
[50, 1080] loss: 0.451
[50, 1110] loss: 0.508
[50, 1140] loss: 0.531
[50, 1170] loss: 0.429
[50, 1200] loss: 0.467
[50, 1230] loss: 0.559
[50, 1260] loss: 0.424
[50, 1290] loss: 0.432
[50, 1320] loss: 0.488
[50, 1350] loss: 0.448
[50, 1380] loss: 0.465
[50, 1410] loss: 0.471
[50, 1440] loss: 0.514
[50, 1470] loss: 0.513
[50, 1500] loss: 0.497
[50, 1530] loss: 0.438
[50, 1560] loss: 0.391
[50, 1590] loss: 0.437
[50, 1620] loss: 0.391
[50, 1650] loss: 0.463
[50, 1680] loss: 0.428
[50, 1710] loss: 0.479
[50, 1740] loss: 0.445
[50, 1770] loss: 0.594
[50, 1800] loss: 0.551
[50, 1830] loss: 0.398
[50, 1860] loss: 0.418
[50, 1890] loss: 0.401
[50, 1920] loss: 0.490
[50, 1950] loss: 0.520
[50, 1980] loss: 0.511
[50, 2010] loss: 0.495
[50, 2040] loss: 0.418
[50, 2070] loss: 0.526
[50, 2100] loss: 0.551
[50, 2130] loss: 0.460
[50, 2160] loss: 0.452
[50, 2190] loss: 0.530
ovft_intermediate_models/epoch49_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.911032
epoch 49 accuracy on train set is: 0.9110323518720466
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	NaN	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.923482

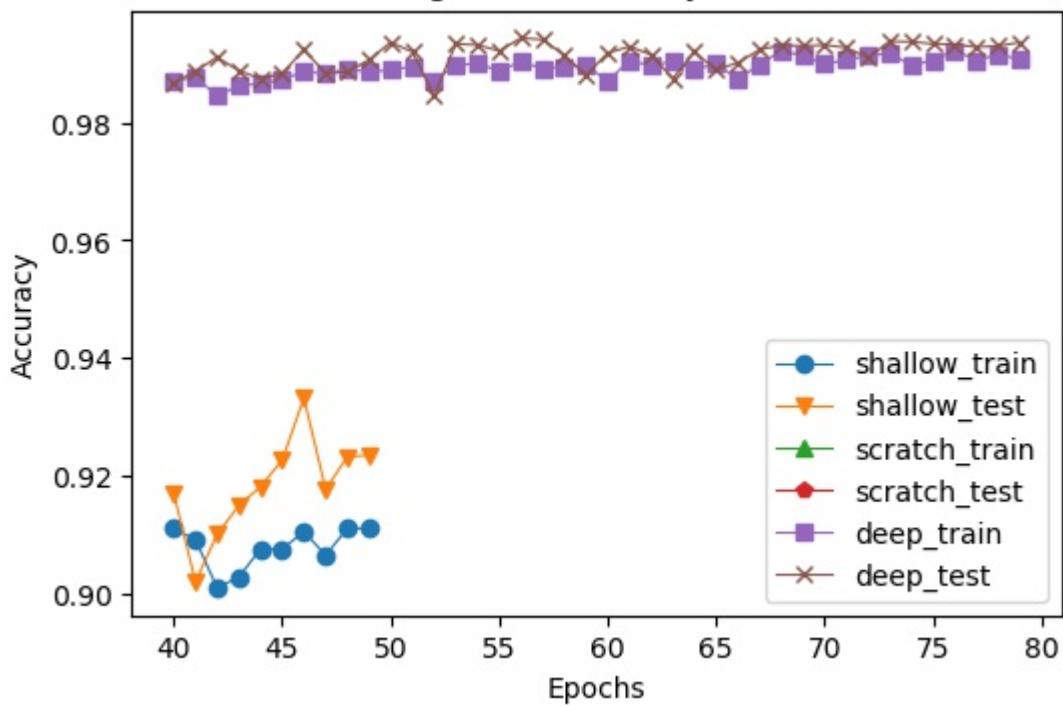
epoch 49 accuracy on test set is: 0.9234823700472555

evaluating on test set takes 16 seconds.

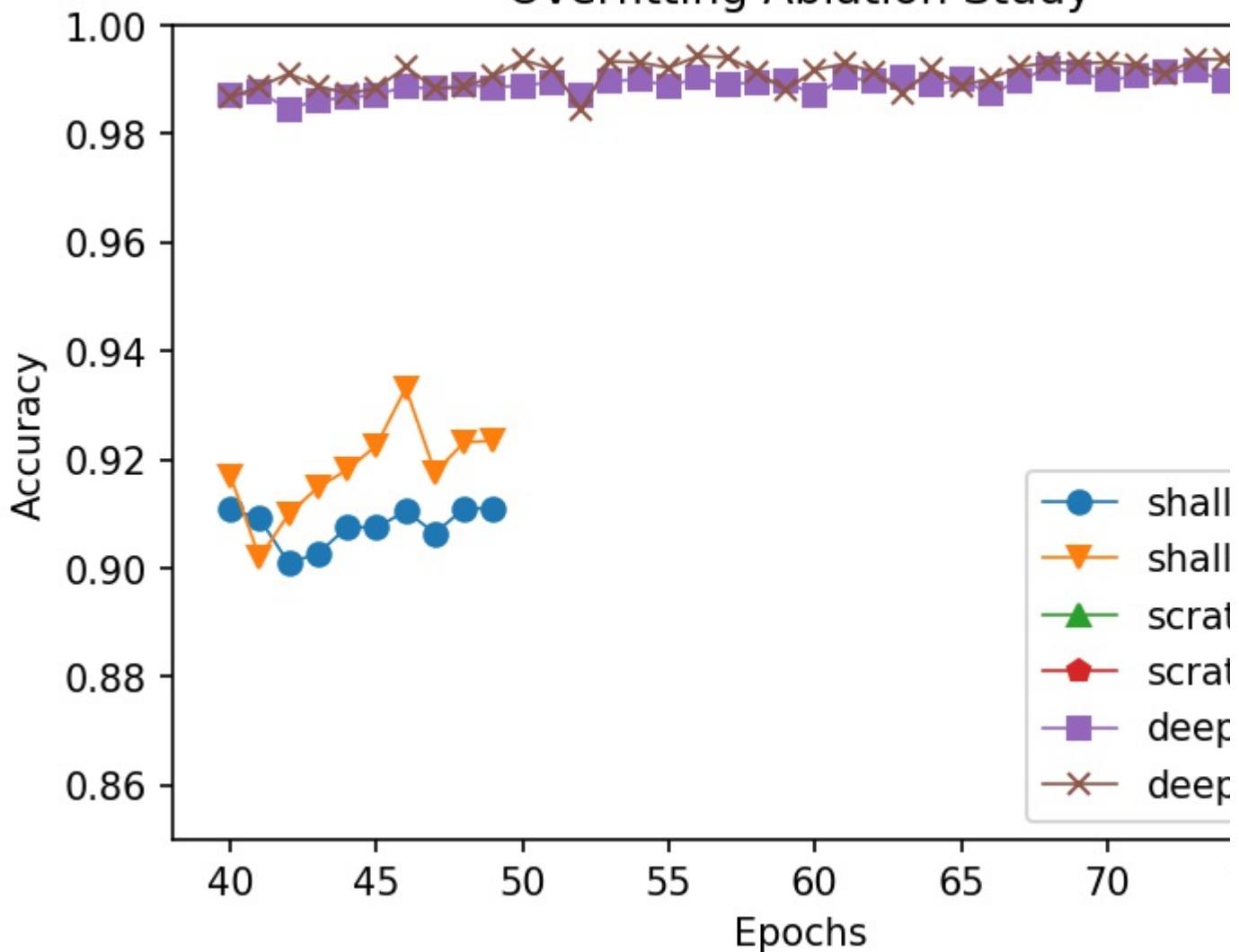
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	NaN	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 49 takes 149 seconds.  
[51,    30] loss: 0.572  
[51,    60] loss: 0.468  
[51,    90] loss: 0.433  
[51,   120] loss: 0.460  
[51,   150] loss: 0.480  
[51,   180] loss: 0.423  
[51,   210] loss: 0.386  
[51,   240] loss: 0.400  
[51,   270] loss: 0.516  
[51,   300] loss: 0.465  
[51,   330] loss: 0.461  
[51,   360] loss: 0.476  
[51,   390] loss: 0.518  
[51,   420] loss: 0.502  
[51,   450] loss: 0.474  
[51,   480] loss: 0.478  
[51,   510] loss: 0.442  
[51,   540] loss: 0.426  
[51,   570] loss: 0.503  
[51,   600] loss: 0.489  
[51,   630] loss: 0.498  
[51,   660] loss: 0.455
```

```

[51, 690] loss: 0.491
[51, 720] loss: 0.428
[51, 750] loss: 0.487
[51, 780] loss: 0.470
[51, 810] loss: 0.466
[51, 840] loss: 0.454
[51, 870] loss: 0.458
[51, 900] loss: 0.457
[51, 930] loss: 0.406
[51, 960] loss: 0.523
[51, 990] loss: 0.544
[51, 1020] loss: 0.489
[51, 1050] loss: 0.495
[51, 1080] loss: 0.430
[51, 1110] loss: 0.505
[51, 1140] loss: 0.495
[51, 1170] loss: 0.465
[51, 1200] loss: 0.492
[51, 1230] loss: 0.474
[51, 1260] loss: 0.515
[51, 1290] loss: 0.420
[51, 1320] loss: 0.382
[51, 1350] loss: 0.418
[51, 1380] loss: 0.425
[51, 1410] loss: 0.451
[51, 1440] loss: 0.458
[51, 1470] loss: 0.527
[51, 1500] loss: 0.436
[51, 1530] loss: 0.481
[51, 1560] loss: 0.511
[51, 1590] loss: 0.508
[51, 1620] loss: 0.496
[51, 1650] loss: 0.484
[51, 1680] loss: 0.456
[51, 1710] loss: 0.578
[51, 1740] loss: 0.512
[51, 1770] loss: 0.485
[51, 1800] loss: 0.448
[51, 1830] loss: 0.458
[51, 1860] loss: 0.542
[51, 1890] loss: 0.494
[51, 1920] loss: 0.374
[51, 1950] loss: 0.507
[51, 1980] loss: 0.425
[51, 2010] loss: 0.455
[51, 2040] loss: 0.386
[51, 2070] loss: 0.532
[51, 2100] loss: 0.514
[51, 2130] loss: 0.484
[51, 2160] loss: 0.471
[51, 2190] loss: 0.547
ovft_intermediate_models/epoch50_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.915463
epoch 50 accuracy on train set is: 0.9154625590694293
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	NaN	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.932933

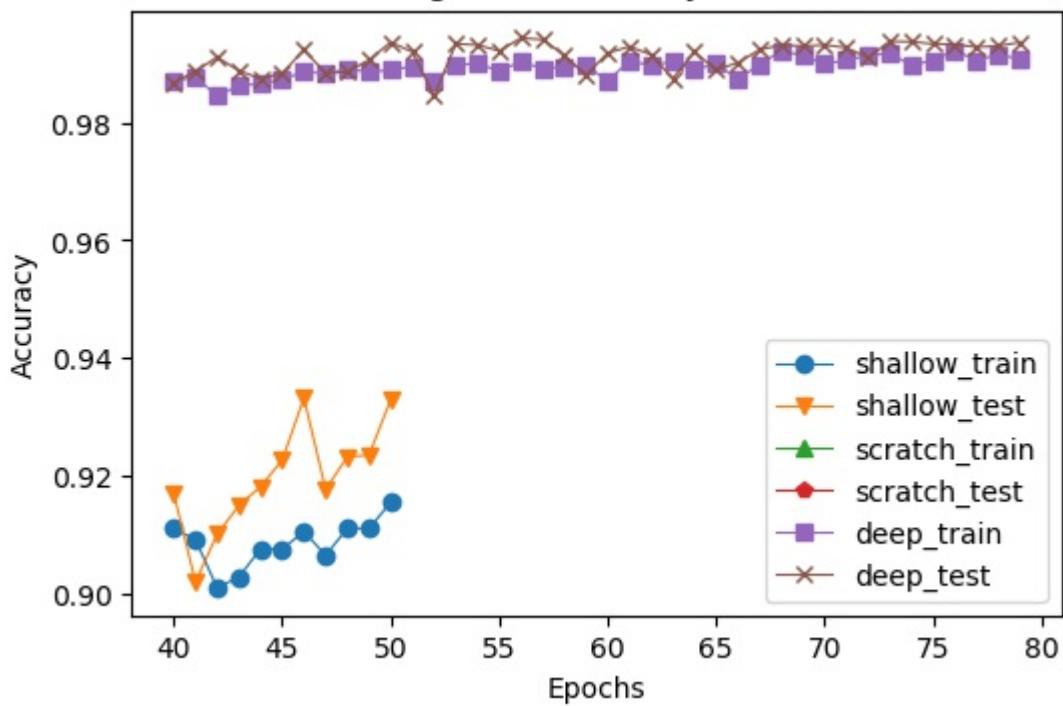
epoch 50 accuracy on test set is: 0.9329334787350054

evaluating on test set takes 16 seconds.

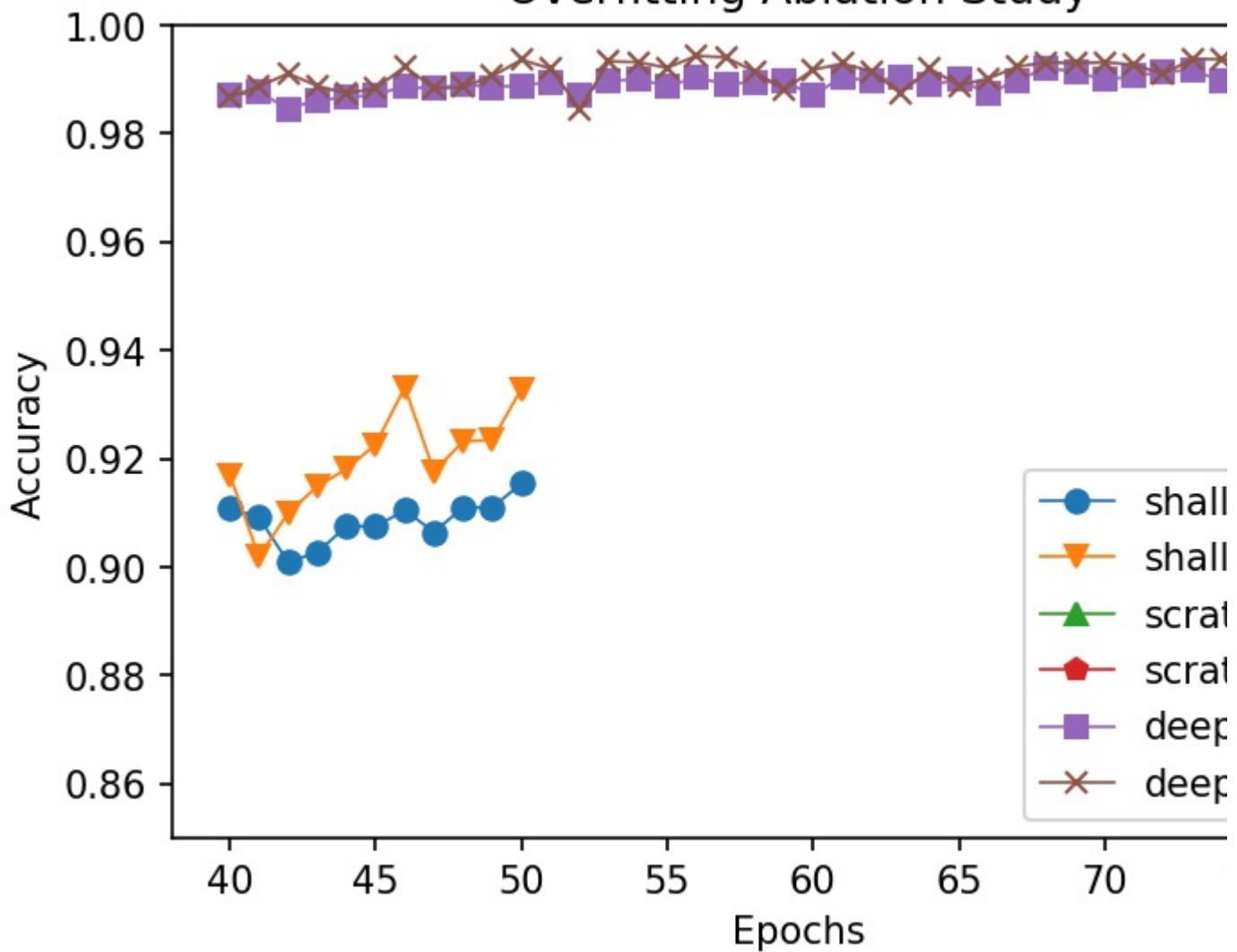
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	NaN	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 50 takes 149 seconds.  
[52,    30] loss: 0.488  
[52,    60] loss: 0.470  
[52,    90] loss: 0.421  
[52,   120] loss: 0.527  
[52,   150] loss: 0.395  
[52,   180] loss: 0.421  
[52,   210] loss: 0.425  
[52,   240] loss: 0.413  
[52,   270] loss: 0.528  
[52,   300] loss: 0.518  
[52,   330] loss: 0.477  
[52,   360] loss: 0.401  
[52,   390] loss: 0.422  
[52,   420] loss: 0.481  
[52,   450] loss: 0.437  
[52,   480] loss: 0.474  
[52,   510] loss: 0.410  
[52,   540] loss: 0.502  
[52,   570] loss: 0.491  
[52,   600] loss: 0.445  
[52,   630] loss: 0.412  
[52,   660] loss: 0.496
```

```

[52, 690] loss: 0.494
[52, 720] loss: 0.349
[52, 750] loss: 0.472
[52, 780] loss: 0.479
[52, 810] loss: 0.497
[52, 840] loss: 0.523
[52, 870] loss: 0.431
[52, 900] loss: 0.457
[52, 930] loss: 0.528
[52, 960] loss: 0.531
[52, 990] loss: 0.435
[52, 1020] loss: 0.548
[52, 1050] loss: 0.441
[52, 1080] loss: 0.464
[52, 1110] loss: 0.546
[52, 1140] loss: 0.522
[52, 1170] loss: 0.438
[52, 1200] loss: 0.410
[52, 1230] loss: 0.354
[52, 1260] loss: 0.471
[52, 1290] loss: 0.440
[52, 1320] loss: 0.440
[52, 1350] loss: 0.468
[52, 1380] loss: 0.384
[52, 1410] loss: 0.511
[52, 1440] loss: 0.467
[52, 1470] loss: 0.423
[52, 1500] loss: 0.420
[52, 1530] loss: 0.441
[52, 1560] loss: 0.485
[52, 1590] loss: 0.467
[52, 1620] loss: 0.409
[52, 1650] loss: 0.429
[52, 1680] loss: 0.469
[52, 1710] loss: 0.517
[52, 1740] loss: 0.484
[52, 1770] loss: 0.497
[52, 1800] loss: 0.487
[52, 1830] loss: 0.513
[52, 1860] loss: 0.521
[52, 1890] loss: 0.369
[52, 1920] loss: 0.512
[52, 1950] loss: 0.504
[52, 1980] loss: 0.437
[52, 2010] loss: 0.433
[52, 2040] loss: 0.445
[52, 2070] loss: 0.424
[52, 2100] loss: 0.438
[52, 2130] loss: 0.418
[52, 2160] loss: 0.451
[52, 2190] loss: 0.479
ovft_intermediate_models/epoch51_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.915917
epoch 51 accuracy on train set is: 0.9159169392948019
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	NaN	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.928117

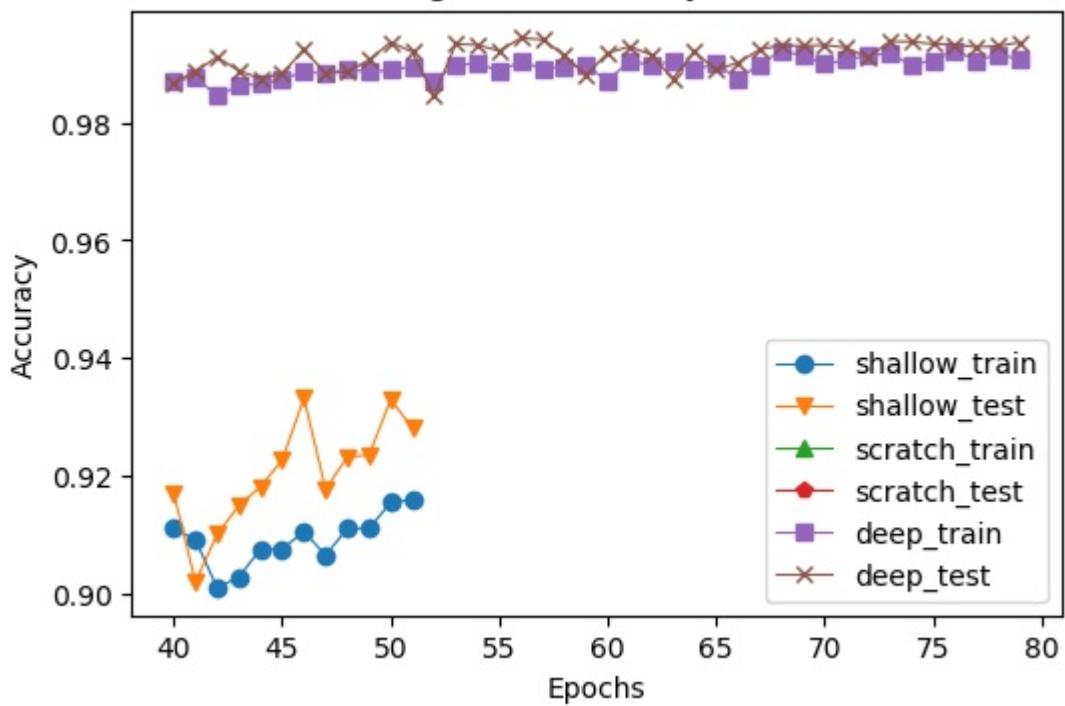
epoch 51 accuracy on test set is: 0.928117048346056

evaluating on test set takes 16 seconds.

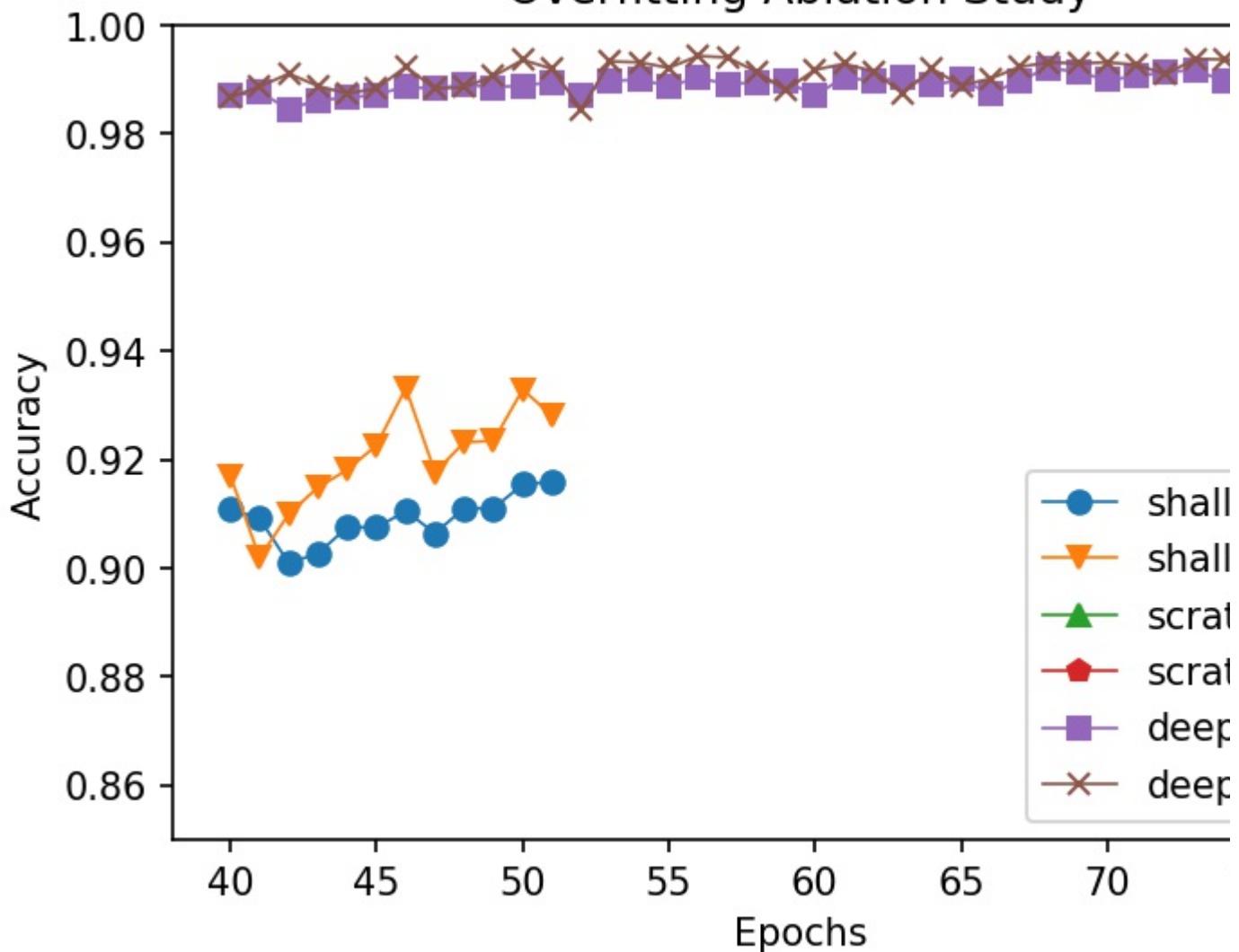
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	NaN	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 51 takes 149 seconds.  
[53,    30] loss: 0.374  
[53,    60] loss: 0.520  
[53,    90] loss: 0.492  
[53,   120] loss: 0.460  
[53,   150] loss: 0.430  
[53,   180] loss: 0.380  
[53,   210] loss: 0.521  
[53,   240] loss: 0.432  
[53,   270] loss: 0.425  
[53,   300] loss: 0.495  
[53,   330] loss: 0.410  
[53,   360] loss: 0.473  
[53,   390] loss: 0.511  
[53,   420] loss: 0.489  
[53,   450] loss: 0.491  
[53,   480] loss: 0.527  
[53,   510] loss: 0.430  
[53,   540] loss: 0.472  
[53,   570] loss: 0.401  
[53,   600] loss: 0.499  
[53,   630] loss: 0.535  
[53,   660] loss: 0.467
```

```

[53, 690] loss: 0.458
[53, 720] loss: 0.381
[53, 750] loss: 0.505
[53, 780] loss: 0.490
[53, 810] loss: 0.439
[53, 840] loss: 0.514
[53, 870] loss: 0.453
[53, 900] loss: 0.564
[53, 930] loss: 0.452
[53, 960] loss: 0.457
[53, 990] loss: 0.407
[53, 1020] loss: 0.433
[53, 1050] loss: 0.522
[53, 1080] loss: 0.421
[53, 1110] loss: 0.439
[53, 1140] loss: 0.387
[53, 1170] loss: 0.467
[53, 1200] loss: 0.602
[53, 1230] loss: 0.548
[53, 1260] loss: 0.436
[53, 1290] loss: 0.450
[53, 1320] loss: 0.457
[53, 1350] loss: 0.486
[53, 1380] loss: 0.462
[53, 1410] loss: 0.418
[53, 1440] loss: 0.521
[53, 1470] loss: 0.516
[53, 1500] loss: 0.416
[53, 1530] loss: 0.399
[53, 1560] loss: 0.414
[53, 1590] loss: 0.497
[53, 1620] loss: 0.466
[53, 1650] loss: 0.601
[53, 1680] loss: 0.496
[53, 1710] loss: 0.343
[53, 1740] loss: 0.440
[53, 1770] loss: 0.496
[53, 1800] loss: 0.477
[53, 1830] loss: 0.371
[53, 1860] loss: 0.529
[53, 1890] loss: 0.465
[53, 1920] loss: 0.563
[53, 1950] loss: 0.385
[53, 1980] loss: 0.451
[53, 2010] loss: 0.491
[53, 2040] loss: 0.455
[53, 2070] loss: 0.380
[53, 2100] loss: 0.500
[53, 2130] loss: 0.496
[53, 2160] loss: 0.412
[53, 2190] loss: 0.396
ovft_intermediate_models/epoch52_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.916212
epoch 52 accuracy on train set is: 0.916212286441294
evaluating on training set takes 66 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	NaN	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.928026

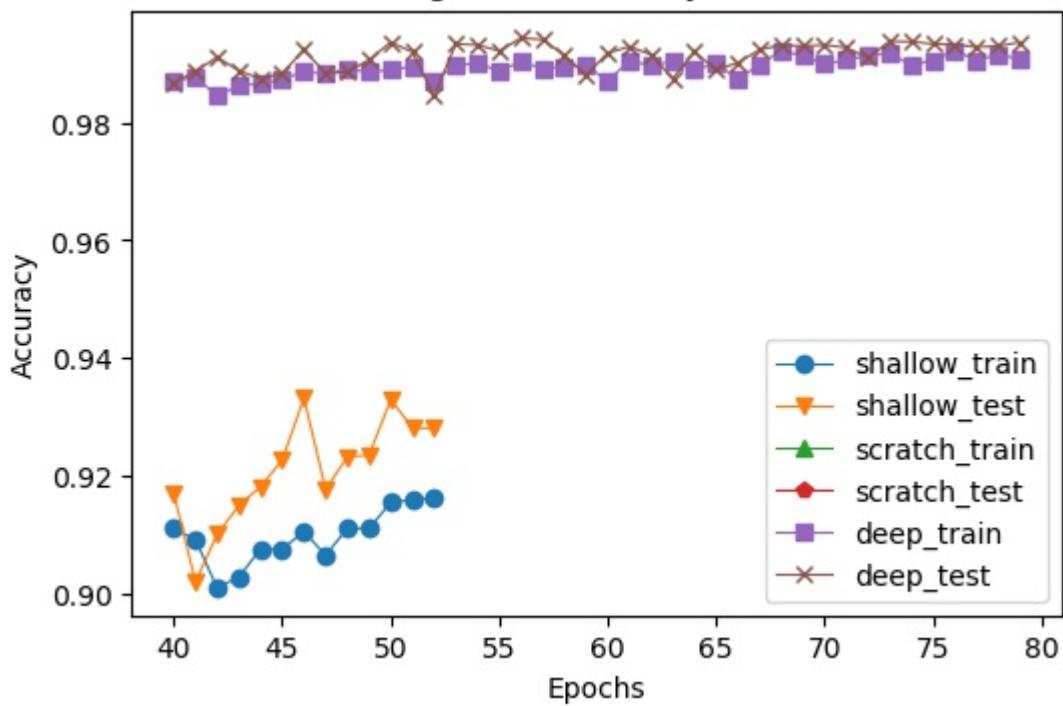
epoch 52 accuracy on test set is: 0.9280261723009815

evaluating on test set takes 16 seconds.

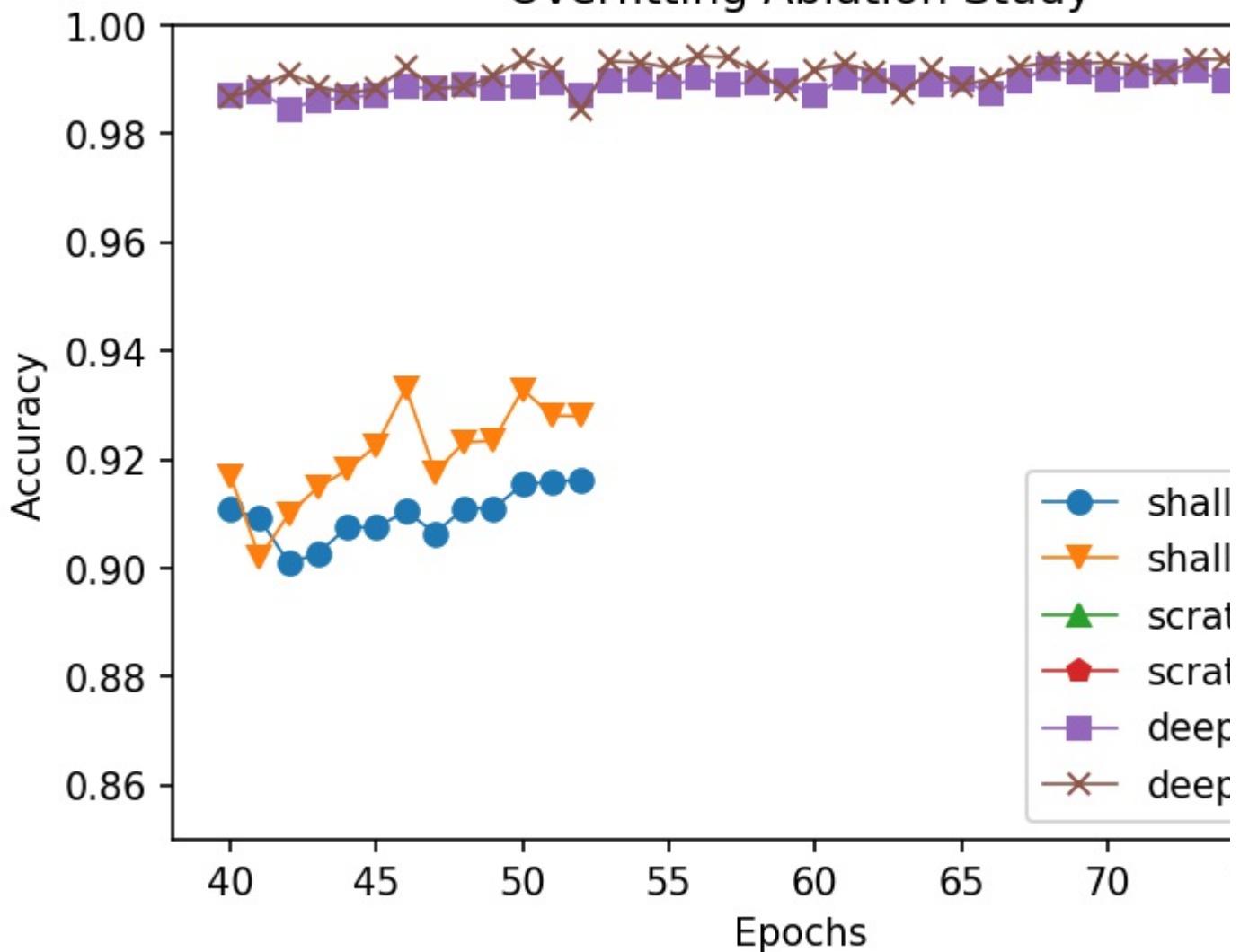
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	NaN	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 52 takes 151 seconds.  
[54,    30] loss: 0.548  
[54,    60] loss: 0.476  
[54,    90] loss: 0.497  
[54,   120] loss: 0.546  
[54,   150] loss: 0.525  
[54,   180] loss: 0.477  
[54,   210] loss: 0.397  
[54,   240] loss: 0.403  
[54,   270] loss: 0.465  
[54,   300] loss: 0.380  
[54,   330] loss: 0.476  
[54,   360] loss: 0.417  
[54,   390] loss: 0.407  
[54,   420] loss: 0.448  
[54,   450] loss: 0.380  
[54,   480] loss: 0.433  
[54,   510] loss: 0.490  
[54,   540] loss: 0.445  
[54,   570] loss: 0.480  
[54,   600] loss: 0.496  
[54,   630] loss: 0.472  
[54,   660] loss: 0.575
```

```

[54, 690] loss: 0.555
[54, 720] loss: 0.506
[54, 750] loss: 0.459
[54, 780] loss: 0.518
[54, 810] loss: 0.419
[54, 840] loss: 0.481
[54, 870] loss: 0.397
[54, 900] loss: 0.491
[54, 930] loss: 0.394
[54, 960] loss: 0.445
[54, 990] loss: 0.386
[54, 1020] loss: 0.457
[54, 1050] loss: 0.470
[54, 1080] loss: 0.429
[54, 1110] loss: 0.446
[54, 1140] loss: 0.387
[54, 1170] loss: 0.551
[54, 1200] loss: 0.501
[54, 1230] loss: 0.480
[54, 1260] loss: 0.509
[54, 1290] loss: 0.447
[54, 1320] loss: 0.485
[54, 1350] loss: 0.514
[54, 1380] loss: 0.401
[54, 1410] loss: 0.533
[54, 1440] loss: 0.446
[54, 1470] loss: 0.482
[54, 1500] loss: 0.482
[54, 1530] loss: 0.504
[54, 1560] loss: 0.472
[54, 1590] loss: 0.527
[54, 1620] loss: 0.495
[54, 1650] loss: 0.538
[54, 1680] loss: 0.582
[54, 1710] loss: 0.543
[54, 1740] loss: 0.468
[54, 1770] loss: 0.454
[54, 1800] loss: 0.533
[54, 1830] loss: 0.476
[54, 1860] loss: 0.516
[54, 1890] loss: 0.472
[54, 1920] loss: 0.484
[54, 1950] loss: 0.468
[54, 1980] loss: 0.439
[54, 2010] loss: 0.382
[54, 2040] loss: 0.544
[54, 2070] loss: 0.510
[54, 2100] loss: 0.437
[54, 2130] loss: 0.475
[54, 2160] loss: 0.523
[54, 2190] loss: 0.507
ovft_intermediate_models/epoch53_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.905762
epoch 53 accuracy on train set is: 0.9057615412577245
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	NaN	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.918939

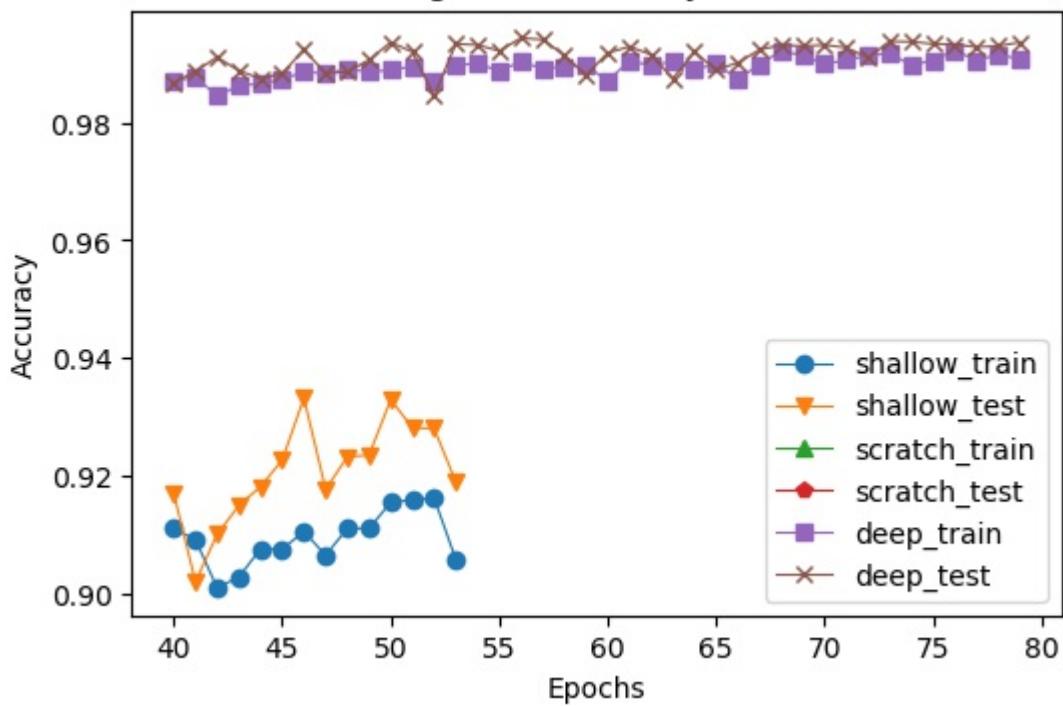
epoch 53 accuracy on test set is: 0.9189385677935297

evaluating on test set takes 16 seconds.

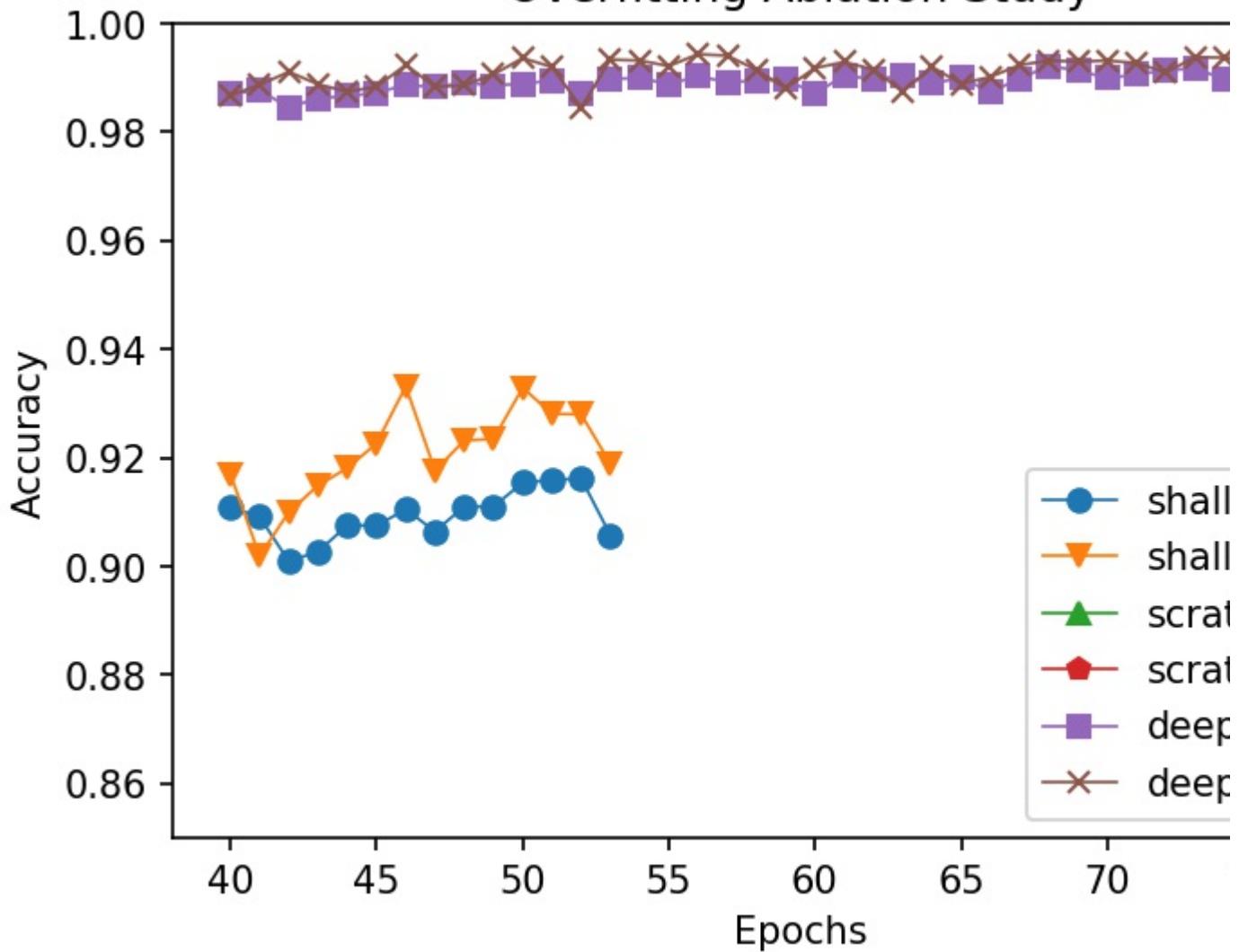
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	NaN	NaN	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 53 takes 150 seconds.  
[55,    30] loss: 0.416  
[55,    60] loss: 0.493  
[55,    90] loss: 0.420  
[55,   120] loss: 0.427  
[55,   150] loss: 0.470  
[55,   180] loss: 0.491  
[55,   210] loss: 0.493  
[55,   240] loss: 0.459  
[55,   270] loss: 0.456  
[55,   300] loss: 0.546  
[55,   330] loss: 0.495  
[55,   360] loss: 0.503  
[55,   390] loss: 0.523  
[55,   420] loss: 0.495  
[55,   450] loss: 0.401  
[55,   480] loss: 0.512  
[55,   510] loss: 0.482  
[55,   540] loss: 0.517  
[55,   570] loss: 0.448  
[55,   600] loss: 0.427  
[55,   630] loss: 0.515  
[55,   660] loss: 0.417
```

```

[55, 690] loss: 0.513
[55, 720] loss: 0.492
[55, 750] loss: 0.399
[55, 780] loss: 0.482
[55, 810] loss: 0.370
[55, 840] loss: 0.403
[55, 870] loss: 0.442
[55, 900] loss: 0.375
[55, 930] loss: 0.419
[55, 960] loss: 0.498
[55, 990] loss: 0.458
[55, 1020] loss: 0.481
[55, 1050] loss: 0.484
[55, 1080] loss: 0.443
[55, 1110] loss: 0.506
[55, 1140] loss: 0.523
[55, 1170] loss: 0.451
[55, 1200] loss: 0.431
[55, 1230] loss: 0.528
[55, 1260] loss: 0.400
[55, 1290] loss: 0.534
[55, 1320] loss: 0.396
[55, 1350] loss: 0.455
[55, 1380] loss: 0.525
[55, 1410] loss: 0.546
[55, 1440] loss: 0.557
[55, 1470] loss: 0.545
[55, 1500] loss: 0.475
[55, 1530] loss: 0.540
[55, 1560] loss: 0.442
[55, 1590] loss: 0.436
[55, 1620] loss: 0.483
[55, 1650] loss: 0.472
[55, 1680] loss: 0.502
[55, 1710] loss: 0.499
[55, 1740] loss: 0.491
[55, 1770] loss: 0.476
[55, 1800] loss: 0.459
[55, 1830] loss: 0.476
[55, 1860] loss: 0.392
[55, 1890] loss: 0.458
[55, 1920] loss: 0.401
[55, 1950] loss: 0.432
[55, 1980] loss: 0.450
[55, 2010] loss: 0.480
[55, 2040] loss: 0.387
[55, 2070] loss: 0.383
[55, 2100] loss: 0.489
[55, 2130] loss: 0.475
[55, 2160] loss: 0.519
[55, 2190] loss: 0.436
ovft_intermediate_models/epoch54_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.912963
epoch 54 accuracy on train set is: 0.91296346782988
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	NaN	NaN	NaN	0.989981	0.993184
55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.913577

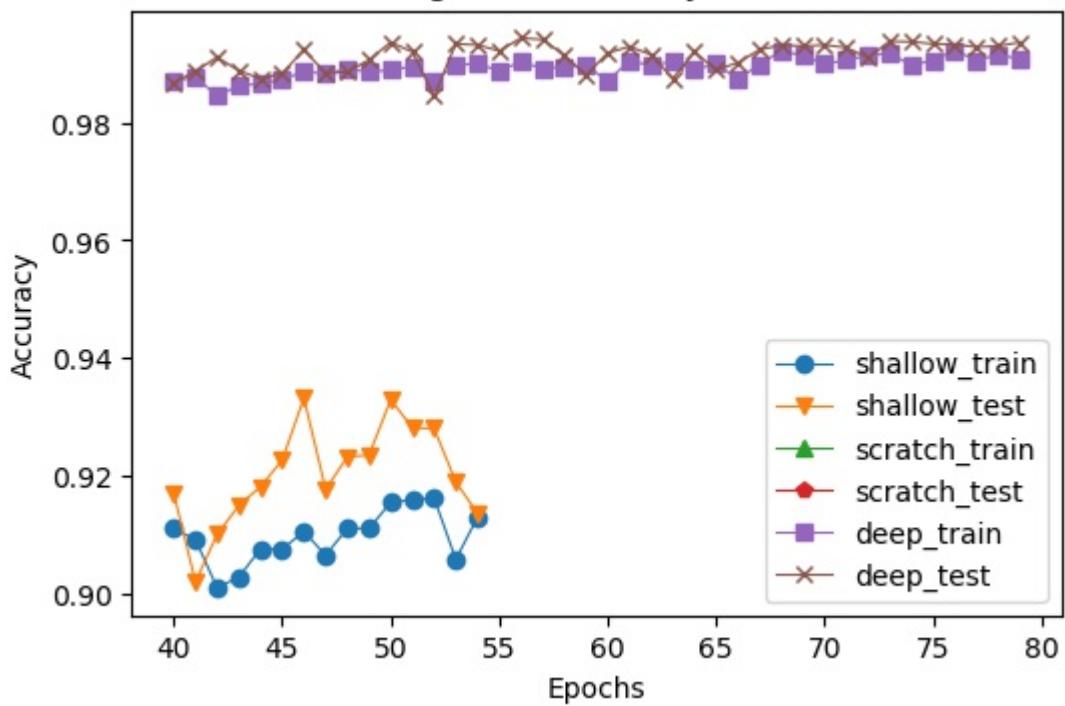
epoch 54 accuracy on test set is: 0.913576881134133

evaluating on test set takes 16 seconds.

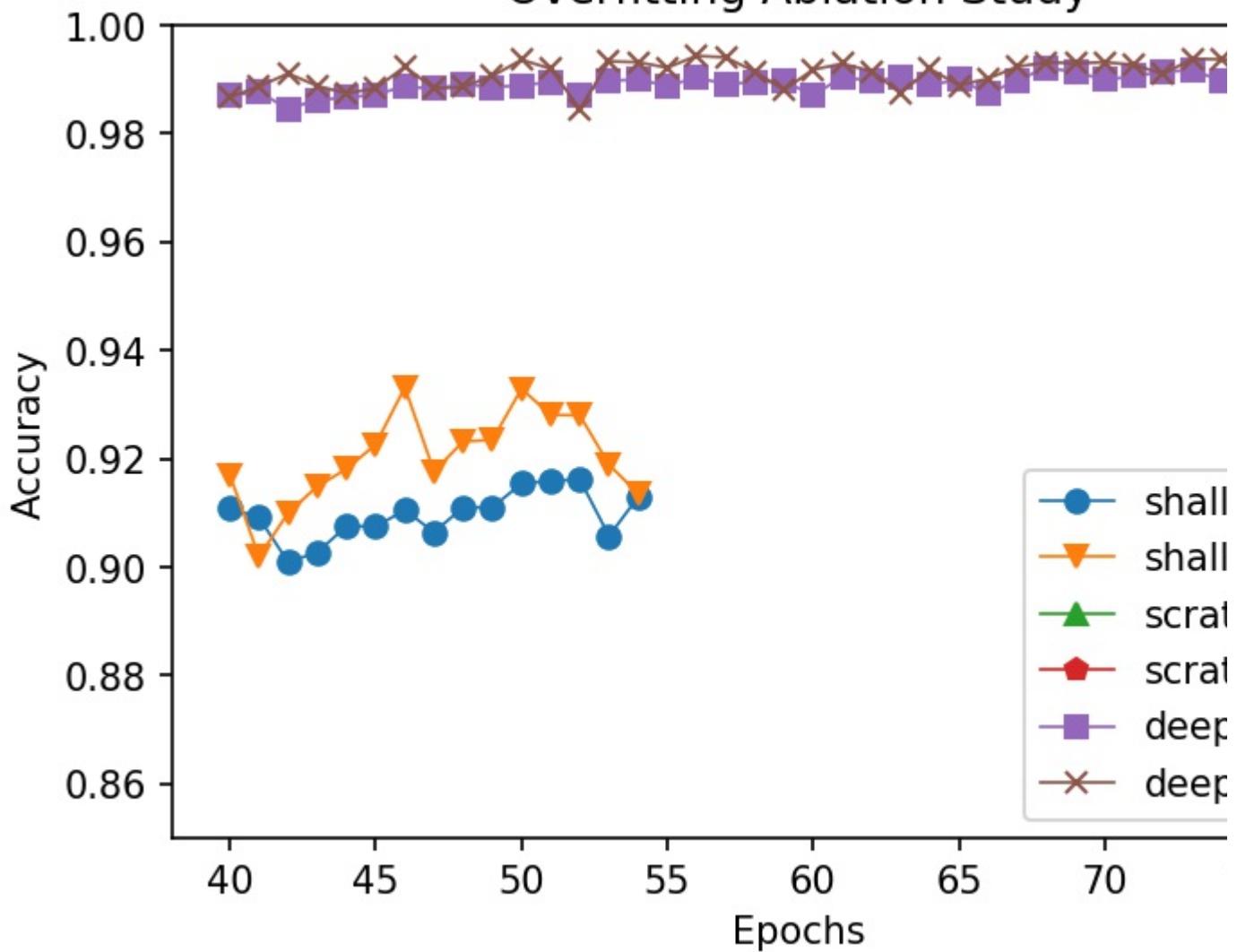
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	NaN	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 54 takes 149 seconds.  
[56,    30] loss: 0.398  
[56,    60] loss: 0.481  
[56,    90] loss: 0.482  
[56,   120] loss: 0.425  
[56,   150] loss: 0.448  
[56,   180] loss: 0.434  
[56,   210] loss: 0.467  
[56,   240] loss: 0.541  
[56,   270] loss: 0.477  
[56,   300] loss: 0.501  
[56,   330] loss: 0.406  
[56,   360] loss: 0.425  
[56,   390] loss: 0.514  
[56,   420] loss: 0.519  
[56,   450] loss: 0.396  
[56,   480] loss: 0.436  
[56,   510] loss: 0.445  
[56,   540] loss: 0.439  
[56,   570] loss: 0.412  
[56,   600] loss: 0.393  
[56,   630] loss: 0.421  
[56,   660] loss: 0.470
```

```

[56, 690] loss: 0.444
[56, 720] loss: 0.459
[56, 750] loss: 0.490
[56, 780] loss: 0.429
[56, 810] loss: 0.478
[56, 840] loss: 0.420
[56, 870] loss: 0.481
[56, 900] loss: 0.489
[56, 930] loss: 0.431
[56, 960] loss: 0.468
[56, 990] loss: 0.486
[56, 1020] loss: 0.516
[56, 1050] loss: 0.432
[56, 1080] loss: 0.358
[56, 1110] loss: 0.462
[56, 1140] loss: 0.551
[56, 1170] loss: 0.486
[56, 1200] loss: 0.445
[56, 1230] loss: 0.418
[56, 1260] loss: 0.478
[56, 1290] loss: 0.430
[56, 1320] loss: 0.456
[56, 1350] loss: 0.522
[56, 1380] loss: 0.530
[56, 1410] loss: 0.408
[56, 1440] loss: 0.424
[56, 1470] loss: 0.554
[56, 1500] loss: 0.493
[56, 1530] loss: 0.429
[56, 1560] loss: 0.502
[56, 1590] loss: 0.468
[56, 1620] loss: 0.413
[56, 1650] loss: 0.466
[56, 1680] loss: 0.510
[56, 1710] loss: 0.481
[56, 1740] loss: 0.428
[56, 1770] loss: 0.483
[56, 1800] loss: 0.437
[56, 1830] loss: 0.510
[56, 1860] loss: 0.455
[56, 1890] loss: 0.391
[56, 1920] loss: 0.519
[56, 1950] loss: 0.507
[56, 1980] loss: 0.420
[56, 2010] loss: 0.444
[56, 2040] loss: 0.492
[56, 2070] loss: 0.507
[56, 2100] loss: 0.394
[56, 2130] loss: 0.487
[56, 2160] loss: 0.452
[56, 2190] loss: 0.435
ovft intermediate models/epoch55_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.906057
epoch 55 accuracy on train set is: 0.9060568884042166
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	NaN	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.918302

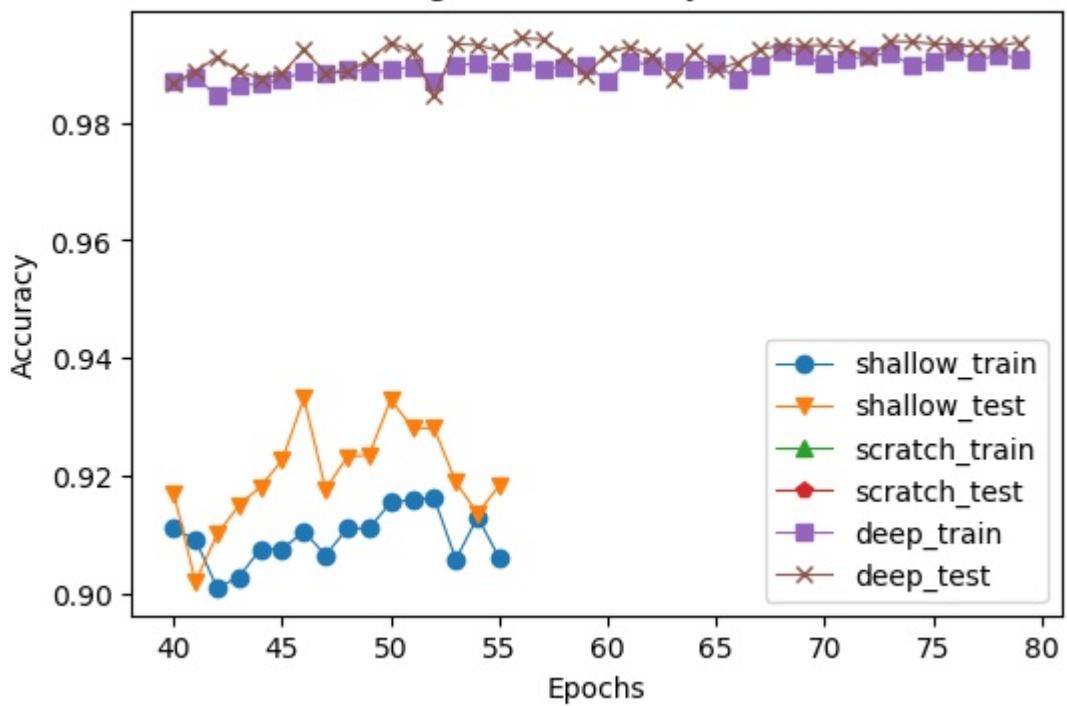
epoch 55 accuracy on test set is: 0.918302435478008

evaluating on test set takes 16 seconds.

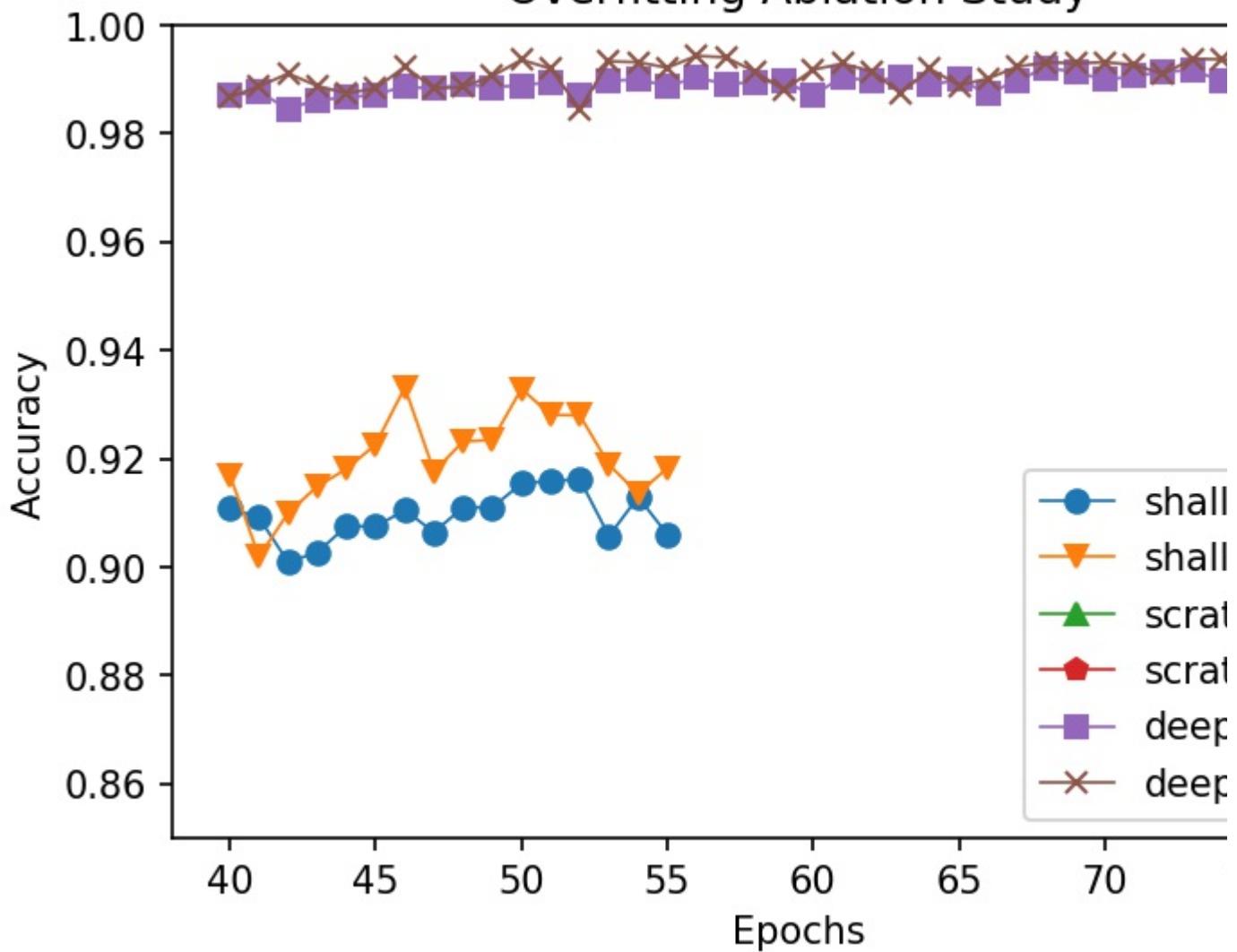
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	NaN	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



epoch 55 takes 150 seconds.

```
[57,    30] loss: 0.392
[57,    60] loss: 0.463
[57,    90] loss: 0.478
[57,   120] loss: 0.480
[57,   150] loss: 0.456
[57,   180] loss: 0.586
[57,   210] loss: 0.508
[57,   240] loss: 0.451
[57,   270] loss: 0.437
[57,   300] loss: 0.520
[57,   330] loss: 0.371
[57,   360] loss: 0.451
[57,   390] loss: 0.540
[57,   420] loss: 0.465
[57,   450] loss: 0.538
[57,   480] loss: 0.489
[57,   510] loss: 0.485
[57,   540] loss: 0.394
[57,   570] loss: 0.434
[57,   600] loss: 0.448
[57,   630] loss: 0.475
[57,   660] loss: 0.466
```

```

[57,   690] loss: 0.479
[57,   720] loss: 0.456
[57,   750] loss: 0.519
[57,   780] loss: 0.496
[57,   810] loss: 0.469
[57,   840] loss: 0.461
[57,   870] loss: 0.372
[57,   900] loss: 0.489
[57,   930] loss: 0.334
[57,   960] loss: 0.500
[57,   990] loss: 0.390
[57,  1020] loss: 0.503
[57,  1050] loss: 0.415
[57,  1080] loss: 0.525
[57,  1110] loss: 0.486
[57,  1140] loss: 0.475
[57,  1170] loss: 0.385
[57,  1200] loss: 0.481
[57,  1230] loss: 0.481
[57,  1260] loss: 0.488
[57,  1290] loss: 0.556
[57,  1320] loss: 0.483
[57,  1350] loss: 0.346
[57,  1380] loss: 0.440
[57,  1410] loss: 0.424
[57,  1440] loss: 0.510
[57,  1470] loss: 0.509
[57,  1500] loss: 0.416
[57,  1530] loss: 0.468
[57,  1560] loss: 0.474
[57,  1590] loss: 0.470
[57,  1620] loss: 0.505
[57,  1650] loss: 0.490
[57,  1680] loss: 0.482
[57,  1710] loss: 0.442
[57,  1740] loss: 0.449
[57,  1770] loss: 0.341
[57,  1800] loss: 0.445
[57,  1830] loss: 0.493
[57,  1860] loss: 0.441
[57,  1890] loss: 0.434
[57,  1920] loss: 0.459
[57,  1950] loss: 0.438
[57,  1980] loss: 0.424
[57,  2010] loss: 0.465
[57,  2040] loss: 0.469
[57,  2070] loss: 0.493
[57,  2100] loss: 0.438
[57,  2130] loss: 0.513
[57,  2160] loss: 0.517
[57,  2190] loss: 0.436
ovft_intermediate_models/epoch56_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.907079
epoch 56 accuracy on train set is: 0.907079243911305
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027  0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	NaN	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.904035

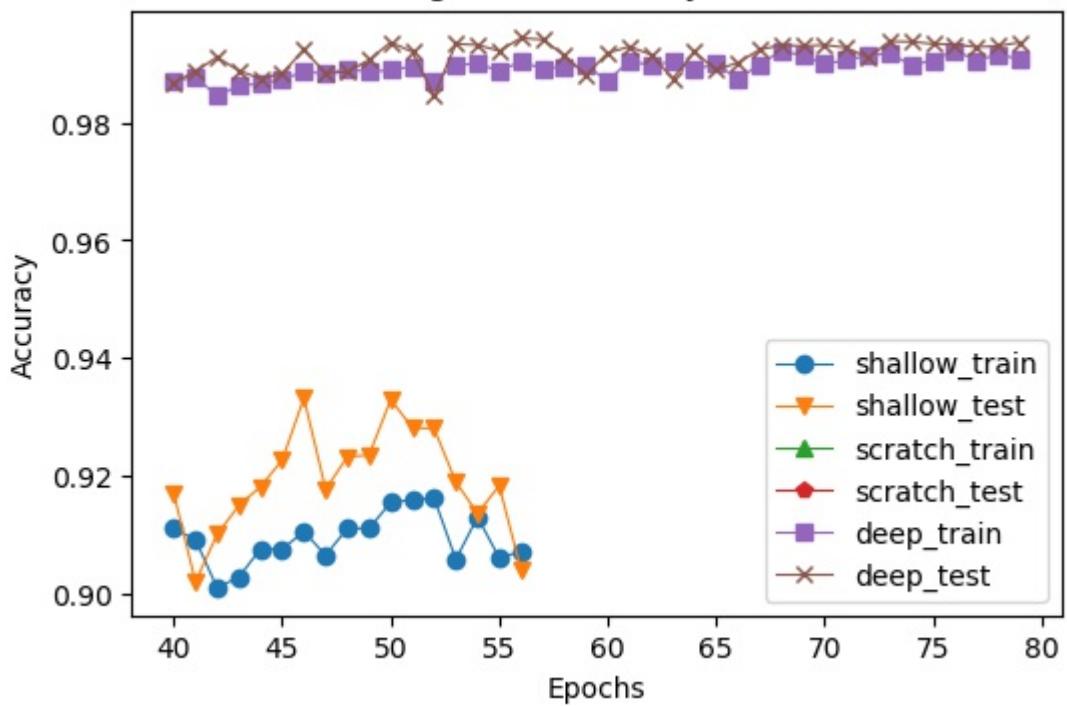
epoch 56 accuracy on test set is: 0.9040348964013086

evaluating on test set takes 16 seconds.

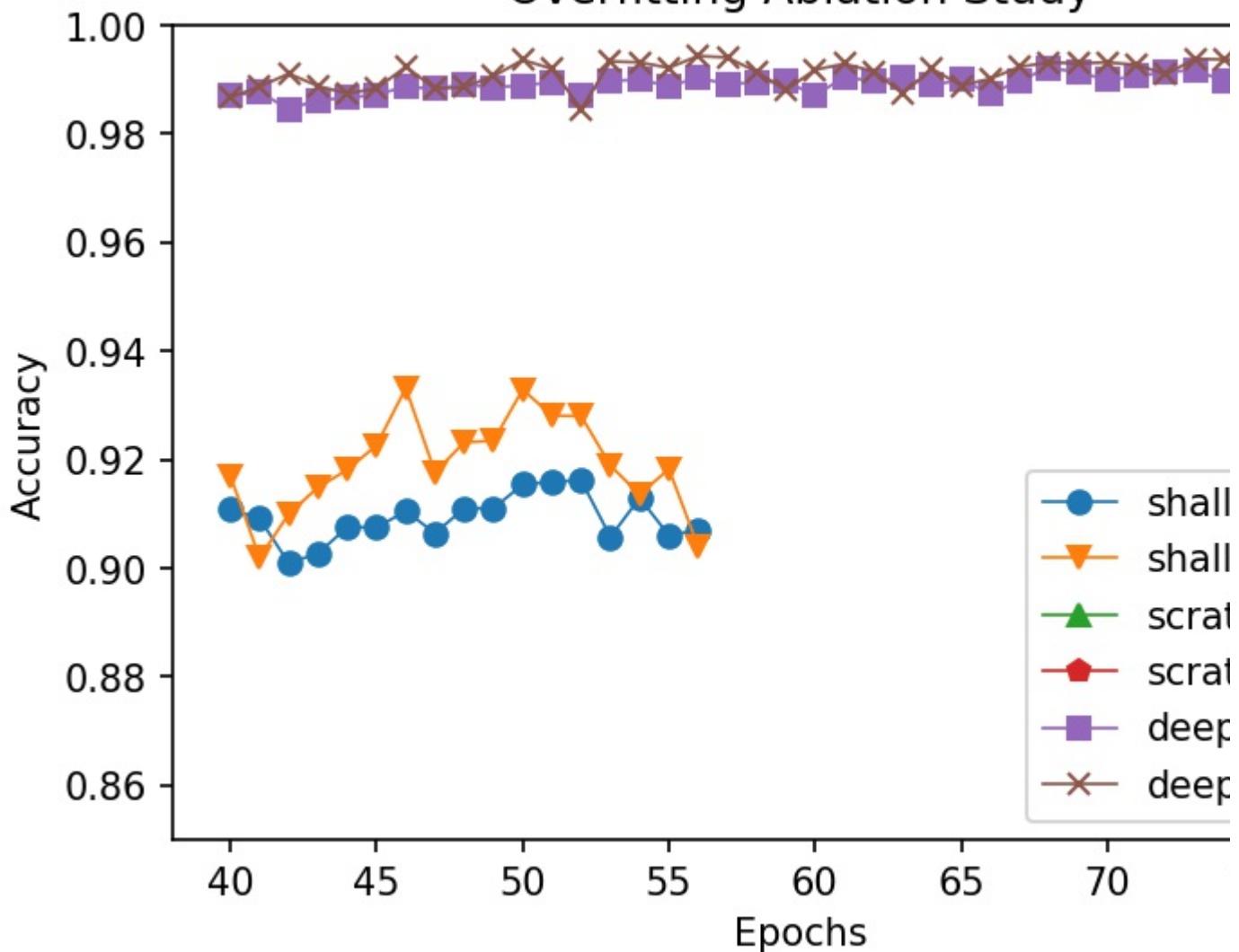
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	NaN	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 56 takes 149 seconds.  
[58,    30] loss: 0.601  
[58,    60] loss: 0.543  
[58,    90] loss: 0.480  
[58,   120] loss: 0.382  
[58,   150] loss: 0.480  
[58,   180] loss: 0.502  
[58,   210] loss: 0.495  
[58,   240] loss: 0.531  
[58,   270] loss: 0.472  
[58,   300] loss: 0.394  
[58,   330] loss: 0.471  
[58,   360] loss: 0.507  
[58,   390] loss: 0.438  
[58,   420] loss: 0.501  
[58,   450] loss: 0.417  
[58,   480] loss: 0.502  
[58,   510] loss: 0.521  
[58,   540] loss: 0.426  
[58,   570] loss: 0.587  
[58,   600] loss: 0.383  
[58,   630] loss: 0.458  
[58,   660] loss: 0.413
```

```

[58, 690] loss: 0.429
[58, 720] loss: 0.455
[58, 750] loss: 0.407
[58, 780] loss: 0.430
[58, 810] loss: 0.481
[58, 840] loss: 0.417
[58, 870] loss: 0.445
[58, 900] loss: 0.418
[58, 930] loss: 0.540
[58, 960] loss: 0.502
[58, 990] loss: 0.403
[58, 1020] loss: 0.537
[58, 1050] loss: 0.451
[58, 1080] loss: 0.572
[58, 1110] loss: 0.435
[58, 1140] loss: 0.567
[58, 1170] loss: 0.474
[58, 1200] loss: 0.462
[58, 1230] loss: 0.515
[58, 1260] loss: 0.452
[58, 1290] loss: 0.495
[58, 1320] loss: 0.520
[58, 1350] loss: 0.385
[58, 1380] loss: 0.415
[58, 1410] loss: 0.454
[58, 1440] loss: 0.461
[58, 1470] loss: 0.475
[58, 1500] loss: 0.468
[58, 1530] loss: 0.410
[58, 1560] loss: 0.446
[58, 1590] loss: 0.487
[58, 1620] loss: 0.426
[58, 1650] loss: 0.510
[58, 1680] loss: 0.488
[58, 1710] loss: 0.450
[58, 1740] loss: 0.441
[58, 1770] loss: 0.429
[58, 1800] loss: 0.427
[58, 1830] loss: 0.492
[58, 1860] loss: 0.502
[58, 1890] loss: 0.418
[58, 1920] loss: 0.594
[58, 1950] loss: 0.409
[58, 1980] loss: 0.519
[58, 2010] loss: 0.363
[58, 2040] loss: 0.479
[58, 2070] loss: 0.458
[58, 2100] loss: 0.516
[58, 2130] loss: 0.552
[58, 2160] loss: 0.569
[58, 2190] loss: 0.478
ovft_intermediate_models/epoch57_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.912327
epoch 57 accuracy on train set is: 0.9123273355143584
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	NaN	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.925573

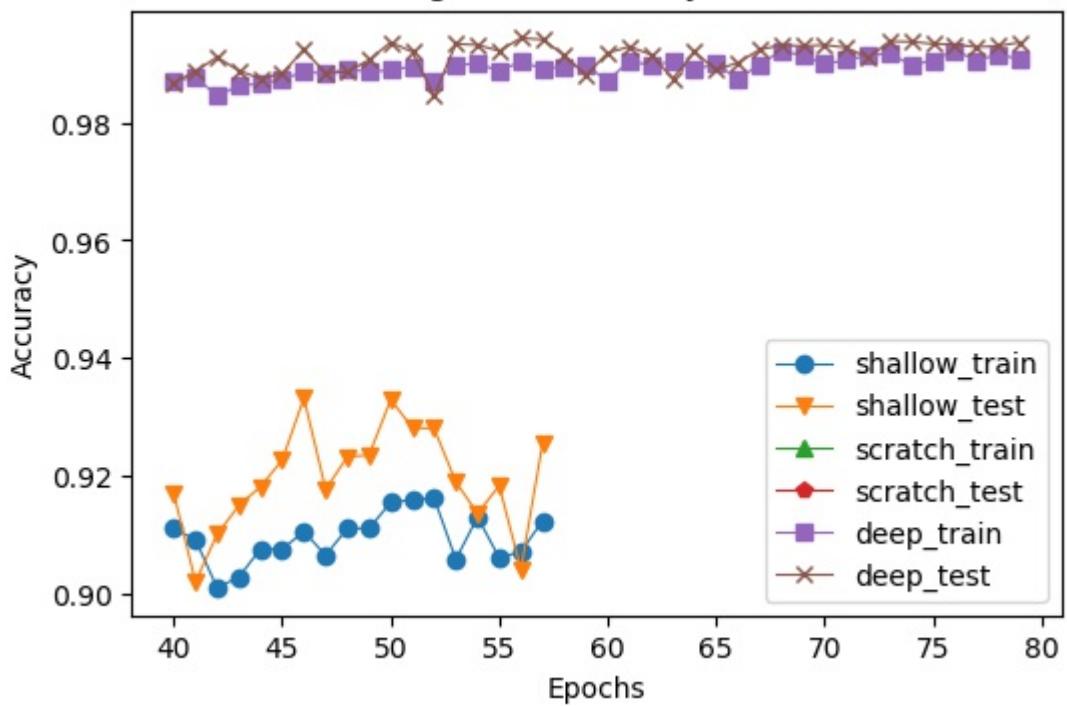
epoch 57 accuracy on test set is: 0.9255725190839694

evaluating on test set takes 16 seconds.

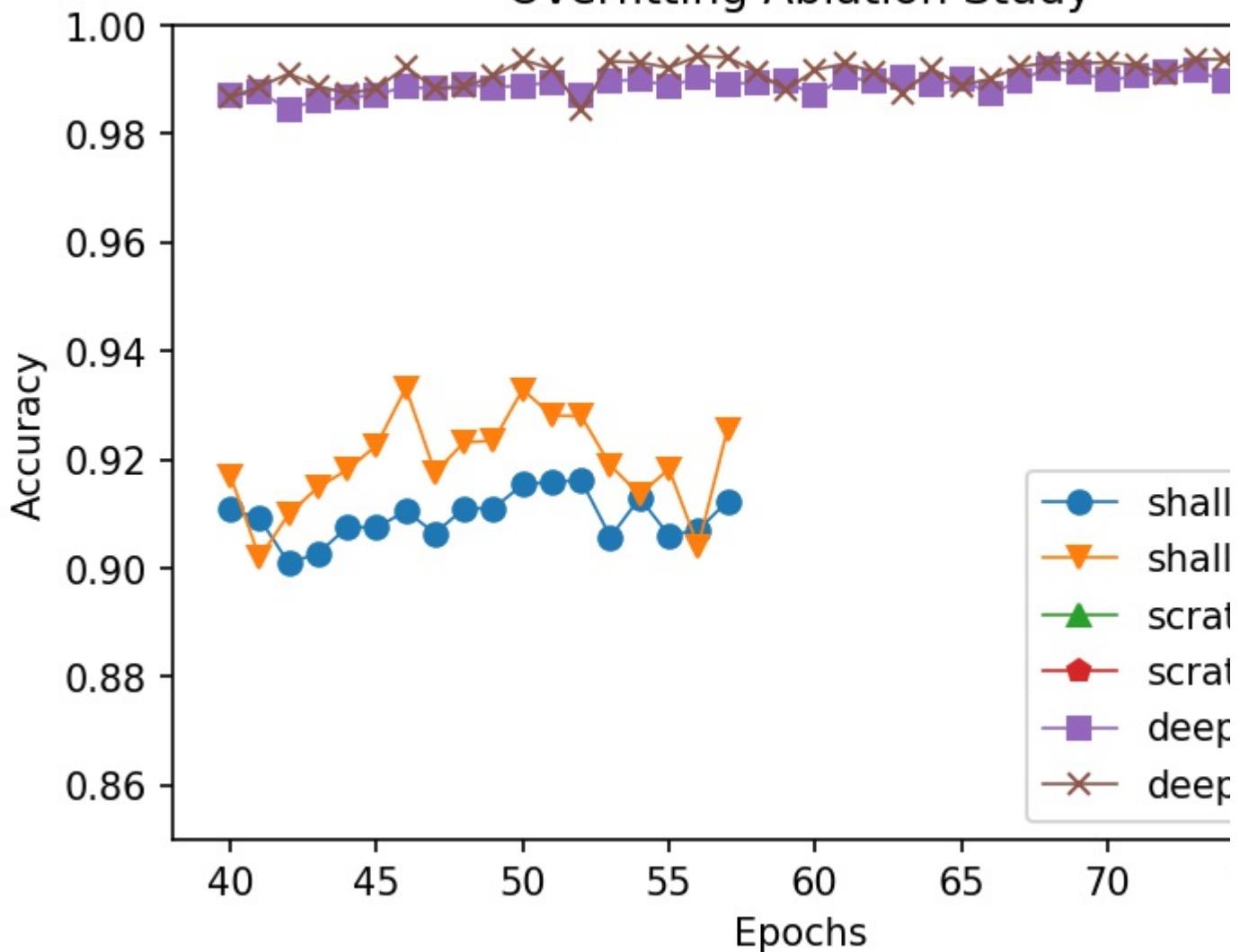
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	NaN	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 57 takes 149 seconds.  
[59,    30] loss: 0.554  
[59,    60] loss: 0.441  
[59,    90] loss: 0.452  
[59,   120] loss: 0.462  
[59,   150] loss: 0.396  
[59,   180] loss: 0.484  
[59,   210] loss: 0.420  
[59,   240] loss: 0.492  
[59,   270] loss: 0.462  
[59,   300] loss: 0.536  
[59,   330] loss: 0.386  
[59,   360] loss: 0.397  
[59,   390] loss: 0.436  
[59,   420] loss: 0.468  
[59,   450] loss: 0.444  
[59,   480] loss: 0.461  
[59,   510] loss: 0.453  
[59,   540] loss: 0.462  
[59,   570] loss: 0.475  
[59,   600] loss: 0.604  
[59,   630] loss: 0.477  
[59,   660] loss: 0.419
```

```

[59, 690] loss: 0.474
[59, 720] loss: 0.400
[59, 750] loss: 0.545
[59, 780] loss: 0.471
[59, 810] loss: 0.462
[59, 840] loss: 0.389
[59, 870] loss: 0.400
[59, 900] loss: 0.412
[59, 930] loss: 0.491
[59, 960] loss: 0.426
[59, 990] loss: 0.518
[59, 1020] loss: 0.555
[59, 1050] loss: 0.429
[59, 1080] loss: 0.516
[59, 1110] loss: 0.449
[59, 1140] loss: 0.494
[59, 1170] loss: 0.452
[59, 1200] loss: 0.423
[59, 1230] loss: 0.492
[59, 1260] loss: 0.448
[59, 1290] loss: 0.410
[59, 1320] loss: 0.479
[59, 1350] loss: 0.456
[59, 1380] loss: 0.449
[59, 1410] loss: 0.502
[59, 1440] loss: 0.420
[59, 1470] loss: 0.441
[59, 1500] loss: 0.454
[59, 1530] loss: 0.488
[59, 1560] loss: 0.464
[59, 1590] loss: 0.467
[59, 1620] loss: 0.490
[59, 1650] loss: 0.459
[59, 1680] loss: 0.486
[59, 1710] loss: 0.465
[59, 1740] loss: 0.383
[59, 1770] loss: 0.503
[59, 1800] loss: 0.467
[59, 1830] loss: 0.463
[59, 1860] loss: 0.418
[59, 1890] loss: 0.502
[59, 1920] loss: 0.566
[59, 1950] loss: 0.501
[59, 1980] loss: 0.622
[59, 2010] loss: 0.524
[59, 2040] loss: 0.483
[59, 2070] loss: 0.482
[59, 2100] loss: 0.513
[59, 2130] loss: 0.438
[59, 2160] loss: 0.499
[59, 2190] loss: 0.363
ovft_intermediate_models/epoch58_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.9141221374045801
epoch 58 accuracy on train set is: 0.9141221374045801
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	NaN	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.923573

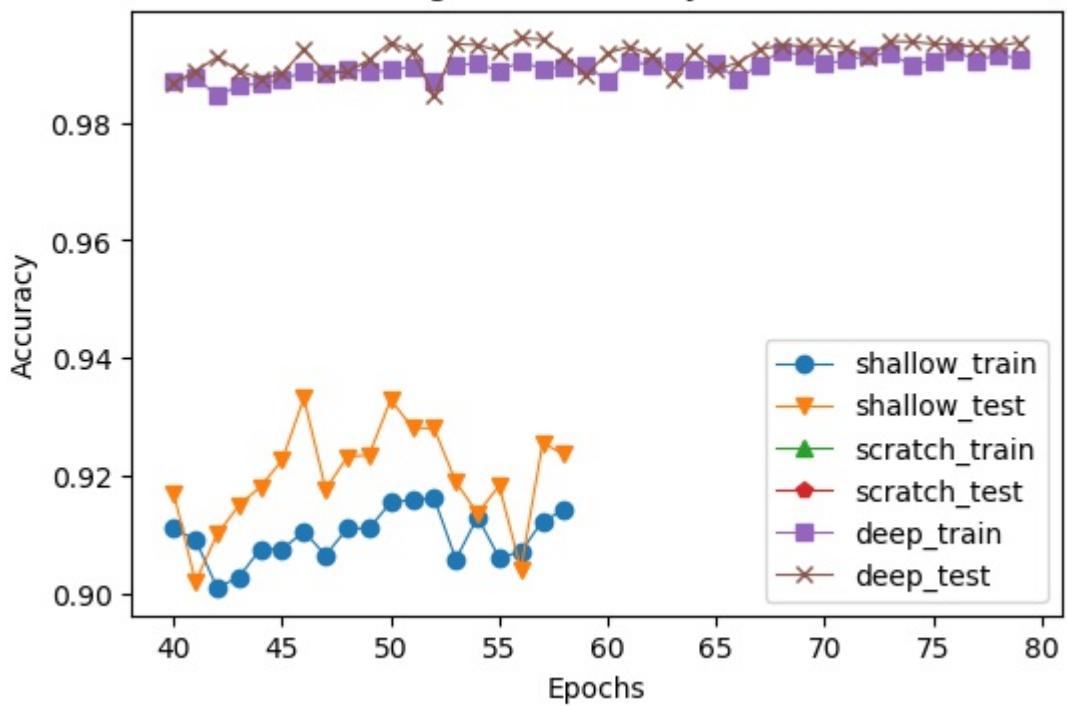
epoch 58 accuracy on test set is: 0.92357324609233

evaluating on test set takes 16 seconds.

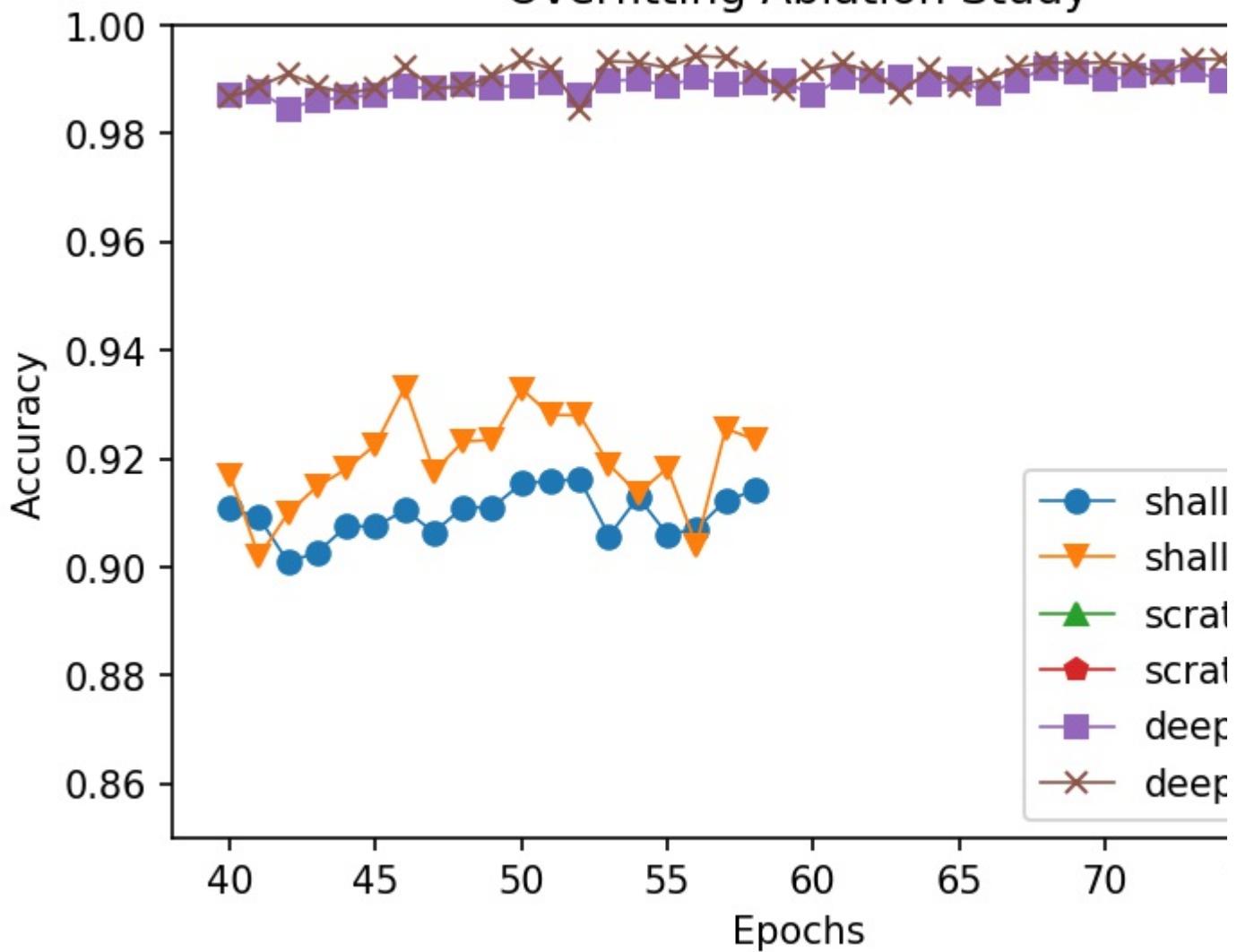
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	NaN	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 58 takes 149 seconds.  
[60,    30] loss: 0.509  
[60,    60] loss: 0.434  
[60,    90] loss: 0.437  
[60,   120] loss: 0.482  
[60,   150] loss: 0.554  
[60,   180] loss: 0.468  
[60,   210] loss: 0.406  
[60,   240] loss: 0.423  
[60,   270] loss: 0.498  
[60,   300] loss: 0.469  
[60,   330] loss: 0.397  
[60,   360] loss: 0.475  
[60,   390] loss: 0.493  
[60,   420] loss: 0.534  
[60,   450] loss: 0.433  
[60,   480] loss: 0.493  
[60,   510] loss: 0.456  
[60,   540] loss: 0.528  
[60,   570] loss: 0.497  
[60,   600] loss: 0.444  
[60,   630] loss: 0.480  
[60,   660] loss: 0.500
```

```

[60, 690] loss: 0.509
[60, 720] loss: 0.341
[60, 750] loss: 0.399
[60, 780] loss: 0.487
[60, 810] loss: 0.448
[60, 840] loss: 0.443
[60, 870] loss: 0.501
[60, 900] loss: 0.493
[60, 930] loss: 0.562
[60, 960] loss: 0.506
[60, 990] loss: 0.469
[60, 1020] loss: 0.490
[60, 1050] loss: 0.441
[60, 1080] loss: 0.482
[60, 1110] loss: 0.489
[60, 1140] loss: 0.494
[60, 1170] loss: 0.422
[60, 1200] loss: 0.416
[60, 1230] loss: 0.478
[60, 1260] loss: 0.489
[60, 1290] loss: 0.439
[60, 1320] loss: 0.409
[60, 1350] loss: 0.420
[60, 1380] loss: 0.426
[60, 1410] loss: 0.427
[60, 1440] loss: 0.456
[60, 1470] loss: 0.448
[60, 1500] loss: 0.468
[60, 1530] loss: 0.365
[60, 1560] loss: 0.499
[60, 1590] loss: 0.431
[60, 1620] loss: 0.420
[60, 1650] loss: 0.512
[60, 1680] loss: 0.481
[60, 1710] loss: 0.325
[60, 1740] loss: 0.482
[60, 1770] loss: 0.466
[60, 1800] loss: 0.498
[60, 1830] loss: 0.429
[60, 1860] loss: 0.449
[60, 1890] loss: 0.472
[60, 1920] loss: 0.507
[60, 1950] loss: 0.495
[60, 1980] loss: 0.507
[60, 2010] loss: 0.463
[60, 2040] loss: 0.547
[60, 2070] loss: 0.500
[60, 2100] loss: 0.425
[60, 2130] loss: 0.542
[60, 2160] loss: 0.469
[60, 2190] loss: 0.508
ovft_intermediate_models/epoch59_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.908738
epoch 59 accuracy on train set is: 0.9087377317339149
evaluating on training set takes 65 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	NaN	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.913395

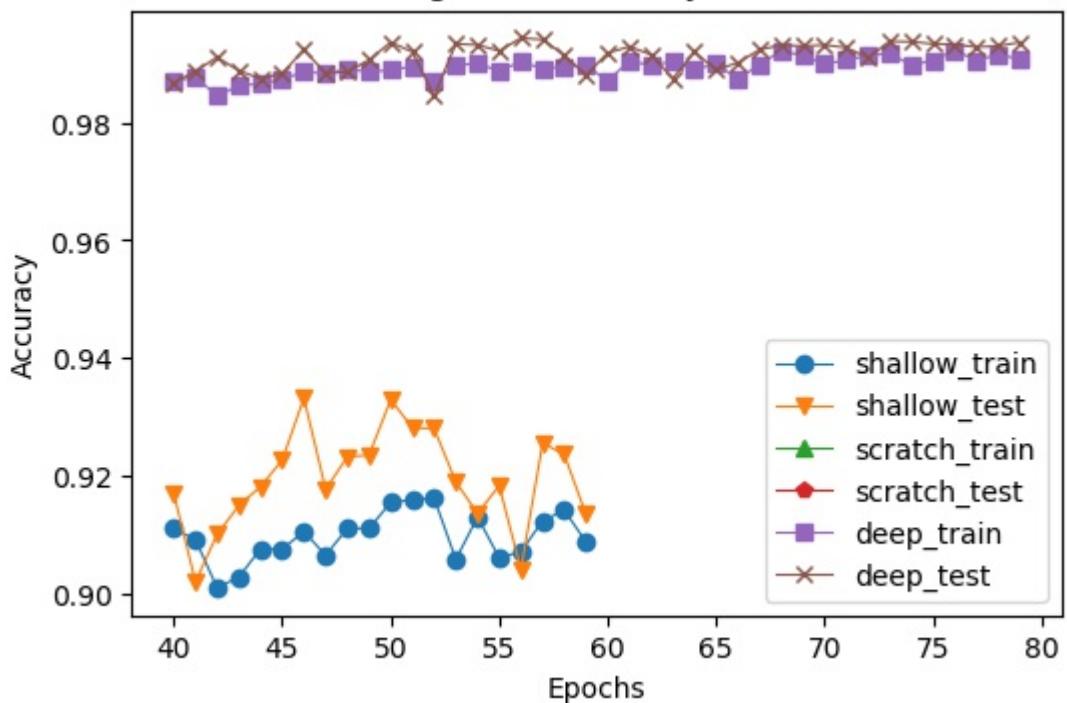
epoch 59 accuracy on test set is: 0.913395129043984

evaluating on test set takes 16 seconds.

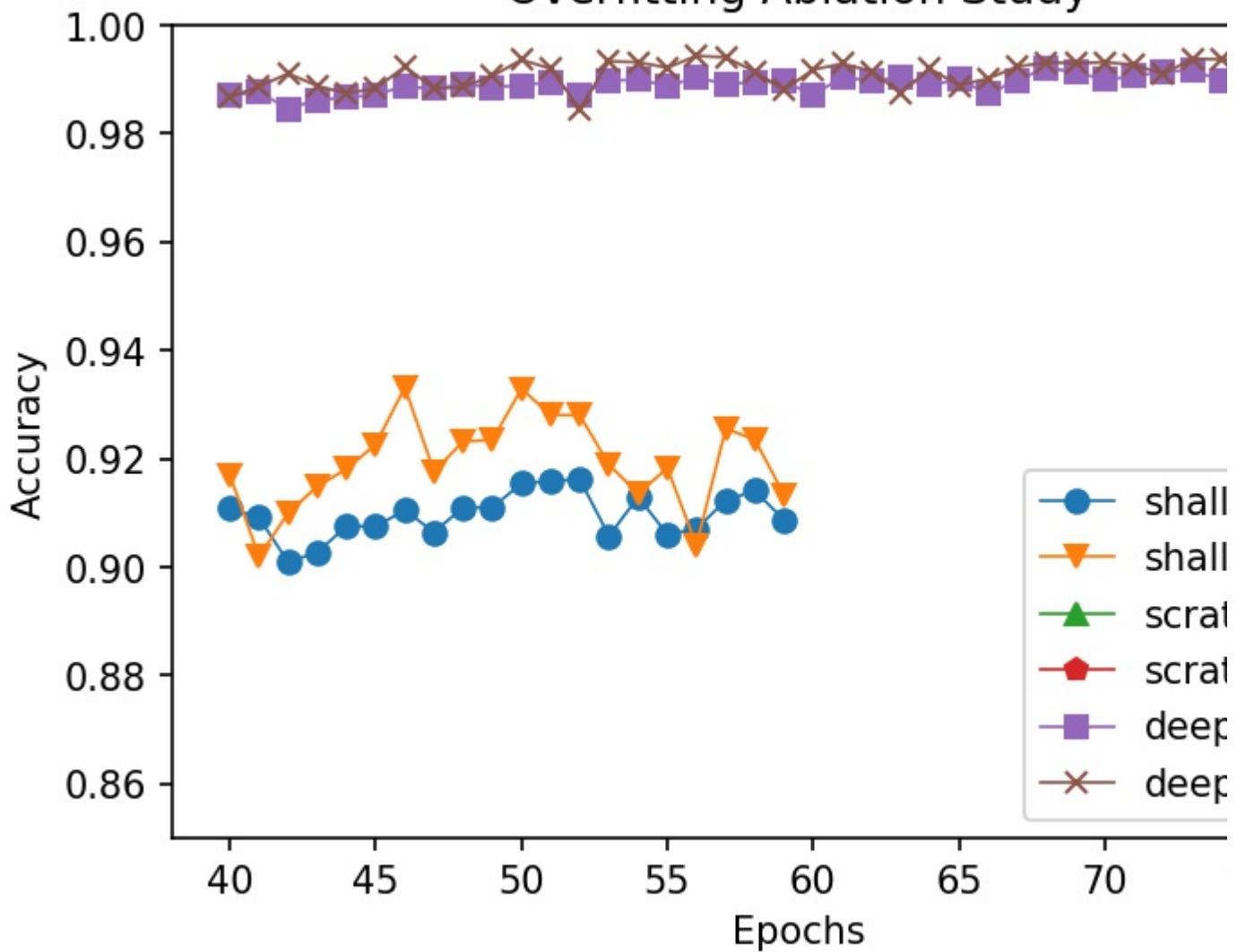
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	NaN	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 59 takes 150 seconds.  
[61,    30] loss: 0.434  
[61,    60] loss: 0.415  
[61,    90] loss: 0.404  
[61,   120] loss: 0.354  
[61,   150] loss: 0.386  
[61,   180] loss: 0.499  
[61,   210] loss: 0.484  
[61,   240] loss: 0.431  
[61,   270] loss: 0.378  
[61,   300] loss: 0.389  
[61,   330] loss: 0.467  
[61,   360] loss: 0.568  
[61,   390] loss: 0.477  
[61,   420] loss: 0.411  
[61,   450] loss: 0.408  
[61,   480] loss: 0.467  
[61,   510] loss: 0.473  
[61,   540] loss: 0.573  
[61,   570] loss: 0.456  
[61,   600] loss: 0.467  
[61,   630] loss: 0.475  
[61,   660] loss: 0.419
```

```

[61,   690] loss: 0.436
[61,   720] loss: 0.420
[61,   750] loss: 0.444
[61,   780] loss: 0.491
[61,   810] loss: 0.451
[61,   840] loss: 0.515
[61,   870] loss: 0.408
[61,   900] loss: 0.483
[61,   930] loss: 0.461
[61,   960] loss: 0.384
[61,   990] loss: 0.478
[61,  1020] loss: 0.474
[61,  1050] loss: 0.478
[61,  1080] loss: 0.464
[61,  1110] loss: 0.482
[61,  1140] loss: 0.527
[61,  1170] loss: 0.452
[61,  1200] loss: 0.423
[61,  1230] loss: 0.536
[61,  1260] loss: 0.398
[61,  1290] loss: 0.507
[61,  1320] loss: 0.477
[61,  1350] loss: 0.423
[61,  1380] loss: 0.480
[61,  1410] loss: 0.524
[61,  1440] loss: 0.473
[61,  1470] loss: 0.417
[61,  1500] loss: 0.426
[61,  1530] loss: 0.460
[61,  1560] loss: 0.403
[61,  1590] loss: 0.408
[61,  1620] loss: 0.419
[61,  1650] loss: 0.471
[61,  1680] loss: 0.493
[61,  1710] loss: 0.368
[61,  1740] loss: 0.403
[61,  1770] loss: 0.488
[61,  1800] loss: 0.504
[61,  1830] loss: 0.429
[61,  1860] loss: 0.477
[61,  1890] loss: 0.397
[61,  1920] loss: 0.368
[61,  1950] loss: 0.381
[61,  1980] loss: 0.518
[61,  2010] loss: 0.455
[61,  2040] loss: 0.471
[61,  2070] loss: 0.494
[61,  2100] loss: 0.478
[61,  2130] loss: 0.446
[61,  2160] loss: 0.463
[61,  2190] loss: 0.488
ovft_intermediate_models/epoch60_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.909965
epoch 60 accuracy on train set is: 0.909964558342421
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027  0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	NaN	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.923755

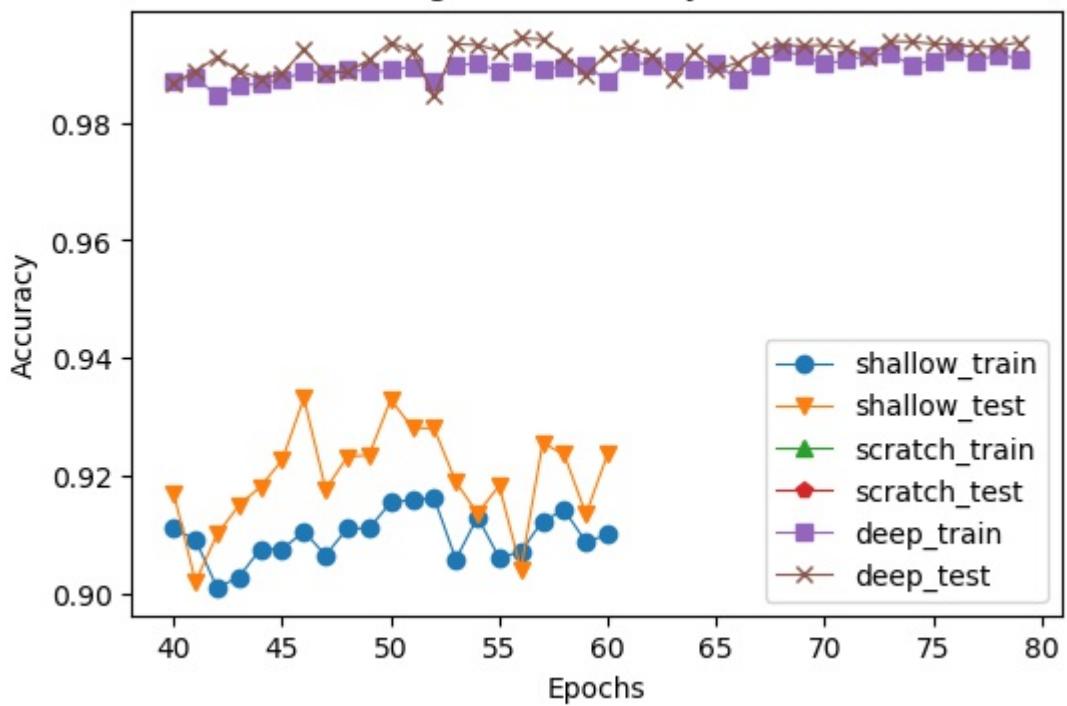
epoch 60 accuracy on test set is: 0.9237549981824791

evaluating on test set takes 16 seconds.

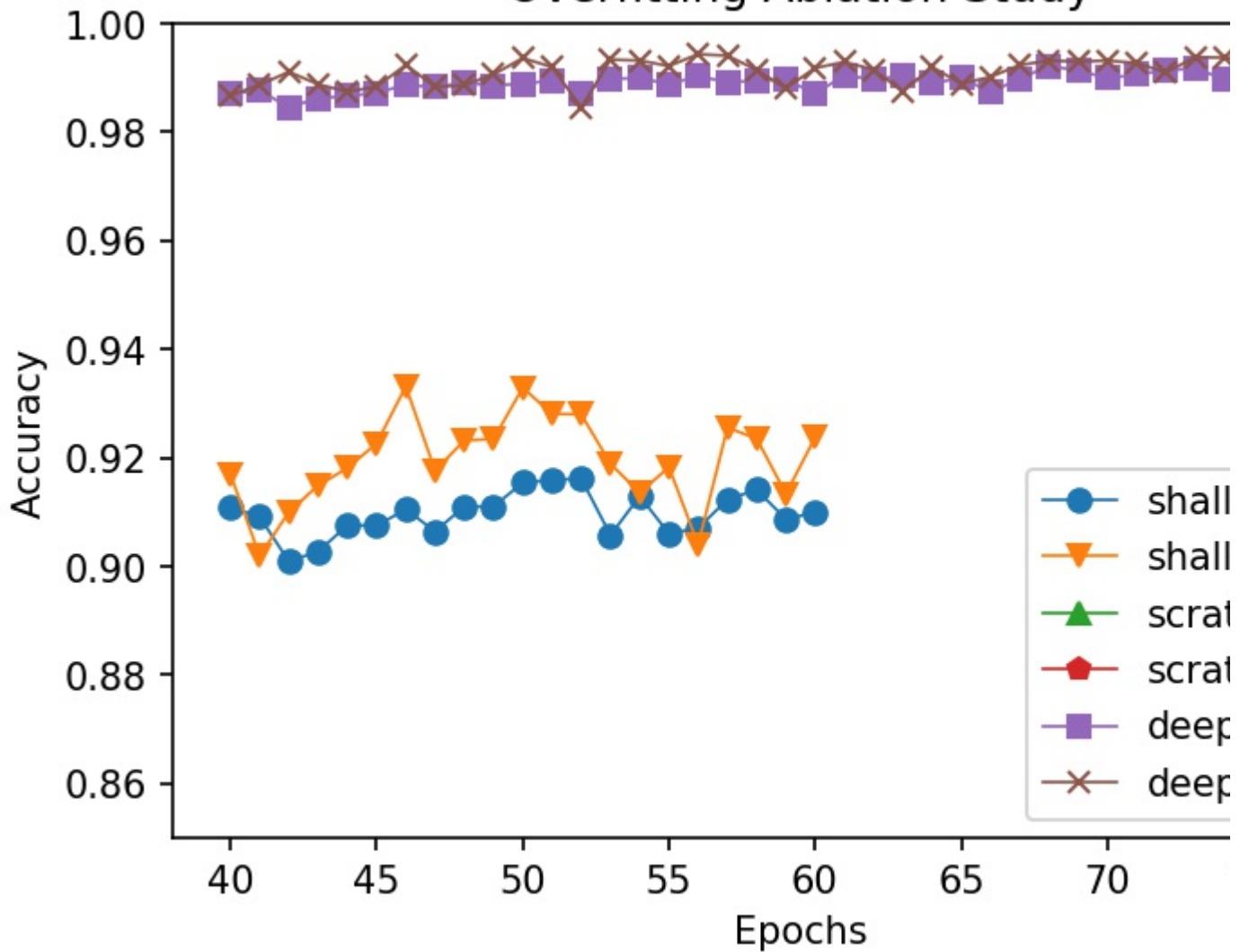
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	NaN	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 60 takes 149 seconds.  
[62,    30] loss: 0.504  
[62,    60] loss: 0.556  
[62,    90] loss: 0.569  
[62,   120] loss: 0.413  
[62,   150] loss: 0.411  
[62,   180] loss: 0.424  
[62,   210] loss: 0.418  
[62,   240] loss: 0.367  
[62,   270] loss: 0.455  
[62,   300] loss: 0.565  
[62,   330] loss: 0.441  
[62,   360] loss: 0.410  
[62,   390] loss: 0.388  
[62,   420] loss: 0.395  
[62,   450] loss: 0.465  
[62,   480] loss: 0.491  
[62,   510] loss: 0.380  
[62,   540] loss: 0.522  
[62,   570] loss: 0.507  
[62,   600] loss: 0.484  
[62,   630] loss: 0.462  
[62,   660] loss: 0.495
```

```

[62, 690] loss: 0.418
[62, 720] loss: 0.540
[62, 750] loss: 0.391
[62, 780] loss: 0.533
[62, 810] loss: 0.463
[62, 840] loss: 0.460
[62, 870] loss: 0.473
[62, 900] loss: 0.498
[62, 930] loss: 0.434
[62, 960] loss: 0.532
[62, 990] loss: 0.456
[62, 1020] loss: 0.485
[62, 1050] loss: 0.510
[62, 1080] loss: 0.536
[62, 1110] loss: 0.466
[62, 1140] loss: 0.444
[62, 1170] loss: 0.483
[62, 1200] loss: 0.506
[62, 1230] loss: 0.474
[62, 1260] loss: 0.518
[62, 1290] loss: 0.534
[62, 1320] loss: 0.423
[62, 1350] loss: 0.484
[62, 1380] loss: 0.387
[62, 1410] loss: 0.542
[62, 1440] loss: 0.444
[62, 1470] loss: 0.393
[62, 1500] loss: 0.455
[62, 1530] loss: 0.462
[62, 1560] loss: 0.439
[62, 1590] loss: 0.438
[62, 1620] loss: 0.444
[62, 1650] loss: 0.445
[62, 1680] loss: 0.445
[62, 1710] loss: 0.397
[62, 1740] loss: 0.539
[62, 1770] loss: 0.470
[62, 1800] loss: 0.509
[62, 1830] loss: 0.423
[62, 1860] loss: 0.426
[62, 1890] loss: 0.477
[62, 1920] loss: 0.482
[62, 1950] loss: 0.434
[62, 1980] loss: 0.500
[62, 2010] loss: 0.492
[62, 2040] loss: 0.480
[62, 2070] loss: 0.429
[62, 2100] loss: 0.439
[62, 2130] loss: 0.425
[62, 2160] loss: 0.413
[62, 2190] loss: 0.412
ovft intermediate models/epoch61_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.912873
epoch 61 accuracy on train set is: 0.9128725917848055
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	NaN	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.915304

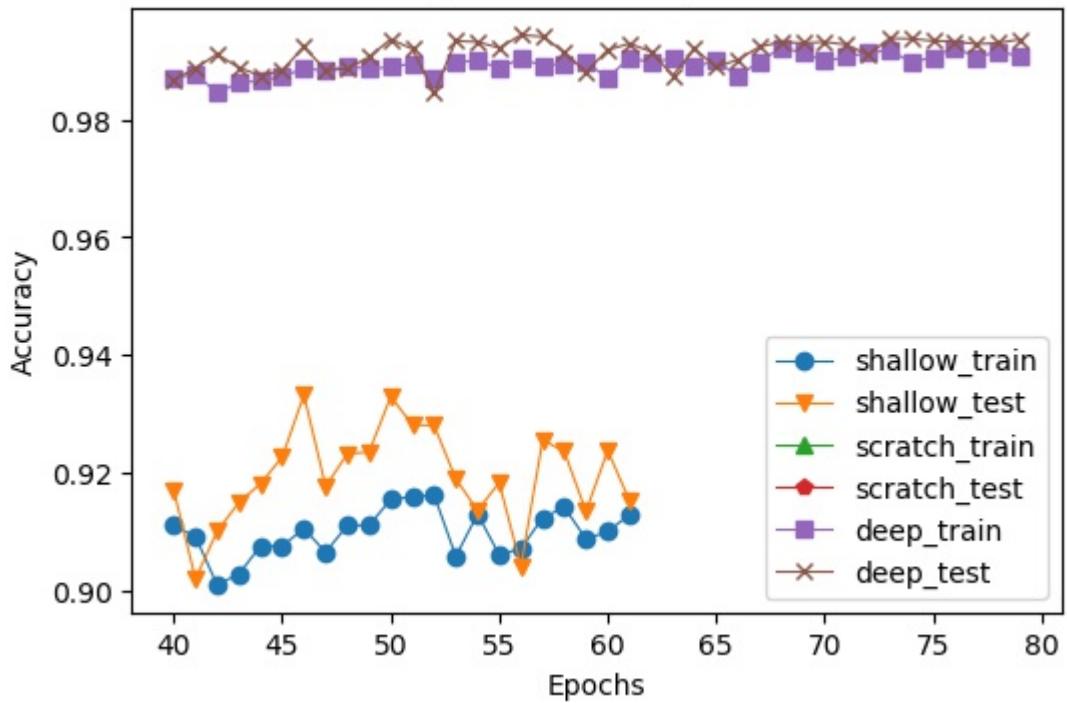
epoch 61 accuracy on test set is: 0.9153035259905489

evaluating on test set takes 16 seconds.

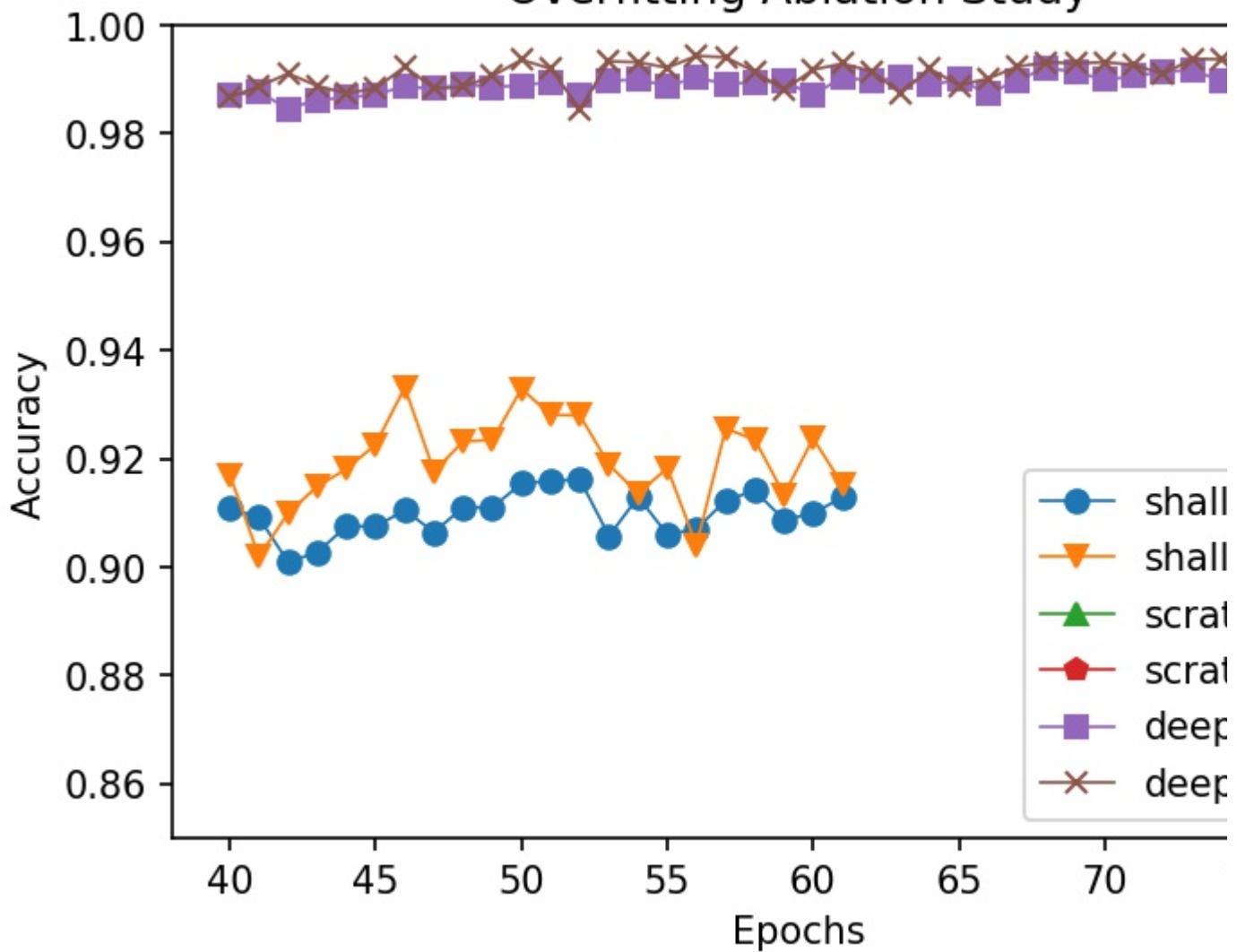
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	NaN	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 61 takes 149 seconds.  
[63,    30] loss: 0.457  
[63,    60] loss: 0.487  
[63,    90] loss: 0.429  
[63,   120] loss: 0.540  
[63,   150] loss: 0.414  
[63,   180] loss: 0.422  
[63,   210] loss: 0.408  
[63,   240] loss: 0.419  
[63,   270] loss: 0.455  
[63,   300] loss: 0.493  
[63,   330] loss: 0.419  
[63,   360] loss: 0.473  
[63,   390] loss: 0.399  
[63,   420] loss: 0.427  
[63,   450] loss: 0.523  
[63,   480] loss: 0.435  
[63,   510] loss: 0.339  
[63,   540] loss: 0.525  
[63,   570] loss: 0.570  
[63,   600] loss: 0.445  
[63,   630] loss: 0.408  
[63,   660] loss: 0.541
```

```

[63, 690] loss: 0.463
[63, 720] loss: 0.506
[63, 750] loss: 0.418
[63, 780] loss: 0.452
[63, 810] loss: 0.500
[63, 840] loss: 0.461
[63, 870] loss: 0.450
[63, 900] loss: 0.510
[63, 930] loss: 0.463
[63, 960] loss: 0.429
[63, 990] loss: 0.563
[63, 1020] loss: 0.407
[63, 1050] loss: 0.488
[63, 1080] loss: 0.413
[63, 1110] loss: 0.492
[63, 1140] loss: 0.571
[63, 1170] loss: 0.443
[63, 1200] loss: 0.559
[63, 1230] loss: 0.422
[63, 1260] loss: 0.472
[63, 1290] loss: 0.513
[63, 1320] loss: 0.430
[63, 1350] loss: 0.522
[63, 1380] loss: 0.416
[63, 1410] loss: 0.595
[63, 1440] loss: 0.522
[63, 1470] loss: 0.449
[63, 1500] loss: 0.489
[63, 1530] loss: 0.521
[63, 1560] loss: 0.451
[63, 1590] loss: 0.449
[63, 1620] loss: 0.464
[63, 1650] loss: 0.485
[63, 1680] loss: 0.477
[63, 1710] loss: 0.410
[63, 1740] loss: 0.393
[63, 1770] loss: 0.405
[63, 1800] loss: 0.495
[63, 1830] loss: 0.379
[63, 1860] loss: 0.447
[63, 1890] loss: 0.436
[63, 1920] loss: 0.367
[63, 1950] loss: 0.422
[63, 1980] loss: 0.409
[63, 2010] loss: 0.473
[63, 2040] loss: 0.415
[63, 2070] loss: 0.383
[63, 2100] loss: 0.488
[63, 2130] loss: 0.511
[63, 2160] loss: 0.429
[63, 2190] loss: 0.424
ovft_intermediate_models/epoch62_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.910283
epoch 62 accuracy on train set is: 0.9102826245001817
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	NaN	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.918848

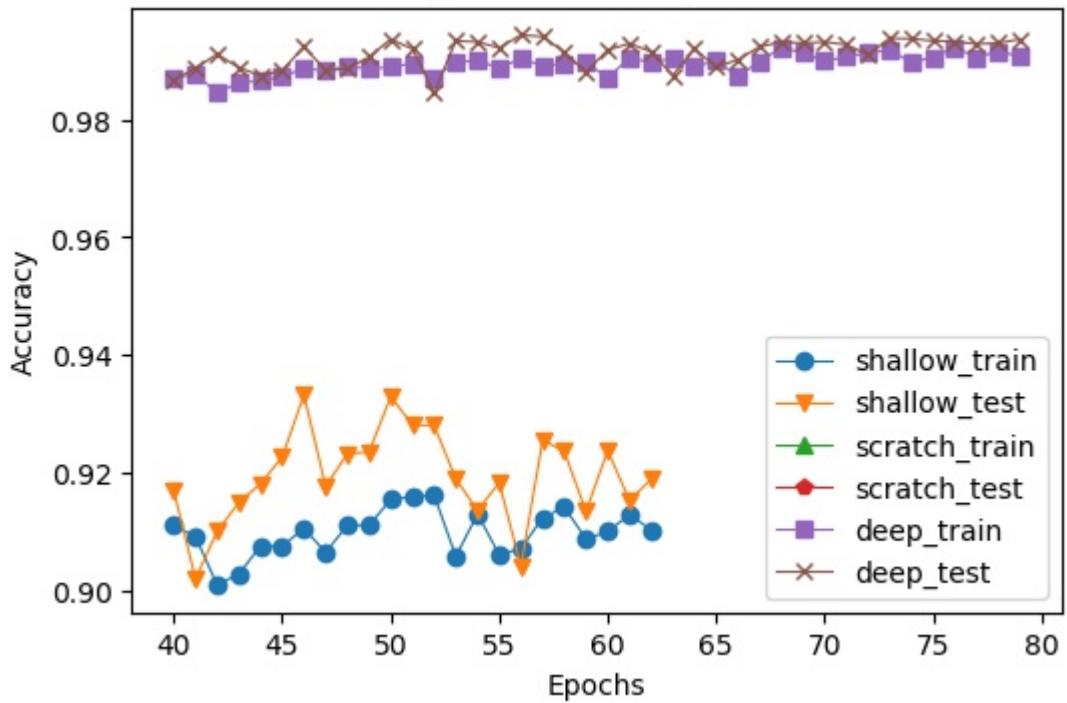
epoch 62 accuracy on test set is: 0.9188476917484552

evaluating on test set takes 16 seconds.

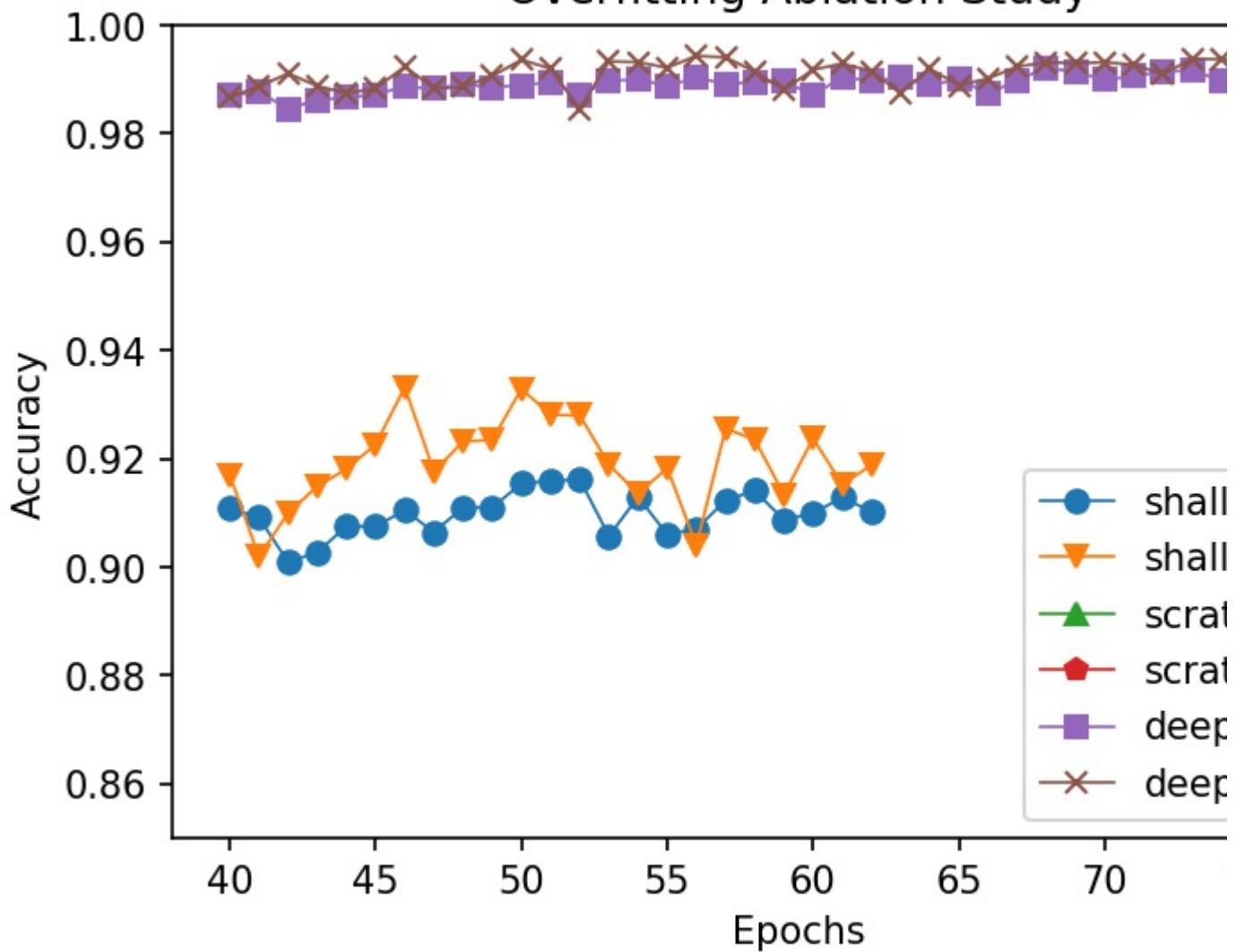
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	NaN	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 62 takes 149 seconds.  
[64,      30] loss: 0.454  
[64,      60] loss: 0.490  
[64,     90] loss: 0.505  
[64,    120] loss: 0.324  
[64,    150] loss: 0.475  
[64,    180] loss: 0.390  
[64,    210] loss: 0.516  
[64,    240] loss: 0.471  
[64,    270] loss: 0.496  
[64,    300] loss: 0.445  
[64,    330] loss: 0.468  
[64,    360] loss: 0.450  
[64,    390] loss: 0.568  
[64,    420] loss: 0.409  
[64,    450] loss: 0.418  
[64,    480] loss: 0.500  
[64,    510] loss: 0.457  
[64,    540] loss: 0.404  
[64,    570] loss: 0.440  
[64,    600] loss: 0.479  
[64,    630] loss: 0.452  
[64,    660] loss: 0.472
```

```

[64, 690] loss: 0.459
[64, 720] loss: 0.396
[64, 750] loss: 0.449
[64, 780] loss: 0.448
[64, 810] loss: 0.452
[64, 840] loss: 0.397
[64, 870] loss: 0.459
[64, 900] loss: 0.386
[64, 930] loss: 0.390
[64, 960] loss: 0.402
[64, 990] loss: 0.479
[64, 1020] loss: 0.478
[64, 1050] loss: 0.448
[64, 1080] loss: 0.425
[64, 1110] loss: 0.486
[64, 1140] loss: 0.410
[64, 1170] loss: 0.404
[64, 1200] loss: 0.417
[64, 1230] loss: 0.437
[64, 1260] loss: 0.525
[64, 1290] loss: 0.483
[64, 1320] loss: 0.497
[64, 1350] loss: 0.499
[64, 1380] loss: 0.453
[64, 1410] loss: 0.444
[64, 1440] loss: 0.521
[64, 1470] loss: 0.453
[64, 1500] loss: 0.434
[64, 1530] loss: 0.573
[64, 1560] loss: 0.473
[64, 1590] loss: 0.504
[64, 1620] loss: 0.504
[64, 1650] loss: 0.428
[64, 1680] loss: 0.500
[64, 1710] loss: 0.449
[64, 1740] loss: 0.500
[64, 1770] loss: 0.492
[64, 1800] loss: 0.465
[64, 1830] loss: 0.412
[64, 1860] loss: 0.470
[64, 1890] loss: 0.452
[64, 1920] loss: 0.548
[64, 1950] loss: 0.488
[64, 1980] loss: 0.475
[64, 2010] loss: 0.477
[64, 2040] loss: 0.467
[64, 2070] loss: 0.601
[64, 2100] loss: 0.458
[64, 2130] loss: 0.454
[64, 2160] loss: 0.475
[64, 2190] loss: 0.458
ovft_intermediate_models/epoch63_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.913372
epoch 63 accuracy on train set is: 0.9133724100327154
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	NaN	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.920756

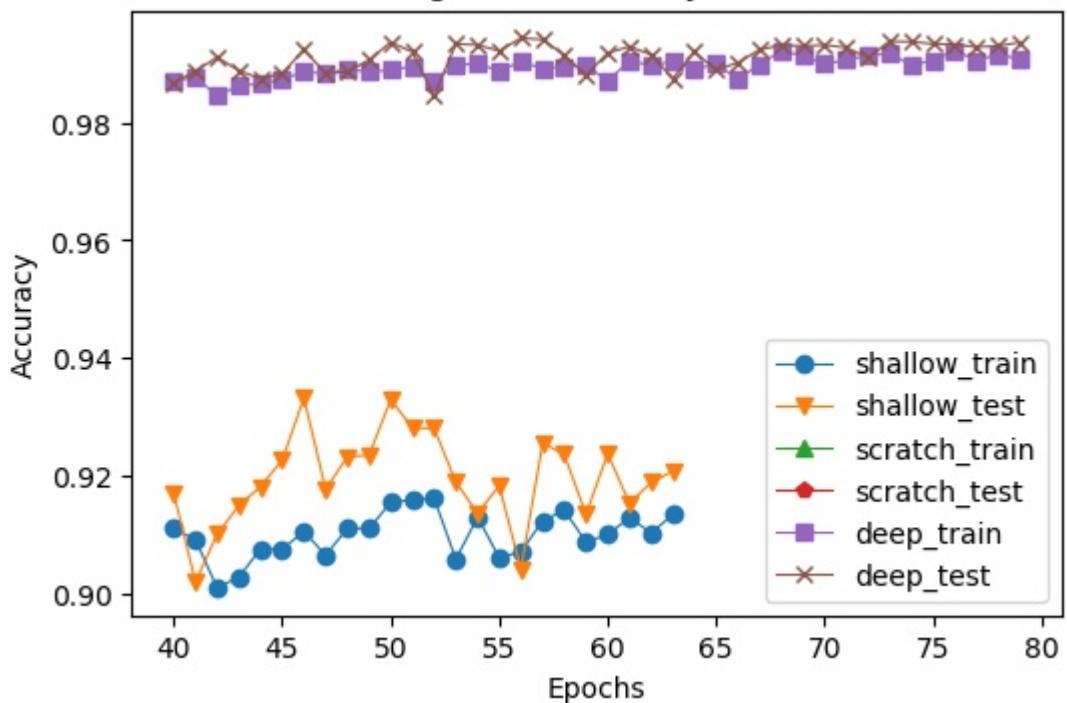
epoch 63 accuracy on test set is: 0.92075608869502

evaluating on test set takes 16 seconds.

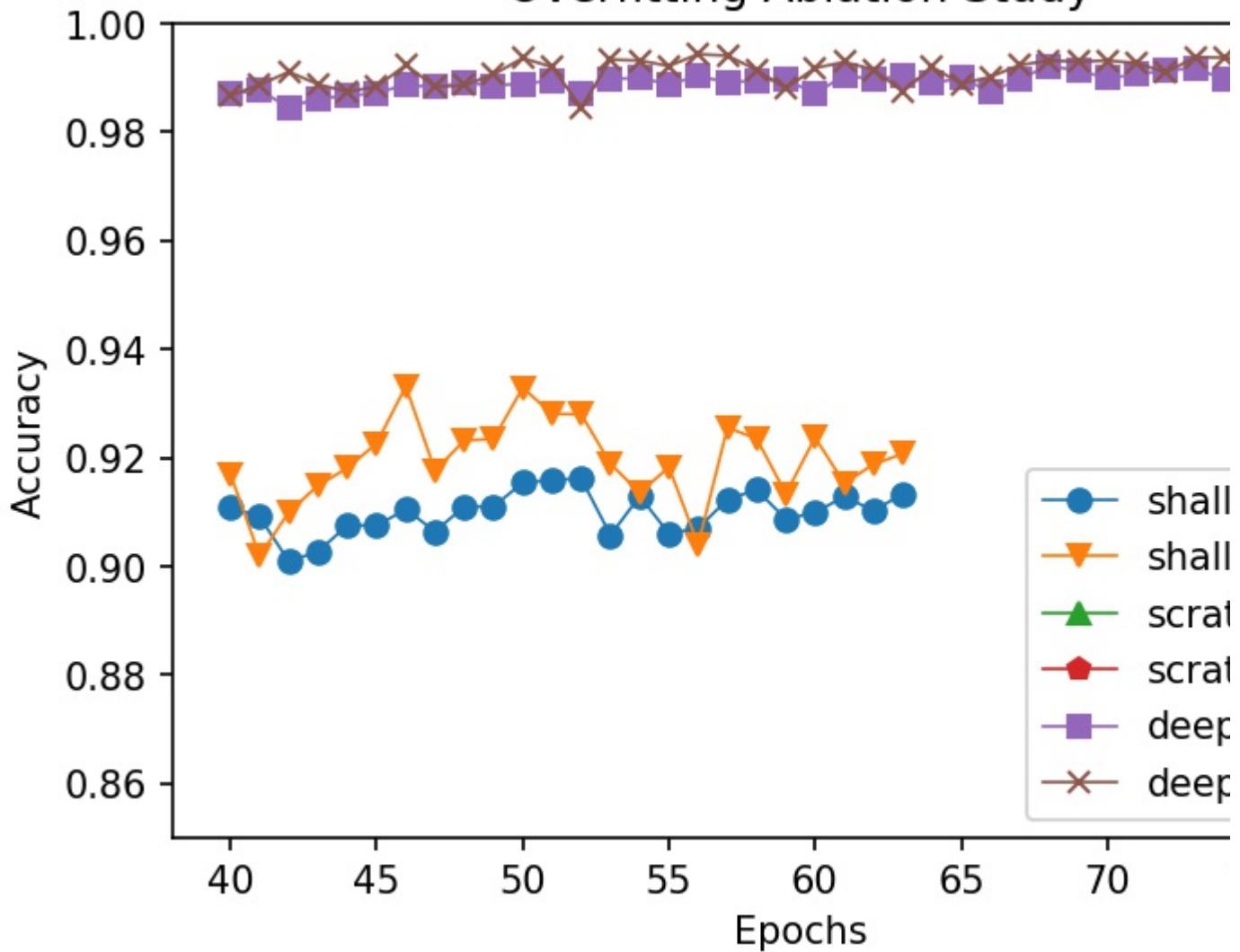
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	NaN	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 63 takes 149 seconds.  
[65,    30] loss: 0.479  
[65,    60] loss: 0.426  
[65,    90] loss: 0.495  
[65,   120] loss: 0.476  
[65,   150] loss: 0.487  
[65,   180] loss: 0.361  
[65,   210] loss: 0.394  
[65,   240] loss: 0.421  
[65,   270] loss: 0.446  
[65,   300] loss: 0.491  
[65,   330] loss: 0.564  
[65,   360] loss: 0.436  
[65,   390] loss: 0.390  
[65,   420] loss: 0.443  
[65,   450] loss: 0.434  
[65,   480] loss: 0.473  
[65,   510] loss: 0.556  
[65,   540] loss: 0.502  
[65,   570] loss: 0.479  
[65,   600] loss: 0.488  
[65,   630] loss: 0.472  
[65,   660] loss: 0.445
```

```

[65, 690] loss: 0.495
[65, 720] loss: 0.529
[65, 750] loss: 0.392
[65, 780] loss: 0.408
[65, 810] loss: 0.462
[65, 840] loss: 0.432
[65, 870] loss: 0.359
[65, 900] loss: 0.494
[65, 930] loss: 0.458
[65, 960] loss: 0.456
[65, 990] loss: 0.428
[65, 1020] loss: 0.418
[65, 1050] loss: 0.439
[65, 1080] loss: 0.385
[65, 1110] loss: 0.460
[65, 1140] loss: 0.462
[65, 1170] loss: 0.433
[65, 1200] loss: 0.496
[65, 1230] loss: 0.428
[65, 1260] loss: 0.549
[65, 1290] loss: 0.510
[65, 1320] loss: 0.446
[65, 1350] loss: 0.441
[65, 1380] loss: 0.449
[65, 1410] loss: 0.387
[65, 1440] loss: 0.393
[65, 1470] loss: 0.446
[65, 1500] loss: 0.442
[65, 1530] loss: 0.369
[65, 1560] loss: 0.501
[65, 1590] loss: 0.420
[65, 1620] loss: 0.496
[65, 1650] loss: 0.410
[65, 1680] loss: 0.446
[65, 1710] loss: 0.464
[65, 1740] loss: 0.492
[65, 1770] loss: 0.449
[65, 1800] loss: 0.395
[65, 1830] loss: 0.403
[65, 1860] loss: 0.409
[65, 1890] loss: 0.378
[65, 1920] loss: 0.488
[65, 1950] loss: 0.416
[65, 1980] loss: 0.432
[65, 2010] loss: 0.512
[65, 2040] loss: 0.518
[65, 2070] loss: 0.394
[65, 2100] loss: 0.373
[65, 2130] loss: 0.426
[65, 2160] loss: 0.510
[65, 2190] loss: 0.443
ovft intermediate models/epoch64_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.909556
epoch 64 accuracy on train set is: 0.9095556161395856
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	NaN	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.919211

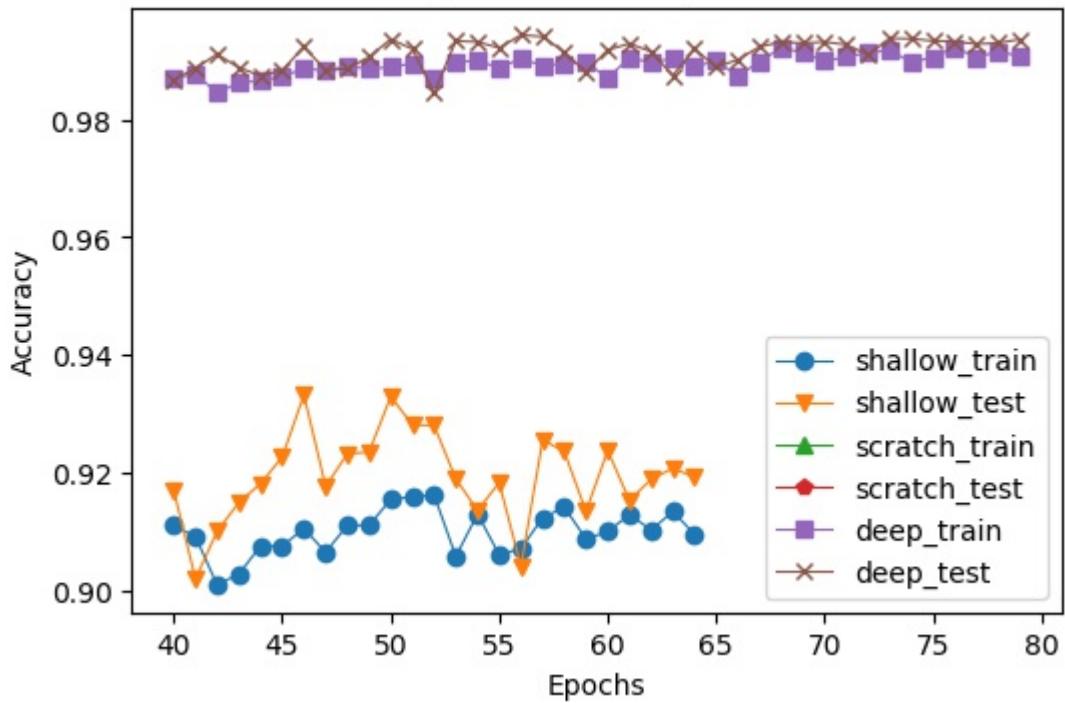
epoch 64 accuracy on test set is: 0.9192111959287532

evaluating on test set takes 16 seconds.

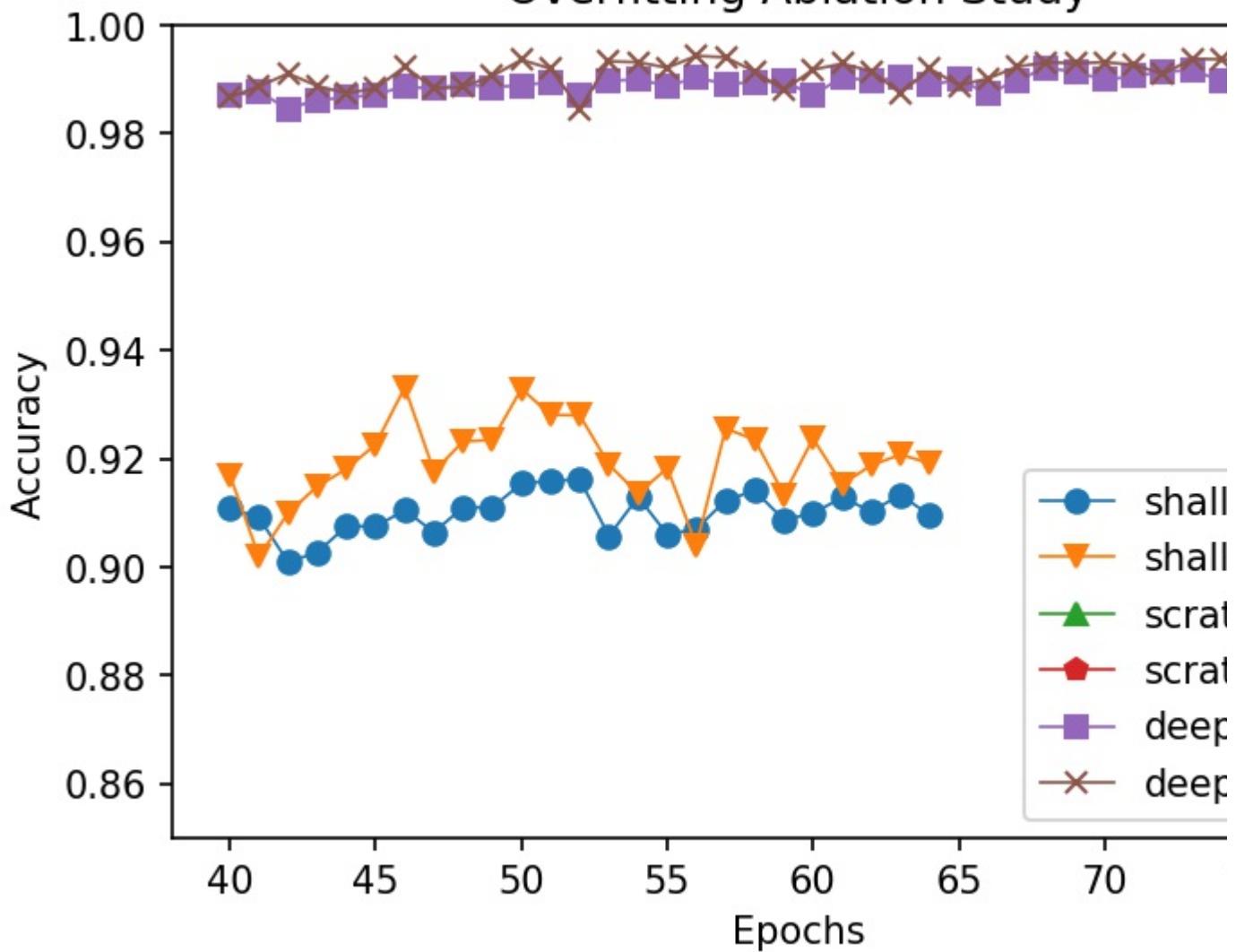
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	NaN	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 64 takes 149 seconds.  
[66,    30] loss: 0.346  
[66,    60] loss: 0.484  
[66,    90] loss: 0.452  
[66,   120] loss: 0.406  
[66,   150] loss: 0.492  
[66,   180] loss: 0.433  
[66,   210] loss: 0.349  
[66,   240] loss: 0.427  
[66,   270] loss: 0.448  
[66,   300] loss: 0.373  
[66,   330] loss: 0.392  
[66,   360] loss: 0.441  
[66,   390] loss: 0.418  
[66,   420] loss: 0.533  
[66,   450] loss: 0.558  
[66,   480] loss: 0.437  
[66,   510] loss: 0.457  
[66,   540] loss: 0.489  
[66,   570] loss: 0.519  
[66,   600] loss: 0.455  
[66,   630] loss: 0.451  
[66,   660] loss: 0.545
```

```

[66, 690] loss: 0.468
[66, 720] loss: 0.453
[66, 750] loss: 0.473
[66, 780] loss: 0.486
[66, 810] loss: 0.438
[66, 840] loss: 0.391
[66, 870] loss: 0.389
[66, 900] loss: 0.428
[66, 930] loss: 0.470
[66, 960] loss: 0.464
[66, 990] loss: 0.493
[66, 1020] loss: 0.461
[66, 1050] loss: 0.363
[66, 1080] loss: 0.535
[66, 1110] loss: 0.419
[66, 1140] loss: 0.407
[66, 1170] loss: 0.466
[66, 1200] loss: 0.519
[66, 1230] loss: 0.429
[66, 1260] loss: 0.466
[66, 1290] loss: 0.446
[66, 1320] loss: 0.453
[66, 1350] loss: 0.408
[66, 1380] loss: 0.451
[66, 1410] loss: 0.367
[66, 1440] loss: 0.476
[66, 1470] loss: 0.483
[66, 1500] loss: 0.403
[66, 1530] loss: 0.456
[66, 1560] loss: 0.436
[66, 1590] loss: 0.438
[66, 1620] loss: 0.446
[66, 1650] loss: 0.439
[66, 1680] loss: 0.526
[66, 1710] loss: 0.477
[66, 1740] loss: 0.416
[66, 1770] loss: 0.535
[66, 1800] loss: 0.454
[66, 1830] loss: 0.538
[66, 1860] loss: 0.534
[66, 1890] loss: 0.441
[66, 1920] loss: 0.435
[66, 1950] loss: 0.391
[66, 1980] loss: 0.483
[66, 2010] loss: 0.425
[66, 2040] loss: 0.429
[66, 2070] loss: 0.491
[66, 2100] loss: 0.461
[66, 2130] loss: 0.476
[66, 2160] loss: 0.431
[66, 2190] loss: 0.422
ovft_intermediate_models/epoch65_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.913713
epoch 65 accuracy on train set is: 0.9137131952017448
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	NaN	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.933660

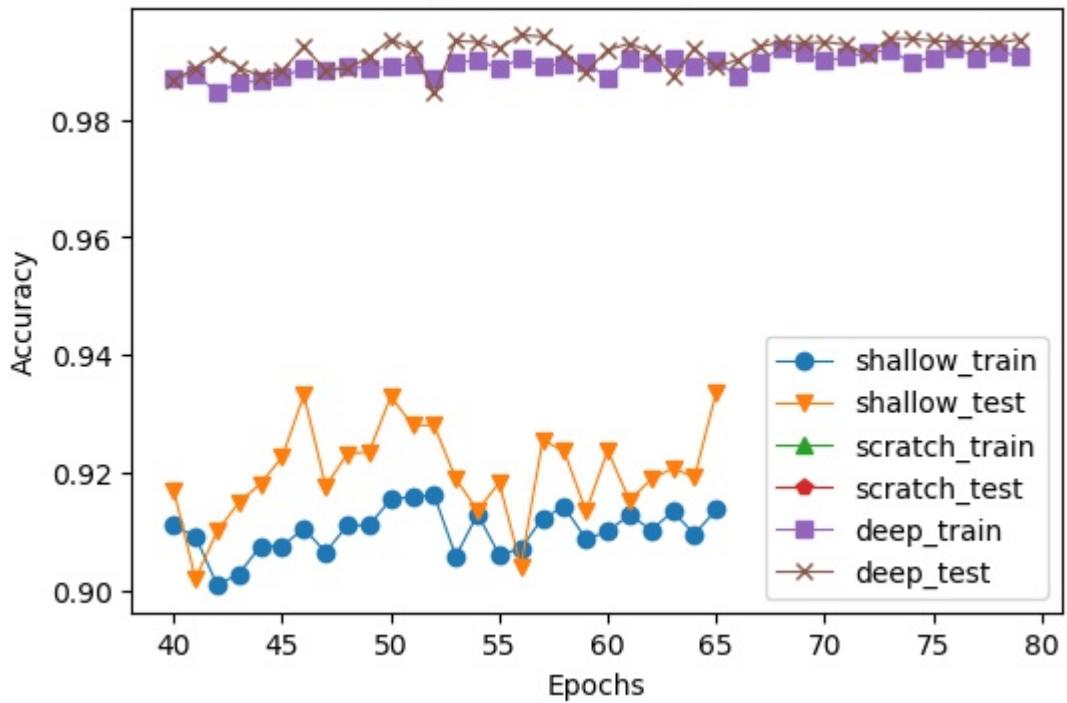
epoch 65 accuracy on test set is: 0.9336604870956016

evaluating on test set takes 16 seconds.

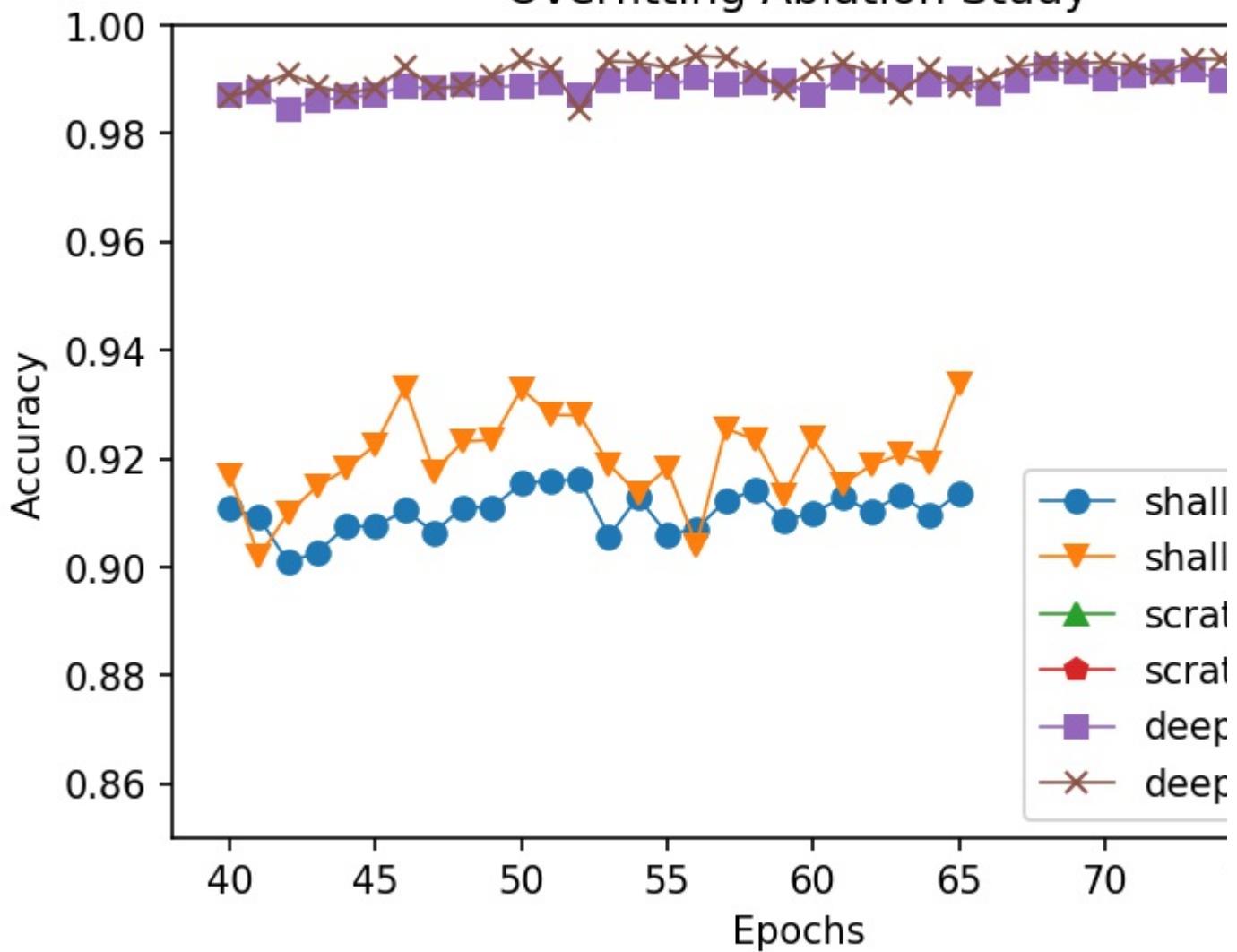
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	NaN	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 65 takes 149 seconds.  
[67,    30] loss: 0.390  
[67,    60] loss: 0.481  
[67,    90] loss: 0.479  
[67,   120] loss: 0.467  
[67,   150] loss: 0.397  
[67,   180] loss: 0.370  
[67,   210] loss: 0.414  
[67,   240] loss: 0.402  
[67,   270] loss: 0.412  
[67,   300] loss: 0.463  
[67,   330] loss: 0.411  
[67,   360] loss: 0.404  
[67,   390] loss: 0.430  
[67,   420] loss: 0.471  
[67,   450] loss: 0.410  
[67,   480] loss: 0.489  
[67,   510] loss: 0.367  
[67,   540] loss: 0.486  
[67,   570] loss: 0.445  
[67,   600] loss: 0.480  
[67,   630] loss: 0.489  
[67,   660] loss: 0.411
```

```

[67, 690] loss: 0.454
[67, 720] loss: 0.495
[67, 750] loss: 0.426
[67, 780] loss: 0.512
[67, 810] loss: 0.427
[67, 840] loss: 0.480
[67, 870] loss: 0.413
[67, 900] loss: 0.445
[67, 930] loss: 0.409
[67, 960] loss: 0.412
[67, 990] loss: 0.492
[67, 1020] loss: 0.482
[67, 1050] loss: 0.372
[67, 1080] loss: 0.423
[67, 1110] loss: 0.523
[67, 1140] loss: 0.382
[67, 1170] loss: 0.549
[67, 1200] loss: 0.450
[67, 1230] loss: 0.465
[67, 1260] loss: 0.502
[67, 1290] loss: 0.406
[67, 1320] loss: 0.559
[67, 1350] loss: 0.439
[67, 1380] loss: 0.476
[67, 1410] loss: 0.560
[67, 1440] loss: 0.555
[67, 1470] loss: 0.427
[67, 1500] loss: 0.465
[67, 1530] loss: 0.473
[67, 1560] loss: 0.474
[67, 1590] loss: 0.412
[67, 1620] loss: 0.468
[67, 1650] loss: 0.440
[67, 1680] loss: 0.426
[67, 1710] loss: 0.498
[67, 1740] loss: 0.438
[67, 1770] loss: 0.464
[67, 1800] loss: 0.424
[67, 1830] loss: 0.431
[67, 1860] loss: 0.440
[67, 1890] loss: 0.416
[67, 1920] loss: 0.524
[67, 1950] loss: 0.473
[67, 1980] loss: 0.444
[67, 2010] loss: 0.441
[67, 2040] loss: 0.456
[67, 2070] loss: 0.481
[67, 2100] loss: 0.434
[67, 2130] loss: 0.433
[67, 2160] loss: 0.518
[67, 2190] loss: 0.515
ovft_intermediate_models/epoch66_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.910555
epoch 66 accuracy on train set is: 0.9105552526354053
evaluating on training set takes 66 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	NaN	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.923210

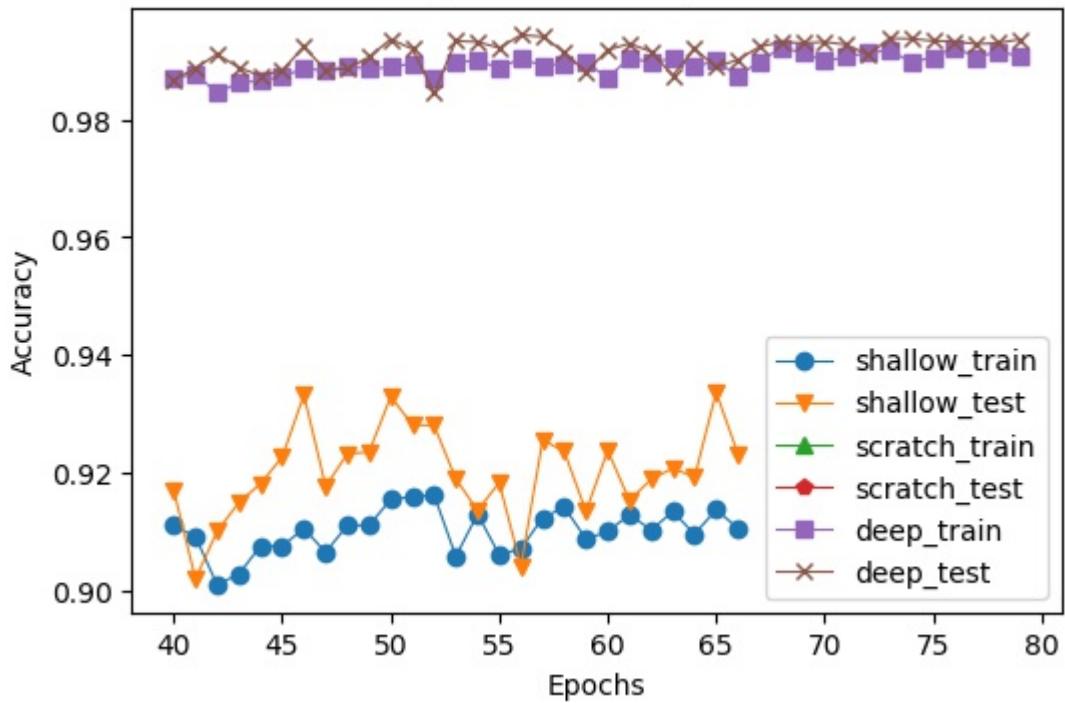
epoch 66 accuracy on test set is: 0.923209741912032

evaluating on test set takes 16 seconds.

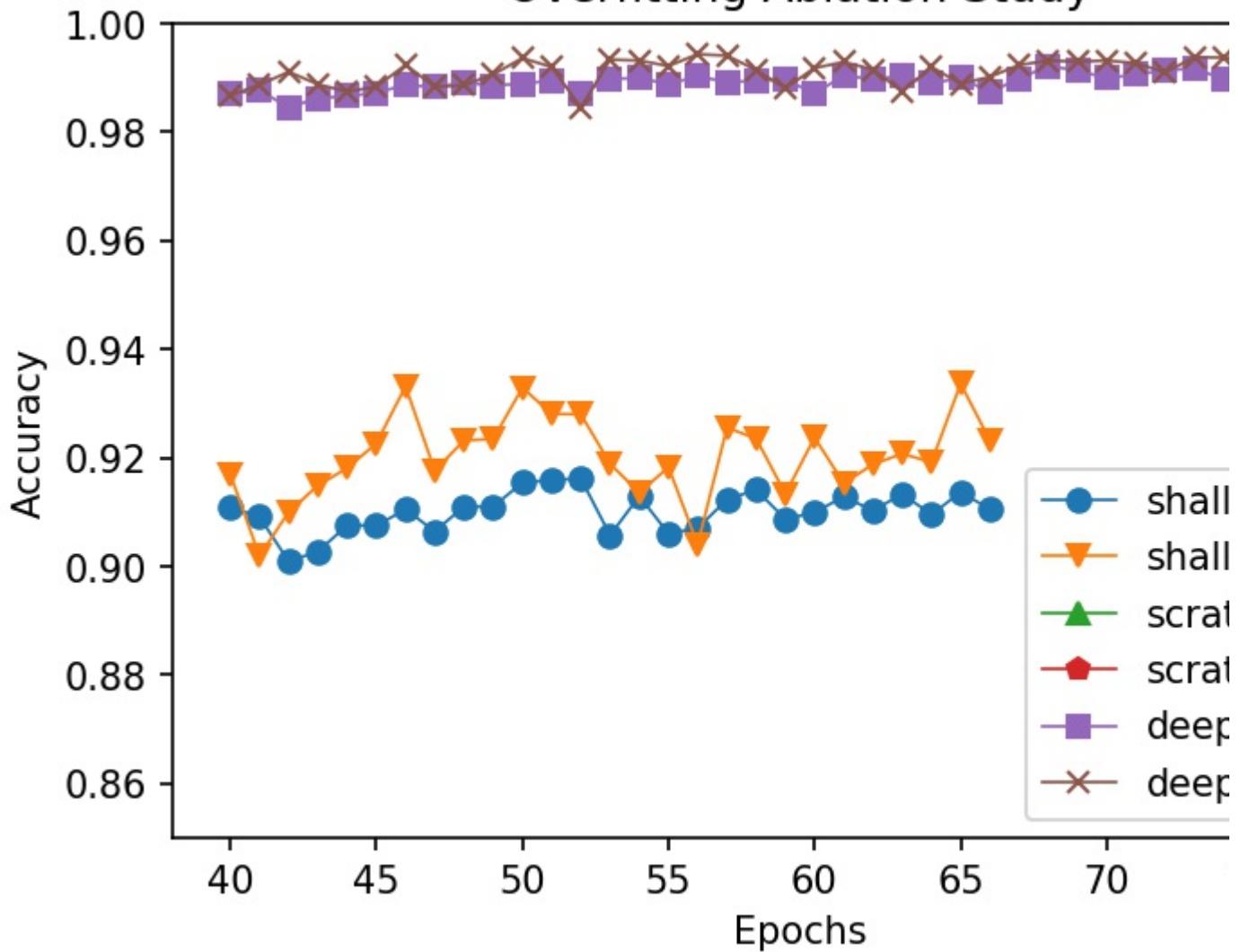
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	NaN	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 66 takes 153 seconds.  
[68,    30] loss: 0.556  
[68,    60] loss: 0.394  
[68,    90] loss: 0.488  
[68,   120] loss: 0.439  
[68,   150] loss: 0.516  
[68,   180] loss: 0.461  
[68,   210] loss: 0.532  
[68,   240] loss: 0.461  
[68,   270] loss: 0.487  
[68,   300] loss: 0.473  
[68,   330] loss: 0.354  
[68,   360] loss: 0.364  
[68,   390] loss: 0.396  
[68,   420] loss: 0.464  
[68,   450] loss: 0.441  
[68,   480] loss: 0.477  
[68,   510] loss: 0.491  
[68,   540] loss: 0.485  
[68,   570] loss: 0.437  
[68,   600] loss: 0.459  
[68,   630] loss: 0.435  
[68,   660] loss: 0.471
```

```

[68, 690] loss: 0.505
[68, 720] loss: 0.475
[68, 750] loss: 0.535
[68, 780] loss: 0.527
[68, 810] loss: 0.457
[68, 840] loss: 0.428
[68, 870] loss: 0.453
[68, 900] loss: 0.409
[68, 930] loss: 0.397
[68, 960] loss: 0.483
[68, 990] loss: 0.460
[68, 1020] loss: 0.432
[68, 1050] loss: 0.387
[68, 1080] loss: 0.445
[68, 1110] loss: 0.497
[68, 1140] loss: 0.427
[68, 1170] loss: 0.415
[68, 1200] loss: 0.466
[68, 1230] loss: 0.458
[68, 1260] loss: 0.389
[68, 1290] loss: 0.498
[68, 1320] loss: 0.530
[68, 1350] loss: 0.510
[68, 1380] loss: 0.544
[68, 1410] loss: 0.473
[68, 1440] loss: 0.388
[68, 1470] loss: 0.494
[68, 1500] loss: 0.446
[68, 1530] loss: 0.447
[68, 1560] loss: 0.485
[68, 1590] loss: 0.500
[68, 1620] loss: 0.436
[68, 1650] loss: 0.495
[68, 1680] loss: 0.488
[68, 1710] loss: 0.396
[68, 1740] loss: 0.393
[68, 1770] loss: 0.482
[68, 1800] loss: 0.386
[68, 1830] loss: 0.494
[68, 1860] loss: 0.410
[68, 1890] loss: 0.527
[68, 1920] loss: 0.433
[68, 1950] loss: 0.465
[68, 1980] loss: 0.414
[68, 2010] loss: 0.440
[68, 2040] loss: 0.409
[68, 2070] loss: 0.485
[68, 2100] loss: 0.512
[68, 2130] loss: 0.431
[68, 2160] loss: 0.440
[68, 2190] loss: 0.433
ovft_intermediate_models/epoch67_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.909919
epoch 67 accuracy on train set is: 0.9099191203198836
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	NaN	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.928299

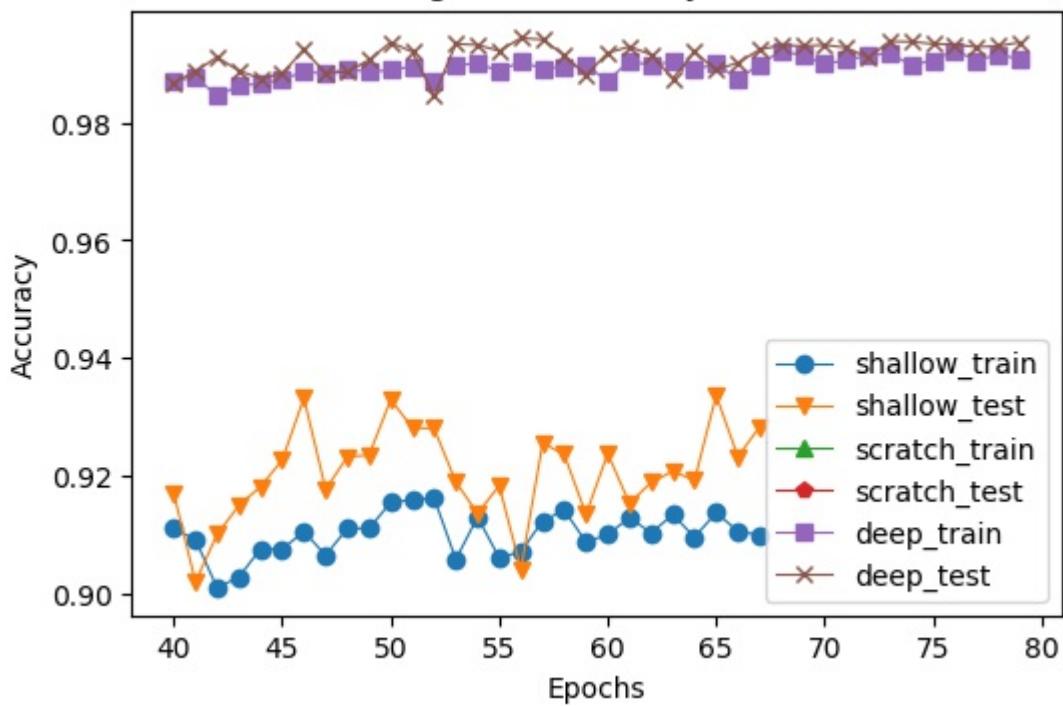
epoch 67 accuracy on test set is: 0.928298800436205

evaluating on test set takes 16 seconds.

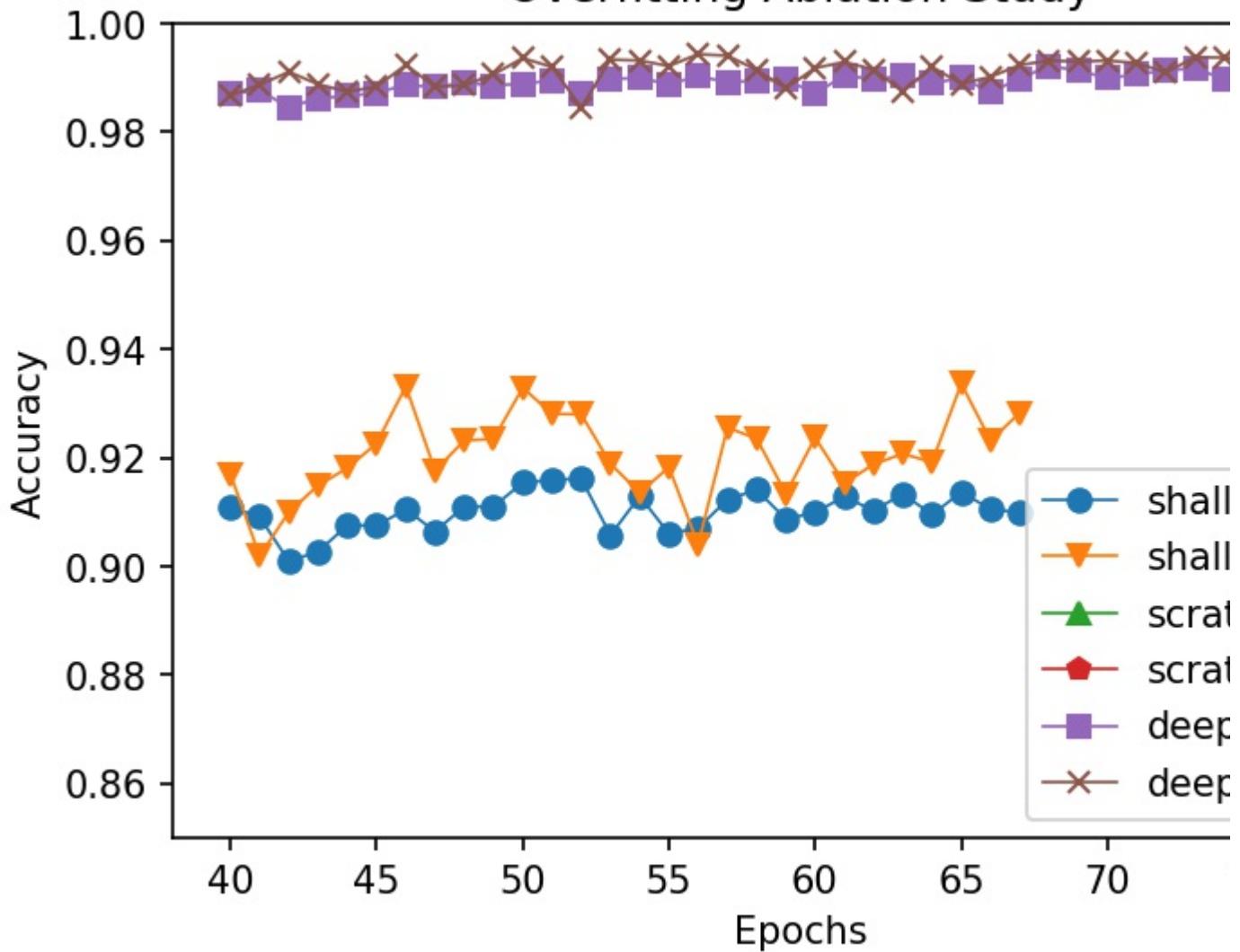
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	NaN	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 67 takes 150 seconds.  
[69,    30] loss: 0.477  
[69,    60] loss: 0.430  
[69,    90] loss: 0.406  
[69,   120] loss: 0.467  
[69,   150] loss: 0.405  
[69,   180] loss: 0.549  
[69,   210] loss: 0.411  
[69,   240] loss: 0.420  
[69,   270] loss: 0.394  
[69,   300] loss: 0.473  
[69,   330] loss: 0.419  
[69,   360] loss: 0.481  
[69,   390] loss: 0.439  
[69,   420] loss: 0.474  
[69,   450] loss: 0.564  
[69,   480] loss: 0.507  
[69,   510] loss: 0.420  
[69,   540] loss: 0.410  
[69,   570] loss: 0.415  
[69,   600] loss: 0.452  
[69,   630] loss: 0.563  
[69,   660] loss: 0.494
```

```

[69, 690] loss: 0.442
[69, 720] loss: 0.440
[69, 750] loss: 0.468
[69, 780] loss: 0.447
[69, 810] loss: 0.501
[69, 840] loss: 0.485
[69, 870] loss: 0.453
[69, 900] loss: 0.432
[69, 930] loss: 0.450
[69, 960] loss: 0.461
[69, 990] loss: 0.422
[69, 1020] loss: 0.518
[69, 1050] loss: 0.433
[69, 1080] loss: 0.508
[69, 1110] loss: 0.471
[69, 1140] loss: 0.504
[69, 1170] loss: 0.508
[69, 1200] loss: 0.432
[69, 1230] loss: 0.442
[69, 1260] loss: 0.453
[69, 1290] loss: 0.424
[69, 1320] loss: 0.484
[69, 1350] loss: 0.476
[69, 1380] loss: 0.408
[69, 1410] loss: 0.461
[69, 1440] loss: 0.455
[69, 1470] loss: 0.545
[69, 1500] loss: 0.453
[69, 1530] loss: 0.424
[69, 1560] loss: 0.517
[69, 1590] loss: 0.443
[69, 1620] loss: 0.391
[69, 1650] loss: 0.429
[69, 1680] loss: 0.450
[69, 1710] loss: 0.460
[69, 1740] loss: 0.511
[69, 1770] loss: 0.548
[69, 1800] loss: 0.457
[69, 1830] loss: 0.443
[69, 1860] loss: 0.423
[69, 1890] loss: 0.518
[69, 1920] loss: 0.628
[69, 1950] loss: 0.478
[69, 1980] loss: 0.471
[69, 2010] loss: 0.445
[69, 2040] loss: 0.571
[69, 2070] loss: 0.434
[69, 2100] loss: 0.487
[69, 2130] loss: 0.439
[69, 2160] loss: 0.521
[69, 2190] loss: 0.433
ovft intermediate models/epoch68_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.916031
epoch 68 accuracy on train set is: 0.916030534351145
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	NaN	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.927754

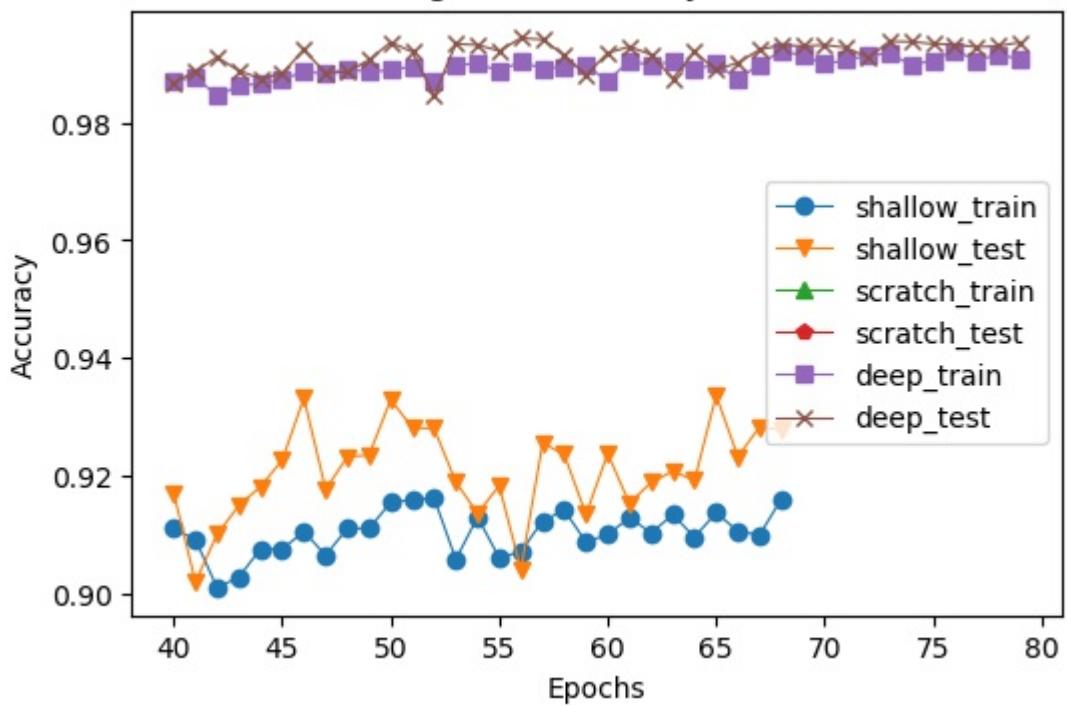
epoch 68 accuracy on test set is: 0.927753544165758

evaluating on test set takes 16 seconds.

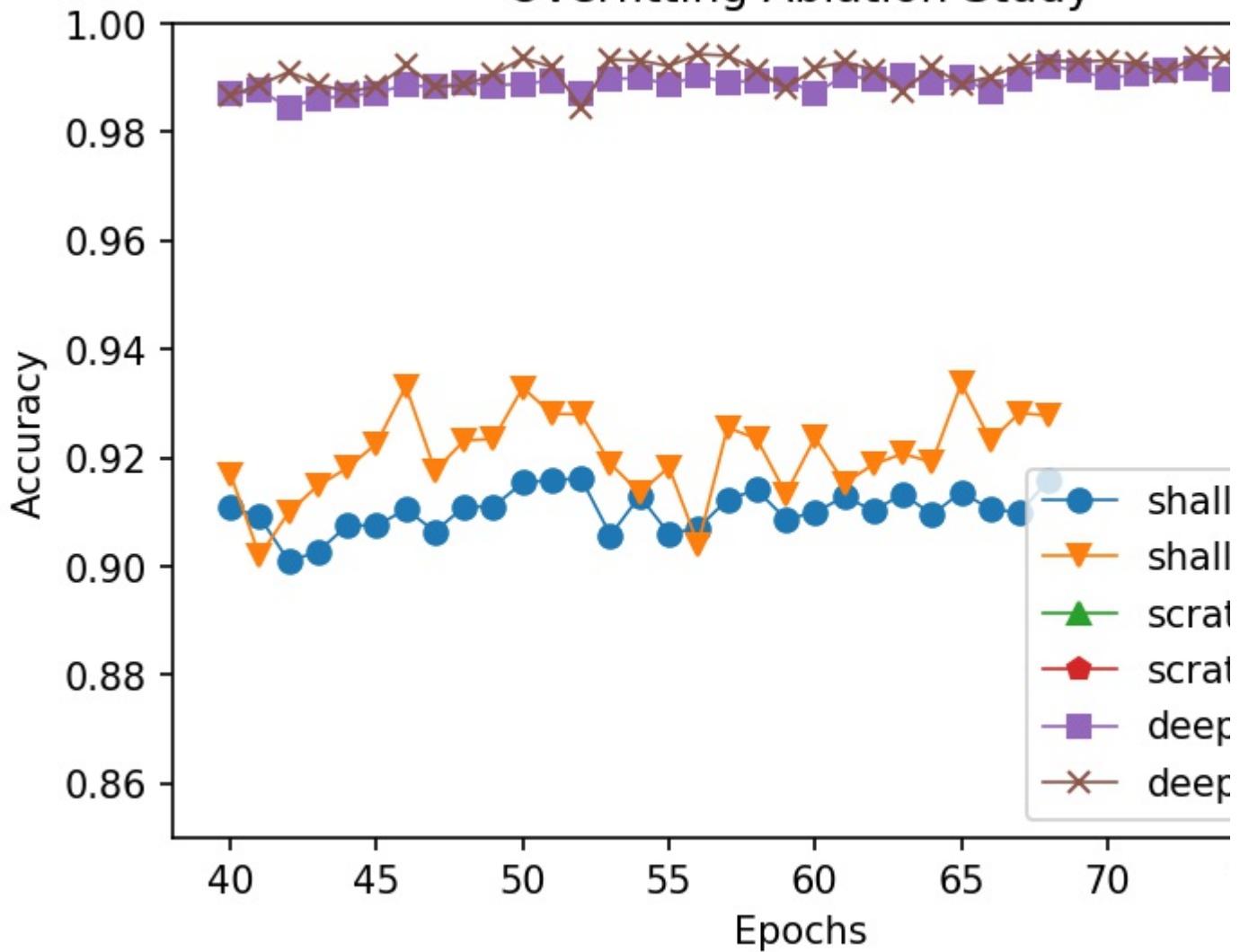
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	NaN	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 68 takes 151 seconds.  
[70,    30] loss: 0.515  
[70,    60] loss: 0.407  
[70,    90] loss: 0.392  
[70,   120] loss: 0.398  
[70,   150] loss: 0.536  
[70,   180] loss: 0.547  
[70,   210] loss: 0.484  
[70,   240] loss: 0.399  
[70,   270] loss: 0.562  
[70,   300] loss: 0.542  
[70,   330] loss: 0.423  
[70,   360] loss: 0.439  
[70,   390] loss: 0.501  
[70,   420] loss: 0.502  
[70,   450] loss: 0.475  
[70,   480] loss: 0.456  
[70,   510] loss: 0.546  
[70,   540] loss: 0.394  
[70,   570] loss: 0.505  
[70,   600] loss: 0.418  
[70,   630] loss: 0.406  
[70,   660] loss: 0.459
```

```

[70, 690] loss: 0.392
[70, 720] loss: 0.427
[70, 750] loss: 0.414
[70, 780] loss: 0.455
[70, 810] loss: 0.453
[70, 840] loss: 0.487
[70, 870] loss: 0.504
[70, 900] loss: 0.415
[70, 930] loss: 0.476
[70, 960] loss: 0.447
[70, 990] loss: 0.477
[70, 1020] loss: 0.513
[70, 1050] loss: 0.504
[70, 1080] loss: 0.518
[70, 1110] loss: 0.424
[70, 1140] loss: 0.454
[70, 1170] loss: 0.540
[70, 1200] loss: 0.474
[70, 1230] loss: 0.499
[70, 1260] loss: 0.420
[70, 1290] loss: 0.421
[70, 1320] loss: 0.470
[70, 1350] loss: 0.442
[70, 1380] loss: 0.549
[70, 1410] loss: 0.450
[70, 1440] loss: 0.466
[70, 1470] loss: 0.496
[70, 1500] loss: 0.501
[70, 1530] loss: 0.398
[70, 1560] loss: 0.504
[70, 1590] loss: 0.417
[70, 1620] loss: 0.462
[70, 1650] loss: 0.456
[70, 1680] loss: 0.440
[70, 1710] loss: 0.414
[70, 1740] loss: 0.382
[70, 1770] loss: 0.432
[70, 1800] loss: 0.496
[70, 1830] loss: 0.447
[70, 1860] loss: 0.483
[70, 1890] loss: 0.465
[70, 1920] loss: 0.469
[70, 1950] loss: 0.457
[70, 1980] loss: 0.438
[70, 2010] loss: 0.407
[70, 2040] loss: 0.475
[70, 2070] loss: 0.440
[70, 2100] loss: 0.469
[70, 2130] loss: 0.584
[70, 2160] loss: 0.494
[70, 2190] loss: 0.465
ovft_intermediate_models/epoch69_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.911123
epoch 69 accuracy on train set is: 0.9111232279171211
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	NaN	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.927572

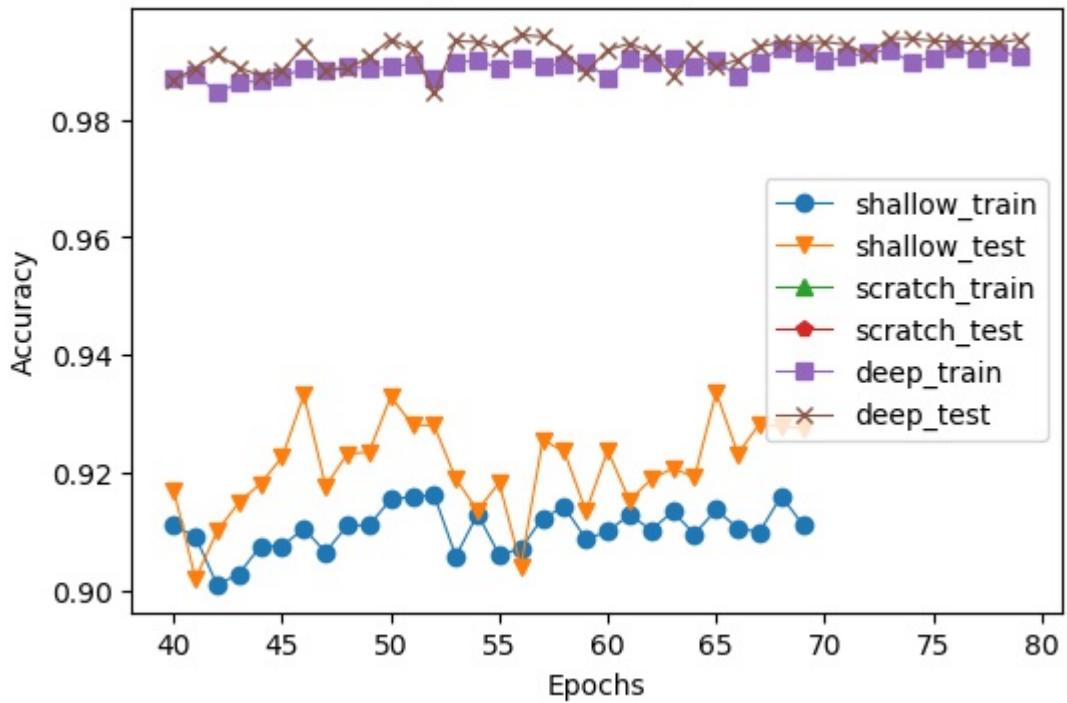
epoch 69 accuracy on test set is: 0.9275717920756089

evaluating on test set takes 16 seconds.

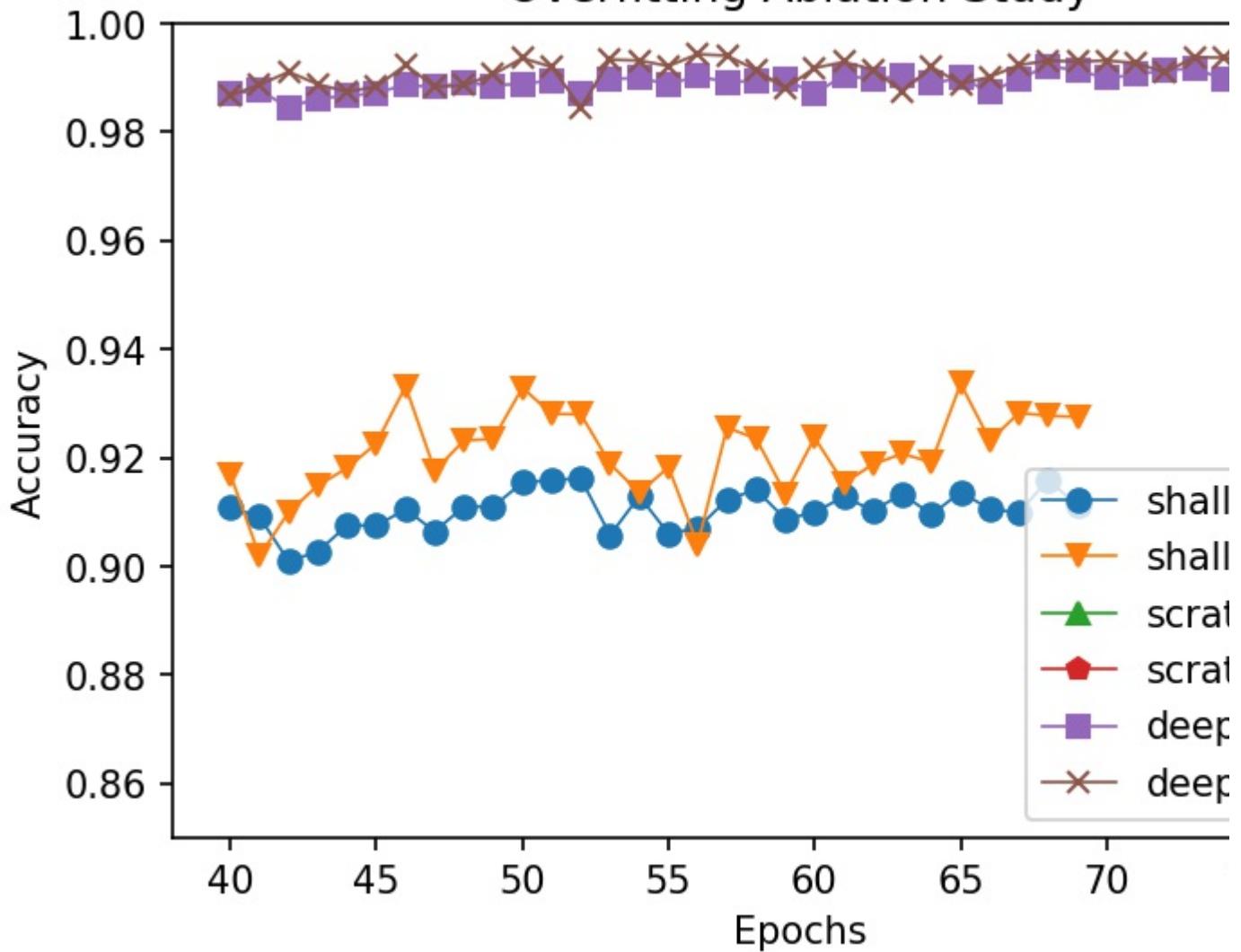
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	NaN	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



epoch 69 takes 150 seconds.

```
[71,    30] loss: 0.435
[71,    60] loss: 0.406
[71,    90] loss: 0.473
[71,   120] loss: 0.507
[71,   150] loss: 0.422
[71,   180] loss: 0.403
[71,   210] loss: 0.488
[71,   240] loss: 0.478
[71,   270] loss: 0.481
[71,   300] loss: 0.486
[71,   330] loss: 0.493
[71,   360] loss: 0.412
[71,   390] loss: 0.542
[71,   420] loss: 0.450
[71,   450] loss: 0.490
[71,   480] loss: 0.505
[71,   510] loss: 0.460
[71,   540] loss: 0.445
[71,   570] loss: 0.415
[71,   600] loss: 0.579
[71,   630] loss: 0.440
[71,   660] loss: 0.495
```

```

[71, 690] loss: 0.443
[71, 720] loss: 0.482
[71, 750] loss: 0.365
[71, 780] loss: 0.405
[71, 810] loss: 0.409
[71, 840] loss: 0.450
[71, 870] loss: 0.485
[71, 900] loss: 0.434
[71, 930] loss: 0.491
[71, 960] loss: 0.470
[71, 990] loss: 0.529
[71, 1020] loss: 0.543
[71, 1050] loss: 0.410
[71, 1080] loss: 0.463
[71, 1110] loss: 0.410
[71, 1140] loss: 0.450
[71, 1170] loss: 0.476
[71, 1200] loss: 0.547
[71, 1230] loss: 0.420
[71, 1260] loss: 0.425
[71, 1290] loss: 0.385
[71, 1320] loss: 0.461
[71, 1350] loss: 0.426
[71, 1380] loss: 0.480
[71, 1410] loss: 0.510
[71, 1440] loss: 0.531
[71, 1470] loss: 0.437
[71, 1500] loss: 0.495
[71, 1530] loss: 0.491
[71, 1560] loss: 0.475
[71, 1590] loss: 0.501
[71, 1620] loss: 0.450
[71, 1650] loss: 0.376
[71, 1680] loss: 0.513
[71, 1710] loss: 0.434
[71, 1740] loss: 0.514
[71, 1770] loss: 0.473
[71, 1800] loss: 0.510
[71, 1830] loss: 0.404
[71, 1860] loss: 0.534
[71, 1890] loss: 0.466
[71, 1920] loss: 0.512
[71, 1950] loss: 0.421
[71, 1980] loss: 0.444
[71, 2010] loss: 0.525
[71, 2040] loss: 0.486
[71, 2070] loss: 0.429
[71, 2100] loss: 0.492
[71, 2130] loss: 0.360
[71, 2160] loss: 0.474
[71, 2190] loss: 0.383
ovft_intermediate_models/epoch70_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.915372
epoch 70 accuracy on train set is: 0.9153716830243548
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	NaN	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.930025

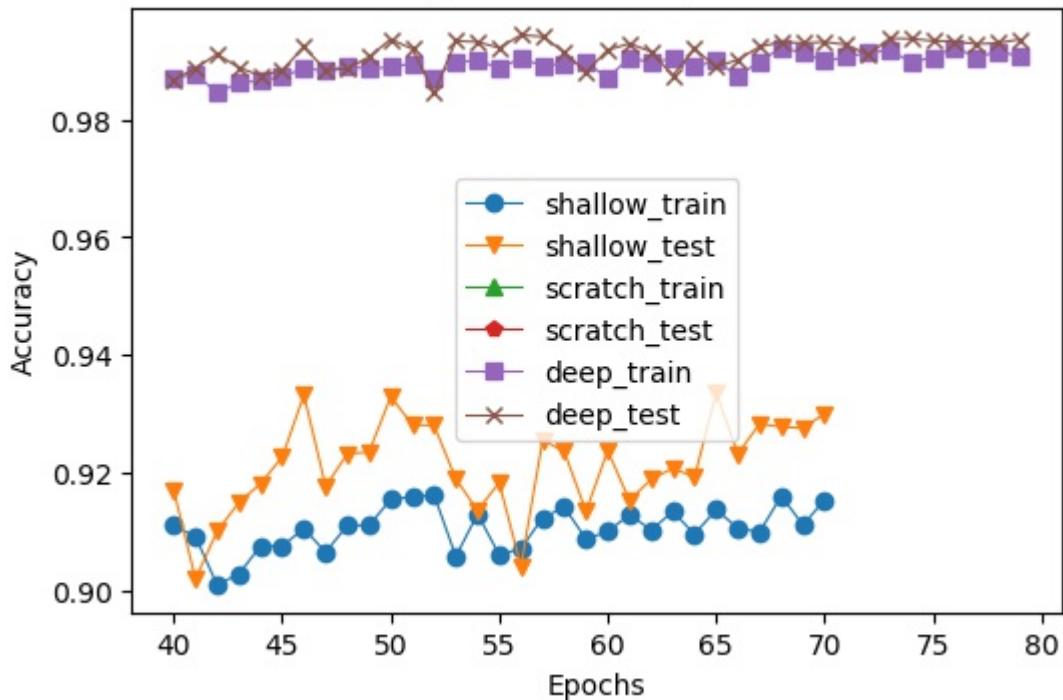
epoch 70 accuracy on test set is: 0.9300254452926209

evaluating on test set takes 16 seconds.

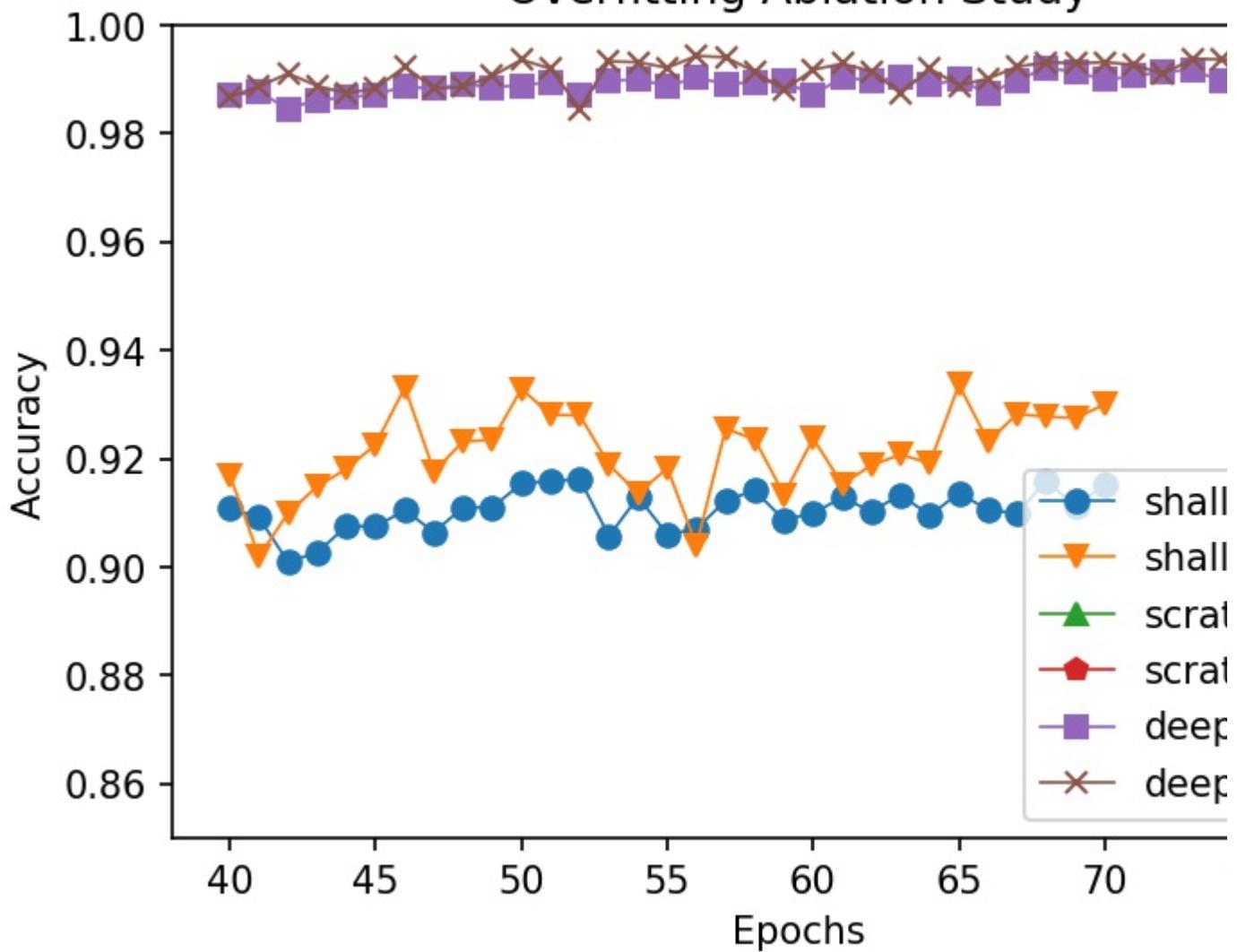
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	NaN	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 70 takes 149 seconds.  
[72,    30] loss: 0.530  
[72,    60] loss: 0.390  
[72,    90] loss: 0.454  
[72,   120] loss: 0.442  
[72,   150] loss: 0.441  
[72,   180] loss: 0.540  
[72,   210] loss: 0.489  
[72,   240] loss: 0.401  
[72,   270] loss: 0.427  
[72,   300] loss: 0.481  
[72,   330] loss: 0.411  
[72,   360] loss: 0.487  
[72,   390] loss: 0.397  
[72,   420] loss: 0.422  
[72,   450] loss: 0.417  
[72,   480] loss: 0.441  
[72,   510] loss: 0.453  
[72,   540] loss: 0.437  
[72,   570] loss: 0.480  
[72,   600] loss: 0.507  
[72,   630] loss: 0.416  
[72,   660] loss: 0.448
```

```

[72, 690] loss: 0.460
[72, 720] loss: 0.463
[72, 750] loss: 0.523
[72, 780] loss: 0.514
[72, 810] loss: 0.497
[72, 840] loss: 0.479
[72, 870] loss: 0.473
[72, 900] loss: 0.449
[72, 930] loss: 0.526
[72, 960] loss: 0.471
[72, 990] loss: 0.447
[72, 1020] loss: 0.406
[72, 1050] loss: 0.419
[72, 1080] loss: 0.451
[72, 1110] loss: 0.557
[72, 1140] loss: 0.523
[72, 1170] loss: 0.511
[72, 1200] loss: 0.462
[72, 1230] loss: 0.479
[72, 1260] loss: 0.514
[72, 1290] loss: 0.531
[72, 1320] loss: 0.481
[72, 1350] loss: 0.426
[72, 1380] loss: 0.536
[72, 1410] loss: 0.586
[72, 1440] loss: 0.519
[72, 1470] loss: 0.394
[72, 1500] loss: 0.462
[72, 1530] loss: 0.567
[72, 1560] loss: 0.436
[72, 1590] loss: 0.404
[72, 1620] loss: 0.517
[72, 1650] loss: 0.411
[72, 1680] loss: 0.549
[72, 1710] loss: 0.471
[72, 1740] loss: 0.434
[72, 1770] loss: 0.457
[72, 1800] loss: 0.503
[72, 1830] loss: 0.386
[72, 1860] loss: 0.413
[72, 1890] loss: 0.480
[72, 1920] loss: 0.514
[72, 1950] loss: 0.453
[72, 1980] loss: 0.464
[72, 2010] loss: 0.526
[72, 2040] loss: 0.439
[72, 2070] loss: 0.417
[72, 2100] loss: 0.513
[72, 2130] loss: 0.447
[72, 2160] loss: 0.516
[72, 2190] loss: 0.464
ovft intermediate models/epoch71_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.916644
epoch 71 accuracy on train set is: 0.916643947655398
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	NaN	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.922755

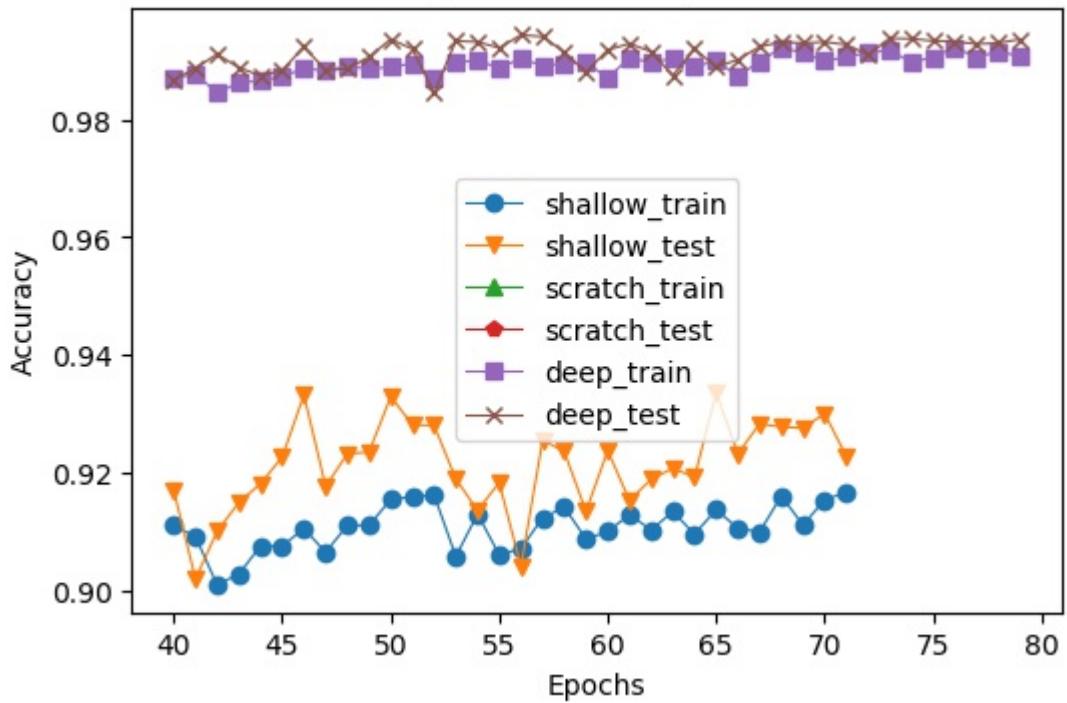
epoch 71 accuracy on test set is: 0.9227553616866594

evaluating on test set takes 16 seconds.

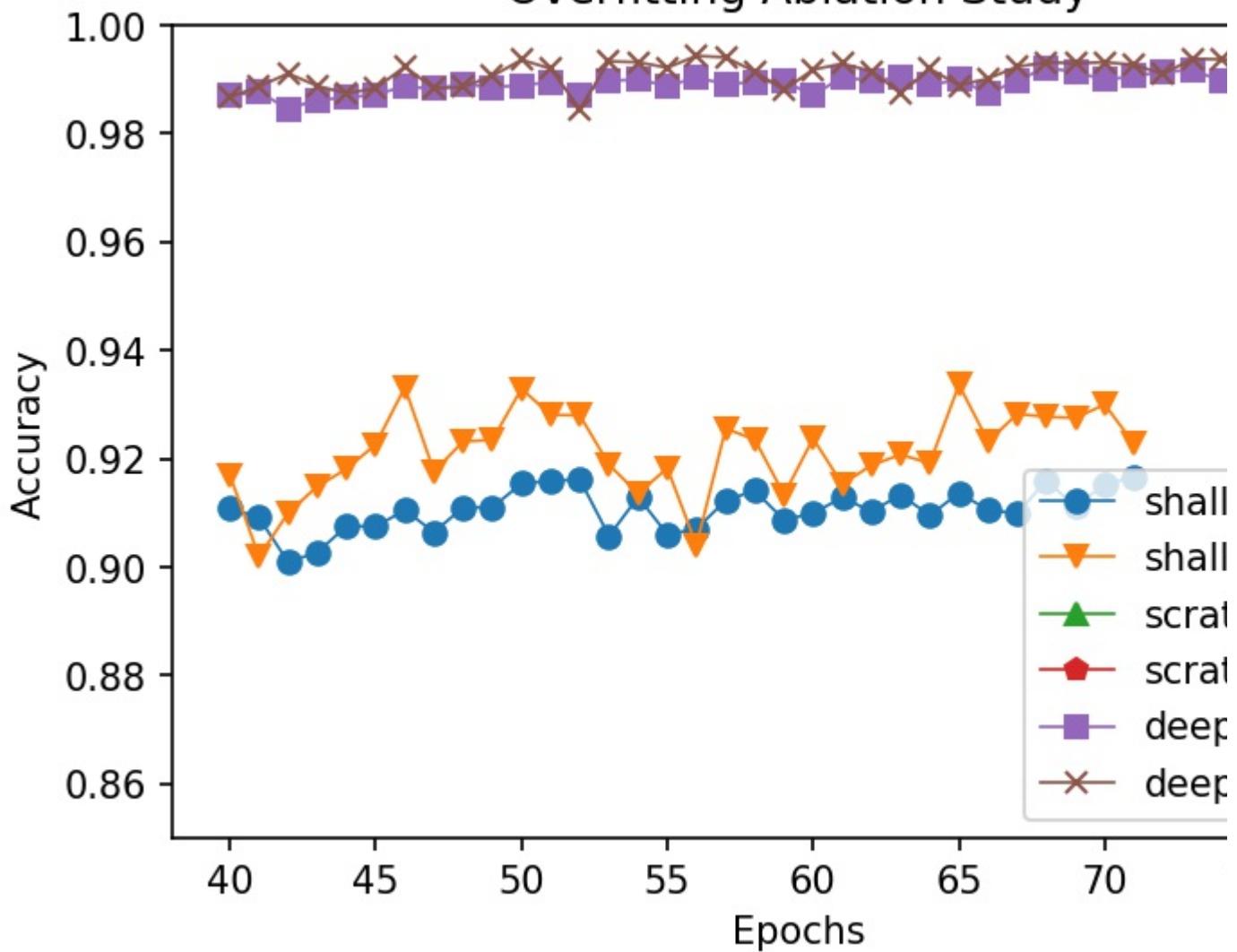
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	NaN	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



epoch 71 takes 150 seconds.

```
[73,    30] loss: 0.447
[73,    60] loss: 0.421
[73,    90] loss: 0.464
[73,   120] loss: 0.501
[73,   150] loss: 0.464
[73,   180] loss: 0.466
[73,   210] loss: 0.421
[73,   240] loss: 0.431
[73,   270] loss: 0.537
[73,   300] loss: 0.428
[73,   330] loss: 0.492
[73,   360] loss: 0.463
[73,   390] loss: 0.418
[73,   420] loss: 0.415
[73,   450] loss: 0.484
[73,   480] loss: 0.539
[73,   510] loss: 0.474
[73,   540] loss: 0.546
[73,   570] loss: 0.477
[73,   600] loss: 0.443
[73,   630] loss: 0.389
[73,   660] loss: 0.348
```

```

[73, 690] loss: 0.480
[73, 720] loss: 0.439
[73, 750] loss: 0.512
[73, 780] loss: 0.432
[73, 810] loss: 0.397
[73, 840] loss: 0.477
[73, 870] loss: 0.442
[73, 900] loss: 0.372
[73, 930] loss: 0.431
[73, 960] loss: 0.461
[73, 990] loss: 0.428
[73, 1020] loss: 0.528
[73, 1050] loss: 0.492
[73, 1080] loss: 0.459
[73, 1110] loss: 0.490
[73, 1140] loss: 0.455
[73, 1170] loss: 0.446
[73, 1200] loss: 0.518
[73, 1230] loss: 0.433
[73, 1260] loss: 0.523
[73, 1290] loss: 0.413
[73, 1320] loss: 0.446
[73, 1350] loss: 0.496
[73, 1380] loss: 0.490
[73, 1410] loss: 0.417
[73, 1440] loss: 0.455
[73, 1470] loss: 0.437
[73, 1500] loss: 0.452
[73, 1530] loss: 0.484
[73, 1560] loss: 0.535
[73, 1590] loss: 0.430
[73, 1620] loss: 0.431
[73, 1650] loss: 0.494
[73, 1680] loss: 0.510
[73, 1710] loss: 0.445
[73, 1740] loss: 0.449
[73, 1770] loss: 0.479
[73, 1800] loss: 0.515
[73, 1830] loss: 0.452
[73, 1860] loss: 0.472
[73, 1890] loss: 0.458
[73, 1920] loss: 0.447
[73, 1950] loss: 0.448
[73, 1980] loss: 0.431
[73, 2010] loss: 0.412
[73, 2040] loss: 0.492
[73, 2070] loss: 0.433
[73, 2100] loss: 0.479
[73, 2130] loss: 0.496
[73, 2160] loss: 0.507
[73, 2190] loss: 0.386
ovft_intermediate_models/epoch72_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.911214
epoch 72 accuracy on train set is: 0.9112141039621956
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	NaN	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.922664

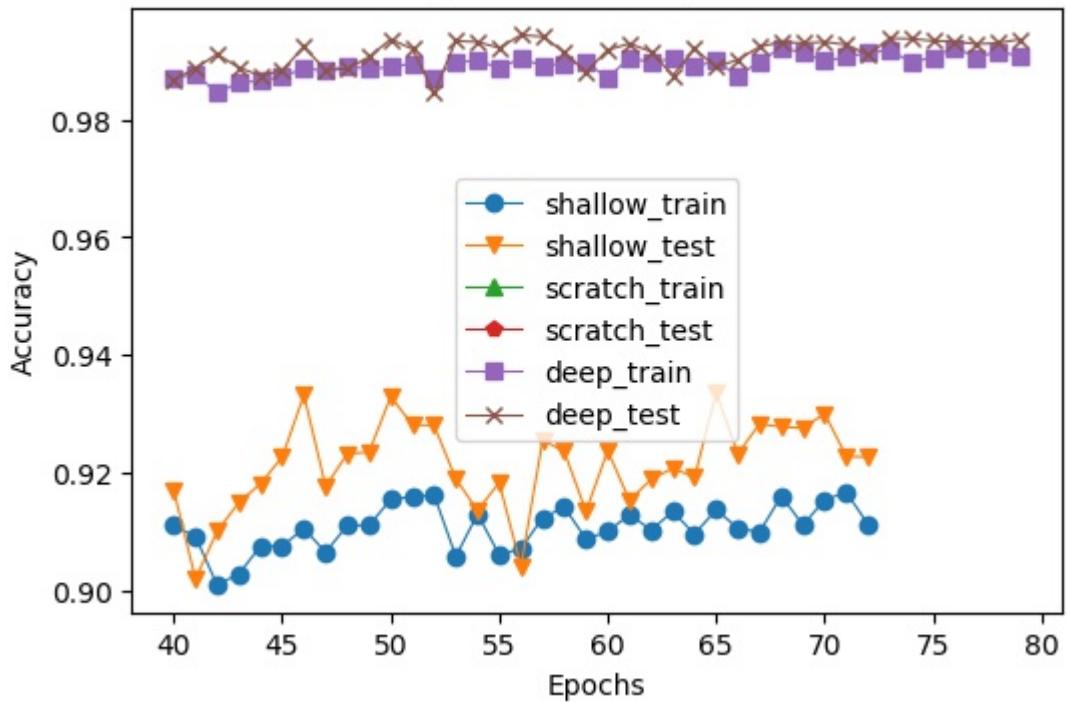
epoch 72 accuracy on test set is: 0.9226644856415849

evaluating on test set takes 16 seconds.

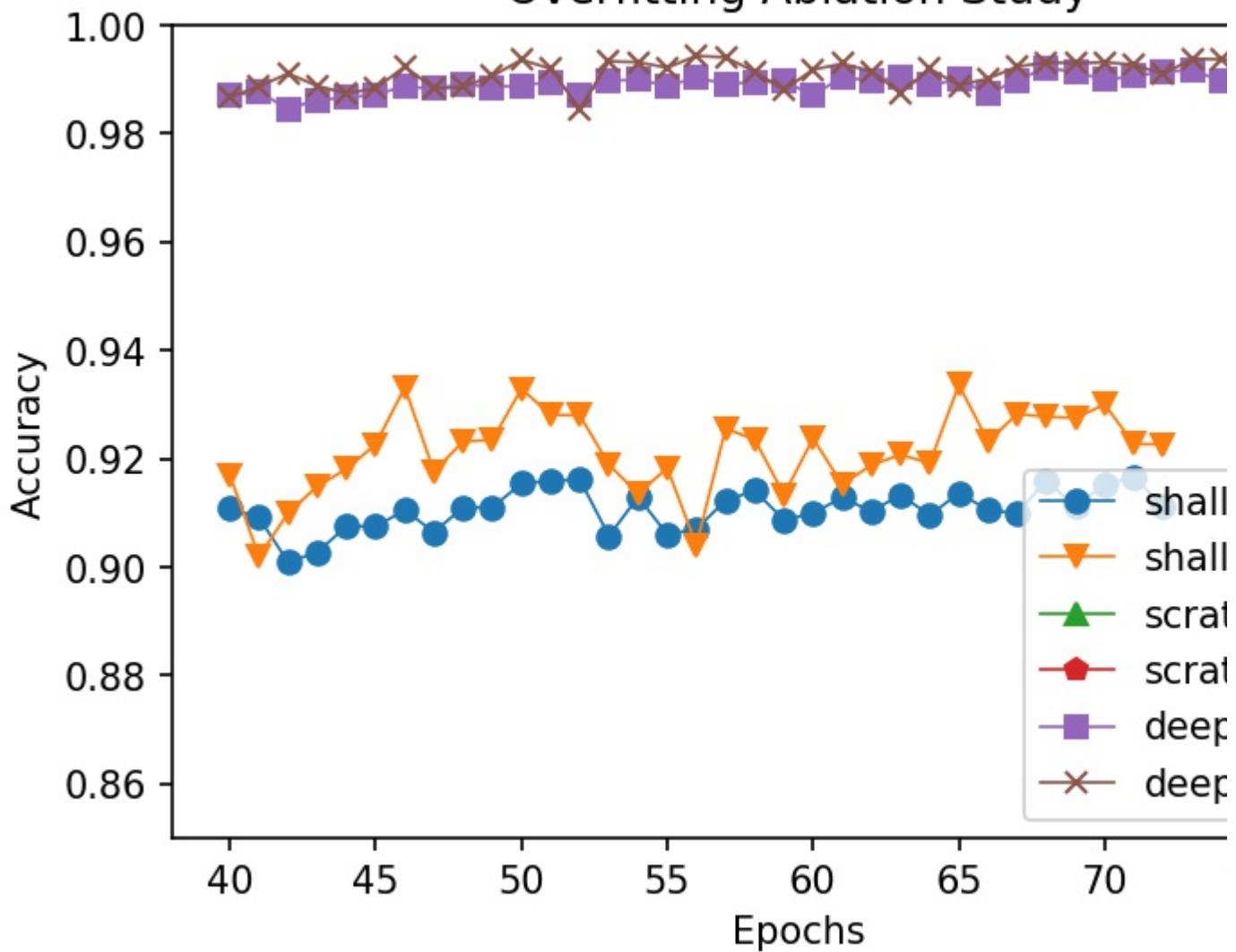
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	NaN	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 72 takes 149 seconds.  
[74,    30] loss: 0.538  
[74,    60] loss: 0.560  
[74,    90] loss: 0.418  
[74,   120] loss: 0.414  
[74,   150] loss: 0.464  
[74,   180] loss: 0.507  
[74,   210] loss: 0.418  
[74,   240] loss: 0.463  
[74,   270] loss: 0.508  
[74,   300] loss: 0.498  
[74,   330] loss: 0.477  
[74,   360] loss: 0.420  
[74,   390] loss: 0.480  
[74,   420] loss: 0.499  
[74,   450] loss: 0.416  
[74,   480] loss: 0.404  
[74,   510] loss: 0.520  
[74,   540] loss: 0.403  
[74,   570] loss: 0.421  
[74,   600] loss: 0.476  
[74,   630] loss: 0.456  
[74,   660] loss: 0.482
```

```

[74, 690] loss: 0.558
[74, 720] loss: 0.391
[74, 750] loss: 0.435
[74, 780] loss: 0.478
[74, 810] loss: 0.460
[74, 840] loss: 0.519
[74, 870] loss: 0.460
[74, 900] loss: 0.488
[74, 930] loss: 0.411
[74, 960] loss: 0.454
[74, 990] loss: 0.399
[74, 1020] loss: 0.455
[74, 1050] loss: 0.462
[74, 1080] loss: 0.454
[74, 1110] loss: 0.393
[74, 1140] loss: 0.481
[74, 1170] loss: 0.392
[74, 1200] loss: 0.456
[74, 1230] loss: 0.443
[74, 1260] loss: 0.447
[74, 1290] loss: 0.607
[74, 1320] loss: 0.511
[74, 1350] loss: 0.490
[74, 1380] loss: 0.403
[74, 1410] loss: 0.509
[74, 1440] loss: 0.423
[74, 1470] loss: 0.412
[74, 1500] loss: 0.468
[74, 1530] loss: 0.413
[74, 1560] loss: 0.406
[74, 1590] loss: 0.420
[74, 1620] loss: 0.451
[74, 1650] loss: 0.470
[74, 1680] loss: 0.519
[74, 1710] loss: 0.499
[74, 1740] loss: 0.478
[74, 1770] loss: 0.470
[74, 1800] loss: 0.438
[74, 1830] loss: 0.402
[74, 1860] loss: 0.442
[74, 1890] loss: 0.452
[74, 1920] loss: 0.383
[74, 1950] loss: 0.467
[74, 1980] loss: 0.394
[74, 2010] loss: 0.500
[74, 2040] loss: 0.472
[74, 2070] loss: 0.411
[74, 2100] loss: 0.463
[74, 2130] loss: 0.400
[74, 2160] loss: 0.396
[74, 2190] loss: 0.477
ovft_intermediate_models/epoch73_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.907102
epoch 73 accuracy on train set is: 0.9071019629225736
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	NaN	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.920302

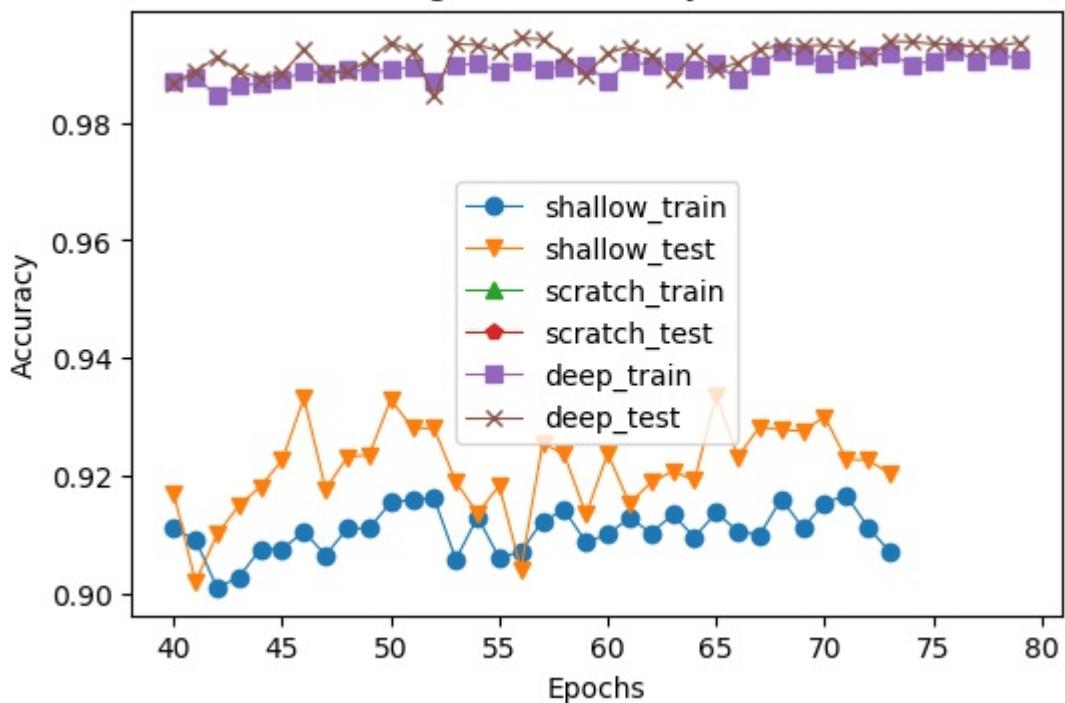
epoch 73 accuracy on test set is: 0.9203017084696474

evaluating on test set takes 16 seconds.

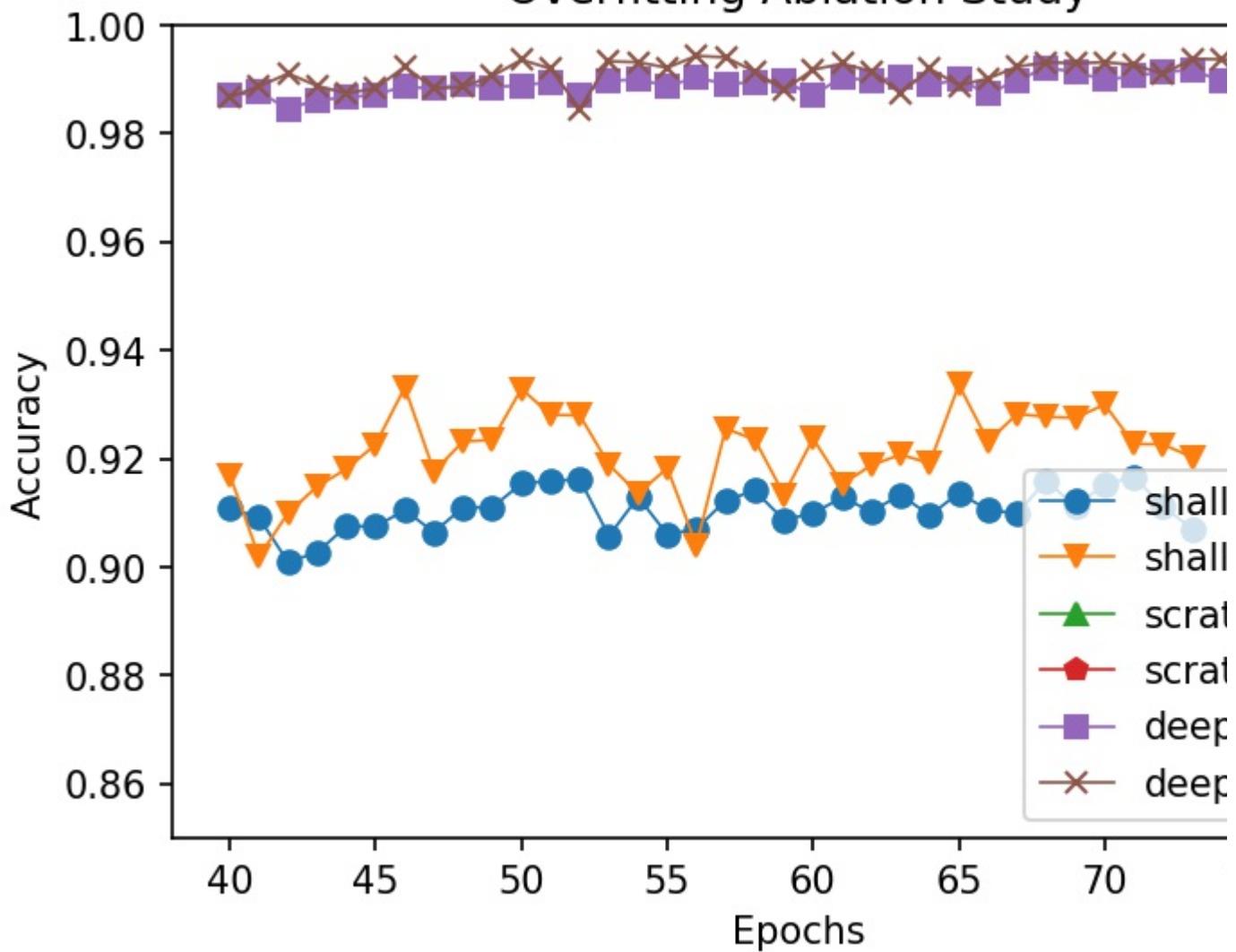
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	NaN	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 73 takes 149 seconds.  
[75,    30] loss: 0.399  
[75,    60] loss: 0.510  
[75,   90] loss: 0.510  
[75,  120] loss: 0.475  
[75,  150] loss: 0.383  
[75,  180] loss: 0.466  
[75,  210] loss: 0.493  
[75,  240] loss: 0.438  
[75,  270] loss: 0.408  
[75,  300] loss: 0.434  
[75,  330] loss: 0.529  
[75,  360] loss: 0.380  
[75,  390] loss: 0.509  
[75,  420] loss: 0.450  
[75,  450] loss: 0.448  
[75,  480] loss: 0.378  
[75,  510] loss: 0.502  
[75,  540] loss: 0.471  
[75,  570] loss: 0.457  
[75,  600] loss: 0.402  
[75,  630] loss: 0.479  
[75,  660] loss: 0.498
```

```

[75,   690] loss: 0.472
[75,   720] loss: 0.516
[75,   750] loss: 0.380
[75,   780] loss: 0.443
[75,   810] loss: 0.467
[75,   840] loss: 0.530
[75,   870] loss: 0.440
[75,   900] loss: 0.478
[75,   930] loss: 0.328
[75,   960] loss: 0.423
[75,   990] loss: 0.532
[75,  1020] loss: 0.486
[75,  1050] loss: 0.430
[75,  1080] loss: 0.503
[75,  1110] loss: 0.374
[75,  1140] loss: 0.549
[75,  1170] loss: 0.459
[75,  1200] loss: 0.451
[75,  1230] loss: 0.470
[75,  1260] loss: 0.425
[75,  1290] loss: 0.399
[75,  1320] loss: 0.491
[75,  1350] loss: 0.497
[75,  1380] loss: 0.494
[75,  1410] loss: 0.446
[75,  1440] loss: 0.489
[75,  1470] loss: 0.481
[75,  1500] loss: 0.441
[75,  1530] loss: 0.489
[75,  1560] loss: 0.517
[75,  1590] loss: 0.519
[75,  1620] loss: 0.538
[75,  1650] loss: 0.389
[75,  1680] loss: 0.512
[75,  1710] loss: 0.409
[75,  1740] loss: 0.453
[75,  1770] loss: 0.372
[75,  1800] loss: 0.390
[75,  1830] loss: 0.436
[75,  1860] loss: 0.458
[75,  1890] loss: 0.426
[75,  1920] loss: 0.439
[75,  1950] loss: 0.561
[75,  1980] loss: 0.449
[75,  2010] loss: 0.429
[75,  2040] loss: 0.462
[75,  2070] loss: 0.517
[75,  2100] loss: 0.473
[75,  2130] loss: 0.468
[75,  2160] loss: 0.469
[75,  2190] loss: 0.484
ovft intermediate models/epoch74_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.909578
epoch 74 accuracy on train set is: 0.9095783351508543
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN    0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	NaN	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.912123

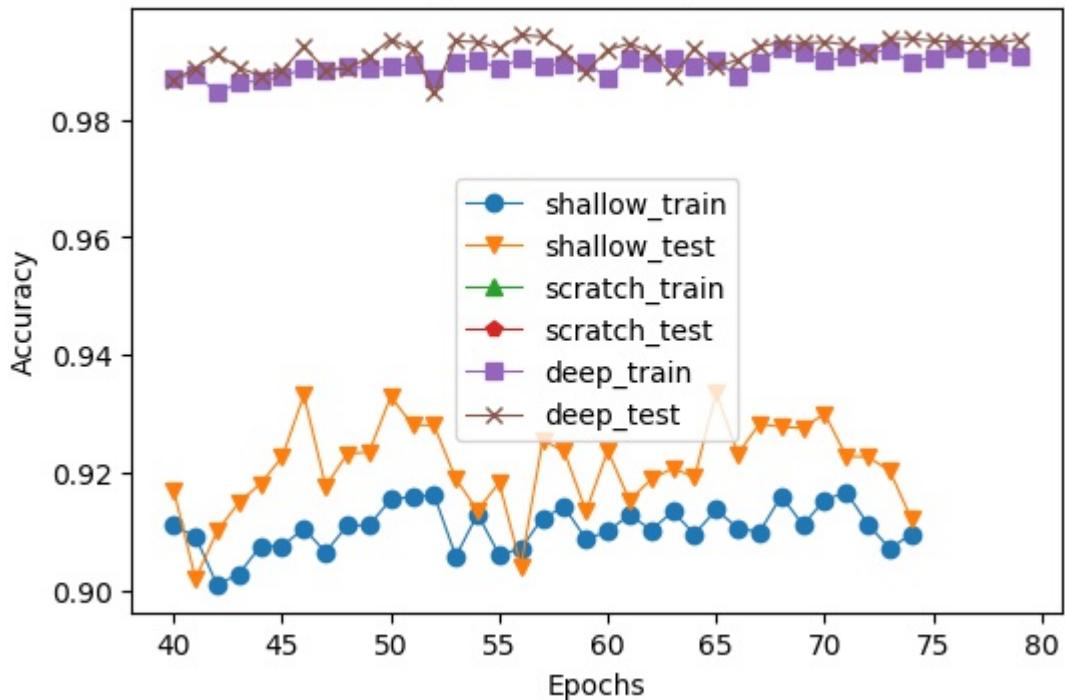
epoch 74 accuracy on test set is: 0.9121228644129408

evaluating on test set takes 16 seconds.

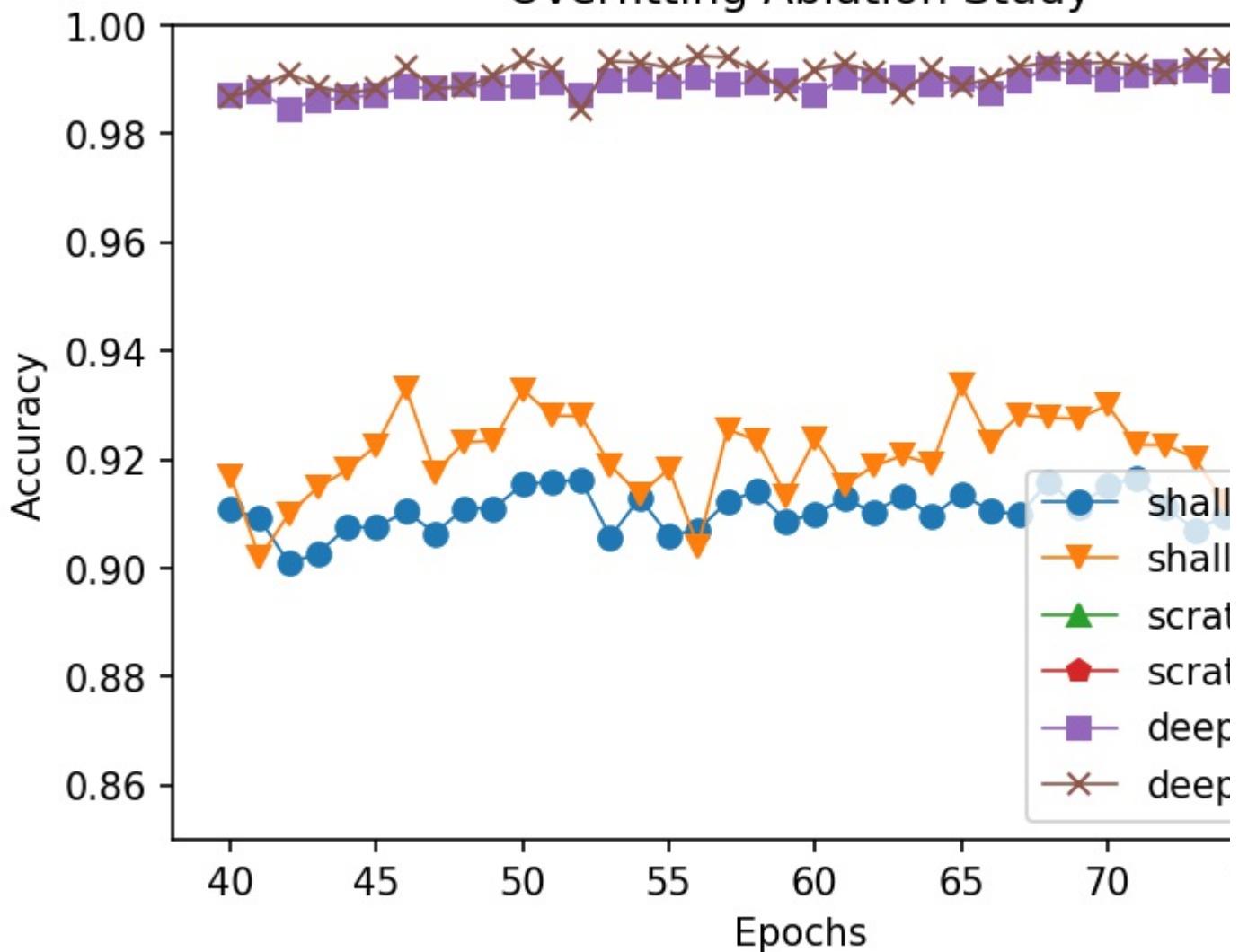
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	NaN	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 74 takes 149 seconds.  
[76,    30] loss: 0.500  
[76,    60] loss: 0.538  
[76,    90] loss: 0.541  
[76,   120] loss: 0.458  
[76,   150] loss: 0.433  
[76,   180] loss: 0.409  
[76,   210] loss: 0.425  
[76,   240] loss: 0.422  
[76,   270] loss: 0.459  
[76,   300] loss: 0.425  
[76,   330] loss: 0.515  
[76,   360] loss: 0.505  
[76,   390] loss: 0.467  
[76,   420] loss: 0.462  
[76,   450] loss: 0.517  
[76,   480] loss: 0.533  
[76,   510] loss: 0.408  
[76,   540] loss: 0.514  
[76,   570] loss: 0.487  
[76,   600] loss: 0.410  
[76,   630] loss: 0.481  
[76,   660] loss: 0.419
```

```

[76, 690] loss: 0.515
[76, 720] loss: 0.434
[76, 750] loss: 0.453
[76, 780] loss: 0.469
[76, 810] loss: 0.495
[76, 840] loss: 0.528
[76, 870] loss: 0.529
[76, 900] loss: 0.537
[76, 930] loss: 0.530
[76, 960] loss: 0.434
[76, 990] loss: 0.487
[76, 1020] loss: 0.427
[76, 1050] loss: 0.508
[76, 1080] loss: 0.491
[76, 1110] loss: 0.492
[76, 1140] loss: 0.438
[76, 1170] loss: 0.483
[76, 1200] loss: 0.490
[76, 1230] loss: 0.501
[76, 1260] loss: 0.479
[76, 1290] loss: 0.510
[76, 1320] loss: 0.451
[76, 1350] loss: 0.465
[76, 1380] loss: 0.463
[76, 1410] loss: 0.432
[76, 1440] loss: 0.465
[76, 1470] loss: 0.469
[76, 1500] loss: 0.416
[76, 1530] loss: 0.474
[76, 1560] loss: 0.455
[76, 1590] loss: 0.440
[76, 1620] loss: 0.409
[76, 1650] loss: 0.461
[76, 1680] loss: 0.449
[76, 1710] loss: 0.515
[76, 1740] loss: 0.491
[76, 1770] loss: 0.458
[76, 1800] loss: 0.454
[76, 1830] loss: 0.438
[76, 1860] loss: 0.415
[76, 1890] loss: 0.500
[76, 1920] loss: 0.414
[76, 1950] loss: 0.499
[76, 1980] loss: 0.447
[76, 2010] loss: 0.439
[76, 2040] loss: 0.447
[76, 2070] loss: 0.496
[76, 2100] loss: 0.425
[76, 2130] loss: 0.419
[76, 2160] loss: 0.429
[76, 2190] loss: 0.458
ovft_intermediate_models/epoch75_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.909578
epoch 75 accuracy on train set is: 0.9095783351508543
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	NaN	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.919575

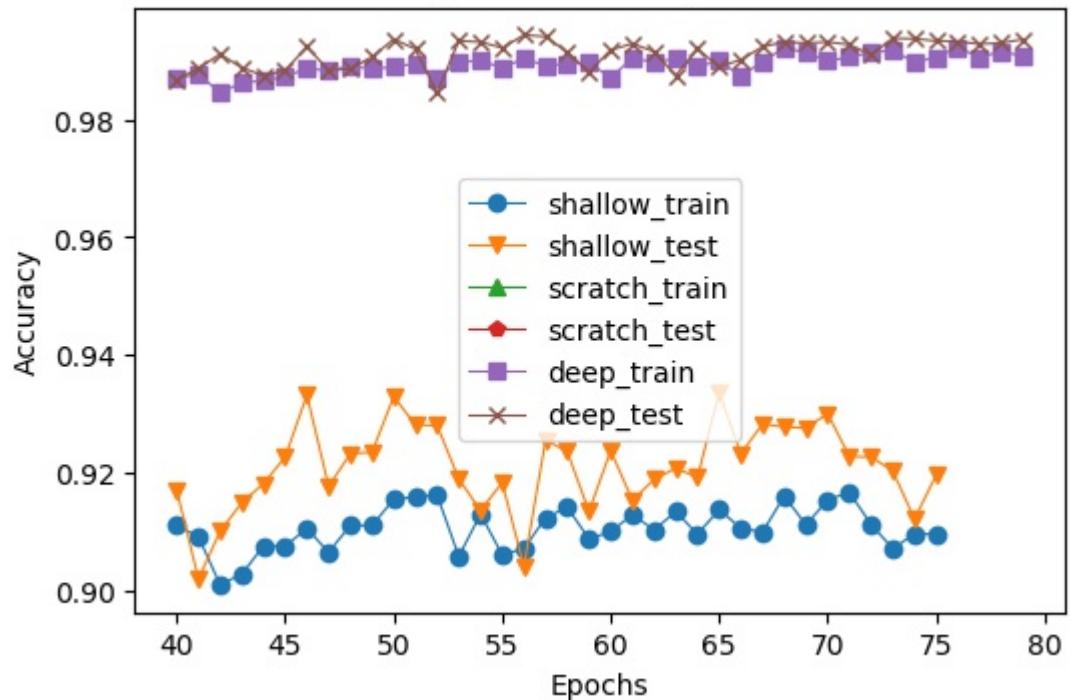
epoch 75 accuracy on test set is: 0.9195747001090513

evaluating on test set takes 16 seconds.

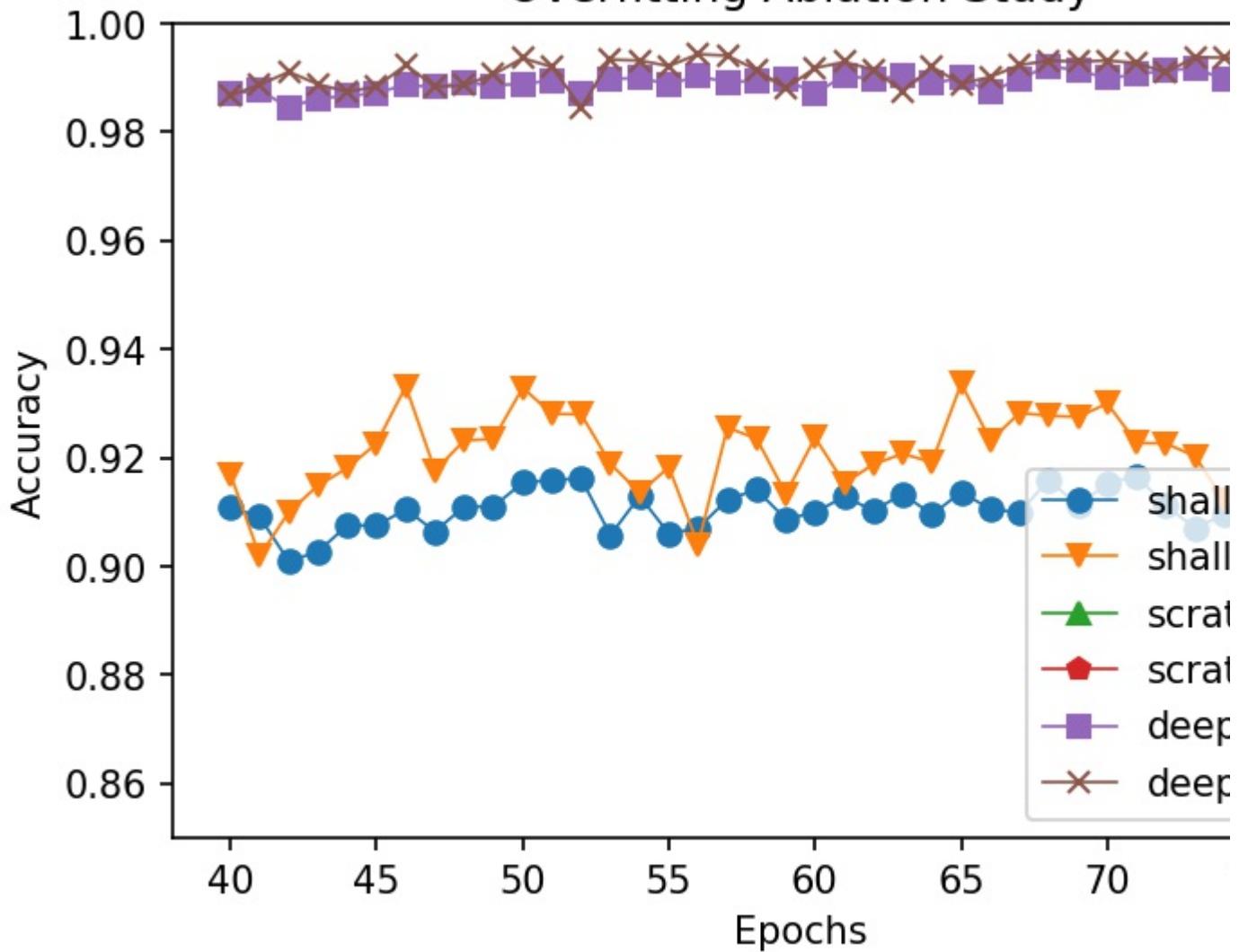
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	NaN	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



epoch 75 takes 149 seconds.

```
[77,    30] loss: 0.401
[77,    60] loss: 0.511
[77,    90] loss: 0.401
[77,   120] loss: 0.379
[77,   150] loss: 0.511
[77,   180] loss: 0.438
[77,   210] loss: 0.425
[77,   240] loss: 0.498
[77,   270] loss: 0.466
[77,   300] loss: 0.441
[77,   330] loss: 0.464
[77,   360] loss: 0.424
[77,   390] loss: 0.437
[77,   420] loss: 0.434
[77,   450] loss: 0.462
[77,   480] loss: 0.426
[77,   510] loss: 0.466
[77,   540] loss: 0.493
[77,   570] loss: 0.432
[77,   600] loss: 0.439
[77,   630] loss: 0.488
[77,   660] loss: 0.440
```

```

[77, 690] loss: 0.363
[77, 720] loss: 0.329
[77, 750] loss: 0.554
[77, 780] loss: 0.526
[77, 810] loss: 0.412
[77, 840] loss: 0.417
[77, 870] loss: 0.542
[77, 900] loss: 0.366
[77, 930] loss: 0.394
[77, 960] loss: 0.427
[77, 990] loss: 0.409
[77, 1020] loss: 0.451
[77, 1050] loss: 0.442
[77, 1080] loss: 0.402
[77, 1110] loss: 0.376
[77, 1140] loss: 0.505
[77, 1170] loss: 0.489
[77, 1200] loss: 0.515
[77, 1230] loss: 0.602
[77, 1260] loss: 0.490
[77, 1290] loss: 0.445
[77, 1320] loss: 0.548
[77, 1350] loss: 0.505
[77, 1380] loss: 0.462
[77, 1410] loss: 0.439
[77, 1440] loss: 0.524
[77, 1470] loss: 0.483
[77, 1500] loss: 0.481
[77, 1530] loss: 0.376
[77, 1560] loss: 0.486
[77, 1590] loss: 0.438
[77, 1620] loss: 0.485
[77, 1650] loss: 0.524
[77, 1680] loss: 0.534
[77, 1710] loss: 0.487
[77, 1740] loss: 0.479
[77, 1770] loss: 0.397
[77, 1800] loss: 0.528
[77, 1830] loss: 0.471
[77, 1860] loss: 0.551
[77, 1890] loss: 0.520
[77, 1920] loss: 0.379
[77, 1950] loss: 0.464
[77, 1980] loss: 0.453
[77, 2010] loss: 0.467
[77, 2040] loss: 0.419
[77, 2070] loss: 0.396
[77, 2100] loss: 0.458
[77, 2130] loss: 0.479
[77, 2160] loss: 0.486
[77, 2190] loss: 0.453
ovft_intermediate_models/epoch76_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.9117366412213741
epoch 76 accuracy on train set is: 0.9117366412213741
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	NaN	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.917757

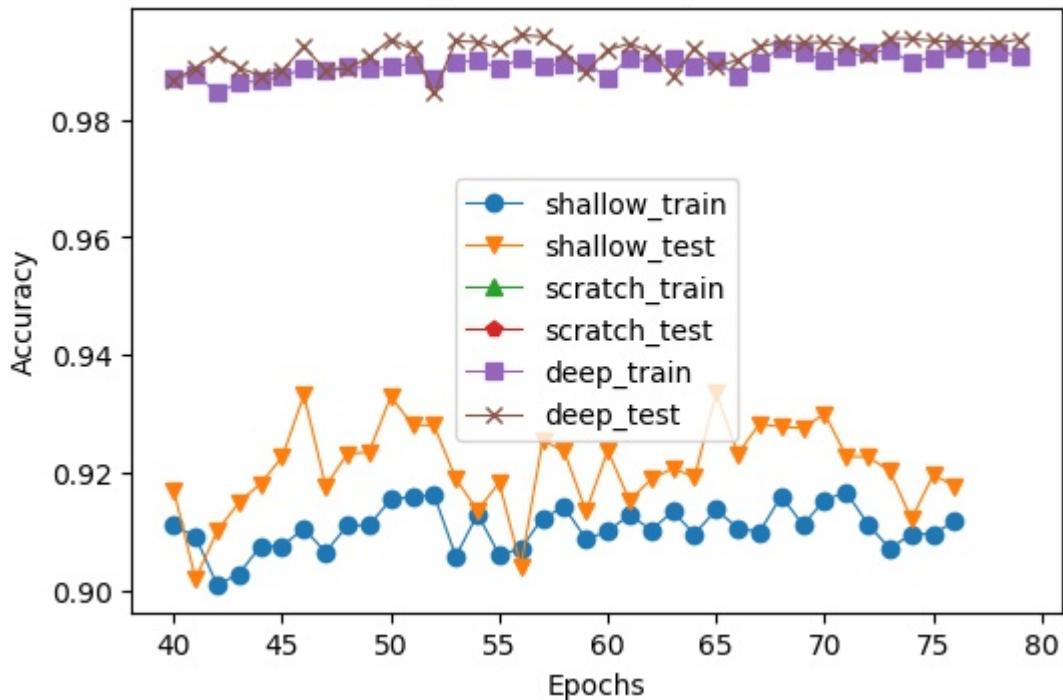
epoch 76 accuracy on test set is: 0.9177571792075608

evaluating on test set takes 16 seconds.

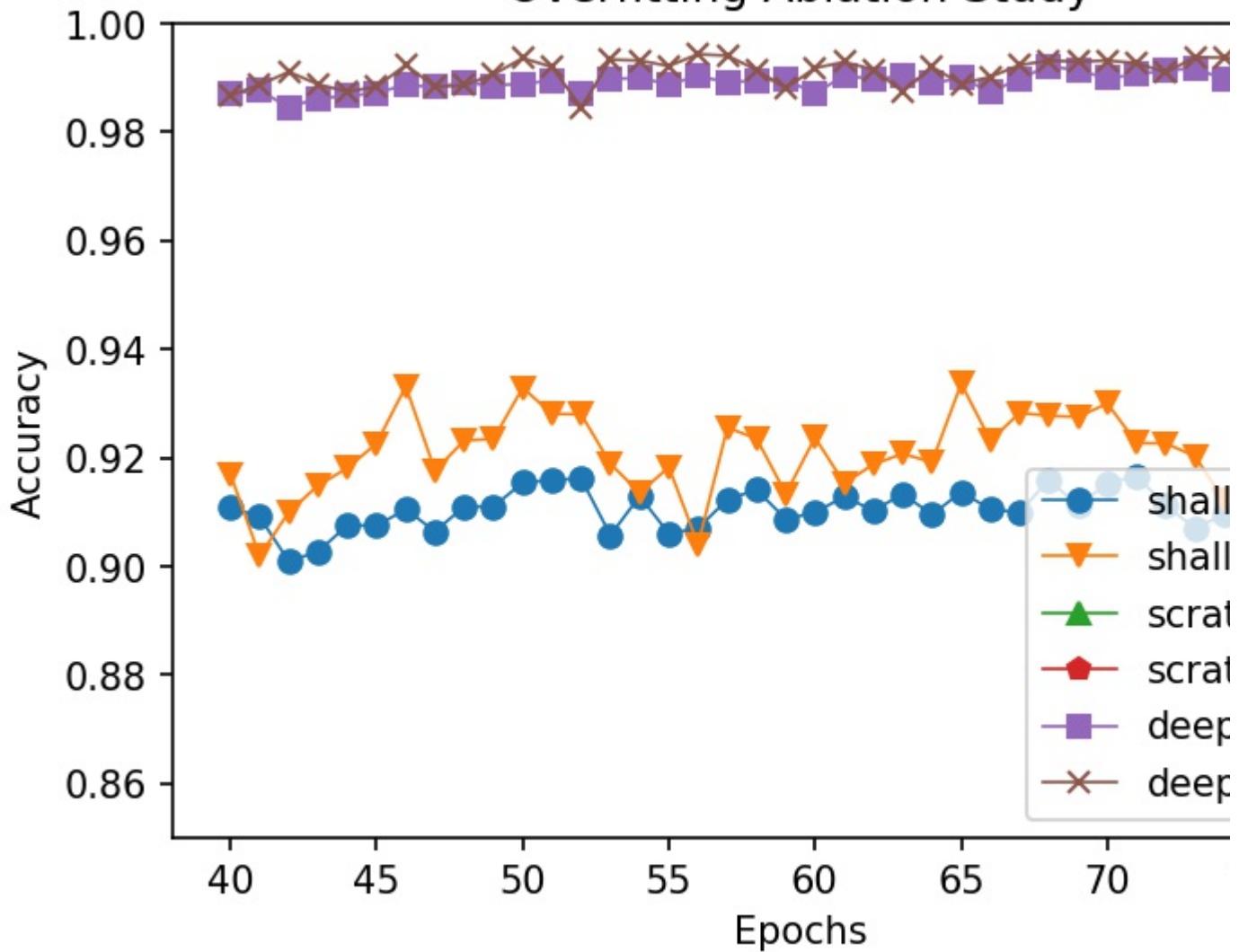
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	0.917757	NaN	NaN	0.992253	0.993275
77	NaN	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 76 takes 149 seconds.  
[78,    30] loss: 0.563  
[78,    60] loss: 0.422  
[78,    90] loss: 0.469  
[78,   120] loss: 0.370  
[78,   150] loss: 0.374  
[78,   180] loss: 0.470  
[78,   210] loss: 0.388  
[78,   240] loss: 0.490  
[78,   270] loss: 0.461  
[78,   300] loss: 0.367  
[78,   330] loss: 0.438  
[78,   360] loss: 0.380  
[78,   390] loss: 0.493  
[78,   420] loss: 0.396  
[78,   450] loss: 0.429  
[78,   480] loss: 0.430  
[78,   510] loss: 0.377  
[78,   540] loss: 0.530  
[78,   570] loss: 0.393  
[78,   600] loss: 0.531  
[78,   630] loss: 0.424  
[78,   660] loss: 0.408
```

```

[78, 690] loss: 0.477
[78, 720] loss: 0.507
[78, 750] loss: 0.491
[78, 780] loss: 0.506
[78, 810] loss: 0.467
[78, 840] loss: 0.508
[78, 870] loss: 0.449
[78, 900] loss: 0.476
[78, 930] loss: 0.376
[78, 960] loss: 0.508
[78, 990] loss: 0.473
[78, 1020] loss: 0.501
[78, 1050] loss: 0.442
[78, 1080] loss: 0.499
[78, 1110] loss: 0.481
[78, 1140] loss: 0.513
[78, 1170] loss: 0.449
[78, 1200] loss: 0.501
[78, 1230] loss: 0.548
[78, 1260] loss: 0.418
[78, 1290] loss: 0.500
[78, 1320] loss: 0.470
[78, 1350] loss: 0.485
[78, 1380] loss: 0.503
[78, 1410] loss: 0.508
[78, 1440] loss: 0.468
[78, 1470] loss: 0.464
[78, 1500] loss: 0.460
[78, 1530] loss: 0.366
[78, 1560] loss: 0.556
[78, 1590] loss: 0.492
[78, 1620] loss: 0.475
[78, 1650] loss: 0.495
[78, 1680] loss: 0.494
[78, 1710] loss: 0.437
[78, 1740] loss: 0.441
[78, 1770] loss: 0.371
[78, 1800] loss: 0.447
[78, 1830] loss: 0.553
[78, 1860] loss: 0.498
[78, 1890] loss: 0.451
[78, 1920] loss: 0.427
[78, 1950] loss: 0.422
[78, 1980] loss: 0.460
[78, 2010] loss: 0.464
[78, 2040] loss: 0.382
[78, 2070] loss: 0.435
[78, 2100] loss: 0.472
[78, 2130] loss: 0.494
[78, 2160] loss: 0.437
[78, 2190] loss: 0.462
ovft_intermediate_models/epoch77_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.916099
epoch 77 accuracy on train set is: 0.9160986913849509
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	0.917757	NaN	NaN	0.992253	0.993275
77	0.916099	NaN	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.925027

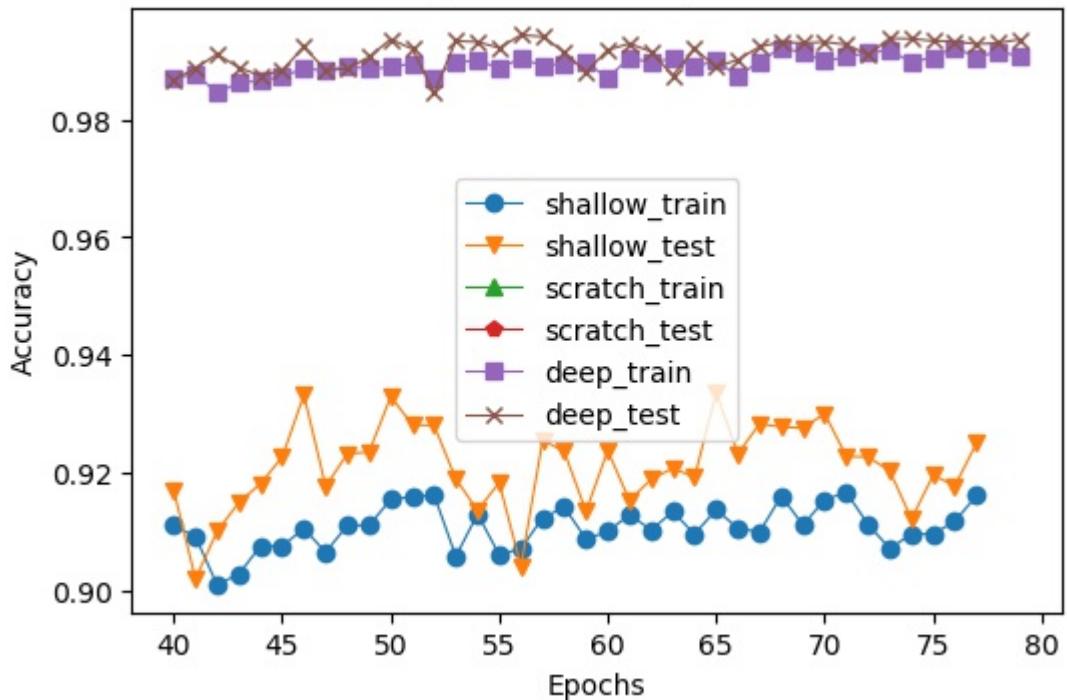
epoch 77 accuracy on test set is: 0.9250272628135223

evaluating on test set takes 16 seconds.

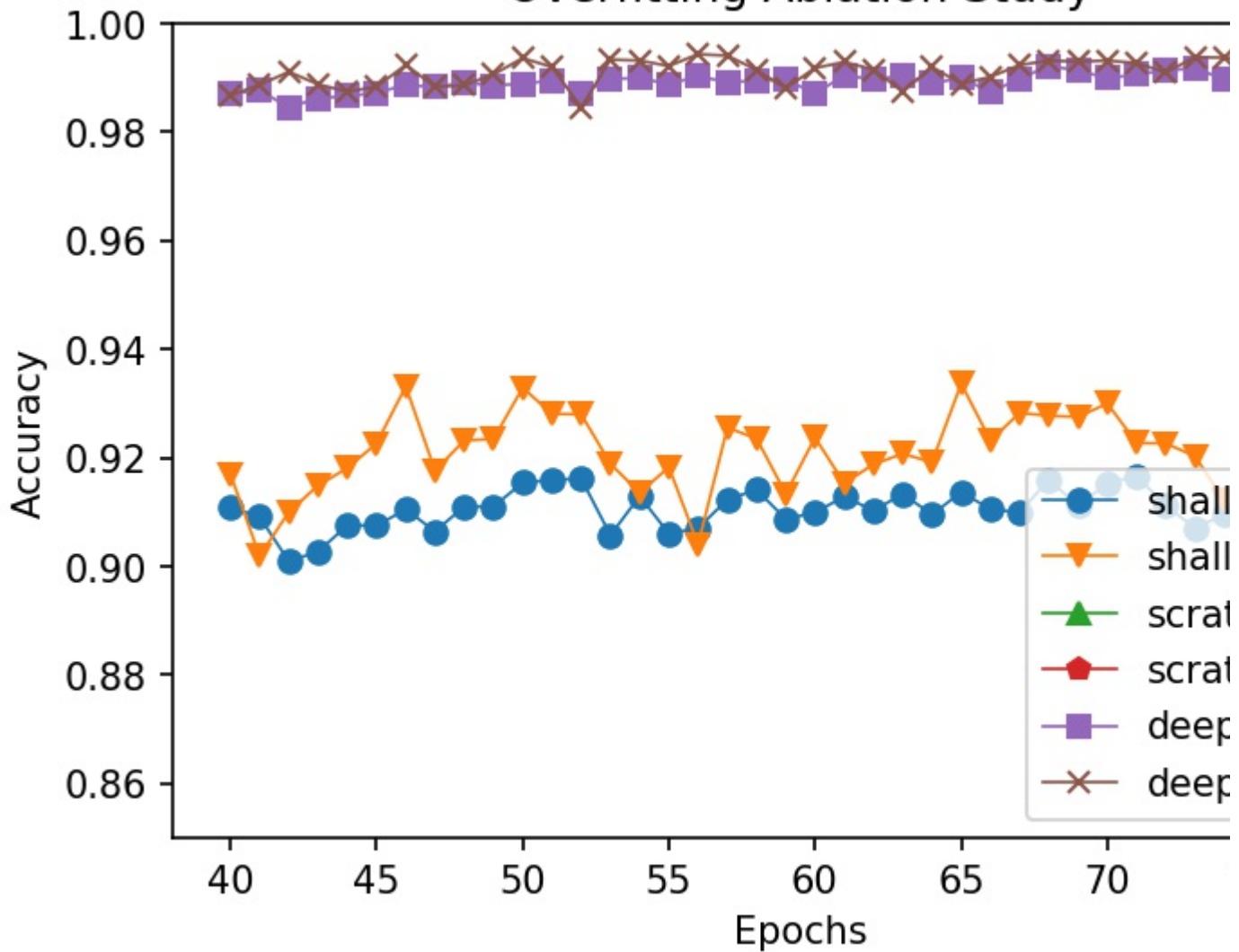
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	0.917757	NaN	NaN	0.992253	0.993275
77	0.916099	0.925027	NaN	NaN	0.990549	0.992821
78	NaN	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 77 takes 149 seconds.  
[79,    30] loss: 0.392  
[79,    60] loss: 0.477  
[79,   90] loss: 0.479  
[79,  120] loss: 0.392  
[79,  150] loss: 0.410  
[79,  180] loss: 0.413  
[79,  210] loss: 0.431  
[79,  240] loss: 0.534  
[79,  270] loss: 0.471  
[79,  300] loss: 0.477  
[79,  330] loss: 0.455  
[79,  360] loss: 0.454  
[79,  390] loss: 0.415  
[79,  420] loss: 0.471  
[79,  450] loss: 0.445  
[79,  480] loss: 0.475  
[79,  510] loss: 0.454  
[79,  540] loss: 0.475  
[79,  570] loss: 0.385  
[79,  600] loss: 0.472  
[79,  630] loss: 0.444  
[79,  660] loss: 0.469
```

```

[79, 690] loss: 0.504
[79, 720] loss: 0.448
[79, 750] loss: 0.425
[79, 780] loss: 0.447
[79, 810] loss: 0.420
[79, 840] loss: 0.392
[79, 870] loss: 0.492
[79, 900] loss: 0.425
[79, 930] loss: 0.444
[79, 960] loss: 0.509
[79, 990] loss: 0.467
[79, 1020] loss: 0.444
[79, 1050] loss: 0.454
[79, 1080] loss: 0.415
[79, 1110] loss: 0.475
[79, 1140] loss: 0.507
[79, 1170] loss: 0.473
[79, 1200] loss: 0.527
[79, 1230] loss: 0.482
[79, 1260] loss: 0.500
[79, 1290] loss: 0.447
[79, 1320] loss: 0.401
[79, 1350] loss: 0.451
[79, 1380] loss: 0.432
[79, 1410] loss: 0.425
[79, 1440] loss: 0.340
[79, 1470] loss: 0.489
[79, 1500] loss: 0.394
[79, 1530] loss: 0.499
[79, 1560] loss: 0.442
[79, 1590] loss: 0.452
[79, 1620] loss: 0.508
[79, 1650] loss: 0.510
[79, 1680] loss: 0.458
[79, 1710] loss: 0.429
[79, 1740] loss: 0.434
[79, 1770] loss: 0.484
[79, 1800] loss: 0.485
[79, 1830] loss: 0.482
[79, 1860] loss: 0.440
[79, 1890] loss: 0.389
[79, 1920] loss: 0.505
[79, 1950] loss: 0.345
[79, 1980] loss: 0.437
[79, 2010] loss: 0.432
[79, 2040] loss: 0.554
[79, 2070] loss: 0.458
[79, 2100] loss: 0.458
[79, 2130] loss: 0.486
[79, 2160] loss: 0.447
[79, 2190] loss: 0.562
ovft_intermediate_models/epoch78_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.915553
epoch 78 accuracy on train set is: 0.9155534351145038
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	0.917757	NaN	NaN	0.992253	0.993275
77	0.916099	0.925027	NaN	NaN	0.990549	0.992821
78	0.915553	NaN	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.917212

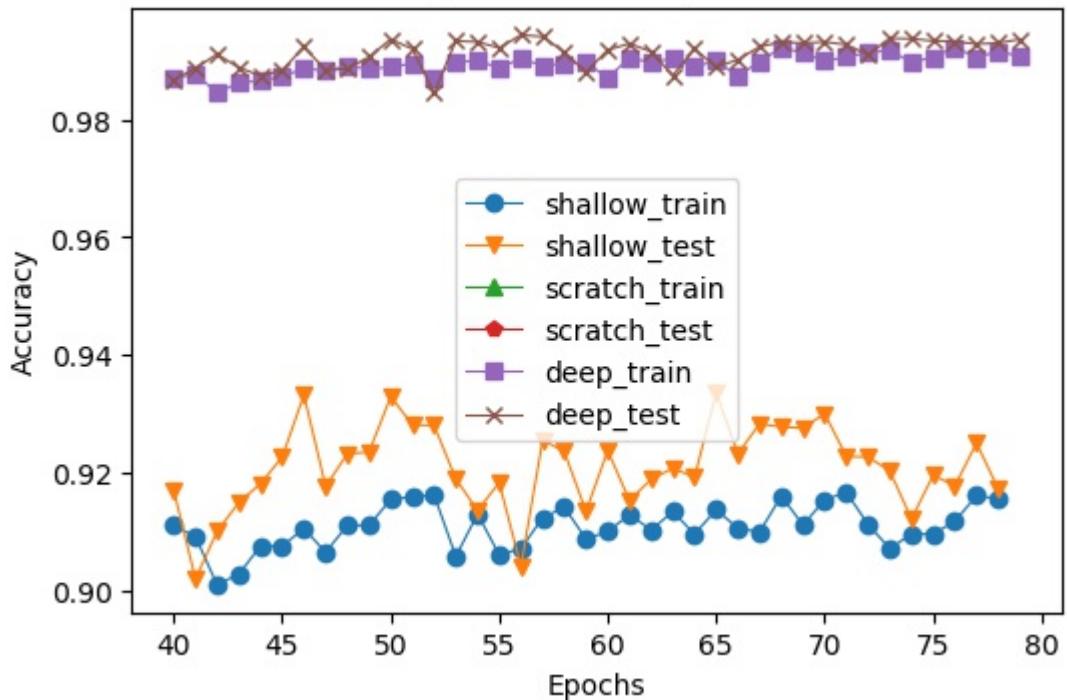
epoch 78 accuracy on test set is: 0.9172119229371137

evaluating on test set takes 16 seconds.

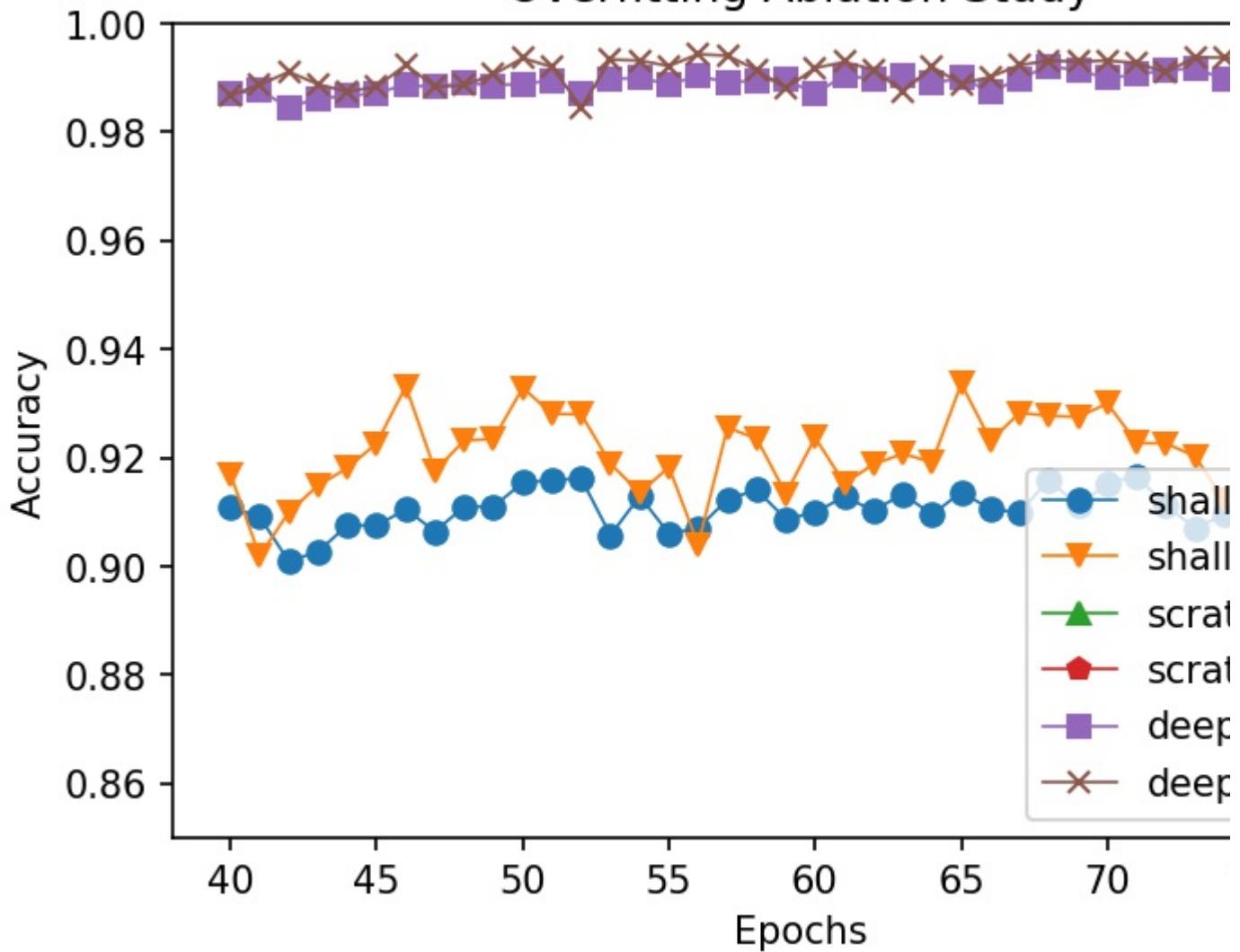
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	0.917757	NaN	NaN	0.992253	0.993275
77	0.916099	0.925027	NaN	NaN	0.990549	0.992821
78	0.915553	0.917212	NaN	NaN	0.991594	0.993003
79	NaN	NaN	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



```
epoch 78 takes 149 seconds.  
[80,    30] loss: 0.485  
[80,    60] loss: 0.530  
[80,   90] loss: 0.460  
[80,  120] loss: 0.424  
[80,  150] loss: 0.394  
[80,  180] loss: 0.483  
[80,  210] loss: 0.442  
[80,  240] loss: 0.502  
[80,  270] loss: 0.446  
[80,  300] loss: 0.374  
[80,  330] loss: 0.505  
[80,  360] loss: 0.481  
[80,  390] loss: 0.438  
[80,  420] loss: 0.520  
[80,  450] loss: 0.372  
[80,  480] loss: 0.440  
[80,  510] loss: 0.442  
[80,  540] loss: 0.430  
[80,  570] loss: 0.475  
[80,  600] loss: 0.462  
[80,  630] loss: 0.485  
[80,  660] loss: 0.470
```

```

[80, 690] loss: 0.446
[80, 720] loss: 0.529
[80, 750] loss: 0.419
[80, 780] loss: 0.431
[80, 810] loss: 0.463
[80, 840] loss: 0.455
[80, 870] loss: 0.486
[80, 900] loss: 0.467
[80, 930] loss: 0.474
[80, 960] loss: 0.418
[80, 990] loss: 0.381
[80, 1020] loss: 0.461
[80, 1050] loss: 0.608
[80, 1080] loss: 0.386
[80, 1110] loss: 0.516
[80, 1140] loss: 0.445
[80, 1170] loss: 0.548
[80, 1200] loss: 0.464
[80, 1230] loss: 0.348
[80, 1260] loss: 0.473
[80, 1290] loss: 0.433
[80, 1320] loss: 0.506
[80, 1350] loss: 0.481
[80, 1380] loss: 0.502
[80, 1410] loss: 0.489
[80, 1440] loss: 0.509
[80, 1470] loss: 0.434
[80, 1500] loss: 0.472
[80, 1530] loss: 0.477
[80, 1560] loss: 0.479
[80, 1590] loss: 0.402
[80, 1620] loss: 0.464
[80, 1650] loss: 0.480
[80, 1680] loss: 0.444
[80, 1710] loss: 0.423
[80, 1740] loss: 0.525
[80, 1770] loss: 0.540
[80, 1800] loss: 0.378
[80, 1830] loss: 0.448
[80, 1860] loss: 0.546
[80, 1890] loss: 0.409
[80, 1920] loss: 0.445
[80, 1950] loss: 0.421
[80, 1980] loss: 0.465
[80, 2010] loss: 0.453
[80, 2040] loss: 0.500
[80, 2070] loss: 0.470
[80, 2100] loss: 0.519
[80, 2130] loss: 0.382
[80, 2160] loss: 0.395
[80, 2190] loss: 0.492
ovft_intermediate_models/epoch79_shallow.pt saved.

evaluating on training set...
Accuracy on test images: 0.911123
epoch 79 accuracy on train set is: 0.9111232279171211
evaluating on training set takes 64 seconds.
shallow_train shallow_test scratch_train scratch_test deep_train deep_test
40      0.910987      0.91703          NaN          NaN      0.987027    0.986732

```

	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184
55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	0.917757	NaN	NaN	0.992253	0.993275
77	0.916099	0.925027	NaN	NaN	0.990549	0.992821
78	0.915553	0.917212	NaN	NaN	0.991594	0.993003
79	0.911123	NaN	NaN	NaN	0.990731	0.993639

evaluating on test set...

Accuracy on test images: 0.930934

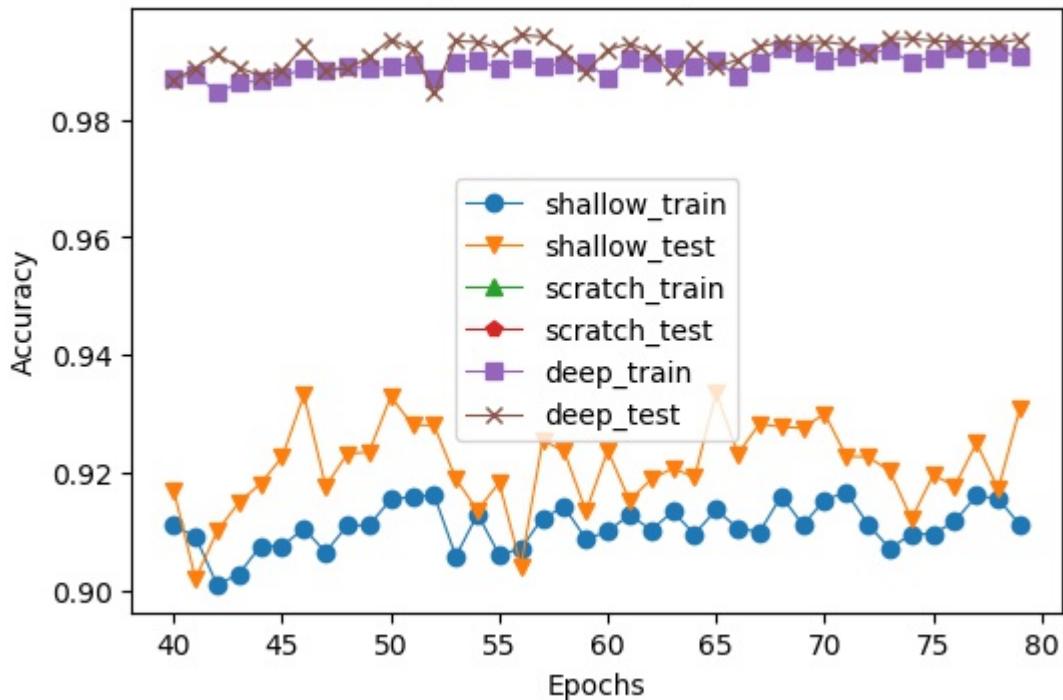
epoch 79 accuracy on test set is: 0.930934205743366

evaluating on test set takes 16 seconds.

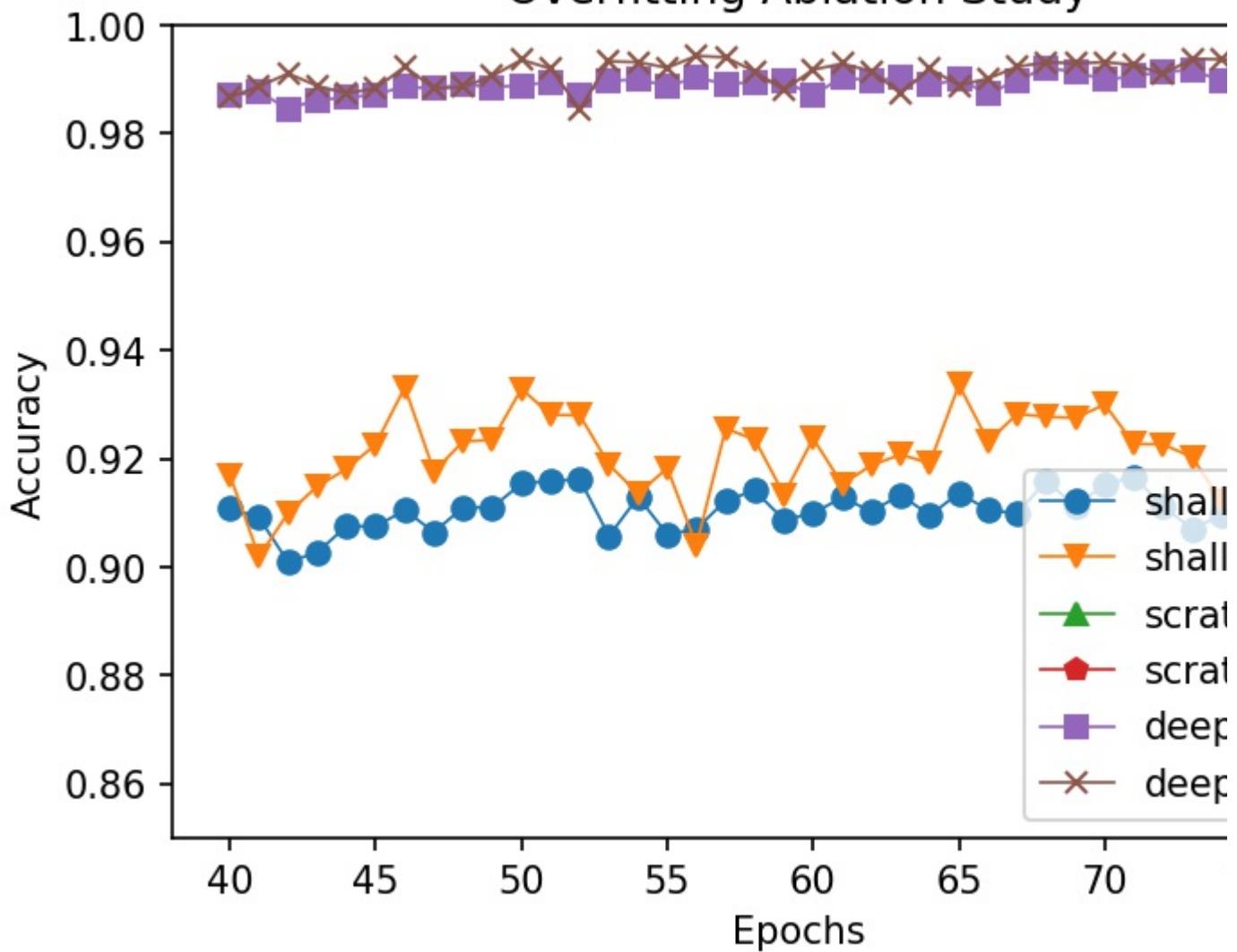
	shallow_train	shallow_test	scratch_train	scratch_test	deep_train	deep_test
40	0.910987	0.91703	NaN	NaN	0.987027	0.986732
41	0.909147	0.901945	NaN	NaN	0.987709	0.988731
42	0.900945	0.910033	NaN	NaN	0.984551	0.991094
43	0.902763	0.91494	NaN	NaN	0.986255	0.988822
44	0.907556	0.918121	NaN	NaN	0.986619	0.987459
45	0.907511	0.922664	NaN	NaN	0.987255	0.988368
46	0.910533	0.933206	NaN	NaN	0.988709	0.992366
47	0.906466	0.917666	NaN	NaN	0.988277	0.988368
48	0.911055	0.92321	NaN	NaN	0.989095	0.98864
49	0.911032	0.923482	NaN	NaN	0.988595	0.99064
50	0.915463	0.932933	NaN	NaN	0.988913	0.993639
51	0.915917	0.928117	NaN	NaN	0.989527	0.992003
52	0.916212	0.928026	NaN	NaN	0.987118	0.984551
53	0.905762	0.918939	NaN	NaN	0.989799	0.993366
54	0.912963	0.913577	NaN	NaN	0.989981	0.993184

55	0.906057	0.918302	NaN	NaN	0.988822	0.992185
56	0.907079	0.904035	NaN	NaN	0.990322	0.994366
57	0.912327	0.925573	NaN	NaN	0.989186	0.994093
58	0.914122	0.923573	NaN	NaN	0.989458	0.991458
59	0.908738	0.913395	NaN	NaN	0.989663	0.988095
60	0.909965	0.923755	NaN	NaN	0.987118	0.991821
61	0.912873	0.915304	NaN	NaN	0.990435	0.992912
62	0.910283	0.918848	NaN	NaN	0.989754	0.991549
63	0.913372	0.920756	NaN	NaN	0.990322	0.987459
64	0.909556	0.919211	NaN	NaN	0.988959	0.992003
65	0.913713	0.93366	NaN	NaN	0.990004	0.988913
66	0.910555	0.92321	NaN	NaN	0.987368	0.990185
67	0.909919	0.928299	NaN	NaN	0.989867	0.992366
68	0.916031	0.927754	NaN	NaN	0.992185	0.993184
69	0.911123	0.927572	NaN	NaN	0.991367	0.993003
70	0.915372	0.930025	NaN	NaN	0.989935	0.993184
71	0.916644	0.922755	NaN	NaN	0.990617	0.99273
72	0.911214	0.922664	NaN	NaN	0.991389	0.991003
73	0.907102	0.920302	NaN	NaN	0.991753	0.99382
74	0.909578	0.912123	NaN	NaN	0.989731	0.99373
75	0.909578	0.919575	NaN	NaN	0.990594	0.993366
76	0.911737	0.917757	NaN	NaN	0.992253	0.993275
77	0.916099	0.925027	NaN	NaN	0.990549	0.992821
78	0.915553	0.917212	NaN	NaN	0.991594	0.993003
79	0.911123	0.930934	NaN	NaN	0.990731	0.993639

Overfitting Ablation Study - whole view



Overfitting Ablation Study



epoch 79 takes 149 seconds.

Finished Training

Evaluating...

In [2]: