

Task 1

Read this week's mandatory paper¹. Explain in up to two sentences: What advantages does capsule network have compared to standard CNN?

1. CNN has an inefficiency of replicating feature detectors on a grid and due to max-pooling it will lose spatial information about the feature.
2. Because of dynamic routing the capsule network is more effective, and it does not throw away information about the precise position of the entity within the searching region.

Task 2

Task 2.2

Report your accuracy on the dev set after training.

Accuracy of simple CNN: 0.769517

Task 2.3

Epoch 00001: val_acc improved from -inf to 0.81253, saving model to .\task2_3.hdf5
Epoch 00002: val_acc improved from 0.81253 to 0.83006, saving model to .\task2_3.hdf5
Epoch 00003: val_acc improved from 0.83006 to 0.83218, saving model to .\task2_3.hdf5
Epoch 00004: val_acc did not improve from 0.83218
Epoch 00005: val_acc improved from 0.83218 to 0.83324, saving model to .\task2_3.hdf5
Epoch 00006: val_acc improved from 0.83324 to 0.83537, saving model to .\task2_3.hdf5
Epoch 00007: val_acc improved from 0.83537 to 0.83590, saving model to .\task2_3.hdf5
Epoch 00008: val_acc did not improve from 0.83590
Epoch 00009: val_acc did not improve from 0.83590
Epoch 00009: early stopping
32/1883 [.....] - ETA: 0s
832/1883 [=====>.....] - ETA: 0s
1664/1883 [=====>....] - ETA: 0s
1883/1883 [=====] - 0s 62us/step
Accuracy (dev) : 0.8359001595101582

32/1883 [.....] - ETA: 0s
864/1883 [=====>.....] - ETA: 0s
1664/1883 [=====>....] - ETA: 0s
1883/1883 [=====] - 0s 63us/step
Accuracy (test) : 0.8353690920645926

Task 2.4

Apply random hyperparameter optimization. Train ten hyperparameter configurations on the training set and evaluate them on the dev set. Then, report the results of your best found configuration on the test set.

test 1:

Params (filters, kernel_size, strides):{'layers': [(37, 2, 1)]}

Epoch 00001: val_acc improved from -inf to 0.60754, saving model to .\task2_4_test_1.hdf5

Epoch 00002: val_acc improved from 0.60754 to 0.73500, saving model to .\task2_4_test_1.hdf5

Epoch 00003: val_acc improved from 0.73500 to 0.78120, saving model to .\task2_4_test_1.hdf5

Epoch 00004: val_acc improved from 0.78120 to 0.79501, saving model to .\task2_4_test_1.hdf5

Epoch 00005: val_acc improved from 0.79501 to 0.79873, saving model to .\task2_4_test_1.hdf5

Epoch 00006: val_acc did not improve from 0.79873

Epoch 00007: val_acc improved from 0.79873 to 0.79926, saving model to .\task2_4_test_1.hdf5

Epoch 00008: val_acc improved from 0.79926 to 0.79979, saving model to .\task2_4_test_1.hdf5

Epoch 00009: val_acc improved from 0.79979 to 0.80244, saving model to .\task2_4_test_1.hdf5

Epoch 00010: val_acc did not improve from 0.80244

Epoch 00011: val_acc improved from 0.80244 to 0.80404, saving model to .\task2_4_test_1.hdf5

Epoch 00012: val_acc did not improve from 0.80404

Epoch 00013: val_acc did not improve from 0.80404

Test 1 Accuracy (dev) : 0.8040361125546444