Theory and Practice of Deep Learning

Chang Jun Qing 1002088

Week 5

Training on all layers of model with no weights loaded
Pretrained model used: Resnet
The following images shows the first five and last four training epochs. In each epoch the model was trained and then validated against the training and validation set respectively.

```
Train Epoch: 1 [5687/5687 (100%)]       Loss: 3.730564

Performance: Accuracy: 146/1224 (11.93%), Loss: 5102.593184

Train Epoch: 2 [5687/5687 (100%)]       Loss: 3.235626

Performance: Accuracy: 318/1224 (25.98%), Loss: 4127.380423

Train Epoch: 3 [5687/5687 (100%)]       Loss: 3.758488

Performance: Accuracy: 306/1224 (25.00%), Loss: 3935.562722

Train Epoch: 4 [5687/5687 (100%)]       Loss: 3.308613

Performance: Accuracy: 352/1224 (28.76%), Loss: 3839.329845

Train Epoch: 5 [5687/5687 (100%)]       Loss: 3.313563

Performance: Accuracy: 479/1224 (39.13%), Loss: 2971.998009
```

Figure 1: First Five epochs

```
Train Epoch: 27 [5687/5687 (100%)]      Loss: 0.082206

Performance: Accuracy: 1011/1224 (82.60%), Loss: 856.458770

Train Epoch: 28 [5687/5687 (100%)]      Loss: 1.326663

Performance: Accuracy: 1033/1224 (84.40%), Loss: 669.895121

Train Epoch: 29 [5687/5687 (100%)]      Loss: 1.197884

Performance: Accuracy: 1024/1224 (83.66%), Loss: 781.961175

Train Epoch: 30 [5687/5687 (100%)]      Loss: 0.049238

Performance: Accuracy: 1034/1224 (84.48%), Loss: 760.473899

Best Loss: 669.8951206207275
```

Figure 2: Last Four epochs

As seen, the best loss on the validation set is 669.895
The following are the training loss over epoch, validation loss over epoch, as well as the validation accuracy over epoch

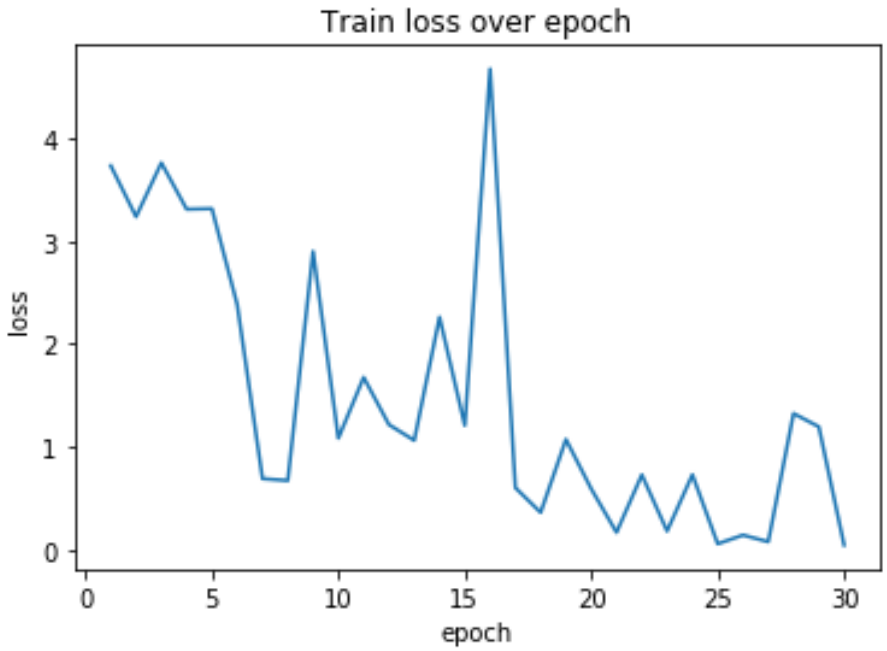

Figure 3: Training loss

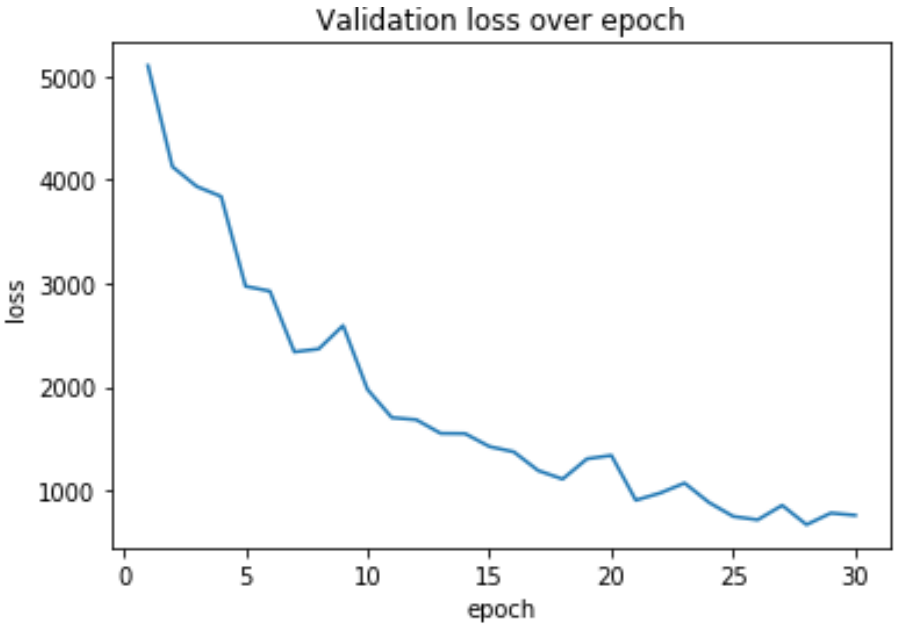

Figure 4: Validation loss

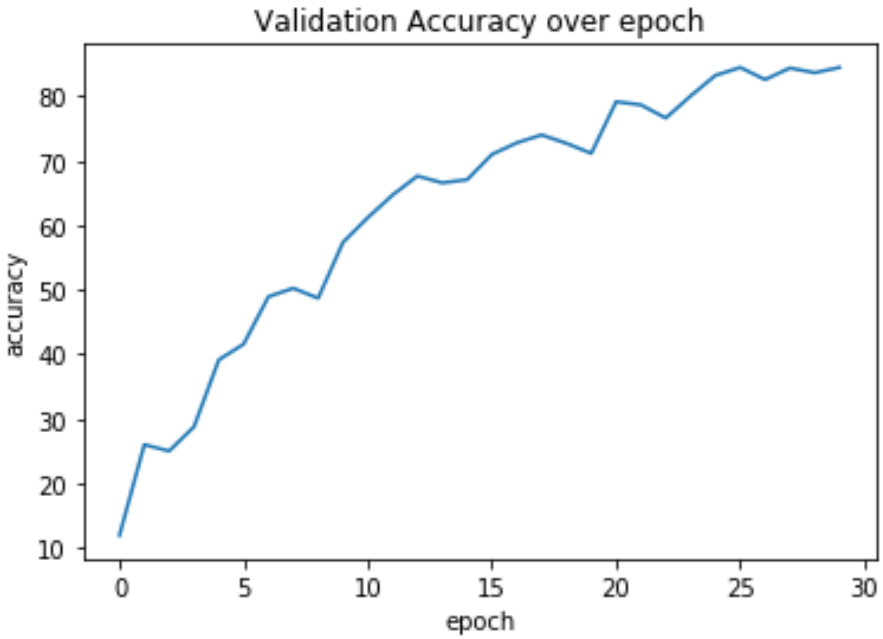

Figure 5: Validation Accuracy

Finally, the model is used to predict the test set accuracy

```
Performance: Accuracy: 1050/1278 (82.16%), Loss: 890.002773
```

Figure 6: Test Accuracy

Training on all layers of models with weights loaded
Pretrained model used: Resnet
The following images shows the first five and last four training epochs. In each epoch the model was trained and then validated against the training and validation set respectively.

```
Train Epoch: 1 [5687/5687 (100%)]       Loss: 3.106496

Performance: Accuracy: 1000/1224 (81.70%), Loss: 802.353889

Train Epoch: 2 [5687/5687 (100%)]       Loss: 0.223955

Performance: Accuracy: 1146/1224 (93.63%), Loss: 294.661448

Train Epoch: 3 [5687/5687 (100%)]       Loss: 0.250084

Performance: Accuracy: 1161/1224 (94.85%), Loss: 200.938255

Train Epoch: 4 [5687/5687 (100%)]       Loss: 0.101890

Performance: Accuracy: 1175/1224 (96.00%), Loss: 179.085724

Train Epoch: 5 [5687/5687 (100%)]       Loss: 0.351579

Performance: Accuracy: 1172/1224 (95.75%), Loss: 181.348999
```

Figure 7: First Five epochs

```
Train Epoch: 27 [5687/5687 (100%)]      Loss: 0.442047

Performance: Accuracy: 1193/1224 (97.47%), Loss: 100.711459

Train Epoch: 28 [5687/5687 (100%)]      Loss: 0.005490

Performance: Accuracy: 1196/1224 (97.71%), Loss: 100.351534

Train Epoch: 29 [5687/5687 (100%)]      Loss: 0.036854

Performance: Accuracy: 1180/1224 (96.41%), Loss: 145.959862

Train Epoch: 30 [5687/5687 (100%)]      Loss: 0.004236

Performance: Accuracy: 1189/1224 (97.14%), Loss: 128.548362

Best Loss: 100.35153388977051
```

Figure 8: Last Four epochs

As seen, the best loss on the validation set is 100.3515
The following are the training loss over epoch, validation loss over epoch, as well as the validation accuracy over epoch

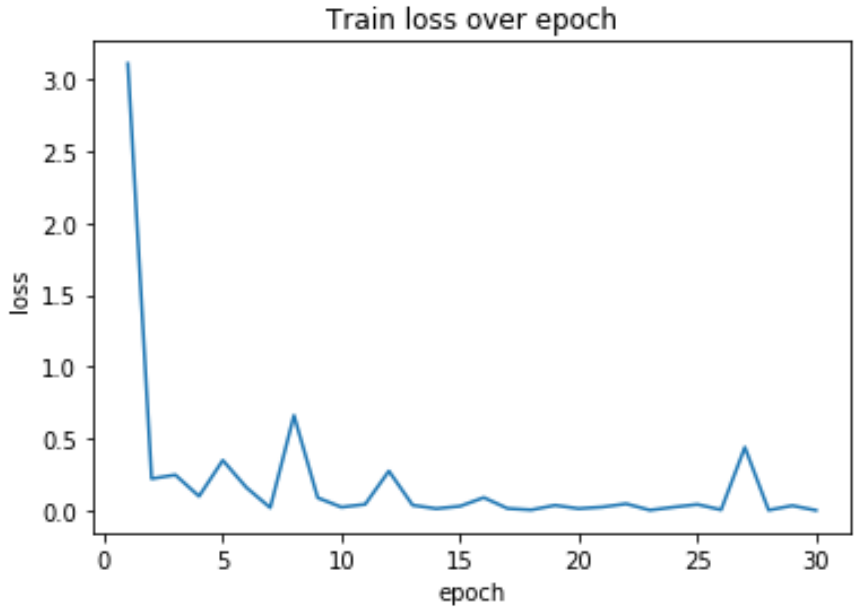

Figure 9: Training loss

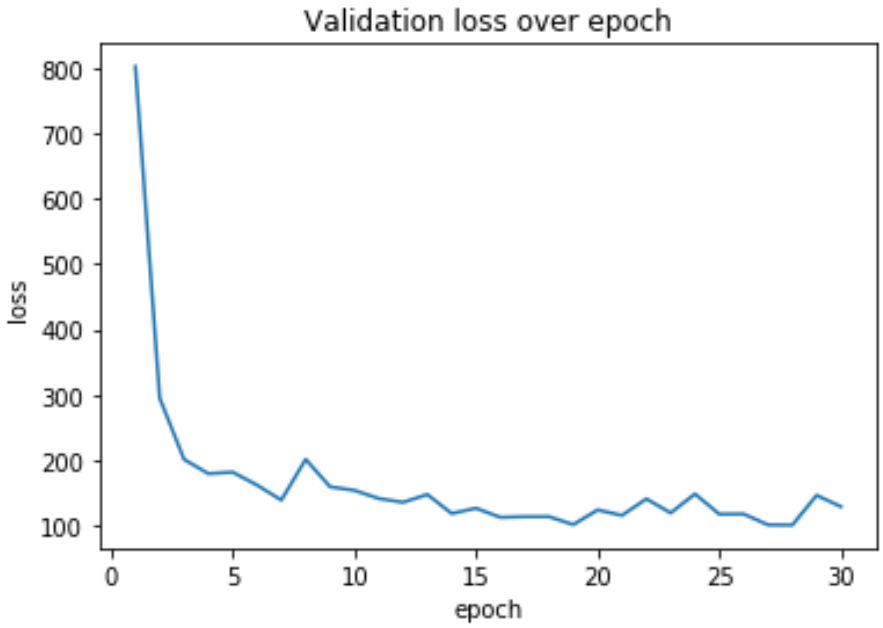

Figure 10: Validation loss

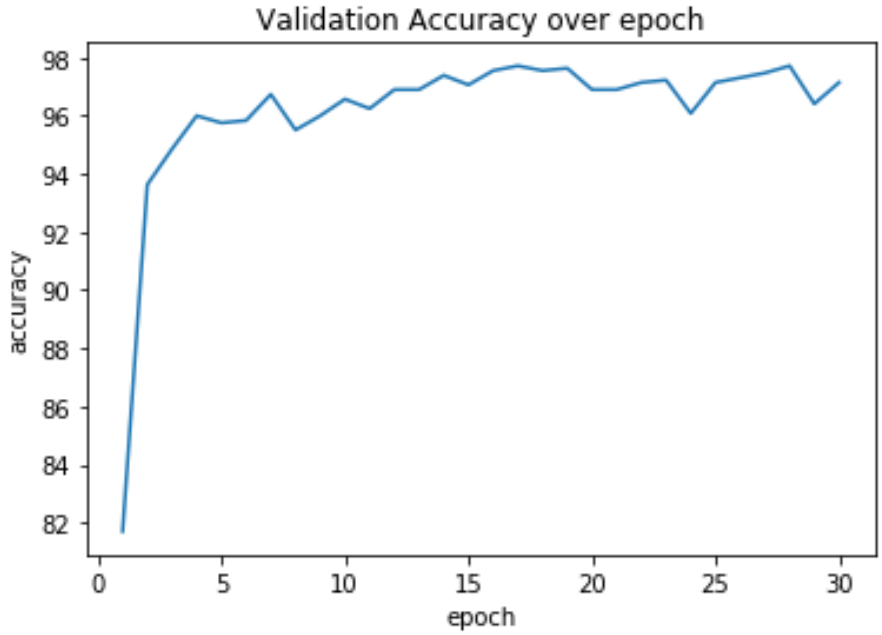

Figure 11: Validation Accuracy

Finally, the model is used to predict the test set accuracy

```
Performance: Accuracy: 1240/1278 (97.03%), Loss: 153.227873
```

Figure 12: Test Accuracy

Training on last 2 layers of models with weights loaded
Pretrained model used: Resnet
The following images shows the first five and last four training epochs. In each epoch the model was trained and then validated against the training and validation set respectively.

```
Train Epoch: 1 [5687/5687 (100%)]       Loss: 1.511928

Performance: Accuracy: 1115/1224 (91.09%), Loss: 408.402533

Train Epoch: 2 [5687/5687 (100%)]       Loss: 0.977143

Performance: Accuracy: 1148/1224 (93.79%), Loss: 262.865331

Train Epoch: 3 [5687/5687 (100%)]       Loss: 2.399775

Performance: Accuracy: 1169/1224 (95.51%), Loss: 203.600915

Train Epoch: 4 [5687/5687 (100%)]       Loss: 0.191064

Performance: Accuracy: 1171/1224 (95.67%), Loss: 180.551285

Train Epoch: 5 [5687/5687 (100%)]       Loss: 0.226763

Performance: Accuracy: 1180/1224 (96.41%), Loss: 157.041892
```

Figure 13: First Five epochs

```
Train Epoch: 27 [5687/5687 (100%)]      Loss: 0.007732

Performance: Accuracy: 1182/1224 (96.57%), Loss: 150.978821

Train Epoch: 28 [5687/5687 (100%)]      Loss: 0.080040

Performance: Accuracy: 1193/1224 (97.47%), Loss: 120.308043

Train Epoch: 29 [5687/5687 (100%)]      Loss: 0.031040

Performance: Accuracy: 1189/1224 (97.14%), Loss: 134.530454

Train Epoch: 30 [5687/5687 (100%)]      Loss: 0.150025

Performance: Accuracy: 1191/1224 (97.30%), Loss: 130.355733

Best Loss: 120.30804300308228
```

Figure 14: Last Four epochs

As seen, the best loss on the validation set is 120.3080
The following are the training loss over epoch, validation loss over epoch, as well as the validation accuracy over epoch

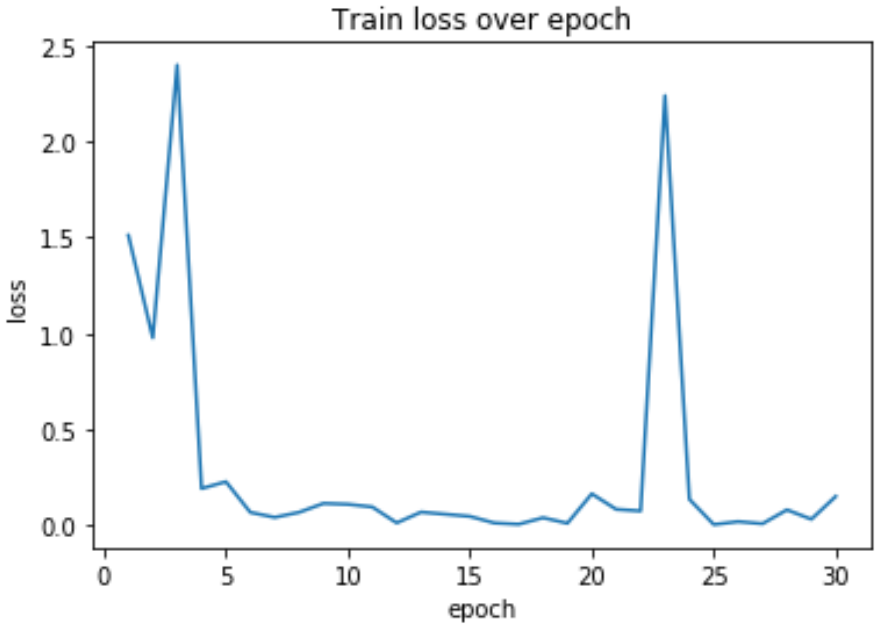

Figure 15: Training loss

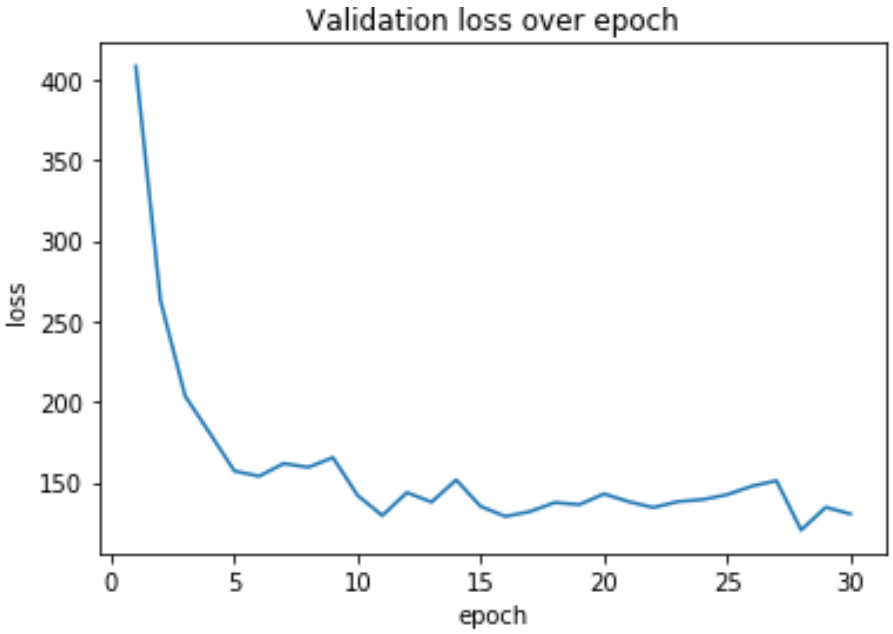

Figure 16: Validation loss

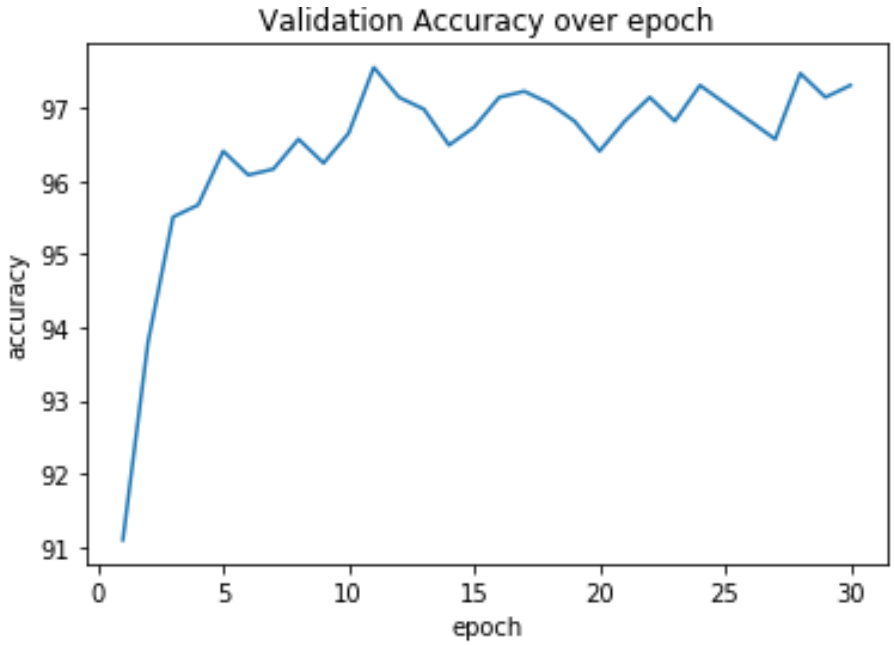

Figure 17: Validation Accuracy

Finally, the model is used to predict the test set accuracy

```
Performance: Accuracy: 1235/1278 (96.64%), Loss: 154.699249
```

Figure 18: Test Accuracy

Summary

Out of 3 models, the best performance is the model trained on all layers with weights loaded. When no weights are loaded, the accuracy steadily increases from 10% since the model has no initial weights to start with. When weights are loaded, accuracy is approximately in the 80s to 90s region. By freezing layers in the last model, the accuracy is lower, possibly due to the fact that the other layers are unable to be trained by the training data.

In models B and C, loss takes a dip before plateauing, and as for A, the loss moves around quite a bit since there are no weights to start with.

Additional Info

Models were trained on Python 3.6.6 and Torch version 1.0.0 with a Windows 10 computer equipped with a Nvidia GTX 1060