

Lab 6 Exercise - Reflections on transfer learning

Luke McClure

29573904

	precision	recall	f1-score	support
Alilaguna	0.43	0.47	0.45	19
Ambulanza	0.38	0.41	0.39	22
Barchino	0.30	0.18	0.22	51
Gondola	0.00	0.00	0.00	3
Lanciafino10m	0.00	0.00	0.00	7
Motobarca	0.11	0.08	0.09	59
Motopontonerettangolare	0.60	1.00	0.75	3
MotoscafoACTV	0.00	0.00	0.00	1
Mototopo	0.77	0.73	0.75	274
Patanella	0.34	0.54	0.42	74
Polizia	0.08	0.13	0.10	15
Raccoltarifiuti	0.40	0.63	0.49	19
Sandoloaremi	0.00	0.00	0.00	3
Topa	0.07	0.03	0.05	29
VaporettoACTV	0.97	0.99	0.98	325
Water	0.95	0.92	0.93	420
accuracy			0.75	1324

Table 1: Classification report of finetuned ResNet50

	precision	recall	f1-score	support
Alilaguna	0.90	1.00	0.95	19
Ambulanza	0.82	0.82	0.82	22
Barchino	0.76	0.25	0.38	51
Gondola	1.00	0.67	0.80	3
Lanciafino10m	0.00	0.00	0.00	7
Motobarca	0.82	0.31	0.44	59
Motopontonerettangolare	1.00	1.00	1.00	3
MotoscafoACTV	0.00	0.00	0.00	1
Mototopo	0.82	0.99	0.89	274
Patanella	0.41	0.84	0.55	74
Polizia	0.67	0.13	0.22	15
Raccoltarifiuti	1.00	0.74	0.85	19
Sandoloaremi	0.00	0.00	0.00	3
Topa	0.00	0.00	0.00	29
VaporettoACTV	0.99	1.00	1.00	325
Water	0.99	0.97	0.98	420
accuracy			0.87	1324

Table 2: Classification report of SVC

1 TRANSFER LEARNING

1.1 Finetuning

When finetuning the ResNet50 model I reasoned that the convolutional layers have already been trained to recognise shapes and

objects, therefore it would be inefficient to try and retrain this part of the network as it would lead to overfitting. This left the final fully connected layers to be fine tuned, the pooling layer was replaced with an average pooling layer and the final fully connected output layer connecting the 2048 inputs to 16 outputs representing the dataset classes.

When retraining with fine tuning, the layers above the tuning's gradients must be frozen so that only the new layers are backpropagated into. It is suggested that the learning rate for this retraining should be small so as to not change the weight too much from their original learned values, but there is some opportunity to experiment with this hyperparameter.

Learning rate	Test Accuracy
1e-4	0.714
1e-3	0.739
1e-2	0.699

It is clear that a learning rate of 1e-3 was optimal for this task, however this is within 1% of the accuracy produced by BetterCNN alone.

1.2 Reflect on the two different approaches

There were two approaches explored in lab 6, finetuning an addition to an already trained and complex network, and training an SVM classifier using the features that were trained from this pretrained ResNet50 model.

Finetuning the ResNet50 network to fit to this problem was relatively slow due to the number of learnable parameters within the fully connected layer of the network, learning over 10 epochs with the ideal found learning rate of $1e^{-3}$ provides the classification report Table 1.

As the Euclidean distance between features produced by the model of the same class tend to be closer than features of difference classes, it is possible to train a Support Vector Machine on the features produced by the ResNet50 model rather than using a fully connected layer to conduct the classification.

This method allows for much faster training than finetuning (sub 20 seconds), and as seen in Table 2 produced a much higher prediction accuracy across the validation set.

In this brief investigation I found that the SVM Classifier is significantly faster than the finetuning deep learning method for creating a classifier on this dataset, it also comes with a 16% prediction accuracy increase over the finetuning method.