

ML2016 HW3 report b02902096

資工四 王浩恩

1. Supervised Learning:

Code很長就不貼上來，我的supervised learning使用了6層convolution layers最後再接兩層fully-connected的Dense，optimizer選用了adadelta，跑大約60~70個epoch，kaggle public的分數可以達到0.68左右，此外我還嘗試過adam， $lr = 1e^{-4}$ 時的結果比 $1e^{-5}$ 跟 $1e^{-6}$ 好，但是仍然遜色於adadelta，所以最後採用adadelta，他真的跑得很快。

2. Semisupervised Learning:

我在supervised的架構之下，多使用了ImageDataGenerator這個函式，並且多跑了五輪15個epoch的training，每一輪我都會先用現有的model去predict unlabeled data，然後把probability > 0.999的unlabeled data當作labeled data併入原本的training data去train，最後public board的結果可以達到0.78左右，private board則達到0.80

3. Semisupervised Learning 2:

我先做了一個fully-connected的autoencoder，把所有照片都先pretrain過後，再丟入CNN中train，得出來的結果很不準，我想可能用convolutional autoencoder的效果可能會更好。

4. Compare and analyze your results

method	supervised learning	self-training with unlabeled data	autoencoder + CNN
best performance	0.68	0.78	<0.3

self-training with unlabeled data的表現明顯優於原本只使用labeled data的表現，除了img_gen的使用以外，應該是因為不停使用現有model去製造新的training data，又同時不停的更新model，形成一個良性循環，讓acc明顯上升。

第三種方法的performance遠低於其他兩種的原因，我認為是我做的autoencoder強迫將 $3 \times 32 \times 32$ 維的資訊壓在256維中，必定會有一些information loss，因此造成training上的失準。