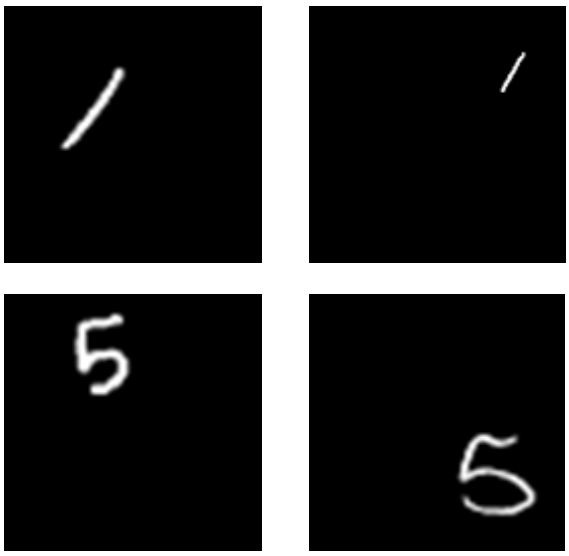
進階機器學習 Advanced Machine Learning

Homework #2
Due 2024 March 20 11:00PM

(一) The dataset ***HW2_MNIST_train.zip*** contains handwritten digit images. Different from the standard MNIST dataset, the resolution of each image is 128 x 128. Each image contains **one** digit from 0 to 9. There are two different sizes of handwritten digits in the image as shown in the following examples. The digit can be located in any location of the image.



The label of each image contains the class of the images with the associated bounding box which includes the (x,y) location, width and height of the digit.

In this assignment, you should design a deep neural network which only has to predict the digit number of the image. You should provide:

1. (15%) The confusion matrix, the top-1 and top-3 accuracy of your training result.
2. (15%) List number of layers and parameters used in your architecture.

   Calculate the value of $\frac{top-1\ accuracy}{\#\ of\ parameter}$. If the metric is higher, you'll get higher score for this sub-problem.
3. (70%) The prediction result of the enclosed ***HW2_MNIST_test.zip***. You have to submit the entire program and a CSV file ***HW2_prob1.csv*** with the following format.

	A	B	C
1	image	class	
2	0000000.png	0	
3	0109539.png	8	

All the files used in this homework can be downloaded from the website:

https://drive.google.com/drive/folders/1a8rGb-6sjMFsWbgCnwAIhNRUDnpy34FY?usp=share_link

(二) Suppose we have an input feature map with a size of 4x4, and the features will pass through a convolution layer with a kernel size of 3x3. The activation function used in convolution is *sigmoid*. The convolution results will be flattened and multiplied by the corresponding weights. The products will be added together to produce a final regression value. Assume the correct label value is *yy*, and the MSE has been used as the loss function. Derive the loss gradient with respect to the weight *w11* and the loss gradient with respect to the input feature *a00* shown in the following figure.