AC50001 Assignment on Image Classification using Keras

This assignment contributes 20% to your mark for AC50001. It might take you 15-20 hours.

Before attempting this assignment, you should have worked through the module lab exercises that use notebooks to train and test classification and convolutional neural networks in Keras.

Getting Set Up

In this assignment you will write notebooks that specify, train, and evaluate models to classify medical images. You will use MedMNIST, a collection of 10 datasets each of which consists of small images that are 28x28 pixels (https://medmnist.github.io/). To help you get started, you are provided with a notebook (written by Jacob Carse) which loads the digital pathology MedMNIST dataset and sets up and trains a convnet on it. You should access this notebook at the Colab link below and copy it to your own account (*Copy to Drive*).

https://colab.research.google.com/drive/1LRlehId853iY2QytjFdZJuNb4mJfdNoY?usp=sharing

Start by running and understanding this notebook. Make sure you select *GPU* under *Change Runtime Type* on the *Runtime* tab; this will enable you train convolutional networks considerably faster than on a CPU.

Your Assignment

You should select two of the MedMNIST datasets, one of them should be a multi-class classification task (not PathMNIST), and the other should be a binary classification task. For each of your two chosen datasets you should develop and evaluate three different methods for classification:

1. A convnet with the same architecture and training regime as in the example Colab notebook provided above. (You should alter the last layer and the loss function as appropriate, depending on the number of classes).
2. A network with dense layers but no convolutional layers, designed to have a total number of parameters comparable to the convnet in 1.
3. A deep network *of your own design*, incorporating design techniques and regularisation methods you have learned about. You should use data augmentation. Try to achieve a good performance on the test data (with a network that does not take days to train!)

You should submit your code, along with text explanations, results, and conclusions, in the form of one or more notebooks on Colab. Notebook text should make clear what you have done and why. The results should be presented, using plots and tables as appropriate, in a way that enables direct comparisons to be made between the performance of the different methods. Based on the evidence provided by these experiments, you should draw conclusions about the relative performance of the methods using appropriate evaluation metrics.

Finally, please include a note of how long you spent on this assignment.

(A marking sheet is provided and will be used to provide feedback on your assignment.)