**Weekly report**

*Work on TFRecords and CASIA dataset*

Went back to basics and tried to increase the efficiency in the way in which we resized a character.

Realised this isn’t necessary, since we only generate the character images once.

Generated the character images again, this time to a larger size (as suggested in literature) of 48x48, including 2 pixels of padding on each side.

Incorporated 4030 unique characters, instead of 3755 as before. This includes alphanumeric characters (a,b,… 1,2,…)

TFRecords requires images are saved as images, not arrays – this was done and each image was labelled.

All .gnt files from the 1.0 dataset can now be loaded, and they are separated into training and test samples as recommended by CASIA

Wrote to TFRecords successfully. Still need to read from these and implement them in a network – and this is the only way we can check if they were written correctly.

*Work on transfer learning*

With transfer learning I have made a script that can load and feed data into a CNN and another script that can load and print out bottlenecks from a CNN. In addition, I have started a script that will read in the bottlenecks I save from a loaded model and uses them to train the final layer of a CNN but with 30 outputs instead of 10.

It is worth mentioning that this work will be good as a rough proof of concept however once we have the dataset complete, we will have to retrain the CNN with 10 outputs so that we can keep training and testing data separate. In addition these scripts will have to be altered to do everything in TFrecords.

**Action points for the next week**

1. Work on the bottleneck script so that it can save the bottlenecks as a dataset.

2. Continue to work on a script to retrain the final layer of a CNN using loaded bottlenecks