COMP90086 – Computer Vision – 2022 S2
Assignment 2 – CNN for Image Classification

Student: Jiahao Chen ID: 1118749

I. CNN Implementation

As figure 1 depicts, the basic CNN model turns out to be overfitted after training on the yoga32 dataset. After implementing regularisation and data augmentation, the problem has been significantly improved as figure 2 shows.

For data augmentation, flipping and cropping and applied, which are two common options. This process can generate more samples for the network.

For regularisation, Lasso regression is applied, which reduces the less important features' coefficient to zero so that only important features are kept.

The modified model has a higher validation accuracy than the original one. However, the modified one tends to be more fluctuated during training. It is probably because the dataset is too small.

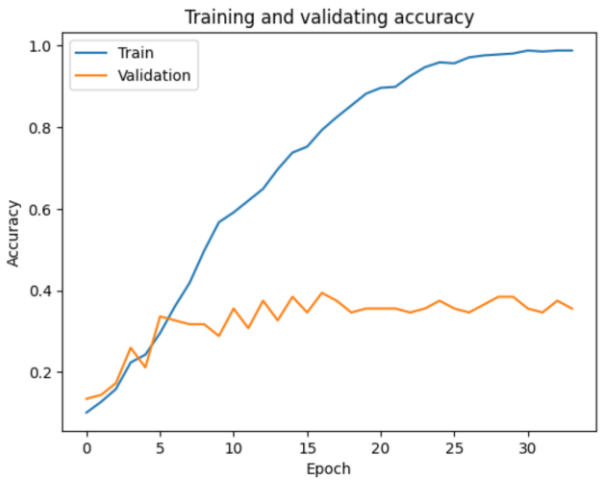

Figure 1: Basic model

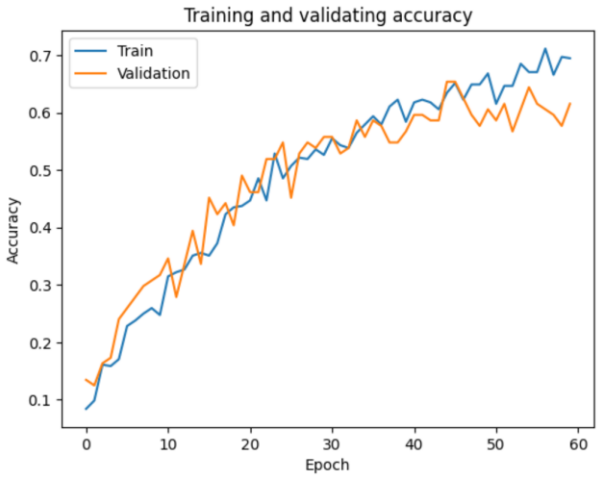

Figure 2: Modified model

II. Error Analysis

The overall accuracy of the CNN model is 0.586. Average accuracy for each of the 10 class is presented in figure 3. Mountain and tree have the highest accuracies while childs and plank have the lowest one.

Class	Accuracy
Bridge	0.43
Childs	0.14
Downward dog	0.71
Mountain	0.86
Plank	0.29
Seated forward bend	0.71
Tree	0.86
Triangle pose	0.71
Warrior1	0.57
Warrior2	0.57

Figure 3: Accuracy for each class

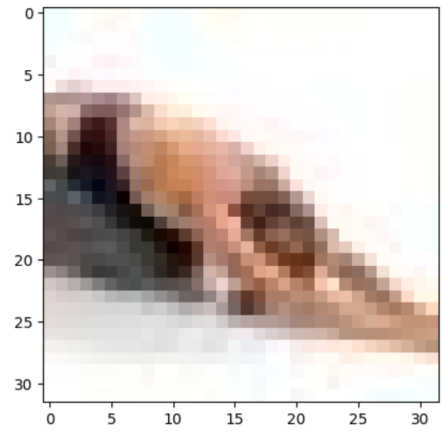

Figure 4: Misclassified childs

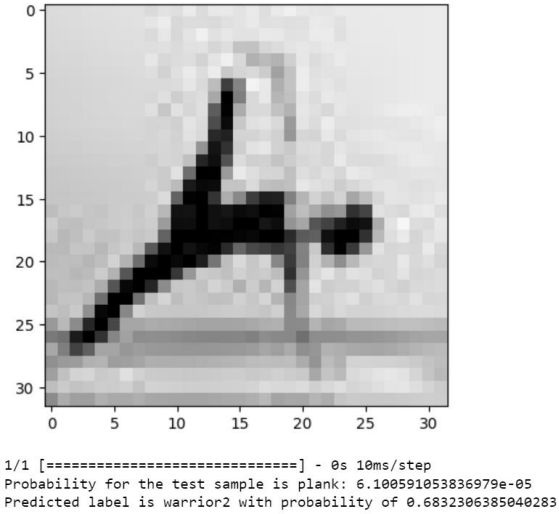

Figure 5: Misclassified plank

From observations, images of mountain and tree almost have very few changes, which may be much easier for the model to find the important features. However, for childs and plank, it

tends to be more complex. For example, as figure 4 shows, it is ambiguous to judge if the image is childs or seated forward bend and both classes have high probabilities. Figure 5 presents a misclassified image which should be plank instead of warrior2. There are a lot of variations in plank so that it is very difficult for the model to know the important features of this class. And the plank in figure 5 is similar to warrior2, which is stretching arms and legs.

There are two possible reasons which lead to this performance. On the one hand, the dataset is too small so that the model cannot obtain enough useful information for complex classes. On the other hand, the images are down sampled to 32 * 32, which may lose some information because some images of different classes seem similar.

III. Visualisation

For each class, a test image is picked to perform nearest neighbour analysis. The results are shown in figure 6. The first column is the test image and the following five images in the same row are the nearest neighbours from the training dataset.

Firstly, some classes can have a lot of similarities. For example, people are standing when doing mountain or tree. And many of them may have similar features on legs, which form a shape of triangle. This ambiguity may hugely confuse the network during training.

Secondly, some classes have many variations. For example, the given test images include 3 kinds of planks. For example, the plank image listed in figure 6 is very similar to warrior from the perspective of the model. For classes with less variations, such as mountain and tree, the model performs much better.

Finally, there are many examples shows that the model is able to detect horizontal reflections, however, it may not be good at predicting pictures with rotations. For instance, neighbours of the seated forward image listed in figure 6 are all different. It seems to be difficult for the model to recognize this class from the front inclined side.

In brief, the model manages to recognize some simple yoga positions. For more complex ones, it still has many problems. Increasing the dataset and adding more layers in the model may help to Improve the performance.


childs
childs
warrior1
childs
plank
seatedforwardbend
downwarddog
downwarddog
downwarddog
warrior1
seatedforwardbend
downwarddog
mountain
mountain
seatedforwardbend
bridge
mountain
mountain
plank
warrior2
warrior2
warrior1
warrior1
warrior2
seatedforwardbend
bridge
plank
downwarddog
plank
mountain
tree
tree
mountain
mountain
mountain
mountain

Figure 6: Nearest neighbours analysis