Abdullateef Shittu [sc19aas]

[Email address]

on Image Classification with  
Convolutional Neural Networks - ImageNet10

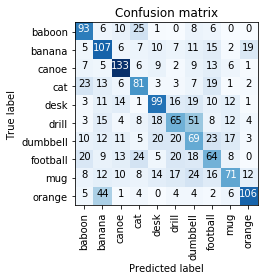
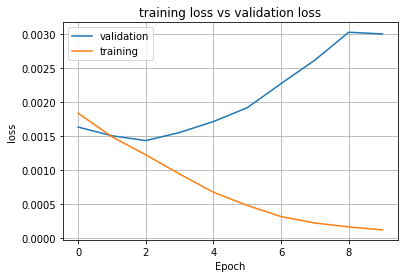
Coursework 1

COMP 5623

**Results to experiments**

2 layered neural network

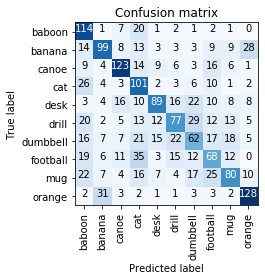
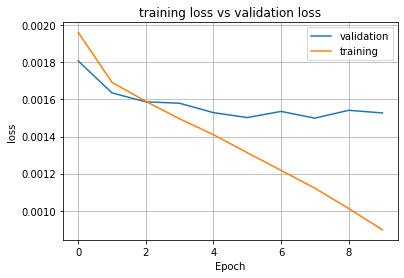
Accuracy 49%

GPU train time 498.216  
  
 

3 layered neural network

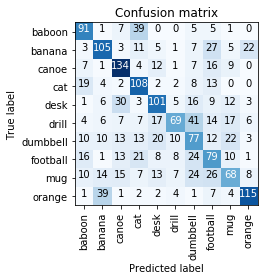
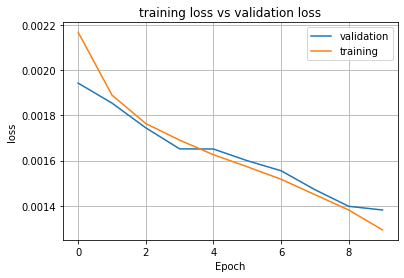
Accuracy 52%

GPU train time 495.471

4 layered neural network

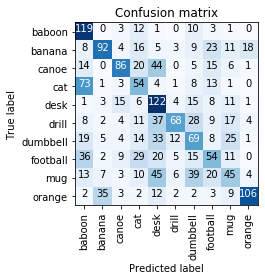
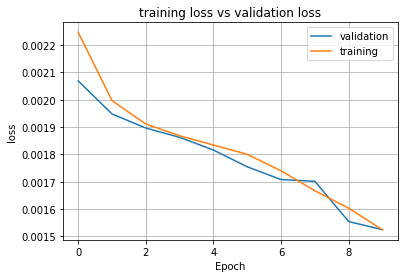
Accuracy 53%

GPU time 493.914  
  
 

5 layered neural network

Accuracy 43%

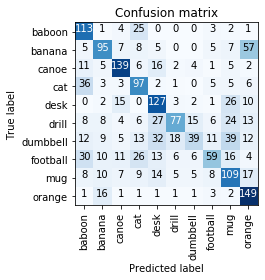
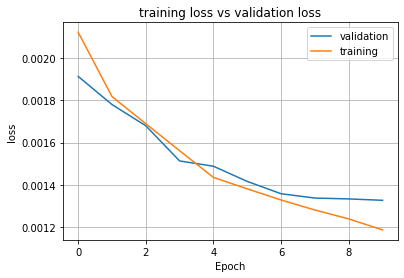
GPU time 498.745

From the results, its evident that the accuracy increased going from using 2 convolutional layers to 3 convolutional layers and also increased going from 3 convolutional layers to 4 convolutional layers, with the architecture of 4 convolutional layers having the best accuracy and then accuracy dropped when using 5 convolutional layers. I concluded that as more maxpool2d was used for every convolution layer , the sizes of the feature maps gradually reduced leading to fewer learning sets, which in turn reduced the accuracy

Max-pool kernel size = 3

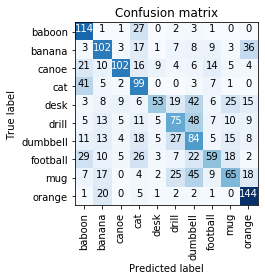
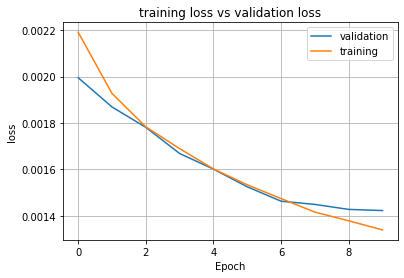
Accuracy 55%

GPU time 504.567  
  
 

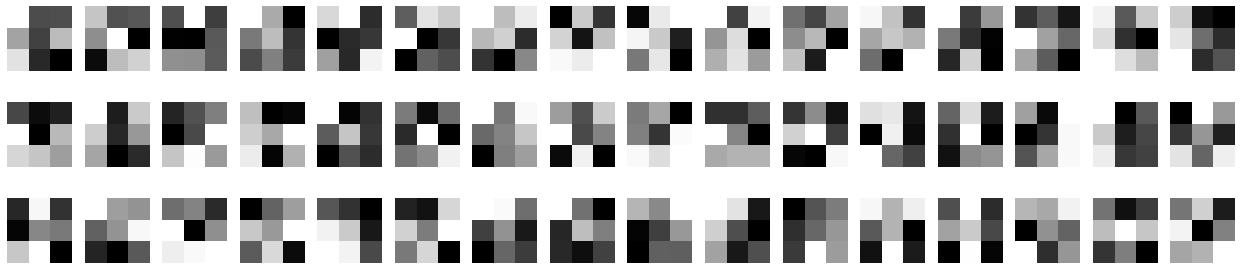
Max-pool kernel size = 4

Accuracy 49%

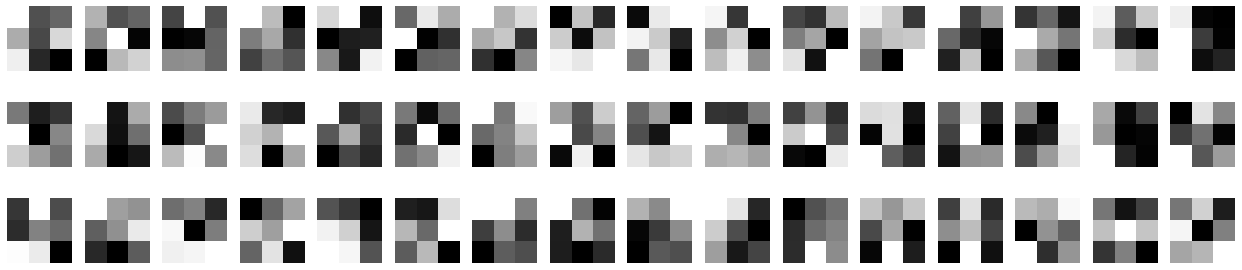
GPU time 493.095

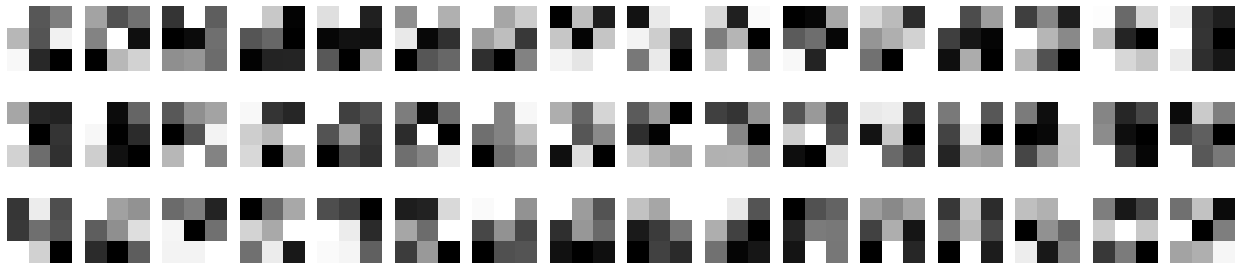
For experiment 2, maxpool2d was the architectural element tested. the accuracy decreased as the kernel size of the maxpooling increased once it got to a certain point. But for the most part an increase in kernel size should increase accuracy. In my experiment though once it got to 4 accuracy reduced.

**Filter visualization**  
  
result of filter visualization before training  


Result of filter visualization halfway during training



Result of filter visualization after training



Visualizing filters at different checkpoints allows us to understand how the CNN learns the complex spatial and temporal dependencies present in the image, which in turn allows us to improve performance of the model.

Investigating the results at three check points, we find that Weight are assigned random values before training(random weight initialisation) and as it learns, the values are updated and thus the visualizations start to change as it visualizes at halfway and then it changes even more when visualized after training.

**Feature maps visualizations**

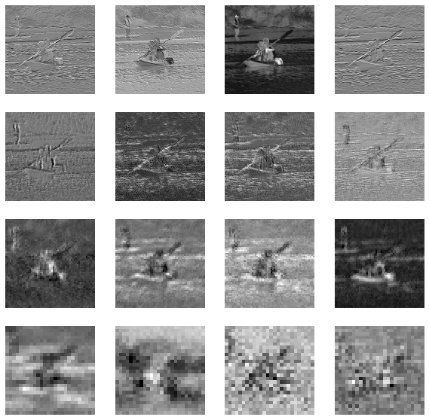
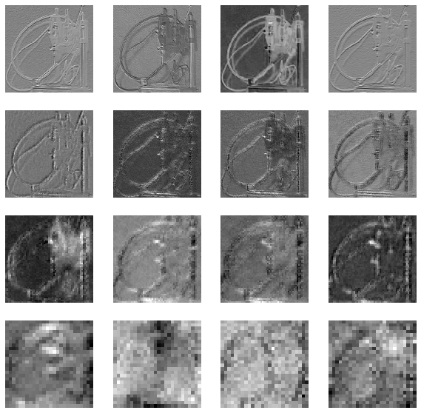
 ****

Image 1 image 2

As we went deeper in the network it became noisier and the pixels became more magnified, edges

**Improving network performance**  
  
The two adjustments investigated to improve the network performance were an increase in number of epochs and an increase in the operation, maxpool2d- kernel size.   
  
for the kernel size of maxpool2d operation, since it extracts the highest number per pool it eventually will increase accuracy as it drops the lesser numbers, reducing the computational cost by reducing the number of parameters to learn.

Increasing number of epochs increases the amount of times the model trains and learns on the set of data thereby increasing the accuracy on unseen data.  
  
A kernel size of 3 and a number of epochs of 30 was selected. The accuracy of the model was 60% the highest performance gotten throughout the coursework!  
  
