

COMP7606 A2 Q5

Layer (type)	Output Shape	Param #
vgg16 (Model)	(None, 7, 7, 512)	14714688
flatten_4 (Flatten)	(None, 25088)	0
dense_7 (Dense)	(None, 512)	12845568
dropout_4 (Dropout)	(None, 512)	0
batch_normalization_1 (Batch Normalization)	(None, 512)	2048
dense_8 (Dense)	(None, 4)	2052
Total params: 27,564,356		
Trainable params: 19,928,068		
Non-trainable params: 7,636,288		

Architecture: The model is fine-tuning on top of VGG16. Layers were trained starting from block5_conv1. Additional layers are:

- One dense layer with 512 units and ReLu: to aggregate previous feature results
- One dropout layer with rate 0.5: to reduce overfitting
- One batch normalization layer: to normalize the activations
- One dense layer with 4 units and softmax: to classify cat, dog, car, motorbike

Input Data: data augmentation was applied to training data.

```
aug_train_datagen = ImageDataGenerator(  
    rescale=1./255,  
    rotation_range=40,  
    width_shift_range=0.2,  
    height_shift_range=0.2,  
    shear_range=0.2,  
    zoom_range=0.2,  
    horizontal_flip=True,  
    fill_mode='nearest')
```

Input Size: I chose 224x224x3 instead of 150x150x3 as the input image size because VGG16 was using this input size, a larger size may have better performance.

Batch Size: 32

Optimizer: RMSprop with learning rate 0.0001.

Epoch: 30

