

1. training: 0.988

Validation: 0.729

test accuracy: 0.72079

I do not believe there is any overfitting since the test accuracy is less than the training, also I think that will about that number of epochs we have enough data to be able to make a good case.

2 Yes it did, it labeled it as dog

3. I do not believe there is any overfitting since the values of the test accuracy is lower than the training. With there being no overfitting I think that those number of epochs is enough to be able to describe the model.

training: .9232

Validation: .6478

test accuracy: .6388

4. training: .6933

Validation: .61599

test accuracy: .61119

I do believe there is overfitting happening since the values of training and accuracy are not far from one another. Therefore more epochs are needed to interrupt this data.

For model 3 when it came to the time of training the epochs it took double the amount of model 2 and it also took longer than model 1.

5. Test accuracy: .77

Training: .78

Final validation: .77

Separable conv. is better for few parameters

Dropout helps reduce overfitting

Batchnorm allow data to train faster

Pooling helps simplify the data that's being represented.

Model 1						
	Layers	Input	Filters	Kernel	Stride	Parameters
4	Conv2D	32x32x3	32	3	2	896
5	BatchNorm	16x16x32	N/A	N/A	N/A	128
6	Conv2D	16x16x32	64	3	2	18496
7	BatchNorm	8x8x64	N/A	N/A	N/A	256
8	Conv2D	8x8x64	128	3	2	73856
9	BatchNorm	4x4x128	N/A	N/A	N/A	512
10	Conv2D	4x4x128	128	3	1	147584
11	BatchNorm	4x4x128	N/A	N/A	N/A	512
12	Conv2D	4x4x128	128	3	1	147584
13	BatchNorm	4x4x128	N/A	N/A	N/A	512
14	Conv2D	4x4x128	128	3	1	147584
15	BatchNorm	4x4x128	N/A	N/A	N/A	512
16	Conv2D	4x4x128	128	3	1	147584
17	BatchNorm	4x4x128	N/A	N/A	N/A	512
18	MaxPooling	4x4x128	N/A	4	4	0
19	Flatten	1x1x128	N/A	N/A	N/A	0
20	Dense	128	N/A	N/A	N/A	16512
21	BatchNorm	128	N/A	N/A	N/A	512
22	Dense	128	N/A	N/A	N/A	1280
23	Total					703552
24						83903744

Model 2

Layers	Input	Filters	Kernel	Stride	Parameters	MACs	Output Size
Conv2D	32x32x3	32	3	2	896	25165824	16x16x32
BatchNorm	16x16x32	N/A	N/A	N/A	128	0	16x16x32
Separable Conv	16x16x32	64	3	2	2400	524288	8x8x64
BatchNorm	8x8x64	N/A	N/A	N/A	256	0	8x8x64
Separable Conv	8x8x64	128	3	2	8896	524288	4x4x128
BatchNorm	4x4x128	N/A	N/A	N/A	512	0	4x4x128
Separable Conv	4x4x128	128	3	1	17664	262144	4x4x128
BatchNorm	4x4x128	N/A	N/A	N/A	512	0	4x4x128
Separable Conv	4x4x128	128	3	1	17664	262144	4x4x128
BatchNorm	4x4x128	N/A	N/A	N/A	512	0	4x4x128
Separable Conv	4x4x128	128	3	1	17664	262144	4x4x128
BatchNorm	4x4x128	N/A	N/A	N/A	512	0	4x4x128
MaxPooling	4x4x128	N/A	4	4	0	0	1x1x128
Flatten	1x1x128	N/A	N/A	N/A	0	0	128
Dense	128	N/A	N/A	N/A	16512	16384	128
BatchNorm	128	N/A	N/A	N/A	512	0	128
Dense	128	N/A	N/A	N/A	1280	1280	10
TOTAL					104106	27280640	