

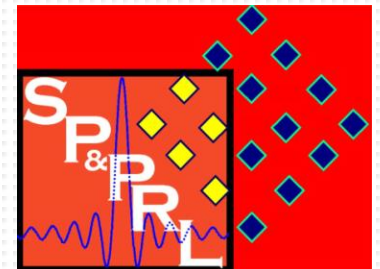
CONVOLUTION NEURAL NETWORKS

- POTHOLE DETECTION

12/7/2018 – ROWAN

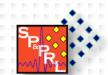
Thai Nghiem

**Signal Processing & Pattern Recognition Laboratory
Electrical & Computer Engineering
Rowan University, Glassboro, NJ**



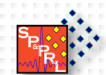
- Problem Definition
- Data Analysis
- Data Preprocessing & Data Augmentation
- First Model
- Second Model
- Misclassification and Why
- Possible ventures

Problem Definition




- **Requirement**– Differentiate between a road that is in a good condition versus a road that has pothole(s) using a 1280x720 RGB image.
- **Constraints**
 - Relatively small dataset (~6000 images)
 - Lack of Memory in GPU
 - Runtime
- **Standards**
 - 70% and 30% ratio
 - K-fold validation & Early stopping
 - Keras wrapper for TensorFlow
 - Images taken from same camera and orientation

Data Analysis



- Dataset will be collected from NJDOT Clinic lead by Dr. Bouaynaya and Dr. Farzan Kazemi.
- Data description

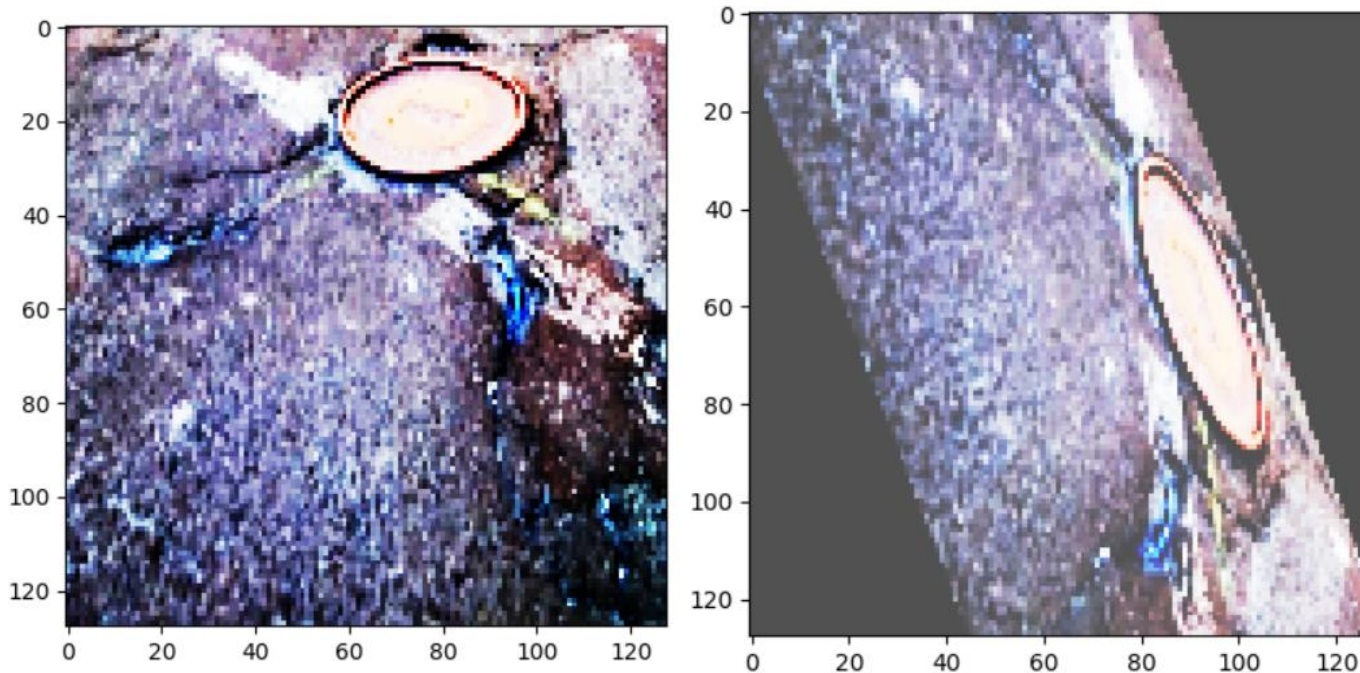
 dataset_information.txt - Notepad

File Edit Format View Help

```
dataset1_01: Only good road, no other attributes
dataset1_02: All badroad (cracks and potholes)
dataset1_03: Only cracks (good road and bad road)
dataset1_04: Only potholes (good road and bad road)
```

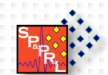
- Good Road (dataset1_01) : 4000 images (128x720 RGB)
- Pothole Road (dataset1_04) : 2000 images (128x720 RGB)

- Equals amount of good road and pothole road images
 - 5,000 images each class for the first model
 - 10,000 images each class for the second model
- Random rotation and blur images



First Model

- The naïve one

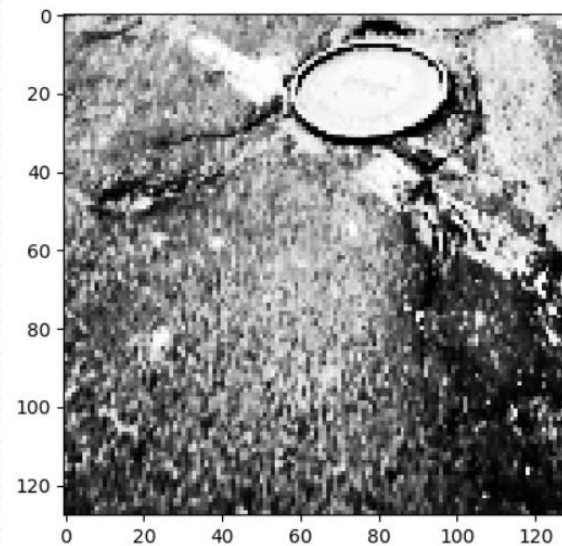


Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 31, 31, 16)	1040
activation_1 (Activation)	(None, 31, 31, 16)	0
conv2d_2 (Conv2D)	(None, 16, 16, 32)	12832
activation_2 (Activation)	(None, 16, 16, 32)	0
conv2d_3 (Conv2D)	(None, 8, 8, 64)	51264
global_average_pooling2d_1 ((None, 64)		0
dropout_1 (Dropout)	(None, 64)	0
activation_3 (Activation)	(None, 64)	0
dense_1 (Dense)	(None, 1024)	66560
dropout_2 (Dropout)	(None, 1024)	0
activation_4 (Activation)	(None, 1024)	0
dense_2 (Dense)	(None, 512)	524800
dropout_3 (Dropout)	(None, 512)	0
activation_5 (Activation)	(None, 512)	0
dense_3 (Dense)	(None, 2)	1026
activation_6 (Activation)	(None, 2)	0

=====
 Total params: 657,522
 Trainable params: 657,522
 Non-trainable params: 0

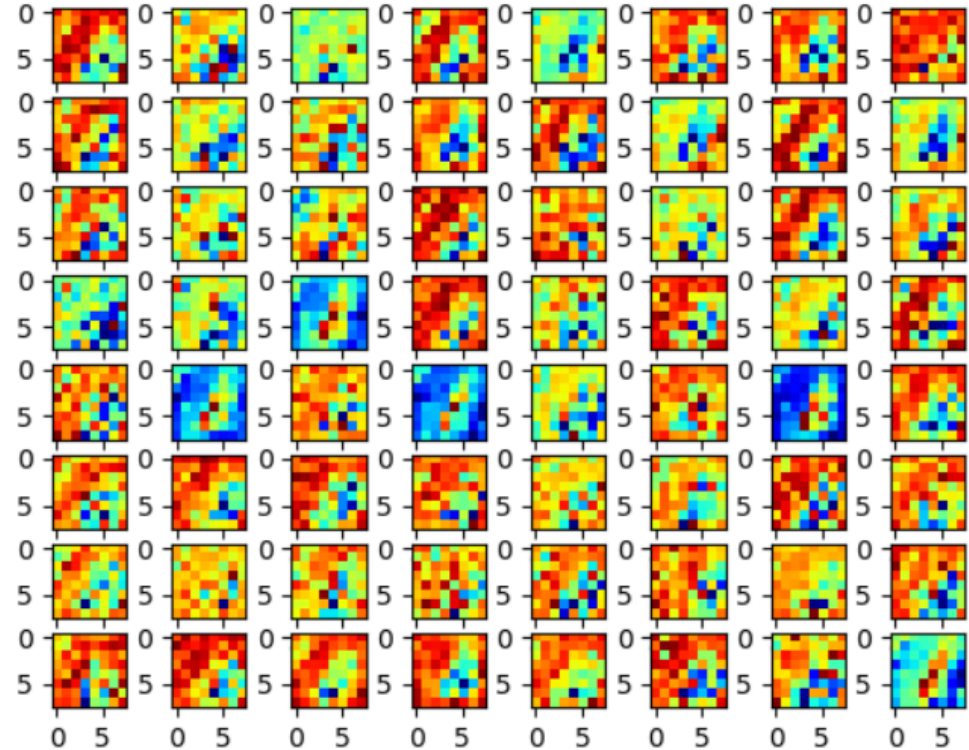
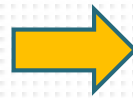
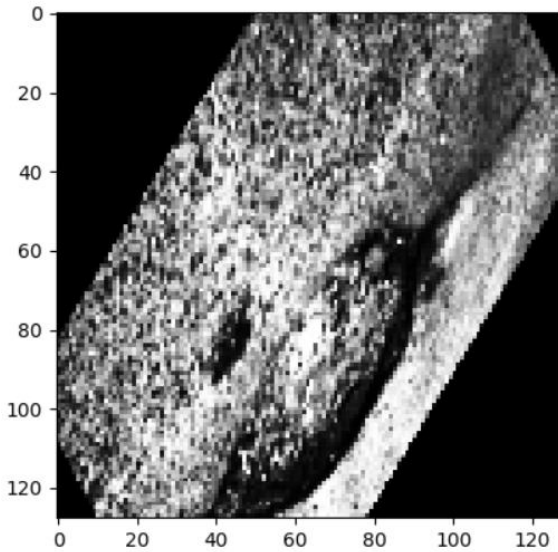
First Model

- Input layer: 128x128 grayscale pictures



- 3 Convolutions layers, each with a subsample
- 2 Hidden FC layers
- Dropout and 'relu'

- Took 16 minutes to train
- Activation layers:

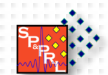


- Result:

```
3000/3000 [=====] - 3s 963us/step
loss: 0.3839658672809601
acc: 0.8446666666666667
```

Second Model

- The new and improved one

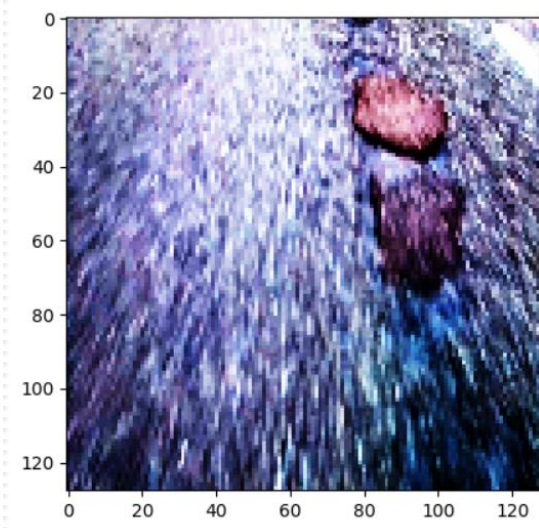


Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 31, 31, 16)	3088
activation_1 (Activation)	(None, 31, 31, 16)	0
conv2d_2 (Conv2D)	(None, 31, 31, 32)	12832
activation_2 (Activation)	(None, 31, 31, 32)	0
global_average_pooling2d_1 ((None, 32)	0
dense_1 (Dense)	(None, 512)	16896
dropout_1 (Dropout)	(None, 512)	0
activation_3 (Activation)	(None, 512)	0
dense_2 (Dense)	(None, 2)	1026
activation_4 (Activation)	(None, 2)	0

Total params: 33,842
 Trainable params: 33,842
 Non-trainable params: 0

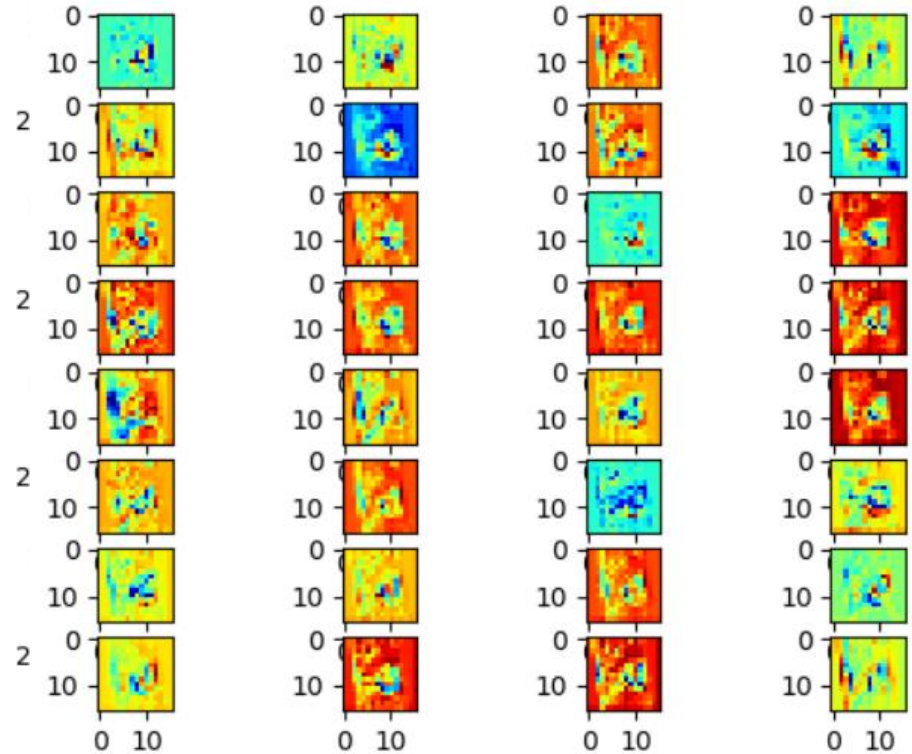
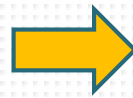
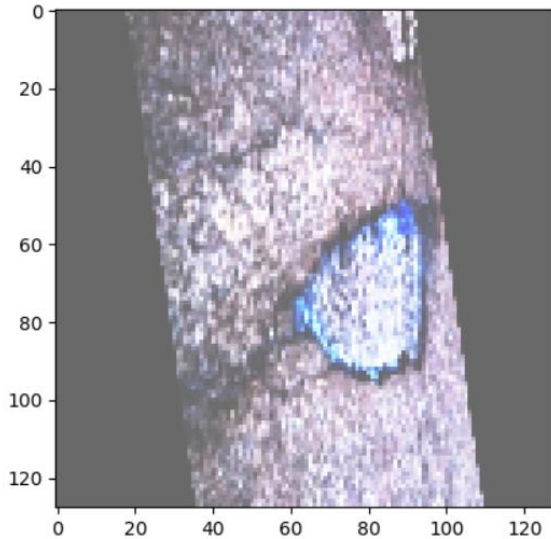
Second Model

- Input layer: 10,000
128x128 RGB pictures



- 2 Convolutions layers, each with a subsample
- 1 Hidden FC layers
- Dropout and 'relu'
- Early stopping

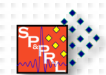
- Took 40 minutes to train
- Activation layers:



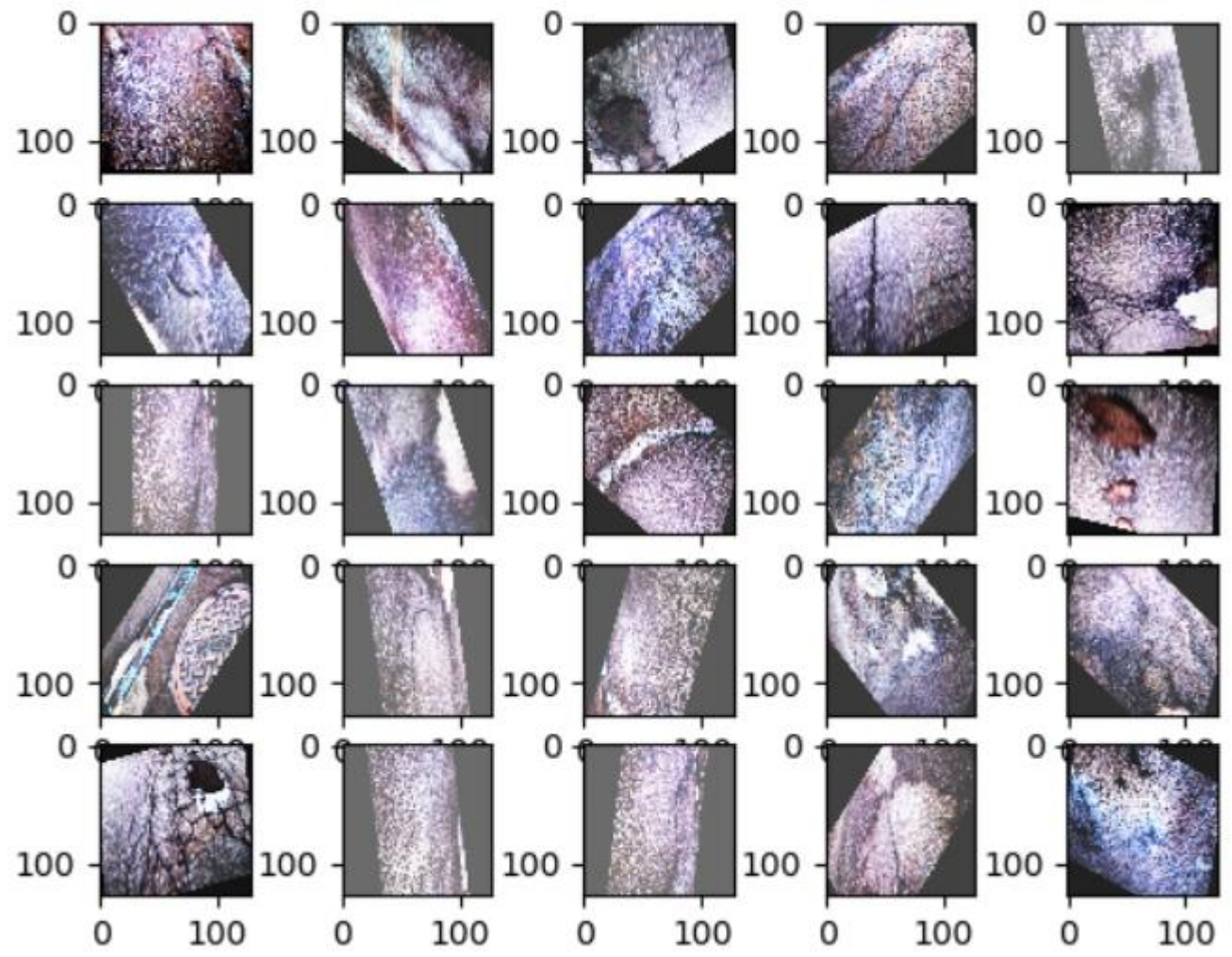
- Result:

```
6000/6000 [=====] - 13s 2ms/step
loss: 0.2043301287094752
acc: 0.926
```

Misclassification & Why



'Good Road' misclassified as 'Bad Road'



- Build Multi-class Classifier in a single image
 - Have images of 1 single class. For example:



Good Road

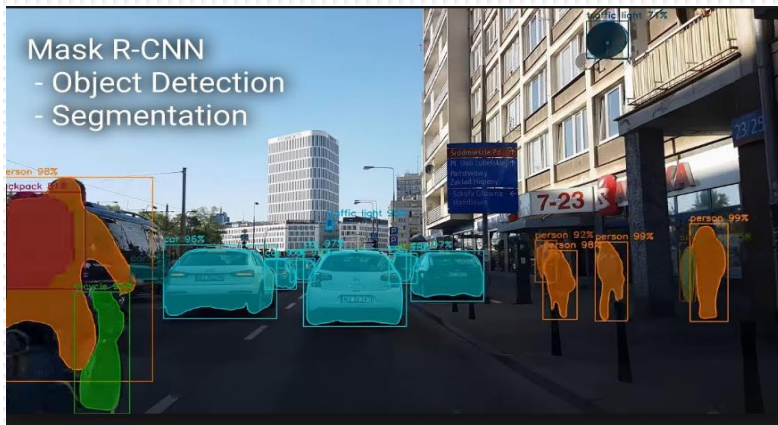


Cracks



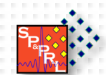
Pothole

- Train algorithms like Mask R-CNN

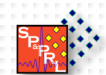


- More data
- Correctly Label Classes

Thanks for listening!

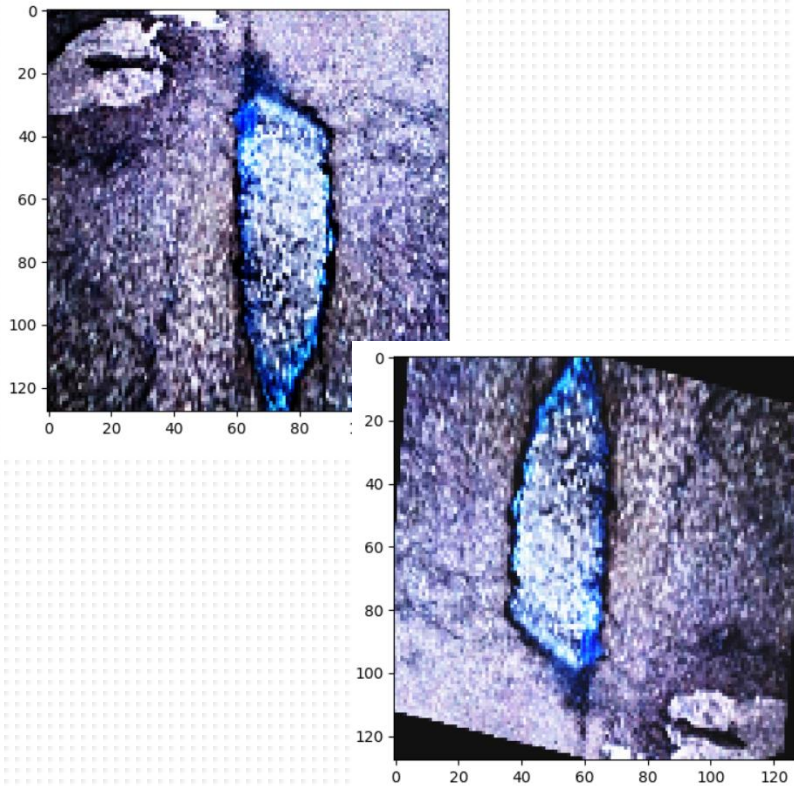


Bonus Slides

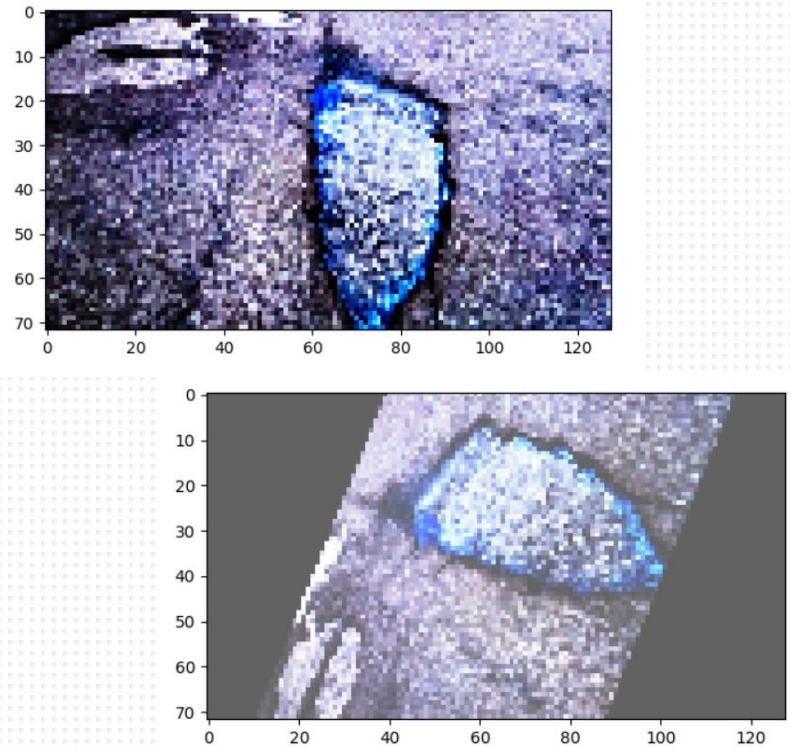


Images Sizes

- 128 x 128



- 128 x 72



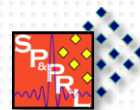
- 128x5000

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 31, 31, 16)	1040
activation_1 (Activation)	(None, 31, 31, 16)	0
conv2d_2 (Conv2D)	(None, 16, 16, 32)	12832
activation_2 (Activation)	(None, 16, 16, 32)	0
conv2d_3 (Conv2D)	(None, 8, 8, 64)	51264
global_average_pooling2d_1 ((None, 64)	0
dropout_1 (Dropout)	(None, 64)	0
activation_3 (Activation)	(None, 64)	0
dense_1 (Dense)	(None, 512)	33280
dropout_2 (Dropout)	(None, 512)	0
activation_4 (Activation)	(None, 512)	0
dense_2 (Dense)	(None, 256)	131328
dropout_3 (Dropout)	(None, 256)	0
activation_5 (Activation)	(None, 256)	0
dense_3 (Dense)	(None, 2)	514
activation_6 (Activation)	(None, 2)	0

Total params: 230,258
 Trainable params: 230,258
 Non-trainable params: 0

```

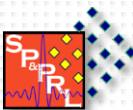
3000/3000 [=====] - 3s 880us/step
loss: 0.3839658672809601
acc: 0.8446666666666667
  
```



- 128x72x3x10000

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 17, 31, 16)	3088
activation_1 (Activation)	(None, 17, 31, 16)	0
conv2d_2 (Conv2D)	(None, 9, 16, 32)	12832
activation_2 (Activation)	(None, 9, 16, 32)	0
global_average_pooling2d_1 (GlobalAveragePooling2D)	(None, 32)	0
dense_1 (Dense)	(None, 1024)	33792
dropout_1 (Dropout)	(None, 1024)	0
activation_3 (Activation)	(None, 1024)	0
dense_2 (Dense)	(None, 256)	262400
dropout_2 (Dropout)	(None, 256)	0
activation_4 (Activation)	(None, 256)	0
dense_3 (Dense)	(None, 2)	514
activation_5 (Activation)	(None, 2)	0
Total params: 312,626		
Trainable params: 312,626		
Non-trainable params: 0		

```
6000/6000 [=====] -
loss: 0.271551072537899
acc: 0.898
```



- 3 classes (Good road, pothole road, and cracks)
- 128x128x3x10000

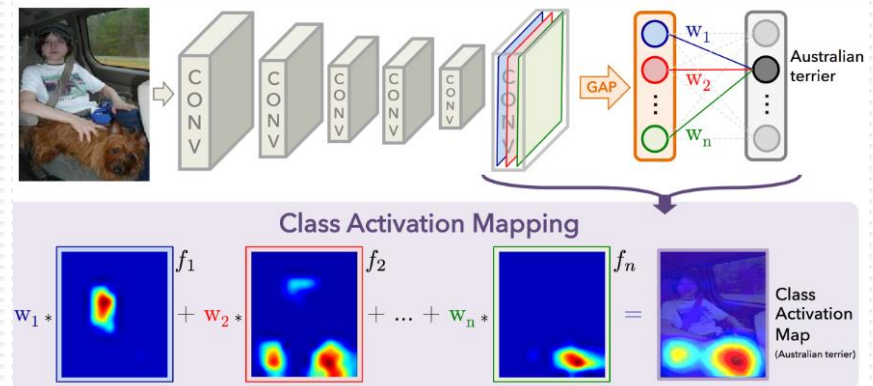
Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 31, 31, 16)	3088
activation_1 (Activation)	(None, 31, 31, 16)	0
conv2d_2 (Conv2D)	(None, 31, 31, 32)	12832
activation_2 (Activation)	(None, 31, 31, 32)	0
global_average_pooling2d_1 ((None, 32)	0
dense_1 (Dense)	(None, 512)	16896
dropout_1 (Dropout)	(None, 512)	0
activation_3 (Activation)	(None, 512)	0
dense_2 (Dense)	(None, 2)	1026
activation_4 (Activation)	(None, 2)	0
=====		
Total params: 33,842		
Trainable params: 33,842		
Non-trainable params: 0		

```
9000/9000 [=====] -
loss: 0.5689032475153605
acc: 0.7632222222222222
```



This image is classified as 'Cracks'

- Only 1 subsample:
 - Reduce total number of parameters \rightarrow Reduce computational complexity \rightarrow Faster
 - While retain information
- 2 Convolution Layer:
 - Avoid overfitting
 - Get more spatial information
- Global Average Pooling 2D
 - Minimize overfitting by reducing the total number of parameters in the model
 - Reduce the spatial dimensions of a three-dimensional tensor \rightarrow Flatten



<https://alexisbcook.github.io/2017/global-average-pooling-layers-for-object-localization/>