Assignment For Skylark drones

Problem:- To detect right angled plastic strips.

Data observations:-

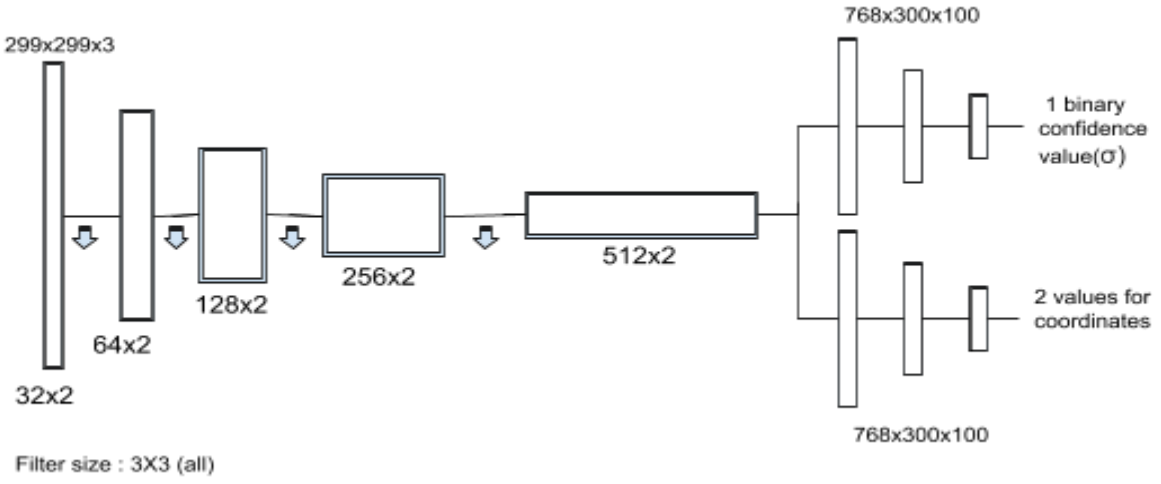
- Images are too less to train any model , and there's generally only one or two strips in images so random crops of training images will create class imbalance in training data.
- For using deep learning models, huge synthetic generation of data is required.
- Strips are not always perfectly right angled and of perfect shape, they are very small and light in some images.
- Images need to be smaller to be used for training. Resizing them leads to loss of quality of strips in the images, hence not recommended.

Possible Solutions:-

- Perform data augmentation by using cropped strips and placing them at random positions in images. Requires huge data.
- Use binarization or white blob detection to help in detecting the shape. Fixed parameter of sensitivity to white objects does not performs good for all the scenes. Only way to solve this is to make the sensitivity a learnable parameter, which is more or less equivalent to convolution layers.

Solution Steps:-

1. I used a network of input size 299x299x3 as the original size of images was very large and strips were relatively too small.
2. Data augmentation is done by splitting up training images into size of 299x299 and randomly placing strips in 40% of those images.
3. This generated approximately 30k images.
4. Built a convolutional neural network using keras as shown below.



5. By observing the training data, it's clear that crops of size 299x299 wouldn't have more than one strip pair.
6. So network has two output heads, one for classification, whether strips appear in the cropped image or not. And the other is for predicting the location of tip of the strips
7. Cross entropy loss is used with sigmoid activation function for classification head, while mean square loss is used with linear activation for regression head.

Results and observations

- Simultaneous training of both heads leads to unstable learning. Separate training is done.
- Network does not generalises well from augmented data set. Better data is required.
- Dropout , L2 regularisation and batch normalisation introduced to help network generalise well
- Images containing white objects get high confidence even if they do not have right angled strips
- Random shaped white objects were introduced in images of negative class to help network better focus on shape of strips rather than focussing on white objects
- Lead to better results , but still a lot of room for improvement.
- Trains to f2 score of 97% on validation data set but performs poorly on nearly half of to images of test set

Sample Outputs (Performance on test images)-

- **True Positives:**
  (Green dot shows output of regression head)

DSC02120.JPG

Confidence:0.93

M1_F1.3_0405.JPG

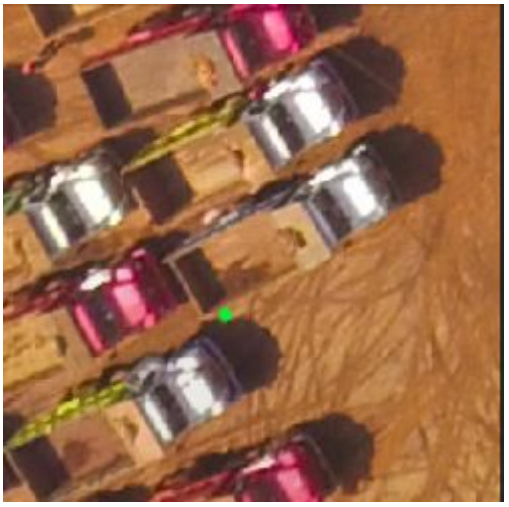
Confidence:0.99

DJI_0086.JPG
(Confidence :0.99)

DSC01590.JPG
(Confidence :0.98)
(Bad regression head output)

- **False Positives**

DJI_0617.JPG
(confidence:0.05 but still maximum in the whole image)



DJI_0422.JPG
confidence: 9.98

Final Results- Model performs good enough for first iteration but does not generalises well enough to be used as production model.

Possible Improvements-

- Better generation of augmented data using backgrounds from other datasets and using random strips cropped from different training images.
- While predicting the output, usage of stride will be better than just splitting the image into segments of 299x299
- Need to use images with more white objects to make network focus on shape rather than color.

Difficulties:-

- If smaller models are used, it's accuracy does not increases after a point in training, while loss continues to decrease. Which probably means that model is trying to remember the data.
- I have 6gb of gpu which makes it difficult to train such large models
- Pretrained models like inception,Xception were tried but that did not improve results

How to run:

- Use output.ipynb notebook file.
- Enter enter any image path in second last cell.
- Run all the cells