**Object detection and faster R-CNN**

Object detection is being able to pick out the interested features in an image and then assigning them the correct label. For example, in an image containing a person, the face can be the interested feature. The detection algorithm should be able to accurately recognise that there is a face in that image then assign it the correct label, which in this case the label will be a face. An image could have multiple features you want the algorithm to correctly detect and label.

For this project a pretrained model of Faster R-CNN will be for object detections in an image. Faster R-CNN is a regional convolutional neural network (CNN). A CNN is a type of artificial neural network that is mostly used for image detection. Faster R-CNN consists of 3 stages. Region proposal network, a classification layer and then a regression layer.

In the first stage locations of possible objects are detected, then the features of these objects are obtained. The classification layer predicts what the object is and then in the regression layer the coordinates of the objects are acquired. At the region proposal stage we input the image into an convolutional layer. The convolutional layer that is used in this network is Resnet-50. Resnet-50 is a 50 layer deep convolutional neural network**.** Convolution is done by sliding filter across the input image to create a feature map. Once that is done pooling occurs where the algorithm tries to decrease the number of features by eliminating pixels with low values.

The output from pooling will be the number of features in an image. Then this is passed through classification and regression layers. In the classification layer all the features goes through softmax to get the probability of each features belonging to each actual class. The regression tells us the size of each features via their coordinates. However this is not needed for this project.

The problem

Nowadays it is normal for individuals to have thousands of images stored on their devices, since most of the images contain arbitrary filename, it is time consuming for users to search for a specific image. In order to access a specific image the user will have to go through each picture one by one. The problem is the more pictures you have the longer it will take for you to find a specific image. By default there is not device that has an application that allows user to sort through their images by the content of the image. Another issue will be being able to detect the objects in the image.

How to solve it

To solve this problem, machine learning can be used to detect objects in an image. The user will use an application where they can upload a folder with multiple images. Then the user will select what type of content they are looking for the image to contain, they can choose a threshold amount which changes the accuracy of the selected categories. The user should be able to select multiple category. Once the folder, category and the threshold has been set, the neural network model will be used to pick out the required images. Predicting objecting in an image is possible thanks to deep learning neural network.

What’s already solved

So far into this project, Pytorch has been selected as the machine learning library. Pytorch has a lot of pretrained models for object detection. The model I will be using for this project is faster-RCNN. Faster-RCNN is built on resnet-50. The model that is being used is a pretrained model, this is because the network has been trained on millions of images from the ImageNet database.

Results

Using faster-RCNN, I have been able to get the predictions of objects in an image and also get the prediction based on the threshold. The results are looking good so far. When an image is inserted the model is able to detect and predict what the objects are in an image. Also there is an percentage of accuracy for each object the model predicts.