

CS-725

Course-Project

Image Segmentation Using MRCNN



Team members

Suman Saurav Panda - 22M0783

Harikrishna Shenoy - 22M0759

Ashish Kumar Pandey -22M1501

Prateek Gothwal -22M0813

Akash Sahoo -22M0797

Dr. Preethi Jyothi (instructor)

Introduction

The task of image segmentation enables machines to separate different objects in the image into individual segments. It transforms an image into something that is easier to analyze and understand. The task of image segmentation usually involves classifying, detecting, and labeling objects.

In this project we are exploring one state of the art method e.g.- **Masked R-CNN** prominently used for instance segmentation tasks. For our experiments we have used this [dataset](#) (manually annotated data set for brain tumor segmentation).

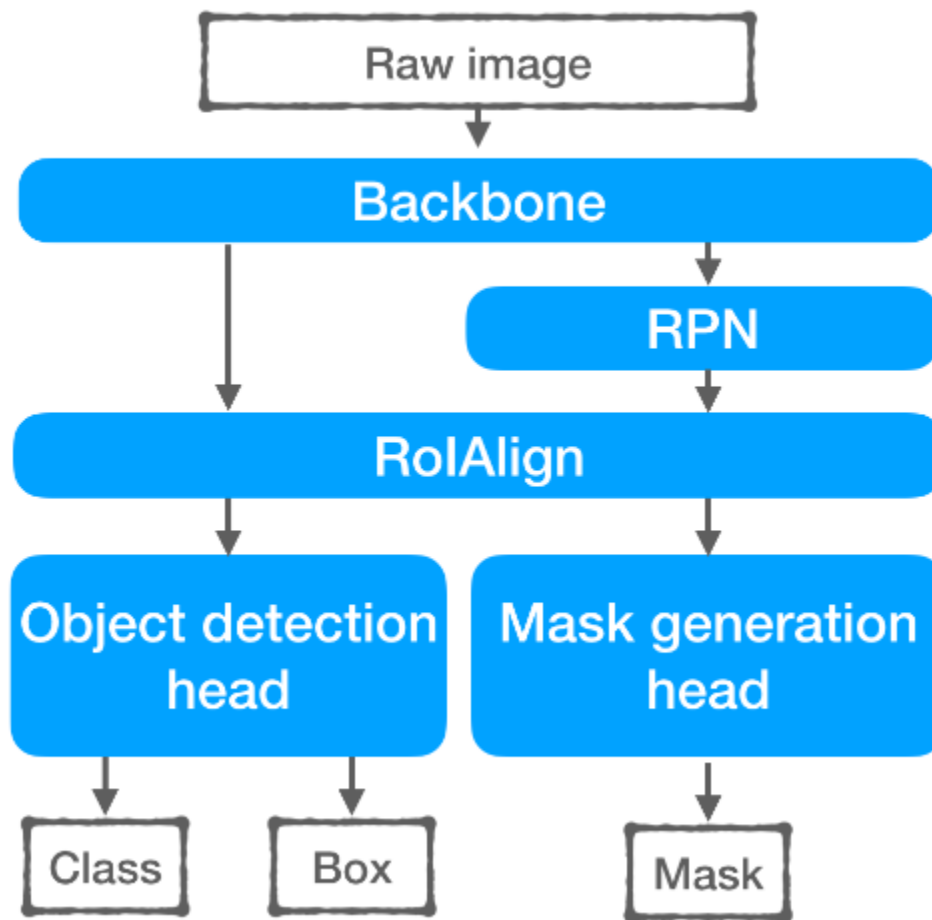
We have used this kaggle [notebook](#) as our starting point for all our experiments on various parameters associated with the algorithm.

Masked R-CNN

Mask R-CNN is an object detection model based on deep convolutional neural networks (CNN) developed by a group of Facebook AI researchers in 2017. The model can return both the bounding box and a mask for each detected object in an image.

Mask R-CNN is built using Faster R-CNN. Faster R-CNN is the precursor model of MRCNN for object detection. It has 2 outputs for each candidate object, a class label and a bounding-box offset, whereas Mask R-CNN is the addition of a third branch that outputs the object mask. The additional mask output is distinct from the class and box outputs, requiring the extraction of a much finer spatial layout of an object.

Brief working of Mask R-CNN



Source - https://www.shuffleai.blog/blog/Understanding_Mask_R-CNN_Basic_Architecture.html

1. **Backbone** - Image is run through the CNN to generate the feature maps.
2. **Region Proposal Network (RPN)** - uses a CNN to generate the multiple Region of Interest(RoI) using a lightweight binary classifier. It does this using 9 anchor boxes over the image. The classifier returns object/no-object scores. Non Max suppression is applied to Anchors with high objectness score
3. **RoI Alignment network** - outputs multiple bounding boxes rather than a single definite one and warps them into a fixed dimension
4. Warped features are then fed into fully connected layers to make classification using softmax and boundary box prediction is further refined using the regression model
5. Warped features are also fed into the Mask classifier, which consists of two CNN's to output a binary mask for each RoI. Mask Classifier allows the network to generate masks for every class without competition among class

Experiments with Mask-RCNN

- Please check the [slides](#) to see the experiment result
- Kaggle notebook for the experiments can be found [here](#)
- Github [repo](#) to make changes to source code of MRCNN for our experiment

Inference

- We compared K-Means clustering with Masked R-CNN. K-Means worked reasonably well to segment images with similar pixel intensity but performed poorly for instance segmentation tasks.
- We have used a pre-trained object detection model (trained on MS-COCO dataset) and used the concept of transfer learning to train our model on the brain-tumor dataset for our subsequent experiments
- Then we experimented with 3 parameters associated with MRCNN algorithm
 - Learning Rate
 - Backbone algorithm
 - Activation function for CNN
- For change in learning rate we didn't observe any significant change to test accuracy
- For the backbone algorithm ResNet-50 outperformed ResNet-101. One of the plausible explanations for this behavior is small train data set where model using ResNet-101 might be overfitting hence poor performance in comparison to ResNet-50
- For the activation function experiment, we compared ELU vs RELU. Our intuition was to give some negative feedback(ELU) for non-tumor examples instead of zero feedback(RELU) and check the test accuracy. But counterintuitively RELU outperformed the ELU in test accuracy.
- We have also performed model training by changing the step size of epochs. But no significant difference was observed.