**Homework 2**

**Segmentation**

Submission Date: 23/11/2022

|  |  |  |
| --- | --- | --- |
| **ID Number** | **Last Name** | **First Name** |
| 324338136 | Yorov | Alexander |
|  |  |  |

# Part 1 – Classic Vs. Deep Learning-based Semantic Segmentation

# Part 2 – Adversarial Images

# Part 3 – Jurassic Fischbach

In this part you are going to analyze a (large) pre-trained model. Pre-trained models are quite popular these days, as big companies can train really large models on large datasets (something that personal users can't do as they lack the sufficient hardware). These pre-trained models can be used to fine-tune on other/small datasets or used as components in other tasks (like using a pre-trained classifier for object detection).

All pre-trained models expect input images normalized in the same way, i.e. mini-batches of 3-channel RGB images of shape (3 x H x W), where H and W are expected to be at least 224. The images have to be loaded in to a range of and then normalized using and .

## 

Load a pre-trained VGG16 with PyTorch.

Load the images and display them.

## 

Pre-process the images to fit VGG16's architecture. What steps did you take?

Feed the images (forward pass) to the model. What are the outputs?

Find an images of a bird/cat/dog on the internet, display and feed it to network. What are the outputs?

# Part 4 – Dry Questions

## 

Explain the difference between ordinary segmentation and semantic segmentation. State an algorithm that performs standard segmentation and an algorithm that performs semantic segmentation.

**In image-segmentation tasks, we identify each pixel’s class.**

A group of pixels belonging to the same class constitutes a segment. Usually, we aim to create segments by isolating objects in an image. By doing that, we change the image’s representation into a new one. Instead of actual pixel values, pixels in a segmented image can be thought of as containing class labels.

Semantic segmentation:

**In semantic segmentation, all the objects that belong to the same class share the label.**

So, if we’re working with autonomous vehicle applications, all pedestrians will receive the same label. The same goes for cars. For instance:



An algorithm that performs a semantic segmentation – **DeepLab**:

DeepLab is a state-of-the-art semantic segmentation model designed and open-sourced by Google. The dense prediction is achieved by simply up-sampling the output of the last convolution layer and computing pixel-wise loss (i.e. using a pixel-wise-softmax). The Deeplab applies **atrous convolution** for up-sample.

Ordinary segmentation:

In ordinary segmentation (sometimes called instance segmentation), each detected object receives its unique label, it classifies pixels into categories based on “instances” rather than classes. An instance segmentation algorithm has no idea of the class a classified region belongs to but can segregate overlapping or very similar object regions on the basis of their boundaries. **We usually implement this type of segmentation when the number of objects or their independence is relevant.** For instance, we may want to count people at a concert. To do so, we need to isolate and differentiate each visitor.

Returning to our example with an autonomous vehicle, each pedestrian and car will receive unique labels (which we represent using different colors):



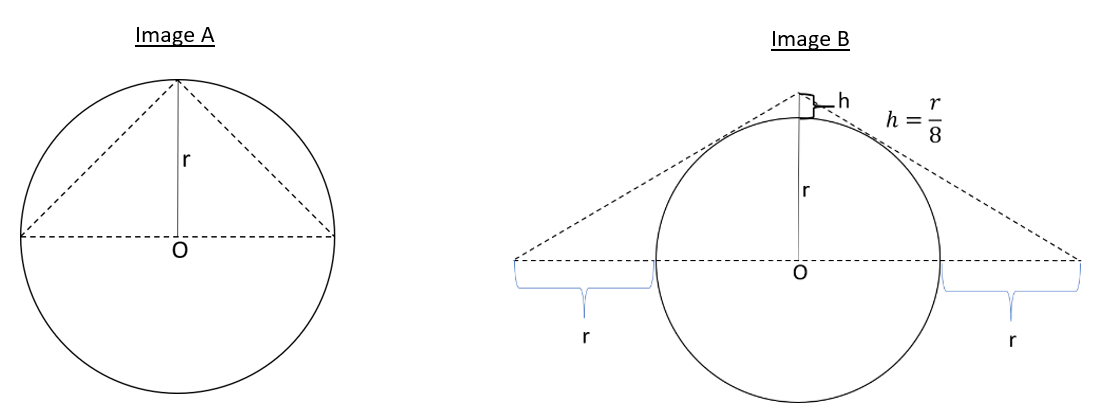
An algorithm that performs a ordinary segmentation – **GrabCut**:

GrabCut is an interactive foreground extraction that segments the desired object by providing a rectangle. You draw a rectangle around the foreground area (make sure your foreground are in side the rectangle entirely!). Next, the algorithm tries to segment the object by applying certain iteration.

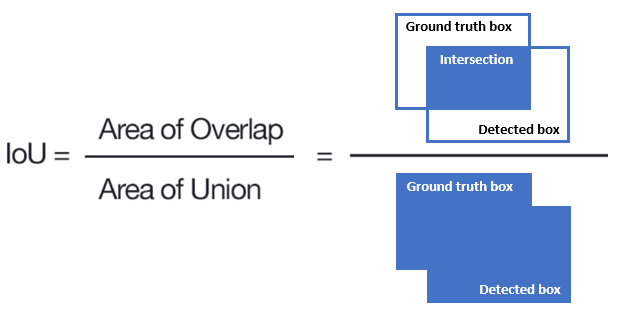
ChatGPT answer to in the Appendices.

## 

Steve designed a segmentation algorithm, where the origin of the algorithm is always triangles. When he tried to segment a circle, he got the results below. Calculate the Intersection Over Union (IoU) in each case (it is enough to find an expression with the given parameters). 'o' indicates the center of the circle. The prediction of the algorithm is marked with a dashed line.

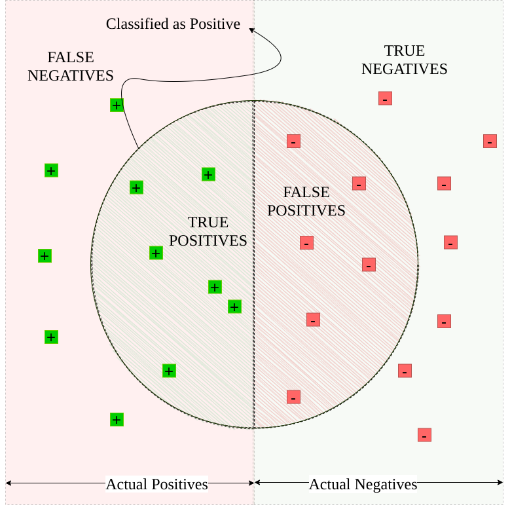


The IoU measures the accuracy of our detections. Given a ground-truth bounding box and a detected bounding box, **we compute the IoU as the ratio of the overlap and union areas:**

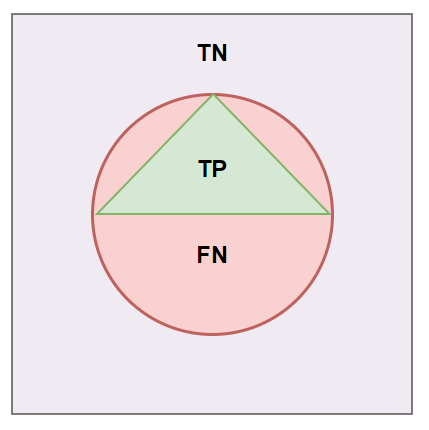


It can be given also as:

Where:



**Image A:**



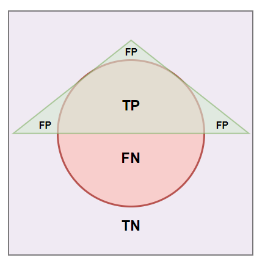
The area of overlap (or TP):

False positive area (FP) is :

False negative area (FN):

IoU:

**Image B:**



The area of overlap (or TP):

False positive area (FP) is :

False negative area (FN):

IoU:

# Appendices

[ChatGPT](https://chat.openai.com/chat)'s answer to Part 4 – Question 1:

Ordinary segmentation is the process of dividing an image into multiple segments or regions, typically with the goal of simplifying the image or highlighting certain objects or features in the image. Standard segmentation algorithms typically use pixel intensity, color, or texture information to group pixels into segments. Examples of algorithms that perform standard segmentation include K-means clustering, mean-shift clustering, and normalized cuts.

Semantic segmentation, on the other hand, involves assigning a class label or category to each pixel in an image, such that each pixel is labeled according to the object it belongs to. This allows for a more detailed and meaningful understanding of the image content. For example, in an image of a street scene, semantic segmentation would label each pixel as belonging to a specific object class such as sky, road, building, or vehicle. Algorithms that perform semantic segmentation typically use convolutional neural networks (CNNs) trained on large annotated datasets. Examples of algorithms that perform semantic segmentation include Fully Convolutional Networks (FCNs) and DeepLab.