Course: Computer Vision 046746
Names: Hadas Ben Atiya, Mor Ventura

EE 046746 - Technion - Computer Vision

Homework 3 – Segmentation

Part 1

1. Display of the images:

2. Methods of segmentation:
   - GrabCut – classic method of segmentation:
     - Method of segmentation based on GMM (gaussian mixture model, for learning and creating the colors pixels distribution) and min-cut (brings to minimum cost of connections between the nodes) which discriminate between the foreground and background, using an initial rectangle set by the user.
     - Advantages:
       - Energy function can be efficiently minimized by mutually applying graph cut algorithm in polynomial time. Hence, it can be used in many real-time applications.

- Easy to use.

○ Disadvantages:

- Depends on initialization (user's accuracy)

○ Results (binary mask multiplied by the image):




- DeepLabv3_Resnet101 – deep learning method for segmentation:
  Deeplabv3-ResNet101 is constructed by a Deeplabv3 model with a ResNet-101 backbone. The pre-trained model has been trained on a subset of COCO train2017, on the 20 categories that are present in the Pascal VOC dataset. It uses atrous spatial pyramid technique.

  ○ Advantages:

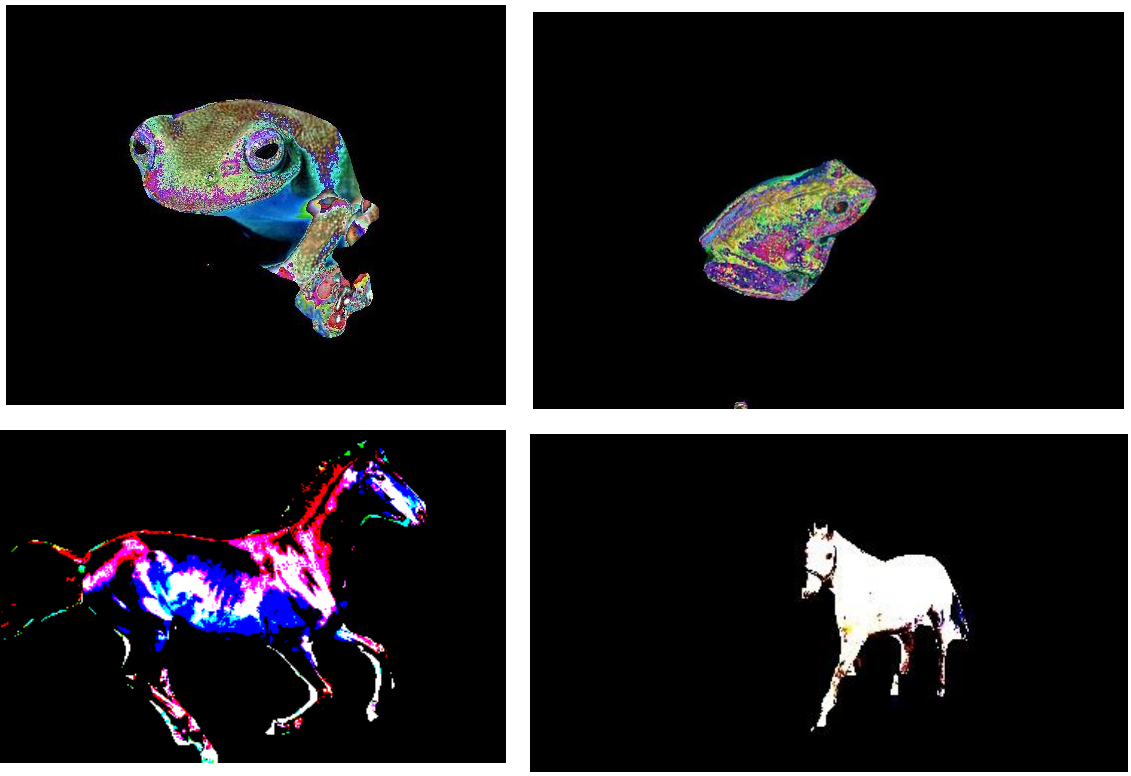
  - No need an initialization or dependency in the user
  - Scale invariant

  ○ Disadvantages:

  - Depend on the commonness of the object in the image (if the model was trained on it or something similar)

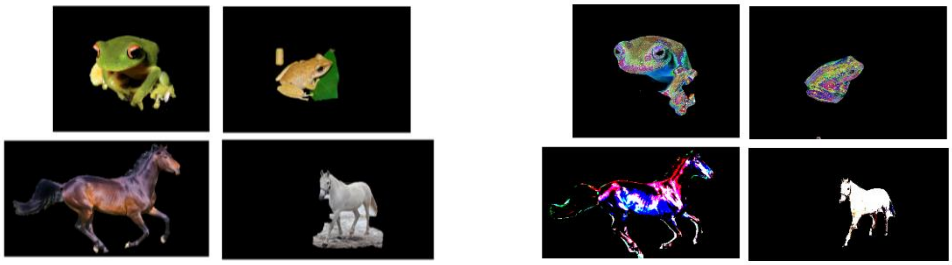
Course: Computer Vision 046746
Names: Hadas Ben Atiya, Mor Ventura

- Results (with color palette):



Comparison:

In general, we can see the results on the deep method are way better than the classic method. GrabCut (Left) VS DeepLabV3 (Right):



- Brown horse – both methods have shown good results. The horse was easy to be distinguished from the background.
- White horse – the white horse was too similar to the ground it was standing on so the classic method was fooled. The deep method showed better result.
- Green frog – here the classic method succeed to contain the object in a more fully way. The deep segmentor didn't consider the yellow leg of the frog as part of it.
- Yellow frog – since the grabCut method based on initial bounding box it had to contain the leaf and to consider it as an object. The deep method segmented it better, without the leaf.

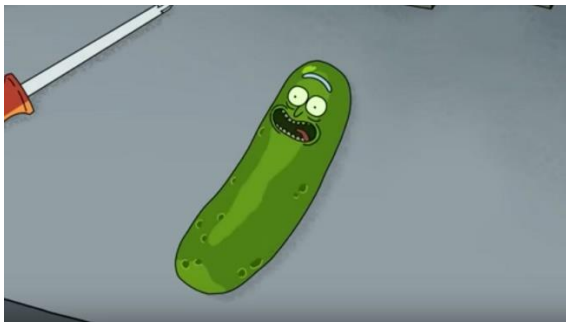
Course: Computer Vision 046746
Names: Hadas Ben Atiya, Mor Ventura

3. Three new images:

   Living being image: (Eden Alene)

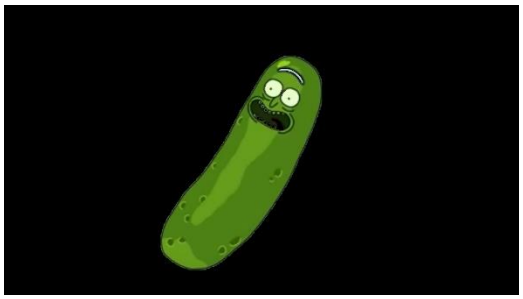
   Commonly used object: (sofa)

Not so commonly used object: (Pickle Rick (Rick & Morty))

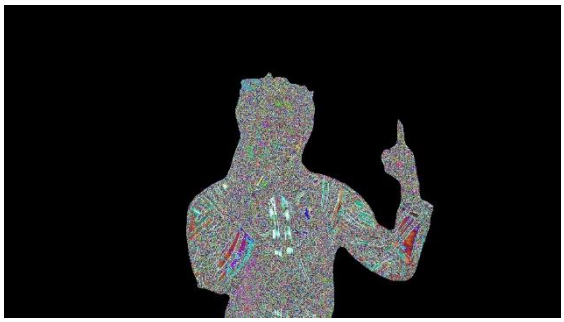
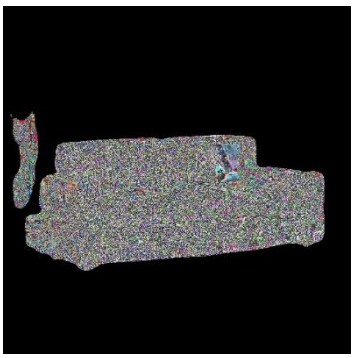


4. Applying each method on the 3 images:
   - Classic method results:





   - Deep Learning method results (with color palette):


Here, we can see similar issues as in the former images. GrabCut method leads to segmentations contain parts of the background but succeeded in the segmenting unknown object (such as pickle-Rick). DeepLabV3 didn't segmented Eden's hair (since it looks different than usual…) and had difficulty in segmenting pickleRick (since it is not common..).

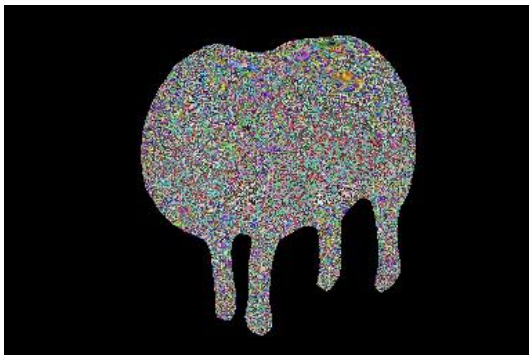
5. In order to fix or at least alleviate the problem of noised segmentation around the edges we can suggest to:
   - Split the segmentation task to ‘body’ and ‘edges’ separately (low frequencies and high frequencies) and then combine the results by some resemblance or space closeness. i.e. work with edge detection as kind of prior to the segmentation.
   - Smoothing the image with some average / gaussian filter for reducing noise and high changes.
   - Changing external and hyper-parameters to fit better the methods

6. We loaded the Wide ResNet 101 _2.

7. We chose the image of sheep on grass:

The net classification is:
["class for sheep.jpg 349: 'bighorn, bighorn sheep, cimarron, Rocky Mountain bighorn, Rocky Mountain sheep, Ovis canadensis',"] **correct.**

8. Using deep learning method and classic method, respectively (left, right):




9. Using the classic method, and padding the outcome for the background size:

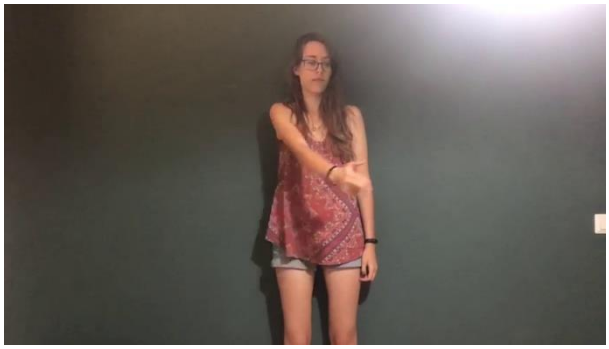
Sheep in a room :D

10. This time the classification net was wrong:
["class for sheep_in_room_.jpg 832: 'stupa, tope',"] **wrong**

Course: Computer Vision 046746
Names: Hadas Ben Atiya, Mor Ventura

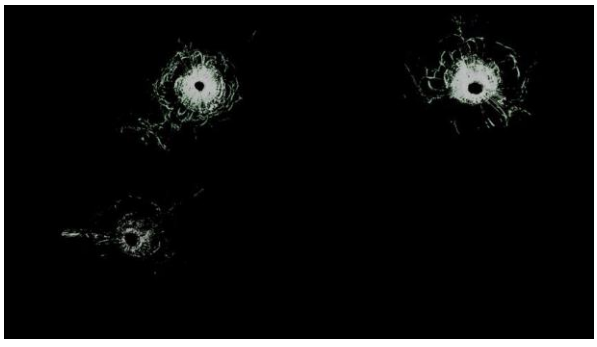
Part 2

1. 2 frames from our video:

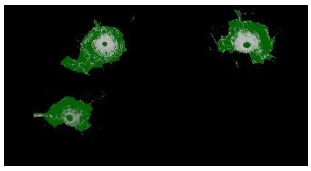



2. 2 frames segmented by deep method:


3. We chose the broken glass video (shots):

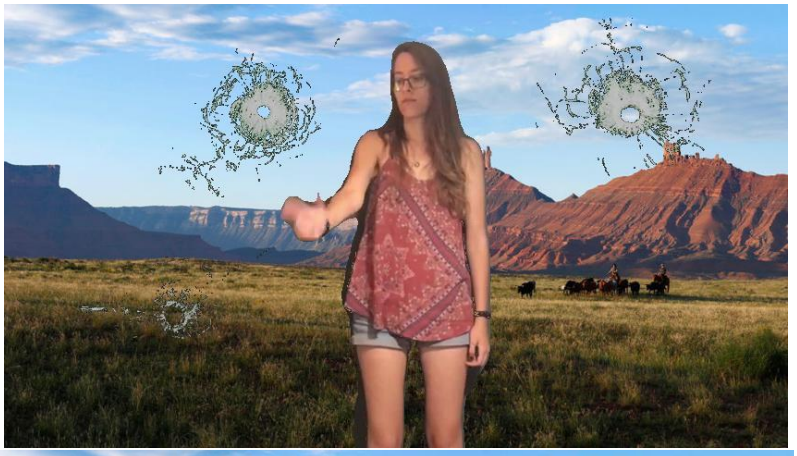


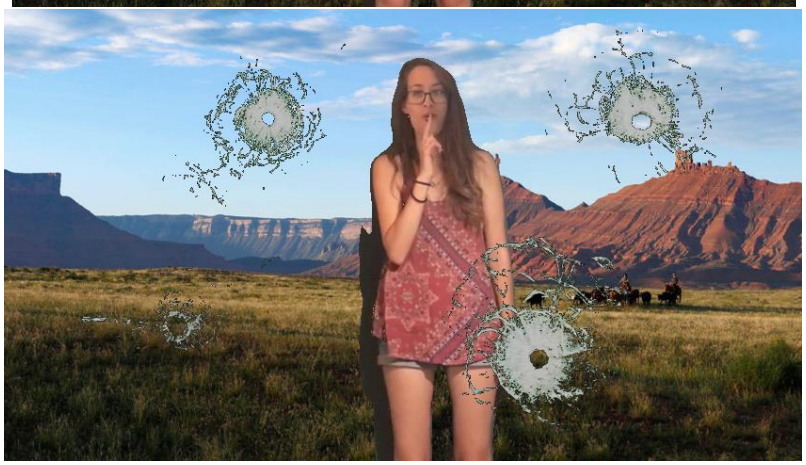

Here we used the classic method, and afterwards we applied threshold on the green channel for better refinement.



Course: Computer Vision 046746
Names: Hadas Ben Atiya, Mor Ventura

4. 2 frames from the final video:





We added part of the classic audio of `The good, the bad and the ugly` theme to the video :)

YouTube link: https://youtu.be/hHQfKbC5s8I

Thanks for the interactive HW! 😊