Simplifying Object Segmentation with PixelLib Library

Ayoola Olafenwa*
Independent AI Researcher
olafenwaayoola@gmail.com

Abstract

PixelLib [1] is a library created to allow easy implementation of object segmentation in real life applications. In this paper we discussed in detail how PixelLib makes it possible for developers to implement semantic segmentation, instance segmentation, and background editing in images and videos with great simplification.

1 Introduction

Object segmentation has a lot of amazing applications in computer vision, such as analysing medical images, road scene segmentation for vision in self driving cars, analysing satellite images and picture editing. Developers usually have a lot of challenges integrating machine learning techniques in their software solutions. One of the main challenges is combining machine learning skills with the normal programming skills. The goal of PixelLib is to ensure developers do not necessarily take stressful learning path to make use of object segmentation to solve machine learning problems in their software solutions. PixelLib makes it possible to implement object segmentation without having to go through the stress of understanding every bit of the theory and mathematics behind it. PixelLib is a flexible library that enables any one with basic knowledge of programming and little or no knowledge of machine learning to easily implement object segmentation successfully.

2 Approach to Object Segmentation

Existing object segmentation libraries like OpenCV [2] and Scikit-image [3] employ traditional techniques like watershed algorithm, region-based and edge detection to implement object segmentation. Performing image segmentation through these techniques can be challenging, because it requires a lot of manual work of finding the best threshold to produce the most efficient segmentation results. PixelLib approaches object segmentation from the perspective of deep learning. PixelLib makes use of efficient deep learning objects' segmentation architecture such as Deeplabv3+ [4] and Mask R-CNN [5] that generate pixel-wise masks directly for objects. PixelLib is based on python programming language framework. PixelLib is a very flexible library as it allows segmentation of objects in images and videos easily using few lines of python code. It

makes use of Deeplabv3+ architecture for semantic segmentation. There are two types of deeplabv3+ models used for semantic segmentation. They are Deeplabv3+ model [6] trained on pascalvoc [7] (a dataset with 20 object categories) and Deeplabv3+ model [8] trained on ade20k [9] (a dataset with 150 object categories). These models are used to perform semantic segmentation of objects in images and videos with PixelLib in five lines of code. Mask R-CNN is used for implementing instance segmentation. Mask R-CNN model [10] trained on coco [11] (a dataset with 80 object categories) is used to perform instance segmentation of objects in images and videos with PixelLib in five lines of code. PixelLib makes use of transfer learning with Mask R-CNN model to train a custom segmentation model that can be used for various applications in solving different machine learning problems, ranging from medical, picture editing to vision in self driving cars.
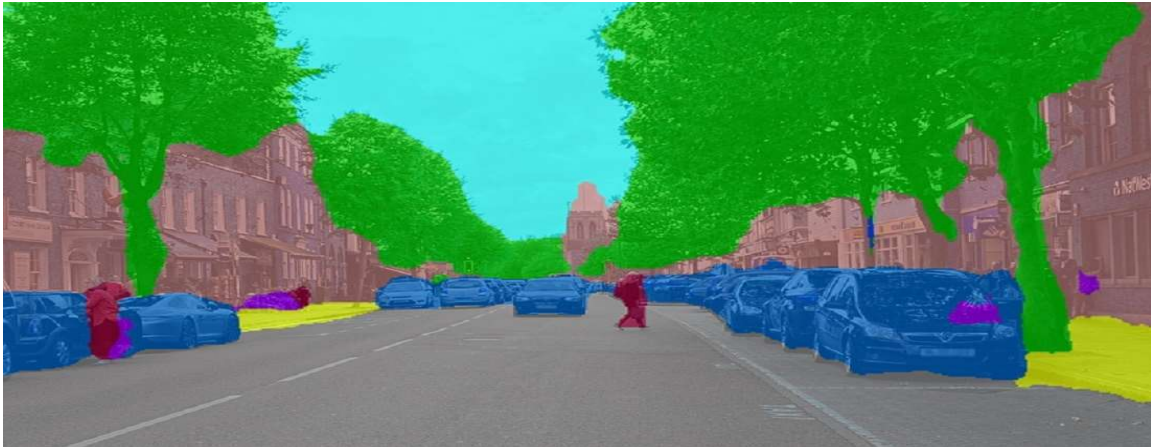
2.1 Segmentation Results from PixelLib
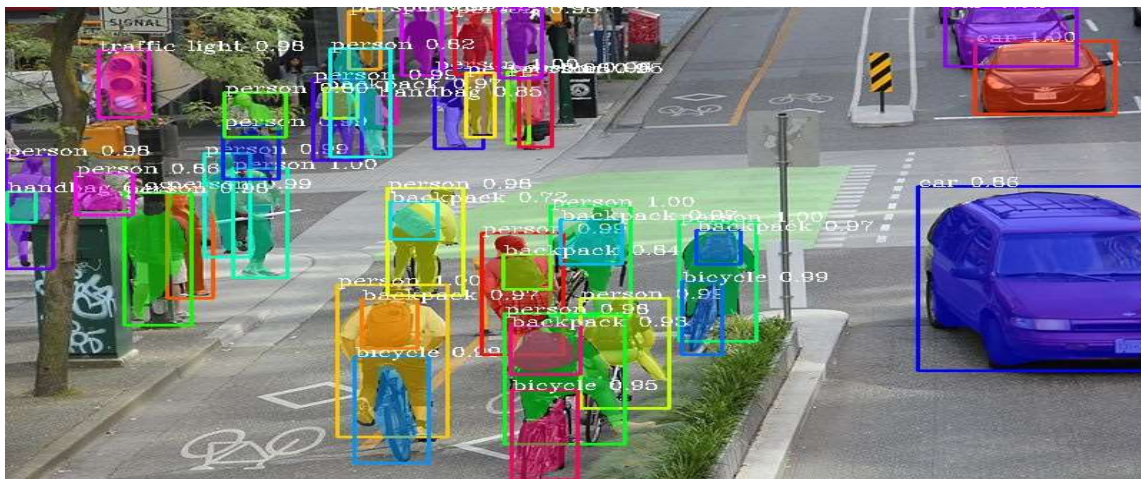
Instance segmentation with a pre-trained Mask R-CNN COCO model.

Instance segmentation with a custom model trained with PixelLib

Semantic segmentation with a pre-trained Deeplabv3+ Ade20k model.

3 Background Editing with PixelLib

PixelLib uses object segmentation to perform background editing in images and videos.

It makes it possible to alter the background of any image and video, using five lines of python code. Special effects supported for background editing in PixelLib are as follows:
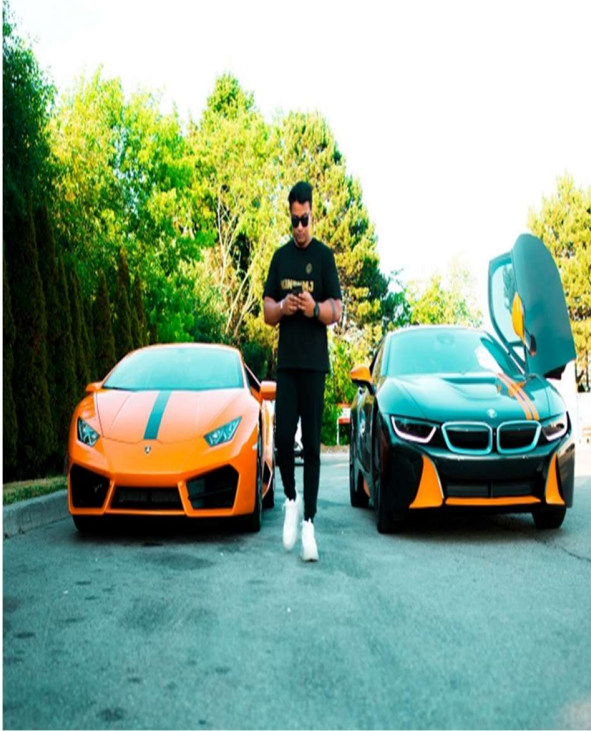
- Creating a virtual background for an image and video
- Blurring the background of an image and a video
- Assigning a distinct color to the background of an image and a video
- Grayscaling the background of an image and a video

3.1 Edited Images' Background Results

Original Image   *Blurred Background*

Original Image   *Grayscaled Background*

Original Image   *Virtual Background*   *Colored Background*

4 Projects using PixelLib

PixelLib was released as a python package on Pypi in May 2020 and as of January 2021 it has over 8000 installs on Pypi. Some of the projects using PixelLib are as follows:

1. A segmentation API using PixelLib to perform semantic and instance segmentation of images on ios. https://github.com/omarmhaimdat/segmentation_api

2. Instance segmentation of live video feed in drone's camera with PixelLib.

https://elbruno.com/2020/05/21/coding4fun-how-to-control-your-drone-with-20-lines-of-code-20-

n/?utm_source=twitter&utm_medium=social&utm_campaign=tweepsmap-Default

3. PixelLib is used to perform object segmentation in an API designed to find similar contents in images for image recommendation.

https://github.com/lukoucky/image_recommendation

Conclusion

In this paper, we introduced a new library that makes use of efficient deep learning architectures for performing image segmentation, video segmentation and background editing in images and videos, using few lines of python code. We displayed results obtained from the library and included some projects using it. Also, PixelLib added the following unique features which are not available in the existing object segmentation libraries:

- Ability to leverage the application of multiple state of the art object segmentation algorithms with great simplification.

- Ability to conveniently train a custom segmentation model.
- Ability to easily integrate background editing in images and videos.

References

[1] PixelLib github's repository.
https://github.com/ayoolaolafenwa/PixelLib

[2] Image segmentation with Opencv.

 https://bit.ly/3kTSaWj

[3] Image segmentation with Scikit-image.

 https://bit.ly/3lUpYDU

[4] Liang-Chieh Chen, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. Arxiv:1802.02611, 2018.

https://arxiv.org/abs/1802.02611

[5] Kaiming He, et al. Mask R-CNN. Arxiv:1703.06870, 2018.

https://arxiv.org/abs/1703.06870

[6] Bonlime, Keras implementation of Deeplabv3+ with pretrained weights.
https://github.com/bonlime/keras-deeplab-v3-plus

[7] Pascalvoc dataset.

http://host.robots.ox.ac.uk/pascal/VOC/

[8] Tensorflow Deeplab model zoo.

https://github.com/tensorflow/models/blob/master/research/deeplab/g3doc/model_zoo.md

[9] Semantic Understanding of Scenes through the ADE20K Dataset. Arxiv:1608.05442, 2016.

https://arxiv.org/abs/1608.05442

[10] Matterport, Mask R-CNN for object detection and instance segmentation on Keras and Tensorflow.

https://github.com/matterport/Mask_RCNN

[11] MICROSOFT COCO: COMMON OBJECTS IN CONTEXT. Arxiv:1405:0312, 2014.

https://arxiv.org/abs/1405.0312