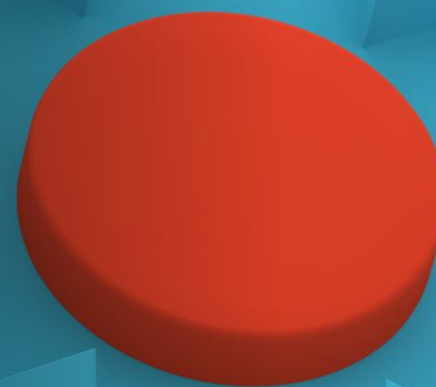# Image Segmentation [Implementation]

Group4

Yufei Li

Yurui Mao

Amey Athavia

Task Introduction

- *Carvana*[1] (online used car startup) Image Masking Challenge

- Goal: develop an algorithm that automatically removes the photo studio background

- Challenge
  - rotating photo with 16 standard images of each vehicle
  - bright reflections and cars with similar colors as the background

- Practical usage
  - Make online used car sales more effective and efficient by displaying car-related pixels only

1. https://www.kaggle.com/competitions/carvana-image-masking-challenge/overview

Carvana Dataset

- Training set
  - 100,064 images (6,254 cars) + corresponding masks
  - Each car has 16 images (different angles)
- Test set
  - 5,088 images (318 cars) + corresponding masks
  - Each car has 16 images

Applied NN Models

U-Net [ronneberger et al. 2015]

- conv 3x3, ReLU
- copy and crop
- max pool 2x2
- up-conv 2x2
- conv 1x1

SegNet [badrinarayanan et al. 2017]

64 features per layer
4 layers
7×7 convolution filter

Encoder

Decoder

$y_k = \frac{\exp(a_k)}{\sum_i \exp(a_i)}$

saved pool indices

FCN [long et al. 2015]

Pytorch, Cuda

Evaluation Results

- Metric

  - Dice coefficient (pixel-wise): $DSC = \frac{2|P \cap GT|}{|P| + |GT|}$

  - IoU (pixel-wise): $IoU = \frac{|P \cap GT|}{|P \cup GT|}$

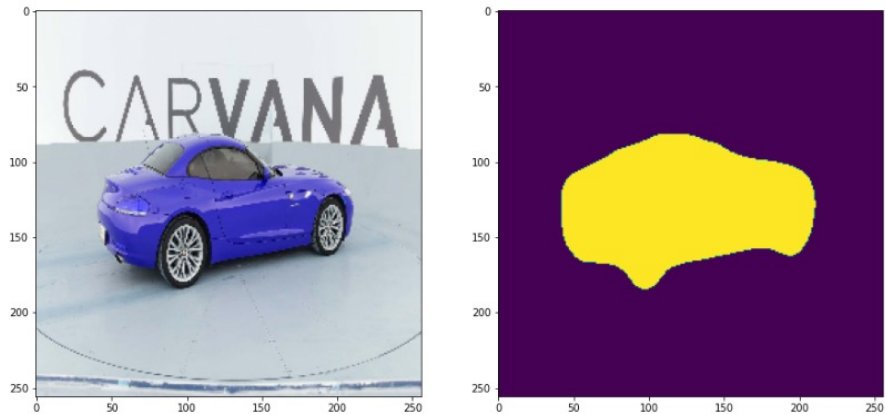 - Accuracy (pixel-wise): $Acc = \frac{|P \cap GT|}{|GT|}$
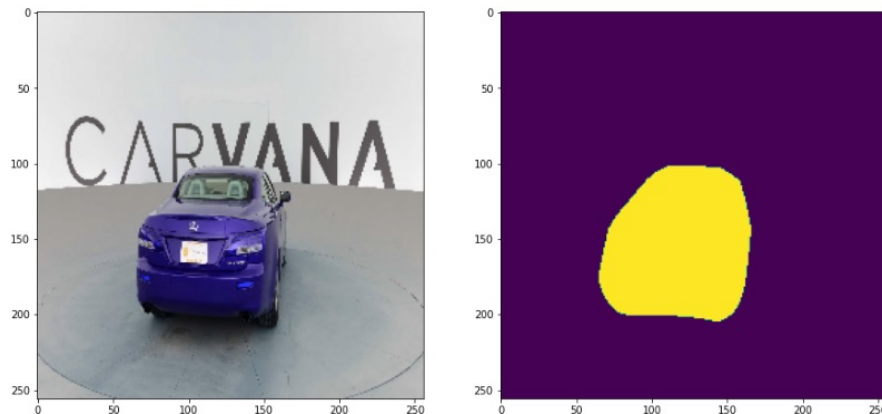
- Results

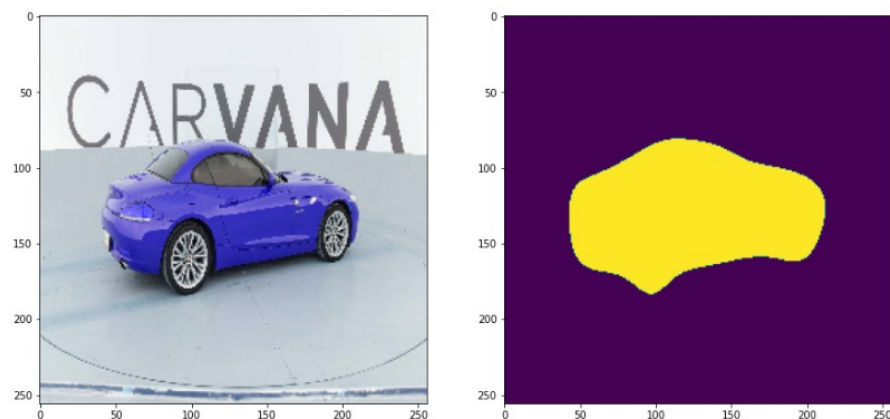| Model	Dice coefficient	IoU	Accuracy
U-Net	**0.9884**	**0.9568**	**0.9760**
FCN	0.9709	0.9512	0.9715
SegNet	0.9509	0.9320	0.9213

Demonstration

U-Net

FSN

SegNet

Conclusion

- Among the three neural network models, U-Net produces the overall best performance on the *Carvana* Dataset.

- Existing methods can well distinguish between object pixels and other pixels, e.g., all achieve more than 90% accuracy and IoU.

Future Work

- We hope to implement more Transformer-based models on practical image segmentation tasks

- We consider other downstream tasks such as Segment salt deposits beneath the Earth's surface.