Image Segmentation for cityscapes

Christopher Marek Jonathan Maxwell Kate Park Osita Arimah

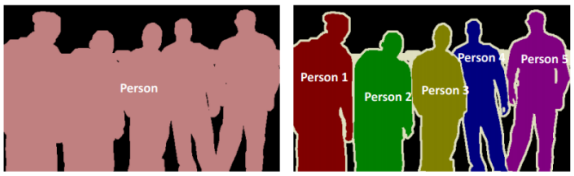
Abstract - This paper analyzes different image segmentation algorithms in a cityscape environment, their efficacy, and the evaluation metrics used to determine accuracy. It finds that a 'best-case' approach produces meaningful evaluations of the methods which are easily comparable.

I. Introduction

Image segmentation is the process of assigning labels to pixels, creating a representation that has a reduced complexity and is more meaningful. Image segmentation is useful in a wide array of fields, including medical imaging and machine vision, but it is critical in autonomous vehicles as they depend on image segmentation for real time scene understanding.

One main problem in the image segmentation field is how to evaluate the correctness of the produced image. Since images can be analyzed in many different ways, it can be difficult to identify a correct answer. Current research has proposed to customize the evaluation metric and segmentation algorithm for the specific needs of the task, but more research is needed in this area. Another key issue is balancing good accuracy without sacrificing computational complexity. For these algorithms to have use in a self-driving vehicle application they must be lightweight.

In this paper we look at comparing and contrasting two image segmentation methods; Thresholding and Edge-Based, these methods work in slightly different ways to produce a semantic segmented image, we also look at the evaluation metrics used in image segmentation and how these metrics can be tuned for the autonomous vehicle field. The evaluation metrics used in semantic segmentation are Intersection over Union, a measure of the overlap between prediction and truth, Pixel Accuracy, a measure of the number of incorrect class pixels, and PSNR.



Semantic Segmentation Instance Segmentation

A comparison of Semantic segmentation and Instance Segmentation.

This study finds that producing an evaluation image using the algorithm on an annotated example rather than a source image creates more meaningful evaluations than evaluating the algorithm on the annotated example. This method of evaluation is a novel way to compare different families of techniques (edge based, thresholding based)

These results are new in the field as there are few studies focused on tuning Edge-based segementation and thresholding-based segmentation for cityscapes. Rather, the majority of prior studies focus more on comparing the different edge-based techniques. Therefore, in comparing Canny edge-based image segmentation against thresholding-based methods, we are able to classify which technique is more efficient.

The impact of our results will help provide insights on which process is best for Image segmentation. A study analyzing Image segmentation challenges states, “Semantic image segmentation for autonomous driving is a challenging task due to its requirement for both effectiveness and efficiency [1a]”. Furthermore, through analyzing and comparing image segmentation techniques, the results of our study will provide clarification on which technique is best in the field. Additionally, this study improves practice by demonstrating how to evaluate segmented images simply and efficiently.

The structure of the report is as follows:
I. Introduction
II. Background & Related Work
III. Methods
IV. Results
V. Conclusions & Future Work

II. Background & Related Work

In the autonomous vehicle field, Image Segmentation is sometimes accomplished using Deep Network architectures which can reach much higher accuracy than traditional algorithms such as Thresholding. However, Deep Network architectures, such as FCNs, also bring a large computational overhead, which can make them non viable for application in self driving cars. Research is actively being done on how to improve the efficiency of these methods [1].

Image segmentation is a task that can be accomplished with a number of different techniques, including Thresholding, Edge-Based, Region-Based, Clustering, and Deep Network Based. These different techniques each perform best in different settings [13].

New developments in image segmentation include the processing of videos rather than images and the processing of colored images rather than solely grayscale. It is evident that this field is still growing and has yet to mature.

Thresholding:
Thresholding is an image segmentation technique that uses application specific thresholds to separate objects in the image from the background. It works by converting a grayscale image into a binary image. “a *binary image*, i.e., one that is simply black and white [2].”. The threshold chosen changes what is selected as the background or the objects of focus. There can be more than one threshold value depending on how many different segmentations are needed by the application. One threshold value segments two regions, the background and the objects [3].

Edge Based:
Edge based image segmentation is a technique that finds the components of an image by relying on the edges to distinguish different components from each other [4]. The primary method of Edge-based image segmentation is Canny Edge detection and it is a popular technique due to its ability to reduce the complexity of an image drastically [5] with little computational overhead.

Other edge-based techniques, such as Roberts and Pewitt, have been implemented in the autonomous vehicle field. In a research paper previously written on Canny Edge based segmentation, the authors write: “In this work, a Canny edges' detector is used to determine frames' edges, and essentially the road boundaries and painted lines [6].” Thus, because of the current development in the field,

it is necessary we test the accuracy of the algorithm.

Quality evaluation of image segmentation algorithms is still the topic of significant research. Fardo writes, “Currently, there is no generic metric that could be applied to any algorithm reliably” [7].

Edge detection is one of the key components in the process of lane detection [8].

In a study that compares the metrics PSNR and SSIM, it is stated that, “ the SSIM seems to be more sensitive to jpeg compression compared to the PSNR” [9]. Important to note when selecting an evaluation metric.

III. Data

To test the image segmentation algorithms we used the cityscapes dataset [10]. The dataset is made up of pictures of street scenes from 50 different cities in varying times of day and weather conditions. With this dataset and its annotated examples we can implement our own image segmentation techniques and evaluate their efficacy.

The following guide provides insight on the design choices that were made to target the dataset’s focus: (https://www.cityscapes-dataset.com/dataset-overview/).

The dataset contains the source images as well as annotated examples which are the truth images. The annotated images were produced by a model which trained for 150 epochs so the accuracy is perfect, however this model would be unrealistic in any application due to time constraints.

Group	Classes
flat	road · sidewalk · parking+ · rail track+
human	person* · rider*
vehicle	car* · truck* · bus* · on rails* · motorcycle* · bicycle* · caravan*+ · trailer*+
construction	building · wall · fence · guard rail+ · bridge+ · tunnel+
object	pole · pole group+ · traffic sign · traffic light
nature	vegetation · terrain
sky	sky
void	ground+ · dynamic+ · static+

The classes present in the dataset.

IV. Data Preprocessing

For both Thresholding and Edge-Based segmentation, and the majority of techniques, the image is first converted to grayscale. This reduces the complexity and computational requirements of the algorithm. After grayscaling, optionally, the image is downscaled or transformed to reduce the computational overhead. Then, the image has a gaussian blur applied to it to reduce the noise in the image, particularly when downsampling.

When using evaluation methods with the segmented images from the cityscapes dataset we use a series of preprocessing steps so the Edge-based and thresholding segmentation techniques are more comparable to the cityscapes ground truth. These steps include, downsampling the image to the same size as our method's result image and then converting it to a 3 channel RGB image then a grayscale image. Additional optional steps may include applying a bitmask to remove specific features such as the roads.

V. Methods

This report aims to compare Thresholding based and Edge-Based Segmentation as well as compare an industry leading Deep-Network

method FCN, analyze how these methods evaluate correctness, and finally discuss how these methods and their evaluation metrics can be tuned to perform on the cityscapes dataset.

A. Research objectives

- Implement Thresholding based image segmentation on the cityscapes dataset in order to change the pixels of the images allowing them to become simpler to analyze. The results
- Implement Edge-Based segmentation on the cityscapes dataset to segment images through edge detection and edge tracking.
- Compare different evaluation metrics for semantic segmentation using Canny and Thresholding
- Produce a suitable evaluation model/metric for comparing the segmented images from the aforementioned techniques.

B. Research methodology

Thresholding

There are two types of thresholding: global and local thresholding. With global thresholding, "the same threshold is applied to all the pixels of an image[11]". Local thresholding is when "a different threshold is applied to different parts of the image, based on the local information of the pixels [11]". Local thresholding is useful for cases when the image has uneven lighting and other irregularities.

All thresholding algorithms use a threshold value, while adaptive threshold algorithms also take a size parameter to describe the size of the local area the threshold is applied to.

There are two types of adaptive thresholding: Mean, and Gaussian. "Mean thresholding is when the threshold value is the mean the local area area. Gaussian thresholding is where the threshold value is the weighted sum of the local area where weights are a gaussian window [12]." The method we are examining is mean thresholding, where the threshold value is the mean of the block size area minus some constant supplied C.

Edge-Based

The Edge-based technique is a popular method for image segmentation as it is lightweight and intuitive. It focuses on the edges of objects in the image which identifies features easily. Due to the edges containing dense information about the objects, Edge-based techniques can reduce the complexity of images remarkably. The most common Edge-based method is Canny Edge Detection [5]. This method and others like it (Prewitt, Deriche, and Roberts) require preprocessing on the image to reduce noise.

The algorithmic steps for the canny edge detection technique are follows:

1. Convert the image to grayscale. The intensity values of the pixels are 8 bit and range from 0 to 255.
2. Convolve the image with a Gaussian function to get a smooth image. The blur removes some of the noise before further processing the image.
3. Apply the first difference gradient operator to compute edge strength then edge magnitude and direction are obtained as before.
4. After the edge directions are known, non-maximum suppression now must be applied. Non-maximum suppression is used to trace along the edge in the edge

direction and suppress any pixel value (sets it equal to 0) that is not considered to be an edge.

5. Apply threshold to the non-maximal suppression image.

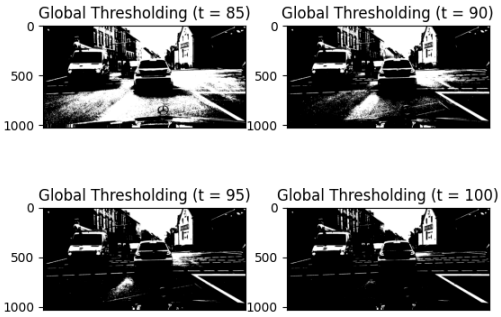
After the canny algorithm is performed, the holes created by the edges are filled (binary fill holes) to create object masks, and the segmented image is displayed.

VI. Experimental Results

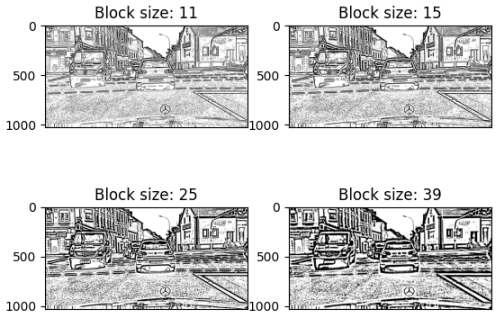
Thresholding

Source Image:

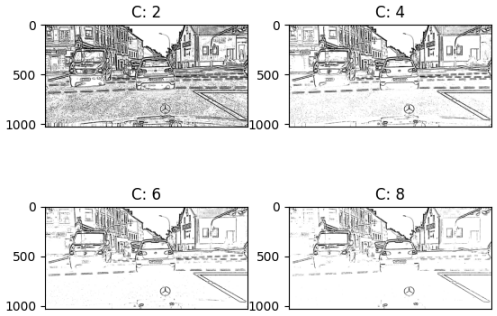
Global Thresholding Applied with threshold values plotted:



Adaptive Threshold (Mean with different block sizes):



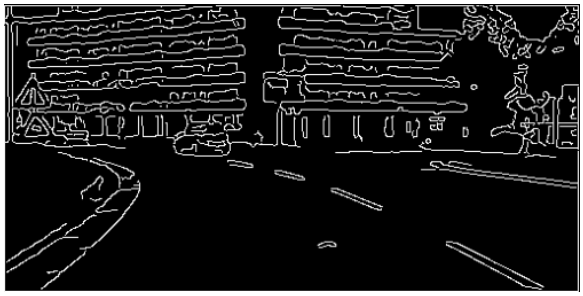
Adaptive Threshold (Mean with different constants): and Block size = 15



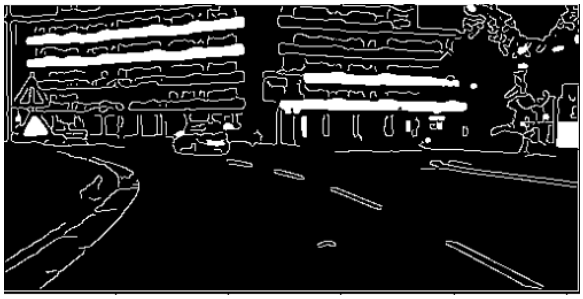
Since the source image has uneven lighting, mean thresholding performs better than global thresholding.

Edge-Based

Source Image



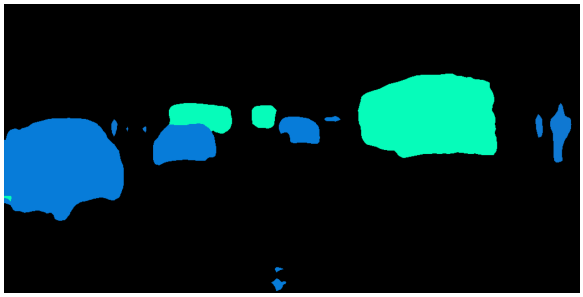
Edges obtained from the image with Canny



Final Segmented Image

FCN, An industry leading method

Source Image



Final Segmented Image

VII. Accuracy Evaluation

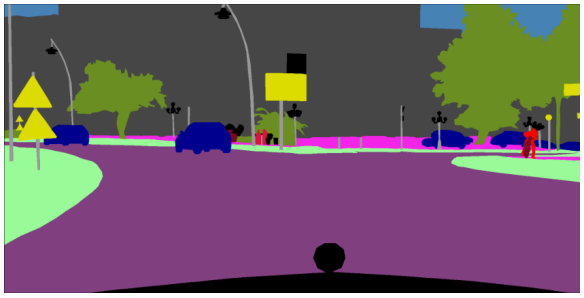
Peak signal-to-noise ratio (PSNR) [7] the ratio of maximum possible power and occurring noise that can affect the representation of the image. In order to compute the quality reconstruction of an image, PSNR can be used to evaluate the quality of segmented images. It is a case where original data is treated as signal and occurring error is treated as noise. The higher the PSNR, the better the quality of the compressed, or reconstructed image.

The PSNR can be expressed as the following:

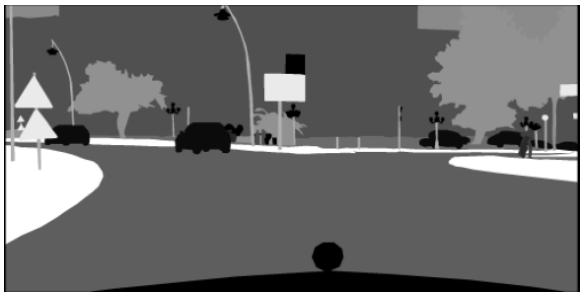
$$PSNR = 10.log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

(Where, MAX = Maximum pixel value of image).

To evaluate the accuracy of the segmented images, we compare the segmented images produced by our Edge-based and Thresholding based methods to the truth segmented image in the cityscapes dataset.



In order to prevent a produced image from scoring higher due to it having similar colors rather than accurate segmentation, we grayscale the image.



This image is then compared against the image produced by the Edge-based segmentation and the Thresholding-based segmentation.

Reconstruction score of Edge-based segmentation: ~8.394

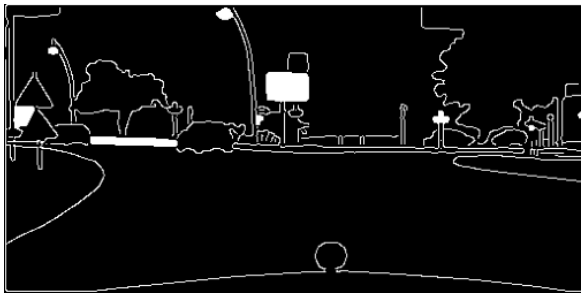
Reconstruction score of Thresholding-based segmentation:

Global, t= 95: ~ -40.21

Adaptive, Blocksize=15: ~ -45.9588

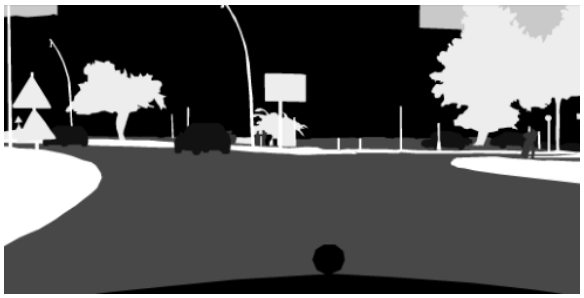
Adaptive, C=6: ~ -46.817

We can also evaluate these algorithms in a different manor. If we create our own version of the truth image by applying the algorithm on the annotated image this would yield an image which represents the best the algorithm could possibly do.



This is the 'best case' image for Edge-Based segmentation. Produced by apply the Edge-based method on the annotated image rather than the source image.

The reconstruction score of the Edge-Based method on the tailored truth image using PSNR is: 21.73426.



This is the evaluation image created by the thresholding-based method on the annotated example rather than the source image.

The global thresholding with t=95 scores: ~ -40.121

Adapative thresholding with Block size=15 scores: ~-44.058

Adaptive thresholding with C=6 scores: ~-42.25

VIII. Conclusion and Future work

Thresholding

Thresholding is one of the simplest image segmentation methods, and it can only output a binary image. For some tasks and for preprocessing it is a suitable method. For best results it should not be used with images with lots of noise or very uneven lighting.

Edge-Based

The algorithm does reduce the complexity of the image, but it retains only the fine edges of the image, so it scores poorly on the PSNR evaluation. However, when tuning the evaluation to the Edge-Based 'best-case' we find that it scores much higher.

It's clear due to the different nature of these images that comparing them can present challenges in producing meaningful results. Our findings indicate that the evaluation metric should be tuned to the algorithm being evaluated in order to meaningfully compare the accuracy of these methods.

Future Work

Future work may include implementing alternate evaluation metrics such as SSIM [9] and comparing both to determine which is better at capturing the measure of 'accuracy'.

IX. References

- [1a.] Papadeas, I.; Tsochatzidis, L.; Amanatiadis, A.; Pratikakis, I. Real-Time Semantic Image Segmentation with Deep Learning for Autonomous Driving: A Survey. Appl. Sci. 2021, 11, 8802. https://doi.org/ 10.3390/app11198802

- [1] Balabanian, Felipe & Da Silva, Eduardo & Pedrini, Helio. (2017). Image Thresholding Improved by Global Optimization Methods. Applied Artificial Intelligence. 1-12. 10.1080/08839514.2017.1300050.

- [2] Thresholding – Image Processing with Python. (2022). Retrieved 8 April 2022, from https://datacarpentry.org/image-processing/07-thresholding/#:~:text=Thresholding%20is%20a%20type%20of,is%20simply%20black%20and%20white

- [3]
- https://arxiv.org/abs/2006.00356
Kanan C, Cottrell GW (2012)
Color-to-Grayscale: Does the Method Matter in Image Recognition? PLoS ONE 7(1): e29740. https://doi.org/10.1371/journal.pone.0029740

- [4] Tyagi, M. (2022). Image Segmentation : Part 1. Retrieved 8 April 2022, from https://towardsdatascience.com/image-segmentation-part-1-9f3db1ac1c50.

- [5] Vadapalli, P. (2022). Image Segmentation Techniques [Step By Step Implementation] | upGrad blog. Retrieved 8 April 2022, from https://www.upgrad.com/blog/image-segmentation-techniques/#2_Edge-Based_Segmentation

- [6] Bounini, Farid & Gingras, Denis & Lapointe, Vincent & Pollart, Herve. (2015). Autonomous Vehicle and Real Time Road Lanes Detection and Tracking. 1-6. 10.1109/VPPC.2015.7352903.

- [7] Fardo, F., Conforto, V., Oliveira, F., & Rodrigues, P.. (2016). A Formal Evaluation of PSNR as Quality Measurement Parameter for Image Segmentation Algorithms.

- [8] Ansari, Mohd & Kurchaniya, Diksha & Dixit, Manish. (2017). A Comprehensive Analysis of Image Edge Detection Techniques. International Journal of Multimedia and Ubiquitous Engineering. 12. 1-12. 10.14257/ijmue.2017.12.11.01

- [9] Horé, Alain & Ziou, Djemel. (2010). Image quality metrics: PSNR vs. SSIM. 2366-2369. 10.1109/ICPR.2010.579.

- [10]https://www.cityscapes-dataset.com/dataset-overview/

- [11] Niblack, Wayne (1986). *An introduction to digital image processing*. Prentice-Hall International. ISBN 0-13-480600-X. OCLC 1244113797.

- [12] C. Gonzalez, R. (2022). Image Thresholding. Retrieved 8 April 2022, from https://amroamroamro.github.io/mexopencv/opencv/thresholding_demo.html

- [13] Acharjya, Pinaki Pratim, Ritaban Das, and Dibyendu Ghoshal. "Study and comparison of different edge detectors for image segmentation." Global Journal of Computer Science and Technology (2012)

- Kaymak, Ç., Uçar, A. (2019). A Brief Survey and an Application of Semantic Image Segmentation for Autonomous Driving. In: Balas, V., Roy, S., Sharma, D., Samui, P. (eds) Handbook of Deep Learning Applications. Smart Innovation, Systems and Technologies, vol 136. Springer, Cham. https://doi.org/10.1007/978-3-030-11479-4_9

- https://openreview.net/pdf?id=S1uHiFyyg

- Pulkit, S. (2022). Image Segmentation | Types Of Image Segmentation. Retrieved 8 April 2022, from https://www.analyticsvidhya.com/blog/2019/04/introduction-image-segmentation-techniques-python/

.