MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL INDIA, 462003

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Semantic Image Segmentation

Minor Project Report

Semester VI

Submitted by:

Aditya Kumar	181112251
Pradhi Anil Kumar Das	181112252
Pushpak Manwani	181112286
Aaron Paul	181112287

Under the Guidance of
Dr. Dehalwar Vasudeo
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL INDIA, 462003

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the project report carried out on

"Semantic Image Segmentation" by the 3rd year students:

Aditya Kumar	181112251
Pradhi Anil Kumar Das	181112252
Pushpak Manwani	181112286
Aaron Paul	181112287

Have successfully completed their project in partial fulfillment of their Degree in Bachelor of Technology in Computer Science and Engineering.

____

Dr. Dehalwar Vasudeo
(Minor Project Mentor)

DECLARATION

We, hereby declare that the following report which is being presented in the Minor Project Documentation Entitled as **"Semantic Image Segmentation"** is an authentic documentation of our own original work and to the best of our knowledge. The following project and its report, in part or whole, has not been presented or submitted by us for any purpose in any other institute or organization. Any contribution made to the research by others, with whom we have worked at Maulana Azad National Institute of Technology, Bhopal or elsewhere, is explicitly acknowledged in the report.

Aditya Kumar	181112251
Pradhi Anil Kumar Das	181112252
Pushpak Manwani	181112286
Aaron Paul	181112287

ACKNOWLEDGEMENT

With due respect, we express our deep sense of gratitude to our respected guide and coordinator **Dr. Dehalwar Vasudeo** , for his valuable help and guidance. We are thankful for the encouragement that he has given us in completing this project successfully.

It is imperative for us to mention the fact that the report of the minor project could not have been accomplished without the periodic suggestions and advice of our project guide **Dr. Dehalwar Vasudeo** and project coordinators Dr. Dhirendra Pratap Singh and Dr. Jaytrilok Choudhary.

We are also grateful to our respected director Dr. N. S. Raghuwanshi for permitting us to utilize all the necessary facilities of the college.

We are also thankful to all the other faculty, staff members and laboratory attendants of our department for their kind cooperation and help. Last but certainly not the least; we would like to express our deep appreciation towards our family members and batch mates for providing the much needed support and encouragement.

ABSTRACT

Image Segmentation is the process by which a digital image is partitioned into various subgroups (of pixels) called Image Objects. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze.Image segmentation is typically used to locate objects and boundaries in images.

Some Applications of Image Segmentation:

- Improved Quality of MRI/Medical images for better detection of diseases and problems
- Segmenting images can help to improve robot vision
- Image segmentation can also be applied to satellite images in order to get better object detection
- Self-driving vehicle autonomous cars must be able to perceive and understand their environment in order to drive safely. Relevant classes of objects include other vehicles, buildings, and pedestrians. Semantic segmentation enables self-driving cars to recognize which areas in an image are safe to drive.

- Iris recognition a form of biometric identification that recognizes the complex patterns of an iris. It uses automated pattern recognition to analyze video images of a person's eye.

- Face recognition identifies an individual in a frame from a video source. This technology compares selected facial features from an input image with faces in a database.

As well as many other application are also possible to be implemented with the help of image segmentation.Our group will try to implement some of these applications of image segmentation.

TABLE OF CONTENTS

1.Introduction

In digital image processing and computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as image objects). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

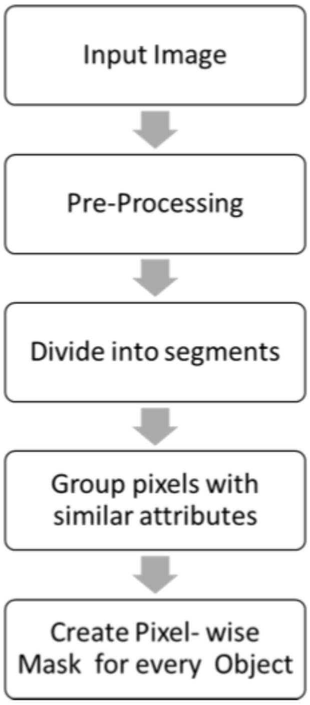
The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like marching cubes.

Basically segmentation is not just a simple procedure, it includes and executes multiple of chores. The images that are lay open to the process of segmentation are segregated into manifold divisions, and the valuable insights would be obtained as a result either as texture, intensity or color. In addition from this the

segmentation take an important part in recognizing the images.

Several techniques are available and more new techniques are devised to perform the procedures of segmentation. But the available and the devised methodologies prove to be less efficient as they are practically incompatible over all sorts of images. So there arises a necessity to identify an effective segmentation procedure, to enhance the performance of the image processing.

The segmentation procedure scopes in simplifying the examination process into eloquent and easily analyzable form. It is vital and the initial procedure in the analyzing an image. The main purpose of the segmentation process is to segregate the images to have the similar attributes. It is essential first step in many of the application like object detection and recognition, retrieval of image based on the content, video surveillance, medical imaging etc. Apart from this the segmentation is the first and the foremost step in developing a multi label classifier construction to classify more than one object found in the image. For this it works by segregating the image into multiple sections and groups the same attributes pixels and creates a pixel-wise mask for all the objects in the image. The figure below shows the stages in the **segmentation process**.



Basically the segmentation process could be classified into two fundamental kinds as local and global segmentation and further based on the image properties as discontinuity detection strategy and similarity detection strategy.The former types of segmentations mentioned are concerned about the particular section or the area of an image and whole image that is encompassed with the huge set of pixels. The later segmentation types based on the image properties segments the images into regions based on the discontinuity and similarity respectively. The several segmentation methods of the image are listed below in figure.2

There are more than a few prevailing practices that are utilized for segmenting images. These methods have their own significance. These modus operandi could come within reach of from dualistic elementary methods of segmentation i.e. region based or the edge based, every single procedures can be applied on dissimilar images to accomplish the necessary segmentation. All these techniques could be classified based on three categories. As (i) structural: works on the structure of the region, (ii) Stochastic: works on the discrete pixel values and (iii) hybrid: Combines both the structural and the stochastic.

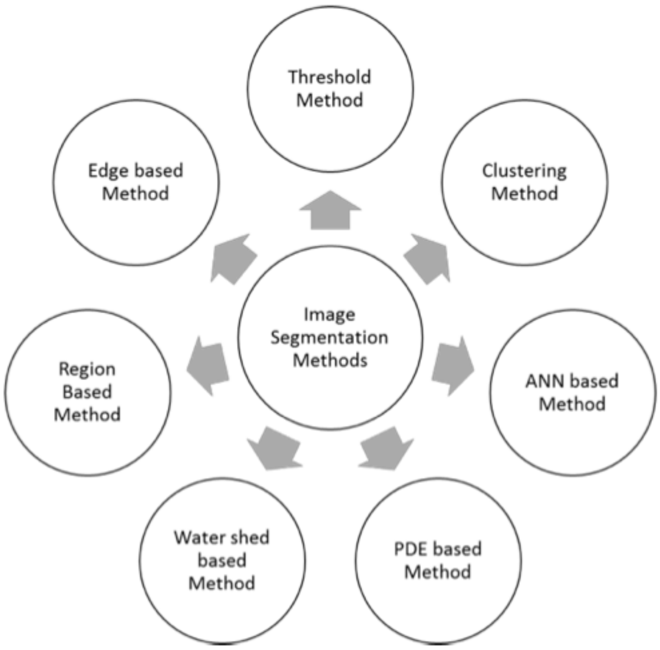


Figure. Image Segmentation Methods

As the digital images acquired are projected to multitudes of distortions that deteriorates the quality of the image, an effective segmentation process becomes a necessary requisite for retaining the valuable data of the images especially in medical imaging.

Some Applications of Image Segmentation:

- Improved Quality of MRI/Medical images for better detection of diseases and problems
- Segmenting images can help to improve robot vision
- Image segmentation can also be applied to satellite images in order to get better object detection
- Self-driving vehicle autonomous cars must be able to perceive and understand their environment in order to drive safely. Relevant classes of objects include other vehicles, buildings, and pedestrians. Semantic segmentation enables self-driving cars to recognize which areas in an image are safe to drive.

- Iris recognition a form of biometric identification that recognizes the complex patterns of an iris. It uses automated pattern recognition to analyze video images of a person's eye.

- Face recognition identifies an individual in a frame from a video source. This technology compares selected facial features from an input image with faces in a database.

As well as many other application are also possible to be implemented with the help of image segmentation.Our group will try to implement some of these applications of image segmentation.

2.Literature Review and Survey

There have been many works done in the area of image segmentation by using different methods. And many are done based on different application of image segmentation.K-means algorithm is the one of the simplest clustering algorithm and there are many methods implemented so far with different method to initialize the centre. And many researchers are also trying to produce new methods which are more efficient than the existing methods, and shows better segmented result.

Some of the existing recent works are discussed here.

Pallavi Purohit and Ritesh Joshi 4 introduced a new efficient approach towardsK-means clustering algorithm. They proposed a new method for generating the cluster center by reducing the mean square error of the final cluster without large increment in the execution time. It reduced the means square error without sacrificing the execution time. Many comparisons have been done and it can conclude that accuracy is more for dense dataset rather than sparse dataset.

Madhu Yedla, Srinivasa Rao Pathakota, T. M. Srinivasa 6 proposed EnhancingK-means clustering algorithm with improved initial center. A new method for finding the initial centroid is introduced and it provides an effective way of assigning the data points to suitable clusters with reduced time complexity. They proved their proposed algorithm has more accuracy with less computational time comparatively original k-means clustering algorithm. This algorithm does not require any additional input like threshold value. But this algorithm still initializes the number of cluster k and suggested determination of value of k as one of the future work.

K.A. Abdul Nazeer, M. P. Sebastian 7 proposed an enhanced algorithm to improve the accuracy and efficiency ofthe k-means clustering algorithm. They present an enhanced k-means algorithm which combines a systematic method consisting two approaches. First one is finding the initial centroid and another is assigning the data point to the clusters.They have taken different initial centroid and tested execution time and accuracy. From the result it can be conclude that the proposed algorithm reduced the time complexity without sacrificing the accuracy of clusters.

Koresh, M. H et al [9] presented the "Computer vision based traffic sign sensing for smart transport." Chandy, Abraham et al [10] elaborated the "A Review on Iot Based Medical Imaging Technology for Healthcare Applications." Bindhu, et al [11] presents the "Biomedical Image Analysis Using Semantic Segmentation." V Kumar, T. Senthil et al [12] proposed the "A Novel Method for HDR Video Encoding, Compression and Quality Evaluation."

3. Proposed work and Methodology

In this section we will mention how each of our application works and the principle behind it.

3.1 Generic Object Segmentation

This folder contains a python file that will convert the given images into segmented images with the help of K-Means Clustering Algorithm

The K-means algorithm is an iterative technique that is used to partition an image into K clusters.The basic algorithm is

1. Pick K cluster centers, either randomly or based on some heuristic method.
2. Assign each pixel in the image to the cluster that minimizes the distance between the pixel and the cluster center
3. Re-compute the cluster centers by averaging all of the pixels in the cluster
4. Repeat steps 2 and 3 until convergence is attained (i.e. no pixels change clusters)

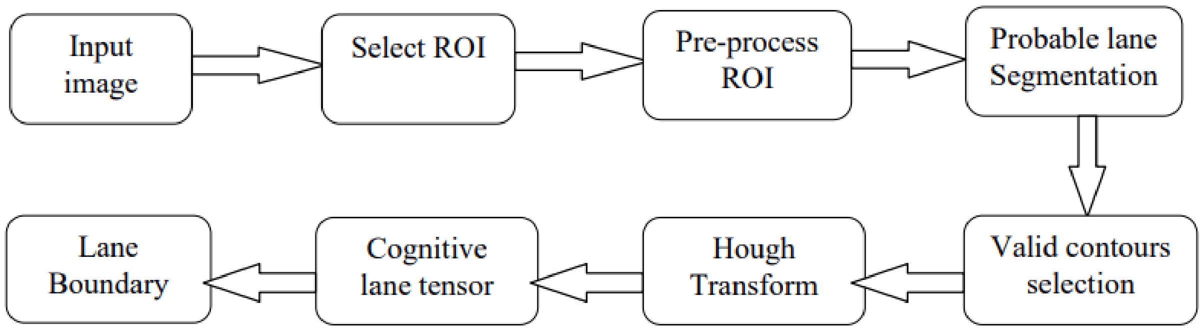
In this case, distance is the squared or absolute difference between a pixel and a cluster center. The difference is typically based on pixel color, intensity, texture, and location, or a weighted combination of these factors. K can be selected manually, randomly, or by a heuristic. This algorithm is guaranteed to converge, but it may not return the optimal solution. The quality of the solution depends on the initial set of clusters and the value of K.

3.2 Medical Image Segmentation

The main principle involved in this project is that the images are segmented with K-Nearest Neighbour algorithm with K=3 as the most common colors in a MRI scan are various shades of black ,white and grey and with just 3 colors the images of the scan will be segmented properly which will provide a lot of useful information to the doctor,such as the place where the tumor is and the place where the mass is most dense etc.

The steps Involved in K-NN algorithm:
1.Choosing the number of Clusters
2.Selecting at random K points for centroid, in our case 40 was passed as the number of neighbors.
3.Assigning each Data point as we say each pixel value closest to the above centroid that further gives us clusters.
4.Now we compute and place the new centroid for each cluster.
5.On the last step we just do the reassignment of the new nearest centroid and if in any case any new reassignment took place we would reiterate the above process.

3.3 Lane Detection Program



A) Selection of Region of Interest: Many papers deal with finding the vanishing point of the image and considering only the region below as ROI. Computing vanishing point is not a trivial task and requires additional computational process. We exploited the mounted camera's property which is so placed that nearly bottom half of the image is the road. By doing so we don’t need to find the vanishing point and the bottom half of the image is taken into account for further processing. It gives a nice approximation considering the computational overhead required for calculating vanishing point.

B) Pre-processing of the ROI: Unlike day time images, night vision barely provides with any color property of the image which would have otherwise provided us with unique colour set for lane. Along with only meagre hue variation and low saturation levels even the brightness variation of such images is very delicate and highly dependent on external lighting (like street lamps, vehicle’s tail/head lamp, signals and the glare because of all these). To get the intensity value of each pixel, we change the color space from RGB to Black and white using the following equation:

Gray = (Red * 0.299 + Green * 0.587 + Blue * 0.114)

This helps in faster computation without rejecting any valuable information. Bad lighting conditions can make the image harder to process and might result in false results. To avoid this, median filtering was applied. A square kernel of window size 5 was taken and passed over the entire image. The pixel lying in the centre was then replaced with the median value in the kernel. It helped in removing any susceptible noise while maintaining edges. The resultant ROI image's contrast and brightness values were modified to enhance the gradient changes and other boundary features.

C) Probable lane marking segmentation: Primary step in the lane marking detection is to identify probable lane marking segments. Proper selection is inevitable as it will serve the basis of further processing. Edge detection techniques (eg. Canny and sobel) and thresholding (global and locally adaptive) were possible approaches but both had their own limitations over here. Canny results in either over-dividing the image region or underdividing it. It doesn't take into account the property of expanded neighbourhood. Edges are detected only based on the immediate surroundings. Dynamic threshold for the entire image using Otsu's method failed because of the varying light conditions in the entire image. Intensity of lane segments themselves varied widely from regions near the headlights/tail lights of the vehicles to shadowed area, whereas adaptive thresholding detected many noise apart from lane markings.

$$d1 = I(x,y) - I(x-\delta, y)$$
$$d2 = I(x,y) - I(x+\delta, y)$$
$$D = d1 + d2 - \left|\left(I(x+\delta, y) - I(x-\delta, y)\right)\right|$$
$$L = 0.5 * I(x,y)$$

$$L(x,y) = \begin{Bmatrix} if(d1 > 0 \ and \ d2 > 0 \ and \ D > L) & 255 \\ else & 0 \end{Bmatrix}$$

The possible lane markings were selected based on the above parameters. Represents the lane width. The lane being brighter in intensity relative to its sides, only if both sides are darker and the sum of the intensity value difference on either side is between a given range then only the pixel is considered as a part of lane segment. The range was calculated using numerous sample point and plotting them. It was found to vary between the 0.25 and 0.75 based on output image's lighting conditions, which after pre-processing the image for better contrast and brightness can be ideally chosen to be around 0.5. Any difference less than that come under noise or non-lane parts. The lane width varies based on distance because of the perspective. Near the base it's maximum whereas near the vanishing point it's minimal. Lane width at any row(r) of the image is calculated using the following formula:

$$\delta_r = min + (max - min) * \frac{(r - r_{vanishingpoint})}{(r_{total} - r_{vanishingpoint})} + \varepsilon$$

max and min represent the maximum and minimal lane width possible in the given image. Keeping value 5 helps in avoiding noises. Max value is dependent on the image size and the mounted camera's position. If the camera is kept very low then because of high perspective and being closer to lane, lane width will be larger near the base as compared to the time when camera is mounted on the top. Minimum by default always remains 0 at vanishing point; otherwise it can be adjusted if required. Once max is set, the above formula can be used to get lane width at varying distances dynamically. Dynamically changing the lane width helps in accurate selection of lanes.

D)Selecting only possible lane segments from above processed image: The above segmentation process often selects other unwanted noise or regions similar to lane (eg, milestones, edges of vehicles, railings, trees, lampposts, headlight glare etc.). We exploited the geometric features of a lane segment and based on its property we selected only valid segments. First of all contours were selected from the above binary image using the [Suzuki85] algorithm1 .Then a minimum area rectangle was drawn around it to get its orientation, length and breadth properties.

E)Hough Transform:Line segments are either continuous or broken based upon where they are marked. Along with this because of poor light conditions and glare from other vehicle's head/tail lamp often continuous lane segments are not obtained. To obtain a proper lane a fixed continuous lane boundary is required. We used the hough transform that serves the purpose well and provides us with the connected lane segment.

F)Cognitive lane tensor: Knowledge of lane which was computed in previous frames can be used to predict the lane in upcoming frames. It can be very useful in regions where either lanes are not marked or marked as false negatives i.e which are undetectable by the above procedure because of any unwanted reason. It can also serve as a tensor such that the currently detected lanes dont deviate drastically. This will help in avoiding consequences of false positives.These are the steps involved in lane segmentation program.

3.4 Road Segmentation using Satellite Images

Get_Data_Generators.py:Builds and returns data generators.

acquire_data.py:Downloads the Massachusetts Roads Dataset or the Massachusetts Buildings Dataset. By changing "link_file" to point at a custom list of links, you can download any other dataset too,but we worked with this dataset.

build_dataset.py; Crops the images into small 256x256 images and divides the data set into training and testing set.

Road_detection_GPU.ipynb:define our model and make a custom loss function and finally testing our model and getting the output.Finally our model is made and it can be used similarly with other data sets.So Detection of road is possible with the help of satellite imagery with the help of our model.

4.Gaps Identified and their Solutions

4.1 Generic Object Segmentation

the gaps identified in this program are for a large value of K (number of cluster) the output image is over-segmented ,an example is given below

(i) Original Image used for Segmentation

(ii) Output image that is over-segmented due to a large K value

This problem can be easily corrected by changing the value of K to a lower value ,but if the value is too low also the segmentation will be not done properly.example is given below when K is too small

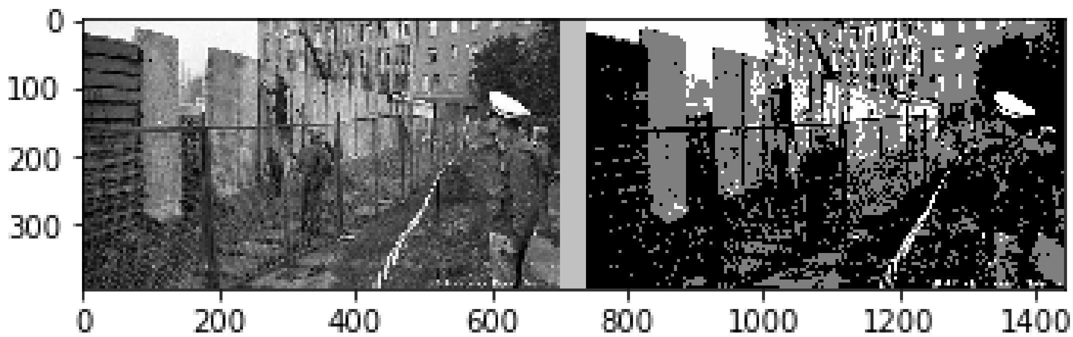
(i) Original Image used for Segmentation

(ii) Output image that is improperly segmented due to a small K and iteration value

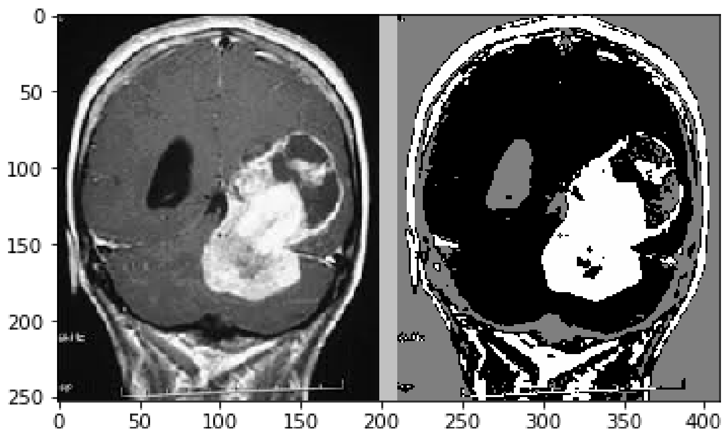
Another problem other than deciding the correct value of K and number of iteration, is that it takes a lot of time to segment a single image (approximately 30 to 40 minutes depending on the value of K and iterations) .To overcome this problem the next program uses KNN algorithm and is implemented in Jupiter notebook to directly show the results instead of waiting for the result during the presentation.

4.2 Medical Image Segmentation

This program works well when the images passed to it are MRI scans and X-rays because it was made to be implemented in this domain and passing any other image will not provide the same quality of result as it does with medical images.



This is an image of Berlin during the world war 2,as you can clearly see the image is not segmented properly and it more distorted than the original image , This is because the application was made with medical images in mind. Below is an example of a medical imaged passed to the program:



We can clearly see than the medical image is properly segmented and we can identify what the injury is and the places where the masses in the brain can be found and where it is most dense.Since it is implemented by Jupiter notebook we can directly see the result instead of waiting for the processing time to get the results(which was a problem in the previous program,its equivalent python code is also provided just in case).

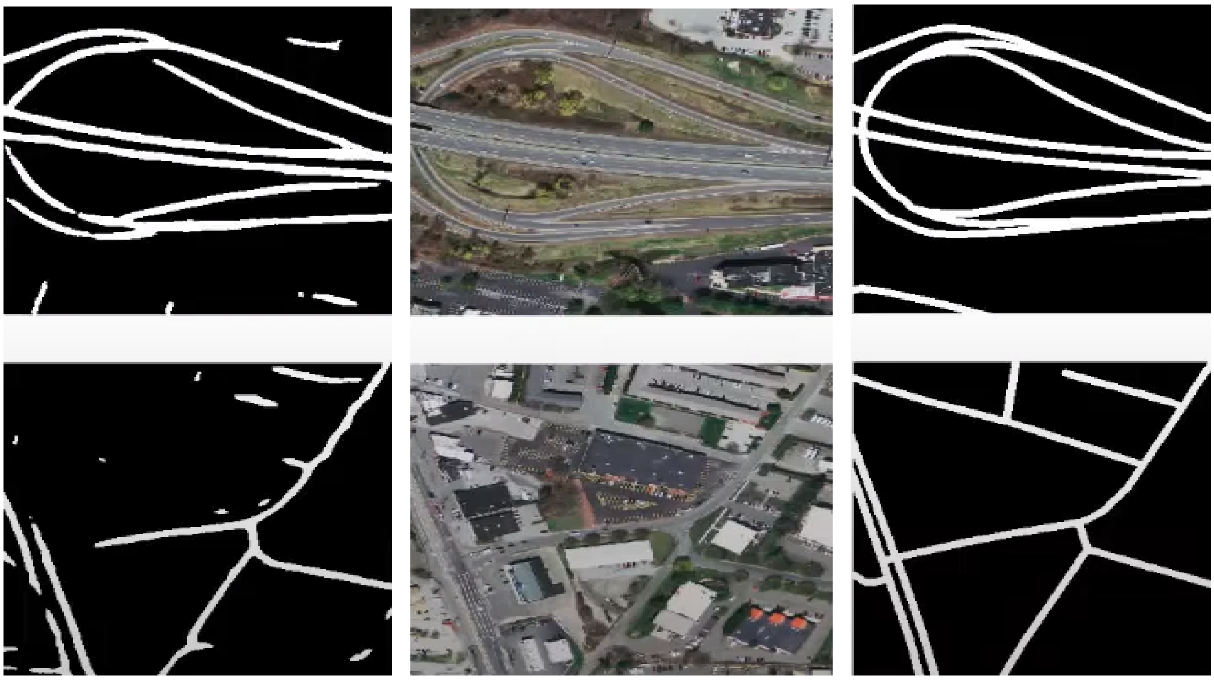
4.3 Lane Detection Program

This program works well when the area is properly lit(daytime) compared to when it is dark(nighttime).another problem encounter was how to decide the lane segment, and the lane segment properties I have taken into consideration were:
· Segment area. Area of segments below minArea threshold denote unwanted object and hence were rejected.
· Ratio of sides. Being a line segment, the ratio of its length to breadth should be greater than 4:1 at least. Only segments with higher ratio were taken into account. Segments having area less than certain threshold but more than minArea can possibly represent small broken centre lane markings and hence ratio for them was brought down to 2:1.
· Orientation. Lane segments by the virtue of their nature are never close to horizontal (unless extremely steep turn is encountered). This property helped us in removing highlighted vehicle bumpers and other segments which were otherwise being treated as false positives. Vertical lane segments are possible only if the vehicle is on the lane, in such cases the lane segment will be near the bottom centre region of the image only. Every other vertical segment detected cannot be lanes hence are discarded. A min area rectangle was bounded to the detected segment. Lane being very close to rectangles, if the segment area was not close to area of bounding rectangle then the segment was rejected.

The lane segments were the resultant segments that passed the above test. Minor false positives still remained but those can be easily avoided by further process.

4.4 Road Segmentation using Satellite Images

One thing that we observed when making our project,we noticed that roads that were surrounded by grass had better performance than the roads that we near parking lots.this can be seen below:



In the above image the first column contains the predicted road, the second column contains the actual image being processed and the final column contains the actual location of the roads.

5.Tools and Technology to be used

The main technologies that was used in each project is mentioned below:

1. Generic Object Segmentation:Python
2. Medical Image Segmentation:Python and Jupyter Notebook
3. Lane Detection Program:C++
4. Road Segmentation using Satellite Images: Jupyter Notebook and Python
5. Website:HTML,CSS,JavaScript and Django

Some of the tools used by our group are:

1. Visual Studio Code- LiveShare

2. GitHub
3. Microsoft Teams

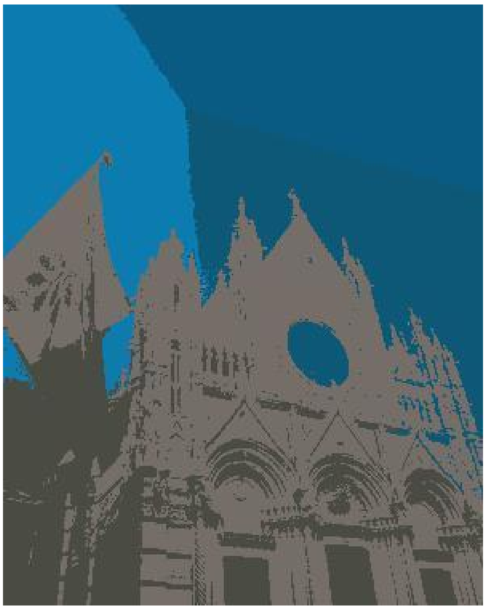
6.Conclusion and Results

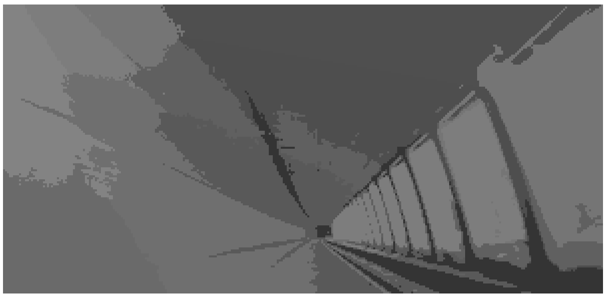
The conclusion of our Minor Project is that there are a lot of applications and projects what can be made in the field of image segmentation that can be use to humankind and our project just scratches the surface of what is possible with the technology available in this field ,it provides a starting point for further development in the future for research in this field.We would again like to thanks our parents and teacher without whom this project could not be completed.

The result of each of our Program are given below:

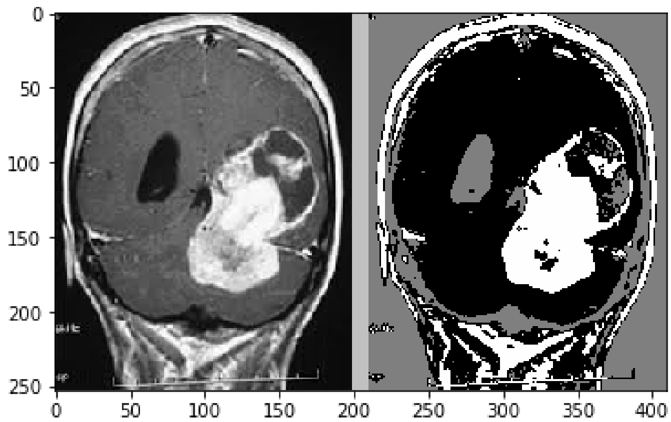
Results of Generic object segmentation program:

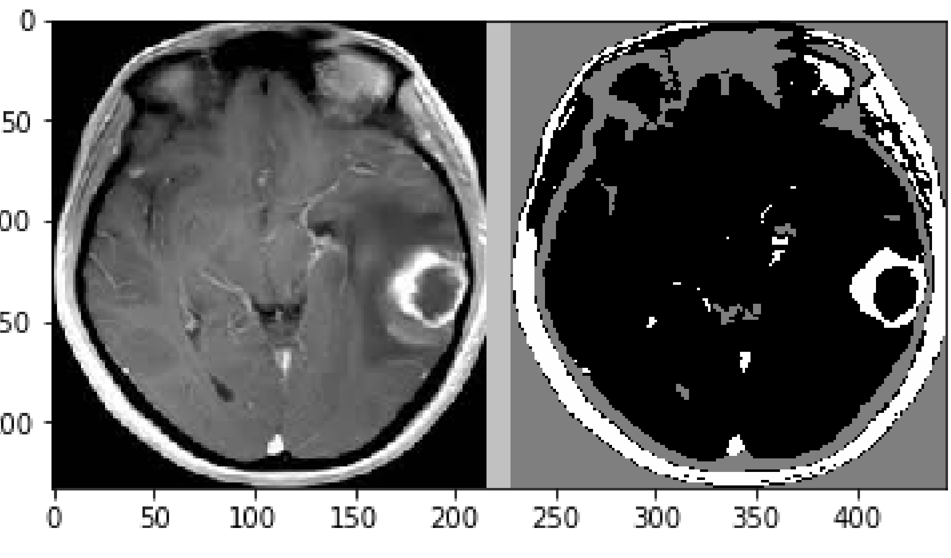


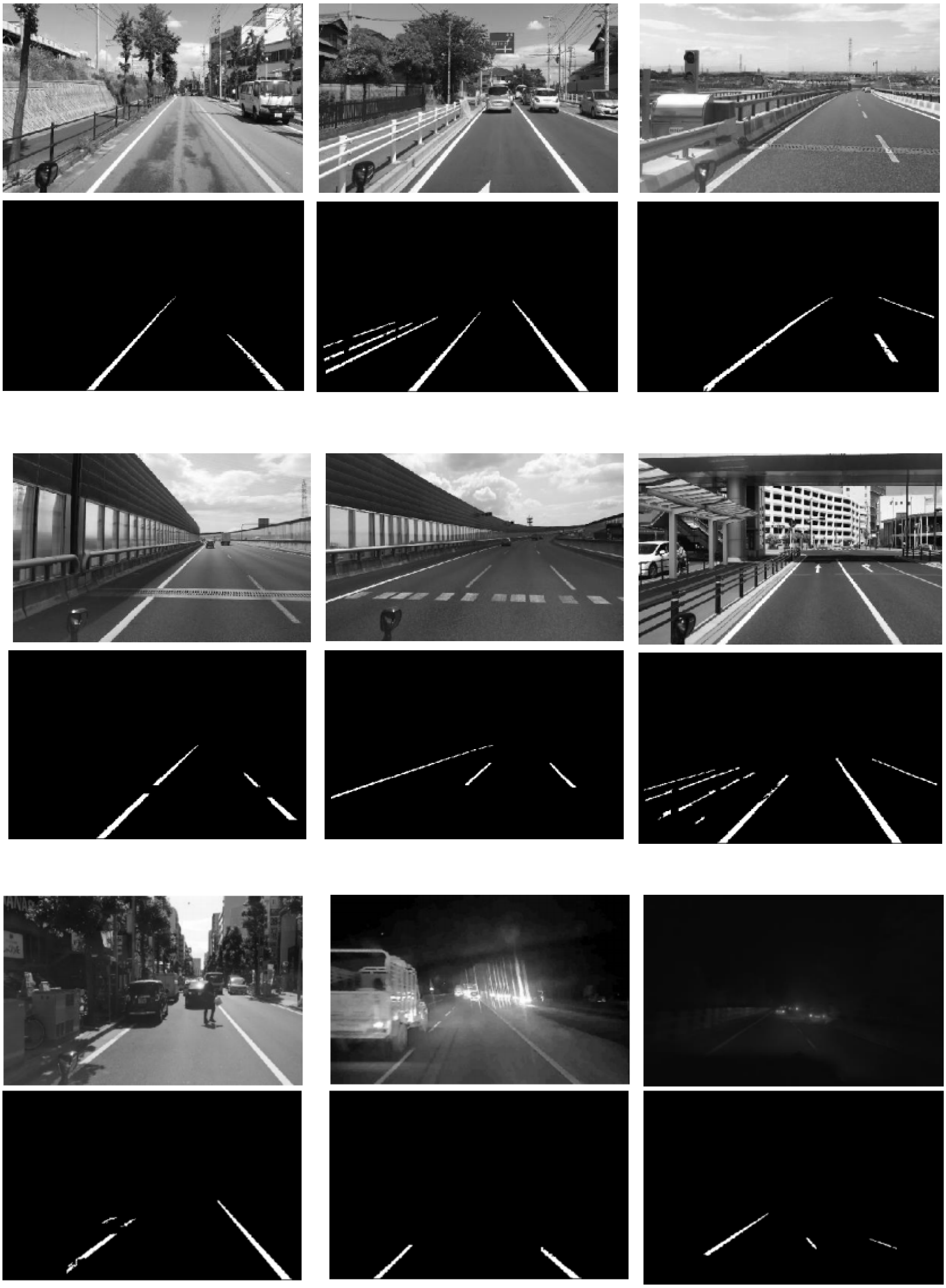


Results of Medical Image Segmentation Program:

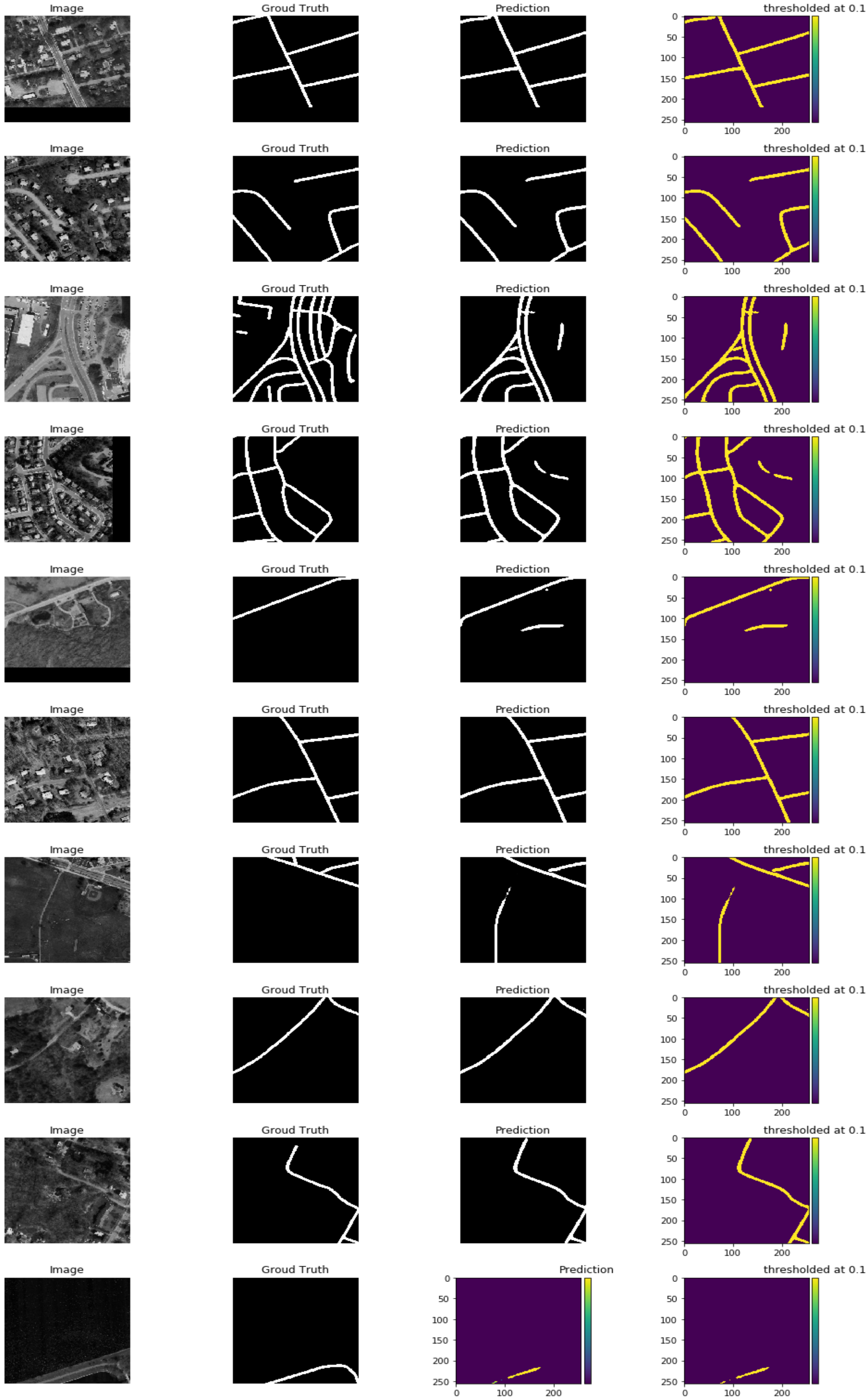




Results of Lane Detection Program:



Result of Road Segmentation using Satellite Images Program:



7.References

https://en.wikipedia.org/wiki/Image_segmentation#:~:text=Image%20segmentation%20is%20typically%20used,same%20label%20share%20certain%20characteristics.

https://www.researchgate.net/publication/340087951_Image_segmentation_Techniques_and_its_application

https://www.cs.toronto.edu/~vmnih/docs/Mnih_Volodymyr_PhD_Thesis.pdf

https://arxiv.org/pdf/1505.04597.pdf

https://www.youtube.com/watch?v=Gs5HlHKqAYQ&ab_channel=MadhurBehl

https://link.springer.com/article/10.1007/s10278-019-00227-x

https://link.springer.com/article/10.1007/s11263-019-01247-4#:~:text=Generic%20object%20detection%20is%20closely,semantic%20segmentation%20which%20does%20not.