Computer Vision for Autonomous Vehicles
Semantic Segmentation using Jetson Nano

Presented by

AVINASH R (122004028)

School of Electrical and Electronics Engineering
SASTRA Deemed to be University

Motivation

- Over the past 100 years, Automobile Industry has seen steady innovation with the help of Technology on par with other fields.

- Self driving cars have taken the world by storm and its efficiency has been proved to be better than calculated.

- Despite its popularity, this tech is yet to be tested and green-lit for Indian roads, and that is the motivation behind this project.

- This demonstration is a scaled down approach to the world of Computer vision that deals in Vehicle Automation.

Objective

- To Apply Convolutional neural networks to perform Image recognition and object detection.

- To Train the application and increase scalability using Transferred Learning technique to make it work in native environment.

- To Experiment with fully-convolutional semantic segmentation networks on a live camera stream.

Methodology

Nvidia Jetson Nano is a compact AI compute Module which can be used to learn and create intuitive Deep Learning based applications.

- The tool further required is a camera which allows to interact with the real world.

- Pre-trained models are used for experimenting with the test data and gather inferences.

- Nvidia GPU Cloud provides models as per the requirements through Docker Container which is deployed onto Jetson Nano.
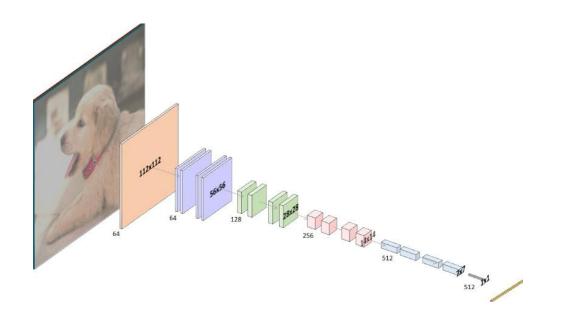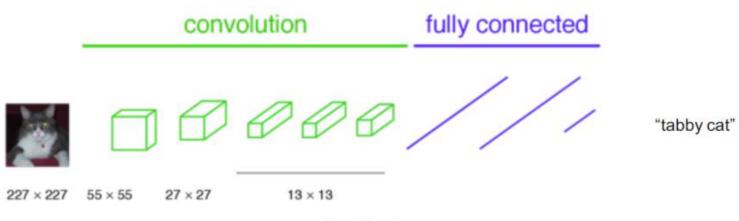
Semantic Segmentation Models

- Cityscapes

assessing the performance of vision algorithms for major tasks of semantic urban scene understanding.

- DeepScene

The DeepScene dataset consists of off-road forest trails and vegetation, aiding in path-following for outdoor robots.

- Multi-Human Parsing (MHP)

Multi-Human Parsing provides dense labeling of body parts, like arms, legs, head, and different types of clothing.

- SUN RGB-D

The SUN RGB-D dataset provides segmentation ground-truth for many indoor objects and scenes commonly found in office spaces and homes.
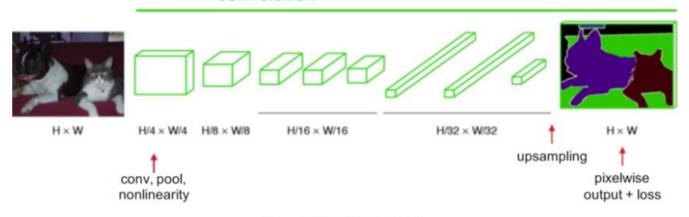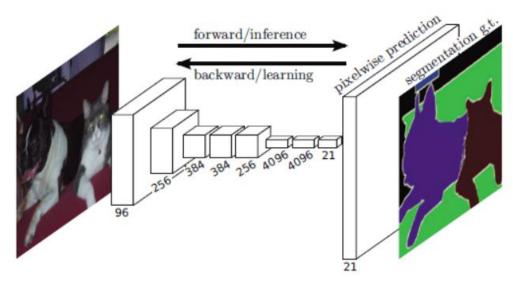
Pre-Trained model architecture

convolution   fully connected
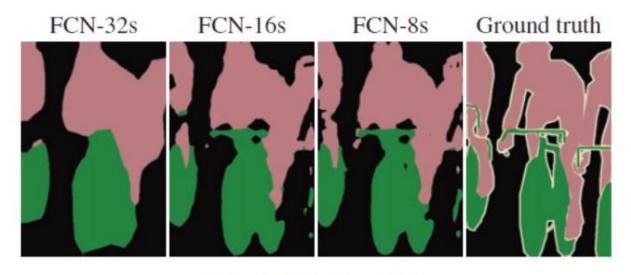
"tabby cat"

227 × 227   55 × 55   27 × 27   13 × 13

Classification

convolution

H × W   H/4 × W/4   H/8 × W/8   H/16 × W/16   H/32 × W/32   H × W

conv, pool, nonlinearity

upsampling

pixelwise output + loss

Upsampling at the last step

forward/inference

backward/learning

pixelwise prediction

segmentation g.t.

96   256   384   384   256   4096   4096   21   21

Feature Map / Filter Number Along Layers

FCN-32s   FCN-16s   FCN-8s   Ground truth

Comparison with different FCNs

FCN-32s result is very rough due to loss of location information while FCN-8s has the best result.

Executed outcome -
Image recognition and object detection

Semantic Segmentation Implementation

Cityscapes Classes

0 void
1 ego_vehicle
2 ground
3 road
4 sidewalk
5 building
6 wall
7 fence
8 pole
9 traffic_light
10 traffic_sign
11 vegetation
12 terrain
13 sky
14 person
15 car
16 truck
17 bus
18 train
19 motorcycle
20 bicycle

DeepScene Classes

0 trail
1 grass
2 vegetation
3 obstacle
4 sky

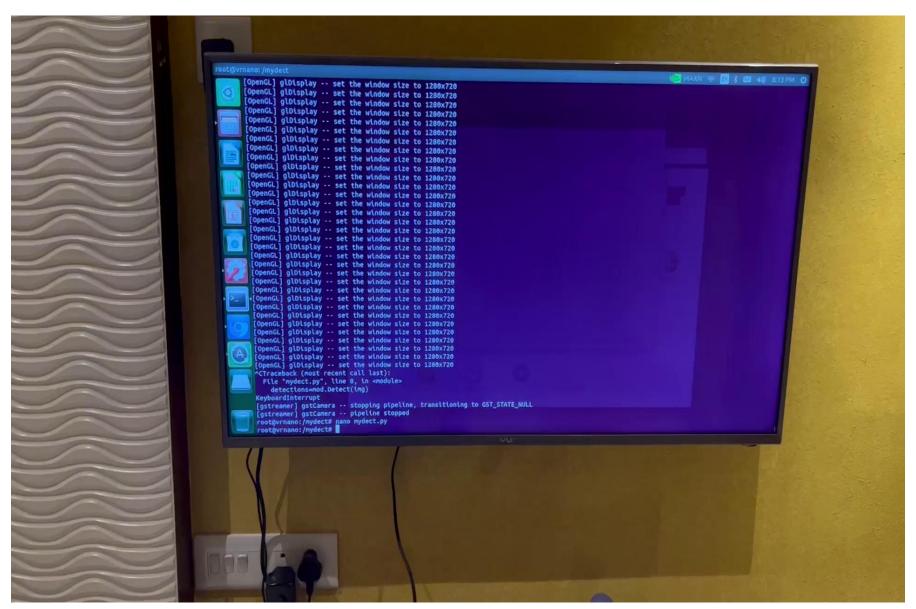# Image Classification and Object Detection

The video attached is an accurate depiction of the of the project. Object detection and Semantic Segmentation is performed and it can be done both real time or statically. Since Jetson NANO's computing power is limited, Complex process like Semantic Segmentation can only reach a certain level of accuracy and it has been achieved by us.

Tabulation

	pixel acc.	mean acc.	mean IU
FCN-32s-fixed	83.0	59.7	45.4
FCN-32s	89.1	73.3	59.4
FCN-16s	90.0	75.7	62.4
FCN-8s	**90.3**	**75.9**	**62.7**

Dataset	Resolution	Accuracy	Frame Rate
Cityscape	512*256	83%	48 FPS
DeepScene	576*320	96%	26 FPS

Accuracy Percentage of various Classes

Classes	Accuracy %
cellphone	95
knive	70
mouse	92
keyboard	84
remote	90
television	88
person	82
bus	99
truck	65
car	80
dog	72
motorcycle	90

Inference

- Image recognition and object detection applications using python by loading pre-trained models is demonstrated.

- Further, Transferred Learning concept is used to train pre-existing models to increase scalability and adapt to native environment.

- Based on in-depth analysis and inferences, developing applications to perform semantic segmentation using image data is demonstrated.

Future Work and Enhancement

- So far this project has achieved autonomous segmented vision solely in the software division. Applying this concept on real environment and testing it's efficiency requires research.

- A 360° view of the surroundings along with accurate calibration is required, involving inputs from the hardware division.

References

- *A Comparative Study of Real-Time Semantic Segmentation for Autonomous Driving*, The 14th IEEE Embedded Vision Workshop, CVPR 2018 At: Salt Lake City, Utah, USA
  https://www.researchgate.net/publication/324866024_A_Comparative_Study_of_Real-Time_Semantic_Segmentation_for_Autonomous_Driving

- *A review of semantic segmentation using deep neural networks*, by Yanming Guo, Yu Liu, Theodoros Georgiou & Michael S. Lew, International Journal of Multimedia Information Retrieval
  https://link.springer.com/article/10.1007%2Fs13735-017-0141-z

- *NVIDIA GPU CLOUD*

  https://www.youtube.com/playlist?list=PL5B692fm6--up8j7qWID9cluY24vky3Xw

- *Getting Started with AI on Jetson Nano*

  https://courses.nvidia.com/courses/course-v1:DLI+S-RX-02+V2/about

- *NVIDIA Computer Vision*

  https://developer.nvidia.com/computer-vision

- *FCN — Fully Convolutional Network (Semantic Segmentation)*

  https://towardsdatascience.com/review-fcn-semantic-segmentation-eb8c9b50d2d1

Thank You