Halifax, Nova Scotia, Canada

CSCI 6505: Machine Learning

Data Generation Report

Instructor
Prof. Sageev Oore

Submitted By:

Mukund Sharma	B00893013
Mayank Sareen	B00899565
Mihir Sanchaniya	B00814744
Sidharth Mahant	B00899439

Description:

We are using Udacity Self Driving Simulator to gather the training dataset of the self-driving car, to validate the feasibility of the model will be using the autonomous mode within this simulator. In just 10 minutes of using the simulator, we were able to gather more than 10 thousand images and the CSV containing the respective steering angle, throttle, brake, and speed of the car.

This is what the initial dataset looks like:

- For a specific point in time, we get the image for three camera angles center, left, and right.

Figure 1: Center Camera Shot

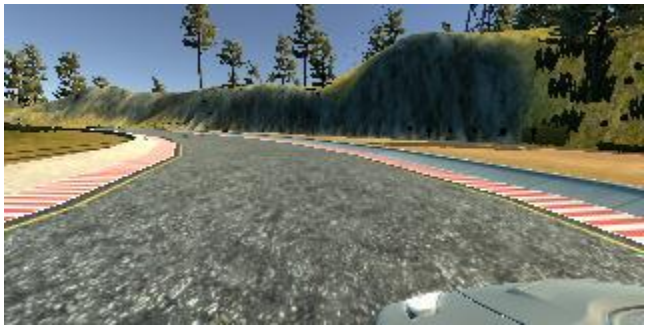


Figure 2: Left Camera Shot

Figure 3: Right Camera Shot

	Center frame	Left Frame	Right Frame	Steering angle	Throttle	Break	Speed
2	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	0	0	8.41E-05
3	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	0	0	7.79E-05
4	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	0	0	7.81E-05
5	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	0	0	7.89E-05
6	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	0	0	8.08E-05
7	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	0	0	7.84E-05
8	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	0.295358	0	0.230116
9	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.09489811	0.654733	0	0.930207
10	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.390594	0.950429	0	1.903474
11	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.6923853	1	0	3.058465
12	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.6260022	1	0	4.204376
13	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.2811006	1	0	5.609113
14	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	1	0	6.754801
15	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	1	0	7.892344
16	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.1534155	1	0	9.015452
17	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.459581	1	0	10.07506
18	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.1697853	1	0	11.17843
19	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	1	0	12.55407
20	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	0	1	0	13.36545
21	D:\Term-2\ML\project\dataset\IMG\center_	D:\Term-2\ML\project\dataset\IMG\left_202	D:\Term-2\ML\project\dataset\IMG\rig	-0.2952306	1	0	14.5899

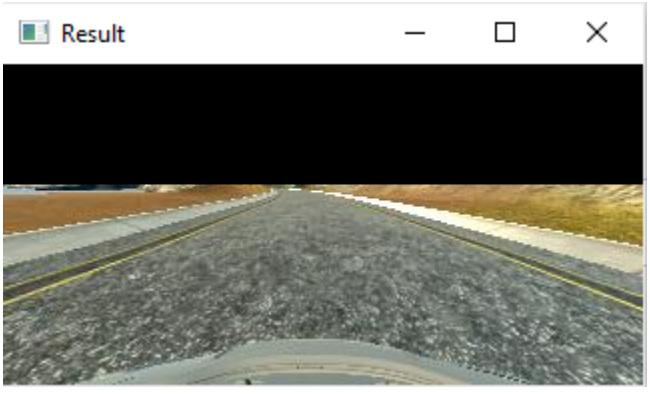
Figure 4: Generated CSV

3 camera shot was initially proposed by Nvidia, which helps in collecting more samples of the same scene and diversifying the dataset. For the purpose of this assignment, we are basically trying to solve a regression problem, all of the above information is nice to have like throttle, brake, and speed but we are focusing on predicting the steering angle which ranges from (-1, +1), wherein -ve denotes left turn angle, +ve means right turn angle and 0 denotes car was running straight.

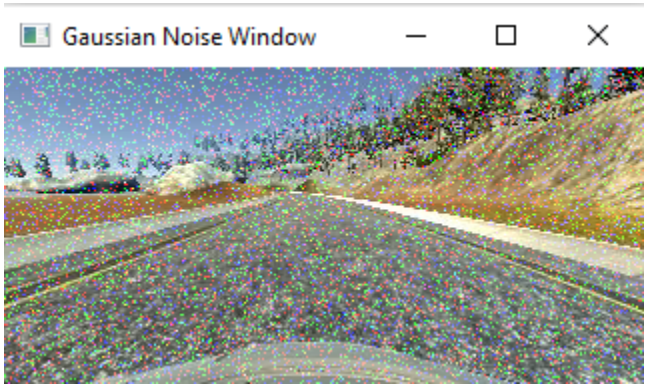
1. Synthetic Dataset Generation

A healthy mixture of natural data and synthetic data has proven to offer great results. However, we will not be testing our model on a real-world dataset, our final model is going to work on a fixed environment that the Udacity simulator provides. To increase the variety of data we are using various data augmentation techniques. For a well-designed synthetic dataset, methods such as Crop, Cutout, ColorJitter, Adding Noise, Filtering, etc are used.

1.1 Crop: As we only need to look at the lines in a frame, we really don't care about the background of the image, the easiest approach is to use OpenCV and make the desired section of the image pixel zero.

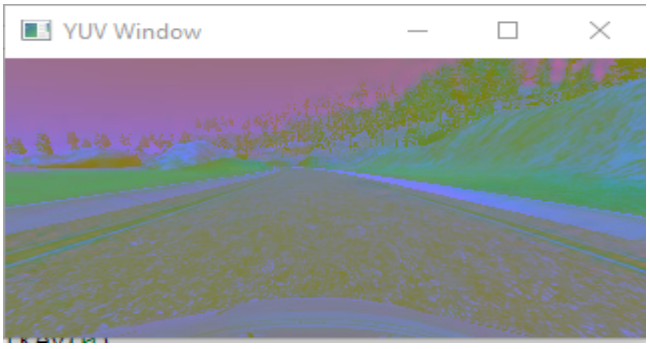


1.2 Adding Noise: Although noise is typically thought of as an unexpected feature in a picture, it can be used for data augmentation in a variety of ways (e.g., Gaussian), and adding noise is relatively simple and beneficial approach to data augmentation in deep learning.



1.3 ColorJitter: It is another simple data augmentation method using which we are randomly changing the brightness, contrast, and saturation of the image.

1.4 YUV Conversion: As Nvidia architecture prefers YUV filter images for training, we are also considering adding YUV filtered images in our training dataset.



1.5 Filtering: Similar to adding noise to an image, filtering is also super easy to implement. We are considering adding one basic type of filtering gaussian (blur).

What are tha parameters you are controlling?

The parameters we are controlling:

- Training data size (Real and Synthetic)
- Number of hidden layers inside different CNN Architectures.
- Activation functions Sigmoid, Tanh, ReLU
- Optimizers, Learning rate, number of Iterations
- Batch size
- Hyperparameter tuning.

How do you get the ground truth?

To train our model, we will be using the combination of real data (generated using the simulator) and the images generated using data augmentation. We are using the **ground truth** that the autonomous mode in the Udacity simulator is driving the car between the lines, i.e. correctly predicting the steering angle which is the goal of this project.

For the initial training, we are planning on using 12,000 images (combining generated data and real data), with 4,000 images each for the center, right, and left angles.

Is this an abstraction of the data described in your original proposal?

In the original proposal also, we were planning on using the Training mode in Udacity simulator and gather a sufficient amount of data, the training mode comes with left turn track only but by simply reversing the car we gathered the right turn track as well. Moreover, the training dataset that we have now is an abstraction of the original proposal.

If this turns out to be too easy, how can you revise it and add complexity to the dataset?

We believe the dataset (Raw + Synthetic) we have for training the model is complex enough for the model to learn and provide a good prediction. With the use of generated data, we have eliminated the risk of overfitting the model. However, we will be carefully monitoring the performance of the model with different sets of training data and make the desired amendments.