Robotic Inference

Reinaldo Ossuna

Abstract—In this project we use the NVIDIA Deep Learning GPU Training System (DIGITS) to retrain several pretained models to classify two different datasets.

Index Terms—Robot, Udacity, DIGITS, NVIDIA, DeepLearning.

1 Introduction

TRAFFIC signs classisification is a challeging problem and there is already many solution proposed in the academicy field [1] [2] [3] [4]. The Germany Traffic Recognition Benchmark (GTSRB) [5] is a large dataset with multi-category classification provided to the academics to study more this field. The goal in this project is to use the already known models, like AlexNet or GoogLeNet to classify tree traffic signs.

2 Background

In the first inference using the supplied dataset,the GoogLenet [6] and the AlexNet [7] was tested with differents optimizers and learning rate decay. In this fase we saw a big difference in both models and how the optimizers and learning decays can interfere in final Inference.

3 Data Acquisition

The dataset was collected using a cellphone camera (Samsung S8 Picture Size 18.5:9, 4032×1960 7.9M) multiple background was used to better augmentation of the images. In average we had 60 images per categorie, this value was not enough for training, for better performace of the model we used some libraries in the python to make some augmentation in the dataset and resize the images to 256×256. In the end we had almost 4500 images per categorie.

Fig. 1. Example of the dataset

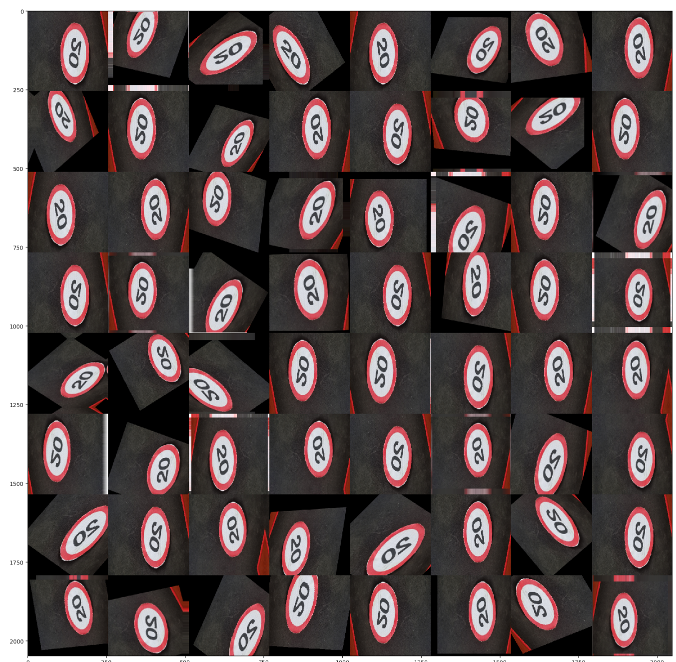


Fig. 2. Example of the aurgemented img

4 Results

given the Table 1 we could observe the model AlexNet with the AdaDelta optimizer and the learning rate decay exponencial with the 0.99 rate has a better Accuracy×Performace. This one was used to classify the Traffic Signs Data set.
The result in the training (Fig:3, Fig:5)was very similar with the provided dataset. A bigger testset is need to provided a better accuracy but all 3 pictures separated to the test have ben right classified.

5 Conclusion

We could see with almost every optimizer the GoogLeNet take more time to training the 10 epochs but it is meet a sable curve a little before. Another point we can put out is the inference time almost in every case the GoogLeNet take more time to infer and this not bring more accuracy to the model. This maybe is related with the size of boths models, GoogLeNet is bigger and more complex Network.

6 FUTURE WORK

- Use the GTSRB dataset instead, this one was many mores classes.
- Detection of the traffic signs and Classification.
- Use the different models, who wad a better perfor-mace in the traffic sign classsification.
- Try to deploy the detection and classification in a car in real time.

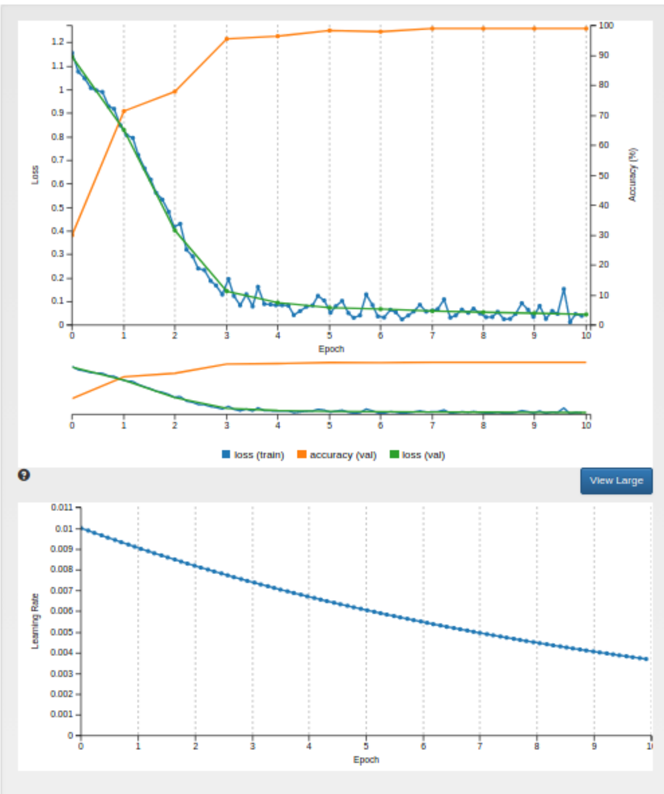
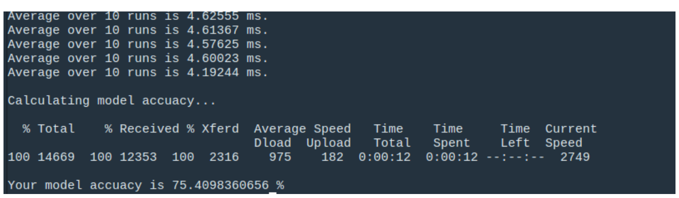


Fig. 3. AlexNet-AdaDelta Training



```
Average over 10 runs is 4.62555 ms.
Average over 10 runs is 4.61367 ms.
Average over 10 runs is 4.57625 ms.
Average over 10 runs is 4.60023 ms.
Average over 10 runs is 4.19244 ms.

Calculating model accuacy...

  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed
100 14669  100 12353  100  2316    975    182  0:00:12  0:00:12 --:--:--  2749

Your model accuacy is 75.4098360656 %
```

Fig. 4. AlexNet-AdaDelta Evalauate

TABLE 1
Models comparative

Model	Optmizer	Decay	Time (10 Epoch)	Epoch steady	Inference Time	Accuracy
GoogLeNet	SGD	Step 33	18	3	7.75	73.77
	Adam	Step 33	6	2	4.1	65.23
	AdaGrad	Step 33	18	3	9.55	75.41
	AdaDelta	Exp .99	19	6	5.33	75.41
AlexNet	SGD	Step 33	6	4	8.55	73.77
	Adam	Step 33	6	1	3.43	67.21
	AdaGrad	Step 33	6	4	5.35	75.41
	AdaDelta	Exp .99	6	7	4.57	75.40

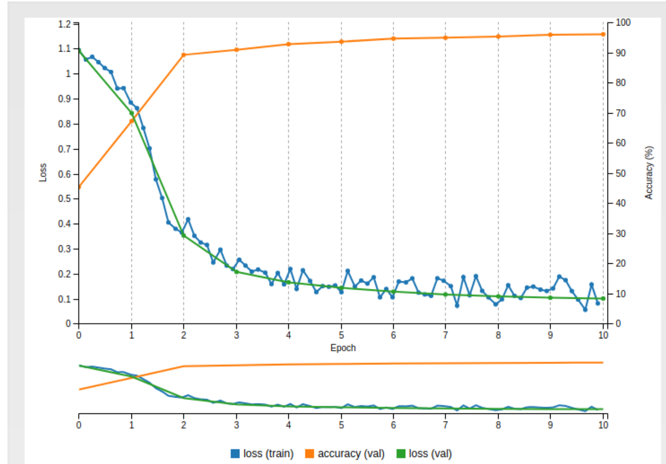


Fig. 5. Traffic Signs AlexNet-AdaDelta Training

REFERENCES

[1] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," *Neural Networks*, vol. 32, pp. 323–332, 2012.
[2] K. Lim, Y. Hong, Y. Choi, and H. Byun, "Real-time traffic sign recognition based on a general purpose GPU and deep-learning," pp. 1–22, 2017.
[3] D. Cireşan, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column deep neural network for traffic sign classification," *Neural Networks*, vol. 32, pp. 333–338, 2012.
[4] A. D. Escalera, L. E. Moreno, and M. A. Salichs, "Road Traffic Sign Detection and Classification," vol. 44, no. 6, pp. 848–859, 1997.
[5] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of Traffic Signs in Real-World Images: The {G}erman {T}raffic {S}ign {D}etection {B}enchmark," in *International Joint Conference on Neural Networks*, no. 1288 in 1, 2013.
[6] N. Iino, W. Takamatsu, A. Iino, Y. Iizuka, and S. Okino, "Going Deeper with Convolutions," 2014.
[7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," 1998.