



Penn



GRASP
LABORATORY

General Robotics, Automation, Sensing & Perception

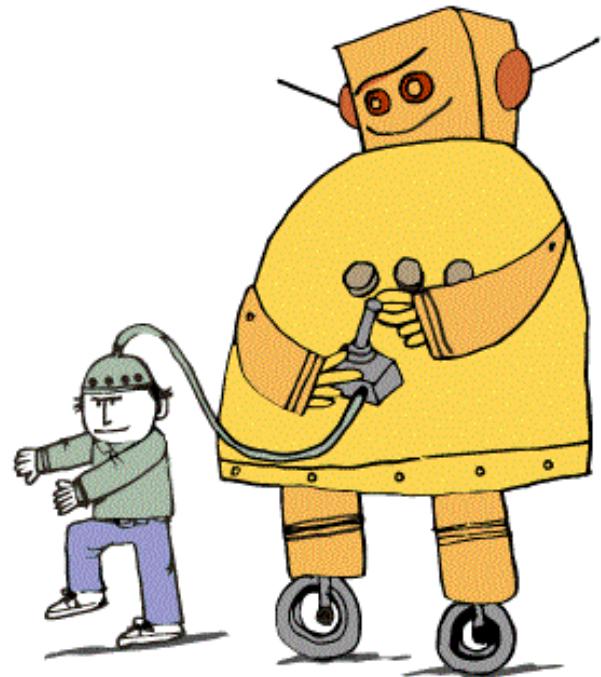
Caffe Framework on the Jetson TK1: Using Deep Learning for Real Time Object Detection

Christopher Alicea-Nieves

Advisor: Camillo J. Taylor, PhD

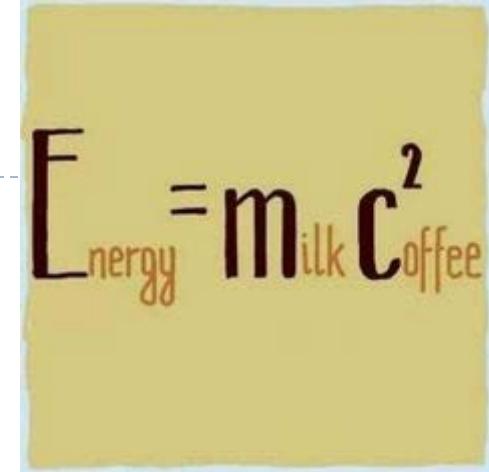
Introduction

- ▶ Machine Learning
 - ▶ Caffe
- ▶ Object Detection
 - ▶ Identify object
 - ▶ Locate object within the frame

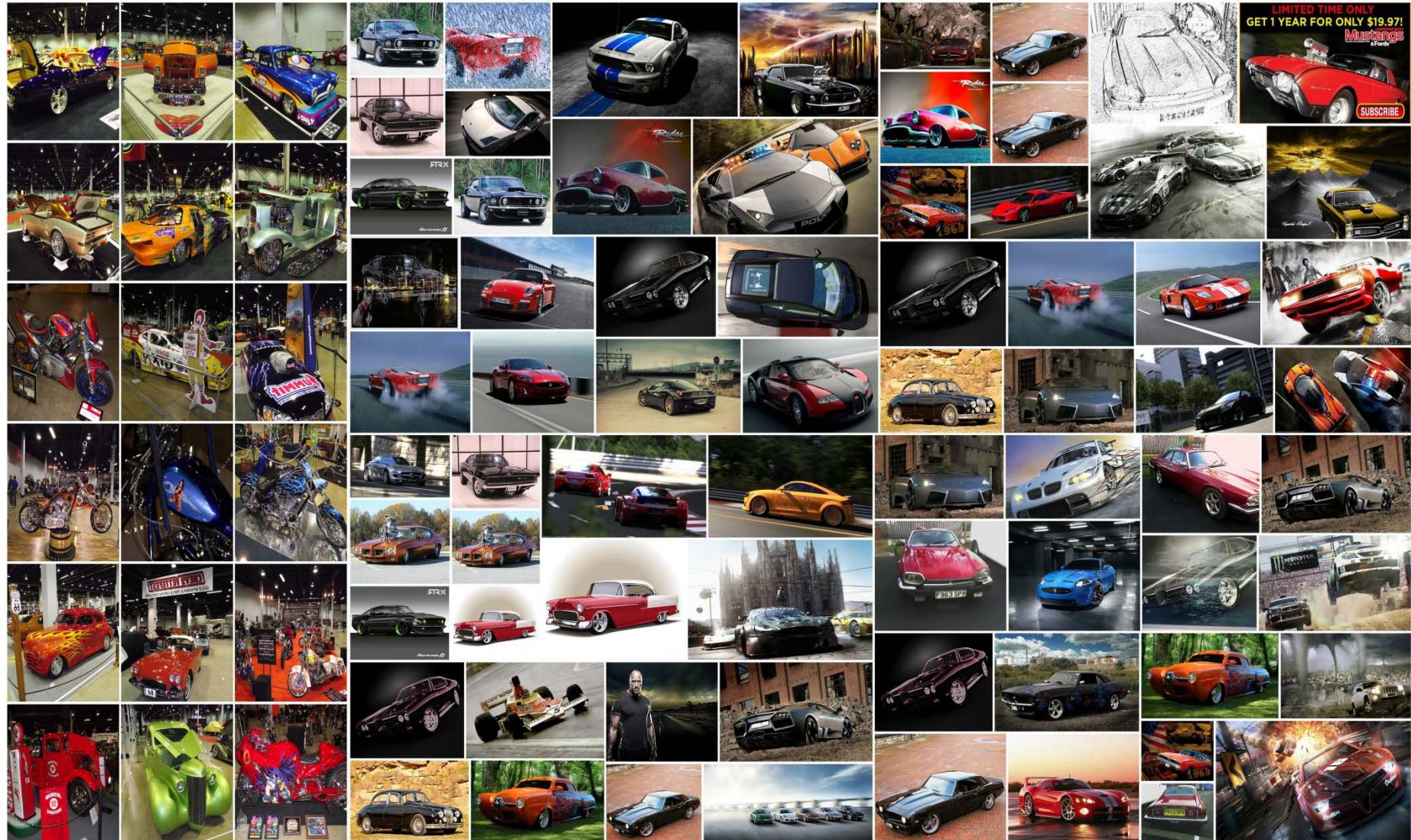


Caffe

- Machine learning framework
 - Developed in UC Berkeley
 - Train and test deep learning algorithms



Caffe

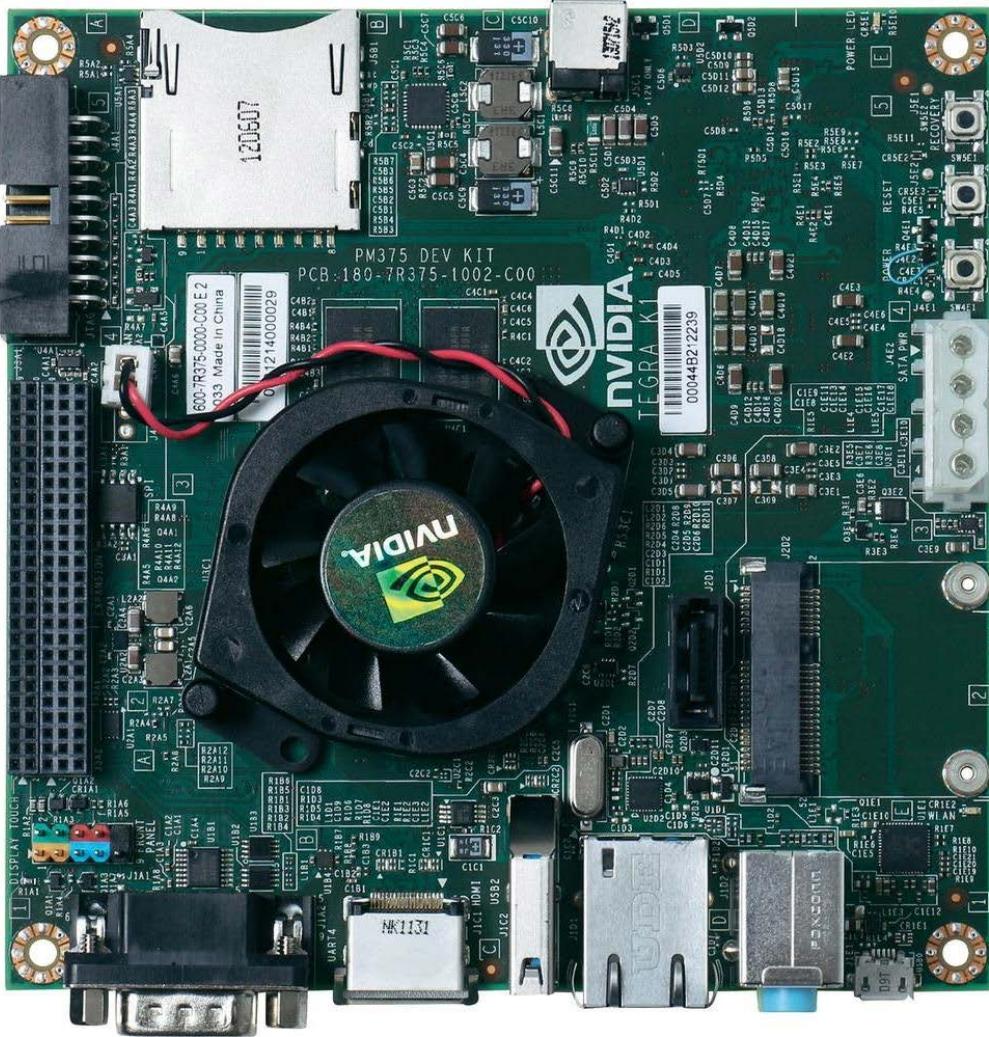


NVIDIA Jetson TK1

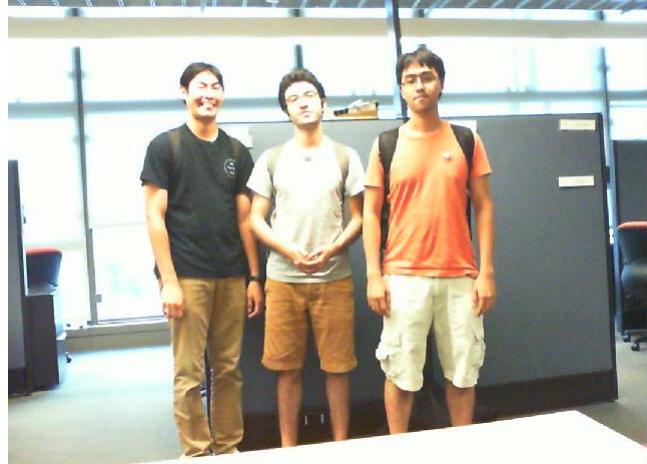
- ▶ Weight – 120g
- ▶ Size – 5" x 5"
- ▶ Power Draw – around 10W
- ▶ On board CUDA-based GPU
- ▶ 2GB of RAM



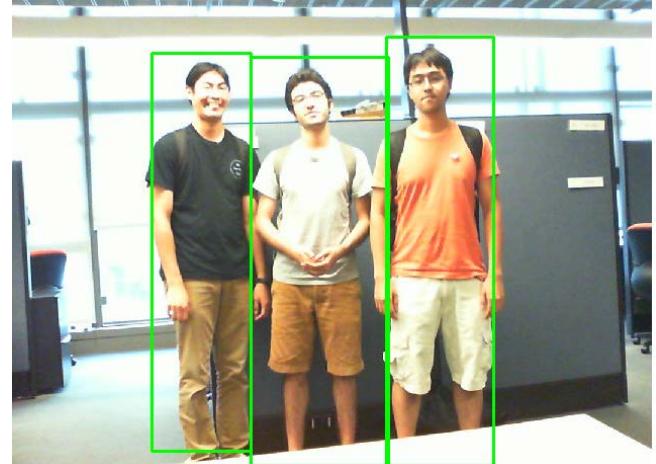
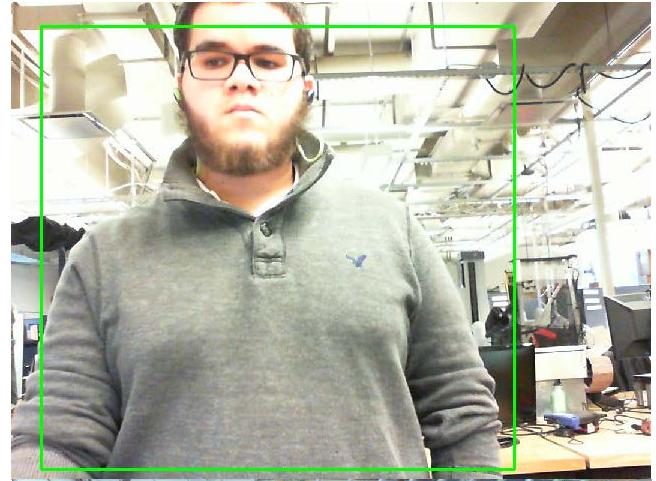
NVIDIA Jetson TK1



Implementation Results



After
Running
Through
A
Convolutional
Neural
Network
using
Caffe



Application Deployment



References

- ▶ Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding.”, arXiv:1408.5093, 2014.
 - ▶ J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, “ImageNet: A Large-Scale Hierarchical Image Database”, IEEE Computer Vision and Pattern Recognition (CVPR), 2009.
 - ▶ Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A., “The Pascal Visual Object Classes (VOC) Challenge”, International journal of computer vision, vol. 88, no 2, p. 303-338, 2010.
 - ▶ N. Dalal, and B. Triggs, "INRIA person dataset, 2011-04-10]", <http://pascal.inrialpes.fr/data/human>, 2005.
 - ▶ A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks”, NIPS, 2012.
 - ▶ R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, CVPR, 2014.
 - ▶ R. Girshick. “Fast R-CNN,” arXiv e-prints, vol. arXiv:1504.08083v1 [cs.CV], 2015.
 - ▶ J. R. Uijlings, K. E. van de Sande, T. Gevers, and A.W. Smeulders, “Selective search for object recognition”, IJCV, 2013.
 - ▶ P. Krähenbühl, and V. Koltun, “Learning to Propose Objects”, CVPR, 2015.
 - ▶ R. J. López-Sastre, T. Tuytelaars, and S. Savarese, “Deformable part models revisited: A performance evaluation for object category pose estimation”, Computer Vision Workshops (ICCV Workshops), IEEE International Conference, 2011.
 - ▶ K. Simonyan, and A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, arXiv preprint arXiv:1409.1556, 2014.
 - ▶ R. Girshick, F. Iandola, T. Darrell, and J. Malik, “Deformable part models are convolutional neural networks”, arXiv preprint arXiv:1409.5403, 2014.
 - ▶ H. O. Song, R. Girshick, S. Zickler, C. Geyer, P. Felzenszwalb, and T. Darrell, “Generalized Sparselet Models for Real-Time Multiclass Object Recognition”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 37(5), 1001-1012, 2015.
 - ▶ N. Dalal, and B. Triggs, “Histograms of oriented gradients for human detection”, IEEE Computer Vision and Pattern Recognition (CVPR), 2005.
-

Questions?

