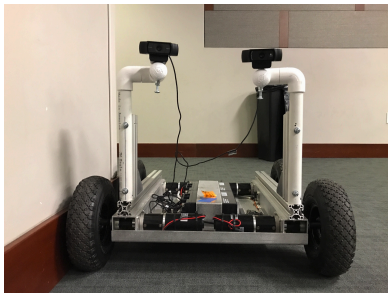
White Paper on Obstacle Detection via a Stereoscopic Camera System, Using Deep Learning

Zhiyang Ong*, Michael Bass†, Khaled Nakhleh‡,
Drupad Khublani§, and Venkata Pydimarri¶
Department of Electrical and Computer Engineering
College of Engineering
Texas A&M University

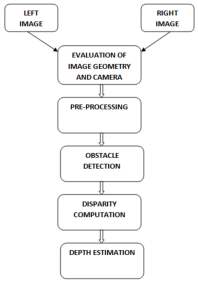
November 28, 2018

1 Project Description

We are enhancing a rover (see Figure 1a) to have stereoscopic vision [4], so that we can carry out data acquisition and binary classification of whether the rover has detected static or dynamic obstacles. To modify the rover, we mounted video cameras and a laptop to sample images from the video cameras, and perform real-time data acquisition by sampling greyscale images from the video cameras. Subsequently, our *Python* program will implement binary classifiers for the the aforementioned tasks. The classifiers will use deep learning [2, 7, 8, 6, 10, 1, 3, 5], specifically an inception-based convolution neural network (CNN) architecture [9]. We plan to enhance the classifer by using stereo matching, disparity map computation, and depth estimation. Figure 1b [4].



(a) Photograph of the enhanced rover



(b) "Flow dragram for obstacle detection" [4]

We aim to perform obstacle detection for a rover traveling at a speed of 20 m/s. We aim to classify static obstacles 30 m away, so that the controller of the rover has time to manipulate the rover away from the obstacles. In addition, we aim to classify moving obstacles at speeds greater than 0.2 m/s.

Fix this description!!!

*Email correspondence to: ongz@acm.org
†Email correspondence to: baklava.akbar@falafel.palestine
‡Email correspondence to: khaled.jkn@gmail.com
§Email correspondence to: dkhublani@tamu.edu
¶Email correspondence to: koustubha26@tamu.edu

2 Work-Breakdown Structure and Project Status

The initial distribution of tasks and project status are described as follows:

1. Michael Bass modified a rover to mount two video cameras and a laptop, so that
2. Khaled Nakhleh develop a *Python script* to sample greyscale images from video cameras mounted on the robotic cyber-physical system.
3. Drupad Khublani and Venkata Pydimarri suggested some techniques, and will examine CNN-based techniques for detection of static and dynamic obstacles.
4. Zhiyang Ong wrote the white paper.

Moving forward, Khaled Nakhleh and Michael Bass will develop *Python* scripts for preprocessing (e.g., decrease the resolution of the images) of the greyscale images from the video cameras, for statistical analysis of the experimental data, and for visualizing experimental data so that we can obtain figures/plots to include in the final report. Drupad Khublani and Venkata Pydimarri will implement the *Python* scripts for deep learning, via a CNN, that will classify the sampled greyscale images into obstacle detected or no obstacle detected. Lastly, Zhiyang Ong would be in charge of writing the final report, with assistance from other team members.

References

[1] Yoshua Bengio, Ian J. Goodfellow, and Aaron Courville. *Deep Learning*. The MIT Press, Cambridge, MA, 2015.

[2] Nikhil Buduma and Nicholas Locascio. *Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms*. O'Reilly Media, Sebastopol, CA, 2017.

[3] Kyunghyun Cho. *Foundations and Advances in Deep Learning*. PhD thesis, Aalto University, Otaniemi, Espoo, Greater Helsinki, Uusimaa, Finland, March 21 2014.

[4] Sai P. Deshmukh and A. M. Deshpande. Detection of obstacles using stereo imaging. *International Journal of Computer Applications*, 131(14):37–42, December 2015.

[5] Ian Goodfellow. *Deep Learning of Representations and its Application to Computer Vision*. PhD thesis, Université de Montréal, Montréal, Québec, Canada, April 2014.

[6] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. The MIT Press, Cambridge, MA, 2016.

[7] Nikhil Ketkar. *Deep Learning with Python: A Hands-on Introduction*. Apress Media, LLC, Berkeley, CA, 2017.

[8] Santanu Pattanayak. *Pro Deep Learning with TensorFlow: A Mathematical Approach to Advanced Artificial Intelligence in Python*. Apress Media, LLC, Berkeley, CA, 2017.

[9] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the* 29^{th} *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, pages 2818–2826, Las Vegas, NV, June 26 – July 1 2016. IEEE Computer Society Press.

[10] Xiaogang Wang. Deep learning in object recognition, detection, and segmentation. *Foundations and Trends in Signal Processing*, 8(4):217–382, 2016.