Obstacle Avoidance for a Teleoperated Robot Through Deep Learning

Gabriel Garcia

Abstract—Deep learning is changing the way that computers see the world, and its application in robotics system is giving great results for different applications. This work proposes using GoogLeNet, a state-of-art classification deep network, to perform classification of images captures using the camera of a mobile robot and then use that information as an obstacle avoidance system. The network was trained with a dataset raised in a laboratory environment and presented satisfactory results.

Index Terms—Mobile robot, obstacle avoidance, deep learning.

1 Introduction

OBSTACLE avoidance is a key problem for mobile robots. Its objective is to allow a robotic system to navigate in an unstructured environment without beating obstacles that could damage the system or the obstacle itself.

This work proposes the use of an inference deep neural network that should classify an image as being obstacle or floor. That network will be used as an obstacle avoidance system in a teleoperated robot dedicated to perform speleological research.

The equipment, shown in Figure 1, which combines innovation with health and safety features, is capable of inspecting underground locations adjacent to existing mining operations, thus reducing employee exposure to risks. The robot is currently in test phase and can detect the threat of falling rocks or the presence of animals, for example. The robot is equipped with a camera and lighting system that affords a real-time view of the underground environment, enabling the operator to conduct simultaneous inspections or surveys for future access, thereby reducing the risks inherent to speleological activities. The equipment operates by remote control and has a battery with three-hour autonomy. The robot has a range of up to 200 meters and, in addition to being water-resistant, can overcome obstacles and traverse rough terrain, fundamental for equipment intended for speleological use.

Although this system has an operator, it is still subject to collisions that can occur from human error controlling it, confusing images or communication issues. Therefore, an obstacle avoidance system can make the robot more reliable and easier to operate.

Fig. 1. Robot in field test.

2 Background

GoogLeNet [1] is a deep convolutional neural network that achieved the state of the art for classification and detection in the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14). The main hallmark of this architecture is the improved utilization of the computing resources inside the network. By a carefully crafted design, it increases the depth and width of the network while keeping the computational budget constant. What makes it ideal to use in a system where we need a small time of inference (less than 10ms to allow on line classification), as in the obstacle avoidance task for our robot.

NVidia Digits provides a pre-trained, easy to set, implementation of GoogLeNet that was used for this project.

3 Data Acquisition

Data is the core of any deep learning system since the network will learn what is important (features) to classify those images from that provided data set. So, it is necessary to create a good dataset, with relevant images taken in different situations and correctly label those.

The first version of our obstacle avoidance system will be developed to prevent collision into a laboratory environment. So, to rise data set, the robot was driven through a laboratory while it was recording images using its embedded camera. Two different set of images were recorded:

- Floor - the robot was driven just in places where it had mainly floor in its front, representing locals where it can navigate.
- Obstacles - to create those samples, the robot was positioned in different locations of the laboratory

with obstacles in its front. Those places are the ones that should be avoided.

In Figure 2 we can see different samples from the created dataset. It is visually easy to determine if the image is formed mainly by floor or by obstacles.

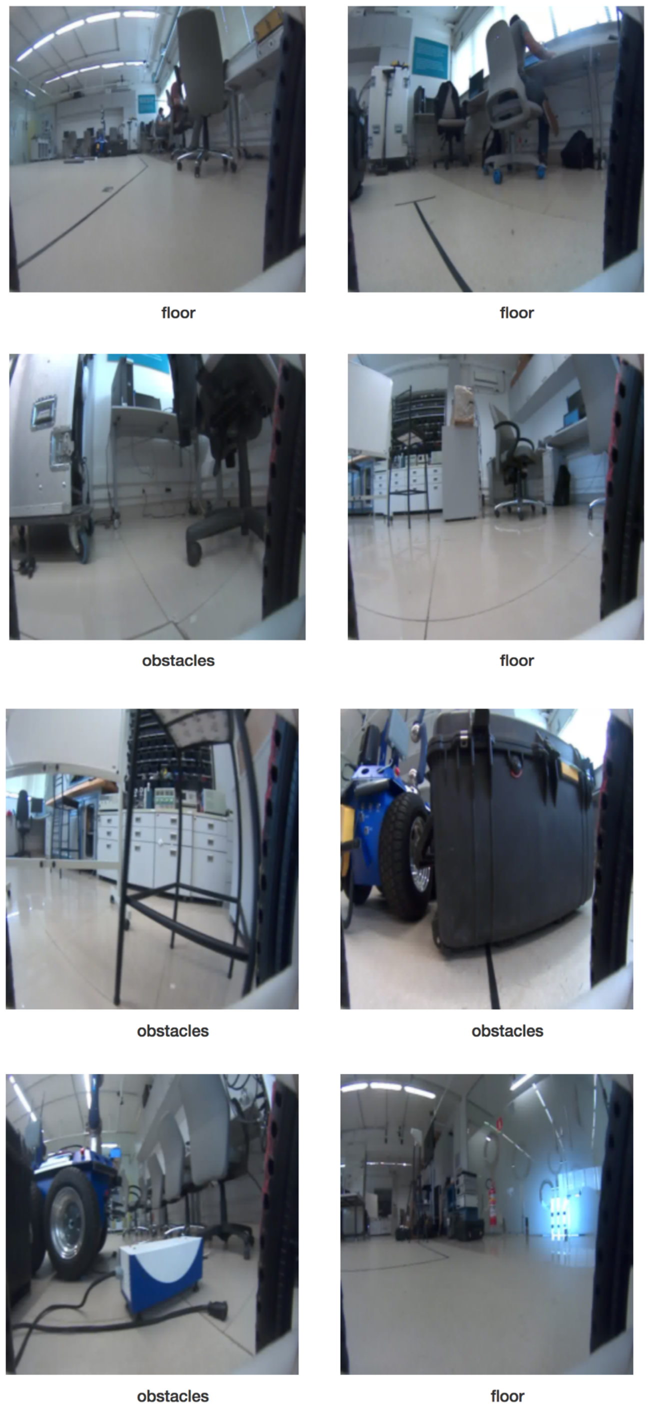

Fig. 2. Samples from dataset.

The raised dataset was preprocessed by Digits and resized to 256 x 256 px, to meet GoogLeNet specifications. The final dataset is formed by 3051 training images and 1016 (25% of the total) validation images. Totalizing 2165 floor images and 1902 obstacle images.

4 Results

The network was trained for 10 epochs with learning rate starting at 0.01 and achieving 0.0001 at the end, as can be seen in Figure 3. Accuracy during training was high, achieving more than 99% accuracy in the 3rd epoch, as shown in Figure 4.

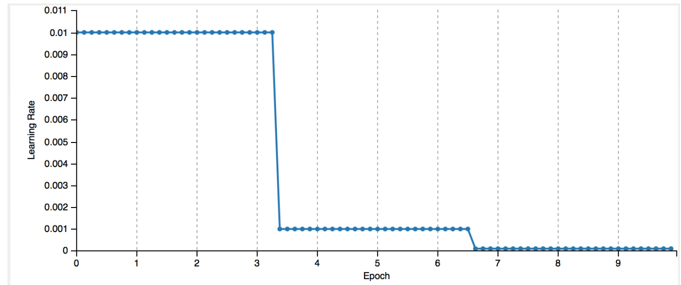

Fig. 3. Learning rate graphic.

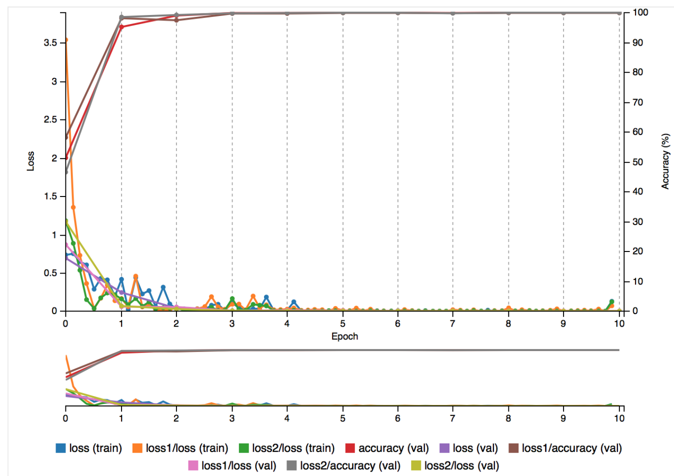

Fig. 4. Accuracy graphic.

Some images were left out of training and validation sets, to perform tests. Figure 5 presents two images, one with floor and another with obstacles, being correctly classified.

For the supplied data, GoogLeNet was trained for 15 epochs with a learning rate of 0.01. The network achieved an accuracy of 78% in the test set, and it is able to classify an image in average 4.8 ms. Therefore, the network meets required specifications (accuracy higher than 75% and inference time lower than 10 ms).

5 Discussion

This network will be applied to a teleoperated system to avoid undesired collisions, therefore, a short time for inference is needed, also a high accuracy classifying obstacles is necessary. Regarding misclassified images, it is better to get false positives obstacles classifications than false negatives floor classifications, since a false negative floor classification can incur in a collision. Summarizing, the network should provide the higher accuracy possible for a short time of inference that should be calculated in the future, based on the robot velocity. The threshold of our system will be the limit for inference time.

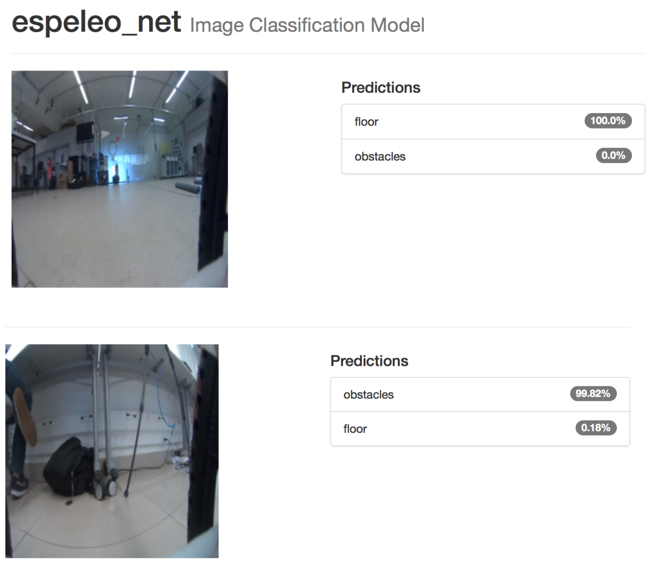

Fig. 5. Tests performed with two images.

For the test dataset, all images were correctly classified. It means that more images should be taken to test the system in different scenarios.

6 Future work

The robotic system is under development, and a prototype version is already performing different field tests. The are not robots focused in cave exploration available in the market, as far as we know, therefore, this project is really feasible as a commercial product. With this successfully proof of concept for obstacle avoidance using deep network, future work should be focused on raising a big dataset with different scenarios and implement that network inside the system that runs the robot.

References

[1] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Computer Vision and Pattern Recognition (CVPR)*, 2015.