# Bachelor in Computer Science
-
# Application for Pedestrian Detection with Intel RealSense Camera Technologies (Short-Version)

**Saturday 11th January, 2020**

Sean Achtatou
University of Luxembourg
Email: sean.achtatou.001@student.uni.lu

Francois Robinet
University of Luxembourg
Email: francois.robinet@uni.lu

## 1. Introduction

This research work, for a semester project in the Bachelor in Computer Science in the University of Luxembourg, is about the study of computer vision. Indeed, the goal of this project is to establish an association between the two following technologies:

- YOLO (application applying convolutional neural networks)
- Intel RealSense Camera (depth camera)

By this association, we would be capable to produce an application, that would possibly be mounted on an autonomous car, having the capability to detect pedestrian(s) on the roads and calculate their distance from the camera.

## 2. Application for Pedestrian Detection with Intel RealSense Camera

### 2.1. Domains

The domain of this project concerns one of the many interesting world of the computer science: *The Artificial Intelligence*. With the evolution of the Artificial Intelligence, and computer power, we are managing to provide application more and more effectively usable for computer vision. Indeed, for this project we are focusing on the study of the computer vision to detect pedestrian(s) on the roads. To achieve this, we would use one of the Deep Learning algorithms, principally used for computer vision, which is the Convolutional Neural Networks.

### 2.2. Objectives

The main objective of this project is to provide a way to be aware, that the association of the technology Intel RealSense Camera, and the YOLO application – based on the use of the Convolutional Neural Networks – would have the possibility to be introduced and usable on an autonomous car, to detect pedestrian(s) on the roads.

As conclusion to this project, we should be convinced that the final application we would have implemented – associating YOLO and Intel RealSense -, could be used and trusted to be introduced into the domain of autonomous vehicles.

## 3. Constraints

The first of the constraints to take into consideration is to have acquired knowledge about Artificial Intelligence, especially into Deep Learning algorithms. By using the Convolutional Neural Networks, we should be aware what type of architecture have been implemented, and how it is processing the images.

The second constraints are consequently knowledge about image processing and image composition. To know how an image is being processed, we need to know how they are structured (i.e. pixels, channels, colors).

The last constraint is the capability to implement multiples applications based on the reuse of YOLO application and the Intel RealSense Camera. All the applications – such as the YOLO application - are being implemented on Python 3. It is

a simple language quite simple to be used and learned.

# 4. Background

# 5. Scientific

## 5.1. YOLO Application

The YOLO application is based on the use of the Convolutional Neural Networks to detect multiples objects into images. This application, allow us to obtain multiples information for each detected objects, such as :

- Position on the image (with a green rectangle surrounding the objects)
- Type of objects having been detected (i.e. person, toothbrush, table, etc.)
- Probability of certitude of the detected objects(i.e. 30% to 99.99%)

For our project, we are reusing the YOLO application to simply detect the pedestrian(s) on the frames.

## 5.2. Intel RealSense Camera

The Intel RealSense Camera is a technology allowing us to obtain a recording of frames with additional features (i.e. color frame, depth frame). The depth frame is a special frame obtained by the use of mounted laser on depth camera, and providing the distance of each different pixel in the environment. Thanks to this information, we would be capable to reuse the depth frame to calculate the distance of the pedestrian(s) from the camera.

# 6. Technical

## 6.1. Image Processing

The image processing is an implemented application applying the re-use of the YOLO application. This application would accept as input pictures of type .jpg/.png, passing them in the Convolutional Neural Networks, and return an output result representing all the detected pedestrian(s) on the picture.

## 6.2. Video Processing

The video processing application is the improved version of the *Image processing*. The application would accept a video, a pass each frame of it into the Convolutional Neural Networks. For each return frame, they would be regrouped in a video, saved, and would be viewable.

## 6.3. Live Processing

The live processing application is the main goal of this project, associating the YOLO application and the Intel RealSense Camera. The application would accept the streaming of a connected USB depth camera (i.e. Intel RealSense Camera). For each set of frame captures (color frame, depth frame). The color would be passed into the Convolutional Neural Networks to detect the pedestrian(s). Whereas, the depth frame would be used to calculate the distance of all the detected pedestrian on the color frame.

# 7. Results

## 7.1. YOLO Application

For the YOLO application, we verified its performance and efficiency to detect the pedestrian(s) on a frame. We have introduced frames of a pedestrian having different poses and executing different actions. After analysis, we have been satisfied that the application is being efficient under certain conditions (i.e. the pedestrian is not detected by executing specific poses) and could be used associated to the Intel RealSense Camera.

## 7.2. Intel RealSense Camera

For the Intel RealSense Camera, we have verified what type of frames we would obtain by streaming the camera. We verified that the color frame and the depth frame were corresponding to each other for a set of stream, and which type of structure the depth frame was made of. We observed that the depth frame, would be great-fully useful to calculate the distance of the pedestrian(s) on the color frames.

## 7.3. Application Detecting Pedestrian with Intel RealSense Camera

The application for the project we had to provide is an association of the YOLO application and the Intel RealSense Camera. We implemented both of them together to obtain an application capable to detect pedestrian(s) on live streaming, and display the distance of the pedestrian(s) detected. The final application produced is well satisfying, and observable to be used on an autonomous car.

## 8. Conclusion

In conclusion of this project, we are assured that the association of the technology Intel RealSense Camera and the YOLO application are applicable in the society of autonomous technology, taking into consideration that their performance are still limited by the time, nevertheless would be rapidly solvable by an improved re-training of the YOLO application.