

Chapter 16

Design of a Real-Time-Integrated System Based on Stereovision and YOLOv5 to Detect Objects

Oumayma Rachidi

ENSAM, Moulay Ismail University of Meknes, Morocco

Ed-Dahmani Chafik

ENSAM, Moulay Ismail University of Meknes, Morocco

Badr Bououlid

ENSAM, Moulay Ismail University of Meknes, Morocco

ABSTRACT

Real-time object detection represents a major part in the development of advanced driver assistance systems (ADAS). Pedestrian detection has become one of the most important tasks in the field of object detection due to the increasing number of road accidents. This study concerns the design and implementation of a Raspberry Pi 4-based embedded stereovision system to detect 80 object classes including persons and estimate 3D distance for traffic safety. Stereo camera calibration and deep learning algorithms are discussed. The study shows the system's design and a custom stereo camera designed and built using 3D printer as well as the implementation of YOLOv5s in the Raspberry Pi 4. The object detector is trained on the context object detection task (COCO) 2020 dataset and was tested using one of the two cameras. The Raspberry Pi displays a live video including bounding boxes and the number of frames per second (FPS).

I INTRODUCTION

Advanced Driver Assistance System (ADAS) is a vehicle-based intelligent safety system aimed towards improving safety in the automobile industry. ADAS technologies are used in many vehicles today, some of which are integrated as a standard equipment. These systems can improve road safety in terms of collision avoidance, protection and post-crash notification. Intelligent speed adaptation (ISA), electronic stability control (ESC), and autonomous emergency braking systems (AEB), represent a good example of ADAS systems that are offering a significant safety potential (European Commission, 2018). The main purpose of ADAS is to reduce road accidents by providing drivers information about objects in front of the vehicle, to take the necessary actions. Object detection represents a vital tool widely used in computer vision to estimate the location of objects in images or videos for ADAS. There have been many studies in the field of computer vision regarding object detection for real time applications. These studies differ in terms of methodology; however, the objective remains the same: find a detector with high accuracy and high speed.

Deep learning, is a machine learning subset that offers solutions in many complex applications, through deep neural networks. These networks, have improved the performance of smart surveillance, smart city, and self-driving cars-based applications in comparison with machine learning methods. One of the first deep neural networks is the CNN (Convolutional Neural Network), which is used for image classification. This network is aimed to recognize many objects such as vehicles, pedestrians, and traffic signs. The main advantage of CNN, is that it automatically extracts features after training on a dataset, without human intervention (Shin et al., 2016).

Nowadays, deep learning models are used in object detection for real time applications. These models can use different architectures to classify and detect objects, and can be divided into two main categories: Two-stage detectors such as R-CNN and Fast-RCNN (Girshick, 2015), and one stage detectors such as the single Shot MultiBox Detector (SSD) (Liu et al., 2016), and the You Only Look Once (YOLO) (Redmon et al., 2015). In the first category, regions of interest are firstly generated, and then fed to a network to apply classification and bounding boxes regression. These detectors are highly accurate, but take time to process images which makes it harder to deploy them in real time applications. In the second category, the object classification and bounding boxes regression are performed using one network. These models have lower accuracy, but achieve high inference speed.

The main goal of a real time object detection application is to find a good trade-off between the accuracy and the speed of the detector. For YOLO detectors, many improvements have been introduced to reach the best performances. Redmon et al. (2015) have created the first version of YOLO in 2015. In this version, a neural architecture based on GoogleNet model was used. The model was compared to Faster-RCNN on PASCAL VOC 2007 dataset, the results shown that the VGG-16 version of Faster R-CNN was 10 mAP higher, but 6 times slower than YOLO. Few years after the release of the first version, YOLOv2 or YOLO9000 was introduced to detect over 9000 categories with high accuracy and speed compared to YOLO. YOLO suffers from a significant number of localization errors and has a low recall in comparison with region proposed-based methods. Thus YOLOv2, was designed to improve localization errors and recall, while maintaining classification accuracy. In this version, Darknet-19 model was used, and the model was trained on ImageNet for classification and both COCO and VOC for detection. On PASCAL VOC 2007, YOLO gets 63,4 Map and 45 FPS on a GeForce GTX Titan X with a resolution of 448x448, while YOLOv2 gets 77,8 Map and 59 FPS with a resolution of 480x480 (Redmon & Farhadi, 2017). In 2018, YOLOv3 was introduced, with some changes in the architecture to make it better than

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/design-of-a-real-time-integrated-system-based-on-stereovision-and-yolov5-to-detect-objects/337464

Related Content

An Opcode-Based Malware Detection Model Using Supervised Learning Algorithms

Om Prakash Samantray and Satya Narayan Tripathy (2021). *International Journal of Information Security and Privacy* (pp. 18-30).

www.irma-international.org/article/an-opcode-based-malware-detection-model-using-supervised-learning-algorithms/289818

Inducing Six-Word Stories From Curated Text Sets to Anticipate Cyberwar in 4IR

Shalin Hai-Jew (2018). *Handbook of Research on Information and Cyber Security in the Fourth Industrial Revolution* (pp. 406-477).

www.irma-international.org/chapter/inducing-six-word-stories-from-curated-text-sets-to-anticipate-cyberwar-in-4ir/206792

Socio-Cultural and Multi-Disciplinary Perceptions of Risk

Yölande Goodwin and Kenneth David Strang (2012). *International Journal of Risk and Contingency Management* (pp. 1-11).

www.irma-international.org/article/socio-cultural-multi-disciplinary-perceptions/65728

Contemporary Financial Risk Management Perceptions and Practices of Small-Sized Chinese Businesses

Simon S. Gao, Serge Orealand Jane Zhang (2014). *International Journal of Risk and Contingency Management* (pp. 31-42).

www.irma-international.org/article/contemporary-financial-risk-management-perceptions-and-practices-of-small-sized-chinese-businesses/115817

Spatio-Temporal Hot Spot Analysis of Epidemic Diseases Using Geographic Information System for Improved Healthcare

Uma V. and Jayanthi Ganapathy (2019). *Emerging Trends and Innovations in Privacy and Health Information Management* (pp. 31-63).

www.irma-international.org/chapter/spatio-temporal-hot-spot-analysis-of-epidemic-diseases-using-geographic-information-system-for-improved-healthcare/228338