Multi-Object tracking and Trajectory Prediction for Autonomous Vehicles

Vinit Awale
AI & Robotics Technology Park
ARTPARK, IISC Bangalore
Karnataka, India
awale.vinit@gmail.com

Naveen Arulselvan
AI & Robotics Technology Park
ARTPARK, IISC Bangalore
Karnataka, India

***Abstract*—This document specifies the implementation details of the perception module using cameras of a self-driving car. This project was completed as a part of the Summer Internship program at ARTPARK, IISC Bangalore, under Prof. Naveen Arulselvan. The perception module consists of multi-object detection, tracking and trajectory prediction. The multi-object detection is based on the *YOLO v5* algorithm [?]. We have used the *Deep Sort* algorithm for multi-object tracking based on the paper "Simple online and realtime tracking with a deep association metric". The trajectory prediction module is implemented using *PEC Net* based on the paper "It is not the journey but the destination: Endpoint conditioned trajectory prediction". Python and PyTorch framework have been used for the code implementation.**

***Index Terms*—Object detection, multi-object tracking, trajectory prediction, YOLOv5, Deep Sort, PEC Net, Autonomous vehicles**

I. Introduction

Perception is a central problem for any autonomous agent, be it humans, robots or self-driving vehicles. This module helps for a smoother and more reliable control of the car using the path-planning module of the autonomous agent. It can also aid in pose estimation. For our project, we have included the following sub-modules for the perception:

- Multi-object detection using the *YOLOv5* algorithm.
- Multi-object tracking using the *Deep Sort* algorithm.
- Trajectory prediction using the *PEC Net* algorithm.

Object detection in the context of autonomous driving refers to detecting the objects present in the scene (making use of the camera sensors on the autonomous vehicle) by making bounding boxes surrounding the detected objects. This is followed by identifying the class of the objects. The family of YOLO (You Only Look Once) models are the most popular object detection models for autonomous driving. The YOLOv5 model is a state-of-the-art object detection model that is capable of detecting 80 classes of objects. The model is trained on the MS COCO dataset, containing over 1.2 million images and over 20,000 bounding boxes for the 80 classes of objects. Multi-object tracking refers to the problem of tracking the objects detected across frames. For this project, we are implementing the *Deep Sort* algorithm for tracking the bounding boxes. Simple Online Realtime Tracking *SORT* is an approach of multi-object tracking using simple and effective algorithms such as Kalman Filter. Including an association metric (using deep learning) for the detected object across frames leads to a more robust and accurate multi-object tracking called Deep Sort.

The problem of trajectory prediction (as is already clear from the name) involves predicting the agents detected by the YOLOv5 model. PEC Net uses an encoder-decoder architecture for predicting the agents detected by the YOLOv5 model. The encoder is a neural network that encodes the input image into a vector of fixed size. The decoder is a fully connected neural network that decodes the vector into a high-dimensional vector. For the implementation of PEC Net, we need an aerial view of the scene. However, in our implementation, we are working with datasets of camera images taken in the ego-centric view. Hence for a complete end-to-end perception pipeline, we need to move the detections of our object detection module to birds-eye view. This is a part of future work.

The code for the project is available at

II. Background and Previous Work

III. Datasets

IV. Procedure, Experiments and Results

V. Limitations

VI. Future work

VII. Acknowledgment

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

VIII. References

Please number citations consecutively within brackets [**?**]. The sentence punctuation follows the bracket [**?**]. Refer simply to the reference number, as in [**?**]—do not use "Ref. [**?**]" or "reference [**?**]" except at the beginning of a sentence: "Reference [**?**] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was

cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [**?**]. Papers that have been accepted for publication should be cited as "in press" [**?**]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [**?**].

REFERENCES

[1]

[2] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[3] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[4] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[5] K. Elissa, "Title of paper if known," unpublished.

[6] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[7] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[8] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.