Review Report 1

Semantic Visual SLAM using deep learning for dynamic scenes

Zhang Xiao Ya , Abdul Hadi Abd Rahman , Faizan Qamar

Scope of research:

Manuscript entitled “Semantic Visual SLAM using Deep Learning for Dynamic Scenes”, addresses the one of common issues involved with vision SLAM approach. Authors proposed an approach for improving the accuracy of traditional SLAM approach by eliminating the dynamic feature point, which one of the main cause in the error accumulation. In the proposed method well known semantic segmentation approach Deeplab V3+ used to identify the dynamic object in the scene. And after identifying, all the feature points of dynamic object removed and only static feature considered for pose estimation. And experimental analysis conducted on benchmark dataset to prove the significance of proposed method.

General Comment:

1. Manuscript written well with good editing standards and sentence ambiguity, grammar error relatively rare. Although study appears to be interesting in automation filed, there are already similar work has been discussed in the published articles. There is nothing new in the objective and research finding highlighted in the manuscript. For reference, please check following few article. Moreover it is well-known issue that, visual and lidar odometry algorithms performance negatively impacted by dynamic object in the scene and there are various approach introduced to successfully address this issue.

A Dynamic Scene Vision SLAM Method Incorporating Object Detection and Object Characterization

An Adaptive ORB-SLAM3 System for Outdoor Dynamic Environments

But it is not clear that, how proposed approach different from existing similar approach ?

2. There are few similar approach found in the literature used different object detection models to eliminate the feature points of dynamic objects. In this manuscript authors choose Deeplab V3+ detection module for identifying the dynamic object in the scene. But it is necessary to discuss, clear motivation and significant advantages found by authors by using Deeplab V3+ model. Why not other model ?

3. Core contribution of this proposed work is relies on, how efficiently proposed approach able to differentiate dynamic feature points from static feature points. More specifically, Section Motion consistency detection discuss in detail. But it is necessary to present some result images depicting the actual process. With typical object detection module used for motion detection and proposed motion consistency algorithm out on same scene.

4. More importantly, as authors mentioned in the abstract proposed approach outperforms traditional visual odometry in terms of accuracy and robustness. It is very necessary to provide the evidence for robustness with proper time compassion analysis.

Also, proposed method should be compared with existing approach which addressed similar issues in the literature. Only comparing with typical approach may not be appropriate.

5. Highlight the specific contributions more clearly at the end of introduction section.

Specific Comments:

1. Figure 2, represents the Deeplabv3+ network architecture, showing input image with cat. Though it is meaningful illustration to describe the proposed SLAM approach. I would strongly recommend modify with the scene (instead cat), more commonly appear in the indoor SLAM.

2. In the Motion consistency detection section, Line# 303, refers Fig 6 mentioning motion consistency detection algorithm illustration. But actual image is number is Fig 7. Please check image numbers before referencing.

3. Maintain the consistency in using phrase, it is used that **Motion consistency** at description, but image subtitle written as **Moving consistency**.

4. Check Fig8, Image subtitles (d), (e) and (f) repeated.

5. Check Cap and Small letter consistency of Subsection heading entire manuscript. It is randomly used.

Extended **c**onvolution and **d**eep **s**eparable **c**onvolution

Effect of **S**emantic **A**nalysis

Motion **c**onsistency **d**etection

Feature **m**atching

Trajectory and **P**ose **E**rrors

6. Table 4, Rotational drift analysis presents observed rotational error. But units mentioned as (m), Suggest authors to mention the unit appropriately.