ABSTRACT

The development of autonomous vehicle technology has driven research in the field of localization and mapping, where the use of sensors such as cameras and Light Detection and Ranging (LiDAR) is becoming increasingly common. The combination of data from these two sensors provides richer information for building accurate three-dimensional environment maps. However, the main challenge is to improve localization accuracy, especially under complex and dynamic conditions. Deep learning methods based on Convolutional Neural Networks (CNN) have been widely used to process sensor data, but there is still room for improvement in improving localization accuracy.

This research aims to improve localization accuracy by modifying the CNN architecture used to process camera and LiDAR data. The Karlsruhe Institute of Technology and Toyota Technological Institute (KITTI) Odometry dataset is used as the basis to test the performance of the model. In this study, several modifications were made to the original CNN, including the addition of Feature Pyramid Network (FPN), Batch Normalization (BatchNorm), and freeze layer. The addition of the FPN architecture helps the model capture features at various scales, while BatchNorm plays a role in accelerating convergence and improving network stability. In addition, the freeze layer is used to prevent learning on some parts of the model so as to reduce overfitting.

The experimental results show that the original CNN obtained an accuracy of 88.09%. After the addition of BatchNorm, the accuracy increased to 88.97%. The addition of FPN gave the best result with an accuracy of 89.58%, while the modification with freeze layer resulted in an accuracy of 86.96%. From these results, it can be concluded that modification of CNN architecture, especially with FPN, can significantly improve the accuracy of camera and LiDAR data-based localization in 3D environments.

Keywords: BatchNorm, Camera, CNN, Depth Map, FPN, KITTI Odometry, LiDAR, Localization, 3D Point Cloud.