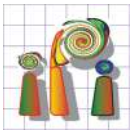

Master Thesis

RGB-Depth Fusion in Panoptic Segmentation

Gottfried Wilhelm Leibniz Universität
Institute for Photogrammetry and Geoinformation

Winter Semester 2024-25

Submitted by:
Arnav Sarkar
Matrikel: 10043042

First Examiner:
apl. Prof. Dr. techn. Franz Rottensteiner
Second Examiner:
Dr. M. Mehltretter

Supervisor:
M.Sc. T. Nguyen

Statement

I declare that this thesis is the result of independent research conducted by me under the guidance of my supervisors. It does not contain the results of any other scientific research that has been published or written by any other individuals or groups, except for those already cited in the thesis. Furthermore, I state that this work in the same or a similar form has not been submitted to an examination authority.

Hannover, 19.03.2025

Place, Date

Signature

Abstract

Depth Aware Panoptic Segmentation addresses the limitations of conventional panoptic segmentation by incorporating depth information to better distinguish between visually similar but spatially separated instances. This thesis investigates novel strategies for effectively integrating depth information into panoptic segmentation frameworks, focusing on enhancing the Feature Pyramid Network (FPN) backbone and exploring various fusion mechanisms. Building upon the Depth aware PanopticFCN as our baseline architecture, we introduce a Feature Aligned Pyramid Network (FaPN) that improves multi-scale feature extraction through better alignment of features in the top-down pathway. This enhancement leads to more refined segmentation outputs, with particularly notable improvements in stuff classes. Additionally, we implement a depth-aware version of coordinate convolutions (CoordConv) that encodes both spatial and depth information, providing a more comprehensive representation of object positioning in three dimensional space. We further explore different fusion strategies, comparing mid-fusion approaches in the bottom-up pathway with fusion in the top-down pathway. Our experiments reveal that while mid-fusion in the bottom-up pathway led to misalignments, fusion in the top-down pathway after proper independent feature alignment yielded better results. In particular, the integration of Adaptive Channel Weighting (ACW) for dynamic feature fusion proved effective in enhancing segmentation accuracy. Experiments conducted on the Cityscapes dataset demonstrate that our proposed enhancements improve panoptic segmentation quality, particularly for stuff classes, while the original Depth-aware PanopticFCN showed stronger improvements for thing classes. The findings contribute to advancing depth-aware segmentation techniques, providing valuable insights for future research in three dimensional scene understanding.

Keywords: Depth Aware Panoptic Segmentation, RGB-D Fusion, Feature Aligned Pyramid Network, Depth-Aware Coordinate Convolution, Adaptive Channel Weighting