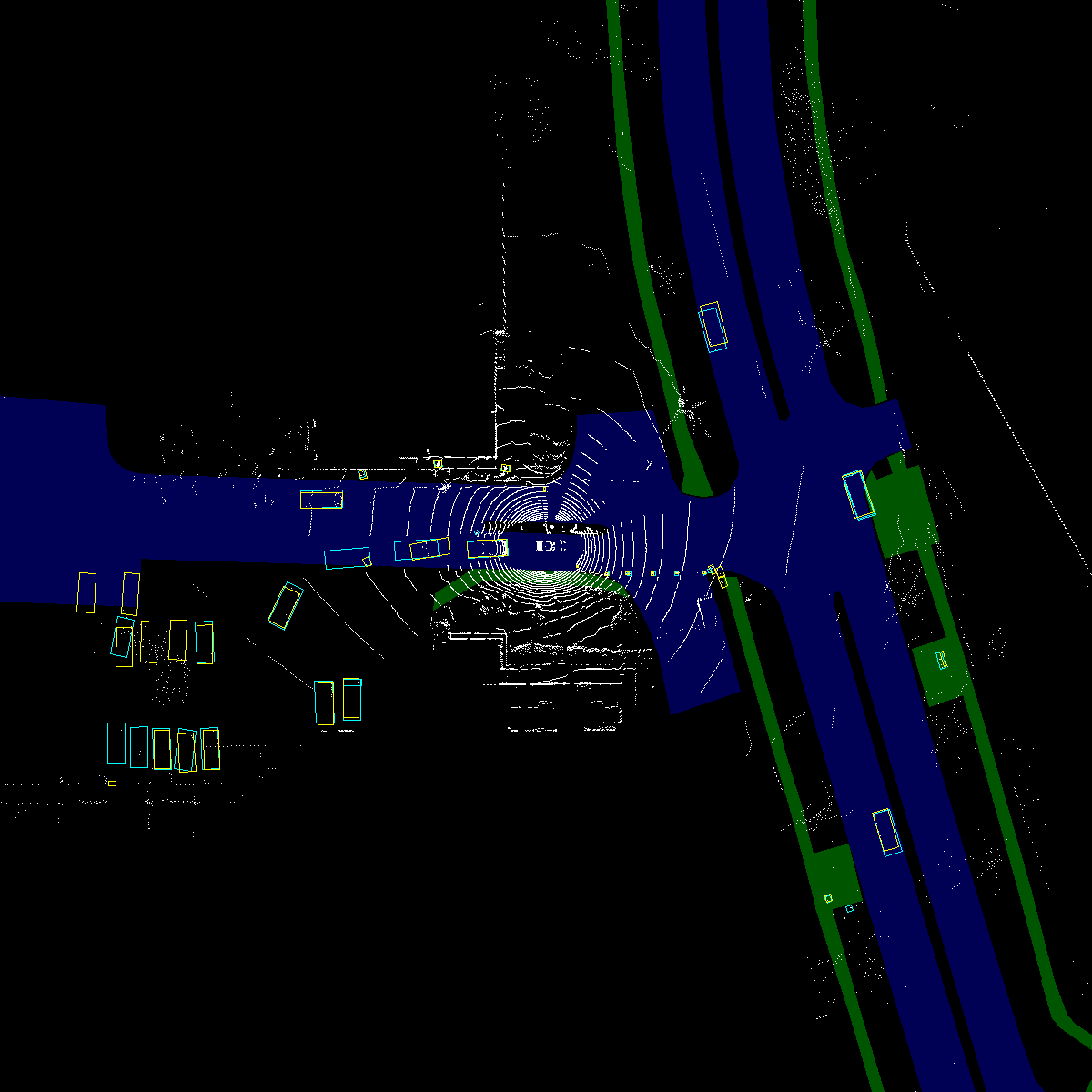
# MapFusion: A General Framework for 3D Object Detection with HDMaps

## Abstract

3D object detection is a crucial perception component in autonomous driving. Recent methods predominantly use Lidar sensors alone or fused with cameras. High Definition Maps (HDMaps), essential for intelligent vehicles, have been underutilized in enhancing object detection tasks. This study introduces MapFusion, a straightforward yet potent framework, to incorporate map data into existing 3D object detector pipelines. The framework features a FeatureAgg module for HDMap feature extraction and fusion, plus a MapSeg module as an auxiliary segmentation head. We integrated MapFusion with three different baseline detectors on a large public autonomous driving dataset, demonstrating significant improvements in mean Average Precision (mAP) across the board.

## I. Introduction

With the rapid advancements in autonomous driving (AD), the need for robust and accurate 3D object detection techniques has become paramount. Traditional camera-based detectors face challenges in accurately capturing 3D distances, making LiDAR sensors essential for precise localization. Despite significant progress in LiDAR-based 3D object detection using point-based and voxel-based approaches, integration of supplementary data sources like HDMaps could further refine detection accuracy.



*Caption: "Detection results comparison" | Explanation: "Figure 1 compares detection results between PointPillars alone and PointPillars with MapFusion. The integration of MapFusion considerably reduces false positives, marked by red circles." | Ref: PDF p.1*

The proposed MapFusion framework offers a unique approach by integrating HDMap data, vastly relegated to post-processing roles, directly into the detection pipeline. Through the FeatureAgg module and MapSeg module, our framework advances the fusion of map features into the detection process, showcasing improved performance in reducing false positives and negatives.

## II. Related Work

Advancements in range sensors and AD technologies have spurred numerous 3D object detection methods. Traditional approaches projected 3D point clouds into 2D representations for processing, but modern methods like Voxelnet, SECOND, and PointPillars leverage deep learning for more efficient 3D object detection directly from point cloud data. However, integrating additional data sources, such as HDMaps, remains less explored despite their potential to significantly enhance detection robustness and accuracy.

## III. Proposed Approach

### A. Overview

MapFusion is strategically designed to be detector-independent, integrating seamlessly into existing frameworks with minimal adjustments. The core components, the FeatureAgg and MapSeg modules, highlight our novel approach in fusing HDMap information with conventional 3D detection methods to enhance overall performance.

[Image missing: page2\_img1.jpg] Caption: "Framework overview" | Explanation: "Figure 2 illustrates the MapFusion architecture. The red dotted box depicts the standard object detection components, while the blue dotted box represents the novel map feature extraction components." | Ref: PDF p.2

The detailed network structures for both the 2D Feature Extractor and the MapSeg module are designed to optimize the integration of spatial and semantic information from HDMaps, providing a comprehensive solution that adapts to various detection tasks.

### B. HDMap Representation

HDMaps offer detailed geometric and topological road data that is crucial for precise object detection. By employing a raster representation of HDMaps centered around the ego vehicle, specific elements like drivable areas and walkways are converted into a format conducive for deep learning applications. This representation is then efficiently processed through our FeatureAgg module, demonstrating a significant impact on detection capabilities.

### C. FeatureAgg Module

This module fuses extracted map features with voxel features using straightforward concatenation, followed by an optional convolutional layer to refine the feature integration, leading to notable performance improvements in detection tasks.

## IV. Experimental Results

The integration of MapFusion with standard detectors such as SECOND, PointPillars, and CenterPoint was evaluated on the nuScenes dataset. Improvements in nuScenes detection score (NDS) and mAP were observed across all baseline methods. Detailed results underscore the effectiveness of MapFusion in enhancing detection precision, particularly in urban driving scenarios.

\*\*Table 1. Evaluation results of MapFusion on nuScenes validation dataset\*\*

| Method | NDS (%) | mAP (%) |  
|:-------|---------:|--------:|  
| SECOND w/o MF | 60.80 | 49.62 |  
| SECOND w/ MF | \*\*62.04\*\* | \*\*50.89\*\* |  
| PointPillars w/o MF | 57.45 | 43.87 |  
| PointPillars w/ MF | \*\*58.95\*\* | \*\*46.66\*\* |  
| CenterPoint w/o MF | 67.13 | 59.43 |  
| CenterPoint w/ MF | \*\*67.97\*\* | \*\*60.61\*\* |

## V. Conclusion

MapFusion represents a significant advancement in the field of autonomous driving by effectively integrating HDMap data into 3D object detection workflows. The framework not only enhances detection accuracy but also paves the way for future research in multi-sensor fusion. The potential to extend this framework to incorporate data from other sensors, such as cameras and radar, offers promising avenues for comprehensive environmental perception and interpretation.

## References

1. E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," IEEE Access, vol. 8, pp. 58443-58469, 2020.  
2. E. Arnold et al., "A survey on 3D object detection methods for autonomous driving applications," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 10, pp. 3782-3795, 2019.  
3. S. Ren et al., "Faster R-CNN: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91-99.