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

## Abstract

3D object detection is a critical perception component in autonomous driving. Traditional approaches often rely solely on Lidar sensors or a fusion with cameras. However, High Definition Maps (HDMaps), a fundamental infrastructure for intelligent vehicles, remain underexploited. This paper introduces a novel framework, MapFusion, which incorporates map information into current 3D object detection pipelines. We develop a FeatureAgg module for HD Map feature extraction and fusion, and a MapSeg module as an auxiliary segmentation head for the detection backbone. MapFusion is detector independent and demonstrates significant improvement in detection performance on a large-scale public autonomous driving dataset.

[Image missing: page1\_img1.jpg] Caption: "Comparison of detection results" | Explanation: "Part (a) shows detection results from PointPillars with red cycles marking false positives, while part (b) shows improved detection with MapFusion reducing false positives." | Ref: PDF p.1

## Introduction

Autonomous driving (AD) technology has garnered significant attention in recent years but continues to face considerable challenges, particularly in perception tasks essential for interpretation and response mechanisms within AD systems. Conventional 2D image-based detectors struggle with accurate 3D object localization due to the lack of depth information, prompting the need for 3D sensors like LiDAR. This paper exploits the potential of HDMaps to enhance the performance of 3D object detection frameworks significantly.

## Methodology

### Architecture Overview

We propose the MapFusion framework designed to integrate HDMap features into existing 3D object detection systems. The framework consists of two key components:

1. \*\*FeatureAgg module\*\* for extracting and fusing features from the HDMaps.  
2. \*\*MapSeg module\*\* as a segmentation head to complement the primary detection backbone by refining the map-based feature inputs.

[Image missing: page2\_img1.jpg] Caption: "Framework overview of MapFusion" | Explanation: "Illustrates the integration of standard 3D object detection and map feature extraction blocks. Detailed mapping between LiDAR point clouds and HDMap data optimizes detection accuracy." | Ref: PDF p.2

### Algorithms and Equations

The core functionality is supported by advanced fusion algorithms that merge input from the LiDAR and map data, using techniques such as 3D Sparse Convolution for feature representation and concatenation operations for feature aggregation:

$$L\_{FeatureAgg} = Concate(L\_{LiDAR}, L\_{Map}) ag{1}$$

### Training Details

The system is trained on the nuScenes dataset, employing data augmentation techniques like random rotation and scaling to simulate various driving conditions, ensuring robustness across different scenarios.

## Experiments & Results

### Datasets and Metrics

We evaluate MapFusion using the nuScenes dataset, focusing on its ability to improve detection metrics such as the nuScenes detection score (NDS) and mean Average Precision (mAP).

### Results

Our framework advances the state-of-the-art for several baseline detectors, showing significant enhancements in detection accuracy and precision.

\*\*Table 1. Performance comparison on nuScenes dataset\*\*

| Detector | 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\*\* |

### Ablation Studies

Investigations into the individual components like FeatureAgg and MapSeg modules validate their effectiveness, with \_FeatureAgg\_ showing slightly higher impact on performance improvements.

## Conclusion

MapFusion showcases a promising avenue for enhancing 3D object detection in autonomous driving by effectively integrating HDMap information. Future work will explore expanding this approach to include other sensor modalities such as radar and camera inputs.