

# 多模態感測融合之 3D 邊框回歸精度及其環境強健性敏感度分析

專題編號：114-2-CSIE-S010

執行期限：114 年第 1 學期至 114 年第 2 學期

指導教授：陳彥霖

專題參與人員：112820023 王承惠

112820045 曾怡瑄

112590058 林宣邑

111590035 李彥勳

## 一、摘要

本專題以自動駕駛場景中的 3D 物件偵測為研究主題，探討 camera-only 與 camera+LiDAR 多模態融合模型在不同影像退化條件下的偵測表現與環境強健性。研究採用 BEVFusion 架構[1]與 nuScenes-mini 資料集[2]進行實驗，並以 mAP、NDS、mATE、mASE、mAOE 等 nuScenes detection metrics 作為評估依據。實驗結果顯示，在乾淨資料下 camera+LiDAR 模型明顯優於 camera-only 模型；在低亮度與高斯雜訊等 camera degradation 條件下，camera-only 模型性能大幅下降，而 camera+LiDAR 模型仍能維持較穩定的偵測能力，顯示 LiDAR 幾何資訊能有效補償影像品質下降所造成的影響。

## 二、緣由與目的

3D 物件偵測是自動駕駛感知系統的重要任務，模型需在不同天候、光照與感測器品質變化下維持穩定表現。nuScenes 提供多感測器自動駕駛資料，可用於 3D detection 評估[2]；相機影像能提供豐富語意資訊，但容易受到低光照、雜訊、模糊等因素影響；LiDAR 則能提供較穩定的空間幾何資訊。本專題目的在於比較 camera-only 與 camera+LiDAR fusion 模型在 camera corruption 下的表現差異，分析多模態融合對 3D 邊框回歸精度與環境強健性的影響。

## 三、研究範圍與目標

本研究以 nuScenes-mini 作為初步實驗資料集，使用 BEVFusion 官方模型進行 clean baseline 與 corruption evaluation。目標建立可重現的 BEVFusion 實驗環境、完成 nuScenes-mini 資料前處理、取得 clean baseline 結果、設計 camera degradation 實驗，並比較 camera-only 與 camera+LiDAR 在不同退化條件下的 mAP、NDS 與 3D box regression error 變化。

## 四、使用技術與方法

本專題使用 Python、PyTorch、MMDetection3D [3]、CUDA extension 與 BEVFusion 進行模型推論與評估，資料集則採用 nuScenes-mini，模型包含 camera-only detector 與 camera+LiDAR fusion detector，影像退化方法則參考 RoboBEV 對 BEV perception robustness 的 corruption 評估概念[4]，包含 Gaussian noise、defocus blur 與低亮度 dark corruption，如圖 1 所示。評估指標採用 nuScenes detection metrics，其中 mAP 與 NDS 用於衡量整體偵測能力，mATE、mASE、mAOE、mAVE、mAAE 分別衡量位置、尺寸、方向、速度與屬性誤差。



圖 1. 不同 camera corruption 條件下之 nuScenes 影像

## 五、研究架構與流程

研究流程分為五個階段：

### (一)第一階段：

建立 Colab GPU 實驗環境，安裝 PyTorch、MMCV、MMDetection3D 與 BEVFusion 相關套件

### (二)第二階段：

完成 nuScenes-mini 資料轉換與 info.pkl 建立

### (三)第三階段：

執行 clean baseline evaluation，取得 camera-only 與 camera+LiDAR 的基準表現

### (四)第四階段：

產生不同 camera corruption 資料集，並重新執行 evaluation

### (五)第五階段：

整理 mAP、NDS 與各項 regression error，分析不同模型在 camera degradation 下的敏感度差異

## 六、實驗結果

實驗結果如表 2 與表 3 所示。Clean baseline 下，camera+LiDAR 在 mAP 與 NDS 皆優於 camera-only；在 Gaussian

noise s5 與 Dark s3/s5 條件下，camera-only 明顯下降，而 camera+LiDAR 仍維持較穩定表現。

表 2. 模型於不同 camera corruption 條件下之 mAP 比較

| 條件                | Camera-only | Camera+LiDAR |
|-------------------|-------------|--------------|
| Clean             | 0.3349      | 0.5746       |
| Gaussian noise s5 | 0.0249      | 0.5276       |
| Dark s3           | 0.1373      | 0.5501       |
| Dark s5           | 0.0000      | 0.5167       |

表 3. 模型於不同 camera corruption 條件下之 NDS 比較

| 條件                | Camera-only | Camera+LiDAR |
|-------------------|-------------|--------------|
| Clean             | 0.3828      | 0.5803       |
| Gaussian noise s5 | 0.0385      | 0.5583       |
| Dark s3           | 0.2064      | 0.5690       |
| Dark s5           | 0.0000      | 0.5547       |

## 七、結論

本專題完成 BEVFusion 於 nuScenes-mini 的環境建置、資料前處理、baseline evaluation 與 camera corruption robustness analysis。實驗結果顯示，多模態融合模型在乾淨資料與影像退化條件下皆優於 camera-only 模型，尤其在低亮度與高斯雜訊等 camera degradation 下，camera-only 模型出現大幅性能下降，而 camera+LiDAR 模型仍能維持穩定偵測能力。此結果說明 LiDAR 幾何資訊能有效補償相機影像品質下降，提升 3D 物件偵測系統的環境強健性。

## 八、参考文献

- [1] Z. Liu, H. Tang, A. Amini, X. Yang, H. Mao, D. Rus, S. Han, “BEVFusion: Multi-Task Multi-Sensor Fusion with Unified Bird's-Eye View Representation,” *IEEE International Conference on Robotics and Automation*, pp. 2774-2781, 2023.
- [2] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, O. Beijbom, “nuScenes: A multimodal dataset for autonomous driving,” *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11621-11631, 2020.
- [3] K. Chen, J. Wang, J. Pang, Y. Cao, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Xu, Z. Zhang, D. Cheng, C. Zhu, T. Cheng, Q. Zhao, B. Li, X. Lu, R. Zhu, Y. Wu, J. Dai, and H. Wang, “MMDetection: Open MMLab Detection Toolbox and Benchmark,” *arXiv preprint*, arXiv:1906.07155, 2019.
- [4] S. Xie, L. Kong, W. Zhang, J. Ren, L. Pan, K. Chen, Z. Liu, “Benchmarking and Improving Bird's Eye View Perception Robustness in Autonomous Driving,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 47, No. 5, pp. 3878-3894, 2025.