

小型 LiDAR 語意分割之後處理優化方法研究

專題編號：114-1-CSIE-S010

執行期限：113 年第 1 學期至 114 年第 1 學期

指導教授：陳彥霖

專題參與人員：

111590452 李晨維

111590453 張竣歲

111590454 洪俊璋

111590455 曾詮淳

一、摘要

本研究聚焦於 3D LiDAR 語意分割的後處理優化，目標是在有限運算資源或邊緣裝置上，於不重新訓練模型的前提下，提升模型輸出結果的一致性與精度。我們選用 RandLA-Net 作為基礎模型，其僅需 1.24M 參數，遠低於多數動輒超過 50M 參數的分割模型，並能在 RTX 2080 Ti (11GB VRAM) 上流暢執行，展現了邊緣部署的可行性。在預測結果上，本研究實作 k 近鄰平滑 (kNN smoothing) 方法，針對低信心點進行鄰域投票修正。實驗結果顯示，kNN 平滑能有效降低孤立錯誤點，並使 mIoU 從 53.9% 提升至 56.4%，證明其為輕量化語意分割系統的有效增強手段。

二、緣由與目的

隨著自駕車與智慧城市的發展，3D LiDAR 語意分割成為環境感知的重要技術。然而，當前許多高效能模型往往仰賴龐大的參數量與運算資源，動輒需要 50M 以上的參數與高階 GPU，這對邊緣裝置與資源受限的部署環境並不實際。

為了解決此問題，本研究採用 RandLA-Net 作為基礎模型，其僅需 1.24M 參數，在維持相對競爭力表現的同時，大幅降低計算與記憶體需求。我們在 RTX 2080 Ti (11GB VRAM) 上進行實驗，展現了其在中等資源環境下的可行性，並嘗試透過後處理優化進一步提升分割效能。具體而言，我們引入 kNN 平滑方法，在不重新訓練模型的前提下，修正預

測結果中的低信心點，增強輸出的一致性與精度。

三、研究範圍與目標

本研究的範圍聚焦於 3D LiDAR 語意分割的後處理優化，主要針對如 RandLA-Net 的輕量化模型。研究目標為，在不重新訓練模型、不增加大量運算負擔的前提下，提升模型在有限運算資源或邊緣裝置上輸出的準確度與一致性。具體目標是透過實作 k 近鄰平滑 (kNN smoothing) 演算法，修正預測結果中的低信心點，並量化其對平均交並比 (mIoU) 指標的提升效果。

四、研究方法

本研究採用 RandLA-Net 作為基礎模型，這是一個僅需 1.24M 參數的輕量級網路，並使用 SemanticKITTI 資料集的官方預訓練權重。實驗在 NVIDIA RTX 2080 Ti (11GB VRAM) 環境下進行，模擬中等資源的硬體限制。

核心方法為 k 近鄰平滑 (kNN smoothing) 後處理演算法，旨在修正模型對低信心點的誤判。具體流程如下：

核心想法：模型輸出中，最大預測機率低於特定閾值 τ 的點（稱為低信心點）較容易被誤判。透過其 k 個鄰近點的標籤來進行投票，可以有效修正這些錯誤。

執行步驟：

首先取得模型對每個點的預測機率分佈與最終分類結果。

設定一個信心閾值 τ （實驗中使用

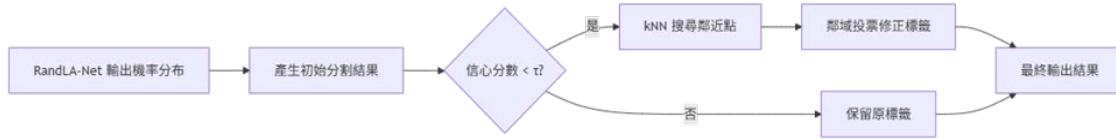


圖 1. 方法流程圖

0.55)，篩選出所有信心分數低於 τ 的點。針對這些低信心點，使用 kNN 演算法搜尋其空間上最鄰近的 k 個點 (實驗中使用 $k=8$)。

對這 k 個鄰近點的標籤進行多數投票，並將投票結果覆蓋該低信心點原本的預測標籤，完成修正。

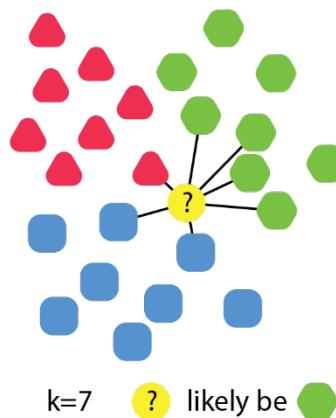


圖 2. kNN 示意圖

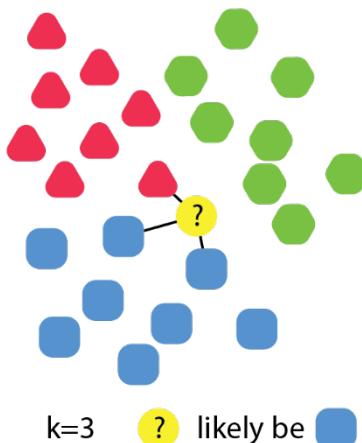


圖 3. kNN 示意圖

五、研究成果

本研究在 SemanticKITTI 資料集的序列 08 測試集上進行了實驗，比較了

基準模型 (Baseline，即未經任何後處理的 RandLA-Net) 與加上 kNN 平滑處理後 ($\tau=0.55$, $k=8$) 的成果。

實驗結果顯示，kNN 平滑演算法能有效提升模型輸出的整體一致性。如表 1 所示，平均交並比 (mIoU) 從基準的 53.9% 提升至 56.4%，增加了 2.5 個百分點。這證明了此後處理方法為一個輕量且有效的增強手段，能在不增加大量資源消耗的前提下，顯著降低孤立的錯誤預測點。

方法	mIoU(%)
Baseline (RandLA-Net)	53.9
+ kNN 平滑	56.4

表 4. 後處理對 mIoU 的影響

六、結論

本研究成功驗證了 kNN 平滑演算法在有限資源環境下，作為一種後處理技術的應用價值。

優點：此演算法設計簡單，計算成本低，可直接嵌入現有的處理流程 (pipeline) 中。[2] 由於其資源友善的特性，在 RTX 2080 Ti (11GB VRAM) 等中低階 GPU 或邊緣裝置上也能流暢執行。

挑戰：方法的成效依賴於參數 τ (信心閾值) 與 k (鄰近點數量) 的設定，若平滑程度過高，可能導致微小物件的細節遺失。

擴展性限制：研究中曾嘗試將此架構擴展至多 GPU 環境以加速流程，但發現原有的 TensorFlow 1.5 基礎架構難以支援。這顯示此方法的擴展需仰賴更現代化的深度學習框架 (如 TensorFlow 2.0 以上版本) 作為基礎。

七、參考資料

- [1] Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., & Gall, J. (2019). SemanticKITTI: A dataset for semantic scene understanding of LiDAR sequences. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).
- [2] Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Gall, J., & Stachniss, C. (2021). *Towards 3D LiDAR-based semantic scene understanding of 3D point cloud sequences: The SemanticKITTI Dataset*. *The International Journal of Robotics Research*, 40(8–9), 959–967. <https://doi.org/10.1177/02783649211006735>
- [3] Q. Hu, B. Yang, L. Xie, S. Rosa, Y. Guo, Z. Wang, N. Trigoni, and A. Markham, “RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
- [4] R. Li, S. Li, X. Chen, T. Ma, J. Gall, and J. Liang, “TFNet: Exploiting Temporal Cues for Fast and Accurate LiDAR Semantic Segmentation,” *arXiv preprint arXiv:2309.07849v3*, Apr. 2024.
- [5] A. Milioto, I. Vizzo, J. Behley, and C. Stachniss, “RangeNet++: Fast and Accurate LiDAR Semantic Segmentation,” *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019. , <https://github.com/PRBonn/lidar-bonnetal>