

國立台北大學資訊工程學系專題報告

基於視覺與導航路徑資訊融合的縱向定位優化：應用於車輛路徑生成

Longitudinal Localization Optimization via Visual and Navigation Route Information Fusion for Vehicle Route Generation

專題組員：許凱翔、林柏均、史博勳、劉冠佑、蔡守恩

專題編號：PRJ-NTPUCSIE-115-004

執行期間：114 年 9 月至 115 年 6 月

Abstract

隨著自動駕駛與擴增實境 (AR) 技術的發展，將導航路線以第一人稱視角視覺化投影至駕駛者畫面，能有效降低認知負荷。然而，現有最佳技術架構複雜且運算量龐大，難以在一般車載晶片上即時運行；而低成本端到端模型在複雜交叉路口又缺乏全局導航路徑指引。為解決此問題，本專題提出一套低成本路線生成系統 (LLOV)，基於地圖拓撲先驗進行顯式幾何對齊。為了有效處理參考架構 OLRA 的局部最佳解與縱向誤差等問題，本系統在初始化階段引入擴展卡爾曼濾波 (EKF) 來平滑定位擾動，並設計跨邊緣排斥損失 (Cross-Edge Loss) 以有效抑制導航路徑不合理穿透實體路線的行為。定量與質性實驗結果表明，本系統在維持低硬體算力的條件下 (於 CPU 上單幀平均推論時間為 33.0 ms, 約 30 FPS)，能將全局平均歐氏誤差收斂至 0.79 m，相較於著名的低成本端到端模型 OpenPilot，顯著提升了中遠距離軌跡預測的幾何精準度與路徑指向的穩定性。

1. 簡介

隨著自動駕駛與擴增實境 (AR) 技術的快速發展，行車路線的生成與視覺化能為駕駛帶來直觀的導航輔助與未來位置預測。本專題的核心目標在於開發一套低成本的擴增實境導航系統，將導航軟體中以地圖視角為中心的原始導航路線，轉化為以駕駛視角為中心的導航路徑，並以「地毯」的形式直接呈現在駕駛者前向相機視角的畫面中。相較於傳統依賴俯視地圖的導航方式，這種第一人稱視角的路線投影在直覺上更能降低駕駛者的認知負荷，進而提升導航過程的直覺性與安全性。美國新創公司 Phiar [10] 過去曾實體展示過這項擴增實境導航功能，但並未公開其核心技術。

近年自動駕駛的軌跡預測技術以深度學習神經網路模型為主流，其透過座標序列所表達的行車軌跡，可進一步轉化為道路節點，用以生成以自車視角為中心的導航路徑。然而，現有的軌跡預測模型在實務應用上仍存在諸多限制。學界中表現優異的最佳技術

(SOTA, 如 UniAD) 由於模型架構複雜且運算量龐大，難以在終端車載晶片上實現即時 (Real-time) 推論；且相關方法必須依賴高精度地圖 (HD Maps)、光學雷達 (LiDAR) [4] 或多視角相機影像 [3, 5, 6]，進而限制了其普及性。另一方面，產業界中低成本的端到端 (End-to-End) 模型 (如 OpenPilot) 雖降低了硬體門檻，但因其僅依賴前視相機影像作為輸入，在面臨複雜道路場景時，單一影像往往難以提供有效的遠距離道路線索；特別是在駛近交叉路口時，系統常會因為缺乏全局導航路徑的指引，而無法準確預測行徑方向。

為結合低成本與高精準度，本專題基於先前開發之 OLRA 系統 (一套基於地圖定位的低成本路線生成方案) 進行改良。OLRA 無需高昂成本的輸入配置，其主要依據前視相機影像獲取車道線資訊，並結合一般導航軟體或開源地圖中普遍可取得的導航路徑資料進行地圖定位評估，藉此避免了先前地圖定位研究對於高精地圖 [12] 或特定地圖屬性資料 [1, 7, 11, 14] 的依賴。相較於學界的 SOTA 模型 (如 UniAD)，OLRA 無需龐大網路；而相較於業界的代表作 OpenPilot，OLRA 則能在維持低成本輸入的同時，藉由地圖拓撲先驗知識，在遠距離與交叉路口取得更佳的預測表現。因此，本專題即基於 OLRA 的系統配置思路與最佳化框架進行效能優化，在提升精準度的同時保持即時性的運算。

在分析大量實際駕駛案例後，我們確認原模型在特定場景下存在兩個關鍵問題：

1. **航向偏移**：若模型在優化過程中僅專注將導航路徑貼齊車道線中心，當初始路徑存在橫向偏差時，系統易透過急遽轉向來快速降低橫向誤差。此現象使模型陷入局部最佳解 (Local Minimum)，導致生成的預測路線與實體車道方向嚴重不符。
2. **縱向誤差導致轉向時機錯誤**：系統初始化定位極易受 GPS 噪聲影響，使預測位置與真實位置產生縱向誤差。此落差會導致導航路線在路口發生提早或過晚轉彎的現象，進而造成投影至相機視角的軌跡出現縱向落差的錯誤指引，產生了導航路線穿越路緣的表現。

針對上述挑戰，本專題提出了一系列改良方案。首先，為解決航向偏移問題，系統修改了對齊損失 (Alignment Loss) 公式並新增航向誤差約束；而針對縱向誤差，則於系統初始化階段引入擴展卡爾曼濾波 (EKF) 來平滑 GPS 擾動，以逼近更真實的 GPS 位置。在此基礎上，我們進一步設計了以視覺為基礎的跨邊緣排斥損失 (Cross-Edge Loss, CRE)，用以抑制導航路徑不合理穿透實體路緣的行為。透過上述在初始化與最佳化階段的擴充，本系統在維持低成本運算的前提下，顯著提升了 AR 導航路徑生成的穩定性與幾何精準度。

2. 相關研究

本節簡要介紹代表性的地圖定位、軌跡預測與車道線預測相關研究，並與我們的系統進行比較。

早期地圖定位研究 [12] 依賴高精度地圖與影像對齊。為降低成本，後續方法改用開源地圖、空拍或街景圖進行匹配 [1, 7, 11, 14]。然而，這類方法多依賴特定地圖屬性，且像是 OrienterNet [11] 等方法多採用隱式特徵對齊，缺乏直觀的幾何解釋性。相較之下，我們僅利用普及的導航路徑進行顯式幾何對齊，不僅降低資料依賴、提升定位穩定性，更讓定位結果具備明確的幾何解釋力。

近年端到端軌跡預測研究逐漸由高成本的 LiDAR 與高精地圖方法 [4]，轉向以視覺為基礎的架構 [3, 5, 6]。我們的系統僅使用導航路徑，在擺脫高成本地圖依賴的同時，也避免了採用多視角影像。儘管近期有研究引入 Transformer 的注意力機制與大型語言模型 (LLM) 的場景高階語意推理，但在低階硬體上難以即時運行；並且，行車路線生成任務的複雜度不同於軌跡預測，我們的系統無需沉重的語意推理，因而更為輕量有效。

相較於學界的自動駕駛研究，美國新創公司 comma.ai 推出的軌跡預測模型僅需單一相機，便能在終端車載設備上即時運行。由於官方未開源訓練程式碼與完整資料集，我們採用 Chen 等人 [2] 的重製版本 OP-Deepdive，並在 nuScenes 資料集上重新訓練。然而，OP-Deepdive 缺乏具可解釋性的中間表徵，且在複雜路口常難以指出正確方向。對此，我們的系統同時結合單相機影像與地圖視角的導航路徑，在產生具幾何解釋性定位的同時，顯著提升了在路口處的預測表現。

在車道線模型的選擇上，我們的系統著重於利用車道線進行定位估計，而非提出新的偵測模型。由於幾何對齊僅需部分車道線資訊，我們優先考量低誤報率與偵測穩定性。相較於 Anchor-based 方法 [8, 13]，Row-wise based 的架構 [9] 具備更低的誤報率。因此，我們選擇採用已開源且經過大量測試的 Cond-LaneNet [9] 作為偵測模型。

3. 方法

3.1. Extended Kalman Filter

在自動駕駛定位系統中，由於車輛的運動學模型包含航向角的三角函數，屬於高度非線性系統，因此本系統引入擴展卡爾曼濾波器 (Extended Kalman Filter, EKF) 進行狀態估計。如圖 1 所示，於原有的 GPS 觀測模組前增加 EKF 前處理方塊，用以融合車輛自身的運動學預測與外部 GPS 訊號，從而提供更平滑且連續的初始位姿。

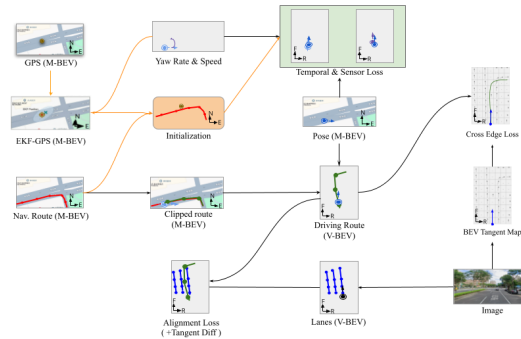


Figure 1. LLOV 系統架構圖，包含 EKF 前處理模組。

EKF 的核心概念是透過在當前狀態估計值附近進行一階泰勒展開 (Taylor Expansion)，計算 Jacobian 矩陣以實現系統局部線性化。在鳥瞰圖 (BEV) 定位系統中，EKF 的狀態向量定義為 $\mathbf{x} = [x, y, \theta]^T$ ，代表車輛在二維平面上的位置與航向角；控制輸入向量為 $\mathbf{u} = [s, \omega]^T$ ，包含由底盤傳感器取得的車速 (Speed) 與偏航率 (Yaw Rate)。

在預測階段 (Prediction Stage)，系統採用常速率常偏航率 (CRCT) 之非線性運動模型。當 $\omega \neq 0$ 時，位置的狀態轉移函數 $f(\cdot)$ 定義如下：

$$x_k = x_{k-1} - \frac{s}{\omega} (\cos \theta_k - \cos \theta_{k-1}) \quad (1)$$

$$y_k = y_{k-1} + \frac{s}{\omega} (\sin \theta_k - \sin \theta_{k-1}) \quad (2)$$

為在卡爾曼濾波框架下傳遞協方差，計算運動模型對狀態向量的 Jacobian 矩陣 $F_k = \frac{\partial f}{\partial \mathbf{x}}$ 。過程噪聲協方差矩陣設定為對角矩陣 $Q = \text{diag}(0.2^2, 0.2^2, 0.1^2)$ ，反映預測模型在位置與航向上的內在不確定性。

在更新階段 (Update Stage)，系統接收來自 GPS 的觀測向量 $\mathbf{z}_k = [x_{gps}, y_{gps}]^T$ 。由於 GPS 僅提供位置觀測而無航向角資訊，觀測矩陣設定為：

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (3)$$

測量噪聲協方差矩陣設定為 $R = \text{diag}(1.0^2, 1.0^2)$ ，以量化 GPS 訊號的誤差範圍。

為適應車道級定位需求，當 EKF 完成狀態向量與協方差更新後，系統會透過 `match_route` 函數獲取與地圖匹配的路徑航向角 (`map_heading`)，並以此替換 EKF 收斂的航向角作為最終輸出姿態。此設定確保了車輛姿態與預定義導航路徑 (En-route) 的絕對一致性。

3.2. Cross Edge Loss

在先前的地圖定位方法中，模型主要依賴本車道線對齊與感測器平滑進行軌跡最佳化。然而，當系統面臨嚴重的縱向定位誤差時，生成的行駛軌跡容易穿透影像中的實體路緣 (如安全島或道路邊界)，產生不合理的轉向行為。為解決此現象，本專題提出跨邊緣排斥損失 (Cross Edge Loss)。此方法透過捕捉真實世界中的物理邊界特徵，於車輛周圍建立虛擬排斥力場；當預測路徑具備穿越實體路緣的趨勢時，該損失值將顯著提高，藉此抑制異常轉向並強制模型修正空間姿態。

跨邊緣排斥損失的演算法整體流程可分為四個主要階段：

邊緣偵測 (Edge Detection)：首先對前向攝影機影像應用邊緣濾波器 (如 Sobel Filter) 提取高頻特徵，以獲取影像的初始邊緣梯度。

BEV 切線影像 (BEV Tangent Image)：接著，將相機視角下的梯度特徵透過透視轉換 (Homography) 投影至車輛鳥瞰圖 (V-BEV) 空間，並透過正規化處理維持空間尺度與物理梯度的一致性。

BEV 切線特徵圖 (BEV Tangent Map)：隨後，將連續的 BEV 梯度影像劃分為固定大小的離散網格，透過加權平均聚合出每個網格內的代表性邊緣切線方向與特徵強度。

損失函數設計 (Loss Design)：最後，計算預測車輛軌跡與特徵網格之間的空間交集。結合雙線性插值與角度約束機制 (過濾平行特徵，最大化正交衝突懲罰)，對穿越路緣的行為給予精確的排斥梯度。

上述四個階段的具體數學模型與實作細節，將於下列子章節依序說明。

3.2.1 BEV Tangent Image

在邊緣偵測前處理部分，為提升邊緣清晰度，經消融實驗測試多種濾波器與色彩通道搭配後證實，各類前處理結果均與單獨使用 Sobel (`ksize=5`) 相近，因此邊緣偵測階段僅採用單獨 Sobel 濾波器。

由於相機視角的影像邊緣受透視效果影響，預測路徑與路緣之間在真實空間的實際角度關係難以準確表達，因此需先將影像轉換至 BEV，在統一的俯視平面上保留幾何關係。

本流程並未先轉換至 BEV 再進行邊緣偵測，而是先於原始影像執行邊緣偵測後再行投影。原因在於透視轉換涉及重採樣與插值，近距離壓縮區域易產生影像失真並削弱梯度資訊。且由於相機座標系至 BEV 座標系轉換涉及非線性幾何變換，轉換後的各像素梯度需經由座標轉換關係重新計算。

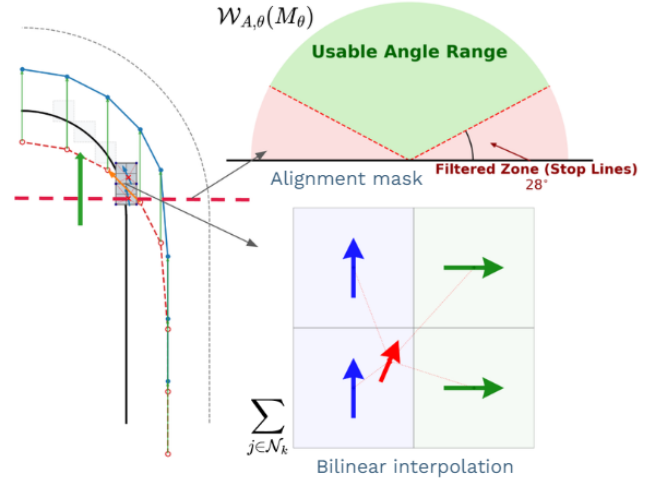


Figure 2. Optimization - Cross Edge Loss 優化示意圖

對於 BEV 影像中 ROI 內任一像素，設座標為 (x, y) 。首先透過 Homography 矩陣 \mathcal{H} ，將其映射至相機影像座標 (u, v) ：

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \mathcal{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

$$(u, v) = \left(\frac{x'}{w'}, \frac{y'}{w'} \right) \quad (5)$$

接著於相機影像中取得該位置之梯度向量 (g_u, g_v) 。因梯度方向物理上垂直於實體邊緣走向，取其正交向量 $(-g_v, g_u)$ 作為切線方向，並於前方取得採樣點座標 (u_s, v_s) ：

$$(u_s, v_s) = (u, v) + (-g_v, g_u) \quad (6)$$

再利用反向 Homography 矩陣 \mathcal{H}^{-1} ，將採樣點投影回 BEV 空間：

$$\begin{bmatrix} u'_s \\ v'_s \\ w' \end{bmatrix} = \mathcal{H}^{-1} \begin{bmatrix} u_s \\ v_s \\ 1 \end{bmatrix} \quad (7)$$

$$(x_s, y_s) = \left(\frac{u'_s}{w'}, \frac{v'_s}{w'} \right) \quad (8)$$

計算位移向量以取得 BEV 空間之邊緣方向 $(d'_x, d'_y) = (x_s, y_s) - (x, y)$ 。由於 Homography 轉換具備位置依賴之非均勻縮放特性，導致強度 $I_{bev} = \sqrt{d_x'^2 + d_y'^2}$ 產生尺度變化。為維持原始影像強度比例，以相機影像梯度強度 $I_{cam} = \sqrt{g_u^2 + g_v^2}$ 進行正規化：

$$(d_x, d_y) = \frac{I_{cam}}{I_{bev}} (d'_x, d'_y) \quad (9)$$

由所有 (d_x, d_y) 構成之向量場即為 BEV Tangent Image。

3.2.2 BEV Tangent Map

將 BEV Tangent Image 劃分為離散網格，每個網格內包含 N 個向量 $(d_x^i, d_y^i)_{i=1}^N$ ，透過加權平均計算網格代表方向。

為避免影像同一特徵線兩側的相反向量於加權平均時互相抵銷，將向量角度範圍限制於 $0^\circ \sim 180^\circ$ 之間，角度大於 180° 則反轉向量。定義網格總強度為：

$$I_{sum} = \sum_{i=1}^N I_i \quad (10)$$

其中 $I_i = \sqrt{(d_x^i)^2 + (d_y^i)^2}$ 。將各向量乘上其強度比例並加總，取得網格代表方向向量 (\bar{d}_x, \bar{d}_y) ：

$$\begin{aligned} (\bar{d}_x, \bar{d}_y) &= \sum_{i=1}^N \frac{I_i}{I_{sum}} (d_x^i, d_y^i) \\ &= \frac{1}{I_{sum}} \sum_{i=1}^N I_i (d_x^i, d_y^i) \end{aligned} \quad (11)$$

3.2.3 Loss Design

鑑於縱向定位誤差常引發實體路緣穿透現象，跨邊緣排斥損失 (\mathcal{L}_{CE}) 旨在透過地圖邊緣建立虛擬排斥力，引導優化過程避開不可行駛區域。

提取特徵網格之梯度強度矩陣 \mathcal{I}_{grad} ，將法向量旋轉 90° 轉換為路緣切線方向角矩陣 M_θ 。同時計算預測軌跡 R_V^t 之線段中點 m_k 、步長向量 ΔR_V 與軌跡方向角 $R_{V,\theta} = \text{atan2}(\Delta x, \Delta y)$ 。 \mathcal{L}_{CE} 計算定義如下：

$$\mathcal{L}_{CE} = \sum_{k=1}^K \sum_{j \in \mathcal{N}_k} \left(\underbrace{\mathcal{W}_{dist}(d_k)}_{\text{Spatial Weight}} \cdot \underbrace{\|\Delta R_V\| \cdot \eta_j}_{\text{Distributed Step}} \cdot \underbrace{\mathcal{W}_{A,\theta}(M_\theta)}_{\text{Alignment Mask}} \cdot \underbrace{\mathcal{L}_{A,\theta}(R_{V,\theta}, M_\theta)}_{\text{Heading Loss}} \cdot \underbrace{\mathcal{I}_{grad}}_{\text{Intensity}} \right) \quad (12)$$

其中 \mathcal{N}_k 為透過雙線性插值取得之相鄰像素點集合。 $\|\Delta R_V\| \cdot \eta_j$ 確保懲罰量平滑分攤至網格； \mathcal{I}_{grad} 數值越大代表物理排斥力越強。為精確懲罰穿越路緣行為並排除停止線等橫向特徵干擾，角度約束分為兩部分：

對齊遮罩 ($\mathcal{W}_{A,\theta}$)：根據邊緣絕對方向角 M_θ 進行空間濾波。由於縱向車道線通常平行於道路，橫向停止線則垂直，透過正弦函數指數衰減自動過濾橫向特徵：

$$\mathcal{W}_{A,\theta}(M_\theta) = \epsilon_{ce} + (1 - \epsilon_{ce}) e^{-\frac{\sin^2(M_\theta)}{\theta_c}} \quad (13)$$

航向損失 ($\mathcal{L}_{A,\theta}$)：評估軌跡撞擊邊緣之嚴重程度。若軌跡與路緣平行則無橫向阻力；若正交穿透則給予最高懲罰：

$$\mathcal{L}_{A,\theta}(R_{V,\theta}, M_\theta) = |\sin(R_{V,\theta} - M_\theta)| \quad (14)$$

最後引入距離衰減權重 $\mathcal{W}_{dist}(d_k) = e^{-d_k/\sigma_d}$ ，使系統優先處理近端之碰撞危機。

4. 實驗結果

4.1. 實驗設置

為確保實驗公平性與可重現性，消融實驗與基準對比均於統一參數下進行（詳見表 1）。軌跡優化階段引入 Cosine Annealing 動態學習率排程確保收斂，並設定跨邊緣排斥 (CRE) 權重為 15.0，以建構穩定邊界約束。

Table 1. 系統核心參數設定表。

模組	參數名稱	符號	數值
EKF	狀態/控制向量	\mathbf{x}, \mathbf{u}	$[x, y, \theta]^T, [s, \omega]^T$
	過程/測量噪聲	Q, R	$\text{diag}(0.04, 0.04, 0.01), \text{diag}(1, 1)$
路徑	步長/平滑/權重	-	4.0, 16, 0.5
LLOV	對齊/傳感器	w_{align}, w_{sensor}	1.0, 0.5
	平滑/排斥力	w_{temp}, w_{ce}	0.1, 15.0
優化	排程/基礎 LR	-	Cosine, $10^{-2} \sim 10^{-3}$

4.2. 量化評估

本節定量分析 LLOV 系統之各項效能指標，並與端到端模型 OpenPilot 及前代純幾何優化方法 (Pure OLRA) 進行對比。

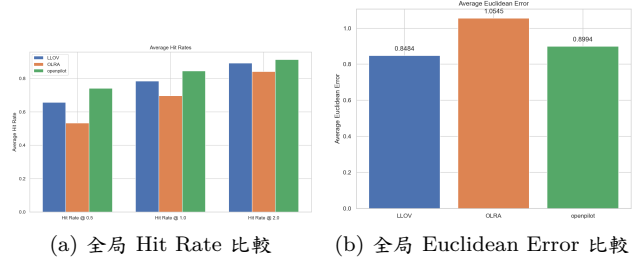


Figure 3. LLOV、Pure OLRA 與 OpenPilot 之全局性能指標對比直方圖。展示了全局平均命中率與幾何歐氏誤差的量化關係。

與基準方法比較 本節將 LLOV 系統與端到端非地圖定位模型 OpenPilot 進行綜合評比。實驗顯示，LLOV 的 Hit Rate 與 Euclidean Error 皆優於 OLRA；與 OpenPilot 相比，LLOV 的 Hit Rate 雖較差，但在 Euclidean Error 上展現出更好的效果。我們認為主要原因有二：一是測試集中直行場景佔多數，使 OpenPilot 在轉彎處的誤差被大幅稀釋；二是 LLOV 高度依賴車道線偵測，較易受感知網路不穩定的影響。

若我們將距離細分以個別評估（圖 4），LLOV 相較 OLRA 仍獲全面提升；雖近距離 (0-10 m) 貼合度微觀表現上略遜於 OpenPilot，但隨距離拉長，地圖先驗與 EKF 濾波的優勢逐漸顯現。特別是在 LLOV 修正 OLRA 的航向偏差與縱向落差後，於 30-40 m 遠距區間的 Hit Rate 與 Euclidean Error 皆優於 OpenPilot，達成更穩健的遠程表現。

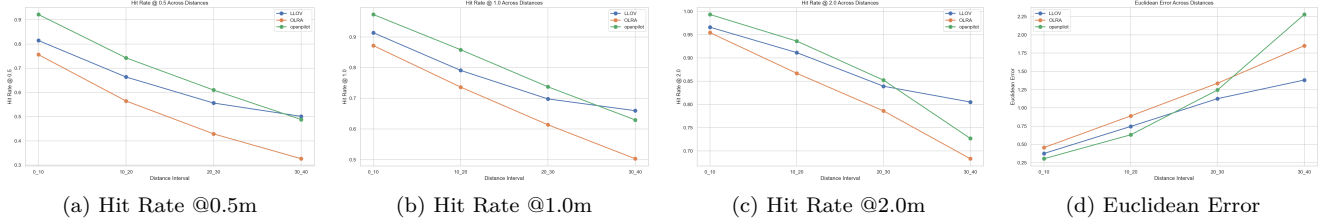


Figure 4. 整體效能評估 (Comparison with OpenPilot)。

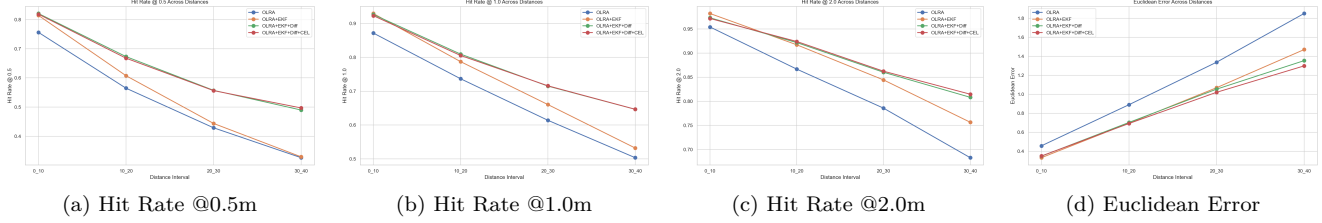


Figure 5. 內部消融實驗 (Ablation Study) 之定量分析。結果證實各模組的疊加能穩步降低預測誤差並提升命中率。

消融實驗 為釐清各模組性能貢獻，設計四階段消融實驗：Pure OLRA (沿用前一代基礎設定，無濾波與學習率排程器)；OLRA + EKF (引入 EKF 進行位姿預處理)；OLRA + EKF + Tangent Diff (啟動切線差機制與動態學習率排程)；LLOV (Full) (加入邊緣排斥力 CRE 的完整模型)。

如圖 5 所示，未經濾波的 Pure OLRA 基準模型易受 GPS 噪聲干擾，全局平均誤差約 1.05 m。加入 EKF 有效抑制噪聲後，全局誤差降至 0.81 m。加入 Tangent Diff 模組後，系統的命中率在各距離區間均出現了顯著的攀升 (微觀指標 Hit Rate @0.5m 從 58.2% 躍升至 65.4%)，證明放寬彎道約束能有效解決軌跡抖動。最終加入 CRE 的 LLOV 模型將全局誤差收斂至 0.79 m，並成功壓制 30-40 m 的遠程誤差發散。

執行時間比較 本專題於同一硬體平台上測試了各方法的單幀平均推論時間 (表 2)。OpenPilot 與 CondlaneNet 高度依賴 GPU 平行運算，具備極低延遲；而 LLOV 於 CPU 運算，單幀執行時間為 33 ms。經進一步分析，主要運算開銷在 CRE 損失計算，若關閉 CRE 執行時間可降至 11 ms。我們在消融實驗發現，關閉 CRE 後精準度雖微幅損失，但整體表現仍超越 OLRA 與 OpenPilot，可作為更輕量的系統版本。

Table 2. 單幀平均推論時間 (Runtime) 之比較。涵蓋感知網路與各優化演算法階段之執行時間。

Method	Runtime (ms)	Platform
OpenPilot (Reference)	~ 4.0	GPU
Condlanenet (Perception)	~ 6.0	GPU
OLRA (Perception)	~ 9.0	CPU
LLOV w/o CRE (EKF + Diff)	10.7	CPU
LLOV (Full w/ CRE)	33.0	CPU

誤差與 Ego-Lane 比例分析 為探討系統對本車車道線 (Ego-Lane) 可見度的依賴程度，我們分析車道線提取比例與平均歐氏誤差之關聯。圖 6 的線性擬合顯示，當 Ego-lane ratio 降低時，歐氏誤差呈上升趨勢。多數場景即便車道線缺失，誤差仍能被 EKF 與 CRE 控制在 2 m 內；但我們觀察到在 Ego-lane ratio 約 0.28 處存在高達 10.78 m 的離群值。這表明在惡劣感知條件下，單一邊界約束可能失效，我們將於下節深入探討此侷限性。

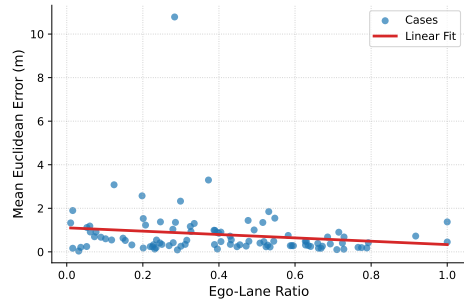


Figure 6. LLOV 模型於不同場景之 Ego-lane ratio 與平均歐氏誤差關聯性。

4.3. 質化評估

為進一步說明定量數據背後之實際軌跡行為，針對高拓撲複雜度場景進行鳥瞰圖質性分析。

右轉彎道修正 (Case 0345)：於缺乏導引線之右轉路口，Pure OLRA 產生 2.30 m 偏移。引入 EKF 修正縱向誤差並整合 Tangent Diff 後成功實現切線對齊，使誤差縮減至 0.15 m，過彎貼合度優於端到端模型。

大跨度左轉 (Case 0638)：寬闊左轉路口中，OpenPilot 與真實軌跡存在 0.67 m 橫向偏移。本系

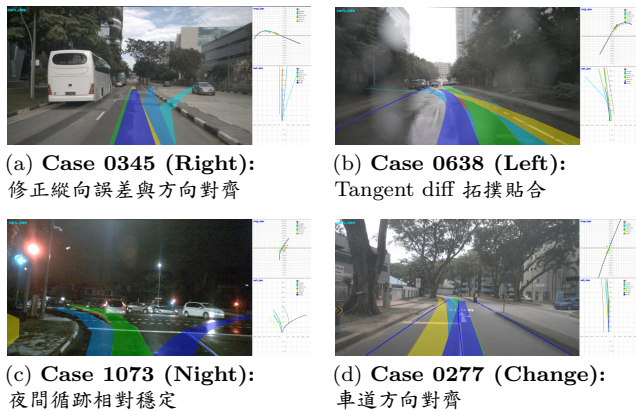


Figure 7. 質性分析視覺化比較 (Qualitative Results)。圖中展示了在四種高複雜度場景下各方法的軌跡分佈。可以觀察到 LLOV 系統在過彎平滑度、夜間穩定性以及複雜車道切換上均展現出優異的性能。

統展現拓撲精確度：Pure OLRA 因全域噪聲導致 3.02 m 誤差；EKF 將其壓制至 0.72 m；Tangent Diff 則解決航向過度修正，使最終誤差收斂至 0.27 m，展現出與真實駕駛行為的貼合。

夜間與複合場景 (Case 1073 & 0277)：夜間視覺特徵退化時，LLOV 結合地圖與 EKF 將誤差控制於 0.49 m 內，視覺依賴度低於 OpenPilot。在複合車道切換路口，系統仰賴切線差機制完成切換，達成 0.35 m 低誤差。

侷限性分析：當車道線消失且受斑馬線或樹影干擾時，邊緣檢測易誤判。過高 CRE 權重會使偽邊緣產生異常排斥力，導致軌跡發生非預期避讓扭曲。此現象凸顯系統缺乏高階語義理解，後續可引入語義分割 (Semantic Segmentation) 過濾障礙物以提升泛化能力。

References

- [1] Gaoya Cao, Florian Damerow, Benedict Flade, Markus Helmling, and Julian Eggert. Camera to map alignment for accurate low-cost lane-level scene interpretation. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 498–504. IEEE, 2016. 1, 2
- [2] Li Chen, Tutian Tang, Zhitian Cai, Yang Li, Penghao Wu, Hongyang Li, Jianping Shi, Junchi Yan, and Yu Qiao. Level 2 autonomous driving on a single device: Diving into the devils of openpilot, 2022. 2
- [3] Zhili Chen, Maosheng Ye, Shuangjie Xu, Tongyi Cao, and Qifeng Chen. Ppad: Iterative interactions of prediction and planning for end-to-end autonomous driving. In *European Conference on Computer Vision*, pages 239–256. Springer, 2024. 1, 2
- [4] Kashyap Chitta, Aditya Prakash, and Andreas Geiger. Neat: Neural attention fields for end-to-end autonomous driving. In *International Conference on Computer Vision (ICCV)*, 2021. 1, 2
- [5] Shengchao Hu, Li Chen, Penghao Wu, Hongyang Li, Junchi Yan, and Dacheng Tao. St-p3: End-to-end vision-based autonomous driving via spatial-temporal feature learning. In *European Conference on Computer Vision*, pages 533–549. Springer, 2022. 1, 2
- [6] Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, Lewei Lu, Xiaosong Jia, Qiang Liu, Jifeng Dai, Yu Qiao, and Hongyang Li. Planning-oriented autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023. 1, 2
- [7] Xiaosong Jia, Chenhe Zhang, Yule Jiang, Songbur Wong, Zhiyuan Zhang, Chen Chen, Shaofeng Zhang, Xuanhe Zhou, Xue Yang, Junchi Yan, and Yu-Gang Jiang. Spatial retrieval augmented autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026. 1, 2
- [8] Dongkwon Jin, Wonhui Park, Seong-Gyun Jeong, Heeyeon Kwon, and Chang-Su Kim. Eigenlanes: Data-driven lane descriptors for structurally diverse lanes. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 17163–17171, 2022. 2
- [9] Lizhe Liu, Xiaohao Chen, Siyu Zhu, and Ping Tan. Condlanenet: a top-to-down lane detection framework based on conditional convolution. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3773–3782, 2021. 2
- [10] Phiar. Ai-powered ar navigation for driving, 2020. 1
- [11] Paul-Edouard Sarlin, Daniel DeTone, Tsun-Yi Yang, Armen Avetisyan, Julian Straub, Tomasz Malisiewicz, Samuel Rota Bulo, Richard Newcombe, Peter Kotschieder, and Vasileios Balntas. Orienternet: Visual localization in 2d public maps with neural matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21632–21642, 2023. 1, 2
- [12] Markus Schreiber, Carsten Knöppel, and Uwe Franke. Laneloc: Lane marking based localization using highly accurate maps. In *2013 IEEE Intelligent Vehicles Symposium (IV)*, pages 449–454, 2013. 1, 2
- [13] Lucas Tabelini, Rodrigo Berriel, Thiago M. Paixao, Claudine Badue, Alberto F. De Souza, and Thiago Oliveira-Santos. Keep your eyes on the lane: Real-time attention-guided lane detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 294–302, 2021. 2
- [14] Teng Wang, Jiawen Li, and Changyin Sun. Dehi: A decoupled hierarchical architecture for unaligned ground-to-aerial geo-localization. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(3):1927–1940, 2023. 1, 2