

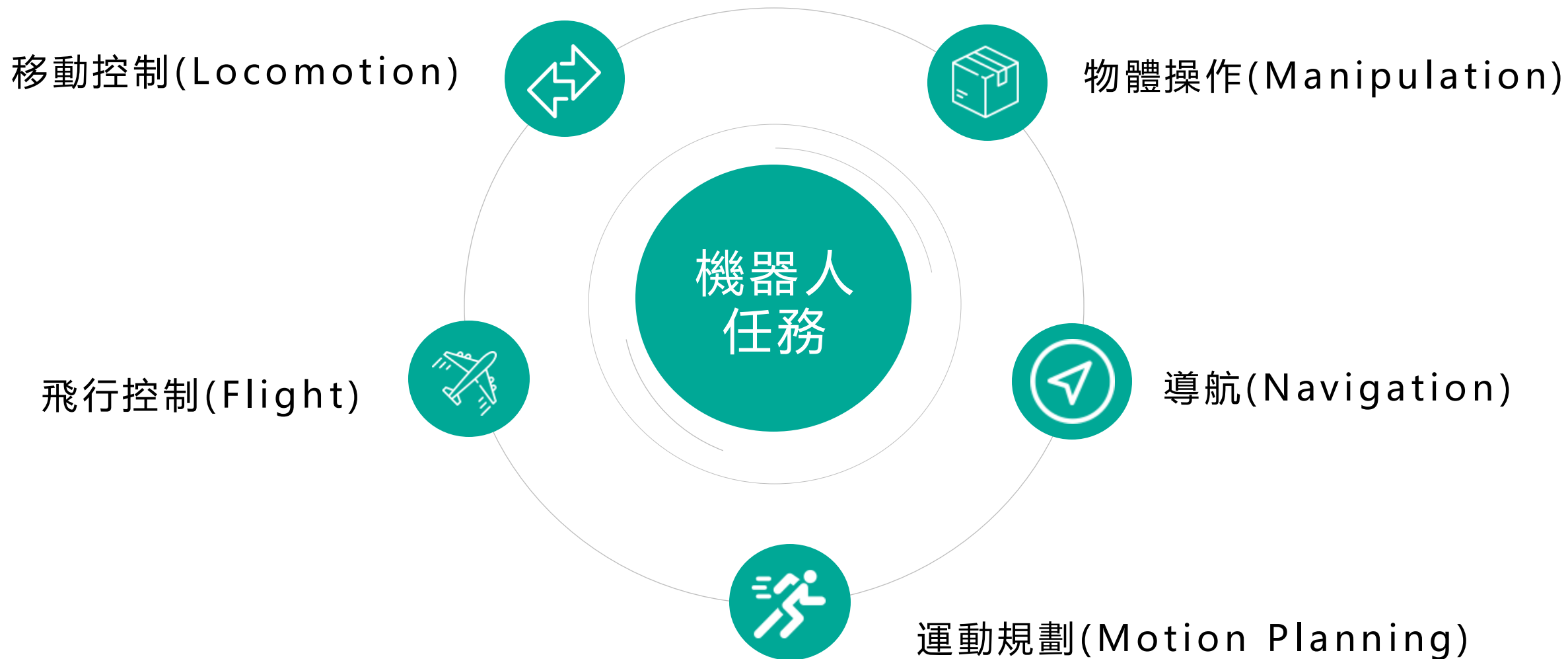


AWS DeepRacer Hyperparameter Tuning with Sim2Real RL

M144020007 詹襄秣

M144020062 郭捷圻

專案背景



挑戰與限制

高樣本複雜度

虛擬環境訓練

訓練成本高

真實差異
(Reality Gap)

專案目標

01

複現論文實驗

- 嘗試多種DeepRacer論文中超參數設定與調整策略

效能最佳化

- 尋找穩定且高效的訓練配置

02

03

Sim-to-Real 遷移研究

- 分析模擬與真實環境間的差異
- 探討縮小Reality Gap方法

建立評估機制

- 了解如何識別可成功遷移的訓練檢查點

04

模 型 建 立



Proximal Policy Optimization 演算法

- 神經網路：Policy Network, Value Network
- 網路架構：3 個卷積層 + 2個全連接層 + 定期更新策略參數(20回合)

參 數 介 紹

Gradient descent batch size

- 每次梯度更新會同時參考n筆過去駕駛經驗，提升訓練穩定性

Discount factor

- 控制模型對未來獎勵的重視程度，影響是否採取較具前瞻性的駕駛策略

Loss function

- 衡量模型預測與實際回饋之間的誤差，引導進行學習與修正

Learning rate

- 決定模型每次更新參數的調整幅度，影響收斂速度與訓練穩定性

SAC alpha

- 追求獎勵最大化同時，鼓勵策略保持一定的探索性，避免過早陷入局部最佳解

實驗設計

- 實驗一：

Gradient descent batch size : 64

Discount factor: 0.99

Loss Type: mean squared error

Learning rate : 0.003

SAC alpha: 0.2

- 實驗二：根據實驗一，加長訓練時間

實驗設計

- 實驗三：

Gradient descent batch size : 128

Discount factor: 0.99

Loss Type: mean squared error

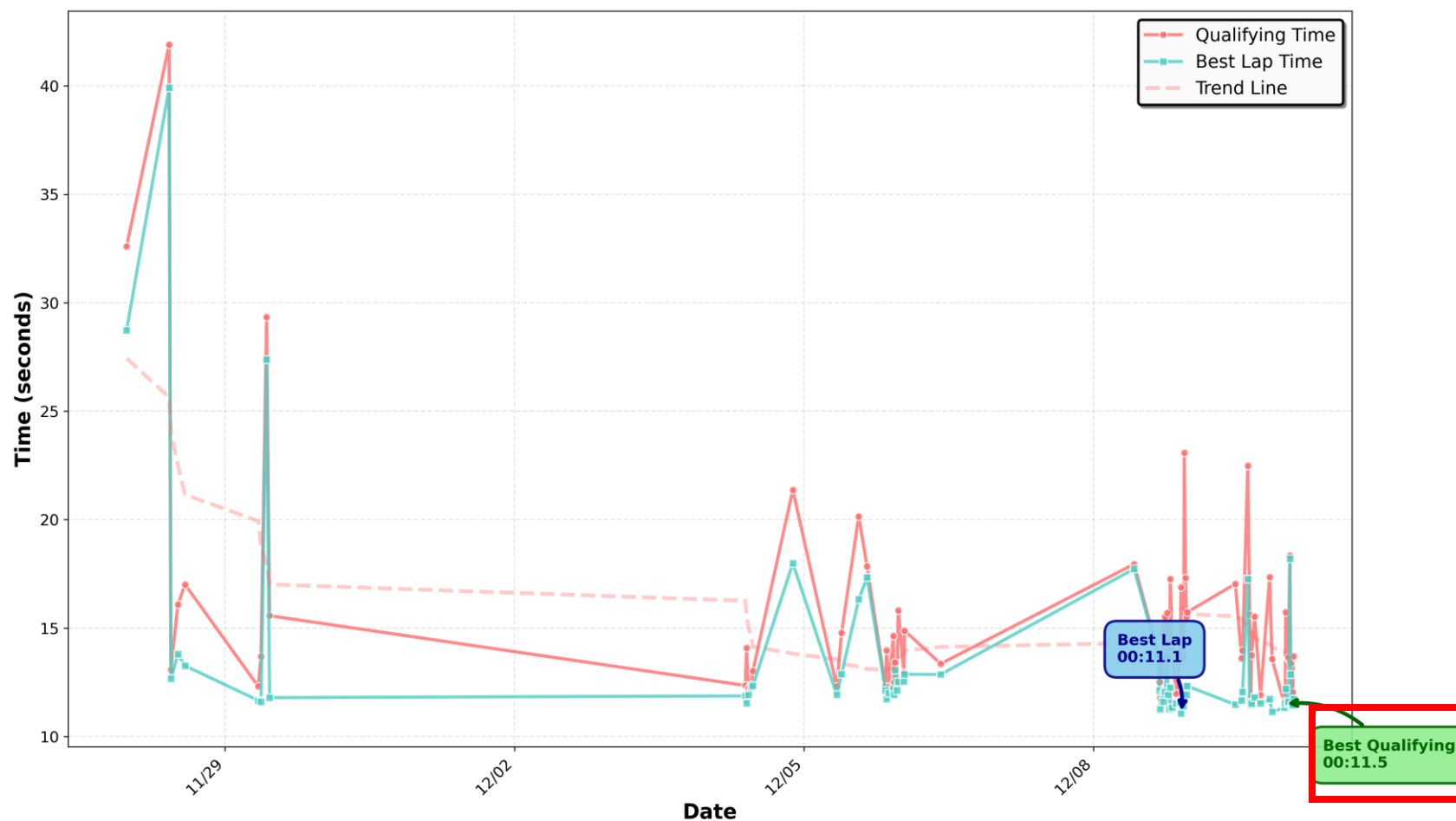
Learning rate : 0.001

SAC alpha: 0.2

實驗結果 - 模型成效

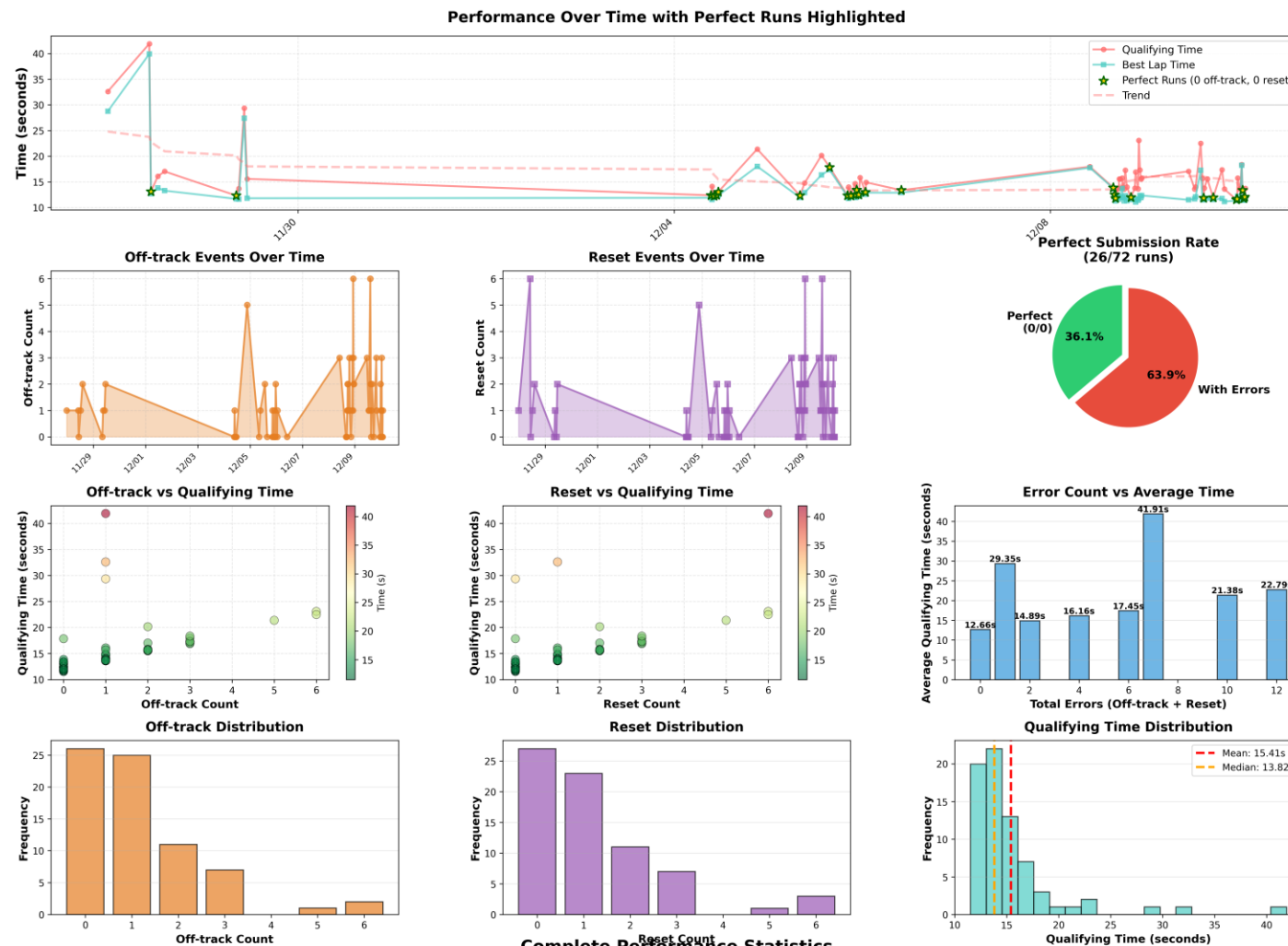
- 總訓練次數：72次

AWS DeepRacer Performance Timeline
72 Submissions



實驗結果 - 模型成效

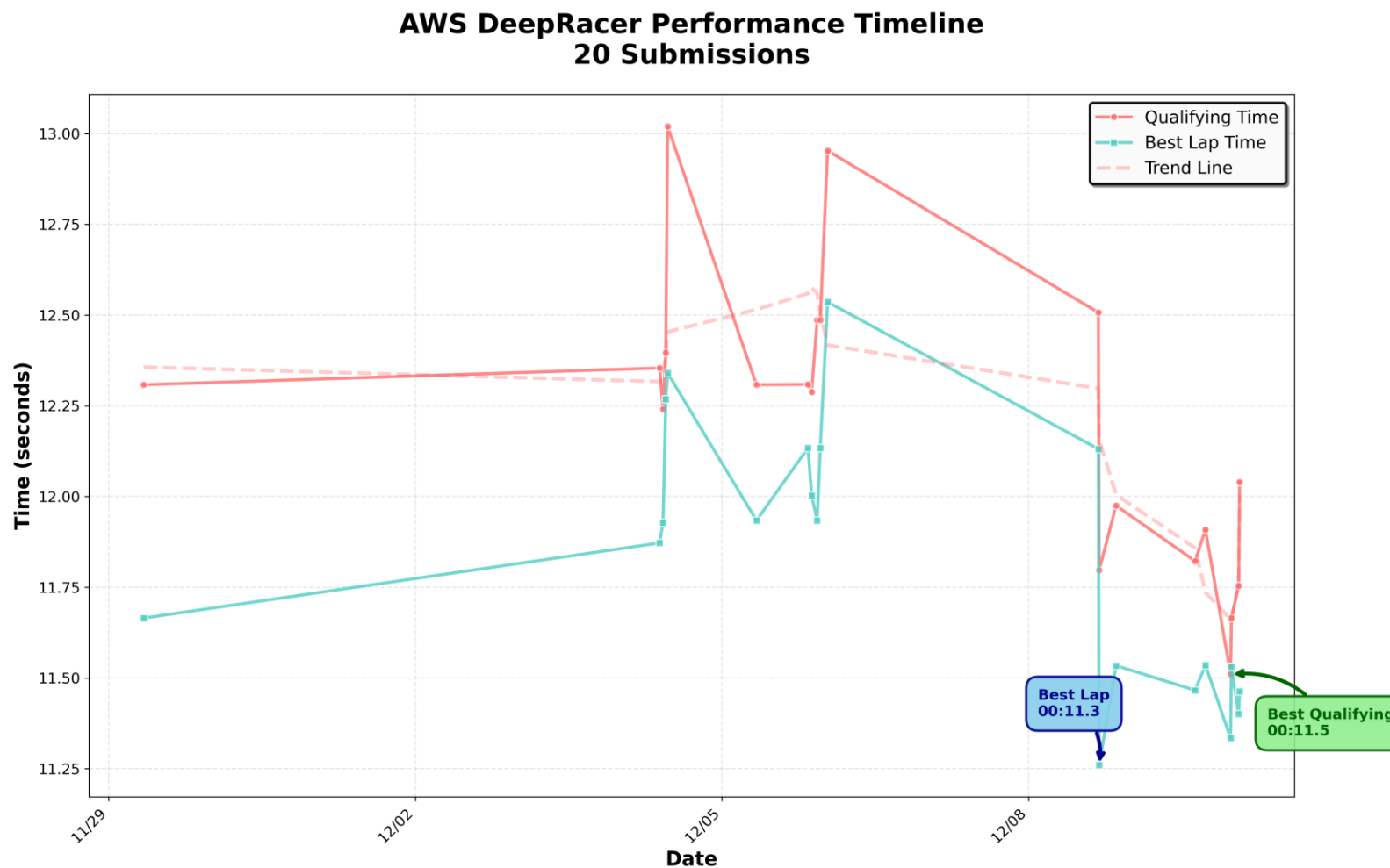
- 總訓練次數：72次



Metric	Qualifying Time	Off-track	Reset	Perfect Runs
Best Overall	00:11.5	0	0	00:11.5
Average	15.413s	1.18	1.24	26 (36.1%)
Total	72 submissions	85 events	89 events	26/72 runs
Max	00:41.9	6 times	6 times	Rate: 36.1%

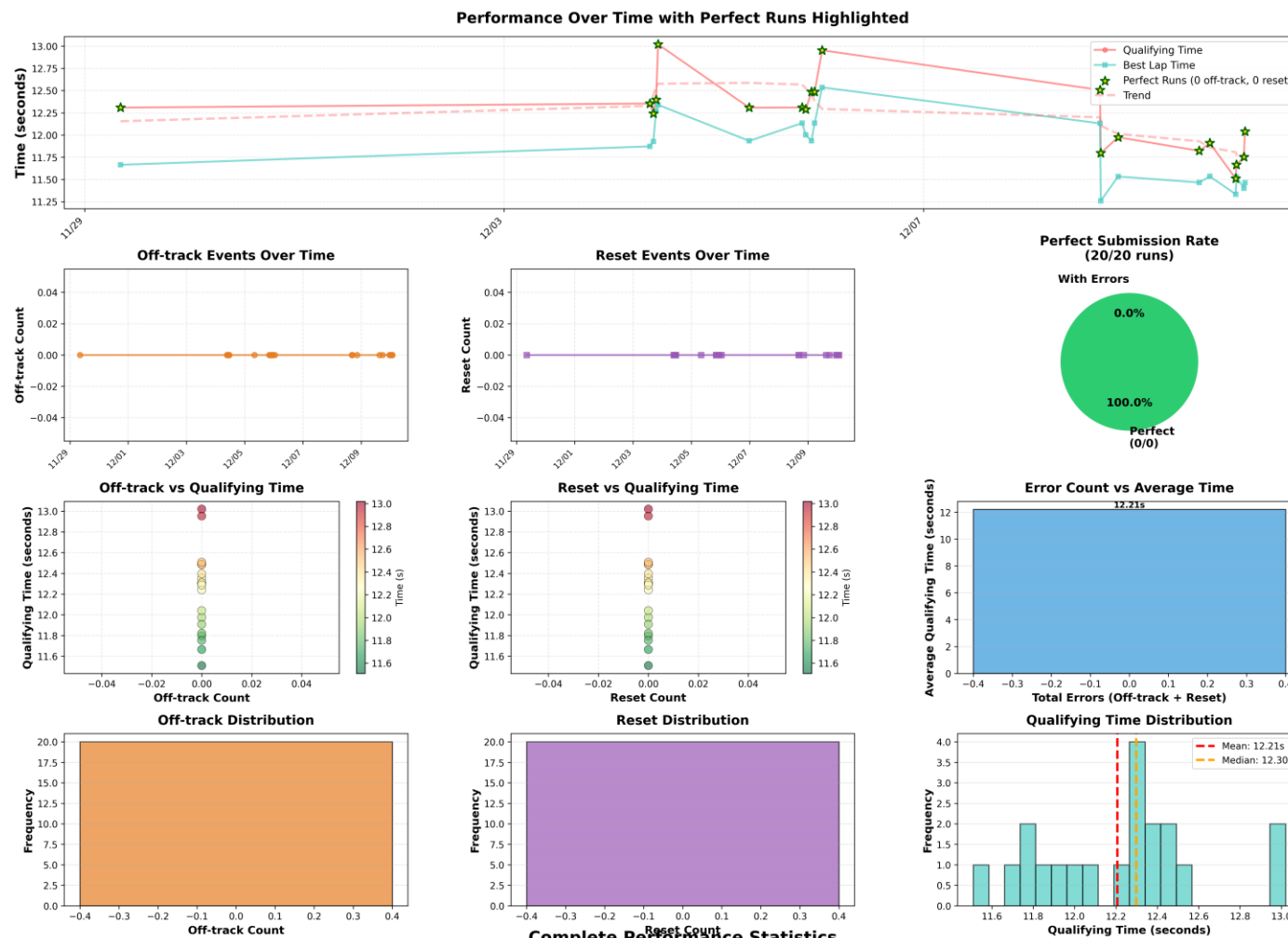
實驗結果 - 模型成效

- 最佳20次模型



實驗結果 - 模型成效

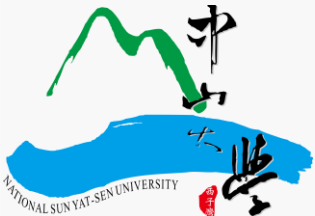
● 最佳20次模型



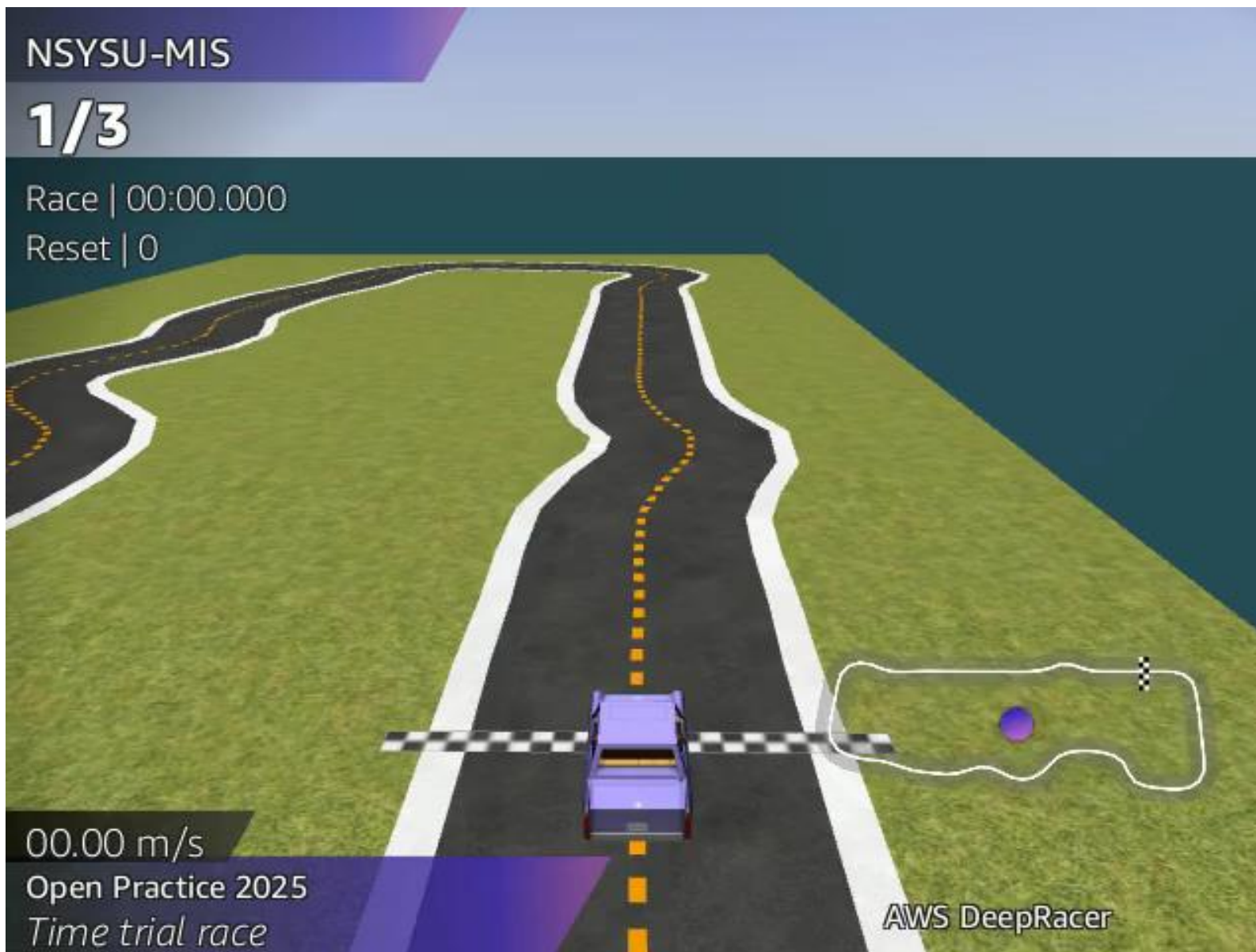
Metric	Qualifying Time	Off-track	Reset	Perfect Runs
Best Overall	00:11.5	0	0	00:11.5
Average	12.206s	0.00	0.00	20 (100.0%)
Total	20 submissions	0 events	0 events	20/20 runs
Max	00:13.0	0 times	0 times	Rate: 100.0%

實驗結果 - 排名

- 臺灣地區：1/261
- 大中華地區：2/422
- 世界：27/6650

 NSYSU-MIS		Your qualifying time 00:11.510		Qualifying model submitted DR2-v2-fine-tune2	
Taiwan		Qualifying time 00:11.510 NSYSU-MIS		My race ranking 1/261	Gap to 1st --
Greater China		Qualifying time 00:11.441 UpgradedModel#5475		My race ranking 2/422	Gap to 1st +00:00.069
World		Qualifying time 00:07.819 rosscomp1		My race ranking 27/6650	Gap to 1st +00:03.691

實驗結果

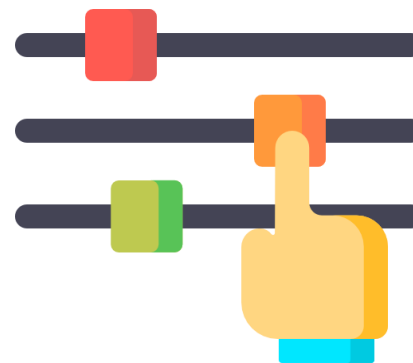


研 究 發 現



獎 勵 函 數 設 計

- 複雜度需適中
- 進度獎勵的重要性
- 速度與轉向的協調



超 參 數 調 整 方 向

- Learning Rate
- SAC Alpha
- Batch Size

挑戰與限制

高樣本複雜度

虛擬環境訓練

訓練成本高

真實差異
(Reality Gap)

未 來 工 作

最佳化Racing Line

- 導入最佳賽車線概念，使用梯度下降或基因演算法計算最佳路徑

精簡Action Space

- 限制轉向角，學習更平滑策略

階段式訓練策略

- 穩定性 → 速度 → 精細調整

Waypoint-based 速度控制

- 實現更細緻的速度管理

Sim2Real 遷移研究

- 探討感測器噪聲模擬、Domain Randomization 等技術

Thanks for listening

