

# AWS DeepRacer Hyperparameter Tuning with Sim2Real RL

M144020007 詹襄林

M144020062 郭健圻

# 專案背景

移動控制(Locomotion)



飛行控制(Flight)



機器人  
任務

物體操作(Manipulation)



導航(Navigation)



運動規劃(Motion Planning)

# 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

## 效能最佳化

- 尋找穩定且高效的訓練配置

複現論文實驗

- 嘗試多種DeepRacer論文中超參數設定與調整策略

02

## 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

03

## 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

Environment

賽道

Agent

賽車

Observation

灰階影像

Action

油門/方向盤

Reward

獎勵機制

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

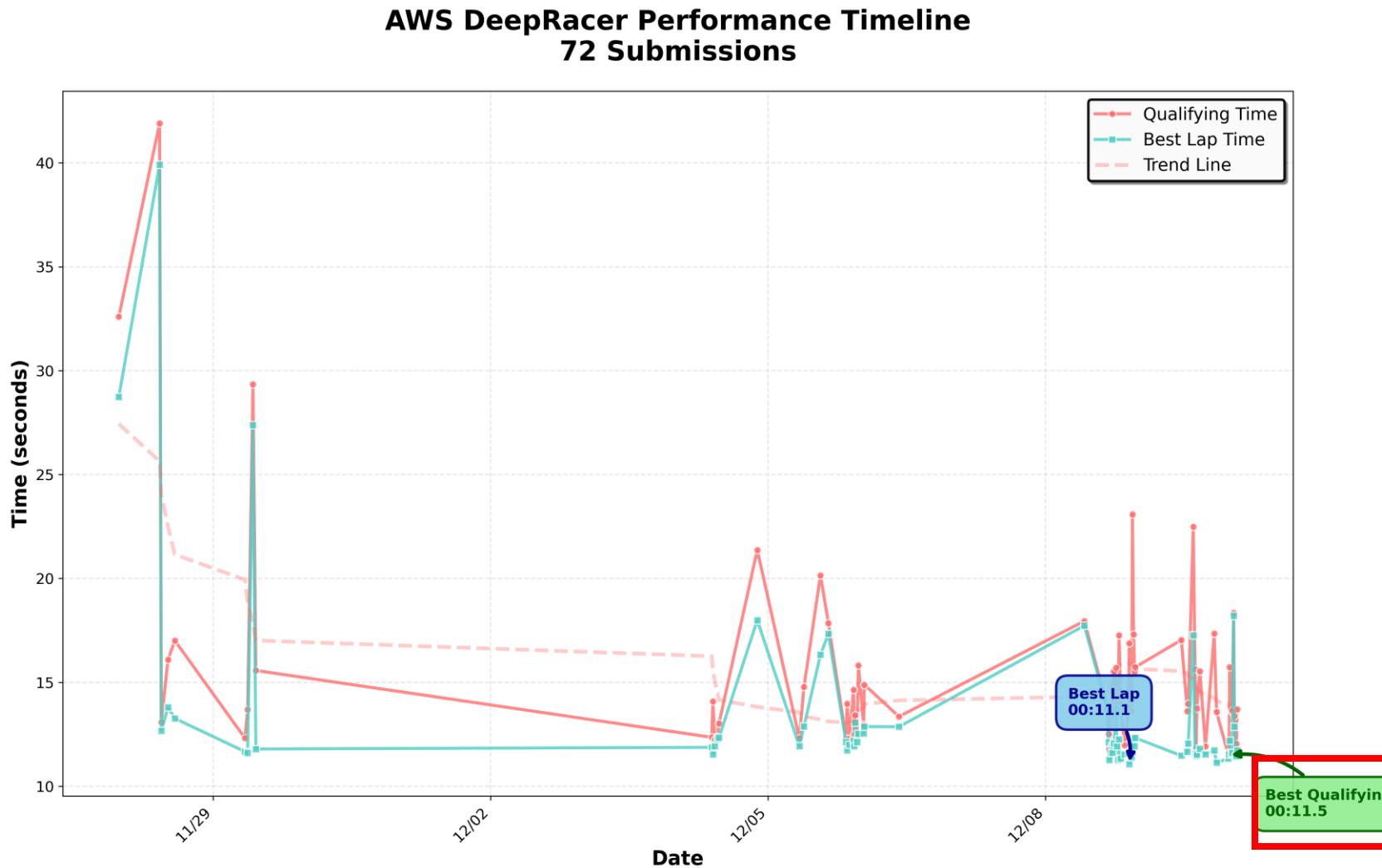
Loss Type: mean squared error

Learning rate : 0.001

SAC alpha: 0.2

# 實驗結果 - 模型成效

- 總訓練次數：72次



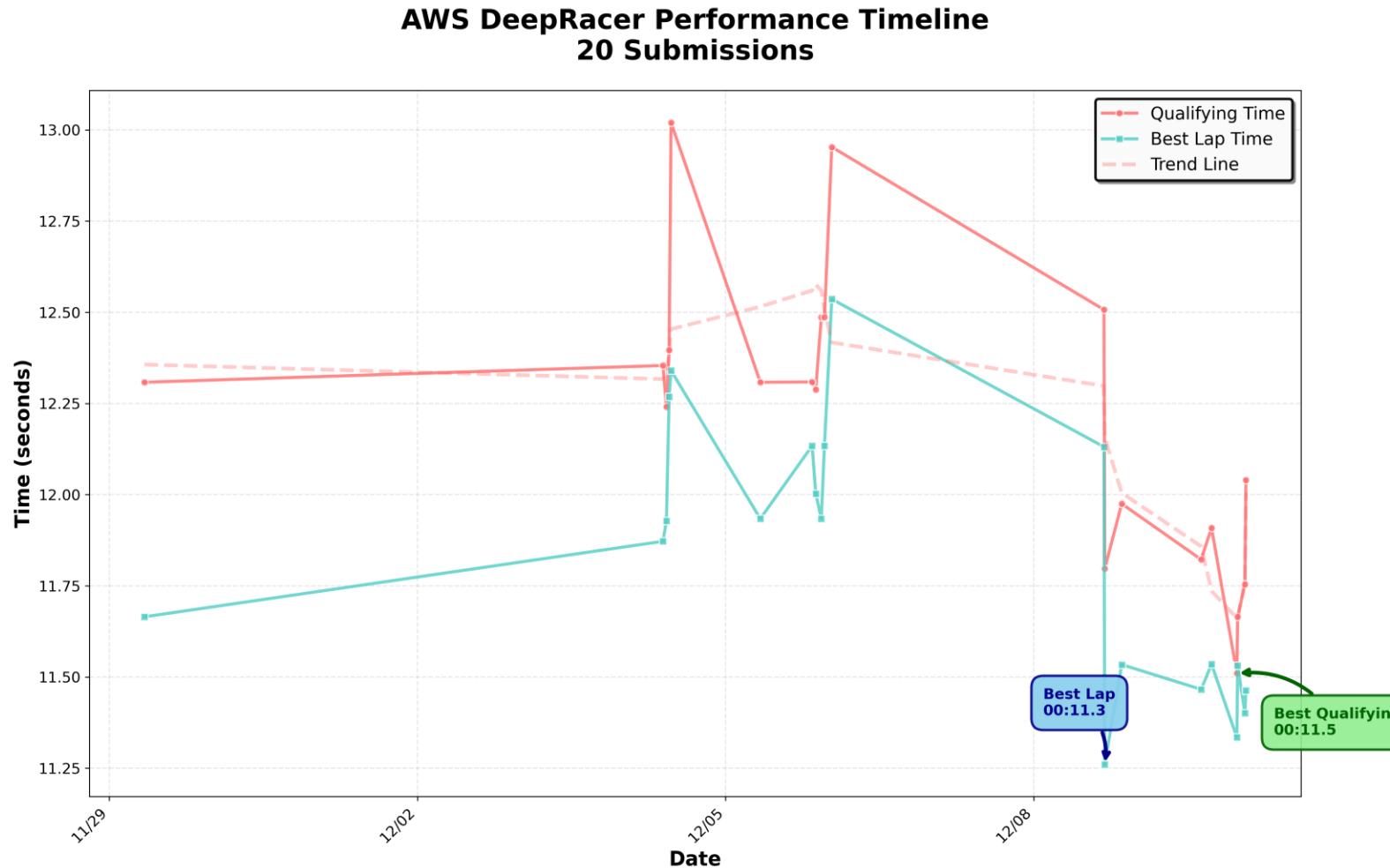
# 實驗結果 - 模型成效

- 總訓練次數：72次



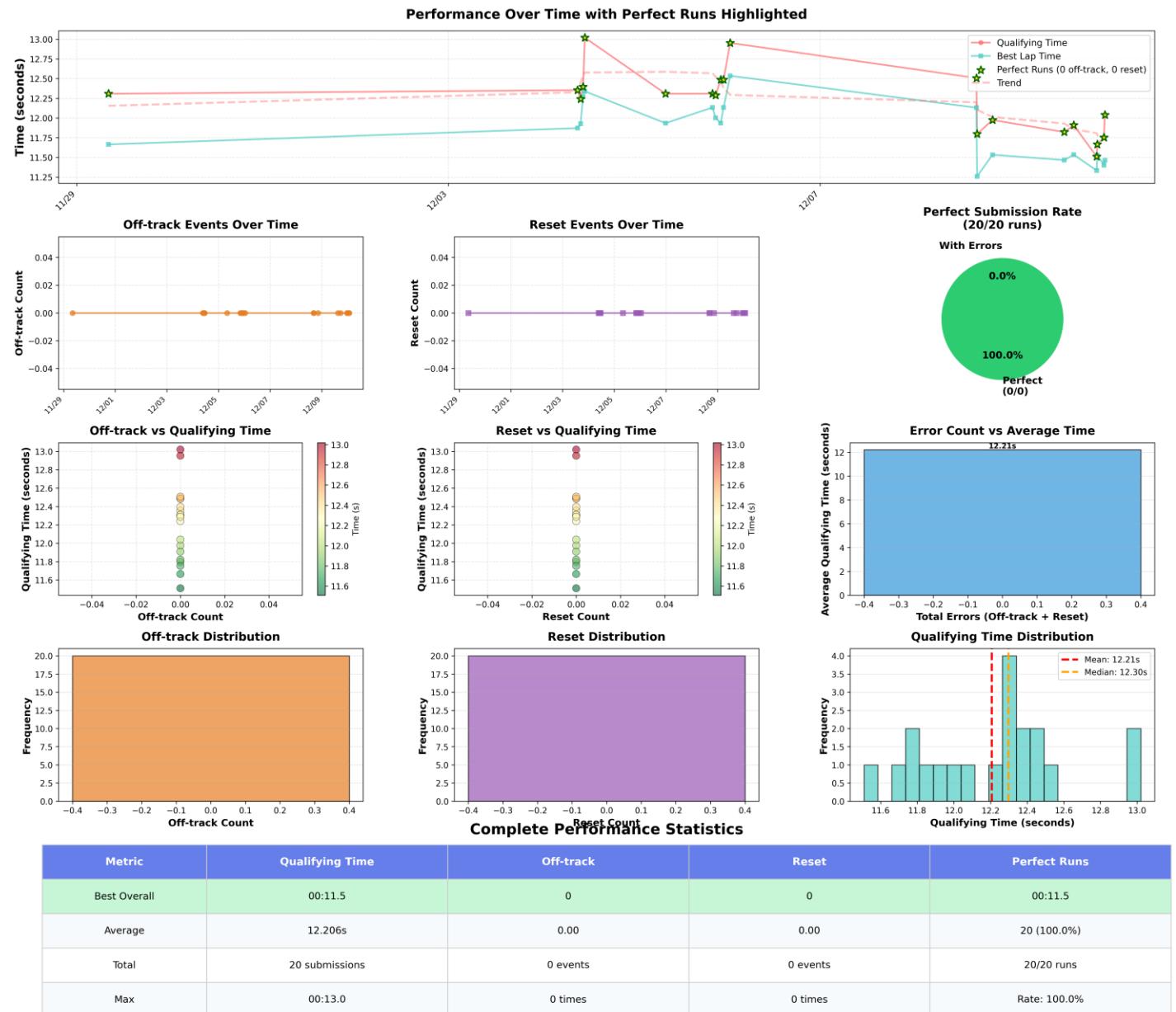
# 實驗結果 - 模型成效

## ● 最佳20次模型



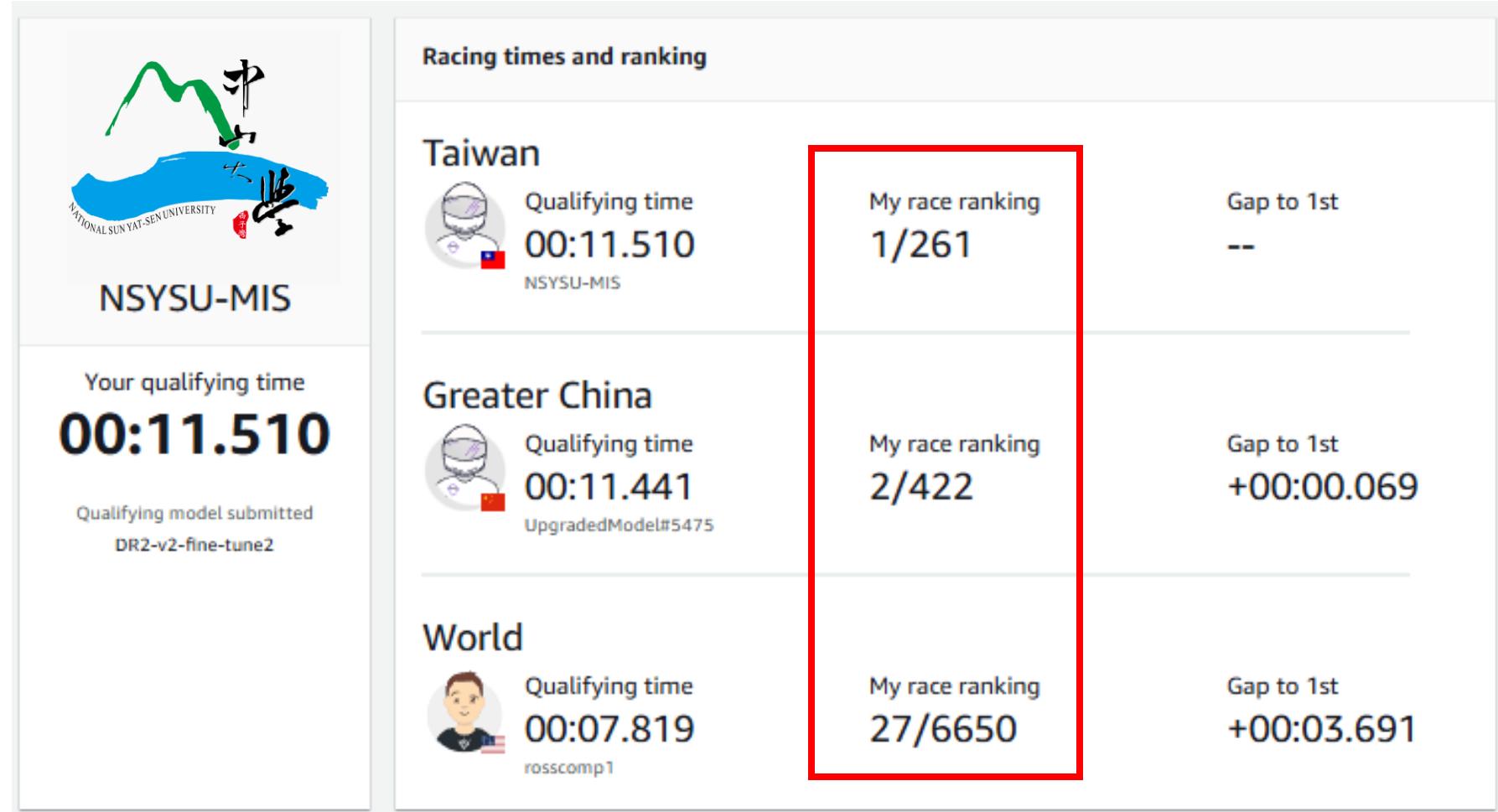
# 實驗結果 - 模型成效

## ● 最佳20次模型



# 實驗結果 - 排名

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



# 實驗結果

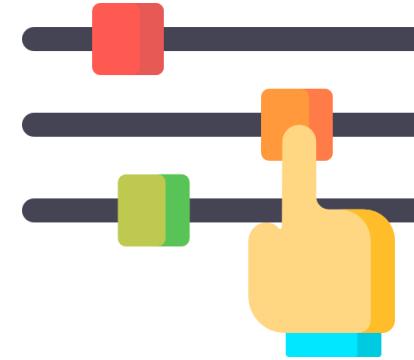


# 研 究 發 現



## 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 来 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

