

國立成功大學資訊工程學系大三專題

基於深度強化式學習之智慧玻璃控制

Smart Glass Control via Deep Reinforcement Learning

指導教授: 蘇銓清 學生: 劉承祐 F74122145

大綱

- 背景與研究動機
- 問題陳述
- 系統架構
- 方法架構
- 採用策略
- 模型訓練結果與成效
- 進階測試資料集
- 進階測試資料集結果分析
- 結論與未來展望

背景與研究動機

- 智慧建築的概念興起
- 傳統規則式控制死板、反饋僵硬
- 使用傳統規則式控制必須取得原始室內光亮度，這一需求使玻璃會出現室內光線閃爍的現象，而閃爍變化量大時對人眼不好正如同ISO標準所強調
- 利用智慧玻璃調整室內光線達成視覺舒適像是一場遊戲，而DRL(深度強化式學習)最擅長遊戲
- 想達成室內光線舒適且保有一定的輕微震盪性模擬自然光

問題陳述

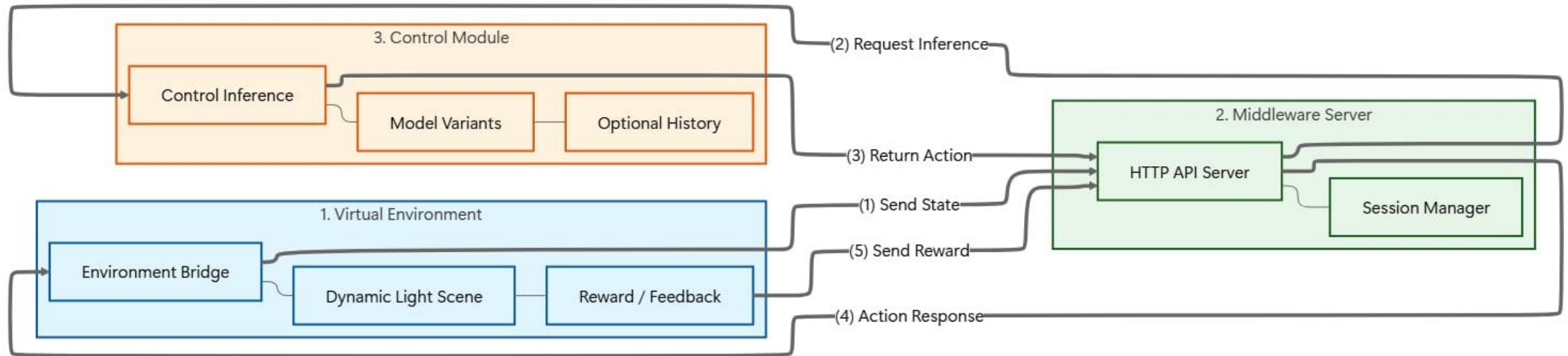
- 模型訓練資料集**不可能**包含所有真實世界情境
- 極端氣候將會是普遍現象，模型能夠適應嗎？
- 什麼樣的模型適合這種工作？
- 我們能怎麼做去壓低模型失誤率？
- 智慧玻璃需要一個表現優良且能在**世界任一角落運行得當**的模型
- 模型如果可以**簡單訓練出良好結果**且**減少成本**會更好

系統架構



- Lux sensor 昂貴且難買，但是引入時序資訊機制需要他才能夠提升模型通用性
- 若不用Lux sensor則通用性可能下降，但成本降低
- 有什麼方法可以提升通用性又降低成本嗎？

方法架構



本專題以虛擬環境和Client-Server架構模擬真實智慧玻璃環境

採用策略

方案一：不需要Lux sensor而降低成本且無須時序資訊機制而模型通用性可能較差

方案二：需要Lux sensor而成本上升但引入時序資訊機制而模型通用性較佳

方案一有8種不同DRL模型

方案二有方案一8種不同DRL模型的對應引入時序資訊機制版本

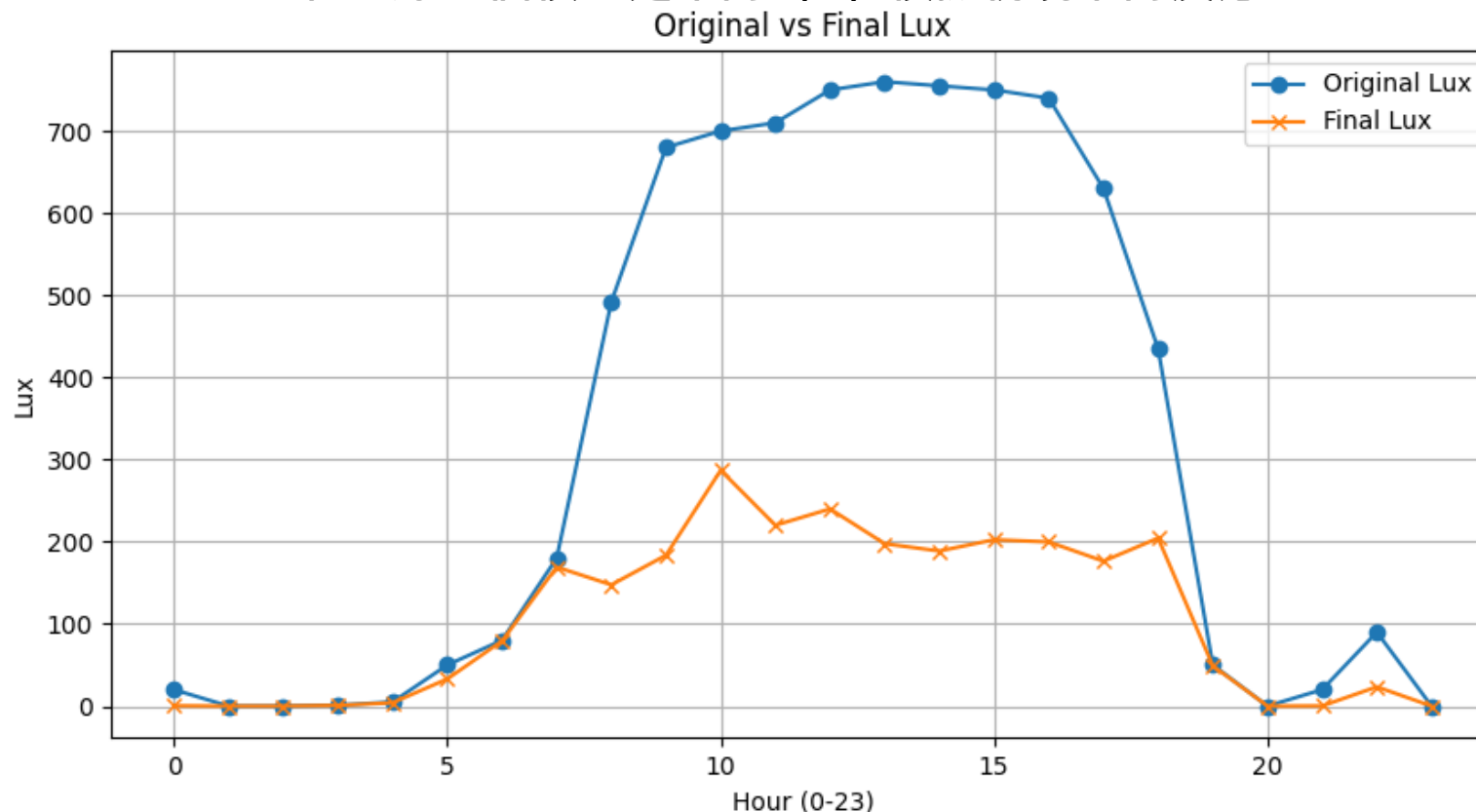
模型訓練結果與成效

在本專題DRL裡環境回饋設為0.0到1.0分，訓練結果顯示：

最終**方案一**DRL訓練**平均約0.84分**的環境回饋，最高分模型有0.88分，而最低的模型0.83分

而最終**方案二**DRL訓練**平均約0.90分**的環境回饋，最高分模型有0.96分，而最低的模型0.85分

下圖為16個模型之中的1位在模擬情境下的反應



Dynamic Demo

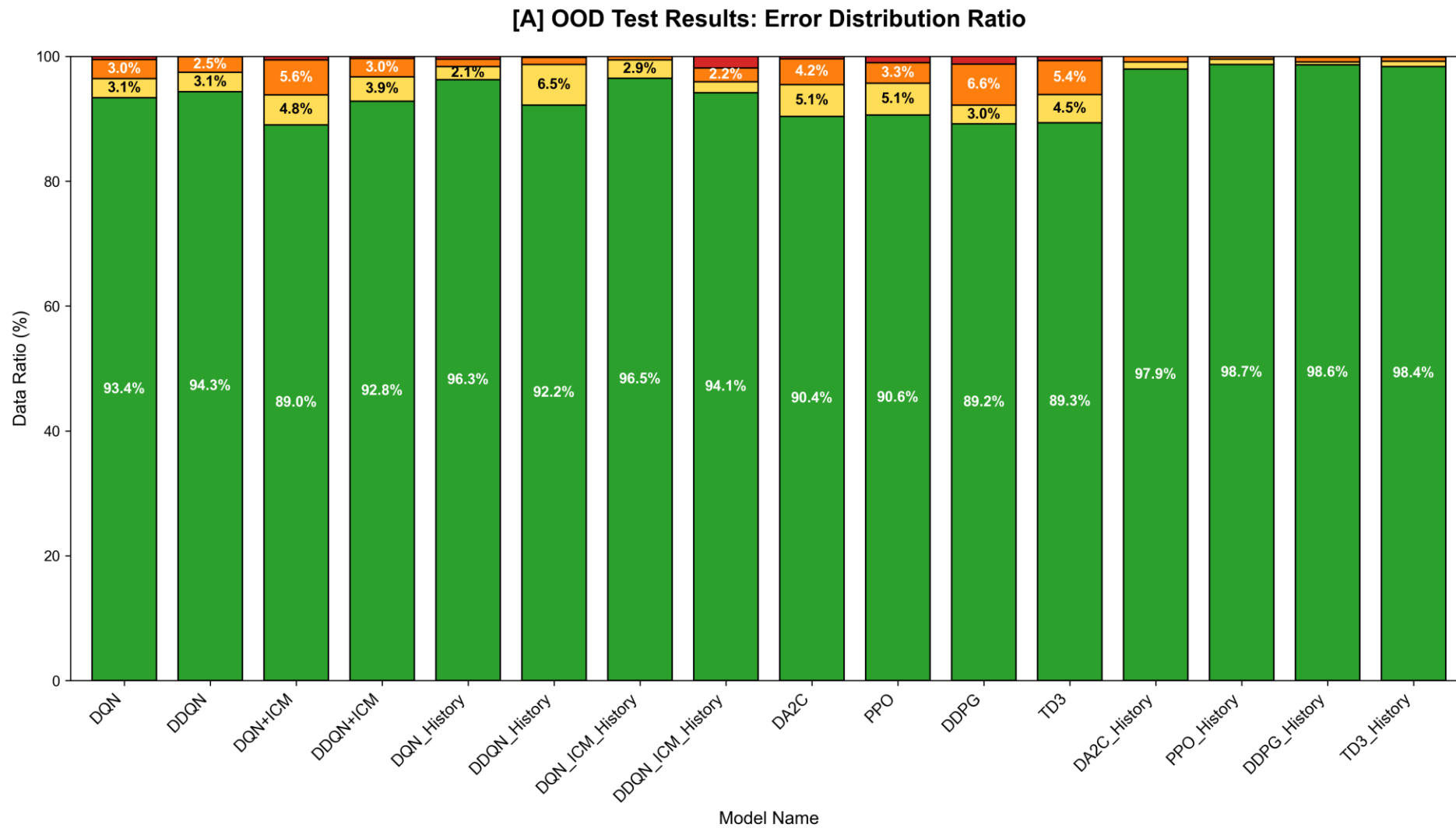
進階測試資料集

Evaluation Datasets and Model Behavior under Progressive Environmental Complexity

Dataset	Environment Type	Noise Intensity	Overall Observation	Implication
A	Baseline Scenario	Low	All methods maintain strong performance	Serves as the reference condition
B	Intermediate Scenario	Moderate	Performance differences begin to appear across methods	Task becomes more sensitive to environmental change
C	Advanced Scenario	High	Robustness outweighs raw inference ability	Methods with stronger adaptability tend to outperform

進階測試資料集結果分析

Dataset A: Baseline



進階測試資料集結果分析

Dataset B and C

- 絕大多數模型**引入時序資訊機制**可以帶來 **5%-20%** 的誤差改善
- 連續控制動作基準模型在極端環境季節下失誤率超過 40%，而引入時序資訊後最低可降至 6% 失誤率
- ICM機制無法幫助模型面對極端氣候探索，甚至對模型帶來破壞性
- 僅方案一DQN比自身方案二通用性效果好

進階測試資料集結果分析

方案一DQN在面對季節或地理平移的資料集時展現出優異的穩健性。在Dataset A與Dataset B中其表現維持在平均水準與多數模型的性能差異不大。然而在處理Dataset C時，方案一DQN的整體誤差率僅有13.9%大幅優於全場平均的28%，位居所有模型之冠。更值得一提的是其誤差組成多屬於「輕度偏移視覺舒適區」的容許範圍。

進一步分析Dataset C的數據可發現，排名第2至第6名的模型不僅整體失誤率較高（第2名較方案一DQN高出約4%），且誤差組成中包含較多「中重度偏移視覺舒適區」的嚴重失誤。對比這幾款模型在Dataset B中的表現，即便它們與方案一DQN的失誤率差距大多在5%以內，方案一DQN依然能將失誤穩定控制在「輕度偏移」的狀態，顯示出其在面對複雜資料分布偏移時，具有更佳的容錯與風險控制能力。

方案一DQN整體表現最佳，且其訓練較其他模型容易、演算法是16個模型中最簡單的、不須Lux sensor而成本低。

結論與未來展望

- 時序資訊機制對大部分模型來說是有有效提升通用性的關鍵技術
- 方案一DQN整體表現最佳且不用時序資訊機制，且其訓練較其他模型容易、演算法是16個模型中最簡單的、不須Lux sensor而成本低。往後研究應當基於DQN延伸。
- 未來須Sim-to-Real 部署將模型移植至實體邊緣裝置，結合真實 lux 感測器與 IoT 設備，驗證實際硬體環境下的控制性能。
- 改良探索機制演算法將會是未來重點。針對環境噪訊干擾問題，探索能區分認知不確定性與環境固有噪訊的改良探索演算法是很重要的，但當前相關技術大多尚未成熟。

Thank you for listening