

When Imitation Learning Outperforms Reinforcement Learning in Surgical Action Planning: A Comprehensive Analysis

Anonymous Author¹ and Anonymous Author²

¹ Anonymous Institution

² Anonymous Institution

Abstract. Surgical action planning requires learning from expert demonstrations while ensuring safe and effective decision-making. While reinforcement learning (RL) has shown promise in various domains, its effectiveness compared to imitation learning (IL) in surgical contexts remains unclear. We conducted a comprehensive comparison of IL versus RL approaches for surgical action planning on the CholecT50 dataset. Our baseline autoregressive transformer achieves strong performance through expert demonstration learning. We systematically evaluated: (1) standard IL with causal prediction, (2) RL with learned rewards via inverse RL, and (3) world model-based RL with forward simulation. Our IL baseline achieves 45.6% current action mAP and 44.9% next action mAP with graceful planning degradation (47.1% at 1s to 29.1% at 10s). Surprisingly, sophisticated RL approaches failed to improve upon this baseline, achieving comparable or slightly worse performance. In surgical domains with expert demonstrations, well-optimized imitation learning can outperform complex RL approaches. This challenges the assumption that RL universally improves upon IL and provides crucial insights for surgical AI development.

Keywords: Surgical Action Planning · Imitation Learning · Reinforcement Learning · Temporal Planning

1 Introduction

Surgical action planning represents one of the most challenging applications of artificial intelligence in healthcare, requiring models to learn from expert demonstrations while ensuring safe and effective decision-making. The question of when to use imitation learning (IL) versus reinforcement learning (RL) in such safety-critical domains remains largely unexplored, despite its importance for practical deployment.

This work provides the first comprehensive comparison of IL and RL approaches for surgical action planning, using the CholecT50 dataset for laparoscopic cholecystectomy. Our key finding challenges conventional wisdom: sophisticated RL approaches fail to improve upon a well-optimized IL baseline, achieving comparable or worse performance across multiple evaluation metrics.

Contributions: (1) First systematic comparison of IL vs RL for surgical action planning, (2) Important negative result showing when RL doesn’t help in expert domains, (3) Comprehensive evaluation framework for temporal surgical planning, and (4) Domain insights about expert data characteristics affecting method selection.

2 Methods

2.1 Baseline: Optimized Imitation Learning

Our IL baseline uses an autoregressive transformer architecture with dual-path training for both current action recognition and next action prediction. The model combines a BiLSTM for temporal current action recognition with a GPT-2 backbone for causal next action prediction.

2.2 RL Approaches Evaluated

Inverse RL with Learned Rewards: We implement Maximum Entropy IRL with sophisticated negative generation to learn reward functions from expert demonstrations.

World Model RL: We develop an action-conditioned world model that predicts future states and rewards given current states and actions.

Direct Video RL: We apply model-free RL directly to video sequences using expert demonstration matching rewards.

3 Results

Table 1 presents our main experimental findings. Our IL baseline achieves 45.6% current action mAP and 44.9% next action mAP, outperforming all RL variants tested.

Table 1. Comparative Performance of IL and RL Approaches

Method	Current mAP	Next mAP	1s Planning	10s Planning
IL Baseline	45.6%	44.9%	47.1%	29.1%
IRL Enhanced	44.2%	43.8%	45.3%	28.7%
World Model RL	42.1%	41.6%	43.8%	27.2%
Direct Video RL	43.9%	43.1%	44.9%	28.1%

4 Discussion

Our results identify several conditions under which imitation learning outperforms reinforcement learning: (1) Expert-optimal demonstrations, (2) Evaluation metric alignment, (3) Limited exploration benefits, and (4) Data sufficiency.

These findings suggest that research resources might be better allocated to optimizing IL approaches rather than developing complex RL systems for surgical planning tasks with expert demonstrations.

5 Conclusion

This work provides crucial insights for surgical AI development by demonstrating that sophisticated RL approaches do not universally improve upon well-optimized imitation learning. Our findings challenge common assumptions about ML method hierarchy and provide practical guidance for surgical AI research resource allocation.

References