



Value-Informed Skill Chaining for Policy Learning of Long-Horizon Tasks with Surgical Robot

Tao Huang*, Kai Chen*, Wang Wei, Jianan Li, Yonghao Long, Qi Dou

The Chinese University of Hong Kong

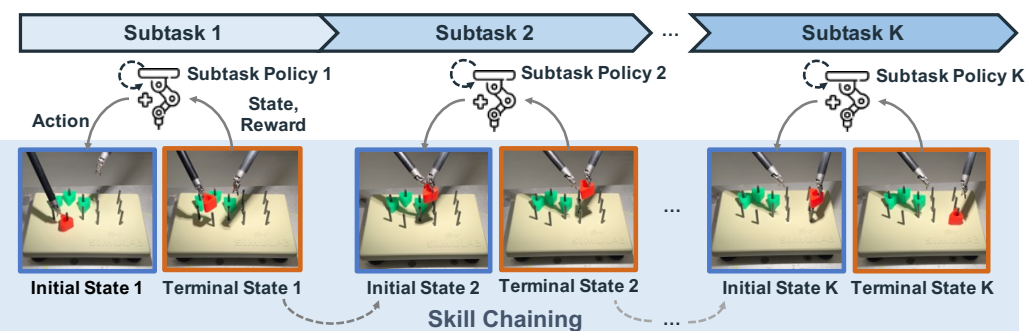
*indicates equal contribution



SCAN ME

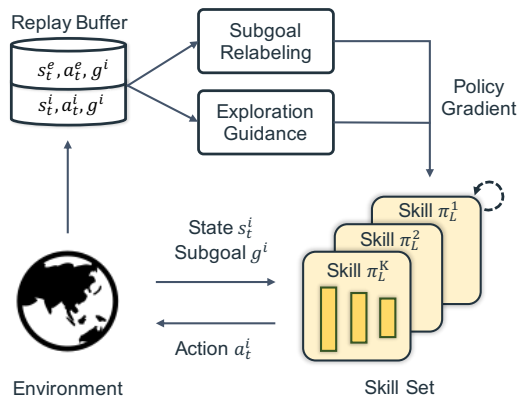


Background

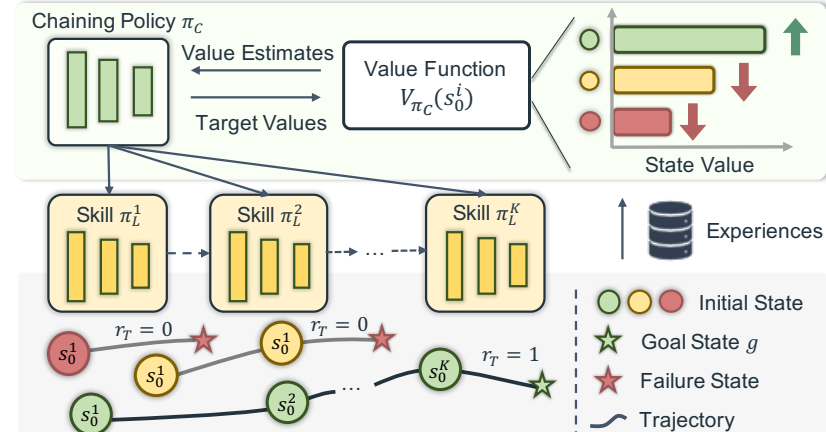


- Skill chaining in decomposable tasks is an effective approach to perform long-horizon manipulation with surgical robot
- Yet, how to chain skill smoothly remains an open problem

Method



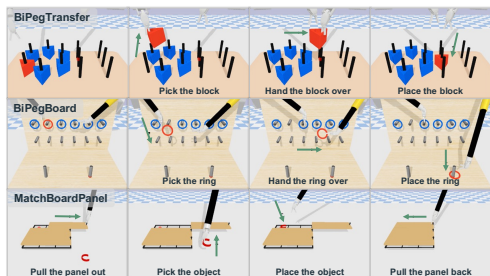
Stage 1: skill learning with demonstrations for solving each subtask separately



Stage 2: skill chaining informed by the learned value function to smoothly connect skills for accomplishing the whole task

Main Results

| Method | BiPegTransfer with 3 subtasks and 100 episode steps | | | BiPegBoard with 3 subtasks and 100 episodes steps | | | MatchBoardPanel with 4 subtasks and 150 episode steps | | |
|----------------|---|------------------------|----------------------|---|------------------------|----------------------|---|------------------------|-----------------------|
| | Succ. Rate (↑) | Subtask Completion (↑) | Rollout Len. (↓) | Succ. Rate (↑) | Subtask Completion (↑) | Rollout Len. (↓) | Succ. Rate (↑) | Subtask Completion (↑) | Rollout Len. (↓) |
| GCBC [36] | 11.27 (±2.85) | 1.40 (±0.05) | 76.26 (±7.31) | 8.26 (±4.78) | 1.33 (±0.08) | 80.22 (±2.55) | 5.50 (±0.99) | 1.36 (±0.03) | 124.47 (±4.67) |
| DEX [19] | 14.03 (±1.45) | 1.55 (±0.11) | 73.42 (±8.24) | 2.60 (±1.15) | 1.23 (±0.04) | 81.76 (±13.27) | - | - | - |
| T-STAR [9] | 67.73 (±5.21) | 2.42 (±0.09) | 65.04 (±8.58) | 65.25 (±9.95) | 2.37 (±0.18) | 56.86 (±10.99) | 45.42 (±4.45) | 2.82 (±0.09) | 94.77 (±3.64) |
| ViSkill (Ours) | 85.24 (±8.42) | 2.73 (±0.15) | 69.19 (±5.72) | 81.76 (±4.08) | 2.67 (±0.07) | 61.21 (±5.22) | 57.09 (±5.68) | 3.07 (±0.12) | 100.39 (±3.66) |



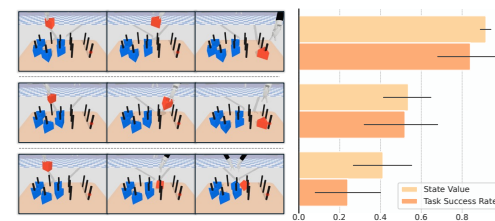
- **Con. 1:** Our method significantly outperform prior IL, flat RL, and skill chaining based methods on three challenging long-horizon surgical robot tasks.
- **Con. 2:** Our method takes more tiny steps to adjust terminal states, ensuring successful execution of all subsequent skills for accomplishing the whole task

Ablation

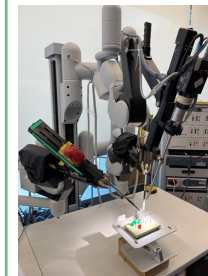
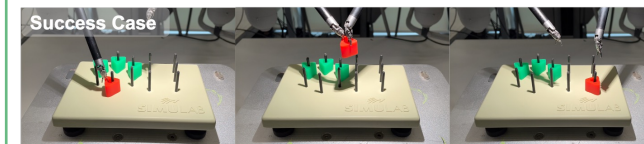
- Ablation on chaining strategy

| Variant | BiPegTransfer | | BiPegBoard | |
|----------------|----------------------|------------------------|----------------------|------------------------|
| | Succ. Rate (↑) | Subtask Completion (↑) | Succ. Rate (↑) | Subtask Completion (↑) |
| ViSkill-DM | 66.23 (±5.20) | 2.37 (±0.09) | 64.13 (±9.95) | 2.23 (±0.15) |
| ViSkill-LDM | 76.20 (±6.56) | 2.56 (±0.10) | 77.50 (±7.90) | 2.59 (±0.16) |
| ViSkill-SR | 61.42 (±4.36) | 2.22 (±0.06) | 60.25 (±6.29) | 2.10 (±0.07) |
| ViSkill (Ours) | 85.24 (±8.42) | 2.73 (±0.15) | 81.76 (±4.08) | 2.67 (±0.07) |

- Analysis of learn value function



Robot Evaluation



- **Con. 3:** The robot sequentially perform the acquired skills with success rate of 80%
- **Con. 4:** The sensitivity to the slight orientation of the block motivates the investigation on sim2real transfer in the future