

Figure 1: Comparison of MICE with CPO in more environments. MICE consistently maintains constraint satisfaction across all tasks, while CPO exhibits a significant number of violations.

Environments	CarGoal1	CarCircle1	PointGoal1	PointCircle1	Ant	HalfCheetah	Humanoid	Hopper
CPO - MICE	61.7%	27.9%	70.7%	19.2%	29.3%	6.4%	12.5%	47.8%

Table 1: The difference in constraint violation rates between CPO and MICE during training.

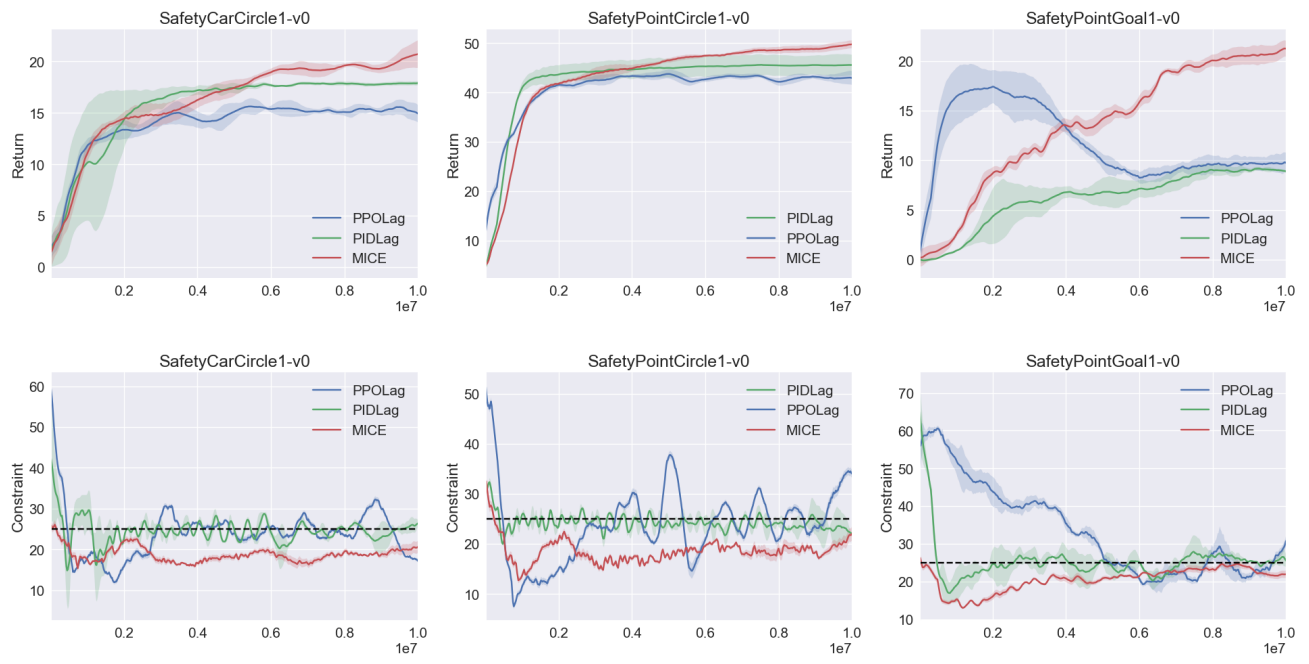


Figure 2: Comparison of MICE under a discount factor of 0.99 with PPO-Lagrangian and PID-Lagrangian across more tasks.