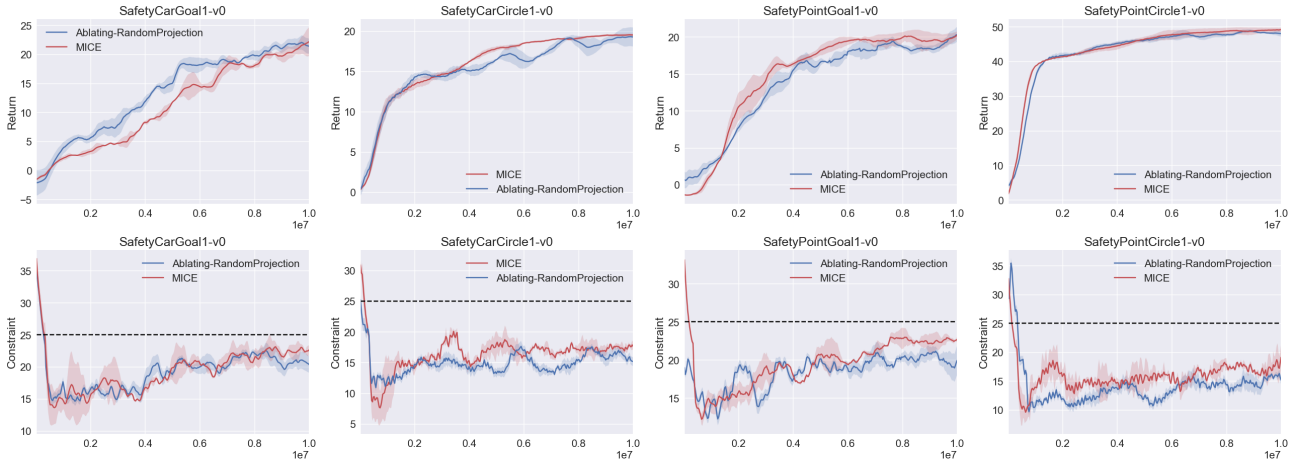

Figure 1: Ablation study of random projection layer in MICE algorithm. Utilizing random projection in MICE does not degrade policy performance or increase constraint violations.

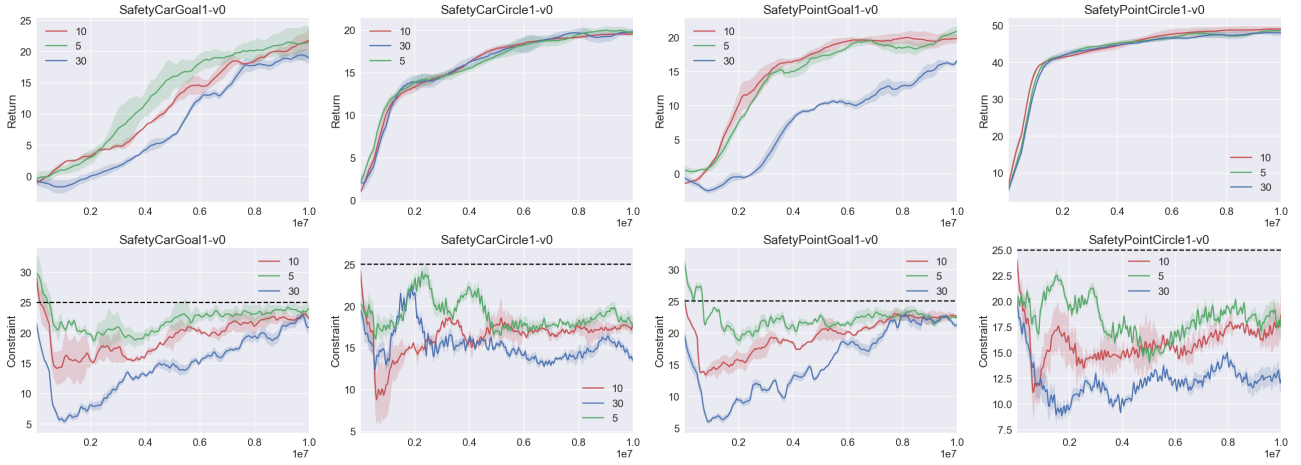

Figure 2: Sensitivity analysis of MICE algorithm for different N_k in KNN. Increasing N_k enhances the safety of the policy but raises computational overhead. In this work, we selected $N_k = 10$ uniformly across all environments to balance safety, performance, and computational efficiency.

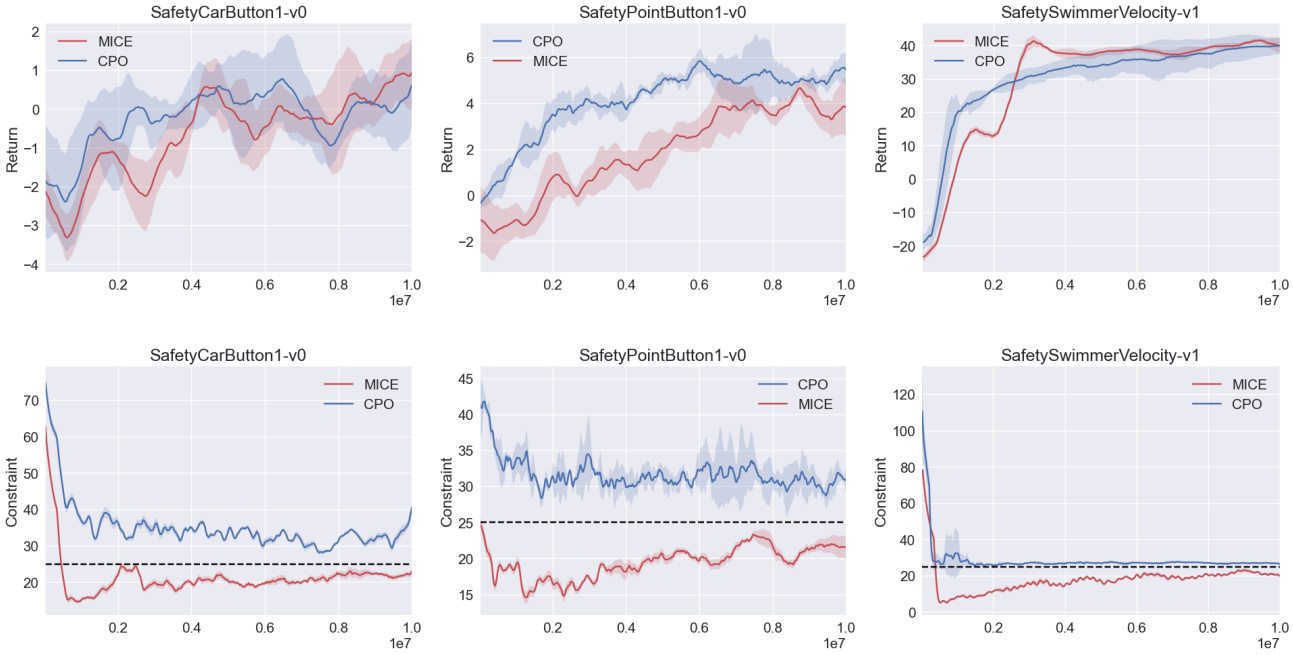

Figure 3: Comparison of MICE with CPO in more environments. MICE consistently maintains constraint satisfaction across all tasks, while CPO exhibits significant violations.

Environments	CarGoal1	CarCircle1	PointGoal1	PointCircle1	Ant	HalfCheetah	Humanoid	Hopper
CPO minus 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, which is computed as the violation rate of CPO minus that of MICE.

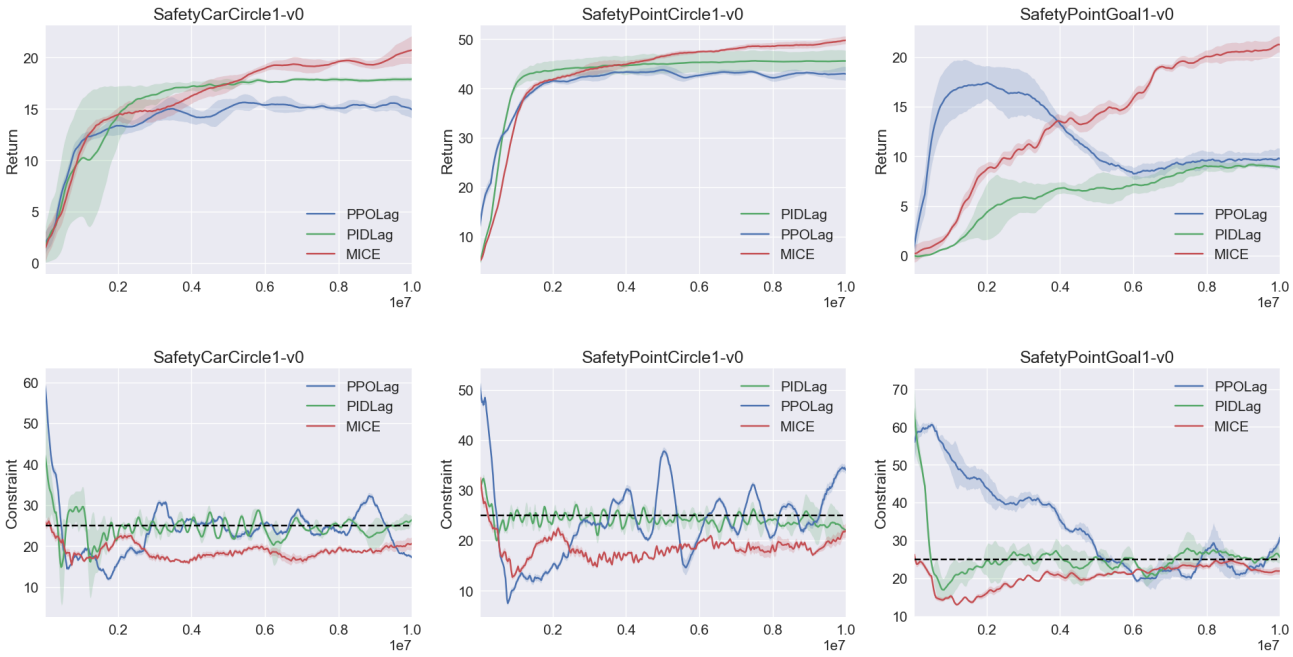

Figure 4: Comparison of MICE under a discount factor of 0.99 with PPO-Lagrangian and PID-Lagrangian across more tasks. MICE consistently achieves constraint satisfaction while maintaining superior policy performance, while PPO-Lagrangian and PID-Lagrangian exhibit significant oscillations.