Learning to Watch: Active Video Anomaly Understanding via Interleaved Policy Optimization

Mengjingcheng Mo · Jiaxu Leng· Xinbo Gao*

[arXiv (2607.00622)] [2-mo.github.io] CQUPT

ICML International Conference On Machine Learning

MOTIVATION

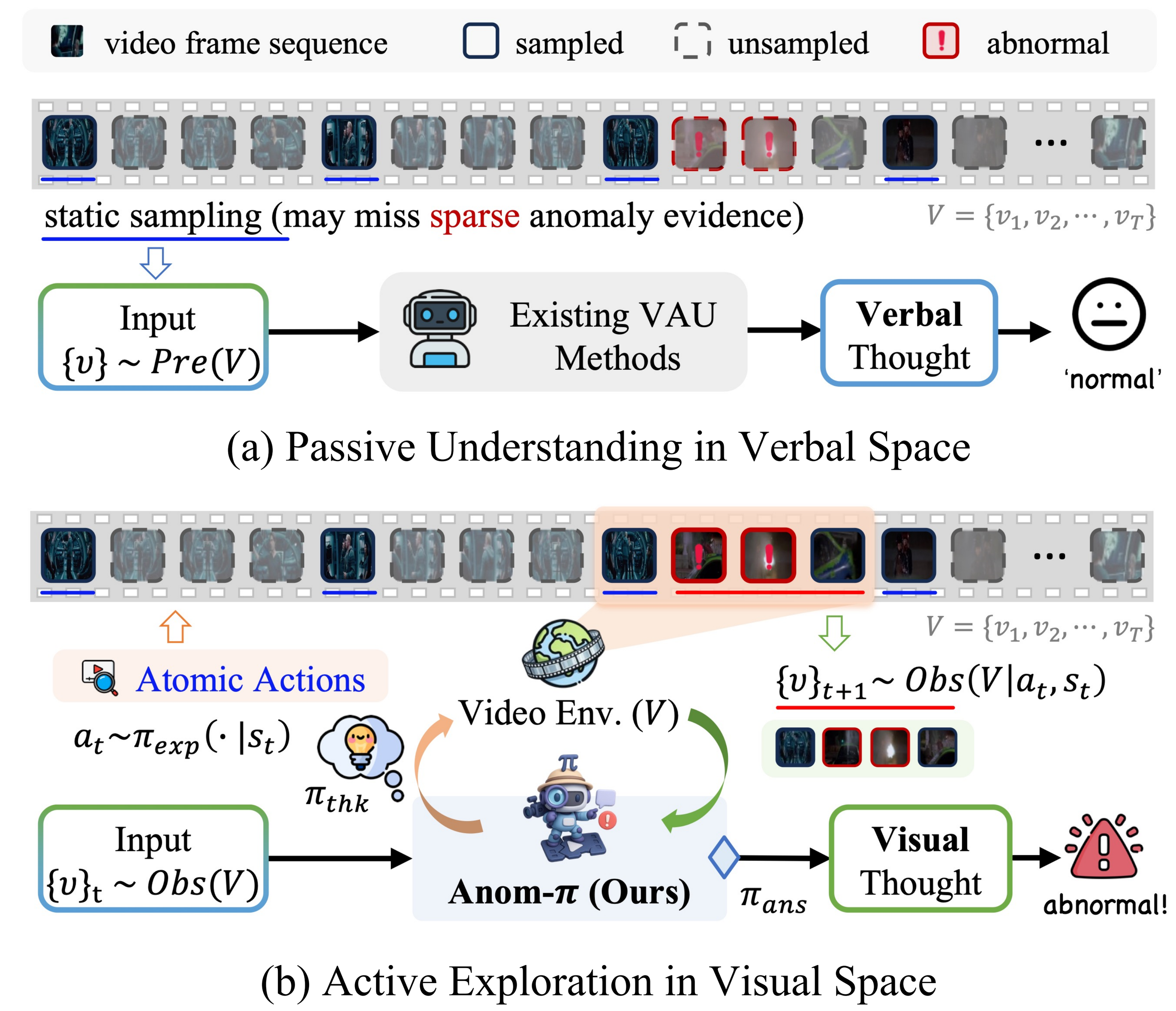

Figure I. Comparison of passive and active evidence acquisition for video anomaly understanding. **(a)** Existing VAU methods use static sampling in the verbal space, which may miss sparse anomaly evidence. **(b)** Anom-π actively explores the video in the visual space to gather informative observations and detect anomalies.

CONTRIBUTIONS

Our main contributions are as follows. **(1)** We propose Anom-π, which reformulates video anomaly understanding as a closed-loop sequential decision-making process with active interaction. **(2)** We design a set of reproducible and composable atomic observation operators, and implement a controllable interact interface via tool calls. **(3)** We propose Interactive Direct Preference Optimization (iDPO), guided by an Active Evidence Inquiry (AEI) utility, to align the interleaved policy π at the trajectory level. **(4)** We validate the effectiveness and generalization of the learned interleaved policy on multiple benchmarks and show that active evidence acquisition improves evidence discovery and decision reliability.

VISUALIZATIONS

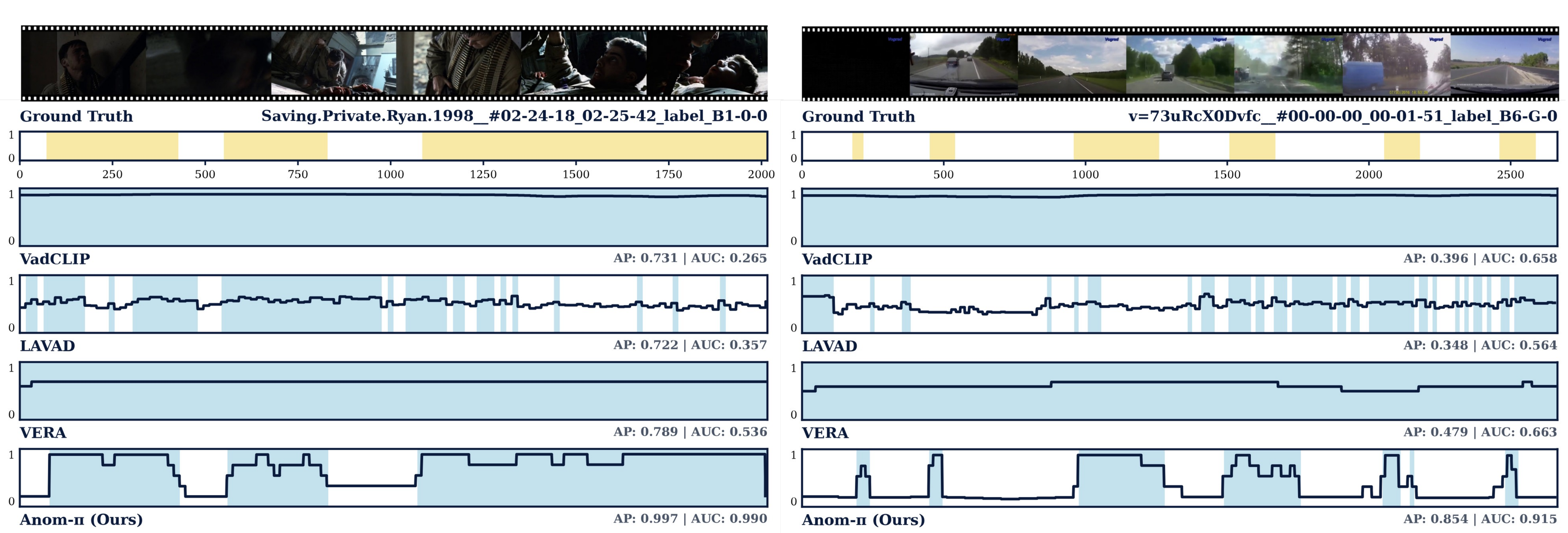

Figure III. Temporal anomaly score curves. We compare Anom-π with VadCLIP, LAVAD, and VERA on three scenarios. Yellow indicates the ground-truth anomalous interval and blue indicates predicted anomalous spans after thresholding the anomaly score at 0.5.

METHOD

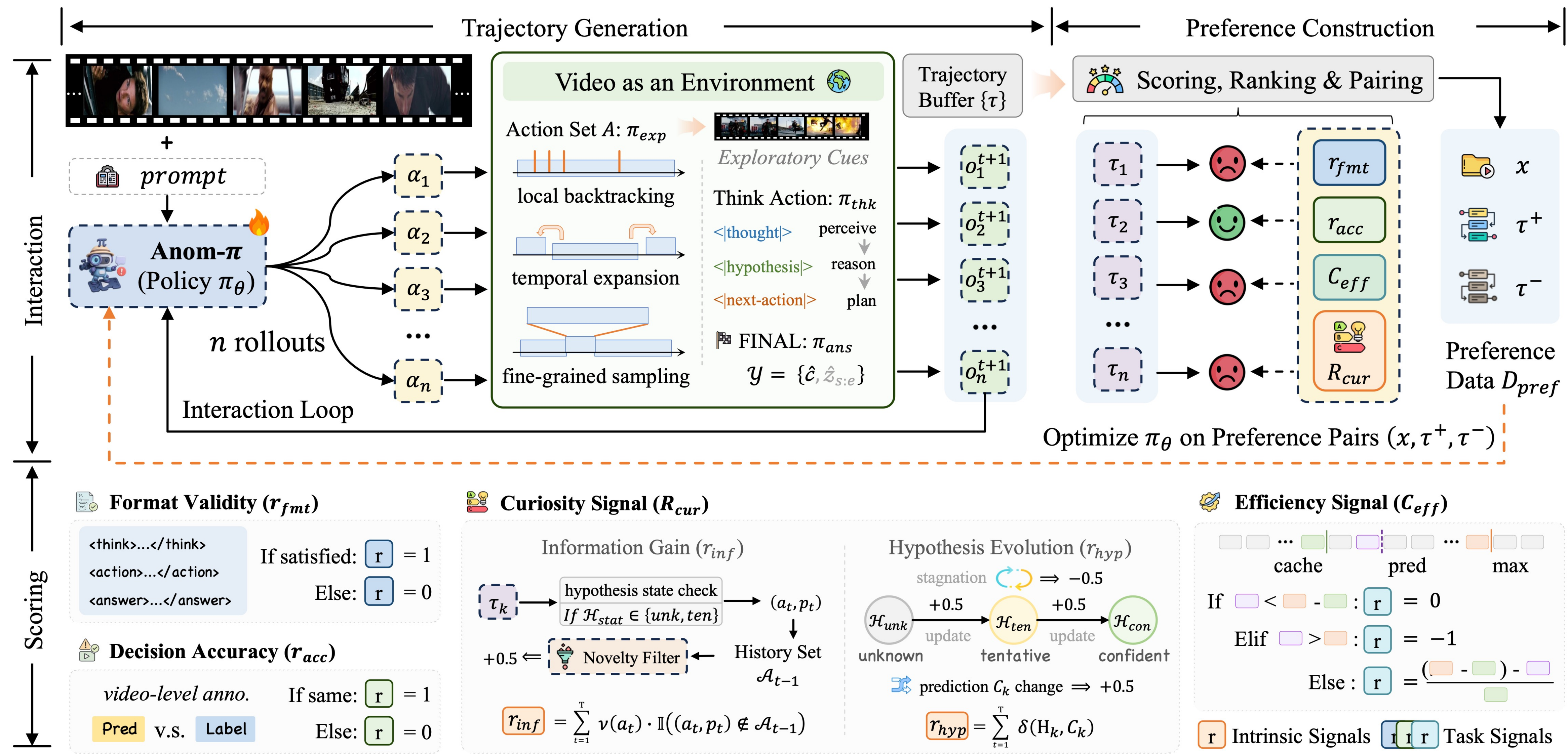

Figure II. Overview of Anom-π. Left: closed-loop active inquiry that alternates deliberation (THINK) with evidence acquisition (BACKTRACK/EXPAND/SAMPLE) before terminating with FINAL to output clip-level hypotheses. Right: trajectory-level preference construction and optimization via iDPO from ranked rollouts.

$$x_t \sim P(\cdot \mid x_{<t}, \{v\}, Q;\ \Theta_{LMM}),\ where\ x_t \in \mathcal{T}_{text} \quad \Rightarrow \quad z_t \sim P(\cdot \mid S_t, V, Q;\ \Theta_{LMM}),\ where\ z_t \in \mathcal{T}_{text} \cup \mathcal{T}_{vis}$$

EXPERIMENTS

Methods	#Params	Expl.	Reas.	Lrn.	Obs.	Multi-Scenario UCF (AUC%)	Multi-Scenario XD (AP%)	Open-Set UB (AUC%)	Complex Scenario CSAD (AUC%)
Passive Specialized Learning									
UR-DMU (Zhou et al., 2023) (AAAI'23)	-	✗	✗	-	✗	86.97	81.66	59.91	-
AED-MAD (Ristea et al., 2024) (CVPR'24)	-	✗	✗	-	✗	-	-	58.50	-
VadCLIP (Wu et al., 2024) (AAAI'24)	-	✗	✗	-	✗	88.02	84.51	-	-
Ex-VAD (Huang et al., 2025) (ICML'24)	-	✓	✗	-	✗	88.29	86.52	-	-
π-VAD (Majhi et al., 2025) (CVPR'25)	-	✗	✗	-	✗	90.33	85.37	-	-
Passive General Understanding									
ZS CLIP (Radford et al., 2021) (ICML'21)	0.3B	✓	✗	✗	Fixed	53.16	17.83	46.20	32.45
LLaVA-1.5 (Liu et al., 2024) (CVPR'24)	13B	✓	✗	✗	Fixed	72.84	50.26	53.71	47.78
LAVAD (Zanella et al., 2024) (CVPR'24)	13B	✓	✗	✗	Fixed	80.28	62.01	64.23	57.26
AnomalyRuler (Yang et al., 2024) (ECCV'24)	17B*	✓	✗	✓	Fixed	-	-	71.90	-
VERA (Ye et al., 2025) (CVPR'25)	8B	✓	✗	✓	Fixed	**86.55**	70.11	-	-
URF (Lin et al., 2026) (NeurIPS'25)	7B	✓	✗	✗	Fixed	84.28	68.07	69.02	-
EventVAD (Shao et al., 2025) (ACM MM'25)	7B	✓	✗	✗	Dyn.	82.03	64.04	-	-
VADTree (Li et al., 2026a) (NeurIPS'25)	8B	✓	✗	✗	Dyn.	84.74	68.85	-	-
PANDA (Yang et al., 2026) (NeurIPS'25)	72B*	✓	✓	✗	Dyn.	<u>84.89</u>	<u>70.16</u>	<u>75.78</u>	<u>73.12</u>
Active Exploratory Reasoning									
Anom-π (ours)	2B	✓	✓	✓	**OnD.**	84.46	**72.29**	**78.75**	**79.18**

Table I. Frame-level performance (AUC/AP) on UCF-Crime, XD-Violence, UBnormal, and CSAD. Expl. indicates whether the method outputs an explanation. Reas. indicates whether explicit reasoning traces are used. Lrn. indicates whether a learned policy is used. Obs. indicates the evidence acquisition setting. An asterisk (∗) indicates a closed-source commercial model.

TRAJECTORYS