



清华大学
Tsinghua University

i-Vision Group

UniGenDet: A Unified Generative-Discriminative Framework for Co-Evolutionary Image Generation and Generated Image Detection

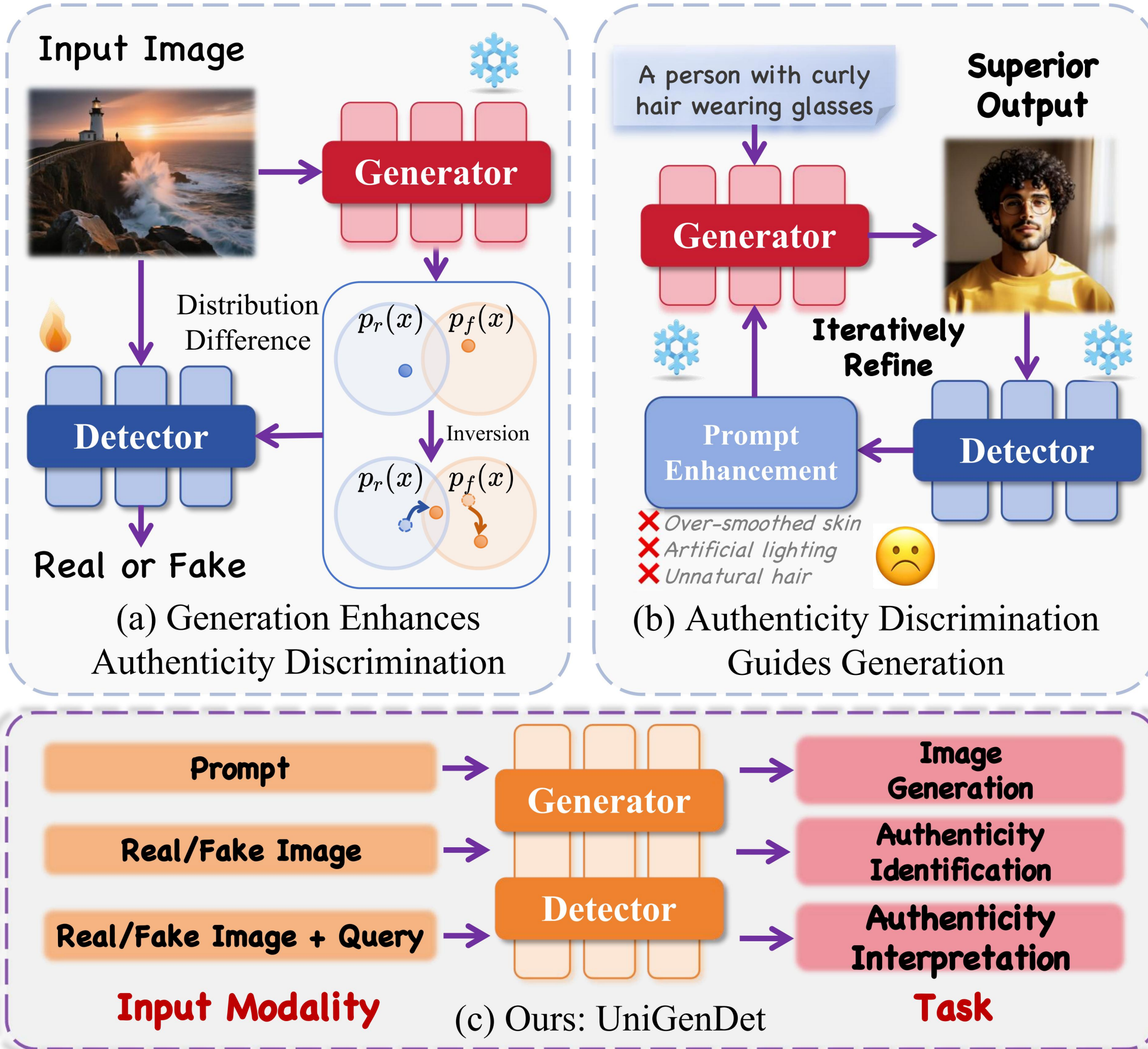
Yanran Zhang, Wenzhao Zheng[†], Yifei Li, Bingyao Yu, Yu Zheng, Lei Chen, Jiwen Lu, Jie Zhou^{*}
Department of Automation, Tsinghua University, China

CVPR
JUNE 3-7, 2026



DENVER
COLORADO

Introduction



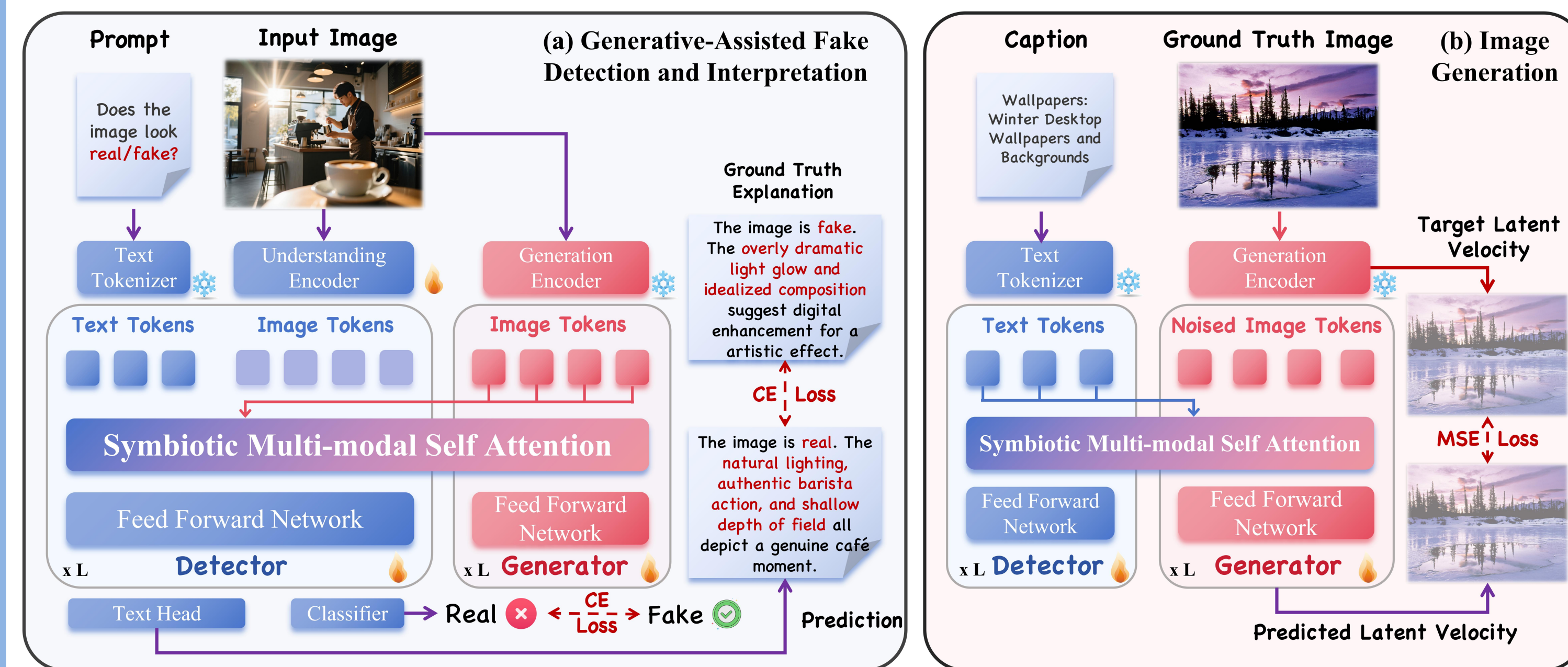
Challenges:

- Generators and detectors evolve **independently** — no closed-loop co-optimization currently exists
- Detectors severely overfit to outputs of known generators and **consequently lag behind** new generative models
- Generators remain unaware of precisely which artifacts make them **easily detectable**

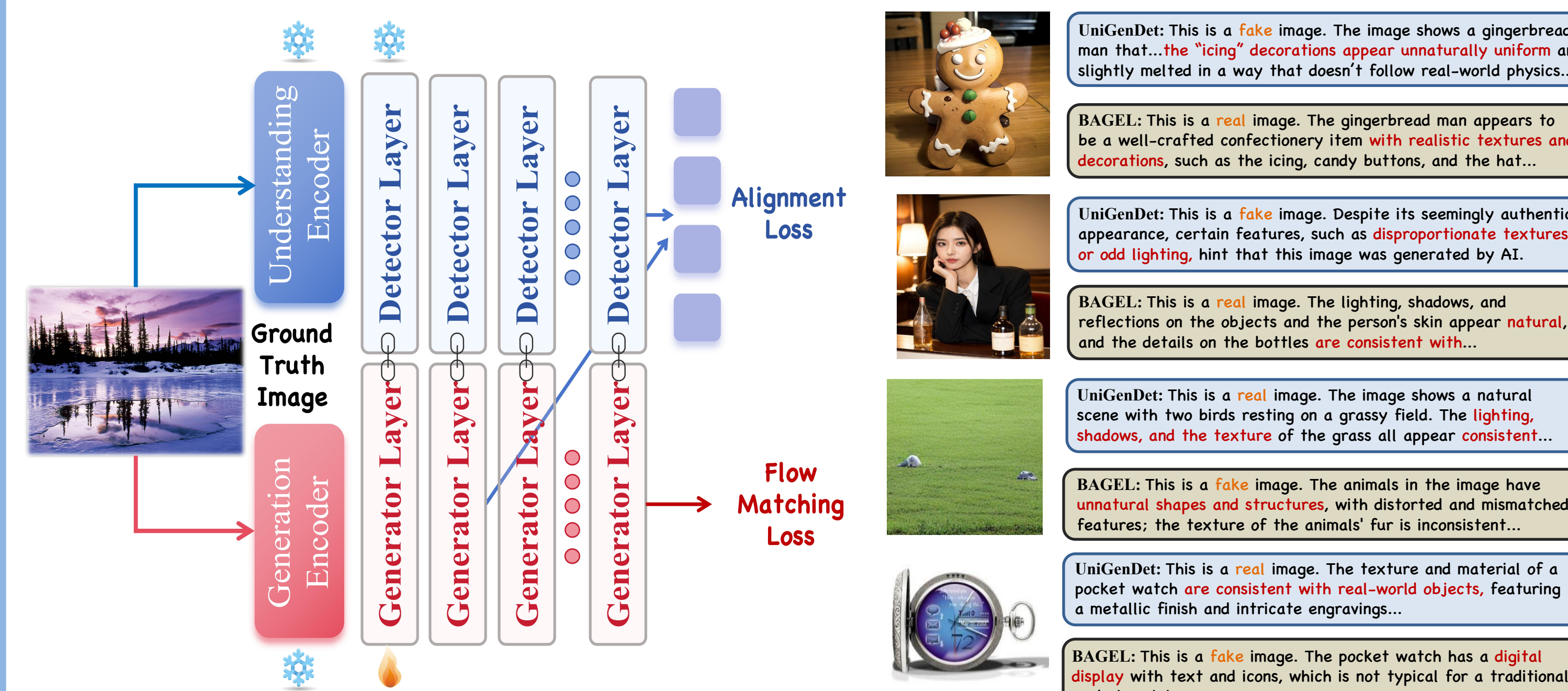
Our Solution — Three Synergistic Advantages

- Task Synergy:** Detection $\leftrightarrow$ Generation mutually improve
- Synchronized:** Evolution Eliminates the "detection lag" problem
- System Efficiency:** Shared parameters, end-to-end training

Methods



- GDUF jointly fine-tunes generation, detection, and explanation in a single framework, letting **generative priors strengthen forensic reasoning and improve interpretability.**



- DIGA injects **detector-aware forensic cues** into the generator through feature alignment, improving realism and reducing detectable artifacts.
- UniGenDet provides more accurate real/fake judgments and more **faithful artifact-level** explanations than BAGEL.

Experiments

- UniGenDet improves robustness and cross-dataset generalization.

Type	Method	FakeClue			
		Acc $\uparrow$	F1 $\uparrow$	ROUGE-L $\uparrow$	CSS $\uparrow$
I	Deepseek-VL2-small [64]	40.4	54.2	17.1	50.4
	Deepseek-VL2 [64]	47.5	54.1	17.2	50.5
	InternVL2-8B [11]	50.6	49.0	18.0	58.1
	InternVL2-40B [11]	50.7	46.3	17.6	55.2
	Qwen2-VL-7B [57]	45.7	59.2	26.6	56.5
	Qwen2-VL-72B [57]	57.8	56.5	17.5	54.4
II	GPT-4o (2024-08-06) [42]	47.4	42.0	13.4	40.7
	BAGEL [14]	40.5	34.1	23.9	46.2
	CNNSpot [58]	43.1	9.8	-	-
	FreqNet [52]	48.7	39.3	-	-
	FatFormer [36]	54.5	45.1	-	-
	UnivFD [41]	63.1	46.8	-	-
	AIDE* [67]	85.9	94.5	-	-
	NPR* [53]	90.2	91.6	-	-
	FakeVLM* [†] [60]	98.6	98.1	32.2	59.5
	Ours*	98.0	97.7	56.3	79.8

Type	Method	LlamaGen Acc.	VAR Acc.	Infinity Acc.	Janus-Pro Acc.	RAR Acc.	Switii Acc.	OM2 Acc.	Mean Acc.
I	CNNSpot [58]	99.9	50.3	50.9	95.7	50.8	56.6	50.1	64.9
	FreDect [18]	99.8	52.9	50.2	88.9	52.5	50.0	57.1	64.5
	Gram-Net [37]	99.6	55.0	52.4	74.5	50.0	57.7	50.1	62.8
	LNP [35]	99.5	49.6	49.8	99.5	49.7	70.3	49.6	66.9
	UnivFD [41]	89.9	80.5	71.7	84.3	88.3	76.0	66.2	79.6
	NPR [53]	100.0	56.9	88.5	93.7	52.3	52.0	63.0	72.3
II	D ³ QE [72]	97.2	85.3	62.9	92.3	91.7	75.3	70.1	82.1
	FakeVLM [60]	98.1	97.7	99.4	99.9	99.9	94.6	90.4	97.1
	UniGenDet	89.4	99.7	99.9	99.7	99.5	98.8	99.5	98.1

Method	Real		Fake		Overall	
	Acc	F1	Acc	F1	Acc	F1
CNNSpot [58]	87.8	88.4	28.4	44.2	40.6	43.3
Gram-Net [37]	62.8	54.1	78.8	88.1	67.4	79.4
Fusing [25]	87.7	86.1	15.5	27.2	40.4	36.5
LNP [5]	63.1	67.4	56.9	72.5	58.2	68.3
UnivFD [41]	89.4	88.3	44.9	61.2	53.9	60.7
AntifakePrompt [6]	91.3	92.5	89.3	91.2	90.6	91.2
SIDA [23]	92.9	93.1	90.7	91.0	91.8	92.4
FakeVLM [60]	98.2	99.1	89.7	94.6	94.0	94.3
UniGenDet	99.0	99.5	97.2	98.6	98.6	99.1

Summary

- A unified framework for the **co-evolution of generation, detection, and explanation.**
- Bidirectional** feedback between generation and detection improves realism and forensic reasoning.
- Toward **trustworthy co-evolution** of generative models and forensic analysis for future AI systems.

- Detection-guided training improves realism and coherence.

	Model	BAGEL	BAGEL+GDUF	UniGenDet
	FID $\downarrow$	22.9	19.4	17.5
Input Prompt	Elegant woman in feather-trimmed hat, timeless profile.			
	Pagoda and Mount Fuji, a timeless Japanese scene.			
BAGEL	A vibrant and wholesome quinoa power bowl.			
	A mirror of pure serenity in the mountains.			
UniGenDet	Sleek wood grain and minimalist brass handles.			
	Jagged peaks tower over serene lake and green meadow.			
BAGEL	Hugh Grant, captured in a moment of quiet intensity.			
	Rainbow over the desert landscape.			
UniGenDet				

- Ablation study

Model	Acc $\uparrow$	F1 $\uparrow$	Rouge-L $\uparrow$	CSS $\uparrow$
w/o GDUF	40.5	34.1	23.9	46.2
w/o SMSA	95.0	94.6	50.9	77.7
UniGenDet	98.0	97.7	56.3	79.8

LinkedIn



Github



Project

