ALBIREO

Adaptive, Energy-Efficient Video Object Detection on the Edge

Amir Taherin[1] José Cano[2] Bin Ren[3] Yanzhi Wang[1] David Kaeli[1]

[1]Northeastern University [2]University of Glasgow [3]William & Mary

Northeastern University


AGX Orin


AGX Thor

Problem: detector cost dominates edge video

Per-frame YOLO26x on BDD100K MOT draws **49% of board power on Thor** (34/70 W) and **72% on Orin** (31/43 W). Temporal redundancy invites frame-skipping, but **content-blind skipping fails**: FixedSkip-2 **−8.6 pp AP@50**, FixedSkip-5 **−28.8 pp**.

Goal: a drop-in efficiency layer

A frame-skipping wrapper that is *all of*: **detector-agnostic**, **training-free**, **codec-free**, **ego-motion robust**, and **energy-efficient on edge SoCs**. No prior system satisfies all five.

Where prior work falls short

Method	DA	NDR	EM	CF	EE	POA
DFF	×	×	✓	✓	×	×
Statues	✓	✓	×	✓	✓	×
CTD	✓	✓	✓	✓	×	✓
FrameHopper	✓	×	✓	✓	×	×
Compressed	✓	×	✓	×	×	×
ALBIREO	✓	✓	✓	✓	✓	✓

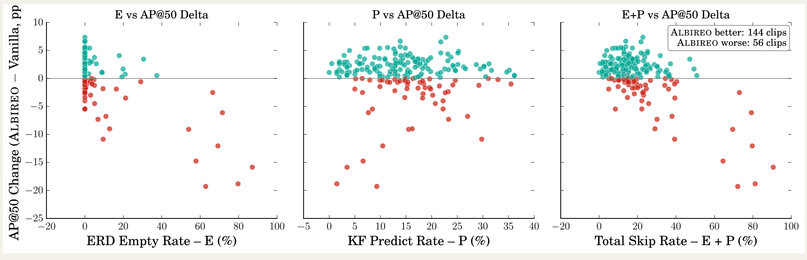
DA: detector-agnostic. **NDR**: no retraining. **EM**: ego-motion. **CF**: codec-free. **EE**: energy on edge SoC. **POA**: per-object adaptive.

Ablation: rescue drives the AP@50 gain

Variant	AP@50	mAP	Skip%	ms/fr
Kill-only (8D)	0.578	0.358	21.5	41.8
+ ERD	0.572	0.355	25.8	42.9
+ 10D KF	0.577	0.407	26.3	42.5
ALBIREO (+rescue)	**0.617**	**0.433**	19.3	46.4

Rescue lifts AP@50 from the 0.572–0.578 no-rescue band to $\mathbf{0.617}$.

Per-clip behavior: P-skips are safe



P-frames stay flat across 0–36%; large drops only on high-E clips (sparse scenes). ALBIREO $\geq$ Vanilla AP@50 on **144/200 (72%)** clips.

Conclusion

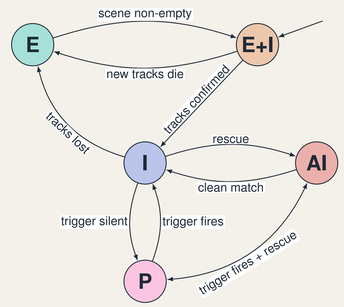
- First **detector-agnostic, training-free, codec-free, content-aware** frame-skipping system measured on commercial Jetson SoCs.
- Rescue dominates the AP@50 gain; ERD pays off only when gated on no-active-track frames.
- $+0.8$ **pp AP@50,** -17.6% **energy,** -24.9% **EDP** on YOLO26x/Thor; consistent across detectors and platforms.

ALBIREO: three composable mechanisms

1. **Uncertainty trigger.** Per-object 10-D Kalman filter; skip the detector when the projected covariance trace $\tau{=}\mathrm{tr}(HPH^\top){<}u$ for every active object. Forward-looking, unlike CTD's stale-detection score.
2. **Detector-flicker rescue.** Confirmed-but-unmatched tracks emit their KF-predicted box when last real-match confidence ≥ 0.5 and rescue count $< R_{\max}{=}2$.
3. **Empty-scene screen (ERD).** When no active tracks exist, a $\sim$14k-parameter CNN gates the detector in $\sim$1.3 ms.

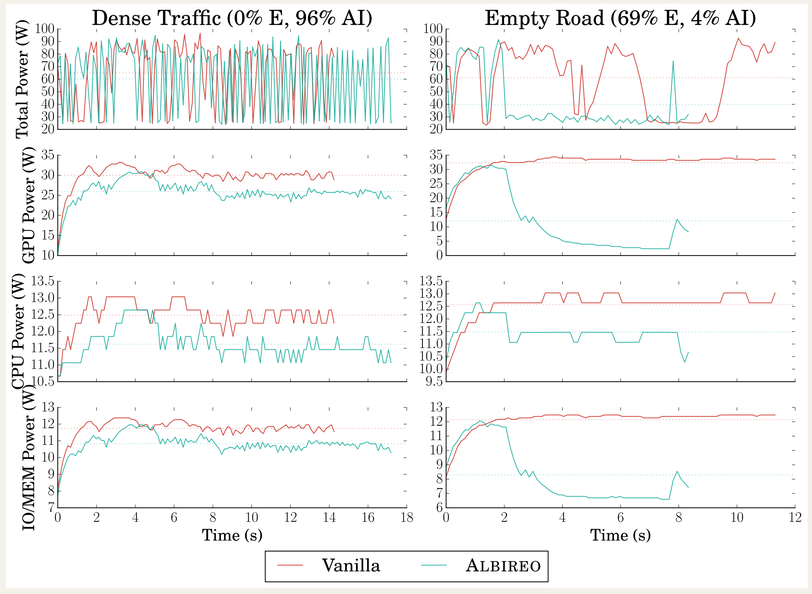
u is the *single* accuracy–efficiency knob; the wrapped detector is never modified or retrained.

Per-frame state machine



Five states: **E** ERD-skip, **P** predict-only, **E+I** ERD-triggered, **I** inference, **AI** augmented inference.

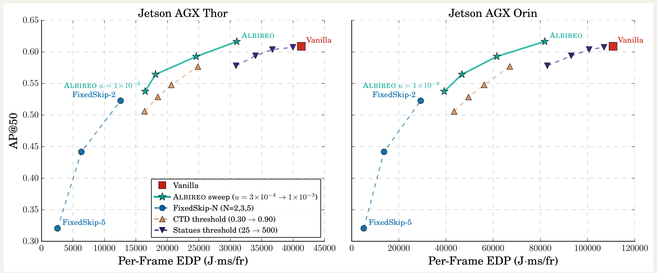
Two operating modes on representative clips



Dense urban (96% AI): rescue preserves AP@50. **Empty road** (69% E): long ERD-only stretches drop GPU power. Same algorithm, two distinct power profiles.

The name. ALBIREO *is a contrasting binary star in Cygnus — bright component ↔ full detector; dim companion ↔ KF prediction.*

Pareto frontier (YOLO26x, BDD100K MOT)



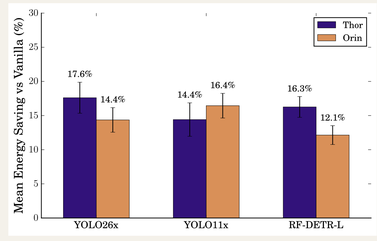
Sweeping u traces the upper Pareto frontier over every evaluated baseline in AP@50 ≥ 0.50, on both Jetson generations.

Headline: strict 3-axis Pareto dominance

Plat.	System	AP@50	Energy (J)	ms/fr	EDP
Thor	Vanilla	0.609	811.5	50.9	41,305
Thor	ALBIREO	**0.617**	**668.6**	**46.4**	**31,023**
	Δ	**+0.8 pp**	**−17.6%**	**−8.8%**	**−24.9%**
Orin	Vanilla	0.609	994.0	111.5	110,831
Orin	ALBIREO	**0.617**	**851.2**	**96.2**	**81,885**
	Δ	**+0.8 pp**	**−14.4%**	**−13.7%**	**−26.1%**

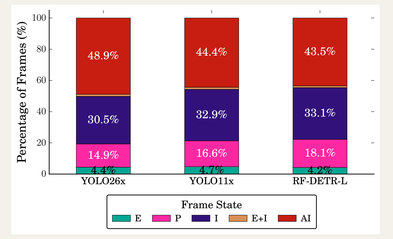
Default $u{=}3{\times}10^{-4}$ (19.3% skip): higher accuracy, lower energy, lower latency *together*.

Detector-agnostic



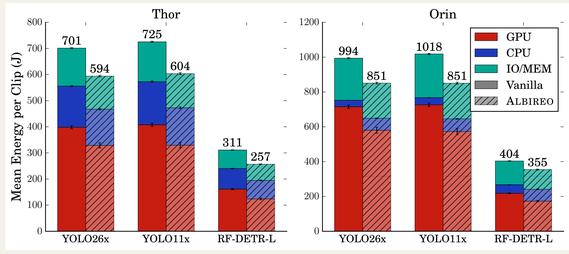
AP@50 within ± 1.2 pp of Vanilla; energy -12 to -18% across all six detector×platform configurations.

Frame-state mix



19–22% skip (E+P); 43–49% AI rate = rescue frequency.

Per-domain energy: GPU drives the savings



GPU energy -17.5% on Thor / -19.0% on Orin. Orin CPU power rises $1.6 \rightarrow 3.6$ W (KF/match/rescue bookkeeping); Thor's 14-core CPU absorbs it.