MobiSys 2020
The 18th ACM International Conference on
Mobile Systems, Applications, and Services

Approximate Query Service on Autonomous IoT Cameras

Mengwei Xu[1], Xiwen Zhang[2], Yunxin Liu[3]
Gang Huang[1], Xuanzhe Liu[1], Felix Xiaozhu Lin[2]

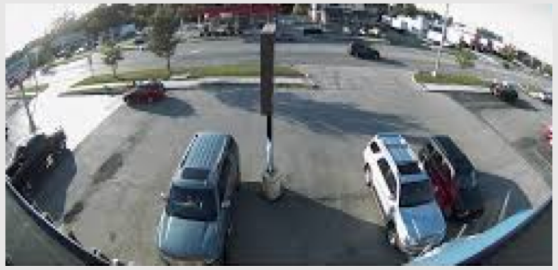
[1]Peking University, [2]Purdue University, [3]Microsoft Research,

PURDUE UNIVERSITY

Microsoft Research

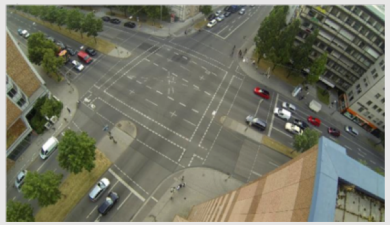
Video Analytics is a Killer App 🔥

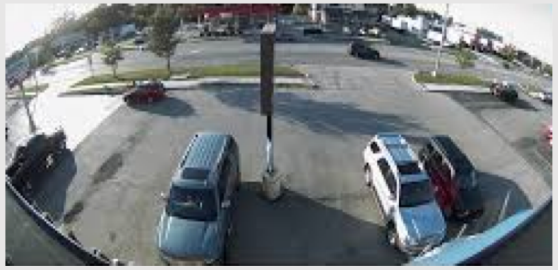
- Busy cross roads
- Retailing store
- Sports stadium
- Parking lots
- …



Video Analytics is a Killer App 🔥

- Busy cross roads
- Retailing store
- Sports stadium
- Parking lots
- …



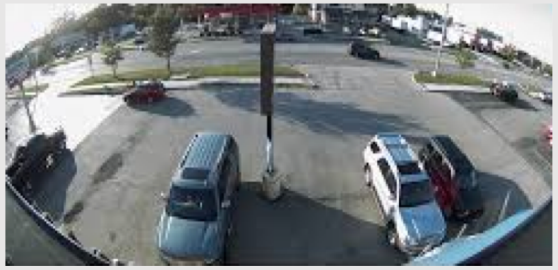


Urban, residential areas

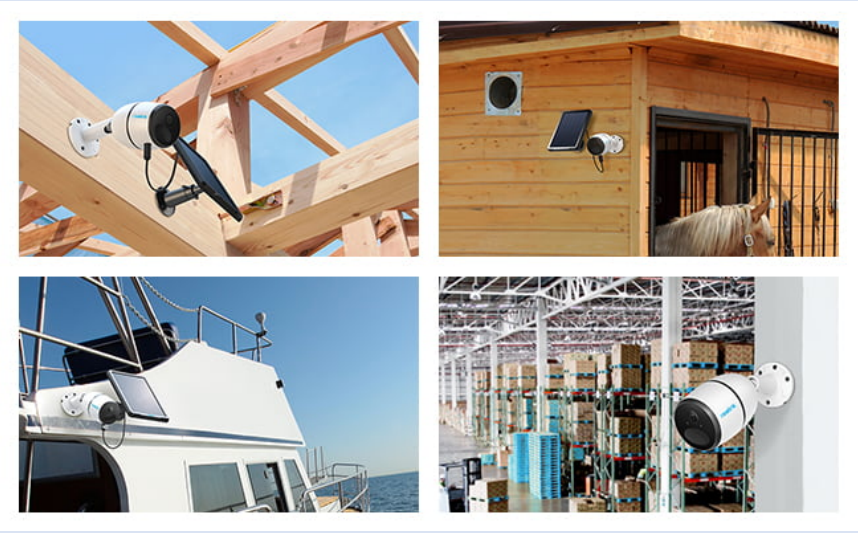
✓ Wired electricity
✓ Good internet

Video Analytics is a Killer App 🔥

- Busy cross roads
- Retailing store
- Sports stadium
- Parking lots
- …



?



- Construction sites
- Cattle farms
- Highways
- Wildlifes
- …

Urban, residential areas

Rural, off-grid areas

Autonomous Camera

- **Energy-independent** and **Compute-independent**

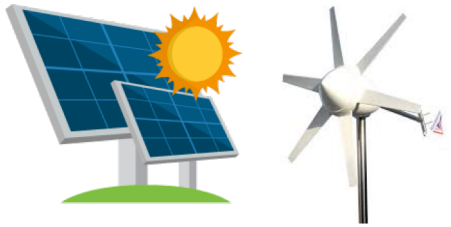
Autonomous Camera

- **Energy-independent** and **Compute-independent**

Commodity SoCs, RPI-like,
chargeable battery

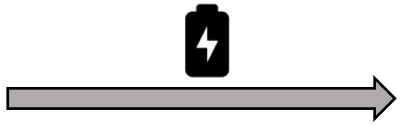
Autonomous Camera

- **Energy-independent** and **Compute-independent**



Small-sized energy harvester

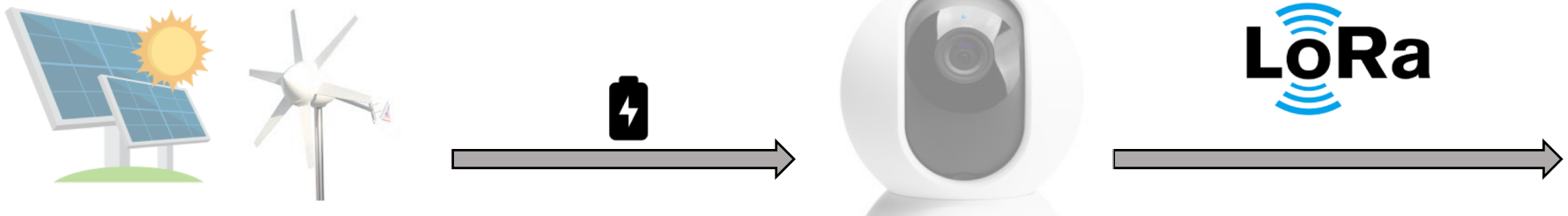
e.g., "10Wh today"



Commodity SoCs, RPI-like, chargeable battery

Autonomous Camera

- **Energy-independent** and **Compute-independent**



Small-sized energy harvester

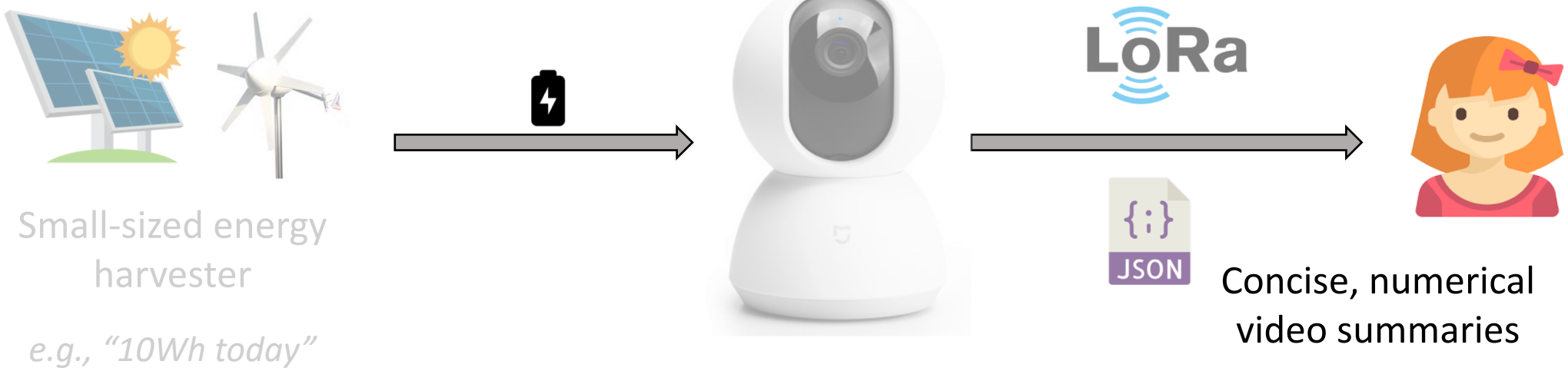
e.g., "10Wh today"

Commodity SoCs, RPI-like, chargeable battery

LPWAN like LoRaWAN,
At most a few Kbps

Autonomous Camera

- **Energy-independent** and **Compute-independent**



Elf for Autonomous Cameras

- Target video query: **object counting**

Elf for Autonomous Cameras

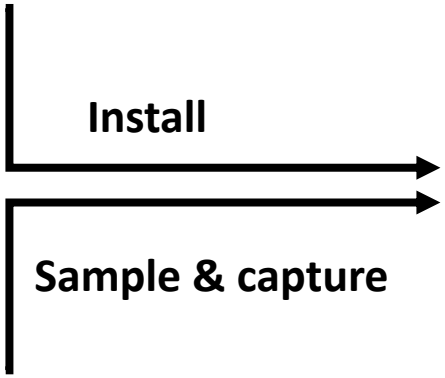
- Target video query: **object counting**

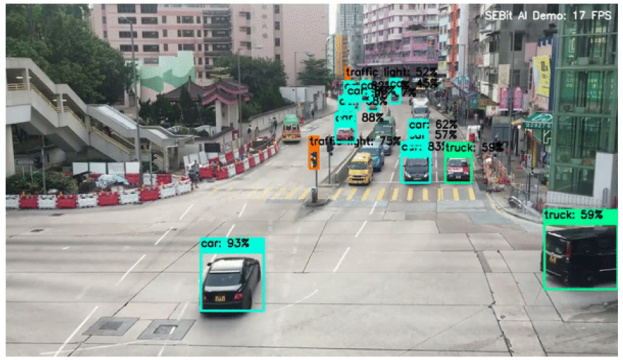
Query: (car, 30 mins)

Elf for Autonomous Cameras

- Target video query: **object counting**

Query: (car, 30 mins)





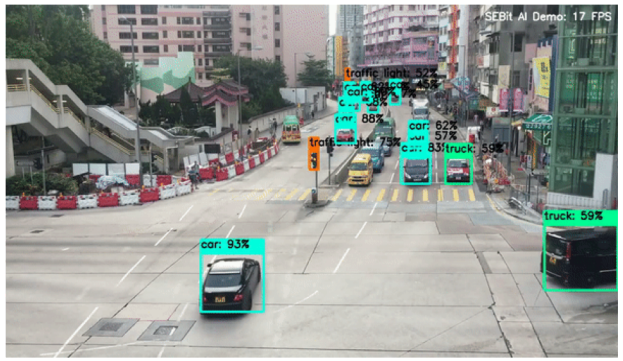
Elf for Autonomous Cameras

- Target video query: **object counting**

Query: (car, 30 mins)

Install

Sample & capture



7:00AM-7:30AM [500 ± 100] Cars

7:30AM-8:00AM [700 ± 140] Cars

8:00AM-8:30AM [800 ± 180] Cars

8:30AM-9:00AM [400 ± 100] Cars

9:30AM-10:00AM [200 ± 80] Cars

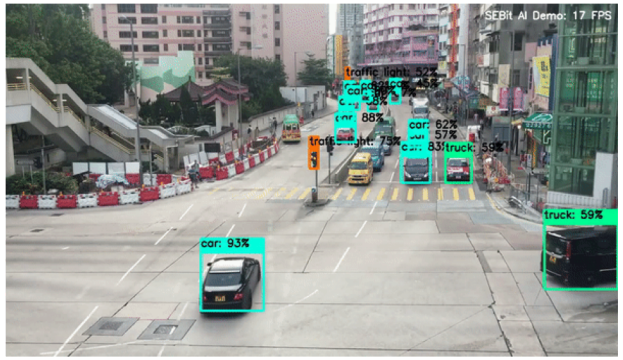
Elf for Autonomous Cameras

- Target video query: **object counting** with confidence interval (CI)

Query: (car, 30 mins)

Install

Sample & capture



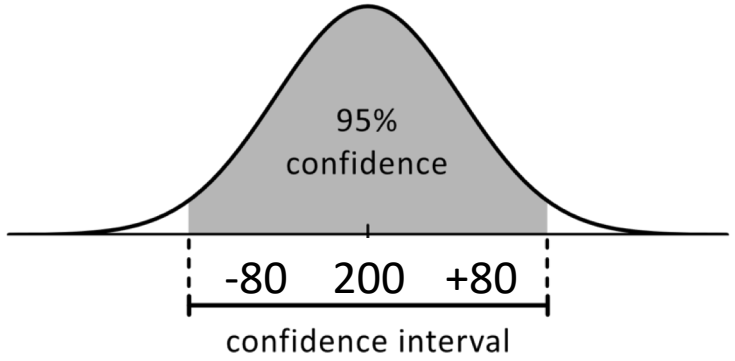
7:00AM-7:30AM [500 ± 100] Cars

7:30AM-8:00AM [700 ± 140] Cars

8:00AM-8:30AM [800 ± 180] Cars

8:30AM-9:00AM [400 ± 100] Cars

9:30AM-10:00AM [200 ± 80] Cars

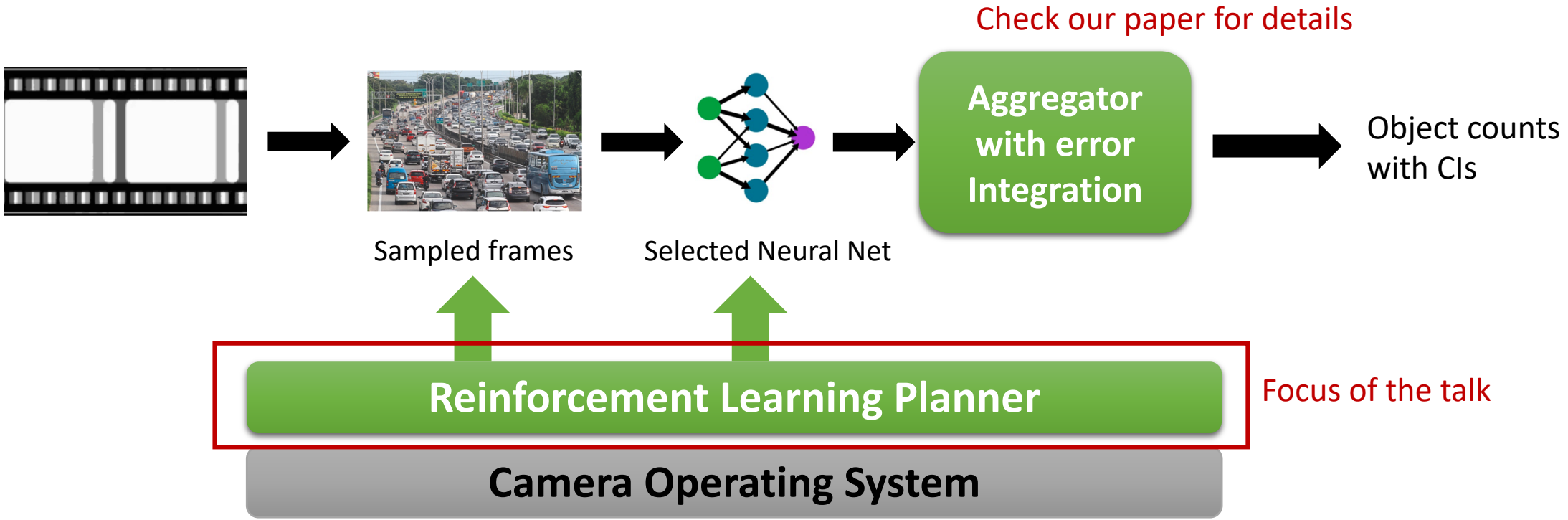


Elf for Autonomous Cameras

- Target video query: **object counting** with confidence interval (CI)

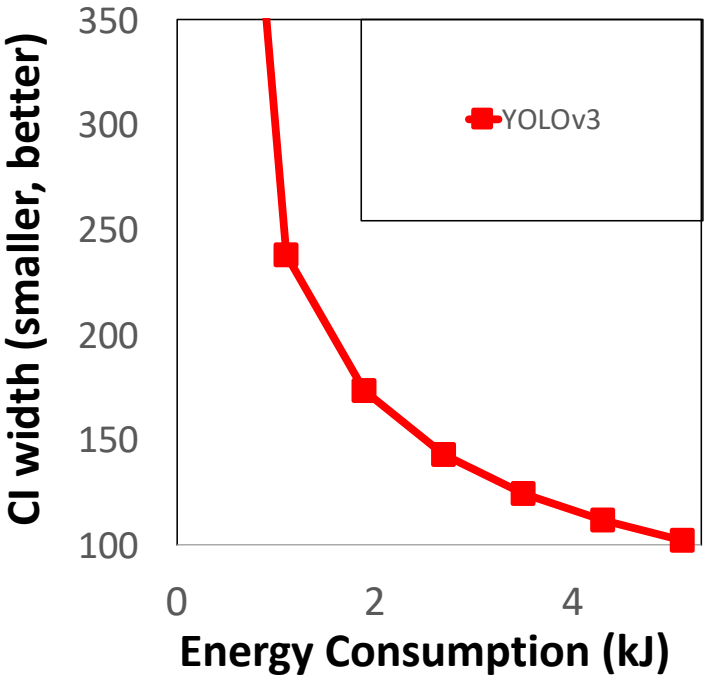
- The central problem: **planning constrained energy for counting**
  - Energy model: a budget that cannot be exceeded in a horizon (e.g., 24 hrs)
  - Trade-offs: frame sampling and NN selection
  - Target: smallest mean CI widths across all (30-min) windows in a horizon

Elf Overview



Elf tech #1: per-window energy/CI fronts

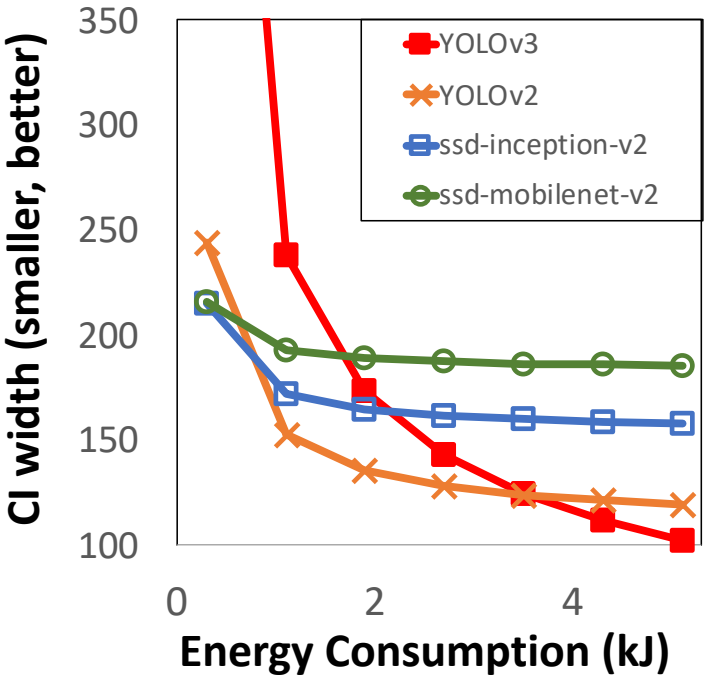
- What's the best count action for a window?
  - A *count action*: determining (1) an NN and (2) # of frames to process



$$\text{Energy Consumption} = E(\text{NN}) * \text{frame\_num}$$

Elf tech #1: per-window energy/CI fronts

- What's the best count action for a window?
  - A *count action*: determining (1) an NN and (2) # of frames to process

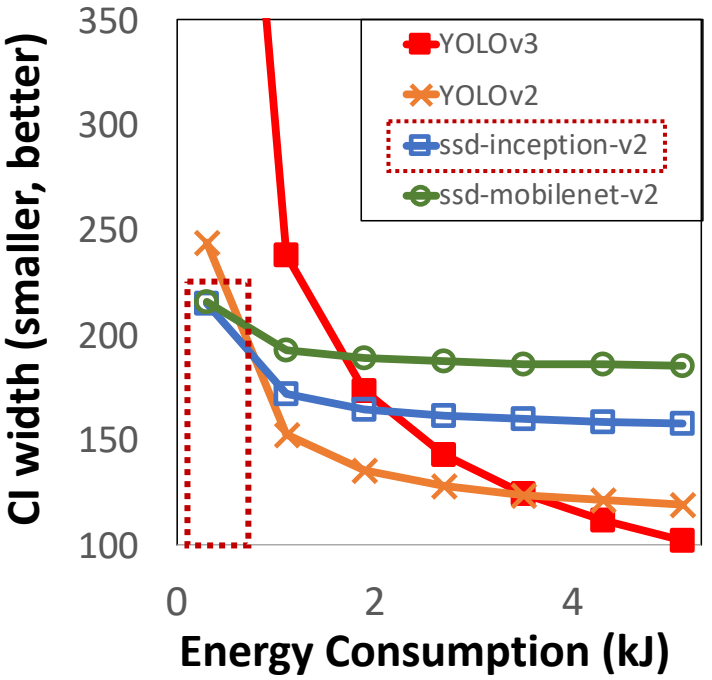


$$\text{Energy Consumption} = E(\text{NN}) * \text{frame\_num}$$

NN Counters	Input	mAP	Energy
YOLOv3 (Golden, GT) [85]	608x608	33.0	1.00
YOLOv2 [84]	416x416	21.6	0.22
faster rcnn inception-v2 [86]	300x300	28.0	0.40
ssd inception-v2 [68]	300x300	24.0	0.08
ssd mobilenet-v2 [88]	300x300	22.0	0.05
ssdlite mobilenet-v2 [88]	300x300	22.0	0.04

Elf tech #1: per-window energy/CI fronts

- What's the best count action for a window? No silver bullet.
  - A *count action*: determining (1) an NN and (2) # of frames to process

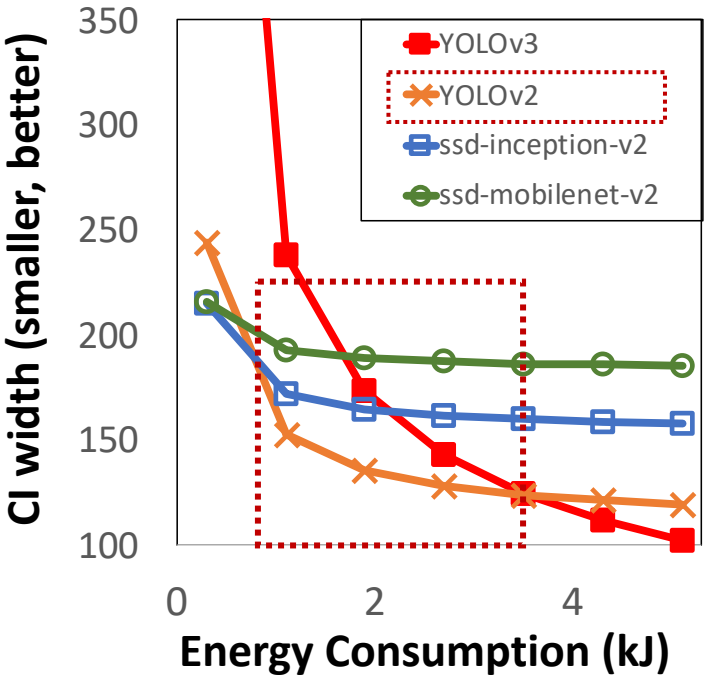


When energy is low: cheaper NNs win

- **Bottlenecked by sampling error (frame quantity)**

Elf tech #1: per-window energy/CI fronts

- What's the best count action for a window? No silver bullet.
  - A *count action*: determining (1) an NN and (2) # of frames to process



When energy is low: cheaper NNs win

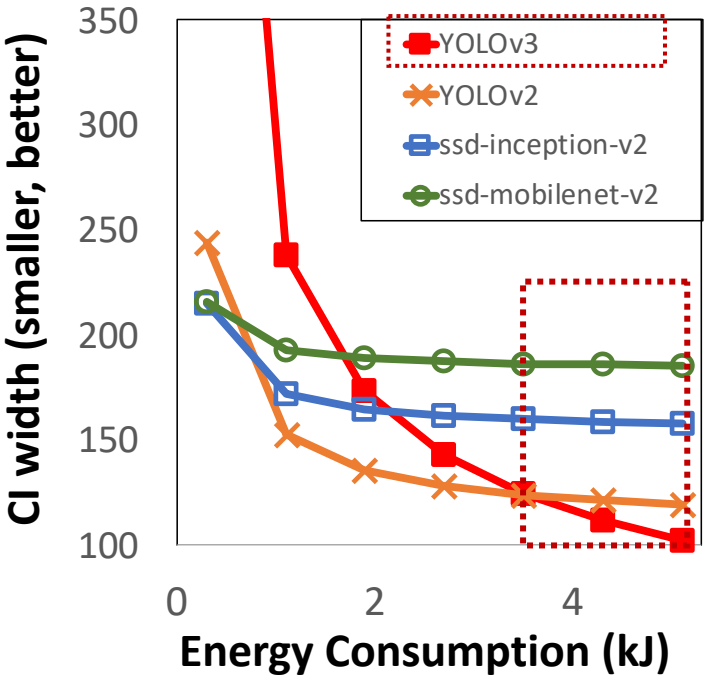
- **Bottlenecked by sampling error (frame quantity)**

When energy is low: more accurate NNs win

- **Bottlenecked by NN error (frame quality)**

Elf tech #1: per-window energy/CI fronts

- What's the best count action for a window? No silver bullet.
  - A *count action*: determining (1) an NN and (2) # of frames to process



When energy is low: cheaper NNs win

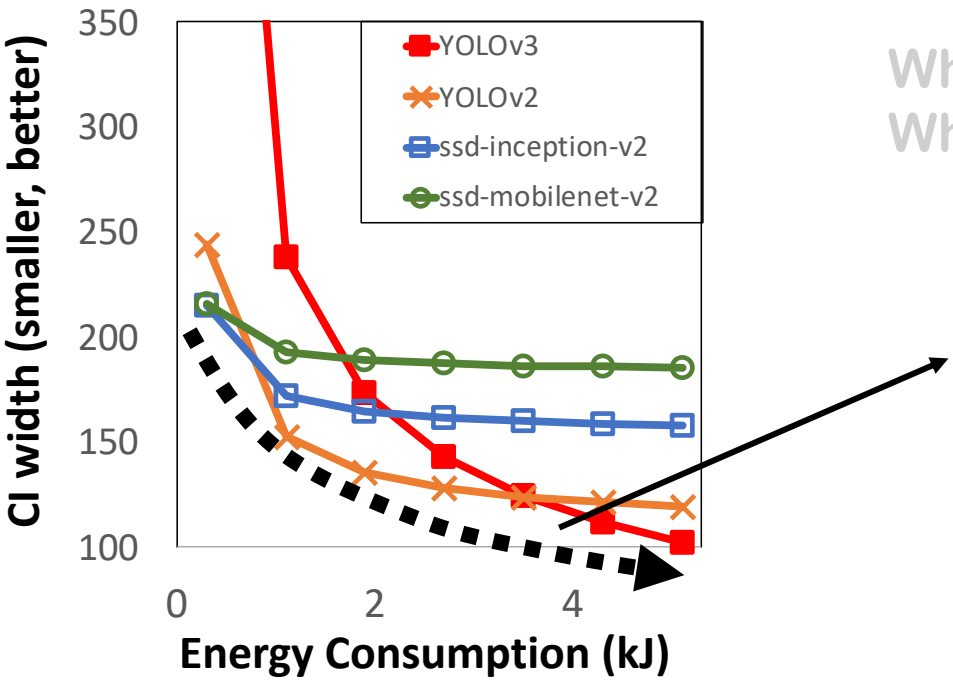
- **Bottlenecked by sampling error (frame quantity)**

When energy is low: more accurate NNs win

- **Bottlenecked by NN error (frame quality)**

Elf tech #1: per-window energy/CI fronts

- What's the best count action for a window? No silver bullet.
  - A *count action*: determining (1) an NN and (2) # of frames to process

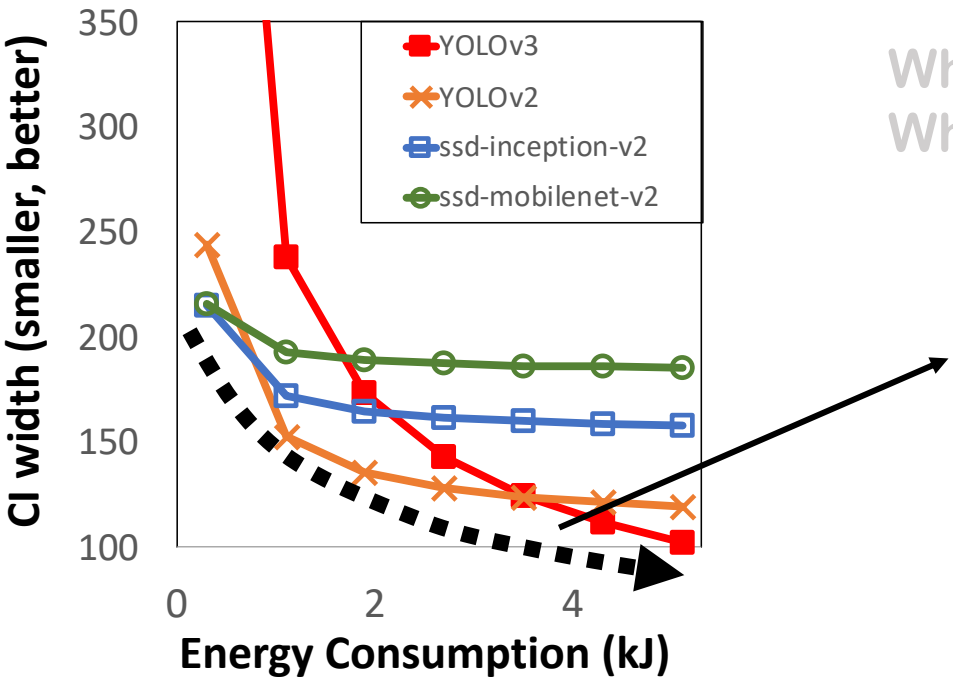


When energy is low: cheaper NNs win
When energy is low: more accurate NNs win

Energy/CI front: *the combination of all "optimal" count actions with varied energy*

Elf tech #1: per-window energy/CI fronts

- What's the best count action for a window? No silver bullet.
  - A *count action*: determining (1) an NN and (2) # of frames to process



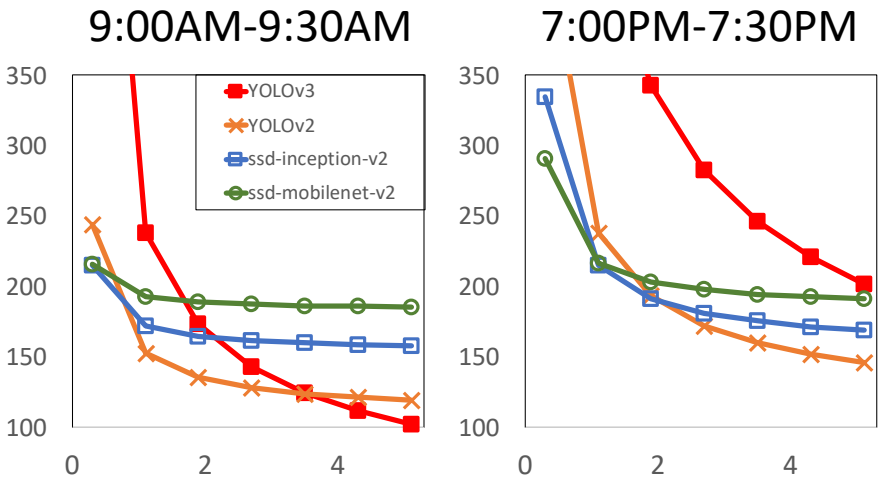
When energy is low: cheaper NNs win
When energy is low: more accurate NNs win

Energy/CI front: *the combination of all "optimal" count actions with varied energy*

- How to construct? Error integration
- Depends on the video characteristics

Elf tech #1: per-window energy/CI fronts

- What's the best count action for a window? No silver bullet.
  - A *count action*: determining (1) an NN and (2) # of frames to process



Different windows have different energy/CI fronts

When energy is low: cheaper NNs win
When energy is low: more accurate NNs win

Energy/CI front: *the combination of all "optimal" count actions with varied energy*

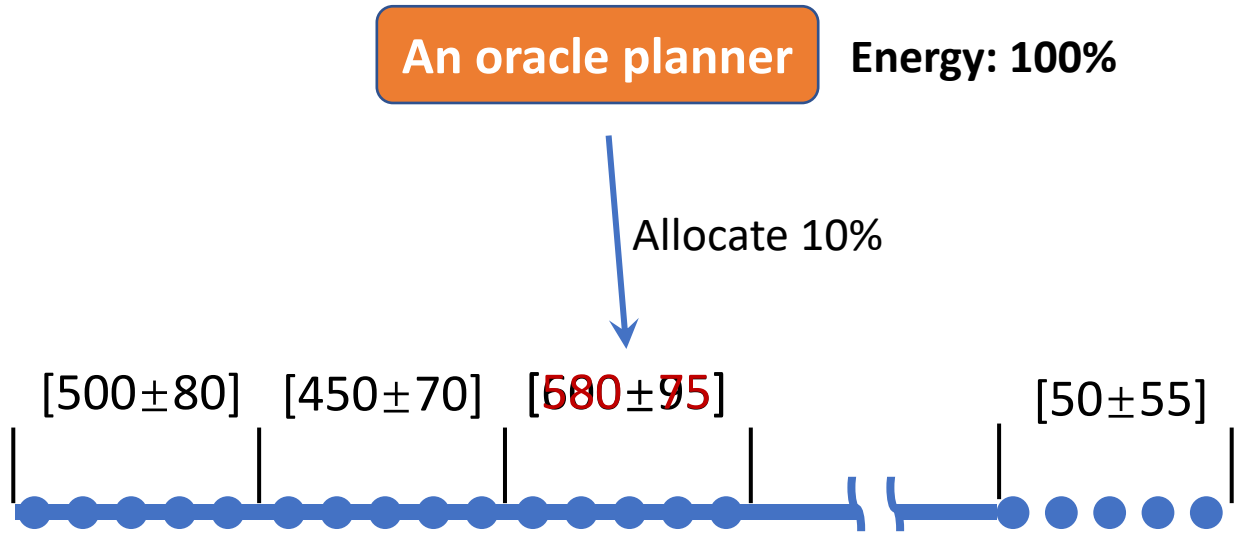
- How to construct? Error integration
- Depends on the video characteristics

Elf tech #2: across-window joint planning

- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts

Elf tech #2: across-window joint planning

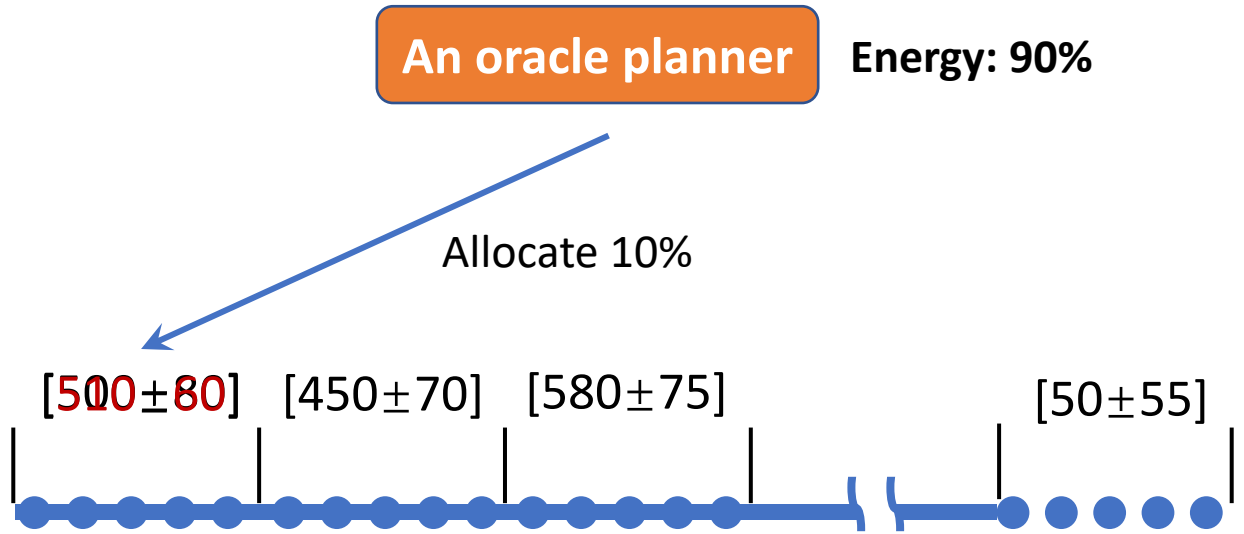
- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts



A greedy approach: giving energy to the window with the most benefit (i.e., CI width reduction).

Elf tech #2: across-window joint planning

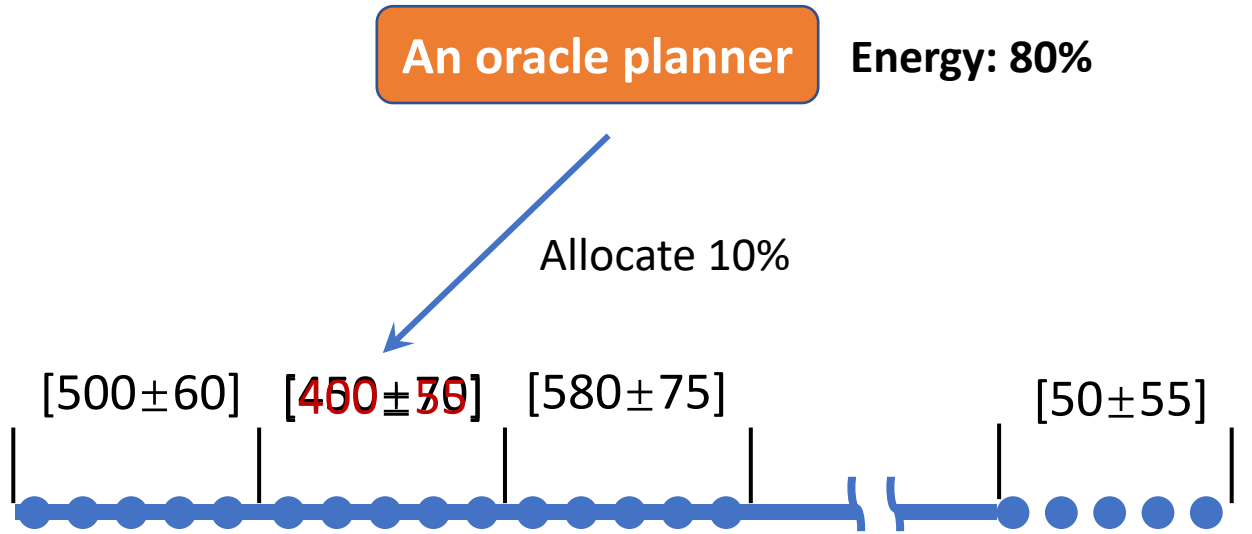
- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts



A greedy approach: giving energy to the window with the most benefit (CI width reduction)

Elf tech #2: across-window joint planning

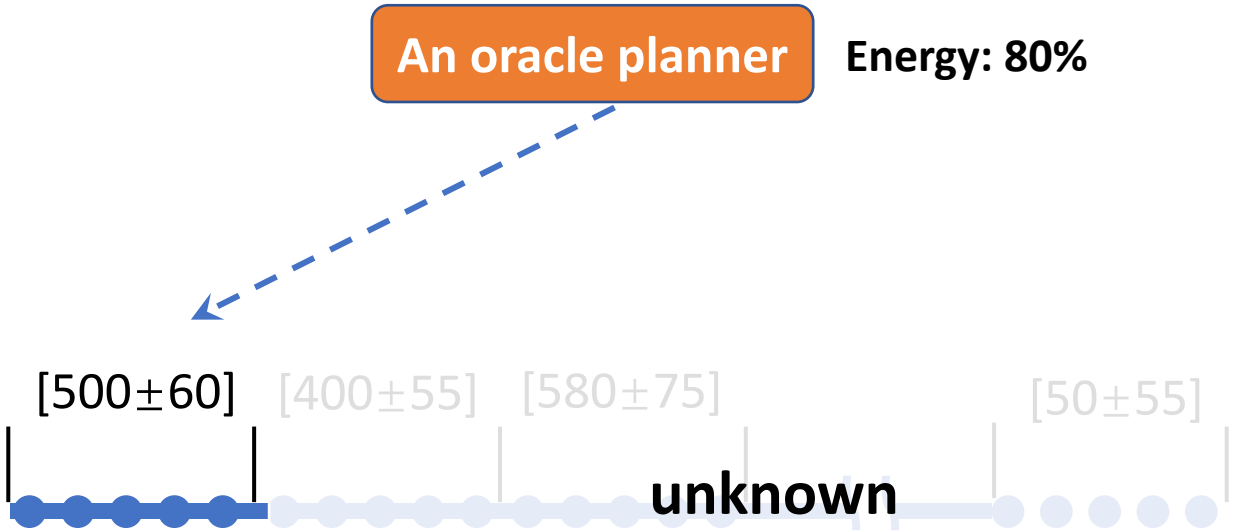
- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts



A greedy approach: giving energy to the window with the most benefit (CI width reduction)

Elf tech #2: across-window joint planning

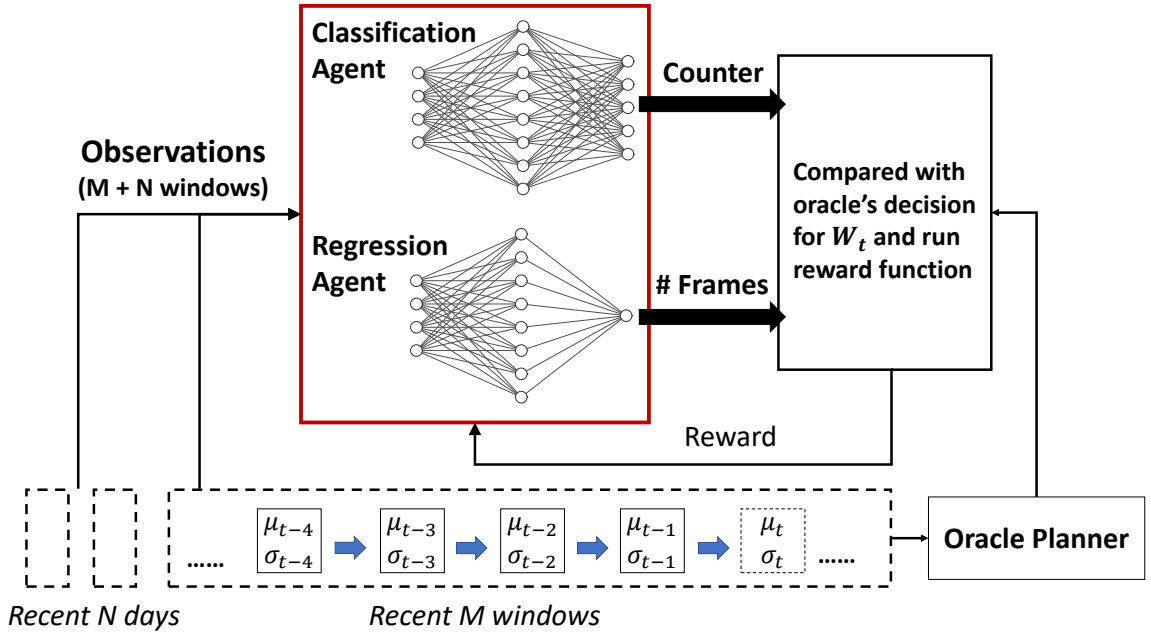
- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts



A greedy approach: giving energy to the window with the most benefit (CI width reduction)

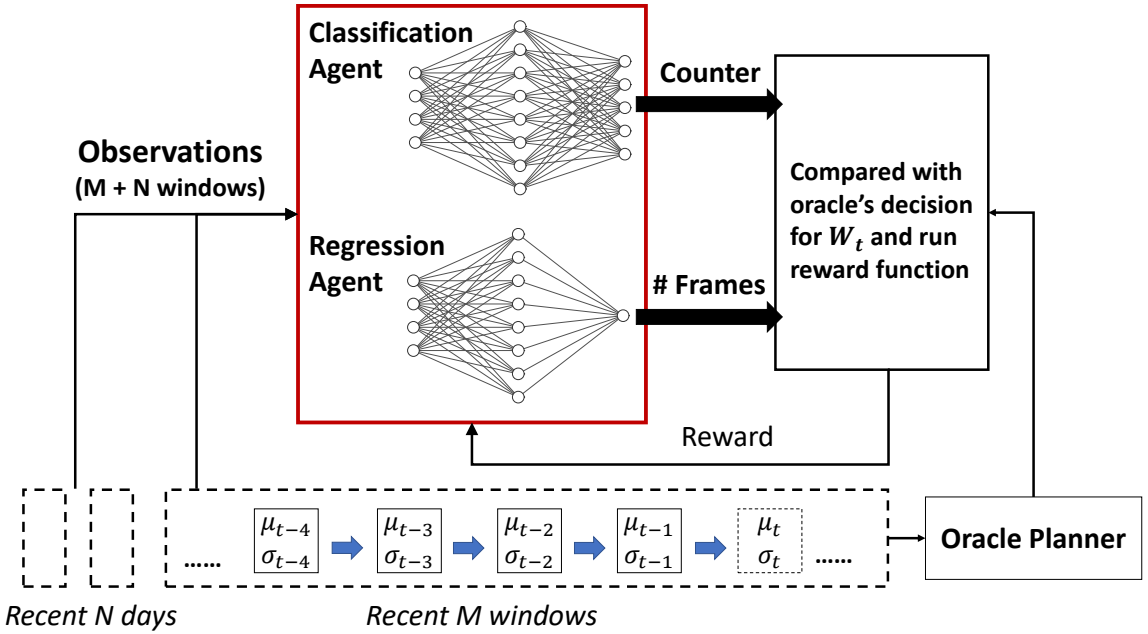
Elf tech #2: across-window joint planning

- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts
  - planned offline
- A learning-based planner: imitating the oracle planner
  - basis: reinforcement learning
  - rationale: daily and temporal patterns



Elf tech #2: across-window joint planning

- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts
  - planned offline
- A learning-based planner: imitating the oracle planner
  - basis: reinforcement learning
  - rationale: daily and temporal patterns
  - offline training -> online prediction
    - Two agents: NN selection and # of frames
    - Observations: knowledge of past windows
    - Penalty: deviation from oracle's decision

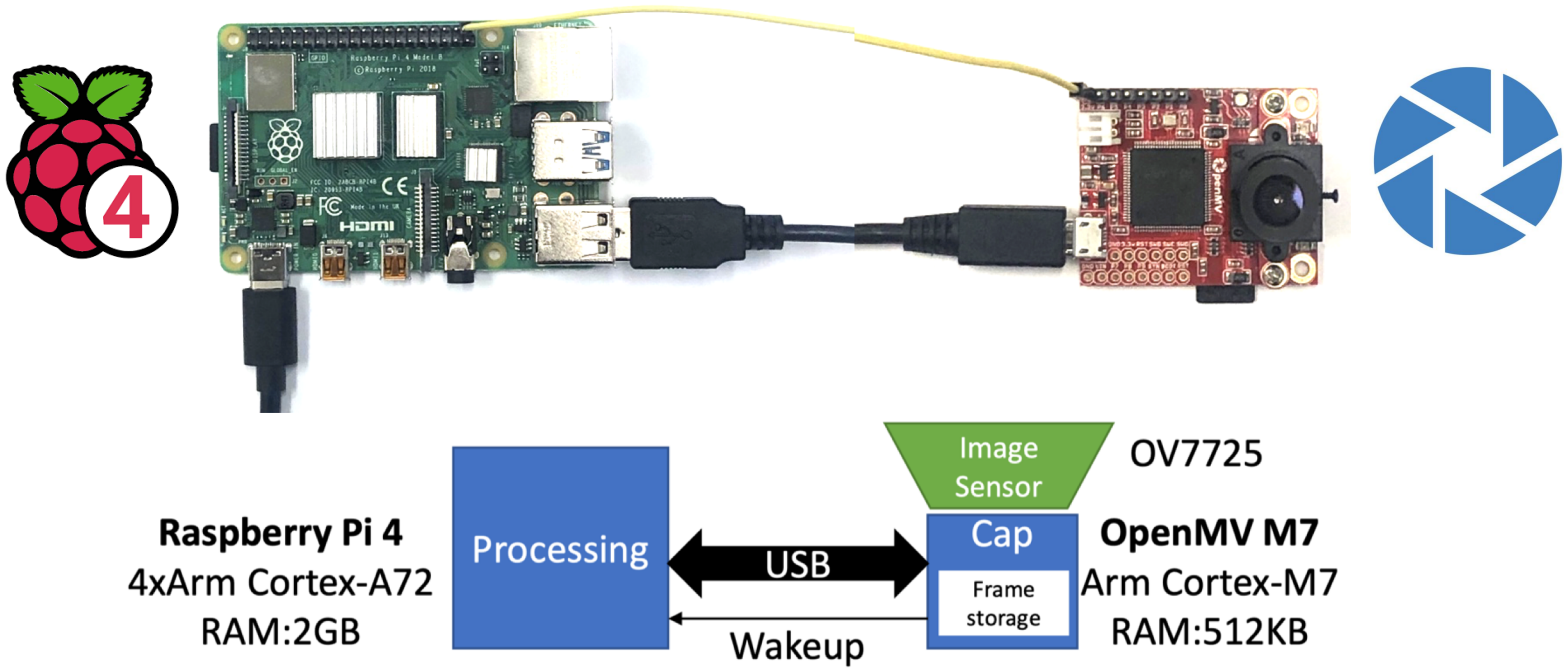


Elf tech #2: across-window joint planning

- An Oracle Planner: best performance but unrealistic
  - knows all energy/CI fronts
  - planned offline
- A learning-based planner: imitating the oracle planner
  - basis: reinforcement learning
  - rationale: daily and temporal patterns
  - offline training -> online prediction
  - Enforce energy budget: make reservation for future windows
    - 30 frames to be statistically meaningful

Elf Implementation

- Capture & processing decoupled for higher energy efficiency
  - Processing batched at the end of each window



Elf Evaluation

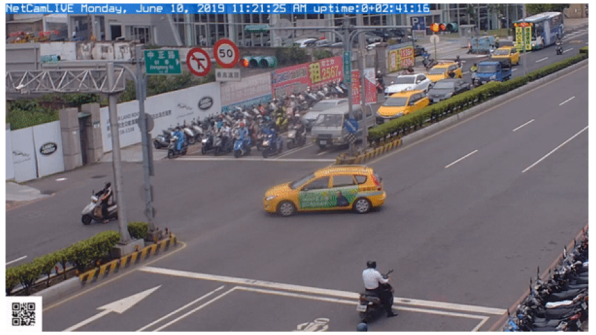
- Over 1,000-hr videos
  - Public, 2-week long each stream
- Baselines
  1. *GoldenNN*: most accurate NN
  2. *UniNN*: one fixed best NN
  3. *Oracle*: offline planned
- Small solar panel
  - 10Wh~30Wh per day

Auburn, AL

Hampton, NY

Jackson, WY

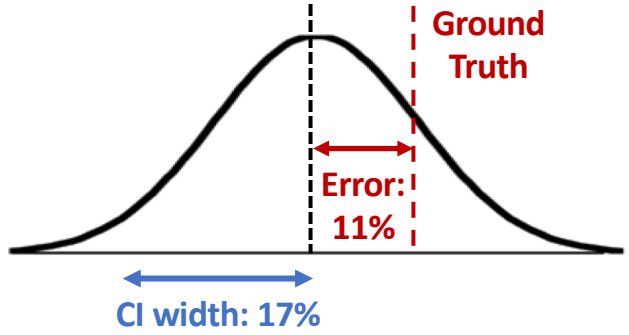
Taipei

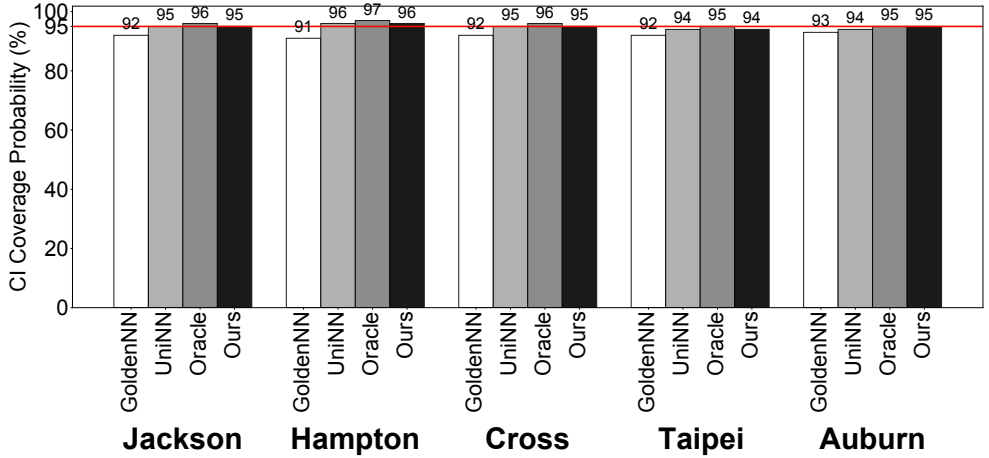


Taipei

Elf Evaluation

- Average: 11% error, valid and 17%-width CI
  - 95% confidence level



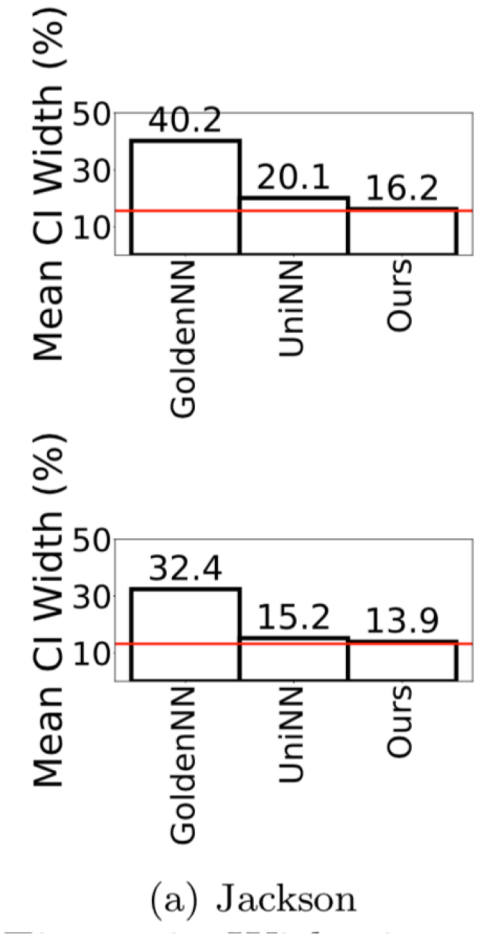


Cis cover ground truth with 95% probability (specified)

Elf Evaluation

- Average: 11% error, valid and 17%-width CI
- Significant improvements over baselines in CI widths
  - 66.6%, 59.8%, and 56.2% smaller over *GoldenNN* (up to 3.4x)
  - 41.1%, 16.6%, and 9.7% smaller over *UniNN*

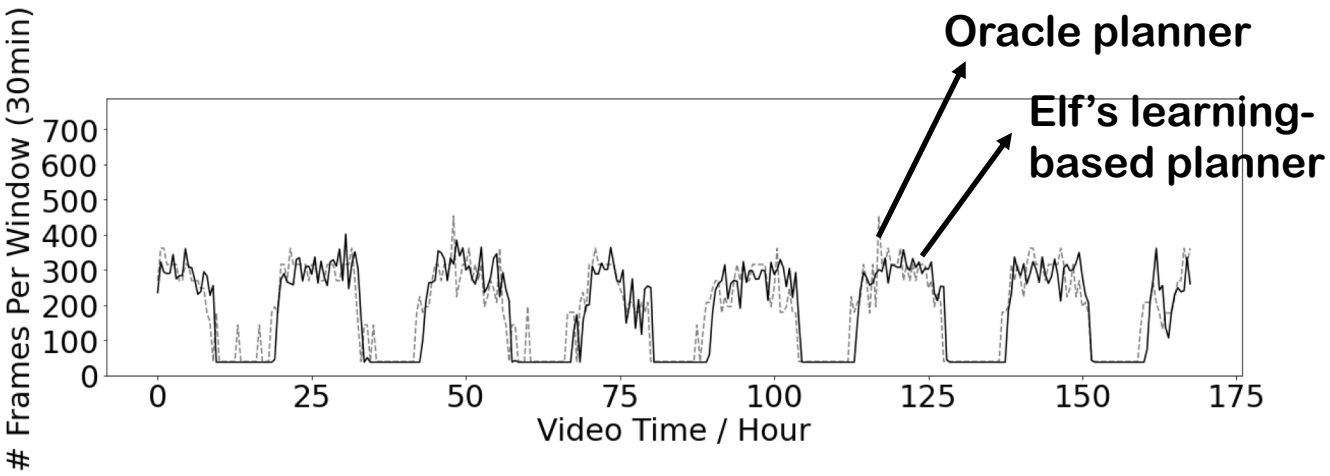
10Wh per day 20Wh per day 30Wh per day

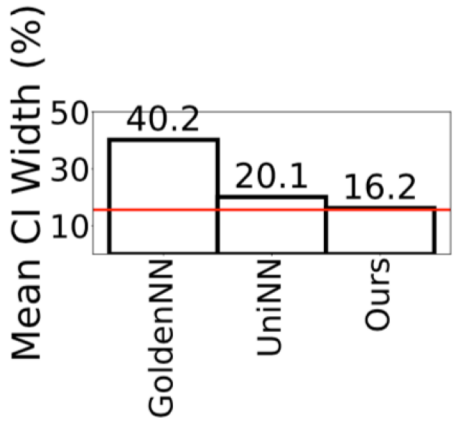


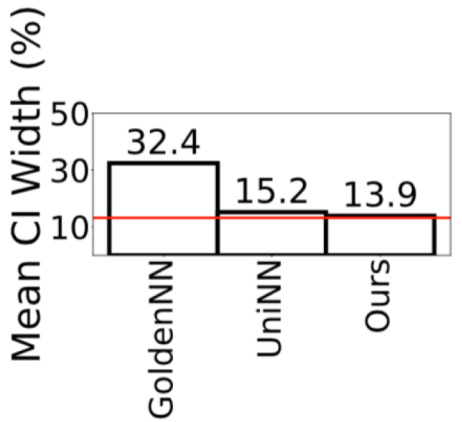
(a) Jackson

Elf Evaluation

- Average: 11% error, valid and 17%-width CI
- Significant improvements over baselines in CI widths
- Very close to *Oracle*
  - < 5% wider CI
  - Well imitating the oracle planner



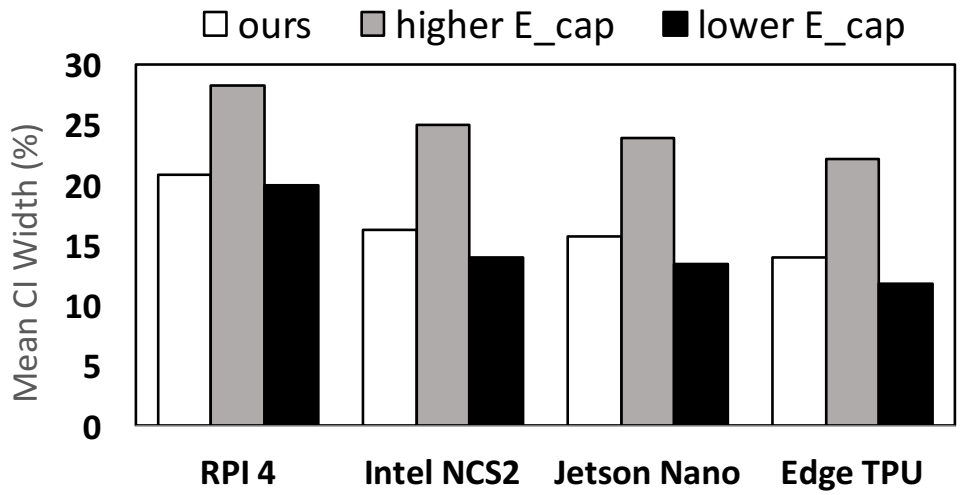




(a) Jackson

Elf Evaluation

- Average: 11% error, valid and 17%-width CI
- Significant improvements over baselines in CI widths
- Very close to *Oracle*
- What if we have AI accelerators?
  - CIs are reduced noticeably (by 22.1%–33.1%)
  - Still cannot process every frame (short of energy)



Summary

- Autonomous camera: expanding the geo-frontier of video analytics
  - Energy-independent and compute-independent
- Elf: the first runtime for autonomous camera
  - Target query: object counting
  - Key idea: count planning per- and across-windows
- Prototyped on heterogeneous hardware
- Evaluated on over 1,000-hr videos
  - 11% error, 17% CI width