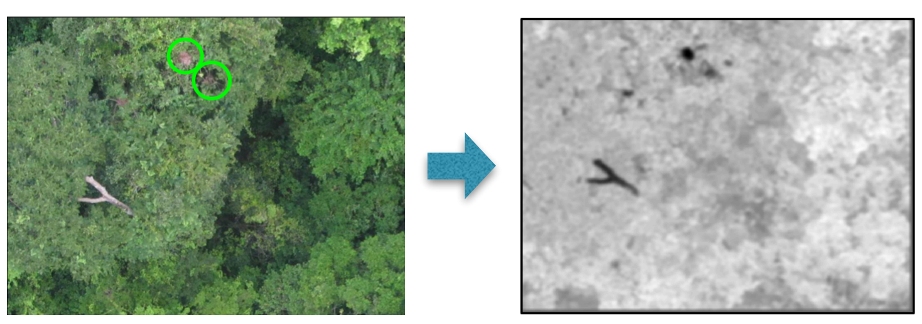
# Active Detection for Biodiversity Monitoring via Adaptive Submodularity

Yuxin Chen, Hiroaki Shioi, Cesar Antonio Fuentes Montesinos, Lian Pin Koh, Serge Wich and Andreas Krause

ETH zürich · liverpool JMU

## TASK: Object detection

Conservation drone · Orangutan (nest)

Orangutan nests detection for biodiversity monitoring

## Including Human in the Loop?

- ✓ Automatic systems are inaccurate.
- ✓ Human labeling is tedious/expensive.
- ✓ How to make the best of both?

## Active Detection as a Seq. Decision Problem

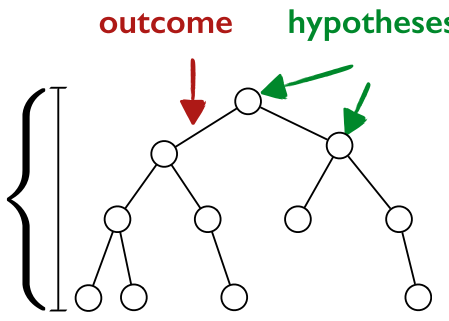
Hypotheses (Obj. at certain location) $\mathcal{H} = \{h_1, \dots, h_n\}$

Outcome is RV : $\mathbf{Y}_{\mathcal{H}} = [Y_1, \dots, Y_n] \in \mathcal{O}$ from $\mathbb{P}[\mathbf{Y}_{\mathcal{H}}]$

Hypothesis-outcome pairs $\mathbf{y}_{\mathcal{A}}$, or $\mathcal{S}(\pi, \mathbf{y}_{\mathcal{H}}) \in \mathcal{H} \times \mathcal{O}$

**Policy** $\pi : 2^{\mathcal{H} \times \mathcal{O}} \to \mathcal{H}$

Worst-case cost of a policy $\text{cost}_{wc}(\pi) = \max_{\mathbf{y}_{\mathcal{H}}} |\mathcal{S}(\pi, \mathbf{y}_{\mathcal{H}})|$

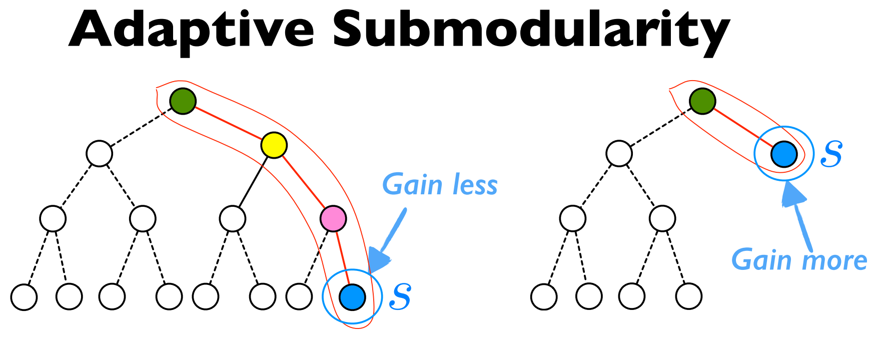



Objective function $f : 2^{\mathcal{H} \times \mathcal{O}} \to \mathbb{R}_{\geq 0}$

The Min-cost Cover Problem

$$\min_{\pi \in \Pi} \text{cost}(\pi), \quad \text{s.t.} \quad f(\mathcal{S}(\pi, \mathbf{y}_{\mathcal{H}})) \geq Q \text{ for all } \mathbf{y}_{\mathcal{H}} \text{ with } P(\mathbf{y}_{\mathcal{H}}) > 0.$$

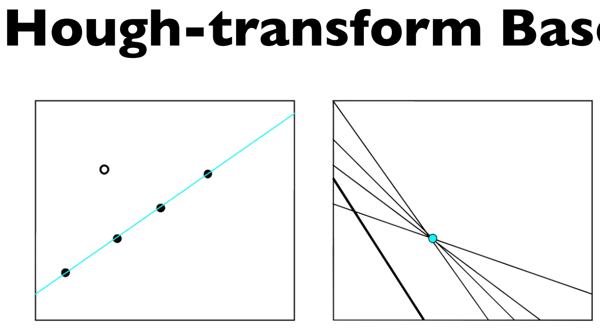
## Adaptive Submodularity




Playing an action at an earlier stage (i.e., at an ancestor) only increases its expected marginal benefit.

taken over its outcome

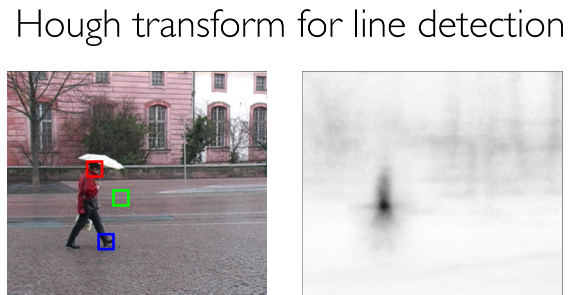
## Hough-transform Based Detector



Hough transform for line detection

Image space (votes, evidences)

Hough space (hypotheses)



Hough forest for pedestrian detection

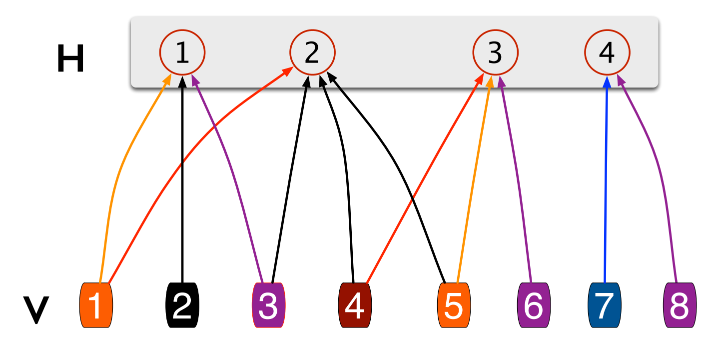
## The Active Detection Framework

Voting elements $\mathcal{V} = \{v_1, \dots, v_m\}$ Hypotheses $\mathcal{H} = \{h_1, \dots, h_n\}$

Interactions between voting elements and hypotheses: $\mathcal{G} = (\mathcal{V}, \mathcal{H}, \mathcal{E})$

**Positive coverage:** Votes can be fully explained /covered by a true hypotheses.

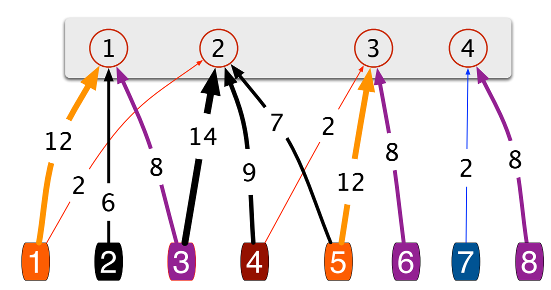
**Negative coverage:** Votes that are similar with false votes should be discounted.



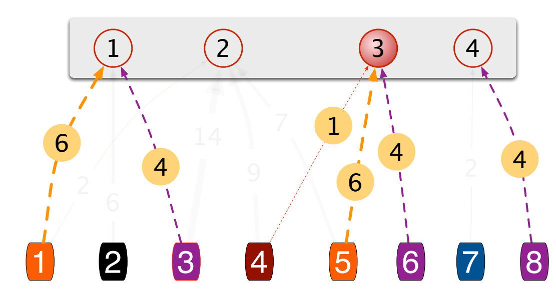

**Binary votes setting**

$$f^{(1)}_{v,h}(\mathbf{y}_{\mathcal{A}}) = \max \begin{cases} 1, & \text{if } \exists h' : (h', +) \in \mathbf{y}_{\mathcal{A}} \wedge (v, h') \in \mathcal{E}; \\ g(v, h, \mathbf{y}_{\mathcal{A}}), & \text{otherwise.} \end{cases}$$

Coverage for edge (v,h) · Discounted (negative) coverage for (v,h): E.g., g=1-0.5^(#similar false votes)



**Real-value votes setting**



Negative edge coverage

$$f_{v,h}(\mathbf{y}_{\mathcal{A}}) = \overbrace{g(v, h, \mathbf{y}_{\mathcal{A}}) \cdot w_{vh}}^{1} + \underbrace{\min \left\{ \max_{h' : (h', +1) \in \mathbf{y}_{\mathcal{A}}} w_{vh'}, \ (1 - g(v, h, \mathbf{y}_{\mathcal{A}})) \cdot w_{vh} \right\}}_{2}$$

weight for edge (v,h)

**1**. weight covered due to negative observations

**2**. remaining weight (i.e., after negative discount) covered due positive observations

## The Objective Function

$$F(\mathbf{y}_{\mathcal{A}}) = \sum_{(v,h) \in \mathcal{E}} f_{v,h}(\mathbf{y}_{\mathcal{A}})$$

**Conditional Marginal Benefit of a hypotheses / detection h**

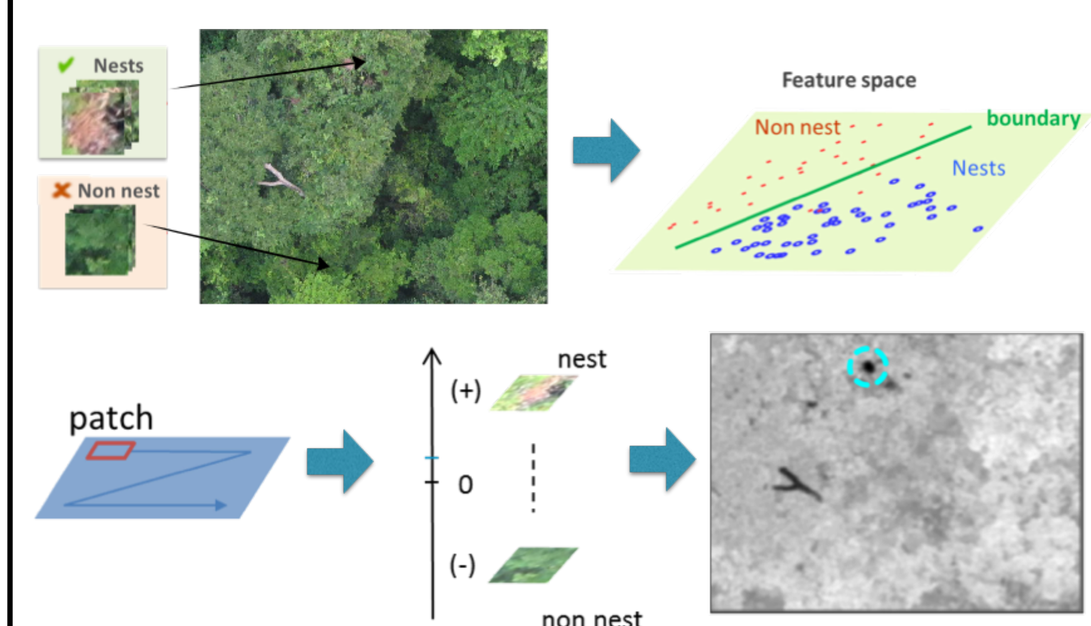
$$\Delta_F(h \mid \mathbf{y}_{\mathcal{A}}) = \mathbb{E}_{\mathbf{y}_{\mathcal{H}}}[F(\mathbf{y}_{\mathcal{A}} \cup \{(h, y_h)\}) - F(\mathbf{y}_{\mathcal{A}}) \mid \mathbf{y}_{\mathcal{A}}].$$

Expectation over all realizations · The utility of a specific realization · Conditioning on previous observations

F is **adaptive submodular** w.r.t. factorial prior P ($Y_1, \dots, Y_n$ ).

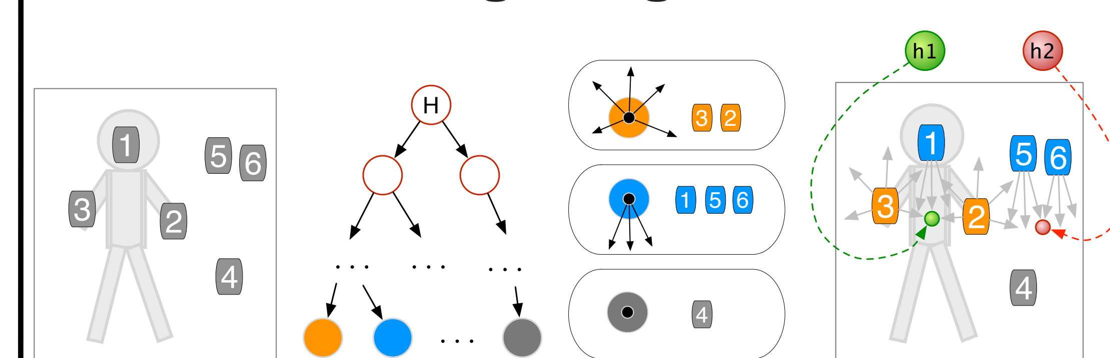
**THM.** Fix Q > 0 and β > 0. Let $C_{greedy}(\pi)$ be the worst-case cost of the greedy algorithm, using a factorial prior P on variables $Y_1, \dots, Y_n$, until it achieves expected value Q - β. Let $OPT_{wc}$ be the worst-case cost of the optimal algorithm. It holds that: $C_{greedy} \leq OPT_{wc}(\ \ln(\ Q/\beta\ ) + 1\ )$. Moreover, $P(\ F(\ y_A\ ) \geq Q\ ) \geq 1 - \beta$.

## Activizing the Base Detector




Orangutan nest detection (UAV-Forest)

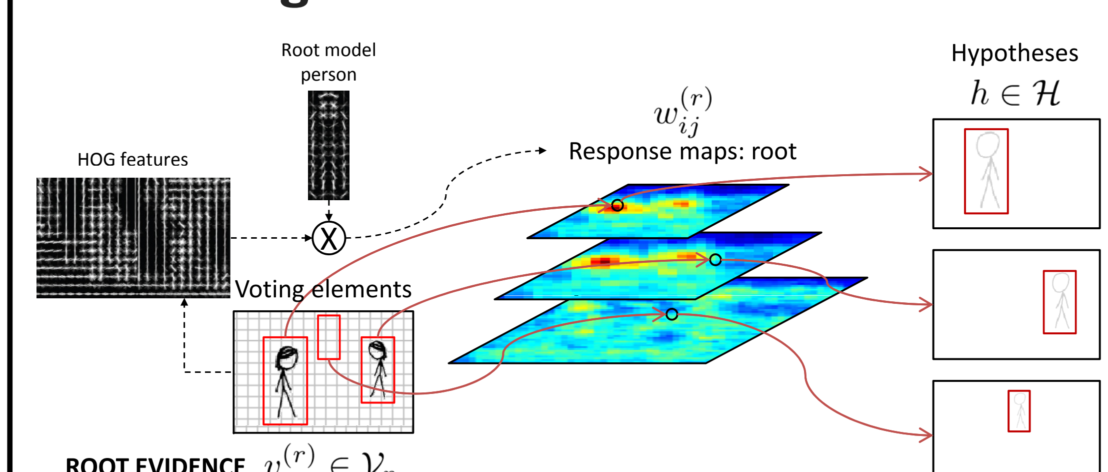
## Activizing Hough Forest




Input image · Hough forest · Clusters · Hough votes

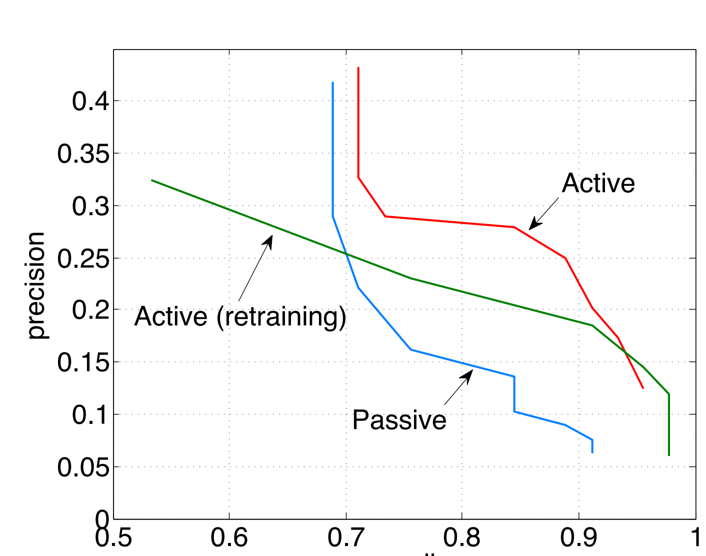
**Hough Forest** as a base detector

## Activizing the Deformable Parts Model

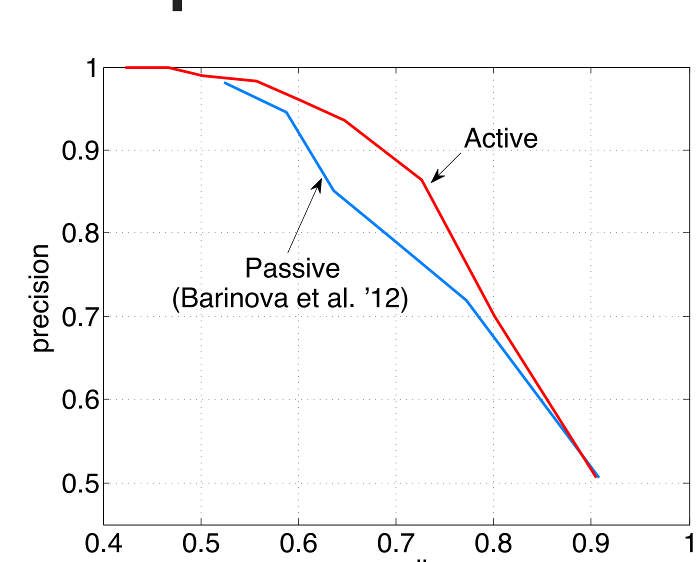



**Deformable Parts Model (DPM)** as a base detector

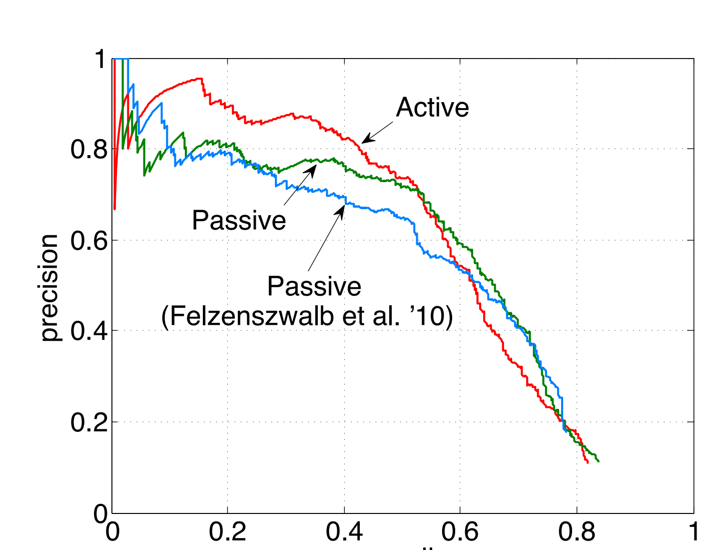
## Experimental Results



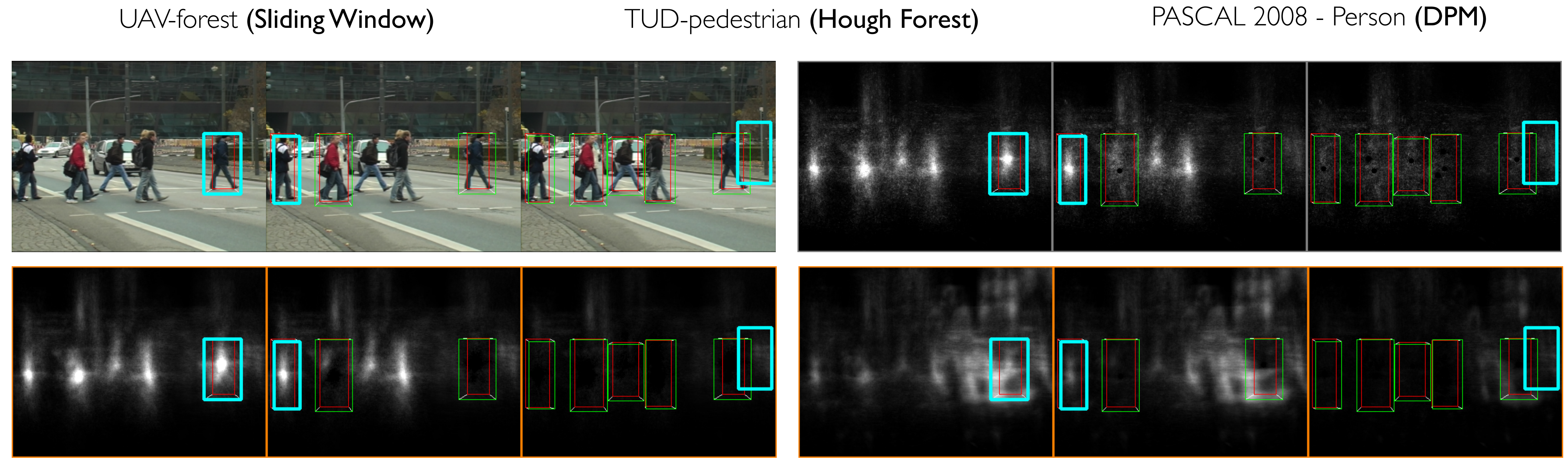

UAV-forest (**Sliding Window**)




TUD-pedestrian (**Hough Forest**)




PASCAL 2008 - Person (**DPM**)



(Upper left) Input image; (Upper right) Response image; (Lower left) Positive coverage; (Lower right) Negative coverage.
Cyan box: current detection; Red boxes: ground-truth labels of pedestrians; Green boxes: detections made by the active detector.