25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Towards Visual Insect Camera Traps

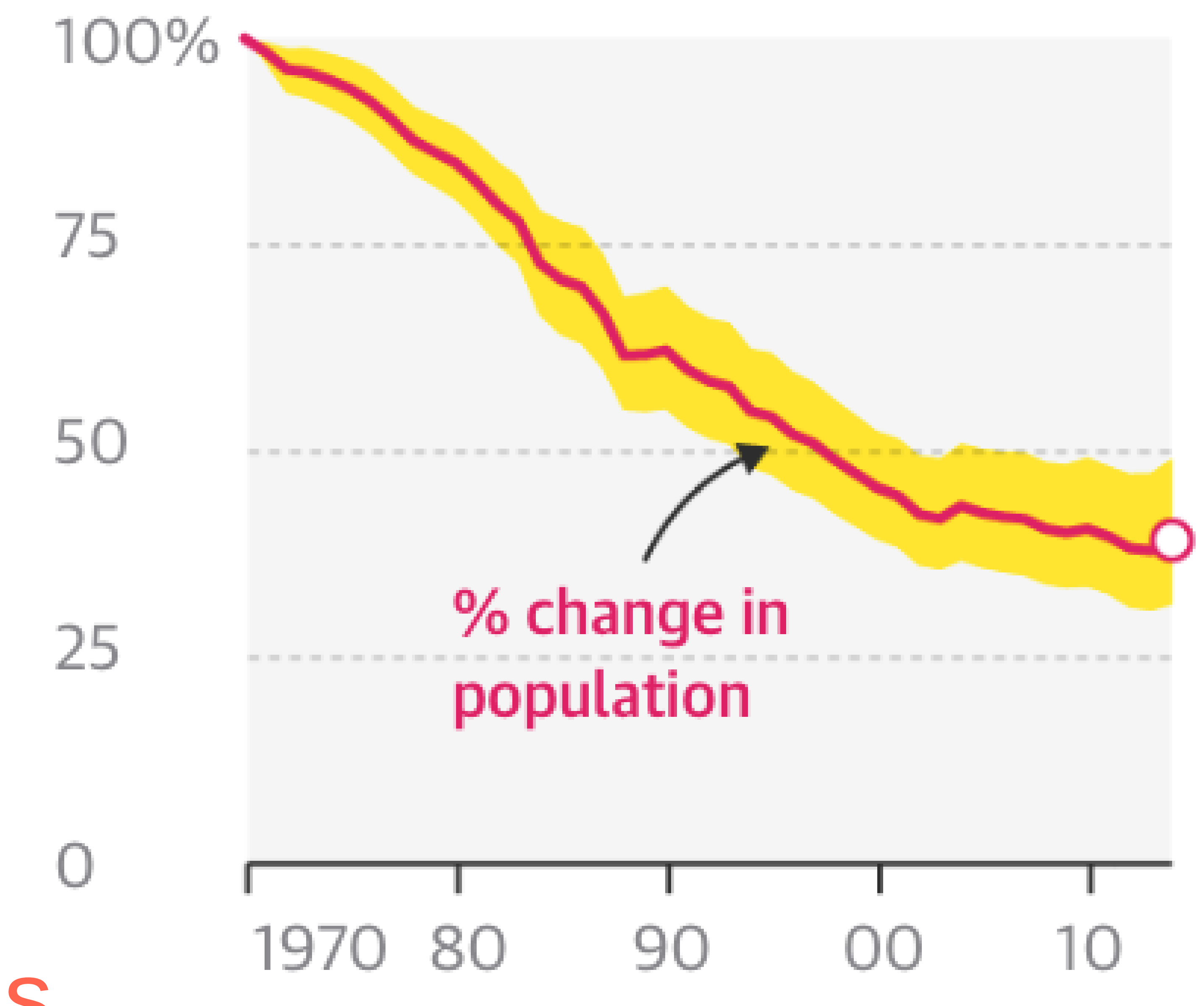
Sebastian Thiele[1*], Lars Haalck[1*], Marvin Struffert[1], Christoph Scherber[2] and Benjamin Risse[1○]

1 University of Muenster, Faculty of Mathematics and Computer Science
2 Centre for Biodiversity Monitoring, Zoological Research Museum Alexander Koenig
* Contributed equally
○ Corresponding author (b.risse@uni-muenster.de)

COMPUTER VISION & MACHINE LEARNING SYSTEMS

Decline in animal populations

- Quantitative data indicates extinction of vertebrates (caused by humans)
- 60% of the vertebrate animals have been lost since 1970
- However: more than 90% of all animal species are invertebrates
- Insects are crucial for the ecosystem
  - 90% of the flowering plants benefit from pollination [1]
  - Insect pollinators promote the production of 75% of major global agricultural crops [2]
  - Insects contribute by ~150 billion € to the global economy [3]
- Monitoring and quantifying invertebrates is more difficult



Living Planet index, WWF/ZSL, 2018

[1] Ollerton, J., Winfree, R. & Tarrant, S. How many flowering plants are pollinated by animals? Oikos 120, 321–326 (2011).
[2] Gallai, N., et al. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. Ecol. Econ. 68, 810–821 (2009).
[3] Eilers, E. J., et al. Contribution of Pollinator-Mediated Crops to Nutrients in the Human Food Supply. PLoS One 6, e21363 (2011).

Decline in insect populations

Plummeting insect numbers 'threaten collapse of nature'

- In the news: “Insects could vanish within a century”
- 41% of global insect species have declined over the past century [1], e.g. bees with 46% or caddisflies with 68%
- Terrestrial and non-terrestrial insect species are affected [2]
- How to quantify the decline of insect populations?

[1] Sánchez-Bayo, F. & Wyckhuys, K. A. G. Biological Conservation 232, 8–27 (2019).

[2] Dirzo, Rodolfo, et al. "Defaunation in the Anthropocene." Science 345.6195 (2014): 401-406.

Insect Camera Trap

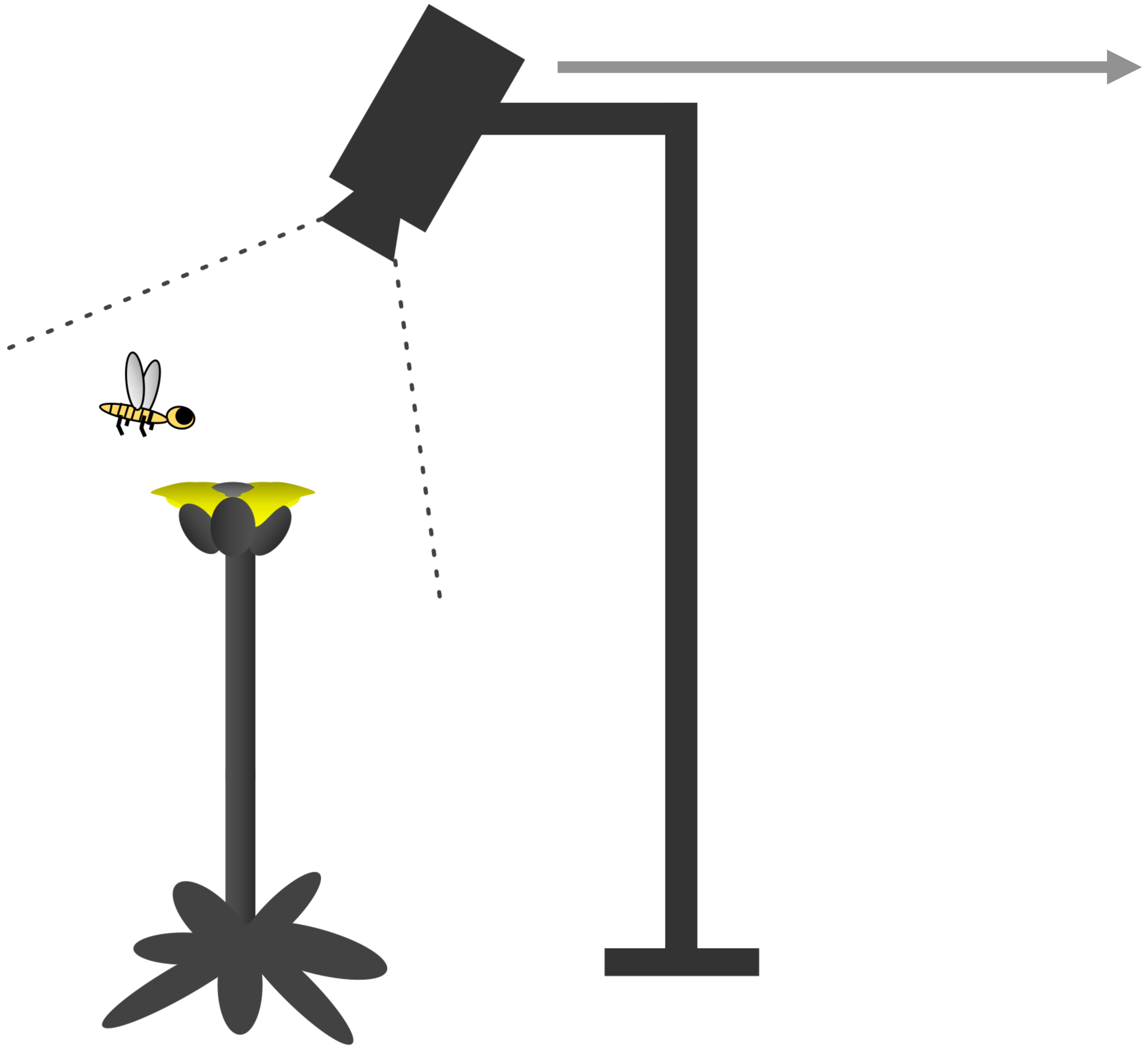
Motivation for a visual insect trap

Malaise Trap

Camera Trap

Insect Camera Trap

Motivation for a visual insect trap



? Save ?

Towards Visual Insect Camera Traps

Dataset

- Common datasets are not well suited as training baselines/benchmarks for our proposed tasks (either insects too big or only image stills).
- We created a new dataset of experimental field plots of $\sim 2 \times 4.5\ m^2$ were recorded using a waterproof outdoor camera (Ricoh WG-50) under suitable weather conditions.
- Overall 16 cameras were installed to monitor different crop combinations in monocultures and mixtures.

Towards Visual Insect Camera Traps

Dataset

- 35,129 frames were extracted, of which 22,008 images included at least one insect.
- For 50% of the frames location information was added by a domain expert using a custom annotation tool, amounting to 14,847 bounding boxes in total.
- With respect to the entire image, the average bounding box had a height of 7.5%, a width of 4.4% and covered an area of 0.4% pixel.

Towards Visual Insect Camera Traps

Spatial vs. temporal cues

Visual Insect Trap Challenges:

small objects

fast motion

cluttered scene

dynamic background

Towards Visual Insect Camera Traps

Insect presence detection (binary)

input

single
frame

no temporal cues…

neural network

output

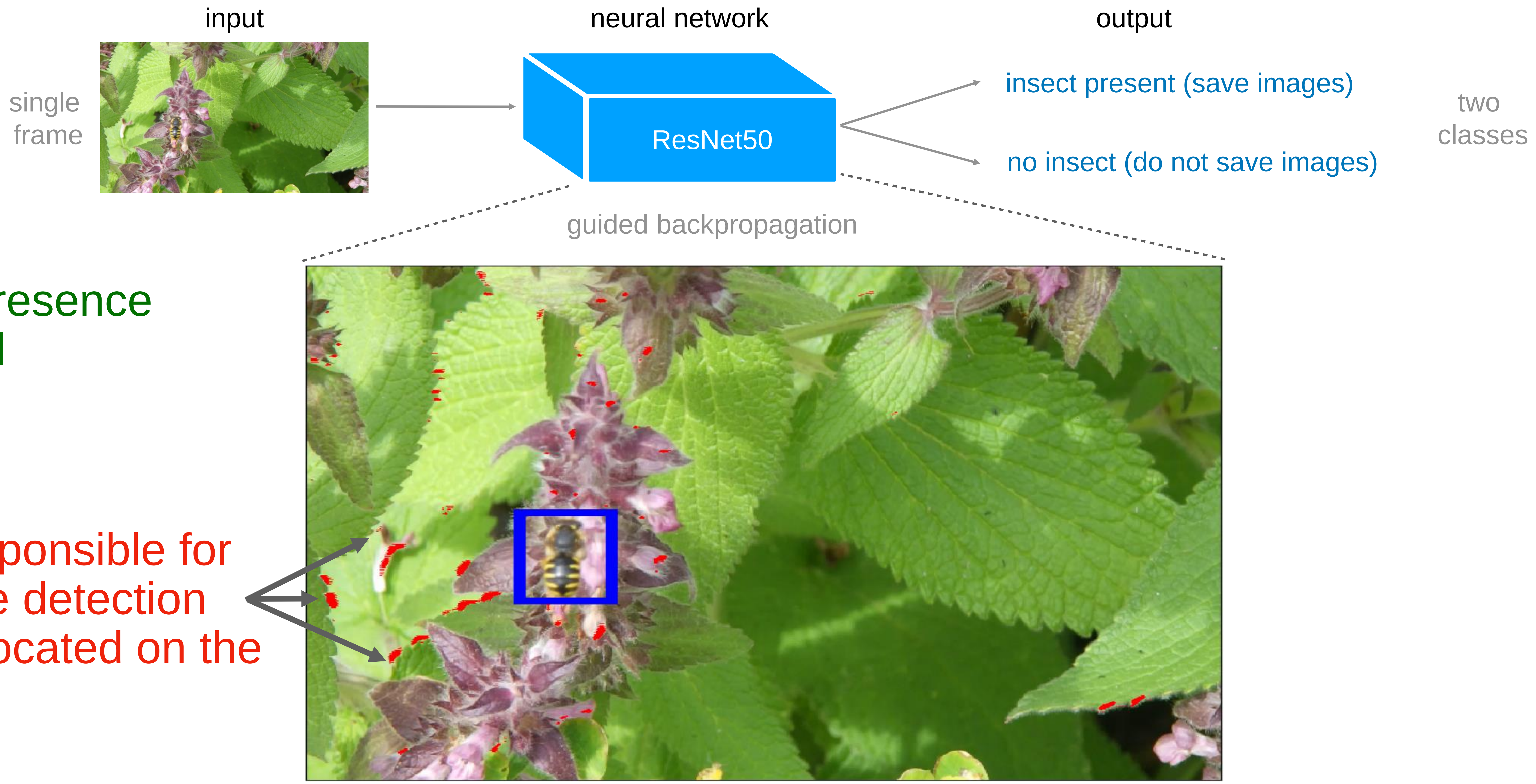
insect present (save images)

no insect (do not save images)

two
classes

Towards Visual Insect Camera Traps

Insect presence detection (binary)

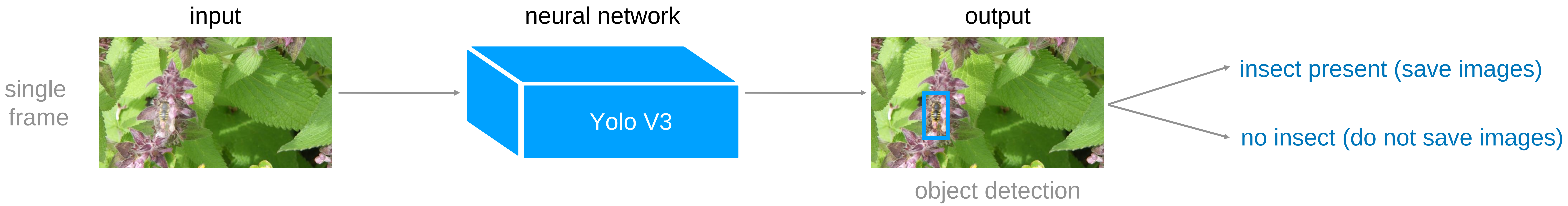


✓ Insect presence detected

- Pixel responsible for presence detection are not located on the insect!

Towards Visual Insect Camera Traps

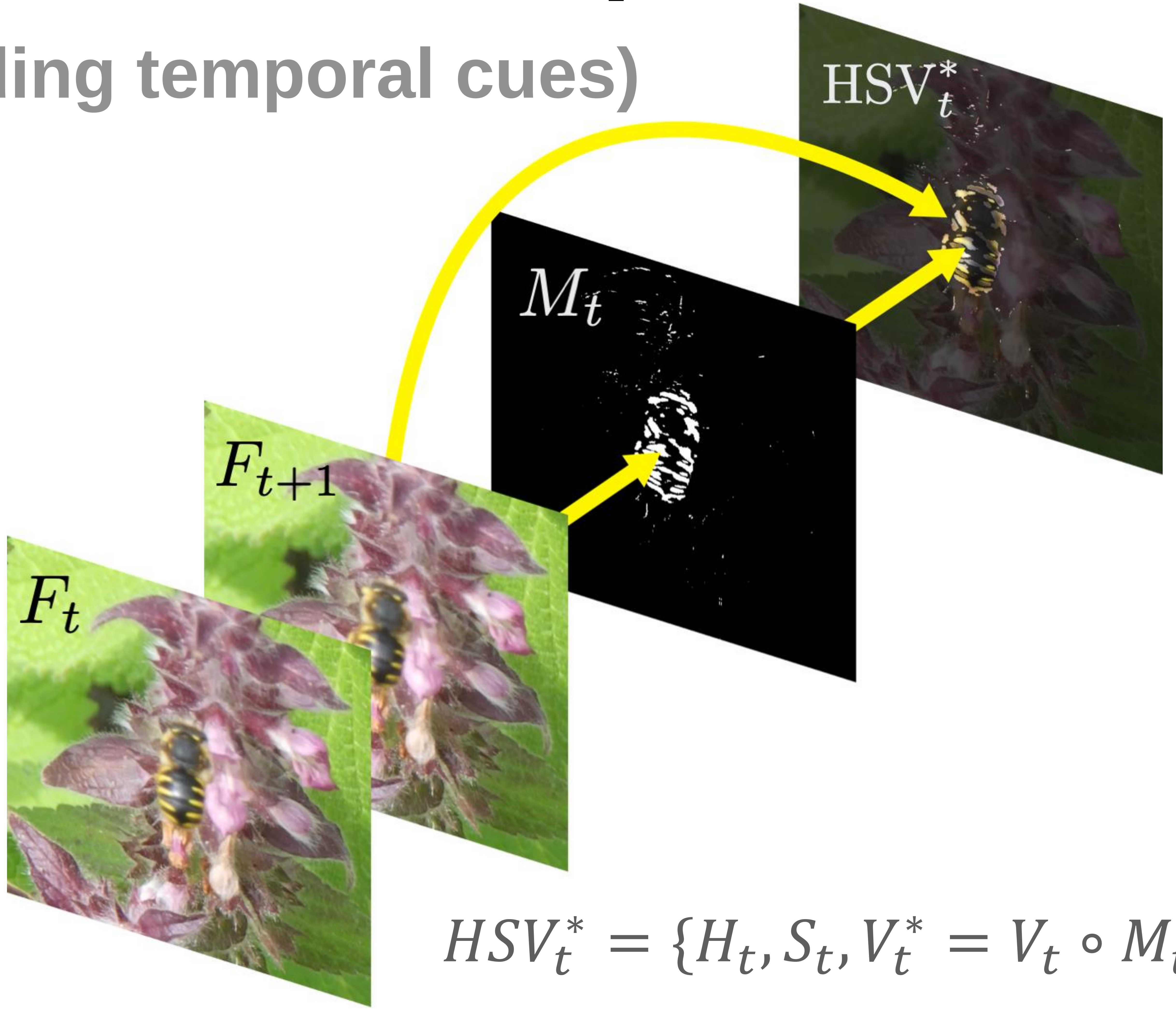
Insect presence detection (bounding box)



- The network is forced to detect the insects themselves, since it will otherwise not fullfill the training target.
- However, additional information through temporal cues in the input is still not utilised.

Towards Visual Insect Camera Traps

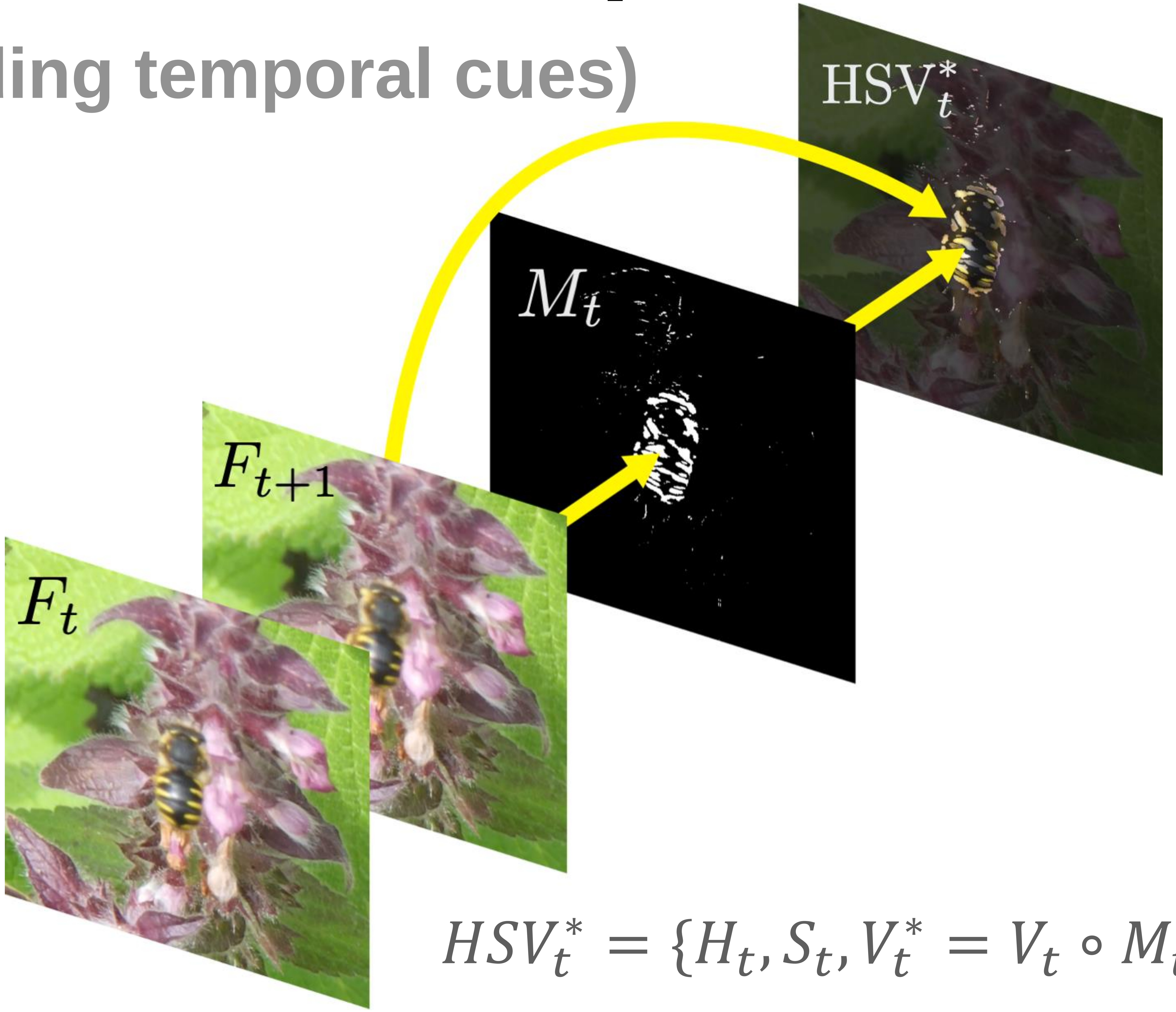
Insect presence detection (including temporal cues)



$$HSV_t^* = \{H_t, S_t, V_t^* = V_t \circ M_t\}$$

Towards Visual Insect Camera Traps

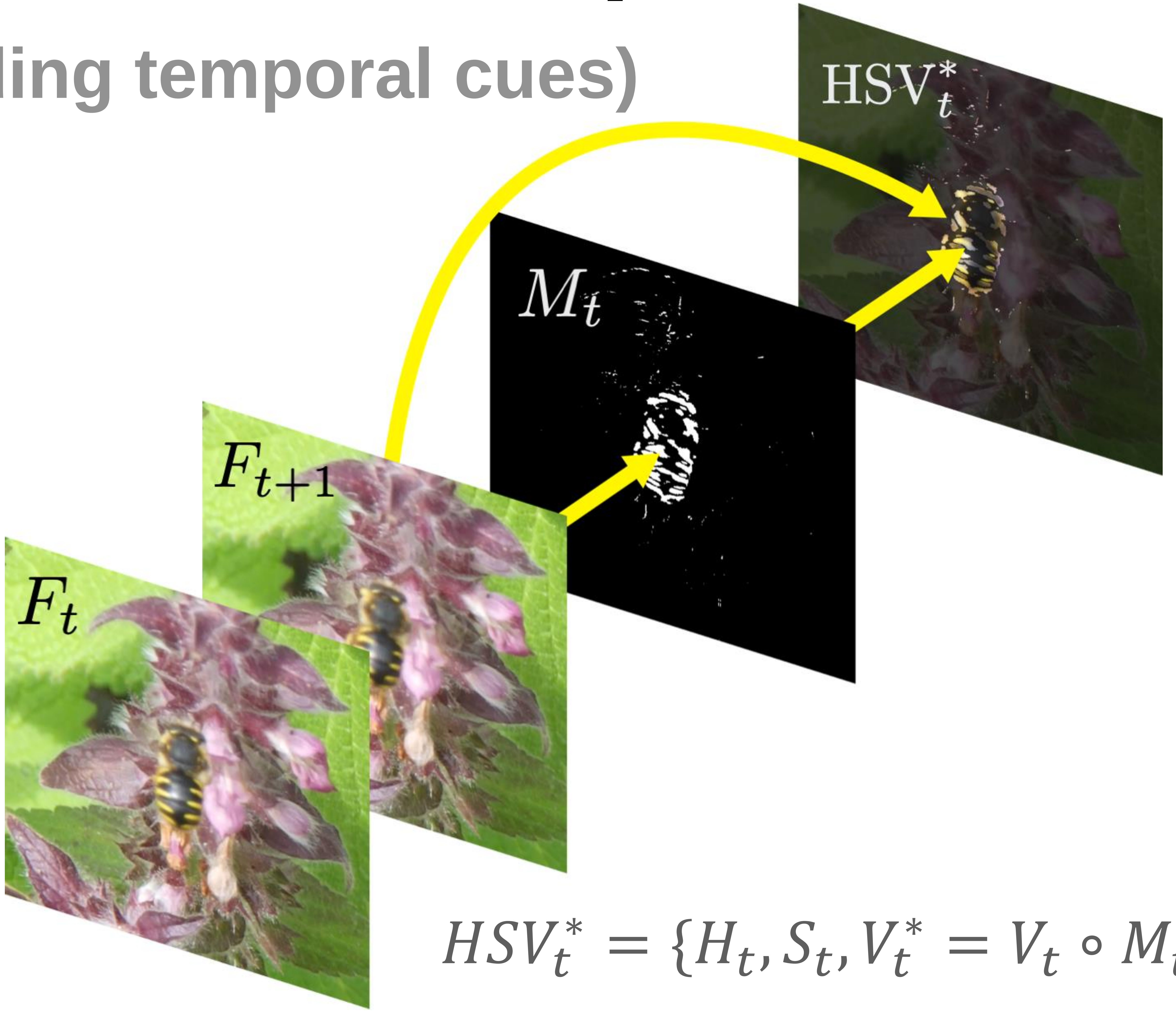
Insect presence detection (including temporal cues)



$$HSV_t^* = \{H_t, S_t, V_t^* = V_t \circ M_t\}$$

Towards Visual Insect Camera Traps

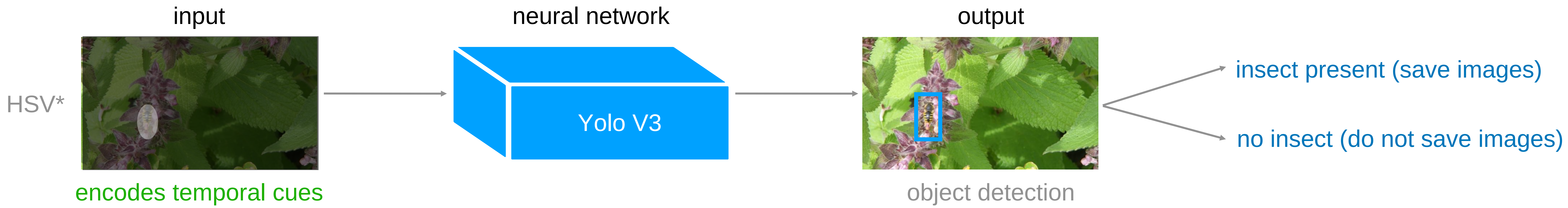
Insect presence detection (including temporal cues)



$$HSV_t^* = \{H_t, S_t, V_t^* = V_t \circ M_t\}$$

Towards Visual Insect Camera Traps

Insect detection accuracy using different models on HSV*



	Architecture	AP	Precision	Recall	F1	Time
RGB	YoloV3	74.15%	96%	70%	81%	2.5 fps
	Faster RCNN	71.95%	91%	80%	85%	0.3 fps
	MobileNet	69.82%	90%	70%	79%	0.45 fps
HSV*	YoloV3	78.07%	92%	72%	81%	2.5 fps
	Faster RCNN	74.39%	86%	82%	84%	0.3 fps
	MobileNet	72.41%	92%	71%	80%	0.45 fps

Towards Visual Insect Camera Traps

Conclusion and Future Work

- Spatial information during training and temporal cues in the input are very helpful to the detection mechanism.
- Test additional state-of-the-art architectures made for sequential data
- Expand dataset (more adverse weather conditions, etc.)
- Improve the hardware (field of view, field of depth, etc.)