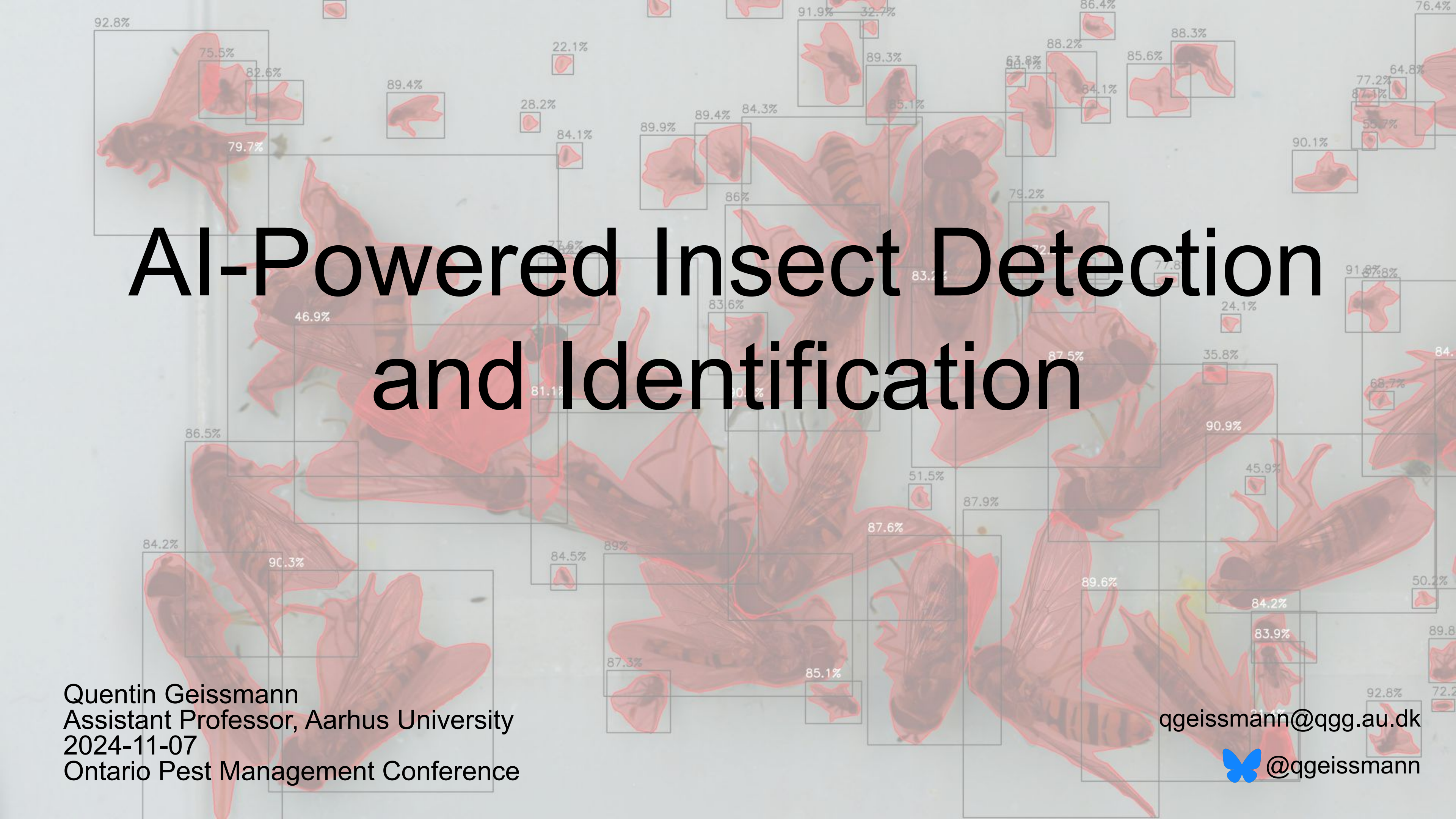


AI-Powered Insect Detection and Identification

Quentin Geissmann
Assistant Professor, Aarhus University
2024-11-07
Ontario Pest Management Conference

qgeissmann@qgg.au.dk
@qgeissmann

The rise of AI in biology

Clinical Infectious Diseases

INVITED ARTICLE

HEALTHCARE EPIDEMIOLOGY: Robert A. Weinstein, Section Editor

Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology

Jenna Wiens[1] and Erica S. Shenoy[2,3,4]

---

remote sensing

MDPI

Perspective

Artificial Intelligence Revolutionises Weather Forecast, Climate Monitoring and Decadal Prediction

Steven Dewitte [1,*], Jan P. Cornelis [2], Richard Müller [3] and Adrian Munteanu [2]

204 | Nature | Vol 588 | 10 December 2020

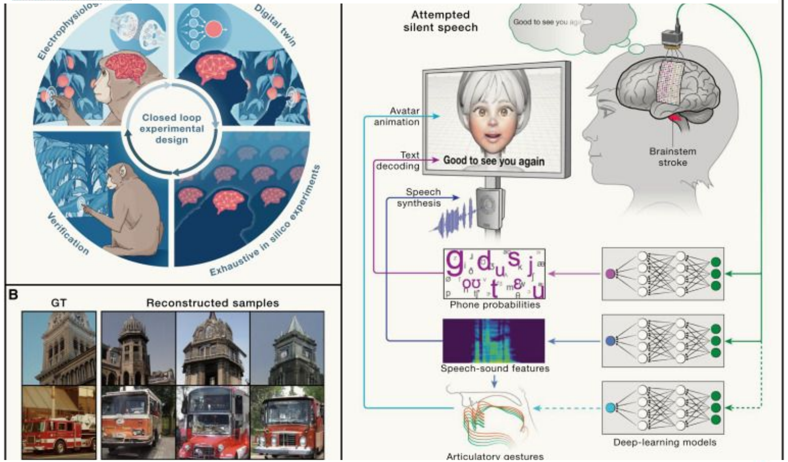
‘IT WILL CHANGE EVERYTHING’: AI MAKES GIGANTIC LEAP IN SOLVING PROTEIN STRUCTURES

DeepMind’s program for determining the 3D shapes of proteins stands to transform biology, say scientists.

The rise of AI in biology

Decoding the brain: From neural representations to mechanistic models

Mackenzie Weygandt Mathis[1,2] · Adriana Perez Rotondo[1,2] · Edward F. Chang[3] · Andreas S. Tolias[4,5,6,7] · Alexander Mathis[1,2]



nature methods

Article

https://doi.org/10.1038/s41592-024-02201-0

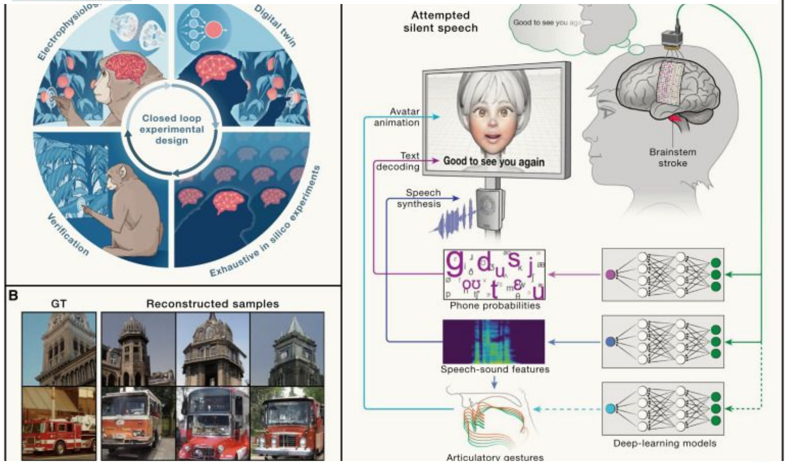
scGPT: toward building a foundation model for single-cell multi-omics using generative AI

The rise of AI in biology

- More data from sensors
- Smart phones
    - Hobbyism (e.g., Raspberry pi)
    - Connectivity
    - More data (images, text, etc)

Decoding the brain: From neural representations to mechanistic models

Mackenzie Weygandt Mathis[1,2] · Adriana Perez Rotondo[1,2] · Edward F. Chang[3] · Andreas S. Tolias[4,5,6,7] · Alexander Mathis[1,2]



nature methods

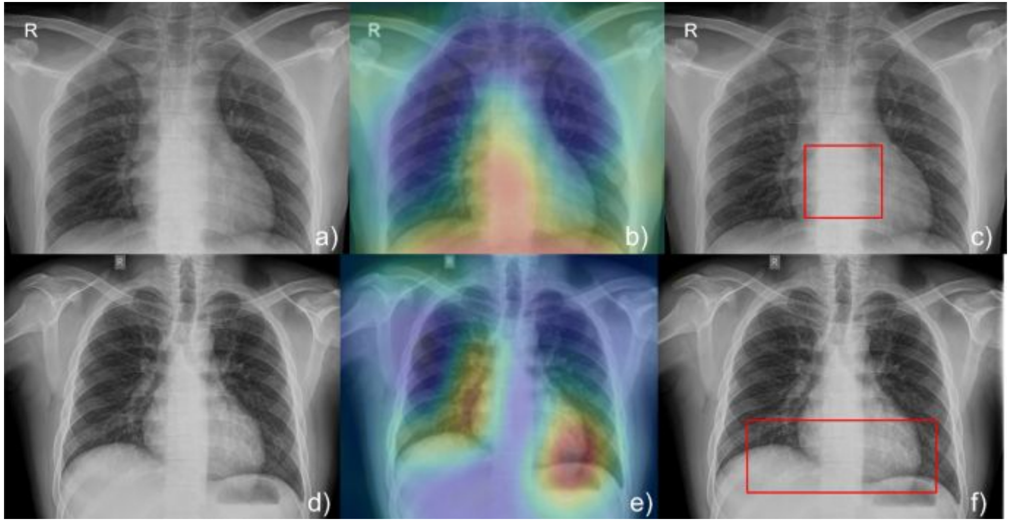
Article

https://doi.org/10.1038/s41592-024-02201-0

scGPT: toward building a foundation model for single-cell multi-omics using generative AI

The rise of AI in biology

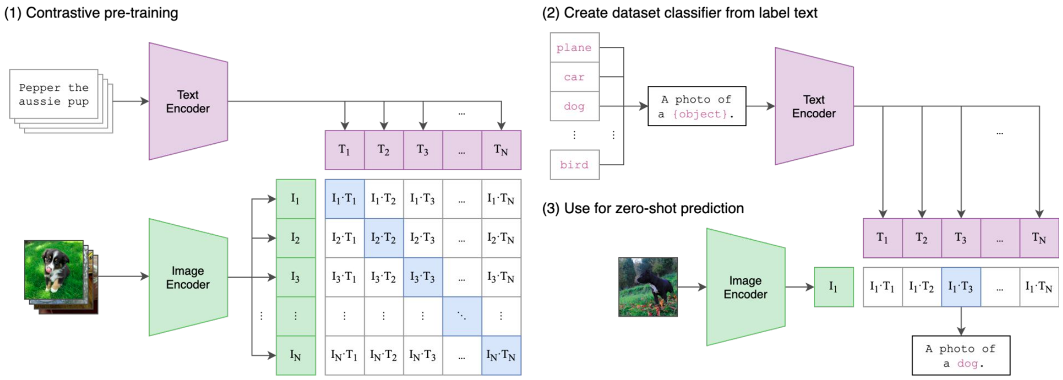
- More data from sensors
- Smart phones
    - Hobbyism (e.g., Raspberry pi)
    - Connectivity
    - More data (images, text, etc)
- Neural Networks > Digital signal processing



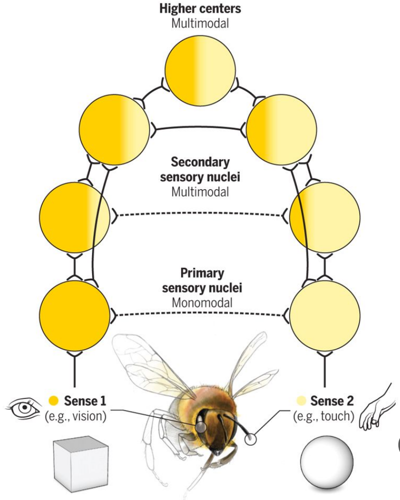
Rangarajan *et al.*, 2021

The rise of AI in biology

- More data from sensors
- Smart phones
  - Hobbyism (e.g., Raspberry pi)
  - Connectivity
  - More data (images, text, etc)
- Neural Networks > Digital signal processing
- From traditional to "fundational"



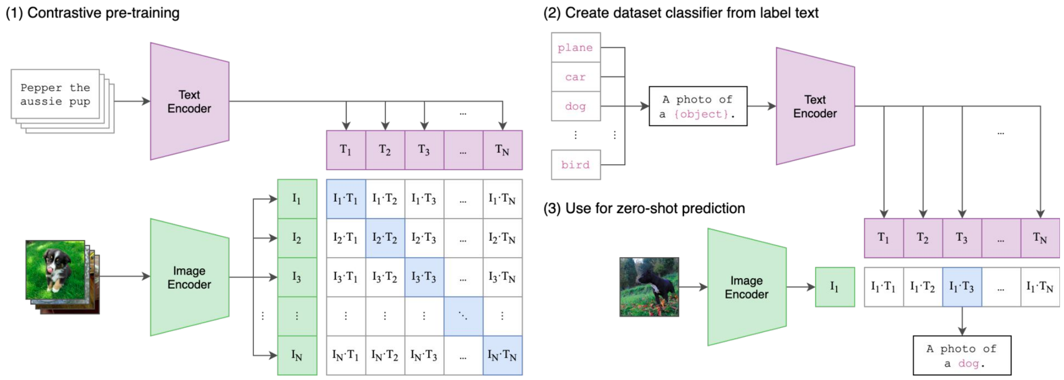
CLIP (OpenAI)



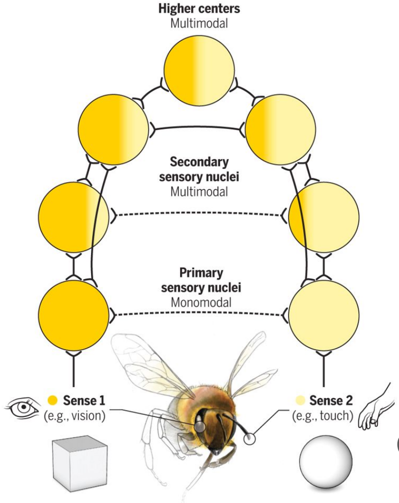
Cross modal learning, Solvi *et al.*, 2020

The rise of AI in biology

- More data from sensors
- Smart phones
  - Hobbyism (e.g., Raspberry pi)
  - Connectivity
  - More data (images, text, etc)
- Neural Networks > Digital signal processing
- From traditional to "fundational"
- Is "AI" just "statistical learning"?
- From explicit to implicit
- Expectations and promises
- Lower cost / higher throughput
- New content / insight



CLIP (OpenAI)

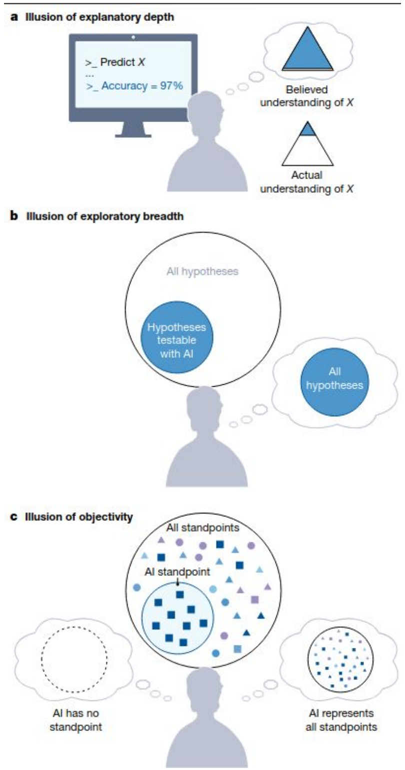


Cross modal learning, Solvi *et al.*, 2020

The rise of AI in biology

- More data from sensors
- Smart phones
  - Hobbyism (e.g., Raspberry pi)
  - Connectivity
  - More data (images, text, etc)
- Neural Networks > Digital signal processing
- From traditional to "fundational"
- Is "AI" just "statistical learning"?
- From explicit to implicit
- Expectations and promises
- Lower cost / higher throughput
- New content / insight
- Fears
  - Black box, “data leakage” & biases
  - Societal and environmental issues

“The proliferation of AI tools in science risks introducing a phase of scientific enquiry in which we produce more but understand less”



Artificial intelligence and illusions of understanding in scientific research

Lisa Messeri[1,4] & M. J. Crockett[2,3,4]

The expectations of AI for entomology

- High expectations /optimism
  - Eng. vs traditional entomology
  - Can we make the most of AI?
- Phenomics
- The robotic entomologist
- The smart insect trap

Deep learning and computer vision will transform entomology

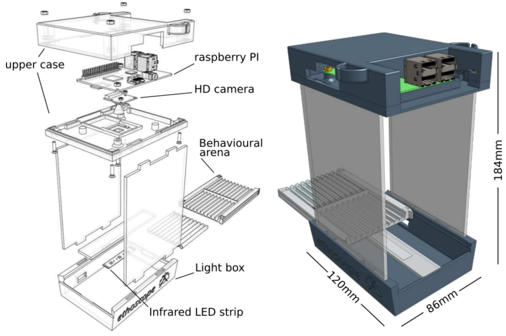
Toke T. Høye, Johanna Ärje, Kim Bjerge, +6, and Jenni Raitoharju Authors Info & Affiliations

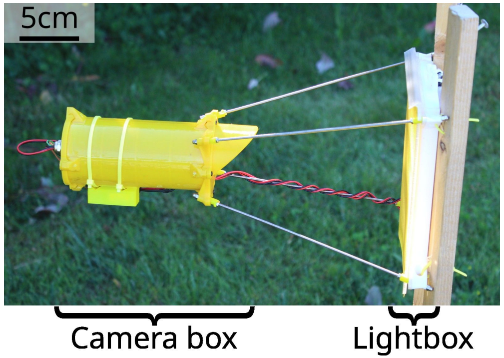
Edited by Matthew L. Forister, University of Nevada, Reno, NV, and accepted by Editorial Board Member May R. Berenbaum October 23, 2020 (received for review March 24, 2020)

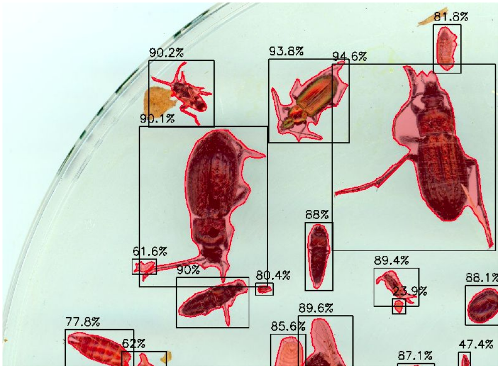
Preti et al. 2020

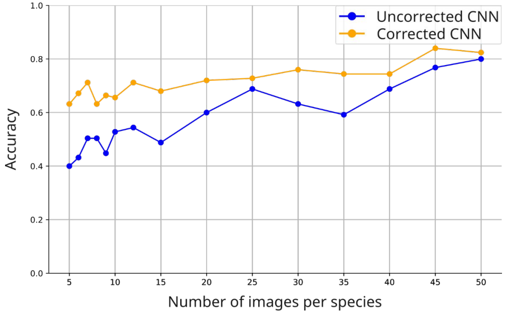
A few examples from my group

- Ethoscope: Insect behaviour in the lab
- Sticky Pi: Monitoring insects captures on sticky cards
- Flat Bug: "Universal" insect detection tool
- Size-aware insect classification
- miniMon: Ultra-low-cost camera









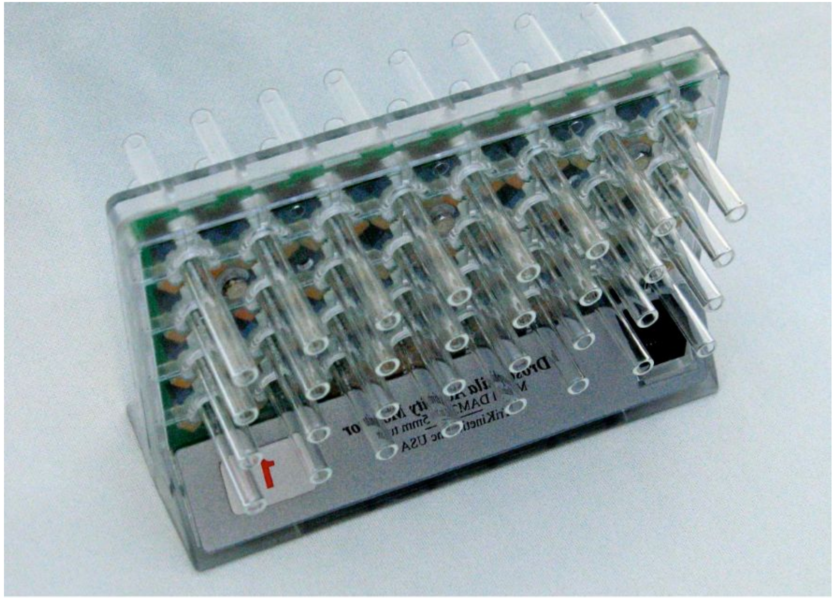
Ethoscope

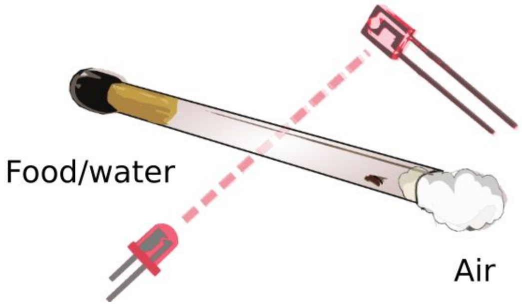
- When/why and how do flies sleep?

Nobel 2017 to Jeffrey Hall (left), Michael Young (centre) and Michael Rosbash (right) for their work on circadian clocks.

Ethoscope

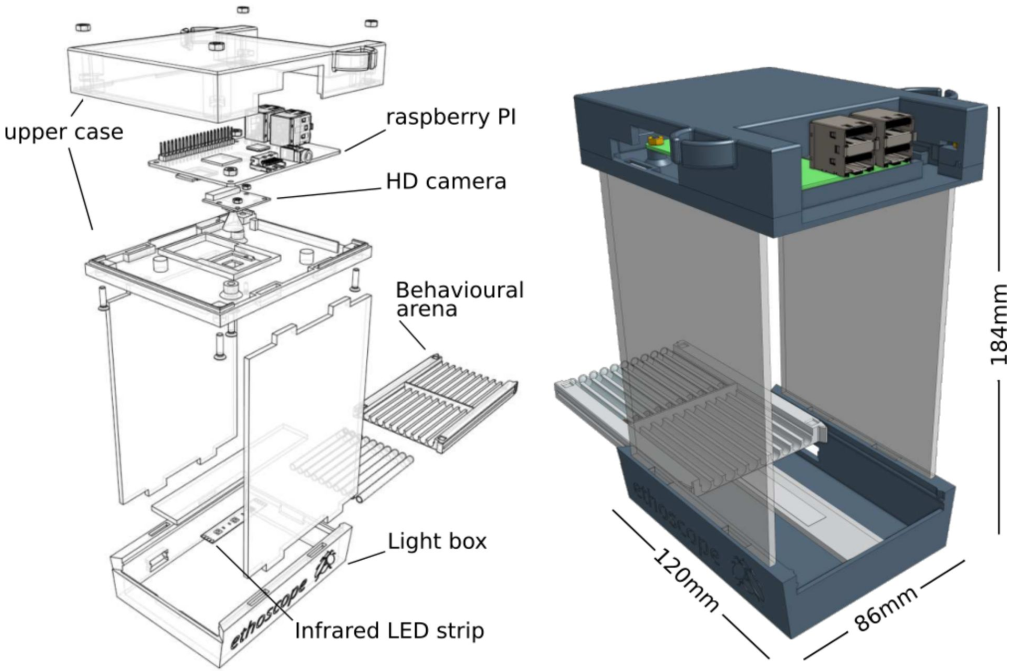
- When/why and how do flies sleep?





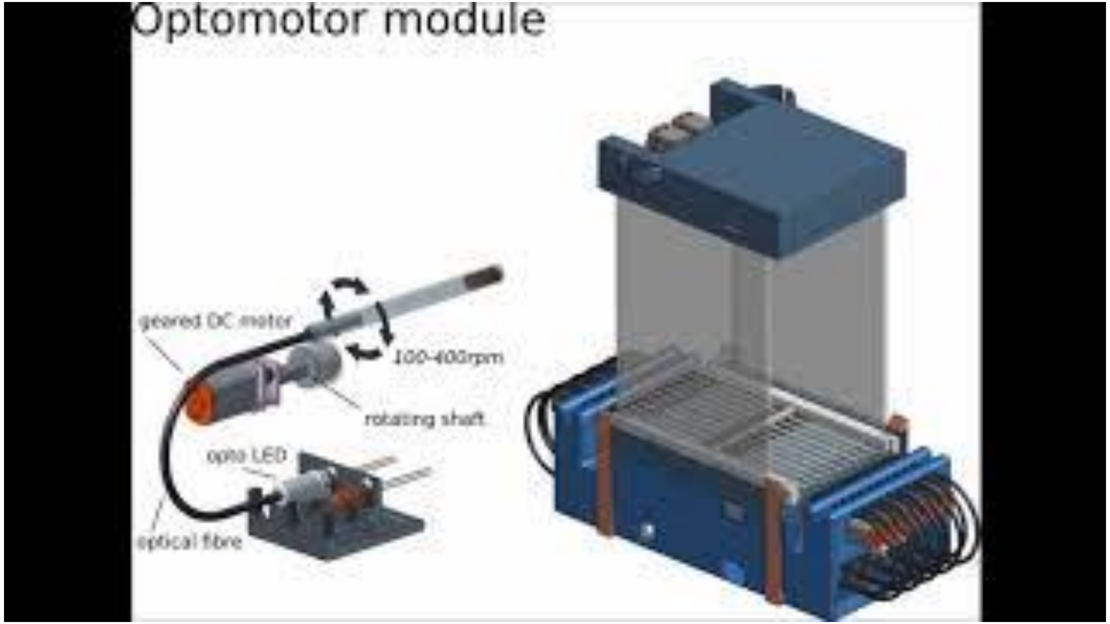
Ethoscope

- When/why and how do flies sleep?
- Microcomputers, cameras and 3d printers
- Movement classification with "image processing"



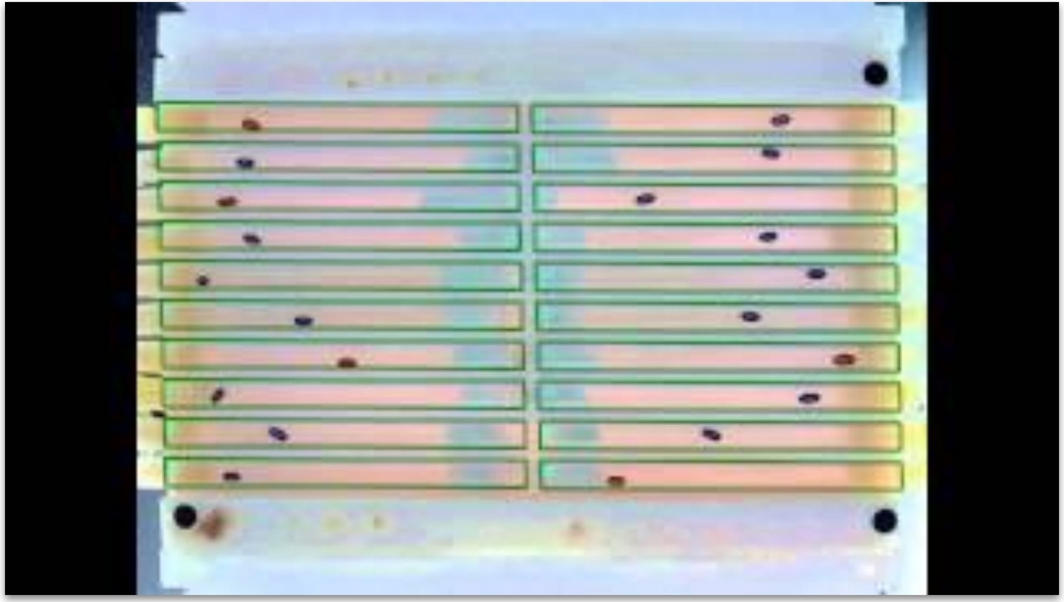
Ethoscope

- When/why and how do flies sleep?
- Microcomputers, cameras and 3d printers
- Movement classification with "image processing"



Ethoscope

- When/why and how do flies sleep?
- Microcomputers, cameras and 3d printers
- Movement classification with "image processing"



Ethoscope

- When/why and how do flies sleep?
- Microcomputers, cameras and 3d printers
- Movement classification with "image processing"
- Robotics and closed-loop system
- All open-source and free
- Many follow ups:
  - R packages
  - Mating-sleep interactions
  - Effect of chronic sleep dep.
  - Olfaction during sleep

Regulation of sleep homeostasis by sexual arousal

Esteban J Beckwith, Quentin Geissmann, Alice S French, Giorgio F Gilestro*

Department of Life Sciences, Imperial College London, London, United Kingdom

SCIENCE ADVANCES | RESEARCH ARTICLE

COGNITIVE NEUROSCIENCE

Most sleep does not serve a vital function: Evidence from *Drosophila melanogaster*

Quentin Geissmann*, Esteban J. Beckwith*, Giorgio F. Gilestro†

Sensory processing during sleep in *Drosophila melanogaster*

Alice S. French, Quentin Geissmann, Esteban J. Beckwith & Giorgio F. Gilestro

Nature **598**, 479–482 (2021) | Cite this article

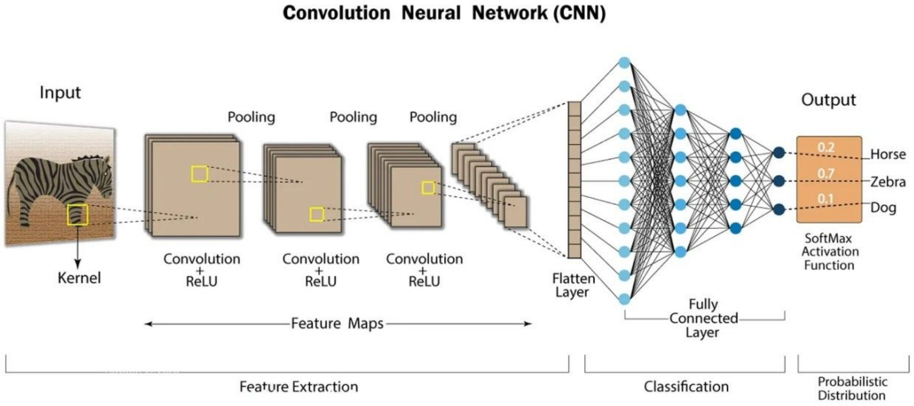
RESEARCH ARTICLE

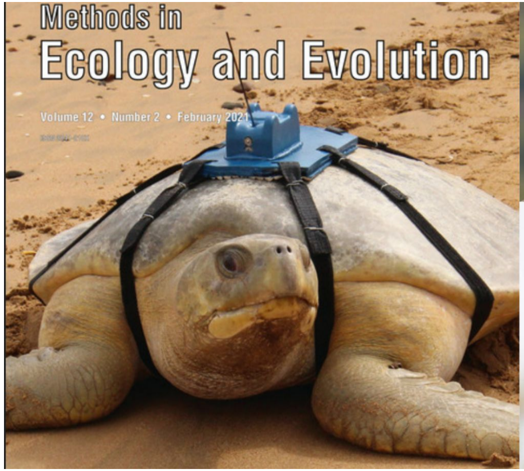
Rethomics: An R framework to analyse high-throughput behavioural data

Quentin Geissmann[1]*, Luis Garcia Rodriguez[2], Esteban J. Beckwith[1], Giorgio F. Gilestro[1]*

Sticky Pi

- From physiology to ecology
- From CV to Deep Learning



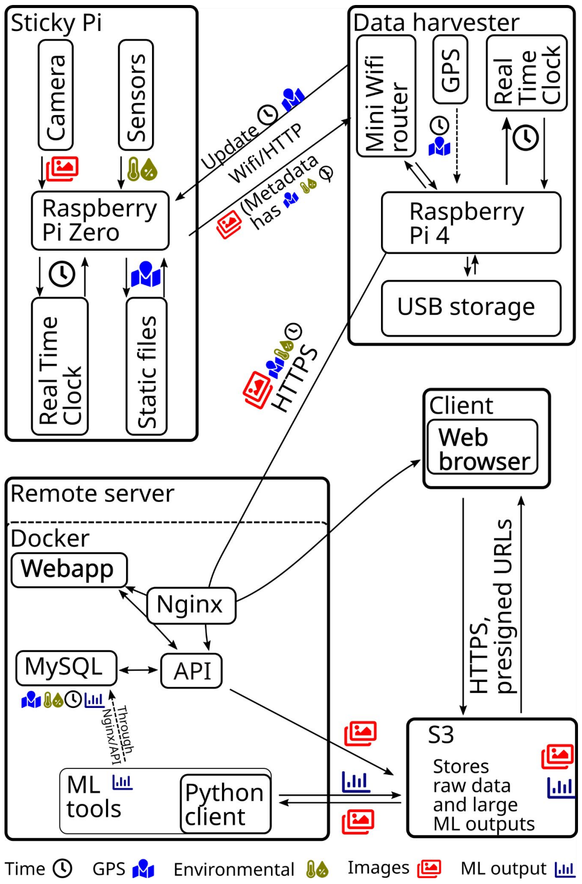


Sticky Pi

- From physiology to ecology
- From CV to Deep Learning
- Can we detect insect automatically, at high frequency, and know when, where and which insects are active?
- Why not put a camera in front of a sticky card?

Sticky Pi

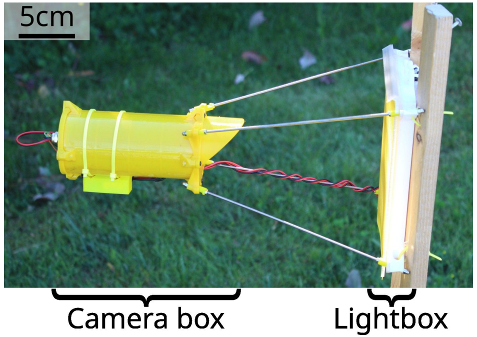
- From physiology to ecology
- From CV to Deep Learning
- Can we detect insect automatically, at high frequency, and know when, where and which insects are active?
- Why not put a camera in front of a sticky card?
- Different parts:
  - Hardware
  - Software
  - App/Cloud
  - Data processing

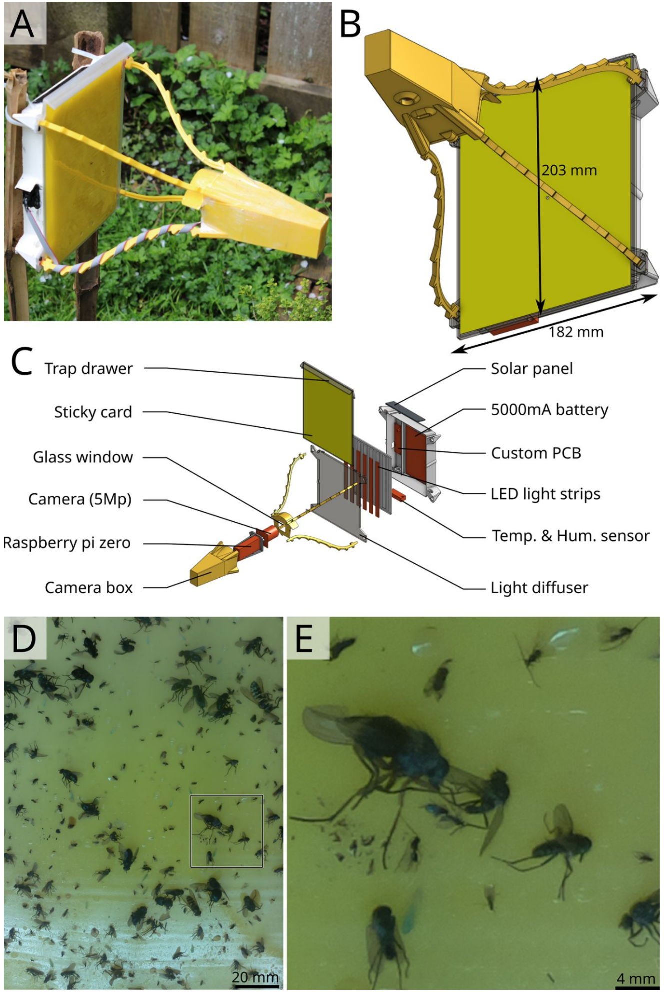


Geissmann *et al.*, PLoS Biol., 2022

Sticky Pi – Hardware

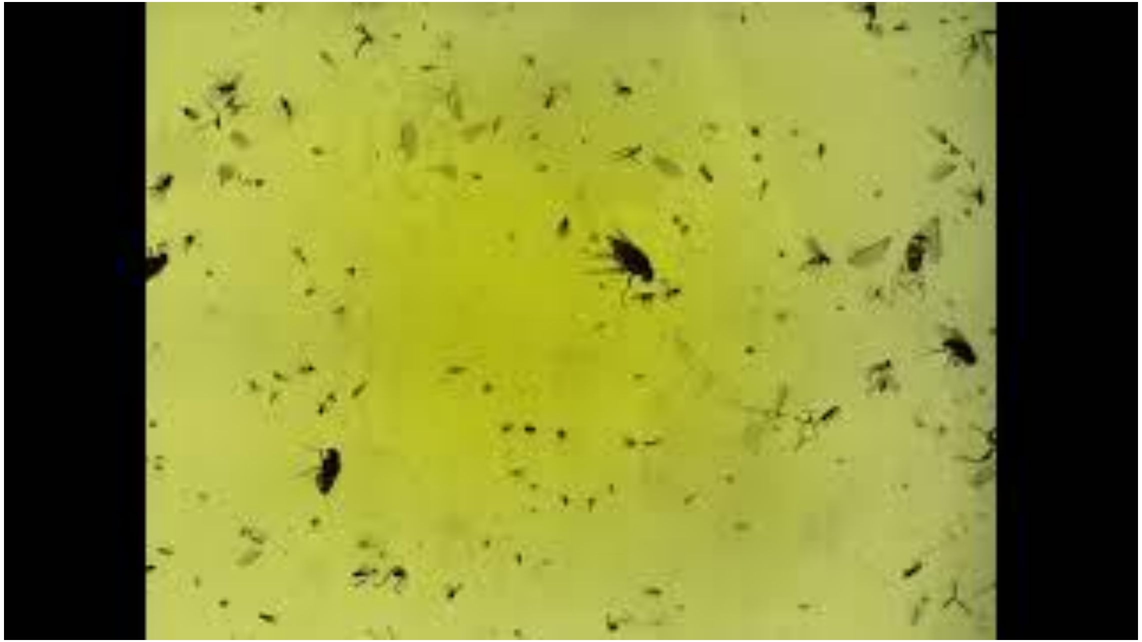
- Picture (16MP) every 20min
- Flash light
- Low-cost (200$)
- Open-source and [documented]
- Battery-powered (2 weeks)
- Temperature and humidity
- Protocol
  - Deployment
  - Retrieval/ card archiving





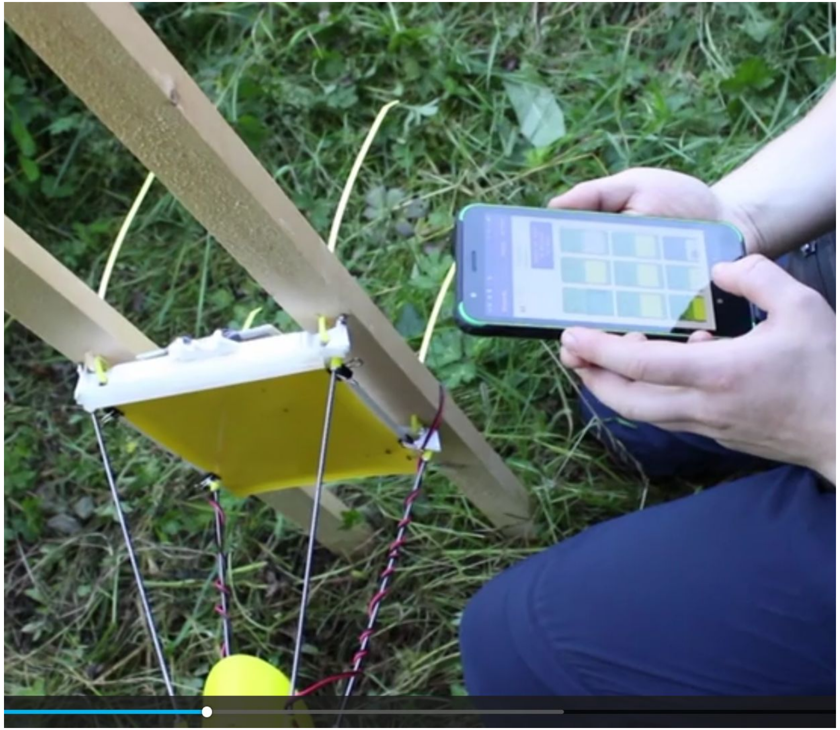
Sticky Pi – Hardware

- Picture (16MP) every 20min
- Flash light
- Low-cost (200$)
- Open-source and documented
- Battery-powered (2 weeks)
- Temperature and humidity
- Protocol
    - Deployment
    - Retrieval/ card archiving



Sticky Pi – App & cloud

- Android App to get data on site
- Send data to custom cloud service
    -



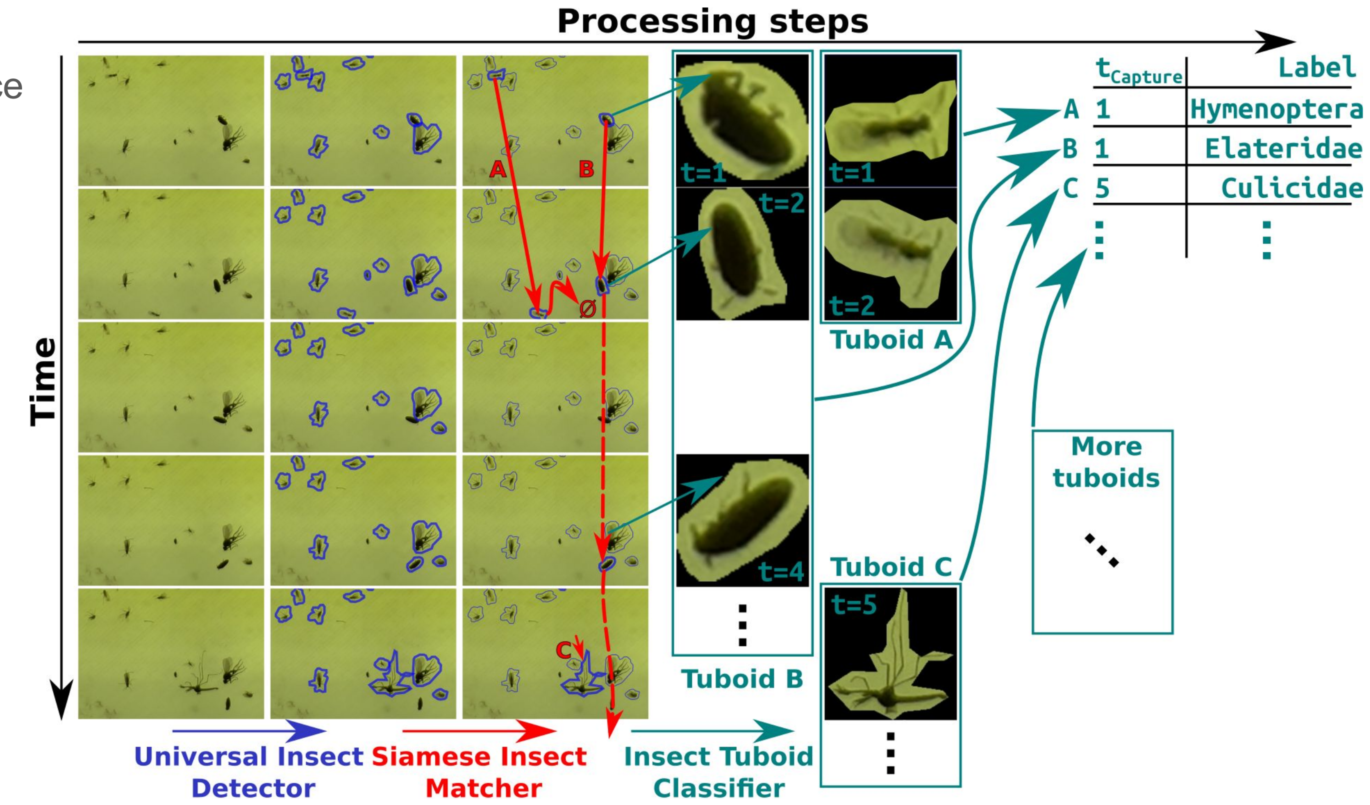
https://doc.sticky-pi.com/user-manual.html

Sticky Pi – App & cloud

- Android App to get data on site
- Send data to custom cloud service
- RShiny website
  - Quality control
  - Centralised interface

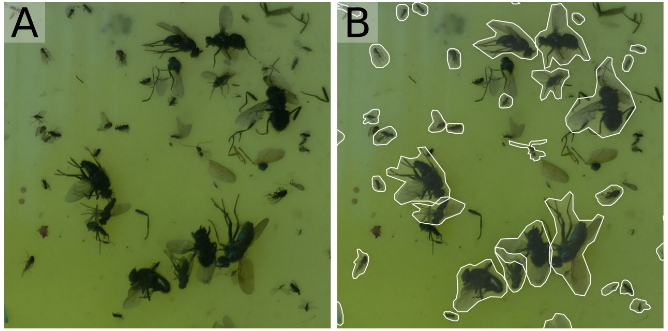
Sticky Pi – Data processing

- 500 imgs/week/device
- Why we cannot "just classify"
- Instead:
  - Detect
  - Track
  - Classify



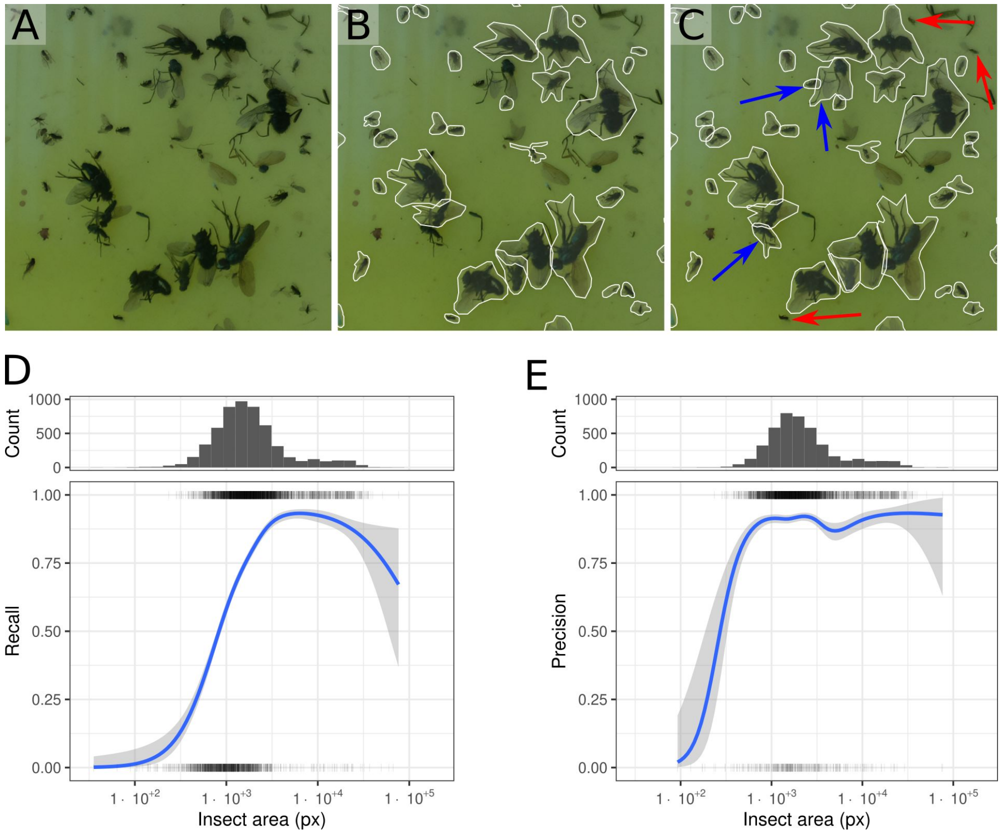
Sticky Pi – Detect

- Define task:
    - “Instance segmentation”
    - 1 "class": insect (vs BG)
- Create dataset
    - Annotate
    - Split (training/validation)
- Train algorithm
    - Augmentation
    - Loss function(s)
- Evaluate



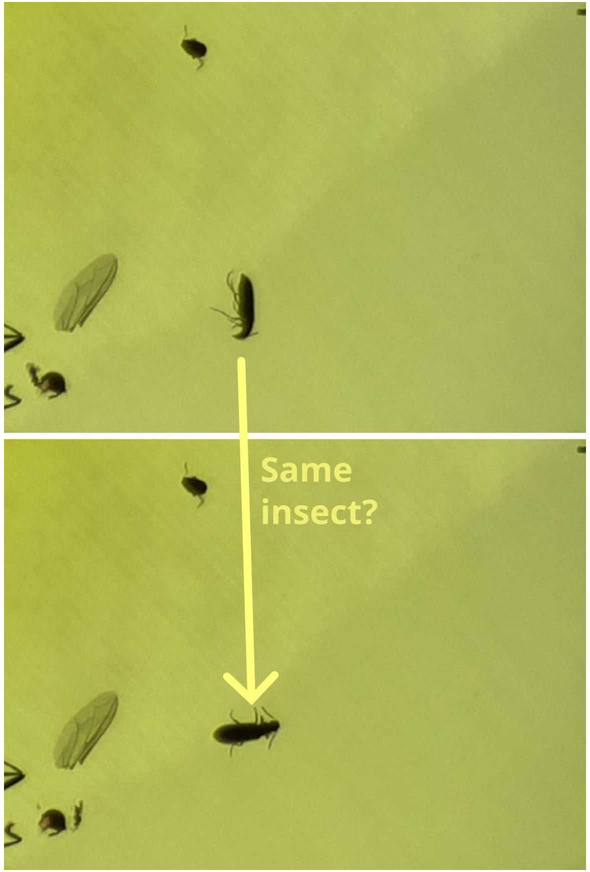
Sticky Pi – Detect

- Define task:
  - “Instance segmentation”
  - 1 "class": insect (vs BG)
- Create dataset
  - Annotate
  - Split (training/validation)
- Train algorithm
  - Augmentation
  - Loss function(s)
- Evaluate



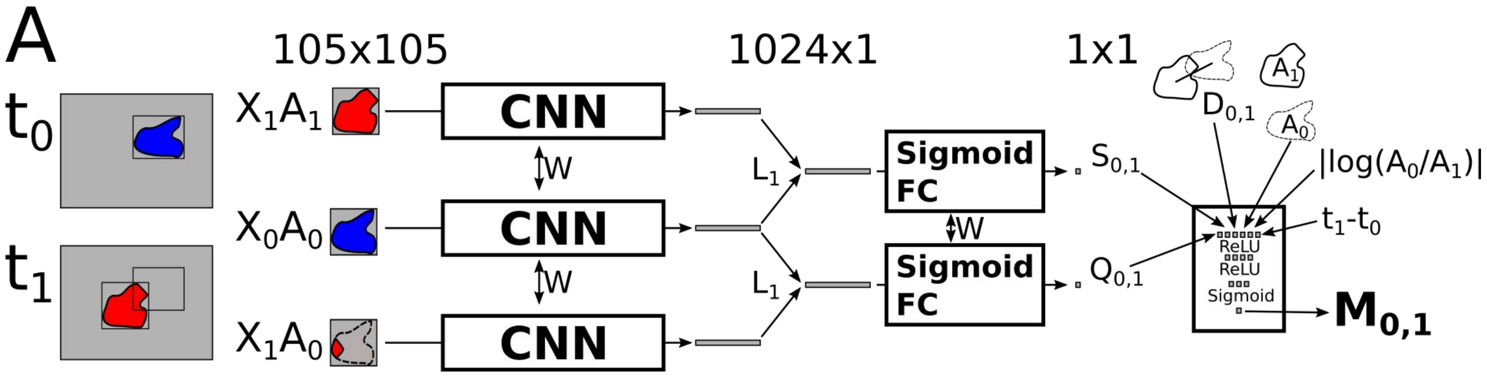
Sticky Pi – Track

- Define task:
  - Is it the same insect?
  - "Stitch" same bugs over time
- Create dataset
  - Annotate
  - Split (training/validation)
- Train algorithm
  - Augmentation
  - Loss function(s)



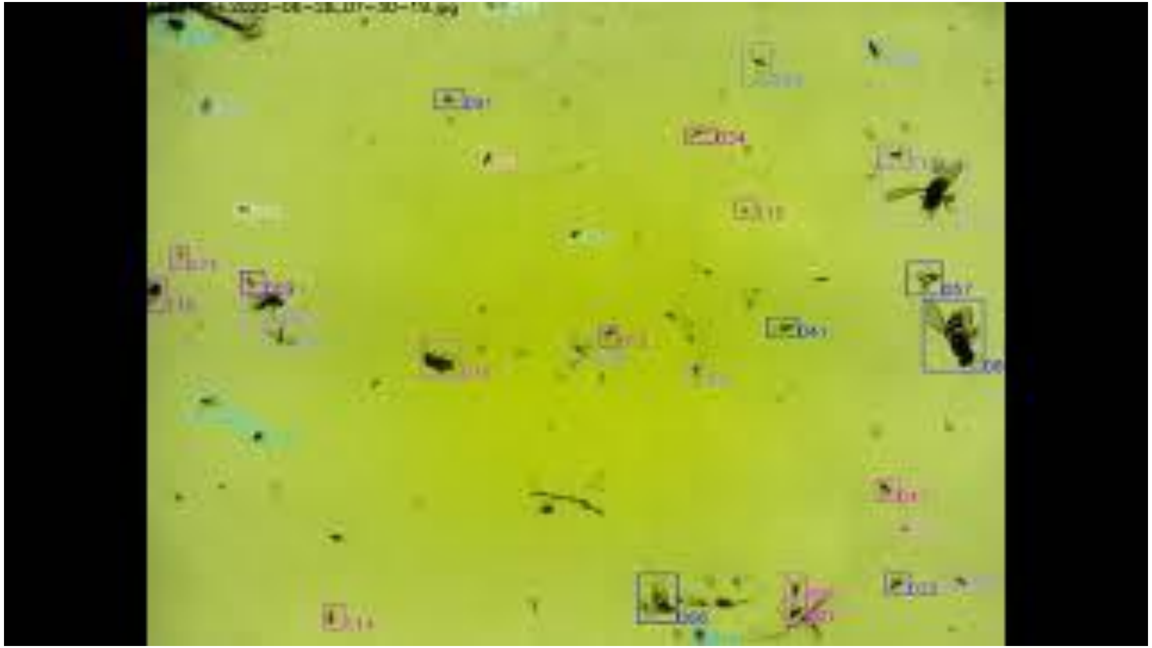
Sticky Pi – Track

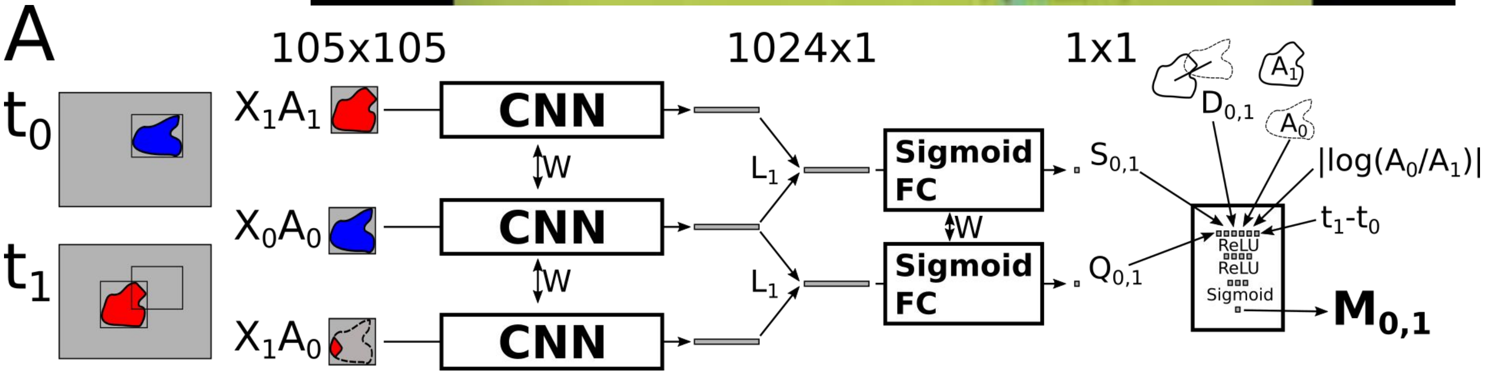
- Define task:
  - Is it the same insect?
  - "Stitch" same bugs over time
- Create dataset
  - Annotate
  - Split (training/validation)
- Train algorithm
  - Augmentation
  - Loss function(s)



Sticky Pi – Track

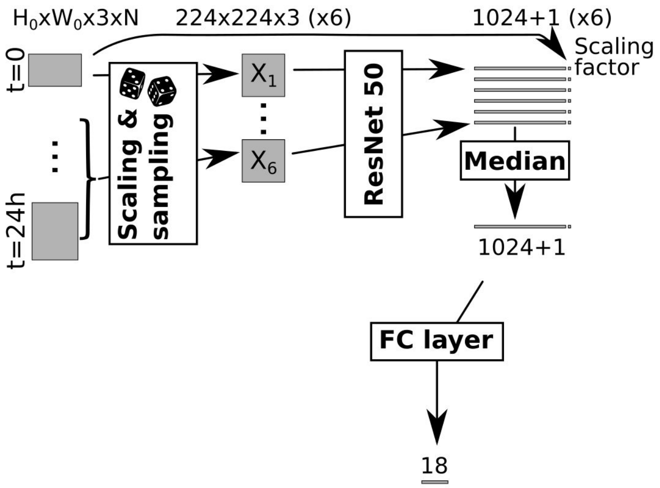
- Define task:
  - Is it the same insect?
  - "Stitch" same bugs over time
- Create dataset
  - Annotate
  - Split (training/validation)
- Train algorithm
  - Augmentation
  - Loss function(s)

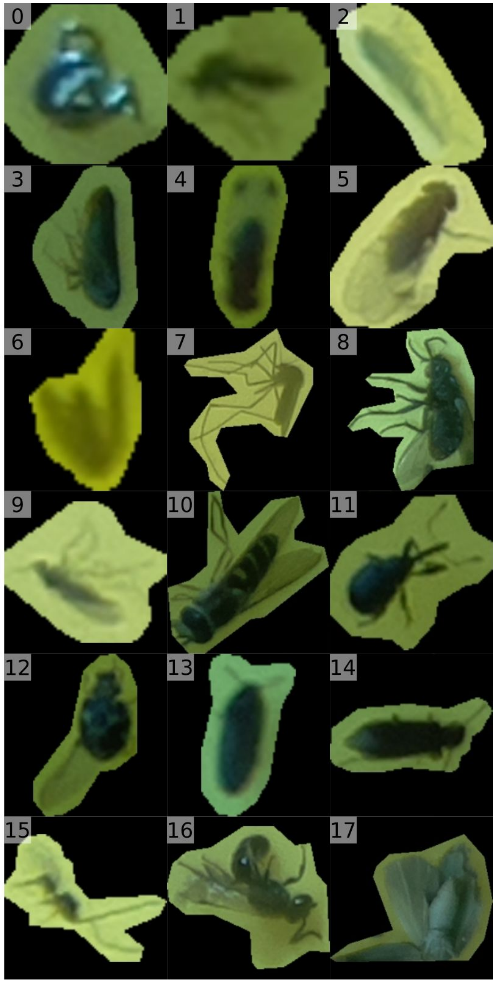




Sticky Pi – Classify

- Define task:
  - Which taxa is it?
- Create dataset
  - Annotate
  - Split (training/validation)
- Train algorithm
  - Augmentation
  - Loss function(s)





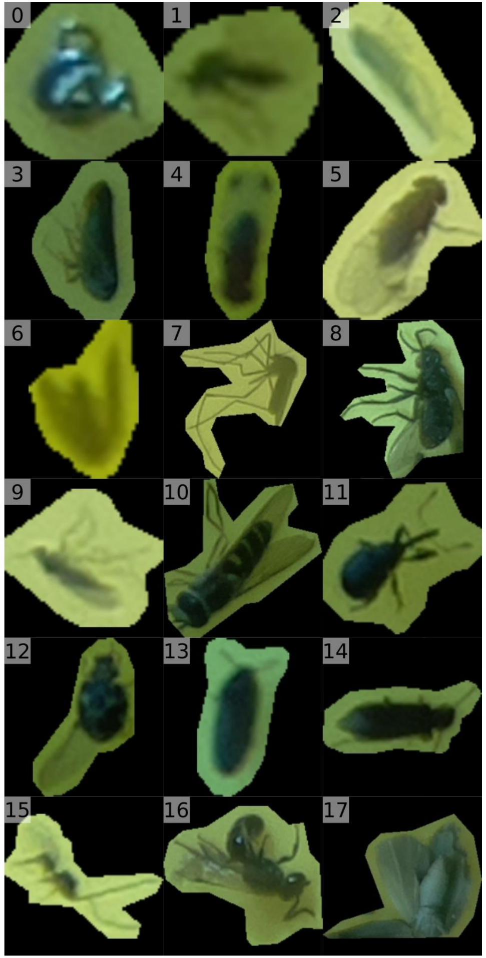
Sticky Pi – Classify

- Define task:
  - Which taxa is it?
- Create dataset
  - Annotate
  - Split (training/validation)
- Train algorithm
  - Augmentation
  - Loss function(s)
- Evaluate

C

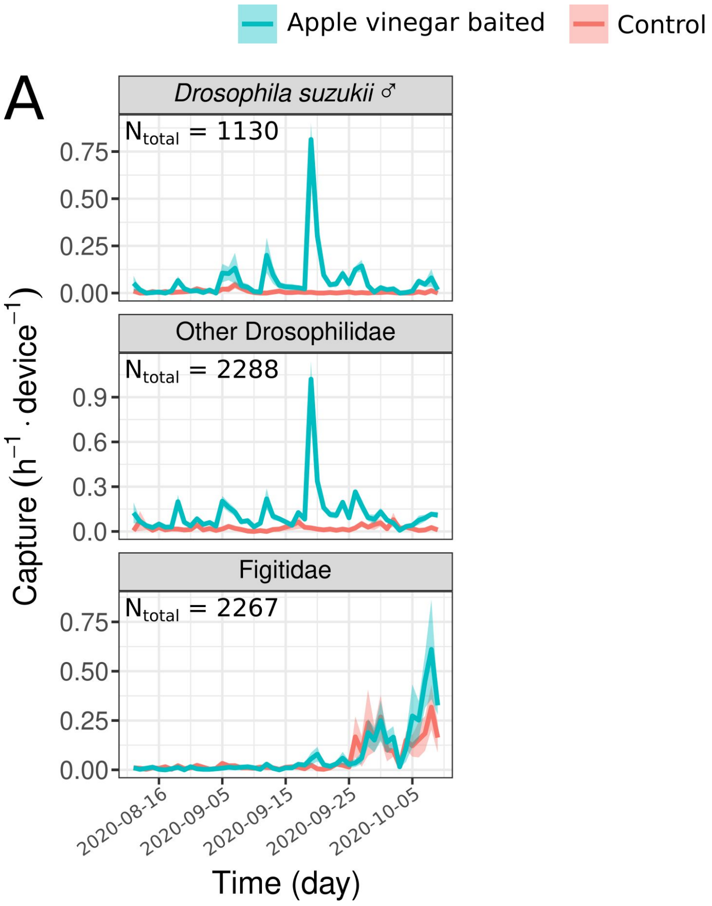
	Precision	Recall	f1-score	N
Background	**0.75**	**0.62**	**0.68**	**159**
Insecta	**0.68**	**0.71**	**0.69**	**220**
Edwardsiana	**0.92**	**0.94**	**0.93**	**95**
Cicadellidae	**1.00**	**0.89**	**0.94**	**9**
Male *D. suzukii*	**0.86**	**0.91**	**0.89**	**47**
Drosophilidae	**0.76**	**0.82**	**0.79**	**99**
Psychodidae	**0.90**	**0.92**	**0.91**	**61**
Culicidae	**0.65**	**0.85**	**0.73**	**13**
Muscidoidae	**0.96**	**0.87**	**0.91**	**86**
Sciaridae	**0.73**	**0.84**	**0.78**	**49**
Syrphidae	**0.67**	**0.75**	**0.71**	**8**
Curculionidae	**1.00**	**0.60**	**0.75**	**5**
Coccinellidae	**0.84**	**1.00**	**0.91**	**21**
Elateridae	**0.92**	**0.92**	**0.92**	**13**
Other Coleoptera	**0.46**	**0.55**	**0.50**	**11**
Figitidae	**0.79**	**0.77**	**0.78**	**70**
Halictidae	**0.57**	**0.80**	**0.67**	**10**
Lepidoptera	**1.00**	**0.33**	**0.50**	**6**

D



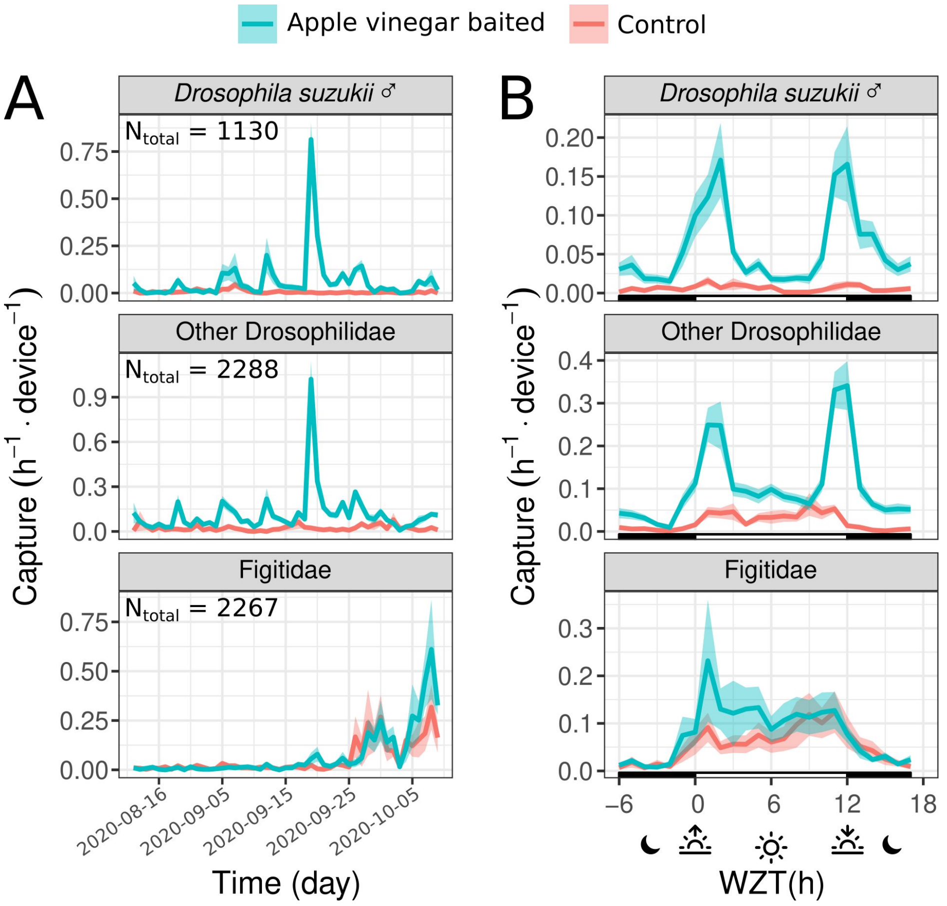
Sticky Pi – Results

- Monitor *D. suzukii* levels over time
  - Retrieve circadian behaviour
  - Seasonal trends
  - Bait kinetics
  - TODO: Weather -> Capture



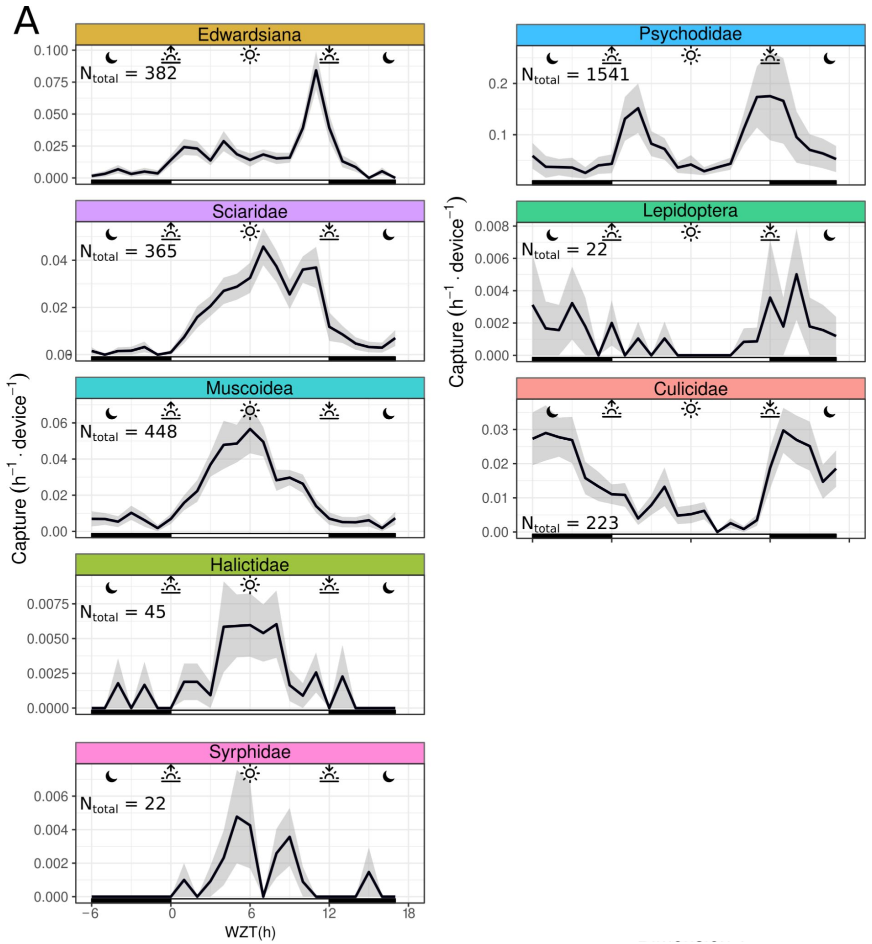
Sticky Pi – Results

- Monitor *D. suzukii* levels over time
  - Retrieve circadian behaviour
  - Seasonal trends
  - Bait kinetics
  - TODO: Weather -> Capture



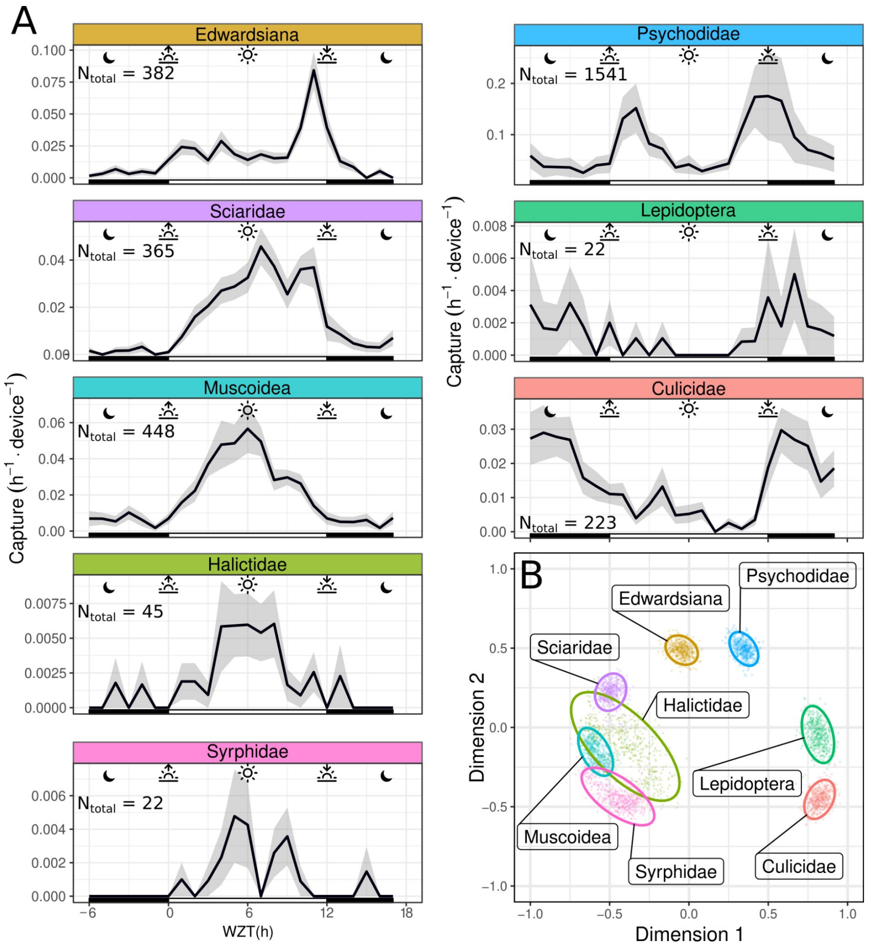
Sticky Pi – Results

- Monitor *D. suzukii* levels over time
  - Retrieve circadian behaviour
  - Seasonal trends
  - Bait kinetics
  - TODO: Weather -> Capture
- Study “temporal niche” partitioning
  - Replace anecdotal observations with numbers



Sticky Pi – Results

- Monitor *D. suzukii* levels over time
  - Retrieve circadian behaviour
  - Seasonal trends
  - Bait kinetics
  - TODO: Weather -> Capture
- Study “temporal niche” partitioning
  - Replace anecdotal observations with numbers



Collaborators at UBC (2018-2022)


Haney Lab


PIEE Lab


Paul Abram


Warren Wong

Flat Bug

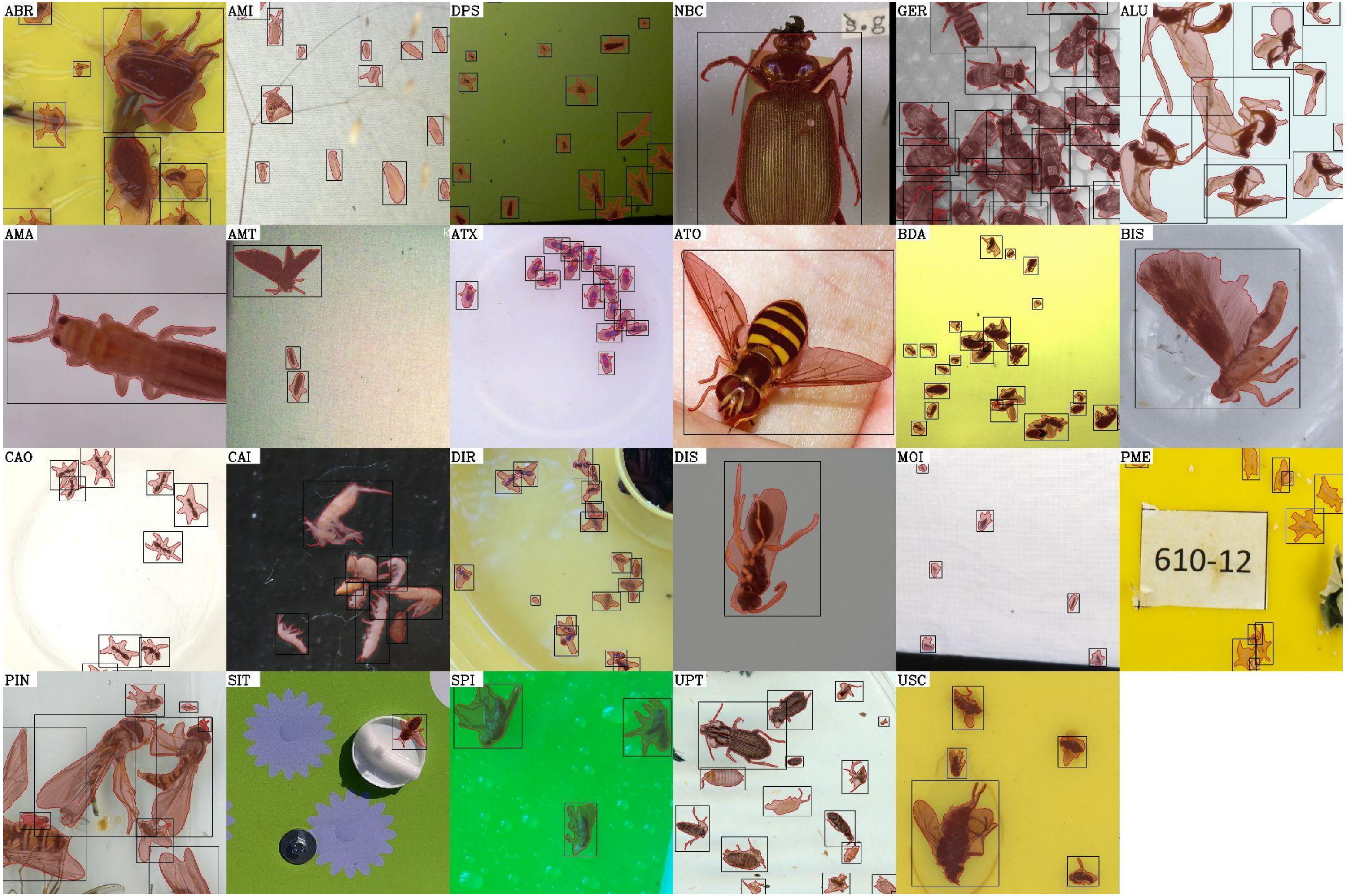
- Dozens of studies create a new Detection tool from scratch, on their one data
- This task is general, so we all reinvent the wheel
- Can we create a larger scope detector?
    - >20 diverse datasets
    - New algorithm (YOLO based)
    - Focus on multiscale (large or small images)
- High performance due to generalisation
- In preparation, preprint in december?

Collab. w/ Toke Høye & Asger Svenning
(AU, Ecoscience)

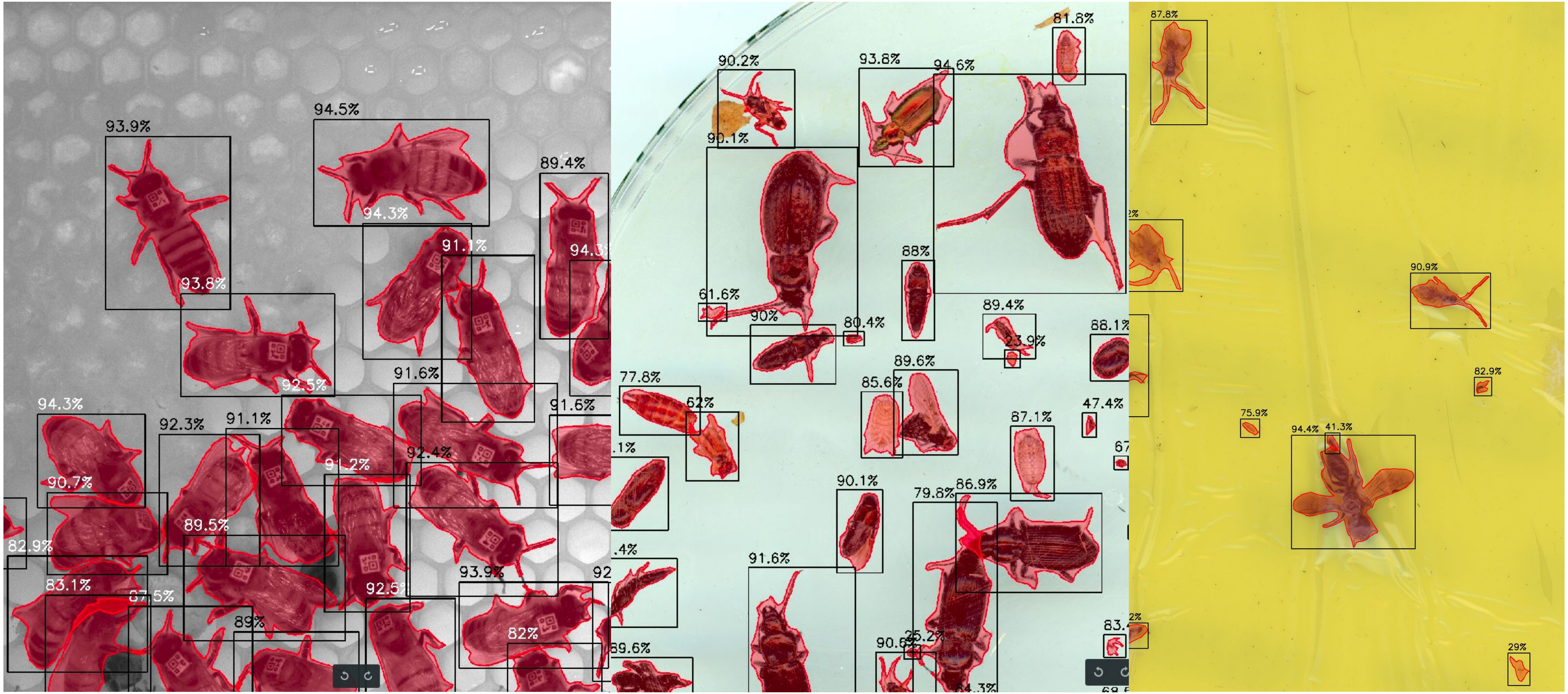
Flat Bug


ABR
AMI
DPS
NBC
GER
ALU
AMA
AMT
ATX
ATO
BDA
BIS
CAO
CAI
DIR
DIS
MOI
PME
610-12
PIN
SIT
SPI
UPT
USC

Flat Bug


ABR
AMI
DPS
NBC
GER
ALU
AMA
AMT
ATX
ATO
BDA
BIS
CAO
CAI
DIR
DIS
MOI
PME
610-12
PIN
SIT
SPI
UPT
USC

Flat Bug



93.9%
94.5%
94.3%
89.4%
91.1%
93.8%
92.5%
91.6%
94.3%
92.3%
91.1%
91.2%
92.4%
90.7%
89.5%
82.9%
83.1%
89%
93.9%
82%
90.2%
93.8%
94.6%
81.8%
90.1%
88%
61.6%
90%
80.4%
89.4%
23.9%
88.1%
89.6%
77.8%
62%
85.6%
47.4%
87.1%
90.1%
79.8%
86.9%
91.6%
89.6%
87.8%
90.9%
82.9%
75.9%
94.4%
41.3%
29%

Size-aware classification

- Out-of-the-box CNNs (e.g., ResNet) ignore size!

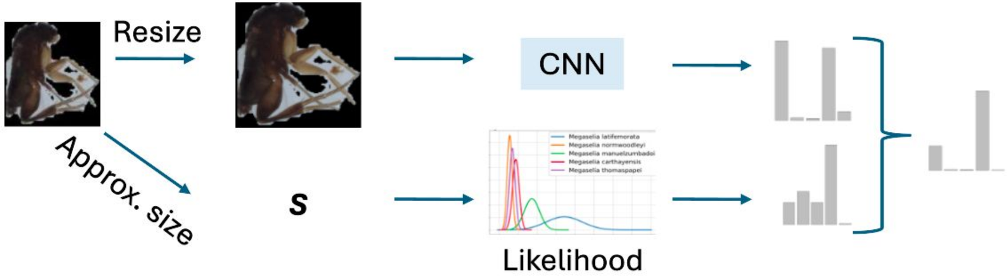
Melika Baghooee (PhD Student)

Size-aware classification

- Out-of-the-box CNNs (e.g., ResNet) ignore size!
- Insect size can help human identification (+ visual feature)
- Hypotheses:
  - Size-aware -> less data hungry
  - Better with class imbalance
  - Better performance when domain shift
  - More explainable

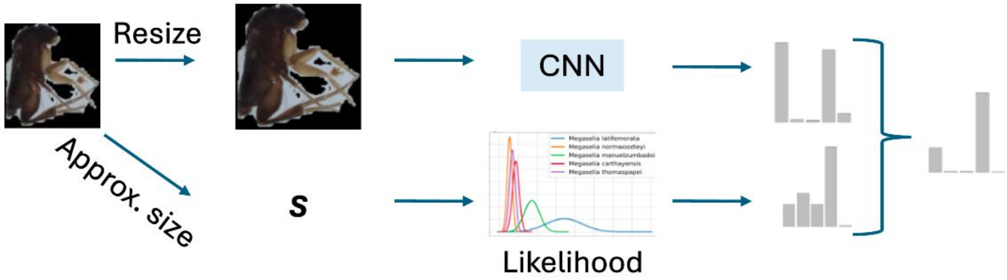
Size-aware classification

- Out-of-the-box CNNs (e.g., ResNet) ignor size!
- Insect size can help human identification (+ visual feature)
- Hypotheses:
  - Size-aware -> less data hungry
  - Better with class imbalance
  - Better performance when domain shift
  - More explainable
- Test on Bioscan1M dataset



Size-aware classification

- Out-of-the-box CNNs (e.g., ResNet) ignor size!
- Insect size can help human identification (+ visual feature)
- Hypotheses:
  - Size-aware -> less data hungry
  - Better with class imbalance
  - Better performance when domain shift
  - More explainable
- Test on Bioscan1M dataset
- Results



miniMon

- Devices are not getting cheaper
- More features:
  - Edge detection
  - Storage
  - Better camera
  - IoT / real time network upload

Eyes on nature: Embedded vision cameras for terrestrial biodiversity monitoring

Kevin F. A. Darras, Marcel Balle, Wenxiu Xu, Yang Yan, Vincent G. Zakka, Manuel Toledo-Hernández, Dong Sheng, Wei Lin, Boyu Zhang ... See all authors

First published: 28 October 2024 | https://doi.org/10.1111/2041-210X.14436

EcoEye: OpenMV H7+ Camera with Sensor Integration for Environmental Monitoring

$450.00 USD

Pre Order

- 1 +

CN Warehouse

Estimated availability Date: Nov 30, 2024

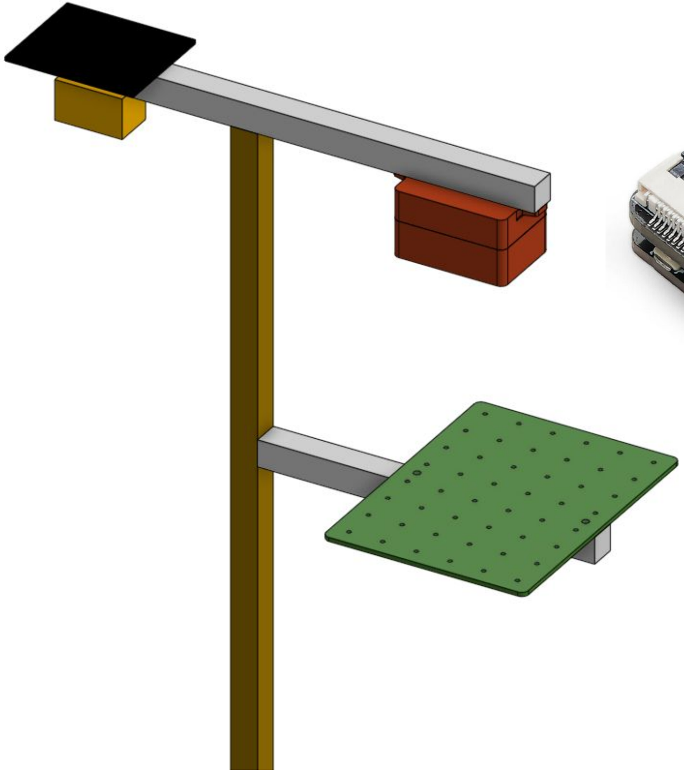
PRE ORDER

PayPal Buy Now

Also Add: $0.00

miniMon

- Devices are not getting cheaper
- More features:
  - Edge detection
  - Storage
  - Better camera
  - IoT / real time network upload
- Instead, we need replication!
  - 10 x $500 vs 100 x 50$?
- Target a $50 customisable timelapse
- Immediate applications
  - Pollinator monitoring
  - Earthworm foraging
  - Citizen science



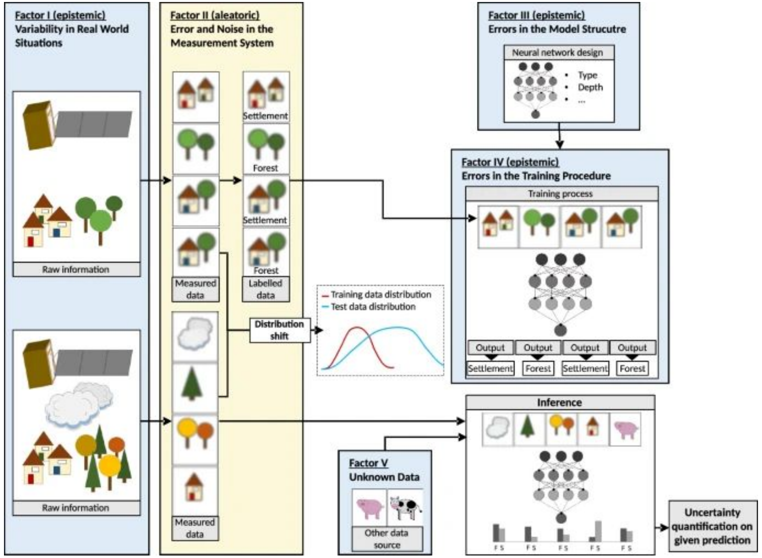
Helena Russello (Postdoc)

What can computational entomology do today

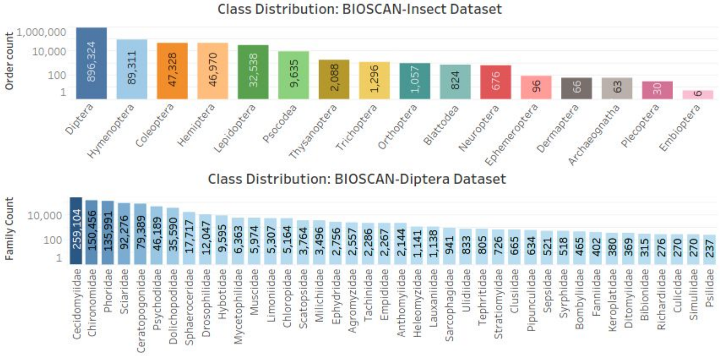
- Automate imaging at a **high cost**
- **In-domain** task automation:
  - Detection
  - Classification
  - Tracking
- “If it seems intuitive and you can teach a kid to do it in 10min”
- Semi-supervised:
  - Discover similarities
  - Soft labels

What computational entomology CANNOT do (yet)

- Difficult with domain shift
  - New camera
  - New field season
  - New background
- Classify “out of domain”
  - New taxa
- Assess how confident it is (often overconfident)
- Heavy class imbalance (1:1000)
- See more than available in the data
- Classify everything to species level
  - The black-box arrogance
  - Many species are very cryptic
  - Optimistic to infer species from single small image



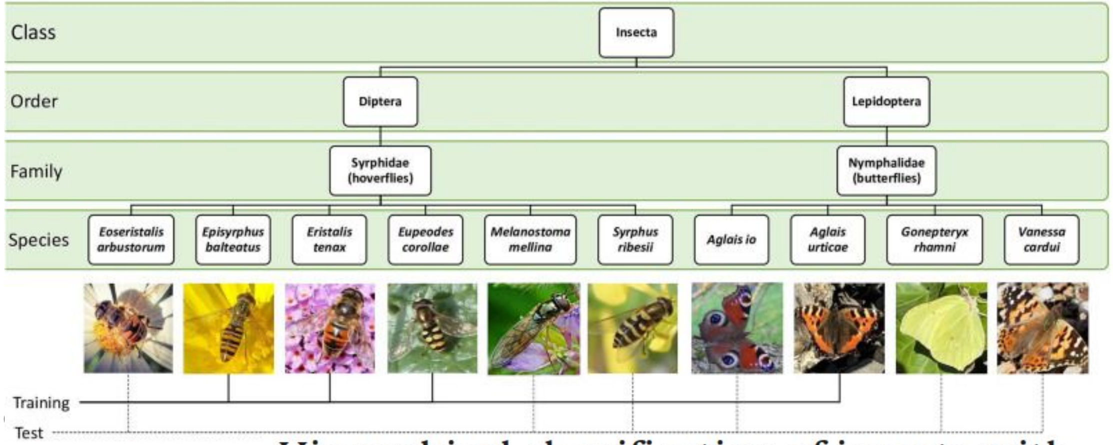
Gawlikowski et al., 2024



Gharaee et al., 2023

Where the is field going

- Edge detection
  - Why: real time / save storage
  - Detection (OK)
  - Classification (no please)
  - Hybrid
- Multimodal
  - Mix sound, images, etc
- Active perception
- Context dependent
  - Include time, space, abundance,
  - Be careful with priors
- Explainability and novelty detection
- Hierarchical classification
  - Insect are “nested classes”
  - Should leverage taxonomy
- Integrate DL output into statistical models
- How to put humans (or dogs?) in the loop



Hierarchical classification of insects with multitask learning and anomaly detection

Kim Bjerge [a], Quentin Geissmann [b], Jamie Alison [c], Hjalte M.R. Mann [c], Toke T. Høye [c d], Mads Dyrmann [a], Henrik Karstoft [a]

The Computational entomology community

- Training school: Aarhus Computational Entomology Summer School (ACCESS)
- Webinars (wild labs)
- InsectAI COST Action (large EU network grant)

WILDLABS.NET

[The conservation technology network]

InsectAI

https://groups.google.com/g/computational-entomology-events

Thanks!

Quentin Geissmann
Assistant Professor, Aarhus University
2024-11-07
Ontario Pest Management Conference

qgeissmann@qgg.au.dk
@qgeissmann