Karlsruhe Institute of Technology

How automation, machine learning, and DNA barcoding can accelerate species discovery in “dark taxa”: Robotics and AI

International Congress of Entomology

Lorenz Wührl, Christian Pylatiuk, Luca Rettenberger, Rudolf Meier

July 19, 2022

Motivation and Goals

- Automation to speed up bio diversity exploring
- Insects are preserved in ethanol: three main problems

sorting

- handling
- size differences
- shape differences

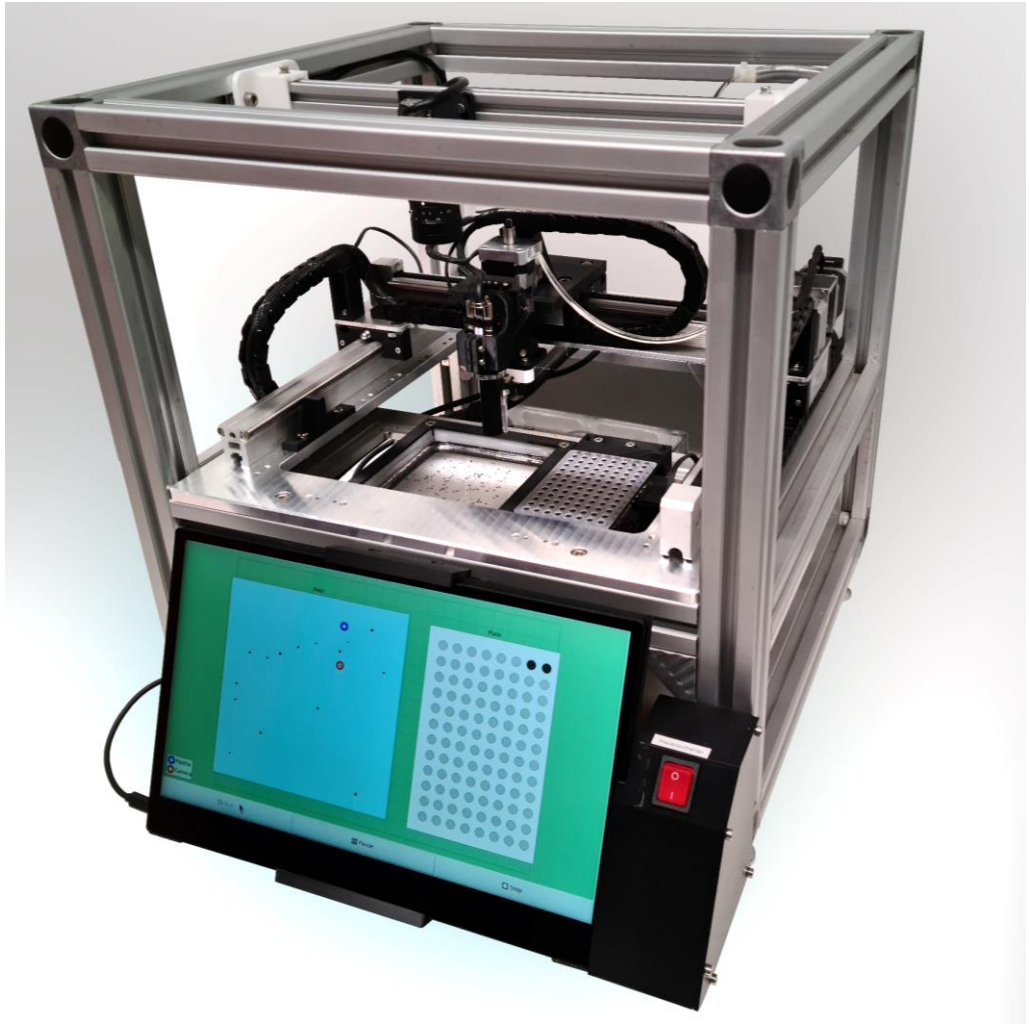
classification

- training data
- small characteristics
- position

biomass

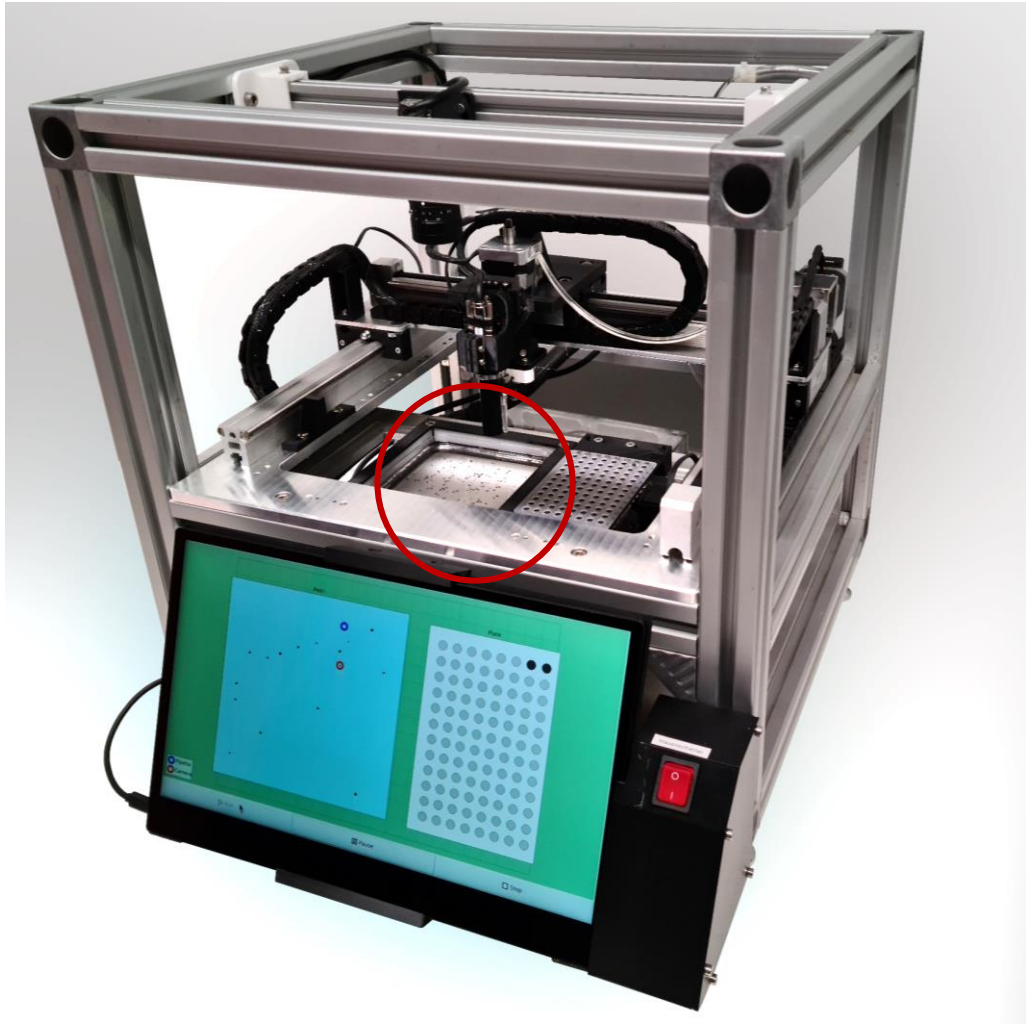
- image-based
- exact
- segmentation

DiversityScanner
Setup



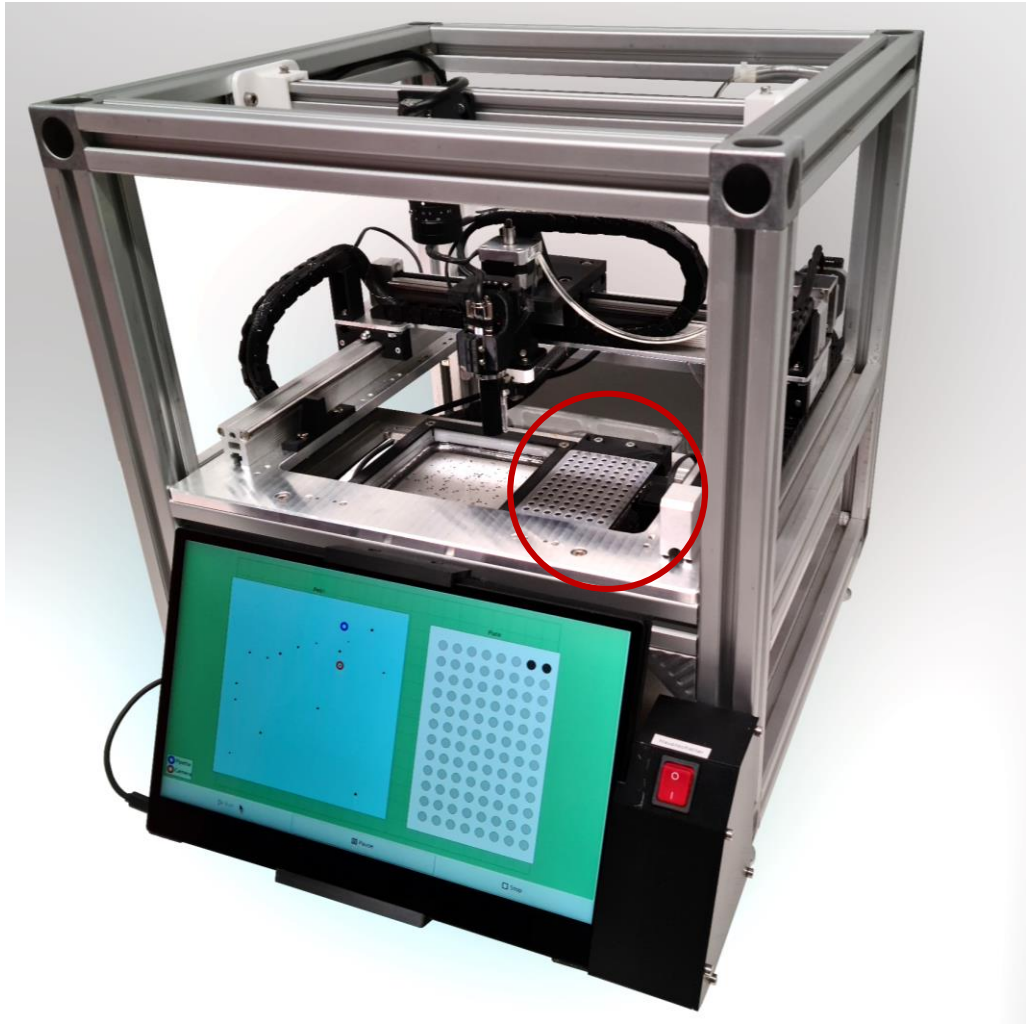
- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup



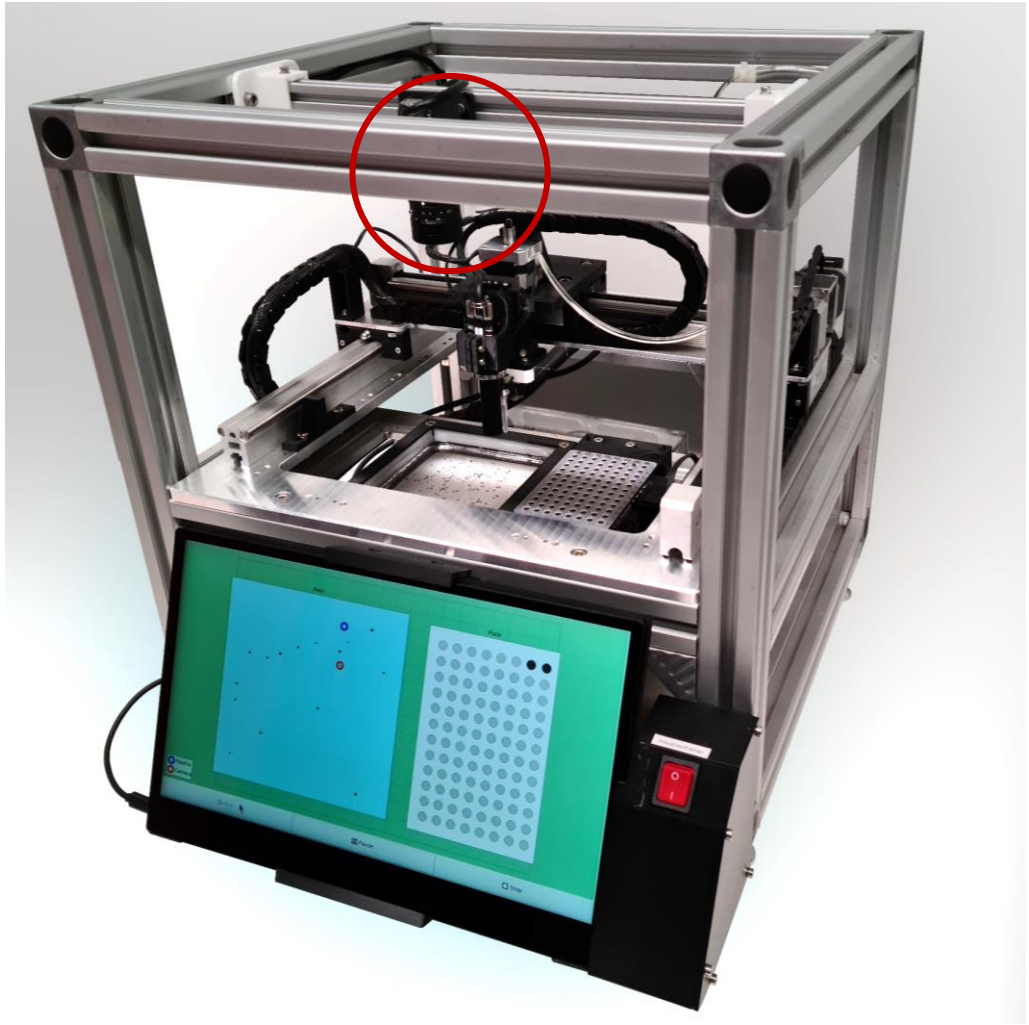
- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup



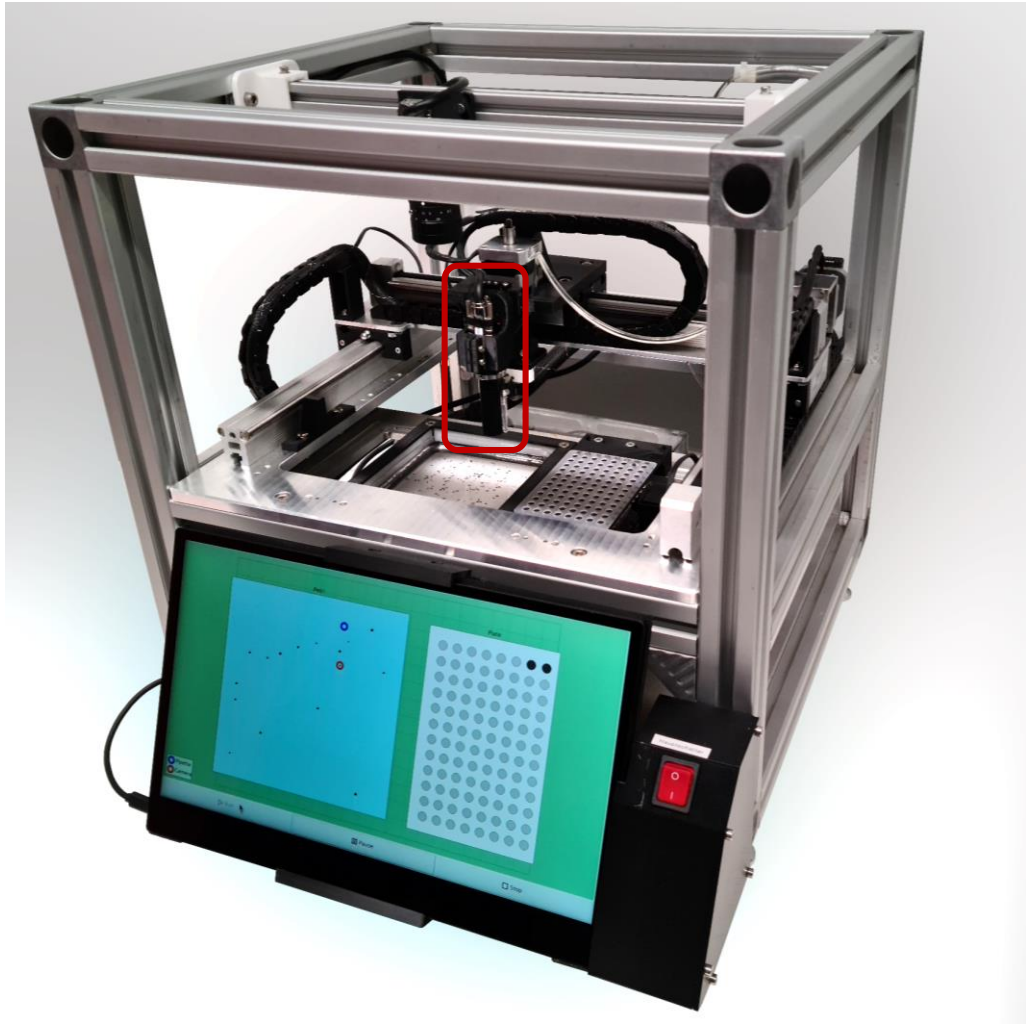
- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup



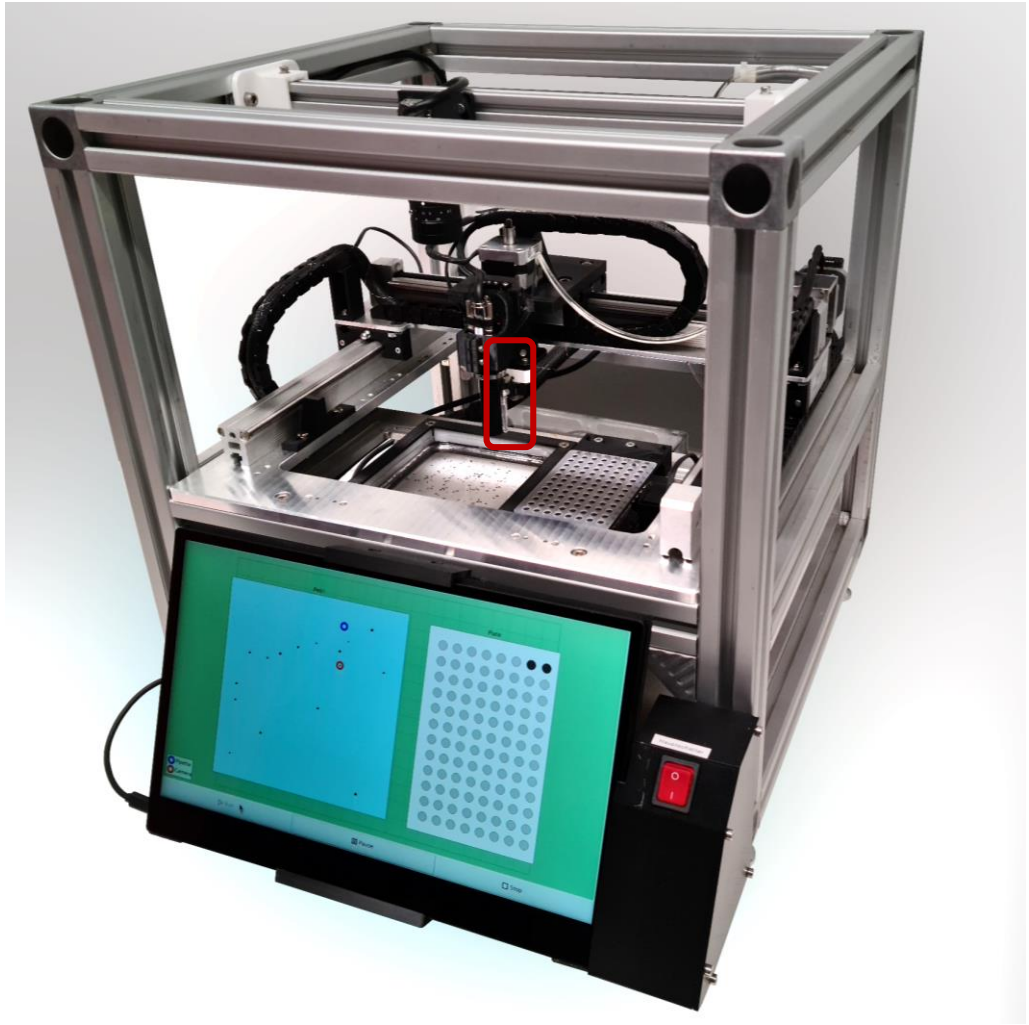
- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup



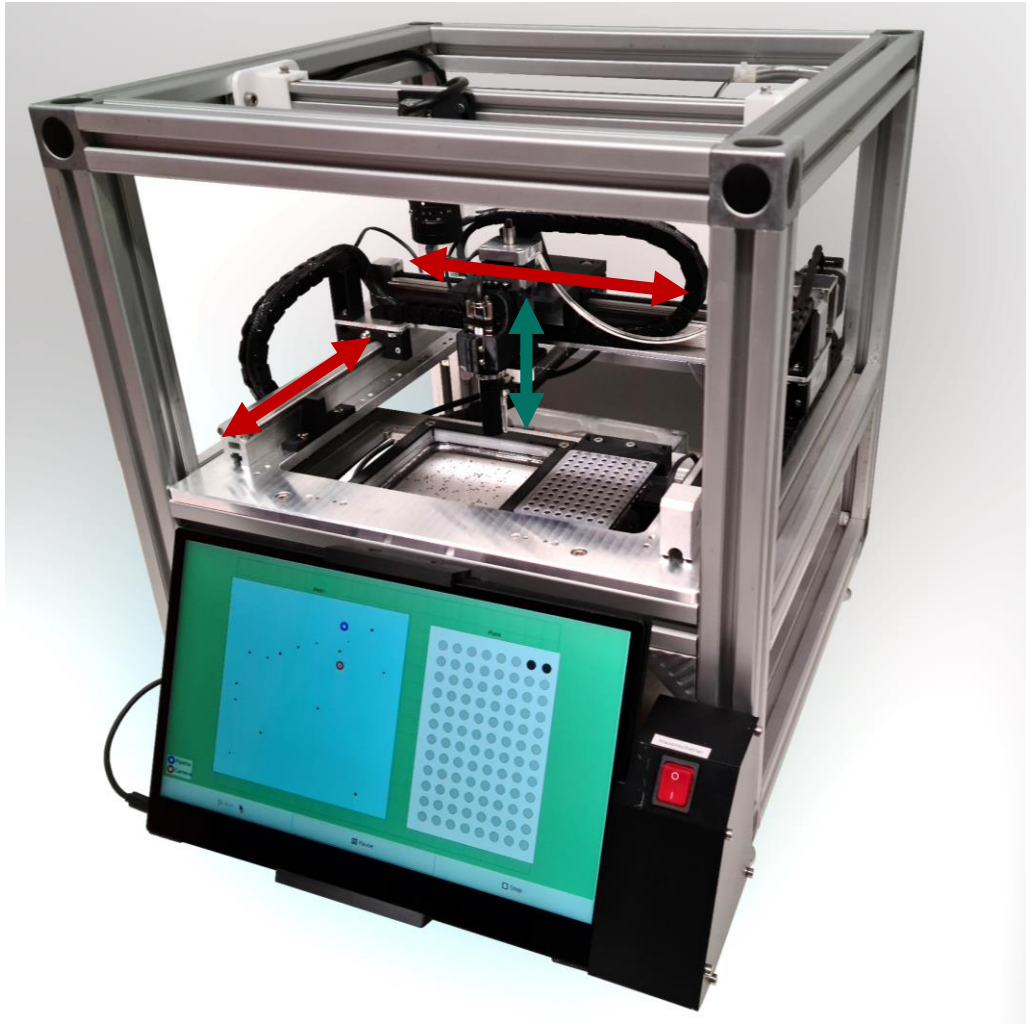
- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup



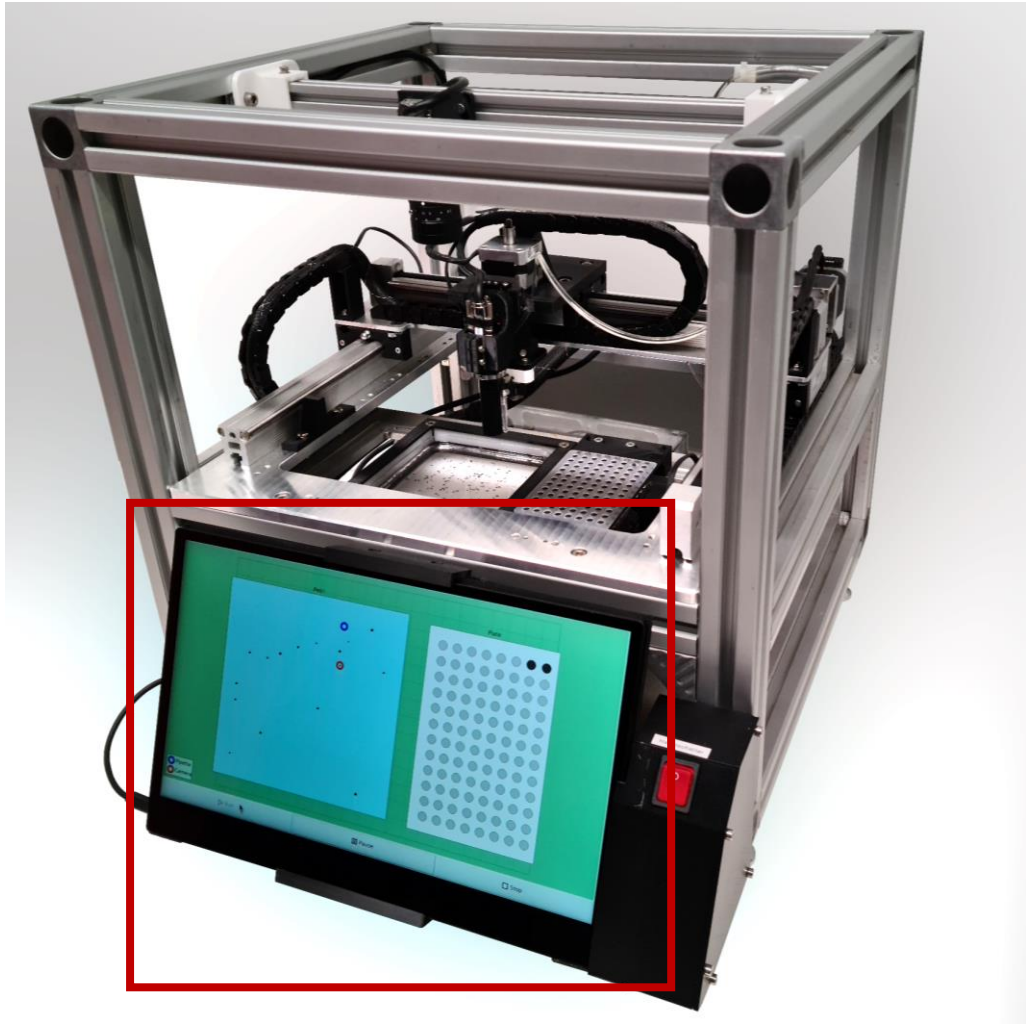
- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup



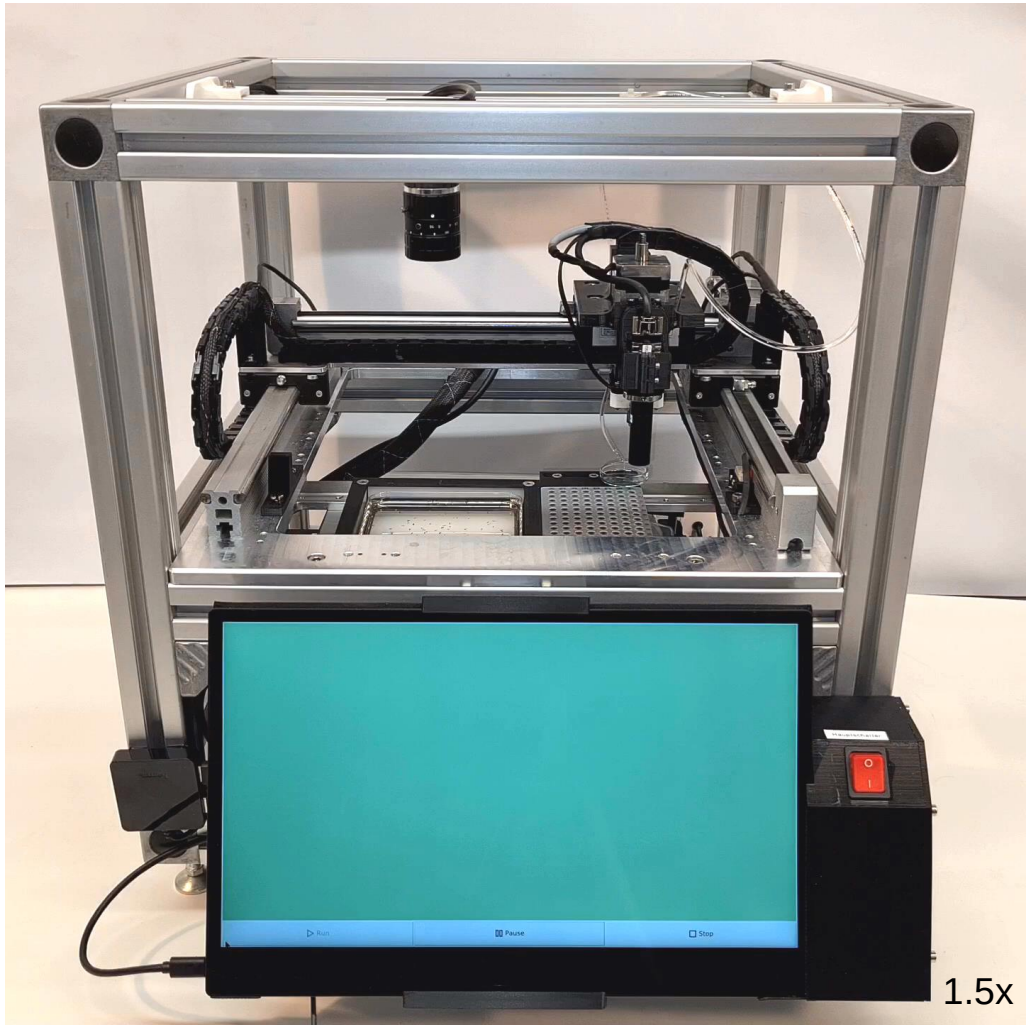
- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup



- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Setup

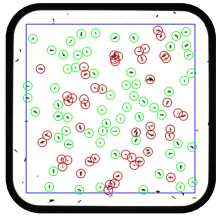


1.5x

- Petri dish for specimens
- 96-well microplate for further processing (e.g. DNA-barcoding)
- Overview camera
- Detail camera
- Automated pipette
- Three linear axes
- Graphical User Interface (GUI)
  - Specimen position
  - Camera and pipette position
  - Filled wells

DiversityScanner
Workflow of the DiversityScanner

Overview Image



Insect positon detection

Detail Image

Detailed specimen image for further processing

Specimen sorting

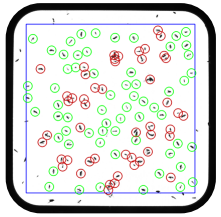
Sorting specimen from Petri dish in a microplate

Specimen classification

Biomass estimation

DiversityScanner
Workflow of the DiversityScanner

Overview Image



Insect positon detection

Detail Image

Detailed specimen image for further processing

Specimen sorting

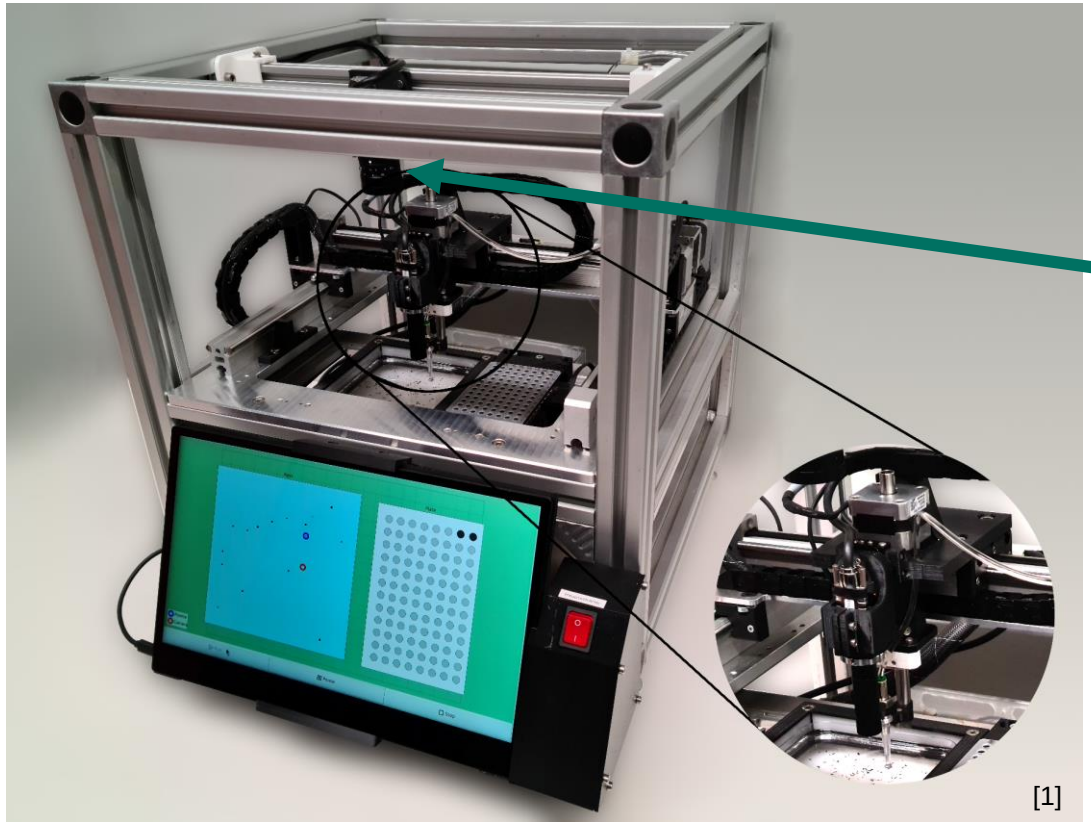
Sorting specimen from Petri dish in a microplate

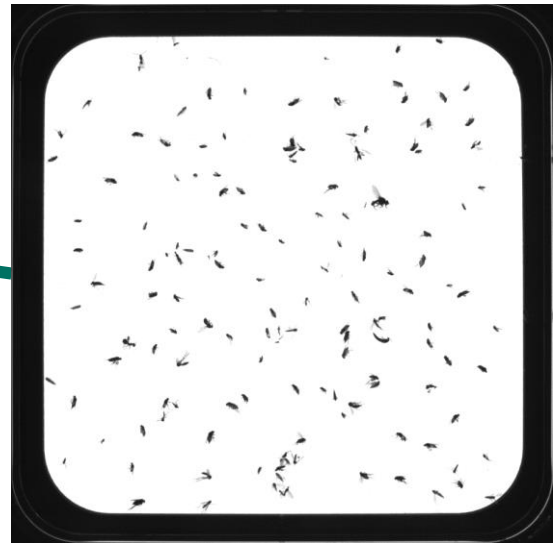
Specimen classification

Biomass estimation

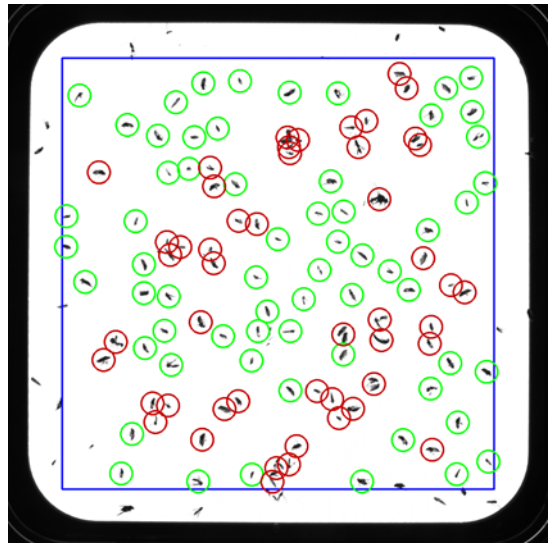
DiversityScanner
Specimen Position Determination

- Overview image to find specimen position in 120 x 120 mm Petri dish


[1]



Overview image of Petri dish

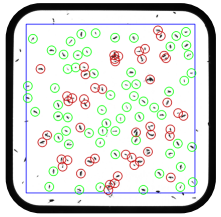


Marked image
green: sortable insects
red: not sortable insects
blue: ROI

[1] Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods" 2022.

DiversityScanner
Workflow of the DiversityScanner

Overview Image



Insect positon detection

Detail Image

Detailed specimen image for further processing

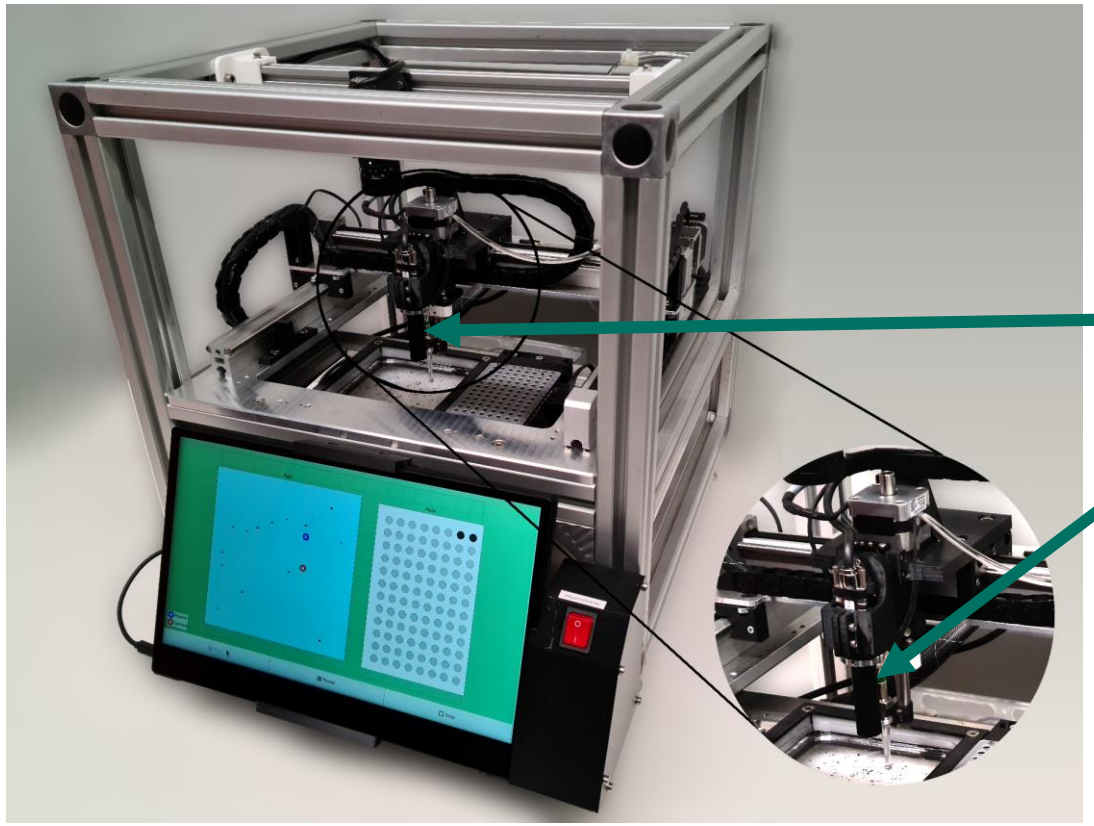
Specimen sorting

Sorting specimen from Petri dish in a microplate

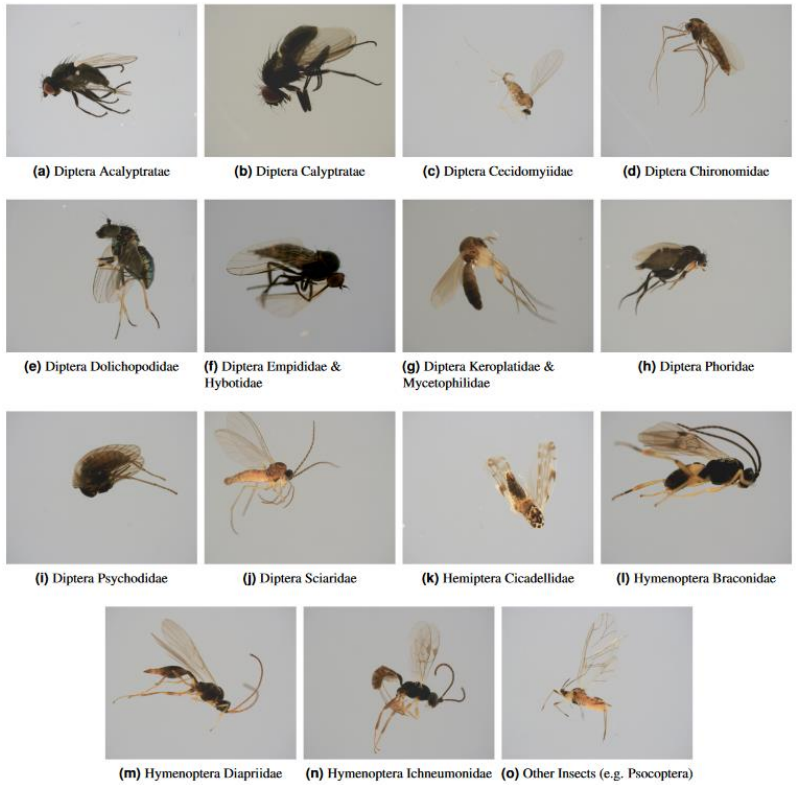
Specimen classification

Biomass estimation

DiversityScanner
Imaging



[1]



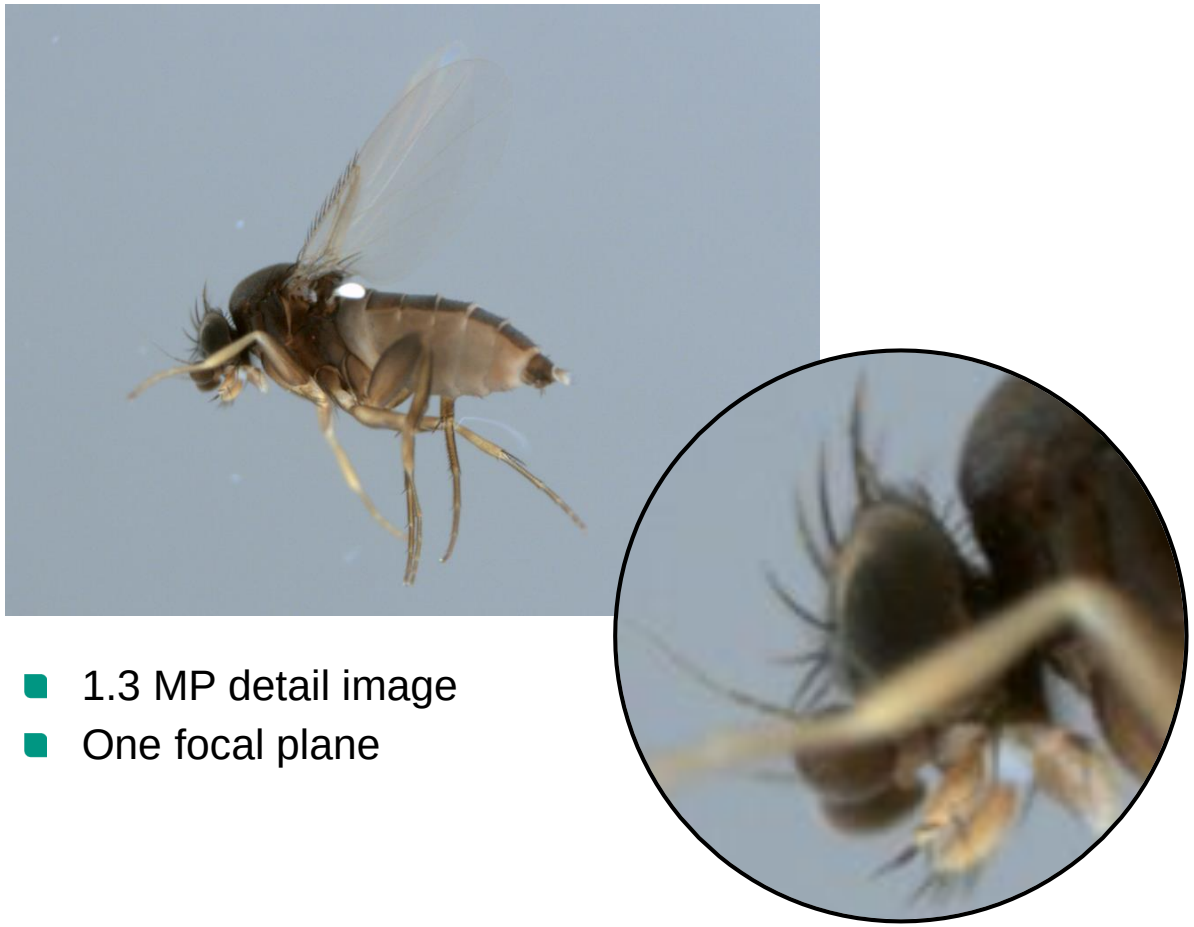
Detail image for classification and volume estimation

[1]

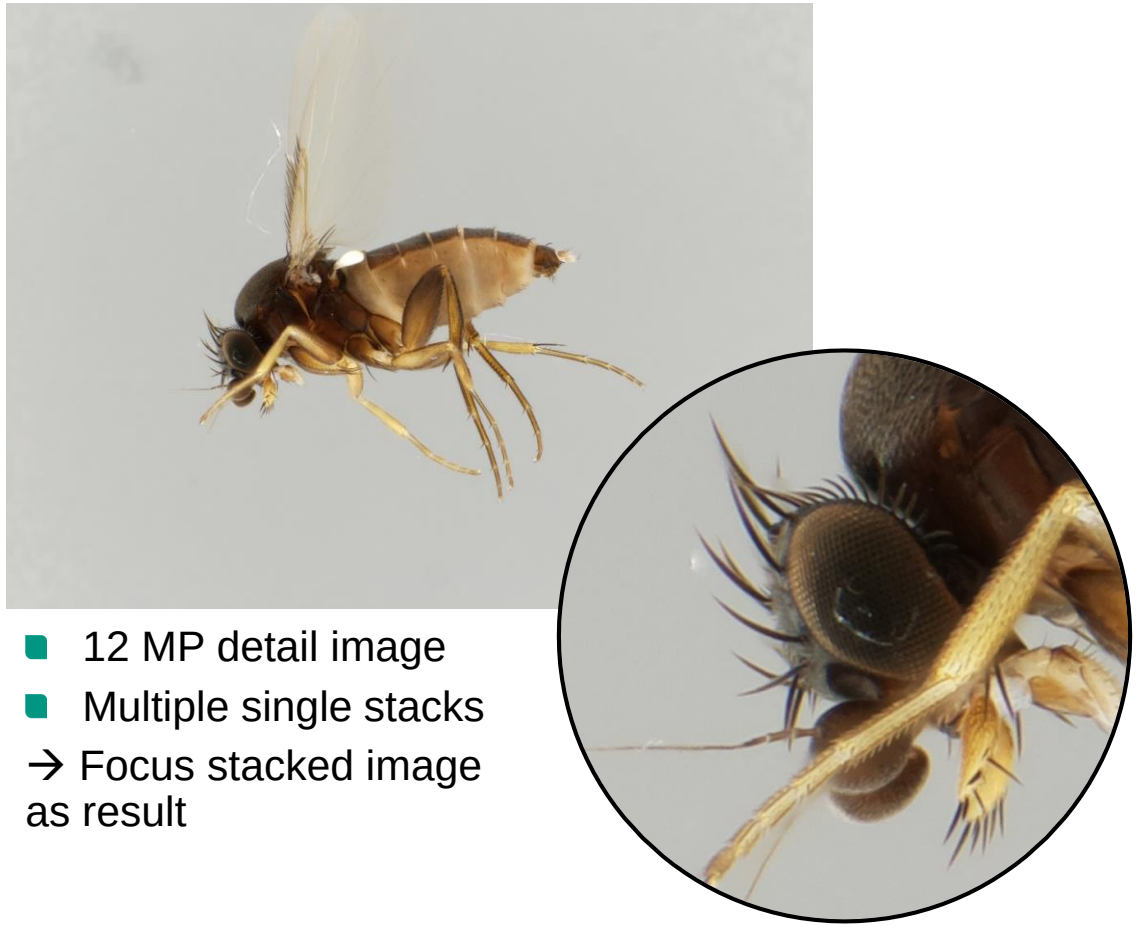
[1] Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods," 2022.

DiversityScanner

Imaging: 4K Camera Upgrade for the DiversityScanner 4K



- 1.3 MP detail image
- One focal plane



- 12 MP detail image
- Multiple single stacks

→ Focus stacked image as result

DiversityScanner

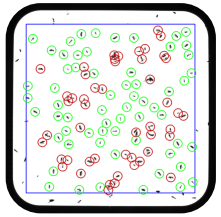
Imaging: 4K Camera Upgrade for the DiversityScanner 4K

Single focus image

Stacked image

DiversityScanner
Workflow of the DiversityScanner

Overview Image



Insect positon detection

Detail Image

Detailed specimen image for further processing

Specimen sorting

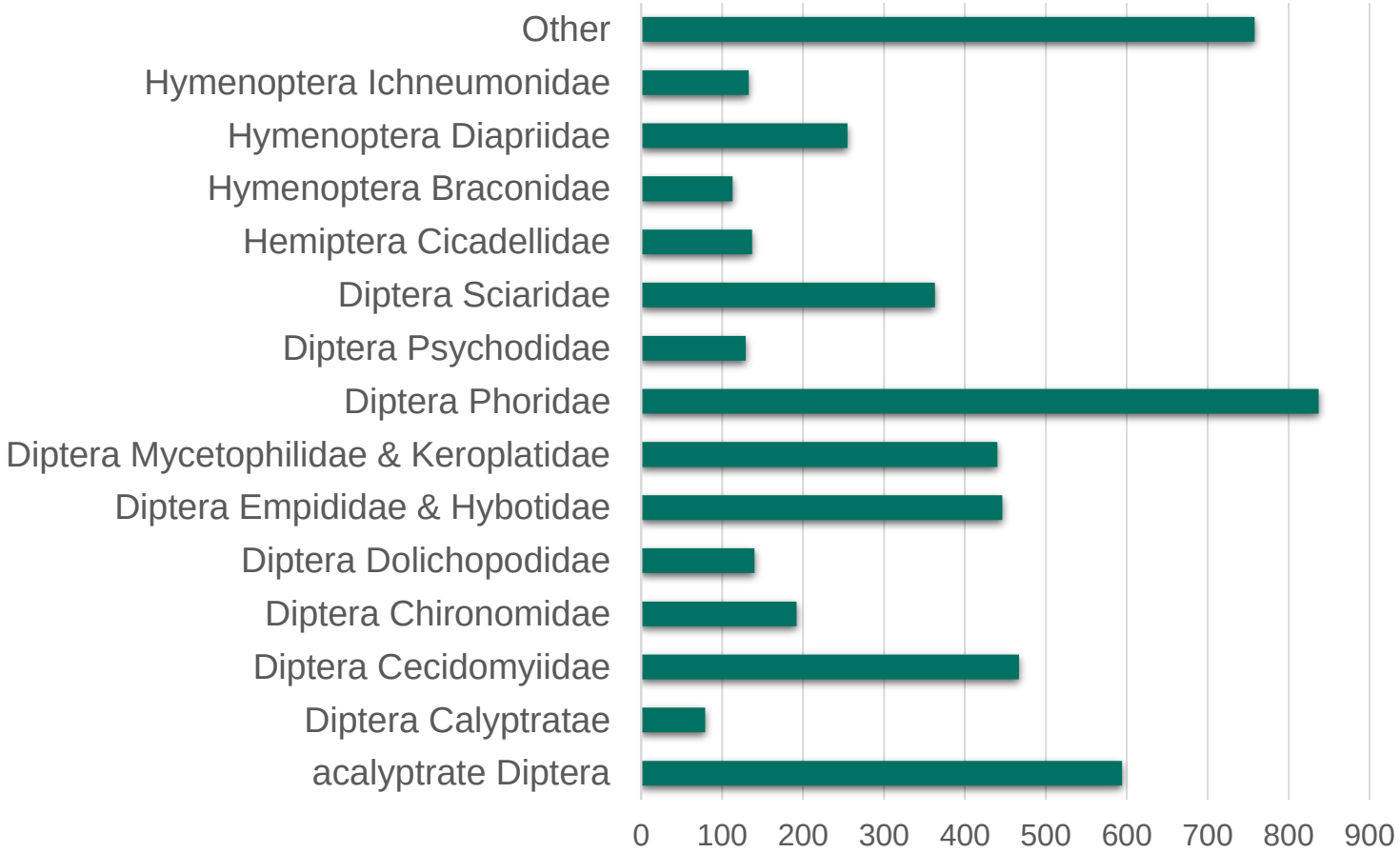
Sorting specimen from Petri dish in a microplate

Specimen classification

Biomass estimation

DiversityScanner
Specimen Classification

Total number of Images per class



- Self generated Dataset with **5083 images** for **15 classes**
- Images from five Malaise traps
  - Classes were selected from which sufficient images were available
- Image data from databases proved to be problematic
  - Often same view
  - Scales, writing or other artefacts in image

Data Availability: https://doi.naturkundemuseum.berlin/data/10.7479/4tbx-qm72

Data from: Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods," 2022.

DiversityScanner
Specimen Classification

- VGG19 – Architecture
- Worst result for:
  - Hymenoptera Ichneumonidae (75%)
- Best result for
  - Hymenoptera Diapriidae (100%)
  - Hemiptera Cicadellidae (100%)

→ **Total accuracy: 91.4%**

Class	Accuracy
Other	**0.810**
Hymenoptera Ichneumonidae	**0.750**
Hymenoptera Diapriidae	**1.000**
Hymenoptera Braconidae	**0.818**
Hemiptera Cicadellidae	**1.000**
Dipt. Sciaridae	**0.922**
Dipt. Psychodidae	**0.895**
Dipt. Phoridae	**0.967**
Dipt. Mycetophilidae & Keroplatidae	**0.991**
Dipt. Empididae & Hybotidae	**0.866**
Dipt. Dolichopodidae	**0.857**
Dipt. Chironomidae	**0.939**
Dipt. Cecidomyiidae	**0.915**
Dipt. Calyptratae	**0.833**
acalyptrate Dipt.	**0.912**

Data from: Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods," 2022.

DiversityScanner
Specimen Classification

- VGG19 – Architecture
- Worst result for:
  - Hymenoptera Ichneumonidae (75%)

Data from: Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods," 2022.

Class	Accuracy
Other	**0.810**
Hymenoptera Ichneumonidae	**0.750**
Hymenoptera Diapriidae	**1.000**
Hymenoptera Braconidae	**0.818**
Hemiptera Cicadellidae	**1.000**
Dipt. Sciaridae	**0.922**
Dipt. Psychodidae	**0.895**
Dipt. Phoridae	**0.967**
Dipt. Mycetophilidae & Keroplatidae	**0.991**
Dipt. Empididae & Hybotidae	**0.866**
Dipt. Dolichopodidae	**0.857**
Dipt. Chironomidae	**0.939**
Dipt. Cecidomyiidae	**0.915**
Dipt. Calyptratae	**0.833**
acalyptrate Dipt.	**0.912**

DiversityScanner
Specimen Classification

- Best results for
  - Hymenoptera Diapriidae (100%)
  - Hemiptera Cicadellidae (100%)

→ **Total accuracy: 91.4%**

Class	Accuracy
Other	**0.810**
Hymenoptera Ichneumonidae	**0.750**
Hymenoptera Diapriidae	**1.000**
Hymenoptera Braconidae	**0.818**
Hemiptera Cicadellidae	**1.000**
Dipt. Sciaridae	**0.922**
Dipt. Psychodidae	**0.895**
Dipt. Phoridae	**0.967**
Dipt. Mycetophilidae & Keroplatidae	**0.991**
Dipt. Empididae & Hybotidae	**0.866**
Dipt. Dolichopodidae	**0.857**
Dipt. Chironomidae	**0.939**
Dipt. Cecidomyiidae	**0.915**
Dipt. Calyptratae	**0.833**
acalyptrate Dipt.	**0.912**

Data from: Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods," 2022.

DiversityScanner
Specimen Classification

- Best results for
  - Hymenoptera Diapriidae (100%)
  - Hemiptera Cicadellidae (100%)

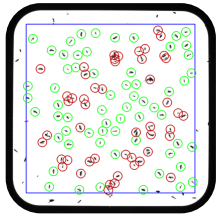
→ **Total accuracy: 91.4%**

Data from: Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods," 2022.

Class	Accuracy
Other	**0.810**
Hymenoptera Ichneumonidae	**0.750**
Hymenoptera Diapriidae	**1.000**
Hymenoptera Braconidae	**0.818**
Hemiptera Cicadellidae	**1.000**
Dipt. Sciaridae	**0.922**
Dipt. Psychodidae	**0.895**
Dipt. Phoridae	**0.967**
Dipt. Mycetophilidae & Keroplatidae	**0.991**
Dipt. Empididae & Hybotidae	**0.866**
Dipt. Dolichopodidae	**0.857**
Dipt. Chironomidae	**0.939**
Dipt. Cecidomyiidae	**0.915**
Dipt. Calyptratae	**0.833**
acalyptrate Dipt.	**0.912**

DiversityScanner
Workflow of the DiversityScanner

Overview Image



Insect positon detection

Detail Image

Detailed specimen image for further processing

Specimen sorting

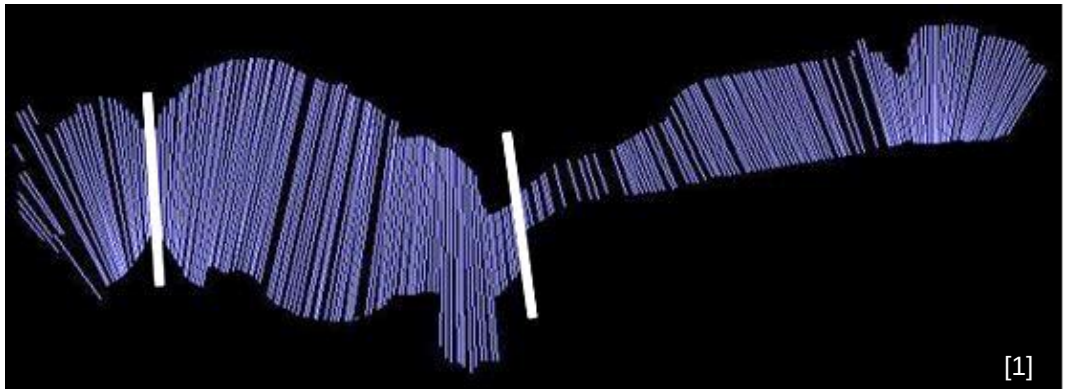
Sorting specimen from Petri dish in a microplate

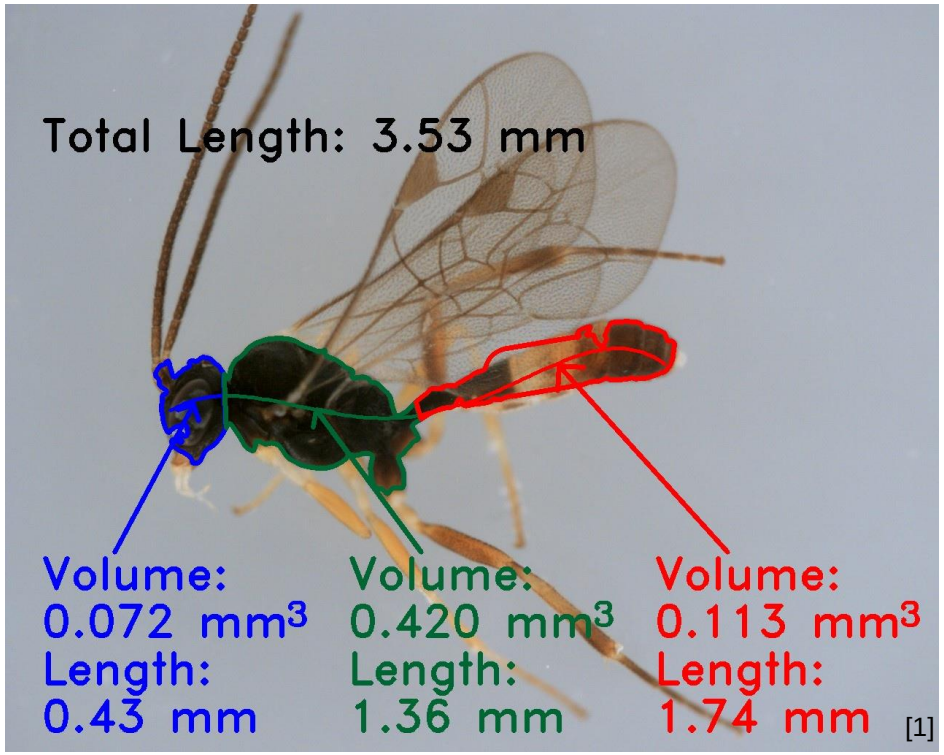
Specimen classification

Biomass estimation

DiversityScanner
Biomass Estimation

- Detail images to estimate biomass and length of individual specimens
- Classical image processing
  - Good results without labeling and training data
- Volume is estimated using circular disks

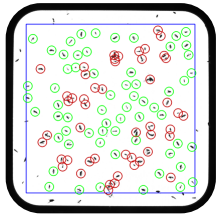




[1] Wührl et al., "Diversityscanner: Robotic handling of small invertebrates with machine learning methods," 2022.

DiversityScanner
Workflow of the DiversityScanner

Overview Image



Insect positon detection

Detail Image

Detailed specimen image for further processing

Specimen sorting

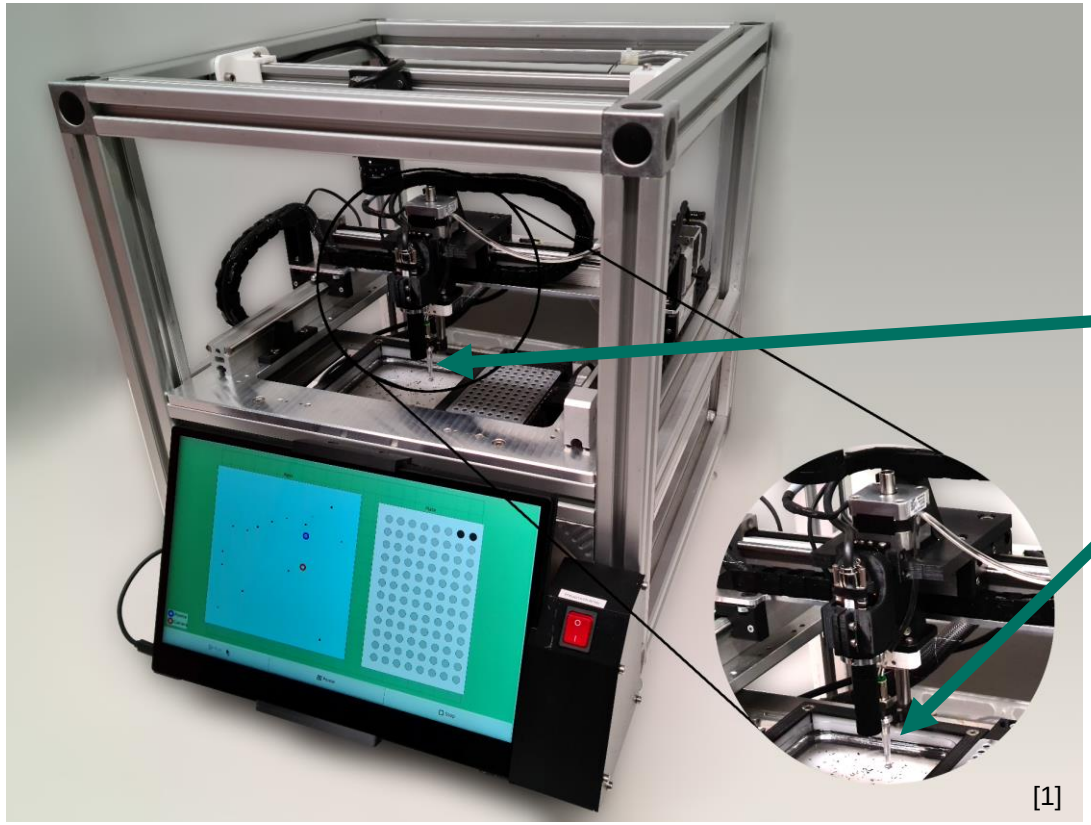
Sorting specimen from Petri dish in a microplate

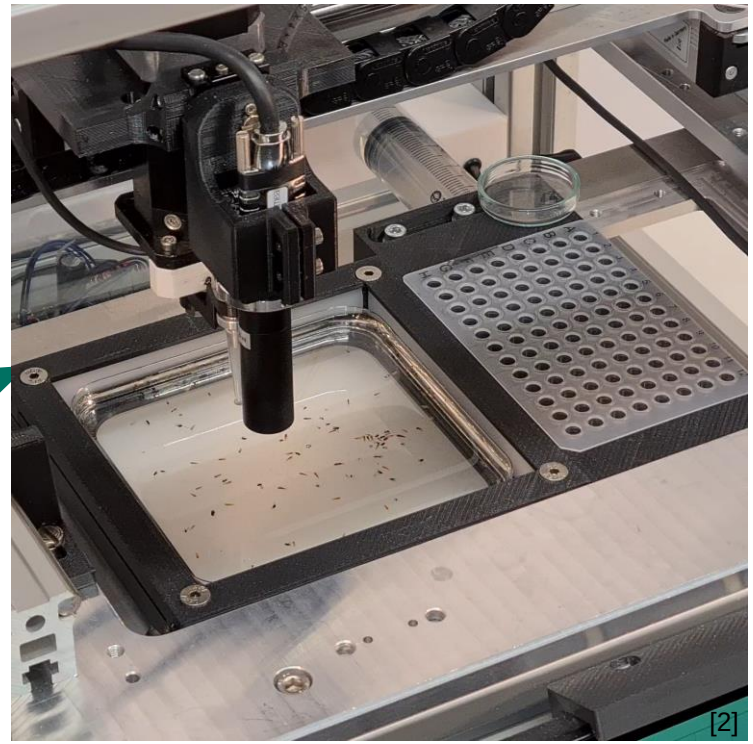
Specimen classification

Biomass estimation

DiversityScanner
Sorting

- Sorting specimens into 96-well microplate (invertebrates < 3mm)


[1]


[2]

Sorting insect into a microplate

[1] Wührl et al., “Diversityscanner: Robotic handling of small invertebrates with machine learning methods,” 2022.
[2] https://www.youtube.com/watch?v=ElJ5VSHa4OI

DiversityScanner
Summary

- Robotics and AI can be utilized to accelerate species discovery in „dark taxa“
- Robotics can be used to perform repetitive and tedious tasks
  - Upscaling and parallelization enable high throughput
- AI enables classification of the specimens for e.g. pre-sorting
- Generated image data for:
  - New training sets
  - Classification to higher taxonomic level

DiversityScanner
Video

https://www.youtube.com/watch?v=ElJ5VSHa4OI