

Title: DrawerDissect: Whole-drawer insect imaging, segmentation, and transcription using AI

Authors: \*Elizabeth G. Postema<sup>1</sup>, Leah Briscoe<sup>1,2</sup>, Chloe Harder<sup>1,3</sup>, George R. A. Hancock<sup>4</sup>, Lucy D. Guarnieri<sup>5</sup>, Tony Eisel<sup>1,6</sup>, Kelton Welch<sup>7</sup>, Nicole Fischer<sup>8</sup>, Christine Johnson<sup>9</sup>, Diego Souza<sup>1</sup>, Dexter Phillip<sup>1</sup>, Rebekah Baquiran<sup>1</sup>, Tatiana Sepulveda<sup>1</sup>, \*Bruno A. S. de Medeiros<sup>1</sup>

1. Field Museum, 1400 S DuSable Lake Shore Drive, Chicago, IL 60605, USA
2. Lake Forest College, 555 N Sheridan Rd, Lake Forest, IL 60045, USA
3. Lane Tech College Prep High School, 2501 W Addison St, Chicago, IL 60618, USA
4. Centre for Ecology and Conservation, University of Exeter, Penryn, TR10 9FE United Kingdom
5. The Ohio State University, Department of Entomology, 2021 Coffey Road, Columbus, Ohio, 43201
6. Roosevelt University, 430 S Michigan Ave, Chicago, IL 60605
7. Ecdysis Foundation, Estelline, SD 57234, USA
8. National Research Collections Australia, Clunis Ross St, Canberra, ACT 2601, Australia
9. American Museum of Natural History, 200 Central Park West at 79th Street, NY 10024, USA

\*Corresponding authors

#### AUTHOR CONTRIBUTIONS

EGP led the pipeline development, imaged drawers, and analyzed data. BASM conceived the pipeline, assisted in the pipeline development, and funded the research. EGP and BASM led the writing of the manuscript. LB, CH, LDG, and TE assisted in the pipeline development, imaged drawers, and analyzed data. LDG, KW, NF, CJ and BASM provided images for training AI models. GRAH developed the workflow for batch color analysis of masked specimens. DS and DP transcribed specimen metadata. RB and TS developed the workflow for batch-uploading data and images to KE EMu. All authors contributed critically to the drafts and gave final approval for publication.

#### DATA AVAILABILITY STATEMENT:

The source code for DrawerDissect and the identification model *Cicindel-ID* are available at [github.com/EGPostema/DrawerDissect](https://github.com/EGPostema/DrawerDissect) and [github.com/de-Medeiros-insect-lab/Cicindelinae\\_ID](https://github.com/de-Medeiros-insect-lab/Cicindelinae_ID), respectively. All images and annotations used to train FMNH roboflow models can be found at [universe.roboflow.com/field-museum](https://universe.roboflow.com/field-museum). Training weights for *Cicindel-ID* are available at [huggingface.co/brunoasm/eva02\\_large\\_patch14\\_448.Cicindela\\_ID\\_FMNH](https://huggingface.co/brunoasm/eva02_large_patch14_448.Cicindela_ID_FMNH).

CONFLICT OF INTEREST: The authors report no conflicts of interest.

## ABSTRACT

1. Many museums curate vast collections of insect specimens. These collections represent invaluable records of biodiversity information, ecological patterns and phenotypic variation. However, traditional methods of imaging specimens and digitizing their metadata are labor-intensive and risk damaging delicate specimens. High-throughput imaging of entire specimen drawers integrated with computer vision artificial intelligence (AI) models can provide a potential solution.

2. We present DrawerDissect, a python-based pipeline for processing high-resolution drawer photographs, and a workflow to use it in entomological collections. By using custom vision models trained in the platform Roboflow and LLM-based transcription with Claude, DrawerDissect can crop and segment specimens from images and extract metadata from specimen labels. DrawerDissect is flexible, tuneable and modular, allowing seamless integration with downstream analyses of phenotypic features (e.g. color, pattern, and size).

3. We validated Drawerdissect by digitizing the Field Museum of Natural History's (FMNH's) entire tiger beetle (family Cicindelidae) collection, processing 13,484 specimens to generate high-resolution dorsal photographs, backgroundless specimen images, and basic body measurements. Geographic data were successfully extracted for 3,648 specimens. To demonstrate the utility of the masked images, we provide an example integration of DrawerDissect with existing image analysis methods in R and ImageJ. Finally, to show the research potential of high-quality specimen images, we trained a species identification model, *Cicindel-ID*, using ~7,000 images of specimens in the genus *Cicindela*.

4. DrawerDissect's novel multi-model AI workflow provides an efficient and reproducible framework that meets the demands of high-throughput digitization of natural history museum collections, unlocking the potential of vast specimen collections for future analyses.

**KEYWORDS:** artificial intelligence, computer vision, digitization, high-throughput imaging, insects, image segmentation, machine learning, museum specimens

## 1 INTRODUCTION

Natural history collections connect the past to the future. Each specimen contains a unique ecological and evolutionary history, and preserving these specimens carries their histories forward into the hands of future researchers - sometimes hundreds of years after their collection. Collections have been leveraged to track the impacts of climate change among plants and animals (Bates et al., 2023; Lister, 2011; Nj et al., 2023); to understand the ecological drivers of diverse morphological traits (Crowell et al., 2024; Holmes et al., 2016); and to estimate the sizes, ranges and compositions of past populations (Davis et al., 2023). They also provide important genomic data that can be leveraged for species identification and evolutionary studies (de Medeiros et al., 2025; Ruane & Austin, 2017). However, unlocking the abundance of natural history collections comes with significant challenges. Metadata digitization and specimen imaging are necessary for collections-based research, but they are infamously time- and labor-intensive, as they traditionally involve handling specimens one at a time. The challenge is magnified for hyper-diverse groups such as insects. In North America, for example, the pace of arthropod specimen digitization cannot keep up with new specimen accessions, let alone existing collections: only ~5% of arthropod specimens have data digitally available, and fewer than 2% have been imaged (Cobb et al., 2019). Given the ecological significance of insects (Losey & Vaughan, 2006; Scudder, 2017), and their perilous position in the face of rapid global change (Wagner, 2020; Yang et al., 2021), there is an increasingly urgent need to mobilize insect specimen data and images.

Most digitization efforts in the past decades have focused on the transcription of specimen label metadata to standardized formats such as DarwinCore (Wieczorek et al., 2012). This has enabled the development of numerous portals and aggregators of biodiversity data, such as the Global Biodiversity Information Facility (GBIF, 2025) and unleashed vast possibilities for research relying on analyzing such metadata. Imaging specimens has often been a secondary goal. For example, at the Field Museum of Natural History (FMNH), we have historically digitized 576,395 arthropod specimens, but only 19,998 (~3.5%) have images available. This means that only a tiny fraction (~0.15%) of an arthropod collection estimated to hold at least 12 million specimens has been photographed. The traditional focus on metadata is justifiable, as images have historically been time-consuming to generate and analyze, and the added value in relation to metadata alone was unclear. While machines can easily summarize millions of text or numeric records, extracting information from specimen images typically requires pre-processing steps like outlining or landmarking - traditionally done by hand (van den Berg et al., 2020; Watanabe, 2018; Zelditch et al., 2012). However, recent advances in artificial intelligence (AI), particularly computer vision (CV) models for object detection, instance segmentation, and keypoint detection, make automating these pre-processing steps possible (Borowiec et al., 2022; Lürig, 2022; Pichler & Hartig, 2023; Weinstein, 2018; Zhao et al., 2024). Other AI models can extract categorical data from specimen images, such as sex (Wilson et al., 2023) or taxonomic identity

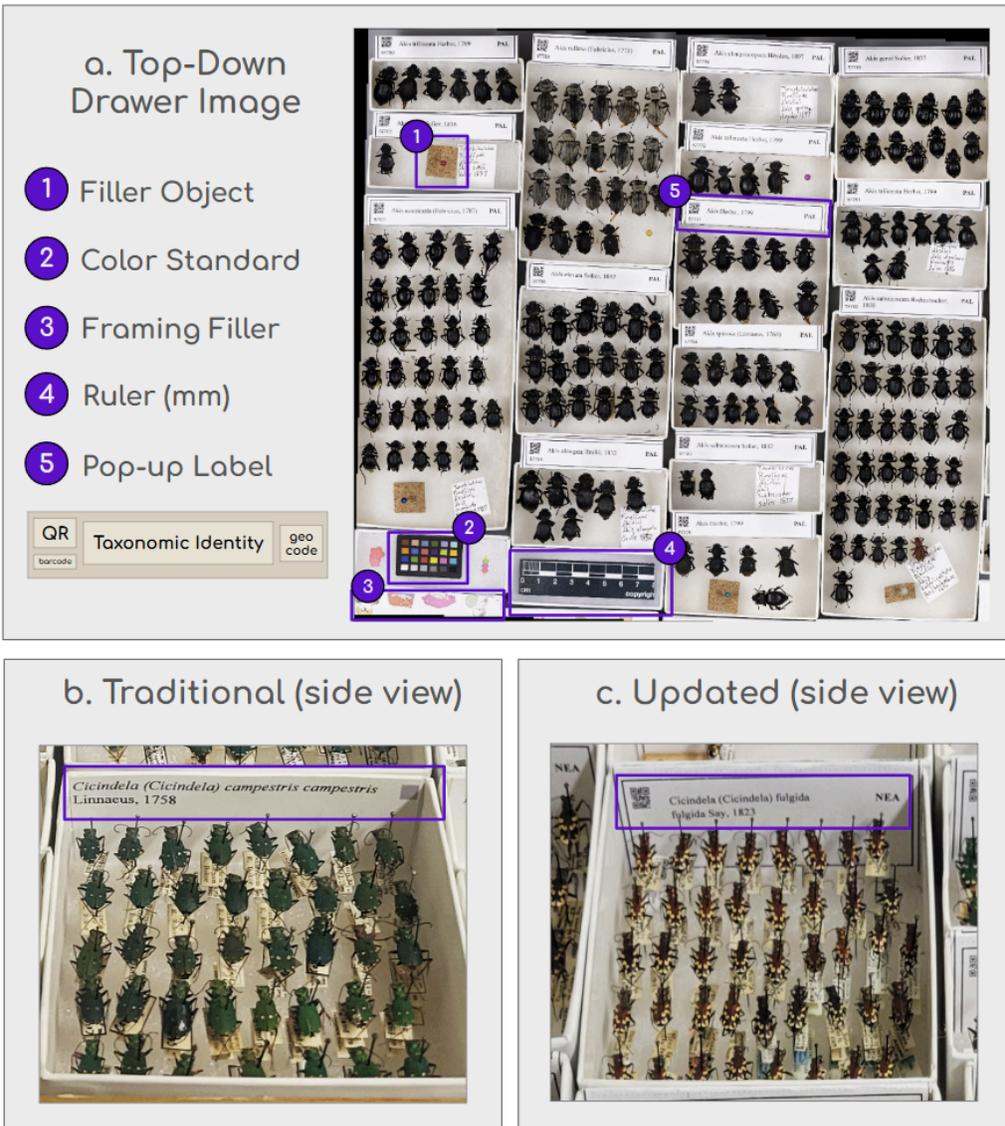
(Wühl et al., 2022). With these new tools, images are not just byproducts of digitization with limited uses, but high-quality sources of data themselves.

Image analysis has significantly improved, but image acquisition still represents a bottleneck that is a focus of active research. There are many existing solutions for high-throughput imaging in collections, depending on the type of organism and method of preservation. For example, many herbaria use conveyor-belt systems for quickly scanning paper-mounted plant specimens (Picturae, <https://picturae.com/>; Sweeney et al., 2018). Bespoke camera rigs have been employed for batch-photographing vertebrate skeletons (Weeks et al., 2023) as well as soil samples containing multiple small invertebrates (Sys et al., 2022). For dry, pinned insects that are stored in drawers, mass-digitization solutions are currently under development (LightningBug, <https://www.lightningbug.tech/prog-orig>; Picturae, <https://picturae.com/>; Steinke et al., 2024), but none are currently in widespread use. No matter how efficient individual specimen handling can be made, a much faster solution is to use the existing organization of specimens in drawers and simply photograph whole drawers at a time (Mantle et al., 2012). Several different approaches to “whole-drawer imaging” have been pursued in the last 10-15 years (reviewed in Holovachov et al., 2014). While some of these initial approaches suffered from image warping and stitching artifacts (Mantle et al., 2012), modern lenses and imaging systems are capable of producing incredibly clear, high-quality images of hundreds of specimens at a time (Fig. 1a).

While high-throughput imaging can be accomplished in a fraction of the time taken to image individual specimens, this creates the challenge of processing bulk images to extract individual-level data. “Drawer” is not the unit of interest for researchers - “specimen” is. A possible solution to this processing bottleneck has been clear for over 10 years. Holovachov et al. (2014) presciently ask, “[c]an specimens in the image be analyzed and identified using a computer algorithm and machine learning?” Looking at the explosion of computer-vision based pipelines designed for specimen digitization in recent years, we can largely respond “yes” to this question in 2025. For example, LeafMachine2 uses a series of computer vision models to extract phenotypic data and transcribe labels from herbarium specimens (Weaver & Smith, 2023); Skelevision segments and measures bones from batch-imaged bird skeletons (Weeks et al., 2023); CollembolAI detects and classifies small invertebrates in photographs of soil samples (Sys et al., 2022); and DiversityScanner combines robotic sorting and imaging of bulk insect samples with a classification model to predict each insect’s family-level identity (Wühl et al., 2022). No doubt this list is incomplete, but from these examples alone, we can draw the conclusion that AI is an effective tool for mass-digitizing collections and extracting specimen-level data (Stenhouse et al., 2025). Applying the same concept to pinned insect collections, we wanted to create an AI-driven, user-friendly pipeline for processing whole-drawer images, with three key outputs:

- (1) Individual specimen images linked to taxonomic identity.
- (2) Outlined specimen bodies without backgrounds (i. e. “masked” specimens) suitable for downstream phenotypic analysis.
- (3) A subset of precise specimen-level collection locations, extracted from any visible label text in a given specimen image.

We call this multimodel pipeline “DrawerDissect.” DrawerDissect relies on a data annotation and AI model training service, Roboflow (Dwyer et al., 2014), for object detection and segmentation. For image-to-text transcription, we feed images to Anthropic’s large language model (LLM), Claude (<https://console.anthropic.com/>), with customizable prompts. We use these subscription-based services rather than open source models specifically because they are more accessible to a broader user base that may lack the coding experience or computing infrastructure to run AI models locally (Heron et al., 2013). Roboflow’s intuitive graphical interface and image labeling tools make training and deploying computer vision models remarkably fast, with no back-end coding required. For transcription, LLMs like Claude are designed to understand conversational (“natural language”) instructions and can handle ambiguities and inconsistencies in text inputs (Naveed et al., 2024). We designed DrawerDissect to be flexible, with built-in toggles for different drawer setups and desired outputs; customizable, with the option to swap between different models, update existing models, and edit LLM prompts; and modular, with steps that can be run independently or combined into unique workflows.



**Figure 1.** (a) A top-down whole-drawer image taken by the GIGAMacro Magnify2, with key objects labeled 1-5: a filler object, color standard, framing filler, ruler, and top-down view of a pop-up label. (b) Standard FMNH unit tray header label. (c) A new pop-up unit tray label, which folds flat when stored.

## 2 METHOD

To demonstrate the utility of DrawerDissect using a real example, we photographed and processed the FMNH's entire collection (over 13,000 specimens) of pinned tiger beetles (Coleoptera: Cicindelidae; Duran & Gough, 2020). Tiger beetles are strikingly colorful insects, often with intricately patterned and iridescent elytra. The ecology and evolution of this group's coloration has been the focus of much research, particularly with respect to background-matching, thermal physiology, morphological convergence, and a unique "pointillistic" method of structural color production (French et al., 2021; Pearson & Vogler, 2001; Schultz & Bernard, 1989; Schultz & Hadley, 1987; Yamamoto & Sota, 2020). To test DrawerDissect's ability to integrate with existing batch color analysis pipelines (Hancock et al., 2025), we use the segmented images generated by DrawerDissect to investigate geographic and climatic patterns of color diversity in two subspecies of a well-represented species in our collection: *Cicindela formosa*. Finally, we use over 7,000 masked images generated by DrawerDissect to train a species ID model, *Cicindel-ID*, to help identify a set of unsorted specimens in the speciose genus *Cicindela*. This model highlights the value of natural history collections as rich data sources for AI-based biological identification tools.

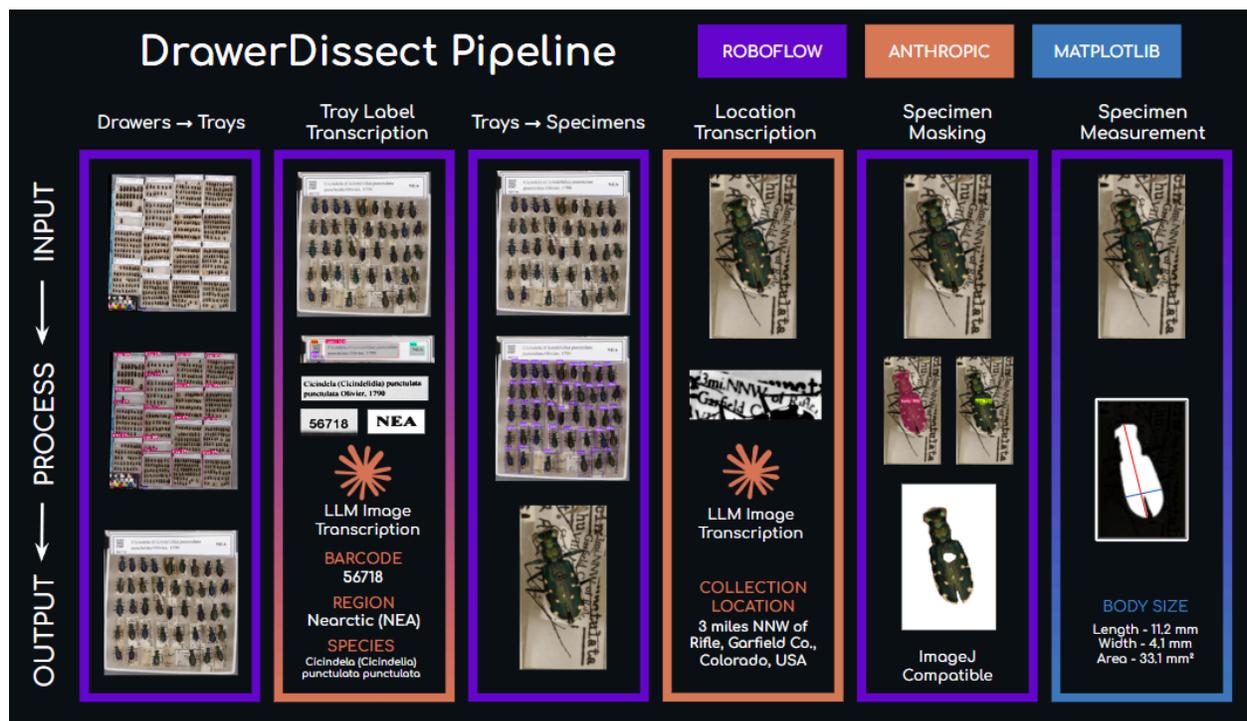
Our workflow involves three steps: (1) imaging drawers, (2) configuring DrawerDissect, (3) running images through DrawerDissect, and (4) optional post-curation. First, we describe our imaging set-up using a GIGAMacro Magnify2 system (Four Chambers Studio, California, USA, <https://gigamacro.com/>), and how the final whole-drawer images are generated. We then give a broad overview of DrawerDissect in terms of installation, configuration, and cost. We also describe how to train the CV models that we use throughout the pipeline. Finally, we describe the steps that occur during a standard DrawerDissect run. A streamlined user guide with quick-start instructions can be found at the DrawerDissect github: [github.com/EGPostema/DrawerDissect](https://github.com/EGPostema/DrawerDissect).

### 2.1 Whole-Drawer Imaging

DrawerDissect was designed with images produced by a GIGAMacro Magnify2 imaging system, and uses Field Museum (FMNH) conventions for insect drawers. These conventions are typical of United States natural history collections, where specimens are grouped into trays ("unit trays") based on shared taxonomic identity and geographical origin (Fig. 1b). Unit trays often have header labels with taxonomic and other information affixed to the tray interior (Fig. 1b). To make the labels visible for imaging, we developed "pop-up" versions that display the same information on both the interior and exterior of the tray. The pop-up portion is visible during imaging (Fig. 1a), but folds down flat when trays are stored (Fig. 1c). Unit trays have also not traditionally received unique identifiers in the FMNH insect collection. As part of this workflow, we developed a system to create unique identifiers for unit trays, which refer to collection locations in the Field Museum KE EMu database (Fig. 1a).

The only pre-curation step needed for drawer imaging is to generate an inventory of unit trays for each drawer, print the new pop-up labels, and replace the old labels with the new. In the FMNH collection, we assign unique identifiers and create 2D barcodes for each drawer using the KE EMu database system, but this is optional. Next, we arrange all unit trays for imaging on the GIGAMacro platform. We aim for the smallest possible footprint to reduce the imaging time. While DrawerDissect can handle natural variation in specimen positioning, we may adjust any specimens that are crooked or overlapping. While we describe our imaging methods based on this standard, DrawerDissect can be adapted to various configurations. At minimum, running the full pipeline requires a reasonably clear, high-quality image of an insect drawer, which can be achieved by a number of different whole-drawer imaging systems (Holovachov et al., 2014).

## 2.2 Overview of DrawerDissect



**Figure 2.** Visual overview of the main image and data outputs of DrawerDissect, with key python packages highlighted.

DrawerDissect is a Python-based pipeline available from GitHub, including documentation on installation and usage at <https://github.com/EGPostema/DrawerDissect>. The package includes test data to familiarize new users with the pipeline's steps, outputs, and organizational structure before they process new images. We specifically aimed to make the pipeline accessible to users who are not expert programmers. Only a rudimentary familiarity with their computer's

command-line interface is required, to run the script and edit a single master configuration file. Overall, DrawerDissect can be installed and run in four simple steps:

1. Create an environment for the program to run (~5 lines of command-line code)
2. Add whole-drawer images to a folder
3. Input API keys from Roboflow and Anthropic
4. Run the full pipeline with a single command

### 2.2.1 Third-party Packages and Dependencies

Our pipeline is designed to run in a Python-based virtual environment (Python 3+) with a suite of required packages that are installed automatically upon installation. In general, we use pandas for generating, merging, and handling dataframes (pandas Development Team, 2024). Our main image processing packages are NumPy (Harris et al., 2020), Pillow (Clark, 2015), and OpenCV (Bradski, 2000). NumPy is used to convert images into arrays, which can then be used to translate specific areas of an image into full-color pixels or censored white pixels. We also use NumPy for some basic mathematical functions, such as calculating angles, distances, and areas in images. OpenCV is used in shape analysis (for the binary masks of specimen bodies) as well as for filtering out accidental partial segmentations. Pillow handles the vast majority of the image-based operations, including resizing, scaling, cropping, color conversion, mask application, drawing bounding boxes, and rotating images. We also use Matplotlib (Hunter, 2007) to plot visual maps of specimen dimensions. For all AI-based processing steps, we utilize Roboflow (Dwyer et al., 2024) and Anthropic (<https://www.anthropic.com/api>) Python packages. The last key package we use with DrawerDissect is pyYAML (<https://github.com/yaml/pyyaml>), which allows us to organize the pipeline with customizable inputs in a single master configuration file (hereafter, *config file*).

### 2.2.2 Pipeline Configuration and Customization

By editing the *config file*, new users can tailor the pipeline for their specific needs. Essential inputs for the *config file* are the user's API Keys for Roboflow and Anthropic. Documentation for how to set up API access through these services can be found at [https://inference.roboflow.com/quickstart/configure\\_api\\_key/](https://inference.roboflow.com/quickstart/configure_api_key/) and <https://docs.anthropic.com/en/api/admin-api/apikeys/get-api-key>. The rest of the *config file* is set up to run using FMNH's defaults. These defaults will work in most cases, but the settings can be edited as-needed for different configurations. We describe these settings in greater detail in the github documentation (<https://github.com/EGPostema/DrawerDissect#user-settings>), but, briefly, users can modify: (1) Roboflow-specific settings, (2) Toggles for transcription steps, (3) Claude-specific settings, and (4) Step-specific memory usage.

### 2.2.3 Training Roboflow Models for Detection and Segmentation

DrawerDissect relies on Roboflow models for all steps requiring vision-only models, such as object detection and segmentation. We have found that general foundation models perform poorly for our tasks, and therefore trained specialized models to annotate particular aspects of images. Reading labels using LLMs, however, is sufficiently accurate without specialized fine-tuning, and in this case no further training was required. Training the detection and segmentation models requires a set of images that have been labeled with the desired output. We use Roboflow’s free labeling tools, such as smart polygon and label assist, to make this step faster. Once annotated, we then split the images into training, validation, and test sets at a standard ratio of 70:20:10. We also employ image augmentations to increase model generalizability with respect to image lighting, background color, subject orientation, etc. (Borowiec et al., 2022).

We trained our models using Roboflow servers. Training can take anywhere between a few minutes and several hours depending on the size and complexity of the model; 2 hours is typical for our largest models with datasets of >3,000 images. When the model is finished training, Roboflow automatically reports the model’s precision and recall, as well as confusion matrices to determine rates of false positives/negatives (Table 1). We trained and have provided public access to all the models used in our pipeline at [universe.roboflow.com/field-museum](https://universe.roboflow.com/field-museum). See Table 1 for model-specific performance, and [github.com/EGPostema/DrawerDissect/supplemental](https://github.com/EGPostema/DrawerDissect/supplemental) for the taxonomic groups included in each of our six public models.

**TABLE 1: Model composition and performance**

| model id<br>(*version)          | # of images<br>(**train, valid, test) | precision | recall | false positive<br>rate | false negative<br>rate |
|---------------------------------|---------------------------------------|-----------|--------|------------------------|------------------------|
| <b>trayfinder-base</b><br>(5)   | (180, 12, 6)                          | 99.6%     | 99.1%  | 0.02%                  | 0.01%                  |
| <b>trayfinder-popup</b><br>(17) | (162, 11, 6)                          | 100.0%    | 100.0% | 0.01%                  | 0.00%                  |
| <b>labelfinder</b><br>(7)       | (3555, 201, 103)                      | 96.0%     | 99.0%  | ***0.02%               | ***0.01%               |
| <b>bugfinder-kdn9e</b><br>(13)  | (4235, 238, 120)                      | 97.1%     | 96.7%  | 0.02%                  | 0.02%                  |
| <b>bugmasker-all</b><br>(5)     | (2035, 113, 62)                       | 99.9%     | 97.3%  | 0.02%                  | 0.02%                  |
| <b>pinmasker</b><br>(6)         | (2210, 126, 67)                       | 91.0%     | 93.0%  | 0.08%                  | 0.07%                  |

\*Specifically, the model version that we used to produce the results in Section 3.1; not necessarily the most current version.

\*\*Includes augmented images, which increases the size of the training split x5.

\*\*\*Rate is averaged across the three detection classes (barcode, geocode, and label).

## 2.2.4 Costs and Alternatives

DrawerDissect uses API keys for paid AI services. By default, users are set up to use public FMNH models via Roboflow for image processing ([universe.roboflow.com/field-museum](https://universe.roboflow.com/field-museum)), while all transcription steps are done via Anthropic. Therefore, the primary costs for using DrawerDissect are associated with API-based inference: i.e., sending an image to a trained model and getting the model's output back. Roboflow pricing plans can be found at [roboflow.com/pricing](https://roboflow.com/pricing), including a free tier. We describe our exact per-specimen cost of unit tray- and specimen-level transcription in Section 3.1.4.

We anticipate that using API services like Roboflow and Claude to run DrawerDissect will be the easiest and most cost-effective method for the average user. These services sidestep the challenges of setting up local servers or cloud computing, at the cost of not being fully open-source. Additionally, running AI models locally requires powerful servers with advanced GPUs, costing tens of thousands of dollars to set up and high energy costs to run. We have not added these options to DrawerDissect for two reasons: (1) this raises the technological barrier to our intended user base and (2) advanced users experienced in AI coding, or large institutional teams, would be able to update DrawerDissect open-source code themselves. These experienced users that could modify DrawerDissect code to use local processing instead of API calls, relying on open-source or otherwise freely available models such as Llama (<https://www.llama.com/>) and DeepSeek (<https://www.deepseek.com/>) as LLM replacements for Anthropic API, and pytorch (<https://pytorch.org/>), fastai (Howard & Gugger, 2020) and the timm library (<https://timm.fast.ai/>) as replacements for the vision tasks.

## 2.3 Running DrawerDissect

DrawerDissect consists of a series of steps (Fig. 2) that can be run automatically, in sequence, with a single command: `'python process_images.py all'`. For a full list of command-line steps, optional command-line arguments for running specific combinations of steps, and/or running DrawerDissect on specific drawers, see [github.com/EGPostema/DrawerDissect#advanced-usage](https://github.com/EGPostema/DrawerDissect#advanced-usage).

### 2.3.1 Object Detection and Segmentation

Drawerdissect uses Roboflow models (Table 1) in sequence to find and outline different objects in images, which we describe below. All models output a JSON file with coordinates that are then used by DrawerDissect functions to crop, measure, filter, or mask images, depending on the

step. Due to file size and format constraints, all TIFs are converted to JPGs before processing, and large images are resized to fit within a 1000x1000px square (maintaining aspect ratio). During cropping steps, the coordinates from downsized images are mapped to the original, full-sized image based on the scaling factor recorded when images are initially resized.

### *trayfinder-base*

This model detects unit trays from drawer images using a Roboflow 3.0 Object Detection model. It is used by default in the current version of DrawerDissect.

Input: Resized whole-drawer images (without pop-up labels).

Output: Bounding box coordinates around all unit trays.

### *trayfinder-popup*

This model detects unit trays from drawer images using a Roboflow 3.0 Object Detection model. This is a FMNH-specific model, *not* used by default in the current version of DrawerDissect.

Input: Resized whole-drawer images (with pop-up labels).

Output: Bounding box coordinates around all unit trays, including the pop-up label.



**Figure 3.** Left: unit trays in a drawer detected by *trayfinder-base*. Right: two examples of label detection by *labelfinder*. Top: *labelfinder* detects three classes of text from an FMNH pop-up label (a 5-digit barcode, the taxonomic identity, and the 3-letter geocode). Bottom: *labelfinder* detects a handwritten species name located within the unit tray.

### *labelfinder*

This model detects three classes of text (taxonomic information, barcodes, and geocodes) if present (Fig. 3), using a Roboflow 3.0 Object Detection model. It is used by default in the current version of DrawerDissect.

Input: Resized unit tray images.

Output: Bounding box coordinates around the locations of tray-level text elements.

### ***bugfinder-kdn9***

This model detects pinned insects (Fig. 4a) using a Roboflow 3.0 Object Detection model. It is used by default in the current version of DrawerDissect.

Input: Resized unit tray images.

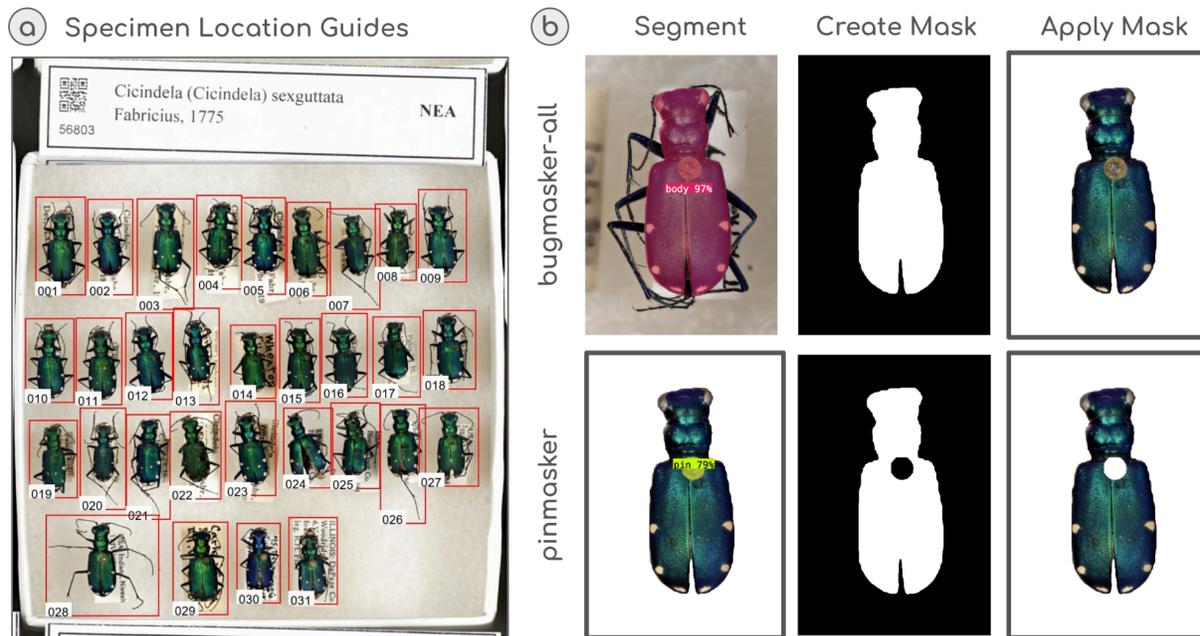
Output: Bounding box coordinates around all pinned specimens, with a 10-pixel buffer.

### ***bugmasker-all***

This model outlines the main body of insects, excluding legs and antennae (Fig. 4b), using a Roboflow 3.0 Instance Segmentation model. It is used by default in the current version of DrawerDissect.

Input: Dorsal images of insects on a neutral background.

Output: Coordinates that represent points along the outline of the specimen's body.



**Figure 4.** (a) Tray guides produced from *bugfinder-kdn9* coordinates. Specimens are automatically numbered from left to right, top to bottom. The number corresponds to the image name produced by DrawerDissect. (b) How Roboflow coordinates from *bugmasker-all* and *pinmasker* (the colored overlays) are translated into binary masks (black and white PNGs) by DrawerDissect, which are then used to create the final masked specimen images.

### ***pinmasker***

This model outlines the specimen pin (Fig. 4b), if present, using a Roboflow 3.0 Instance Segmentation model. It is used by default in the current version of DrawerDissect.

Input: Masked, body-only dorsal images of insects with a solid white background.

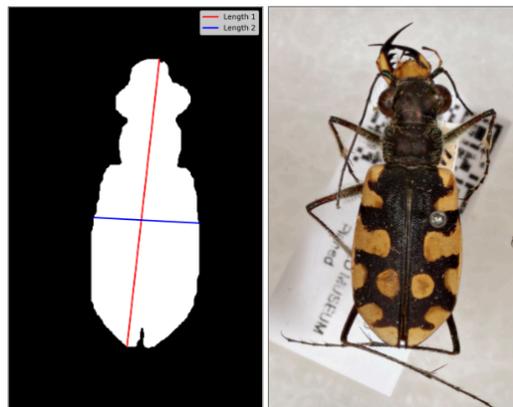
Output: Coordinates that represent points along the outline of the pin.

### 2.3.2 Generating Masks and Transparencies

To create transparent versions of the masked specimens, coordinates from *bugmasker-all* and *pinemasker* are converted into numerical arrays, which are then used to generate binary masks: black and white images (PNGs) where all background pixels are black and all body pixels are white (Fig. 4b). We then run the masks through a filtering step to remove partial segmentations. The binary mask is applied to the full-color specimen, such that white portions of the mask clip out the full-color specimen, while the black portions are turned transparent.

### 2.3.3 Measuring Specimens

For each specimen mask, we get two length measurements: *len1*, which is the greatest distance between any two points on the outline, and *len2*, calculated as the maximum distance perpendicular to *len1*. These measurements are good estimates for length and width though the script is intentionally agnostic to orientation. We also calculate the area, in pixels squared, of the body mask. To verify these measurements, *DrawerDissect* outputs a map of *len1* and *len2* for each specimen mask (Fig. 5). If users know the pixel:unit ratio for their camera, *len1*, *len2*, and area can be converted for each specimen.



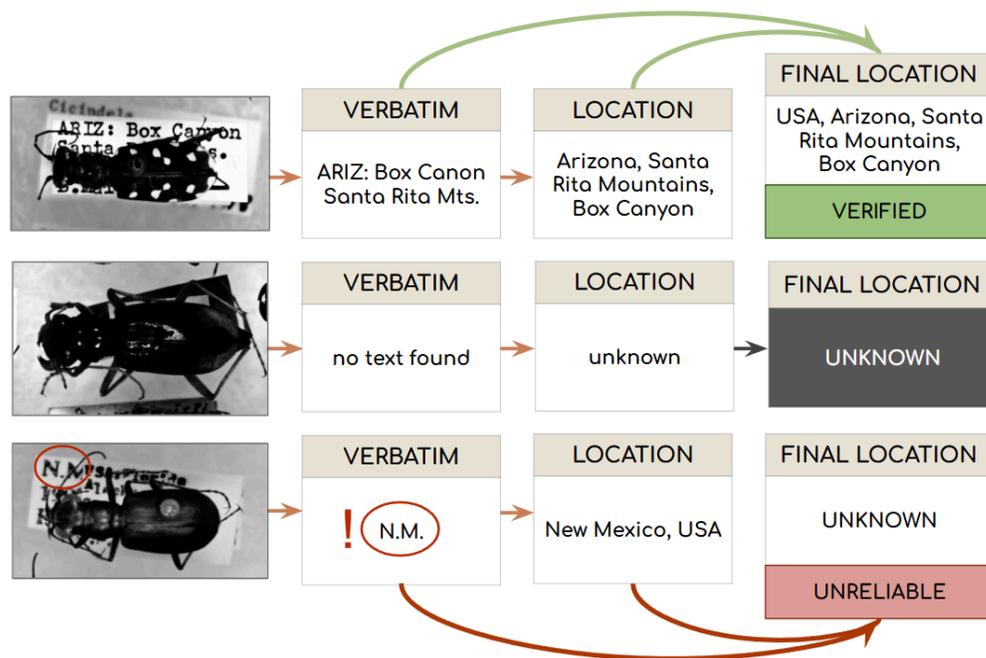
**Figure 5.** A map of *len1* (red) and *len2* (blue) for the specimen on the right, generated by *DrawerDissect*.

### 2.3.4 Claude (LLM) text transcription

To transcribe image text, we use a combination of Claude API and customizable prompts. For the results in Section 3.1, we used the model *claude-3-7-sonnet-20250219*. Before transcription, all image inputs are preprocessed to enhance text visibility by converting to grayscale and increasing contrast. Both handwritten and typed material can be transcribed (Fig. 3), though transcription errors are more likely with handwriting. We review the results of all AI transcription steps manually.

Unit tray label text is transcribed from cropped text images generated by *labelfinder* (Fig. 3). We have three separate prompts for each type of transcription, which are tailored to how these different types of information are structured. These prompts can be edited by users for their specific label structure in the *config file*, though our default setting to detect and transcribe taxonomic identity will likely work for most cases.

In many cropped specimen images, text from the top label (usually containing geographic information) is visible. We tested a novel method of metadata extraction based solely on this fragmented information, using a multistep approach (Fig. 6). The prompts for each of these steps can be edited in the *config file*. The outputs of this process are transcriptions of the verbatim text, an initial location estimate, the model's final location estimate, and an assessment of the estimation's quality (Fig. 6). The quality ranks are defined as: (1) **verified**, meaning the location estimate is logical given the verbatim text, (2) **unreliable**, meaning there is not sufficiently specific verbatim text to justify the location estimate, and (3) **unknown**, meaning that there was not enough verbatim text to result in any location estimate at all.



**Figure 6.** Our three-step process for LLM-based transcription of specimen location labels. Top: a successful transcription, resulting in a verified location with a standardized final location (despite a typo in the verbatim text). Middle: an unknown location and final location due to a lack of visible text in the image. Bottom: an unreliable location due to the verbatim text's lack of geographic specificity, and thus an unknown final location.

### 2.3.5 Data outputs and databasing

The final step of the pipeline compiles a time-stamped folder with data summaries at the drawer, tray, and specimen level for a given run. At the drawer level, we provide a simple summary of the number of trays, specimens, and masked specimens per drawer. DrawerDissect also generates a lookup table for tray-level information, such as barcode, geocode, taxonomic identity, specimen count, and masked specimen count. The last dataset merges all specimen-level data that were generated by the pipeline with tray-level data.

To upload the specimen images and their metadata to KE EMu, we print specimen labels with FMNH-INS numbers and an associated QR code. We apply these labels to each pinned specimen, using the tray guides (Fig. 4b) to match the DrawerDissect specimen image (e.g. spec\_001) to the FMNH-INS number. Lastly, we run a Python script to reorganize data to fit the metadata structure expected by KE EMu, and batch-import the images plus metadata.

### 2.4 Use Case 1: Batch color analysis of the big sand tiger beetle

Masked specimens are a key output of DrawerDissect. These images are useful for quickly extracting and analyzing phenotypic information from large numbers of specimens. To test the utility of these images, we targeted a specimen- and data-rich taxa within the FMNH tiger beetle collection: the big sand tiger beetle (*Cicindela formosa*). *C. formosa* has long been noted to possess a variety of color patterns across its range, with multiple purported subspecies (Pearson et al., 2006). There are clear regional variations in this species' maculations (the distinctive white markings on the elytra) which are often used to justify subspecies (French et al., 2021; Gaumer, 1977; Pearson et al., 2006). The FMNH tiger beetle collection includes >300 *C. f.* specimens collected across North America, and is particularly rich in two semi-sympatric subspecies: *C. f. formosa* and *C. f. generosa*. Due to a previous digitizing effort, the FMNH has records of the collection locations (most with coordinates) for all specimens of these subspecies. Given the uncertainty of the subspecies' classification and their regional diversity in the FMNH collection, we used this complex to show how DrawerDissect's outputs can be used in comparative phenotypic analyses. Given the use of color and pattern characteristics to identify *C. f. formosa* versus *C. f. generosa*, we might expect the two subspecies to significantly differ in appearance, even in regions where they overlap. However, the difference in color and maculation coverage between the subspecies could also be explained by local adaptation to certain biotic and/or abiotic pressures. In this case, we might expect the subspecies to differ *only* in the non-sympatric portions of their ranges, and for specimens from both subspecies to vary along environmental gradients that are known to impact animal coloration (Postema et al., 2023).

### 2.4.1 Phenotype extraction and statistical analysis

We used an existing ImageJ batch-analysis pipeline (method from Hancock et al., 2025) to quantitatively measure CIELab color and pattern geometry in the masked *C. formosa* specimens ( $n = 374$ ). As this phenotyping method outputs 36 different chromatic and geometric measurements, we used principal component analysis (PCA) to reduce the dimensionality of our dataset and to identify the main components of variation for our two tiger beetles. All statistical analyses were carried out using R (version 4.3.2). We used linear mixed models with PCs 1-3 as the dependent variables and both subspecies and sympatry as factors. We used the latitude and longitude data for each specimen to create a binary variable for sympatry: specimens were considered to be sympatric if they were within 3-degrees of one another. To further test the visual discreteness of the two subspecies we used k-means clustering of PCs 1-3 set to two clusters to determine whether or not the two subspecies could be naively identified within the dataset. Only specimens with coordinates were used in these analyses ( $n = 347$ ). To compare the effect of latitude, mean annual temperature (celcius), and mean annual precipitation (mm) on the mean CIELab L\* value (lower = darker) and the mean CIELab a\* value (greater = redder), we again used linear mixed models. Mean annual temperature and precipitation measurements were collected from a guide to North American level III ecoregions (Wiken et al., 2011); we then mapped each specimen's collection location (coordinates) to ecoregion using the Ecoregion Locator tool (<https://bplant.org/>). For all of our analyses, numeric variables were rescaled to a mean of 0 and a standard deviation of 1. The year the specimens were collected (to account for wear) and subspecies were used as random effects for the ecocline analyses. Residual plots were used to verify that all models met assumptions, and where required, data was transformed to match a normal distribution, e.g. square-root transformation.

### 2.5 Use Case 2: Training a classification model to identify *Cicindela* specimens

A major contribution of batch-imaging and image processing tools like DrawerDissect is their ability to supply taxonomic identification AI models (Spiesman et al., 2021; Sun et al., 2021; Truong & Van der Wal, 2024; Welch & Lundgren, 2024), and other curation tools, with the large amounts of training data they require. In our collection of tiger beetles, the majority of specimens (7000+) belong to the genus *Cicindela*, representing 239 unique species and subspecies across 2 subgenera (*Cicindela* and *Cicindelia*). There are also a small number of specimens in our collection (12) that were identified to the genus *Cicindela* but not to species. As a proof of concept, here we train an image model to identify species and subspecies of *Cicindela* based on the masked images we produced with DrawerDissect.

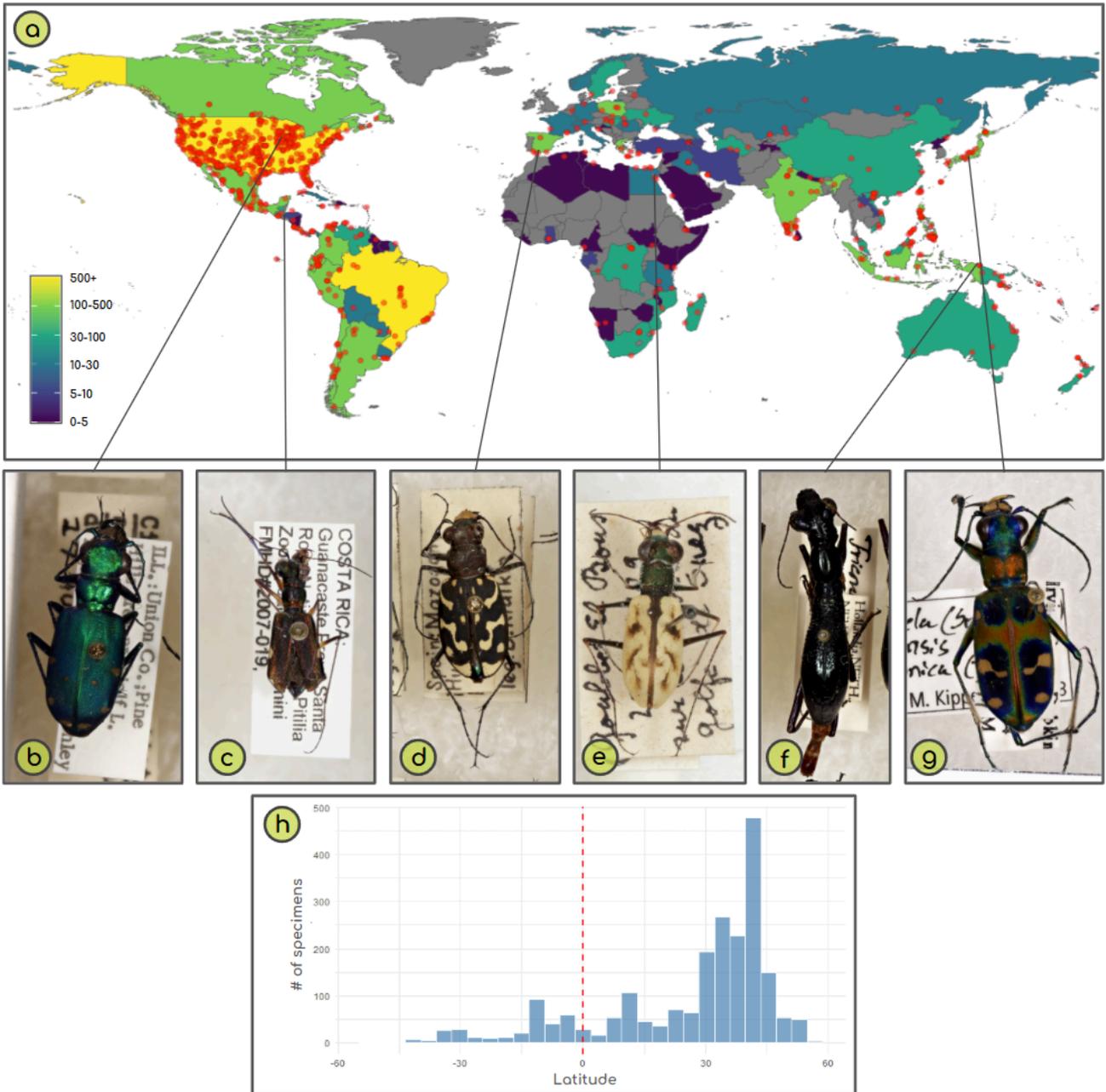
To train and evaluate a taxonomic classification model for *Cicindela*, we split all imaged and identified specimens into training, validation and test sets at an approximately 70:20:10 ratio. We

included taxa with single images in the training set but not validation or test sets. Our goal was to train a multilabel model that could simultaneously identify a given specimen to species and subspecies. The input data to the models were individual masked tiger beetle images with white backgrounds, associated with all species and subspecies labels. We primarily used the Python libraries pytorch (Paszke et al., 2019), fastai (Howard & Gugger, 2020), and timm (Wightman, 2019) for deep learning, and computations were performed on 2 NVIDIA RTX A5000 GPUs with 24GB of memory each. One of the primary challenges with any biological dataset for taxonomic identification is data imbalance: some species are very abundant and represented by hundreds of specimens in our dataset, while others are rare and represented by unique specimens (Figure 15a). To circumvent this challenge, we applied methods similar to de Medeiros et al. (2025), who performed taxonomic identification based on an imbalanced dataset of images representing genome composition (Asprino et al., 2025). Additional details on our model architecture, augmentations, and our trained model weights can be found at [huggingface.co/brunoasm/eva02\\_large\\_patch14\\_448.Cicindela\\_ID\\_FMNH](https://huggingface.co/brunoasm/eva02_large_patch14_448.Cicindela_ID_FMNH) and [github.com/de-Medeiros-insect-lab/Cicindelinae\\_ID](https://github.com/de-Medeiros-insect-lab/Cicindelinae_ID). We evaluated the final model using the test set, with a confidence threshold of 0.5 to make predictions. Finally, we also predicted labels for the 12 unknown samples. Before using the model to identify the 12 unknown samples, we manually identified each specimen using guides to the *Cicindela* for comparison (Table 3).

### 3 RESULTS

#### 3.1 Digitizing the FMNH tiger beetle collection

With DrawerDissect, we were able to produce a complete inventory of the FMNH tiger beetle collection. We processed the collection into 13,496 separate images of intact specimens obtained from 44 drawers of pinned tiger beetles from around the world (Fig. 7). All specimen photos are linked to taxonomic and broad biogeographic metadata. The majority of these specimens (13,484) were successfully masked and measured (length, width, and area, in mm). Over a quarter (3,627) of the specimen images are associated with more specific collection locations, as a result of automated transcription via DrawerDissect, manual transcriptions for one specimen per tray, or additional metadata transcriptions from a previous databasing effort. Our collection contains 56 unique genera and 663 unique species of tiger beetle. This represents nearly a third of the ~2,300 known species globally (Gough et al., 2018). FMNH tiger beetles were collected from at least 1,002 unique locations across 91 countries, and all continents excluding Antarctica (Fig. 7). The collection is strongly biased towards the Northern hemisphere (64.4% from the Nearctic). Among the specimens with transcribed locations, the vast majority were from the United States. The collection also contains a large proportion of Neotropical specimens (18.8%), particularly from Brazil.



**Figure 7.** (a) A world map showing the approximate number of FMNH specimens per country. The per country counts are estimated from the number of specimens in a given tray and the countries of representative specimens from the same tray whose locations were determined (manually or via LLM). Yellow = 500+ specimens, green = 100-500, teal = 30-100, blue = 10-30, dark blue = 5-10, purple = 1-5, and gray = 0. Specific known collection sites with coordinates are marked with red points. Middle: (b) *Cicindela sexguttata* from Pine Hill, IL, USA. (c) *Odontocheila iodopleura* from the Pitilla Zoological Station, Costa Rica. (d) *Lophyra flexuosa* from Mazagón, Spain. (e) *Hypaetha singularis* from Ghoubbet-El-Bous, on the coast of the Gulf of Suez in Egypt. (f); *Tricondyla aptera aptera* from Jayapura, Indonesia; (g) *Cicindela chinensis japonica* from Ōdai, Japan. Specimens not to scale. (h) specimen counts by absolute latitude (n = 2159).

### 3.1.1 Unit tray cropping and transcription

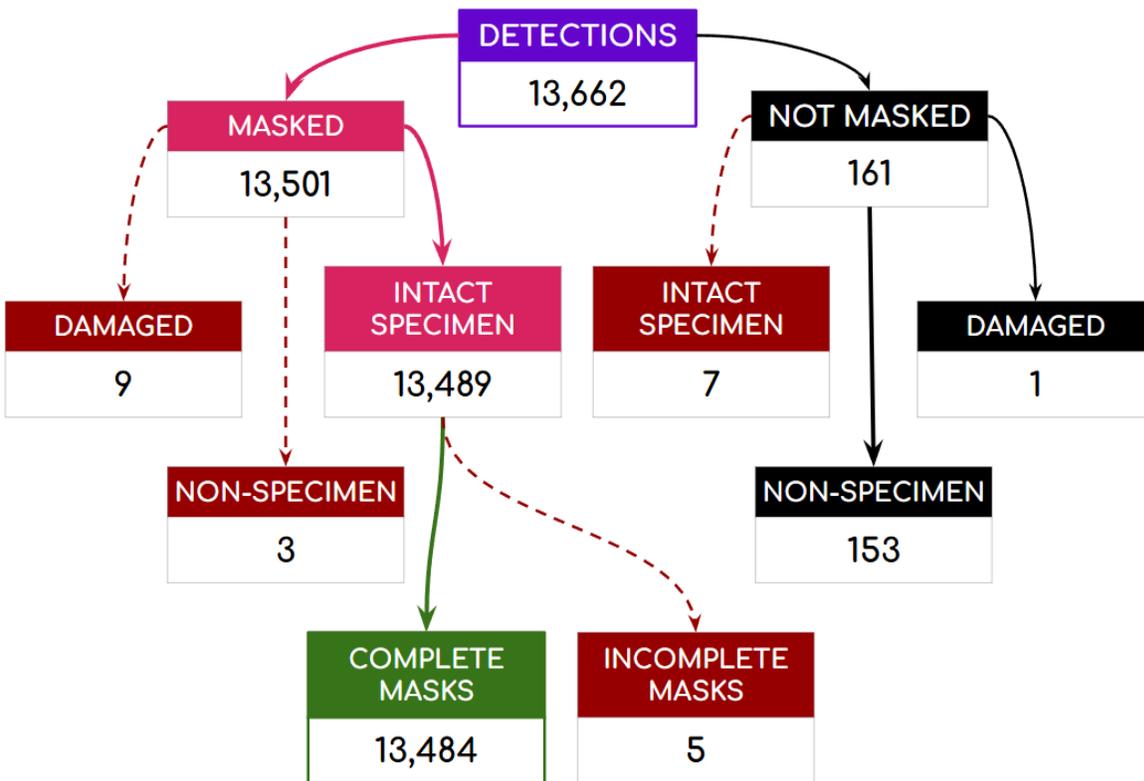
Out of 941 trays, *trayfinder-popup* made no detection errors. However, 36 trays (3.9%) required manual transcription of the taxonomic identity, geocode, and/or barcode. For the majority of trays, this happened because *labelfinder* failed to detect the information on the label (Fig. 3), meaning there were no cropped images for Claude to transcribe. In four instances, label information was detected, but the bounding box did not fully encompass the label information - thus some text had to be transcribed manually. We observed only two instances of taxonomic transcription typos at the tray level: the genus *Manticora* was misspelled as *Mantichora*, and the genus *Distipsidera* was misspelled as *Distintipsidera*. For barcodes, there were three instances where the numbers 5 and 6 were erroneously switched, likely due to their similar shape. For geocodes, there were 9 transcription errors that resulted in the incorrect biogeographical realm. All text in this transcription step was typed.

### 3.1.2 Specimen detection

Overall, *bugfinder-kdn9e* made 13,662 specimen detections. 153 (1.1%) of these were false positives, that is, non-specimens that were erroneously labeled as specimens. This is a slightly higher rate of false positives than the model's roboflow test set (0.02%; Table 1). For the most part these detections could be filtered out using the results of the masking and measuring steps, as non-specimens tend to produce either an oddly-proportioned mask or no mask at all. To check for missed specimens, we visually scanned all 930 tray guides and counted any specimens that lacked a bounding box. No specimens were missed by the model, meaning the rate of false negatives was 0% (lower than the roboflow-calculated rate of 0.02%; Table 1).

### 3.1.3 Specimen masking

Out of the 13,662 cropped specimen photos generated from *bugfinder-kdn9e*, *bugmasker-all* found masks for 13,501 images (Fig. 8). Of the 161 images where no mask was found, 154 were true negatives, meaning they did not contain a specimen, or contained a heavily damaged specimen. Only 7 were false negatives, where no mask was found despite the image containing a complete, intact specimen. In the masked set, we found 12 false positives, where a mask was applied to a non-specimen (3) or heavily damaged specimen (9). In 5 cases a mask did not fully outline an intact specimen. Our actual rates of false positives (0.09%) and false negatives (0.05%) for our set of tiger beetles were below the rates reported for the model in the confusion matrix generated by Roboflow (Table 1) based on the test set with a broader taxonomic diversity.



**Figure 8.** Flowchart showing the outcomes of DrawerDissect masking for the tiger beetle collection. Solid lines indicate correct responses, such as masking intact specimens or failing to mask non-specimens. Dotted lines indicate incorrect responses, such as masking damaged specimens/non-specimens, or failing to mask intact specimens

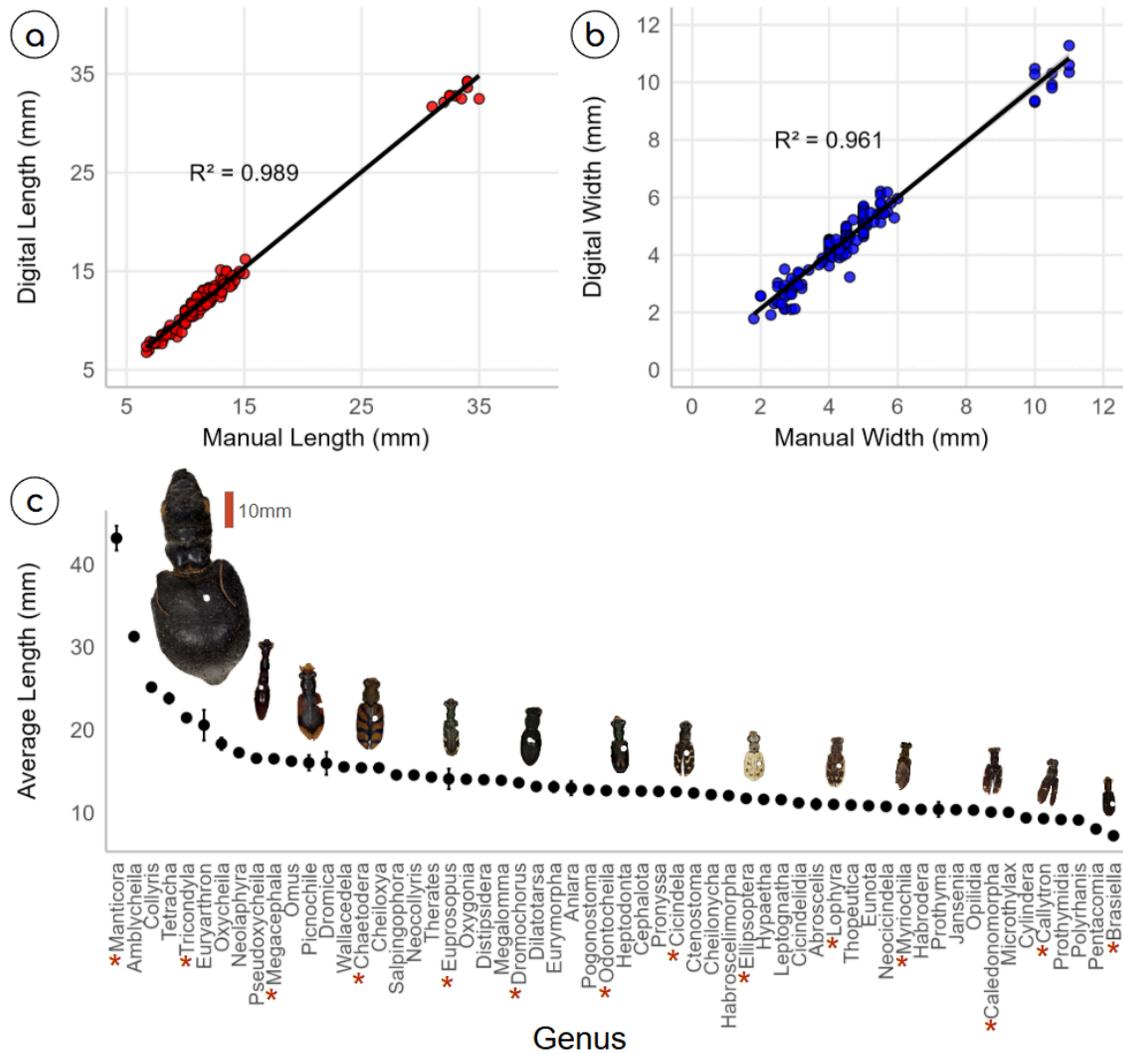
### 3.1.4 Specimen-level label transcription

For the majority of specimens (80.3%), no location could be determined from the image alone, likely because either no text was visible, the specimen did not have a label with location information, or the location text was too fragmented (or too non-specific) to make a reliable determination (Fig. 6). In total we were able to reconstruct locations for 2,475 specimens (18.2%). This included both perfect transcriptions from the text available (9.4% of images), and cases where there was enough context from the LLM’s verbatim text transcription that a location could be determined manually by a human reviewer even if the model failed to make a determination (8.8% of images). Additionally, the frequency of false positives - incorrect location estimates that were incorrectly judged as ‘reliable’ by the LLM - was low, at 1.6% of all images. Of the 425 unique collection locations extracted from the specimen images by AI, the majority (69.6%) could be linked to a set of centroid coordinates (with up to a 20mi radius of uncertainty) on Google Maps. Running all available LLM-based steps on our set of >13,000 tiger beetle images cost ~\$100. For a per-specimen breakdown of cost and speed, see Table 2.

**TABLE 2: Time and cost of specimen transcription**

| cost per specimen (USD) | cost per 1000 specimens (USD) | processing time (1 specimen) | processing time (1000 specimens) |
|-------------------------|-------------------------------|------------------------------|----------------------------------|
| \$0.00757               | \$7.57                        | 2.7 seconds                  | 45 minutes                       |

**3.1.5 Size measurements**



**Figure 9.** (a) The relationship between manual and digital length for a set of randomly selected specimens ( $n = 143$ ), both in mm. (b) The relationship between manual and digital width for the same set of specimens. (c) The average body length, in mm, +/- SE of each genera in the FMNH collection ( $n = 13,484$ ). \*Indicates genus with an associated specimen icon (generated by DrawerDissect), ordered left to right by size. All specimen icons are to scale and are positioned roughly above their corresponding genus.

We were able to measure all specimens that were properly masked (Fig. 8), generating length, width, and area in mm/mm<sup>2</sup> for 13,484 specimens. To assess the accuracy of the digital measurements, we compared them against a randomly selected subset of specimens ( $n = 79$ ) that we hand-measured using calipers (Fig. 9a-b). We found that the digital and manual measurements were highly correlated (length:  $R^2 = 0.989$ , width:  $R^2 = 0.961$ ). Body area, length, and width varied both within and across genera (Fig. 9c). The largest specimens were *Manticora* collected from South Africa (average length: 43.1mm) while the smallest were in the Neotropical genus *Brasiella* (average length: 7.2mm).

### 3.1.6 Imaging and processing time

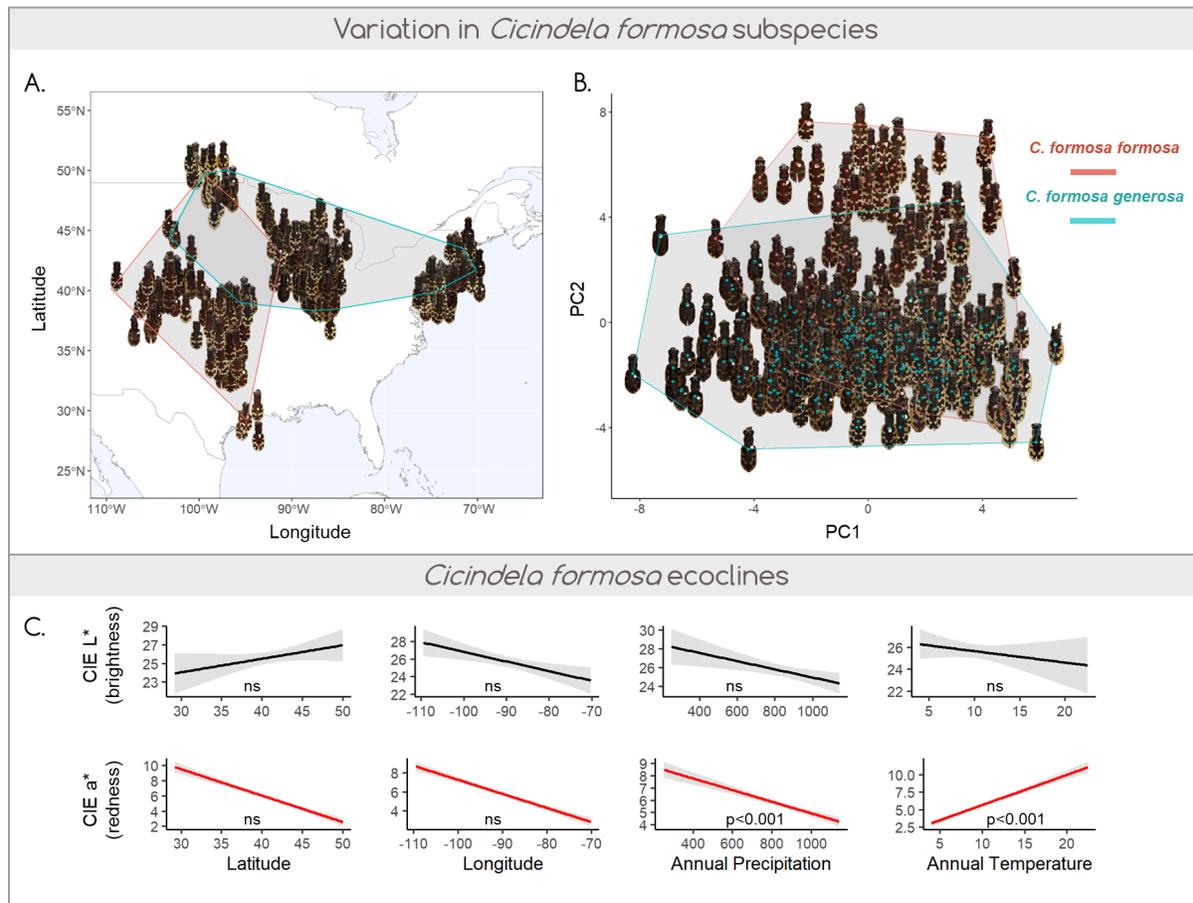
Imaging all 44 drawers took approximately 2 weeks, including moving drawers, inventorying unit tray labels, and printing and replacing header labels (70 hours). This is equivalent to ~190 specimens imaged per work-hour. On a Windows computer with an AMD Ryzen™ 7 Processor 7800X3D CPU and 32GB of RAM, DrawerDissect is able to fully process (i.e. produce all possible DrawerDissect outputs) a 8GB drawer image in 16 minutes. For 44 drawers it took ~10 hours to fully process; batch image processing runs are more efficient per drawer due to parallel processing. We spent an additional 40 hours on data cleaning and validation, and another 35 hours to manually transcribe location information for one specimen per unit tray - an optional step to supplement DrawerDissect's location estimates. In total, the full workflow (imaging, processing, and data-cleaning) for the full tiger beetle collection took about 155 hours. This translates to ~3.5 hours per drawer or ~40 seconds per specimen. This is significantly faster than manually photographing, measuring, and outlining specimens.

### 3.2 Use Case 1: Color variation between and within subspecies of *Cicindela formosa*

Our PCs comprised 53% of variation within our coloration metrics. PC1 (22%) primarily consisted of luminance with higher values indicating lighter, more contrasting maculation; PC2(19%) consisted of CIE a\* (red-green) with higher values indicating redder coloration; and PC3 (12%) consisted of lower spatial frequency patterns. Both subspecies varied considerably in PC1 (Fig. 10b). Meanwhile, *C. f. formosa* on average were redder in their body coloration than *C. f. generosa* (Higher PC2, Fig. 10c). For both subspecies classifications, specimens at the highest latitudes had maculations that were unusually wide (Fig. 10a).

For PC1, sympatry had a significant effect ( $B = 1.68$ ,  $SE = 0.62$ ,  $t\text{-value}_{287} = 2.72$ ,  $p=0.007$ ), likely due to increased maculation size in Northern regions where populations overlapped. For PC2, *C. f. generosa* were less red than *C. f. formosa* ( $B = -5.12$ ,  $SE = 0.369$ ,  $t\text{-value}_{304} = -13.93$ ,  $p<0.001$ ). Individuals of both subspecies were less red in sympatric regions ( $B = -2.74$ ,  $SE = 0.430$ ,  $t\text{-value}_{336} = -6.37$ ,  $p<0.001$ ), but less so for *C. f. generosa* ( $B = 2.77$ ,  $SE = 0.604$ ,  $t\text{-value}_{342} = 4.60$ ,  $p<0.001$ ). For PC3, no significant differences were observed. When using k

means clustering to assess variation in tiger beetles, cluster 1 comprised 63 of the 97 *C. f. formosa* and cluster 2 comprised 244 of the 250 *C. f. generosa*. Naïve clusters largely followed the same geographic gradient as the museum classifications. Luminance was not significantly affected by any of the geographic or environmental factors included in our analysis, while redness increased in regions with lower mean annual precipitation ( $B = -0.53$ ,  $SE = 0.123$ ,  $t\text{-value}_{245} = -4.33$ ,  $p < 0.001$ ) and higher mean temperature ( $B = 0.53$ ,  $SE = 0.133$ ,  $t\text{-value}_{259} = 3.97$ ,  $p < 0.001$ ) (Figure 10c).



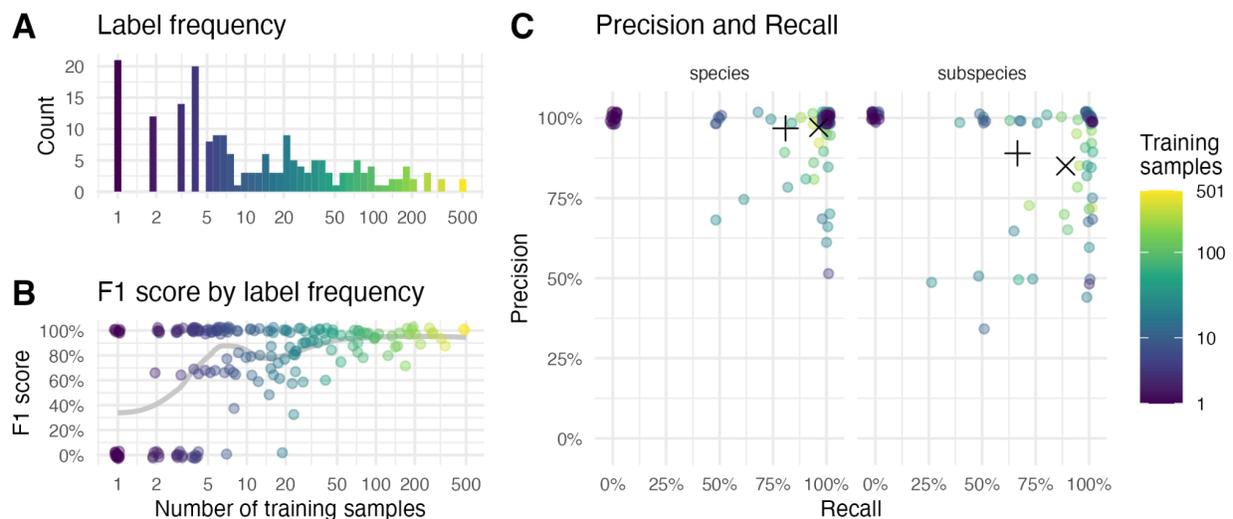
**Figure 10.** A) geographic distribution of collected big sand tiger beetles, *Cicindela formosa*, subspecies *C. f. formosa* in the west and *C. f. generosa* in the east. Polygons show the convex hulls of the subspecies ranges. B) The phenotypic ranges of the subspecies expressed by PC1 (darker to lighter) and PC2 (redness). C) linear model plots for CIE L\* (luminance) and CIE a\* (greater values = redder) for biogeographic variables: latitude, longitude, mean annual precipitation (mm), and mean annual temperature (°C). Significance is labeled on the plot (ns = non-significant).

Overall, we found that the two subspecies are significantly different in appearance (*C. f. formosa* are significantly redder). However, this distinction decreases where specimens overlap in range. Further genetic analysis would be required to determine whether the similarity is due to hybridization between the subspecies and/or convergent evolution of less red elytra within

similar environments (French et al., 2021). The pattern of increased redness in hotter, more arid environments was mainly driven by the large populations of *C. f. formosa* in the western and southwestern United States. A thermoregulatory (heat-reducing) function of increased redness is unlikely, as metallic red portions of *C. f. formosa* elytra show increased IR absorbance compared to white portions; however, the thermodynamic properties have not been compared between redder versus browner elytra (Schultz & Hadley, 1987). Elytral cover may instead be linked to local background-matching camouflage, as in other tiger beetle species (Yamamoto & Sota, 2020). This hypothesis is supported, anecdotally, by the presence of strikingly red soils in certain hotter, more arid *C. f. formosa* collection sites (Kansas, Texas, New Mexico, parts of Utah and Colorado; nrcs.usda.gov soil colors of the US dataset).

### 3.3 Use Case 2: Results of *Cicindel-ID* specimen identification

Overall, species and subspecies of *Cicindela* could be identified with high accuracy even for modest sample sizes in the training set. The average F1 score, which balances precision and recall, increases with availability of samples in the training set (Fig. 11b). However, most taxa resulted in high prediction accuracy even with modest sample sizes, with high precision throughout and recall more variable across labels (Fig. 11c). Specifically, we observed 97.0% precision and 96.4% recall for species when averaged over specimens (Fig. 11c), decreasing to 96.8% and 80.9%, respectively, when averaging over labels. For subspecies, these numbers were somewhat smaller, 85.0% precision and 85.0% recall averaging over specimens and 89.0% precision and 66.5% recall when averaging over taxa. This means that most specimens can be predicted with very high accuracy, while the model may fail to recognize rare species and subspecies.



**Figure 11.** Taxonomic identification based on an image classification model trained on images of *Cicindela*. (a) Distribution of species and subspecies label frequencies (log scale), showing high imbalance. (b) F1 score increases with sample size used in training. Points represent taxonomic labels,

including both species and subspecies, and the gray trend line shows smoothed mean scores. (c) Precision and recall for each taxon (circles), with averages across samples (X), and across labels (+). A small jitter was added to individual labels in (b) and (c) to allow observation of overlapping records.

Using *Cicindel-ID* to identify unknown samples reinforces this pattern. Among the 12 unknown samples, 10 were common species correctly predicted according to our identifications (Table 3). Of the two samples with no prediction, one was possibly a subspecies rare in the FMNH collection and not present in the training set (*C. transbaicalica japonensis*; Shiyake, 2017). The model correctly failed to make a prediction in this case. We identified the other unknown sample as *C. sylvicola*, which included 25 specimens in the training set. The trained model has a moderate accuracy in identifying this particular taxon (75% precision, 60% recall), so this seems to be a case of a false negative.

**TABLE 3.** Model versus human identifications for 12 previously unidentified samples of *Cicindela*. (?) Indicates an uncertain human ID.

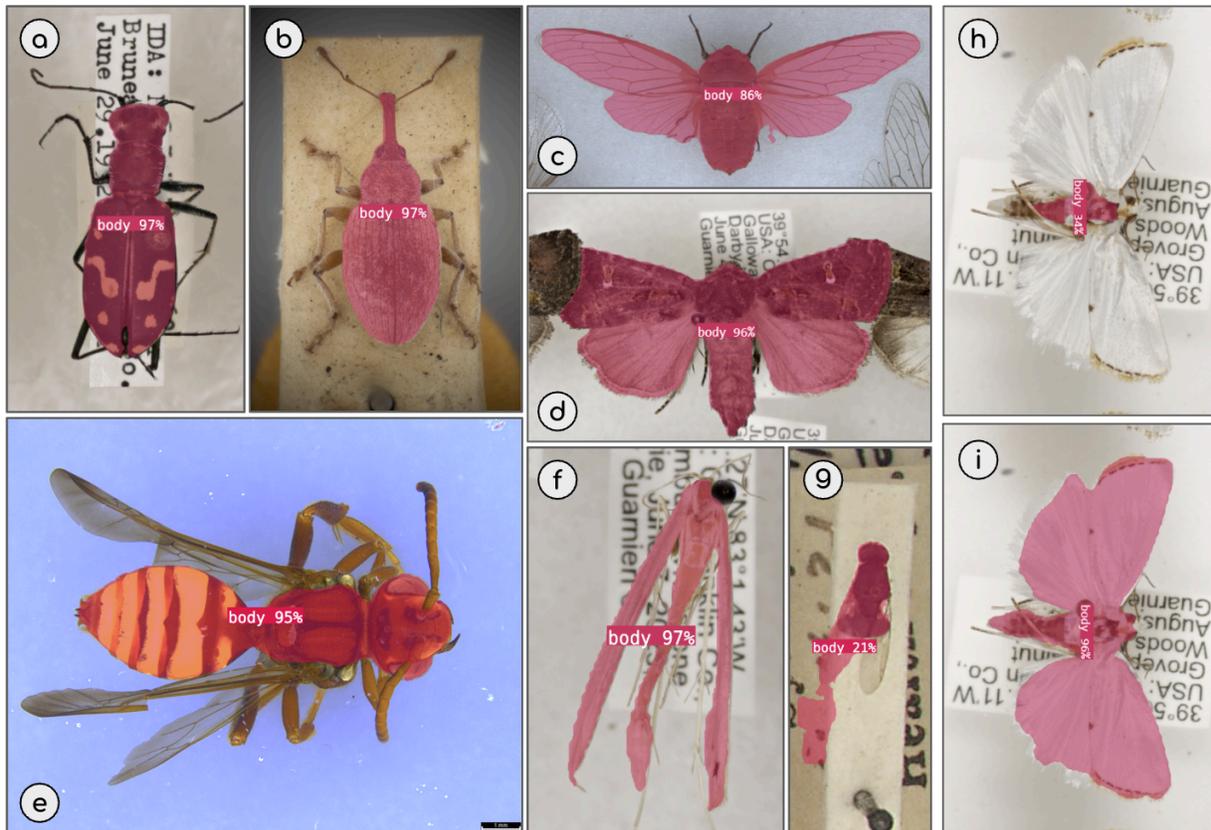
| # of samples | <i>Cicindel-ID</i> predictions     | Human identification                    | Identification source         |
|--------------|------------------------------------|---|-------------------------------|
| 7            | “ocellata”,<br>“ocellata_ocellata” | <i>C. ocellata ocellata</i>             | Pearson et al., 2006          |
| 2            | “sylvicola”                        | <i>C. sylvicola</i>                     | Trautner & Geigenmüller, 1987 |
| 1            | “Repanda”,<br>“repanda_repanda”    | <i>C. repanda repanda</i>               | Pearson et al., 2006          |
| 1            | N/A                                | <i>C. transbaicalica japonensis</i> (?) | Shiyake, 2017                 |
| 1            | N/A                                | <i>C. sylvicola</i>                     | Trautner & Geigenmüller, 1987 |

#### 4 DISCUSSION

DrawerDissect is the first usage of general-purpose equipment for mass digitization of whole pinned insect drawers yielding specimen-level images, masks, and data. Using DrawerDissect, we were able to successfully image, segment, and extract data (taxonomic, morphometric, and geographic) from 13,484 pinned insect specimens in two weeks. This is a marked increase in digitization efficiency at the FMNH. For comparison, in the last 20 years, fewer than 20,000 insect specimens were imaged in our entire pinned collection. Batch-imaging combined with AI processing is the clear way forward for large-scale digitization projects in natural history collections (Stenhouse et al., 2025; Weaver & Smith, 2023; Weeks et al., 2023). Through imaging services like Picturae, for example, several herbaria have already been imaged in full (e.g. the Smithsonian, which houses ~3.8 million specimens), and AI workflows have been

developed to further process the outputs of such digitization efforts (de la Hidalgo et al., 2022; Herbst et al., 2025; Weaver & Smith, 2023). For insects, few alternatives exist. There are ongoing developments on mass-digitization of fluid-preserved insect collections based on bespoke imaging and robotic equipment (Wührl et al., 2022). Picturae (<https://picturae.com/services/entomology-digitization/>) also offers a solution for pinned insects including metadata digitization that purportedly can reach a rate of 5,000 specimens daily. However, it involves custom robotic equipment, a large installation footprint, and individual handling of each specimen. InSelect (Hudson et al., 2015) is another option for processing whole-drawer images into single-specimen images. This tool is excellent for highly standardized specimens, but lacks customizable AI models that perform well on many types of variably-positioned specimens. Additionally, InSelect does not include built-in tools for image segmentation or text transcription.

Using a series of fine-tuned computer vision models (Table 1), DrawerDissect can produce high-quality images of specimens that exclude the background and insect pin (Fig. 4b). This automates a normally time-consuming image-preparation step that is necessary for most morphological analyses (Curlis et al., 2022; Van Belleghem et al., 2018; Weaver & Smith, 2023; Weeks et al., 2023; Weller et al., 2024). DrawerDissect's masking step (model id: *bugmasker-all*) is able to segment pinned insects of various shapes, sizes, colors, orientations, and degrees of overlap with other specimens (Fig. 12). Given the taxonomic composition of the training data, DrawerDissect currently performs best on beetles, moths, and morphologically similar groups (Fig. 12a-g). However, retraining existing models to identify new taxa (or to improve predictions for underperforming taxa; Fig. 12h-j) is simple: for example, to train *bugfinder-all* to detect various moth families, we only had to annotate ~100 new images. We anticipate that the models we or others develop for DrawerDissect will only improve in taxonomic and morphological generalizability as more training data are added.



**Figure 12.** Example *bugmasker-all* predictions for various taxa photographed with different camera systems. The pink overlay shows the location of the specimen body and model confidence is shown as a percent. (a) A tiger beetle (Coleoptera: Cicindelidae) from the FMNH collection, photo by EGP/LB/CH/BM, (b) a weevil (Coleoptera: Curculionidae) (c) a cicada (Hemiptera: Cicadidae) from the Australian National Insect Collection (ANIC) (d) A noctuid moth, overlapping with other specimens (Lepidoptera: Noctuidae), photo by LDG/TE (e) a wasp (Hymenoptera: Vespidae), photo by KW, (f) a plume moth (Lepidoptera: Pterophoridae), photo by LDG/TE, (g) a small fly (Diptera) from the ANIC. (h) An earlier version (v5) of *bugmasker-all* had difficulty detecting and segmenting a set of grass moths (Lepidoptera: Crambidae) with white wings. (i) Adding training data that targeted underperforming categories significantly improved model performance (v8).

One limitation of our approach is that the metadata collected is fragmentary. For example, DrawerDissect was only able to automatically transcribe collection locations for ~18% of our tiger beetles, and these locations are often partial (i.e. capturing only the state or county, but missing more precise location data). However, there are ongoing efforts to optimize the transcription of museum labels in terms of prompt engineering, model selection, and cost effectiveness (Herbst et al., 2025). We also demonstrate that this kind of data can still have research uses. As insect collections tend to have more metadata records than specimen images (Cobb et al., 2019), a possible strategy for batch-imaging is to target previously digitized groups that lack images, as we did in the case of *Cicindela formosa*. The combination of AI-driven phenotype extraction and traditional label transcription has the potential to produce rich datasets

to investigate the drivers of insect morphological diversity. For example, we were able to extract accurate length and width measurements from all 13,464 masked specimens (Fig. 9), as well as detailed color and pattern data from a subset of big sand tiger beetles (*Cicindela formosa formosa* and *C. f. generosa*; Fig. 10a-b). Using the latter dataset, combined with pre-digitized location records and climate data, we found significant differences both between subspecies and along environmental clines (Fig. 10b-c).

Species identification models (Borowiec et al., 2022; Spiesman et al., 2021; Sun et al., 2021; Welch & Lundgren, 2024) are another promising use case for DrawerDissect's masking step, as we demonstrate with the *Cicindel-ID* (Fig. 11). *Cicindel-ID* was able to achieve high precision (97%) and recall (96.4%) for species in the genus *Cicindela*, and 85.0% precision and recall for *Cicindela* subspecies (averaged across specimens). Masked museum specimens offer both quality and quantity as training data for species ID models: they are standardized, censored to avoid shortcut-learning (Geirhos et al., 2020; Weaver & Smith, 2023), taxonomically diverse, and intraspecifically varied. Backgroundless specimens can also be composited onto naturalistic backgrounds to simulate in situ photographs (Sun et al., 2021), potentially expanding training sets for field image classification.

AI, in the form of both tailored CV models and LLMs, is a powerful tool that is changing the scale and nature of biological research (Borowiec et al., 2022). In DrawerDissect, AI can quickly perform repetitive tasks that are relatively simple from a human perspective, such as transcribing text, detecting specimens, and outlining objects. Chaining these simple-but-specialized models together, in an assembly-line fashion, produces powerful workflows for processing batch-imaged specimens (Weaver & Smith, 2023). The outputs of these “multi-model” pipelines can then be fed into existing image analysis pipelines or used to train new CV models (Chan et al., 2019; Lürig, 2022; Van Belleghem et al., 2018; van den Berg et al., 2024; Weller et al., 2024), opening the door to truly high-throughput digitization and analysis of museum specimens.

## ACKNOWLEDGEMENTS

We thank the Grainger Bioinformatics Center for funding this research and providing computational support, as well as the Collaborative Imaging Lab and its manager, Stephanie Ware. We also thank the FMNH for funding support through the Science Innovation Award. This work was made possible by several FMNH internship programs, including the Women in Science Internship and the Prince Fellowship. We thank Gene Cooper, Jacob Witt, Janeen Jones, and Matt Wadleigh for their technical assistance and programming advice. We are grateful to Rick Ree for his insightful comments on the manuscript, to Tom Schultz for his expertise on tiger beetle coloration, to Maureen Turcatel for help with collection management, and to Matt Buffington for discussions on the value of whole-drawer imaging.

## REFERENCES

- Dwyer, B., Nelson, J., Hansen, T., et. al. (2024). Roboflow (Version 1.0) [Software]. Available from <https://roboflow.com>. computer vision.
- Shiyake, S. (2017). How to identify Tiger Beetles. Osaka Museum of Natural History.
- Asprino, R. C., Cai, L., Yan, Y., Flynn, P. J., Marinho, L. C., Duan, X., Anderson, C., Lyra, G. M., Davis, C. C., & de Medeiros, B. A. S. (2025). A curated benchmark dataset for molecular identification based on genome skimming. *Scientific Data*, *12*(1), 906. <https://doi.org/10.1038/s41597-025-05230-2>
- Bates, J. M., Fidino, M., Nowak-Boyd, L., Strausberger, B. M., Schmidt, K. A., & Whelan, C. J. (2023). Climate change affects bird nesting phenology: Comparing contemporary field and historical museum nesting records. *Journal of Animal Ecology*, *92*(2), 263–272. <https://doi.org/10.1111/1365-2656.13683>
- Borowiec, M. L., Dikow, R. B., Frandsen, P. B., McKeeken, A., Valentini, G., & White, A. E. (2022a). Deep learning as a tool for ecology and evolution. *Methods in Ecology and Evolution*, *13*(8), Article 8. <https://doi.org/10.1111/2041-210X.13901>
- Chan, I. Z. W., Stevens, M., & Todd, P. A. (2019). pat-geom: A software package for the analysis of animal patterns. *Methods in Ecology and Evolution*, *10*(4), 591–600. <https://doi.org/10.1111/2041-210X.13131>
- Crowell, H. L., Curlis, J. D., Weller, H. I., & Davis Rabosky, A. R. (2024). Ecological drivers of ultraviolet colour evolution in snakes. *Nature Communications*, *15*(1), 5213. <https://doi.org/10.1038/s41467-024-49506-4>
- Curlis, J. D., Renney, T., Davis Rabosky, A. R., & Moore, T. Y. (2022). Batch-Mask: Automated Image Segmentation for Organisms with Limbless or Non-Standard Body Forms. *Integrative and Comparative Biology*, *62*(4), Article 4. <https://doi.org/10.1093/icb/icac036>
- Davis, C. L., Guralnick, R. P., & Zipkin, E. F. (2023). Challenges and opportunities for using natural history collections to estimate insect population trends. *Journal of Animal Ecology*, *92*(2), 237–249. <https://doi.org/10.1111/1365-2656.13763>
- de la Hidalga, A. N., Rosin, P. L., Sun, X., Livermore, L., Durrant, J., Turner, J., Dillen, M., Musson, A., Phillips, S., Groom, Q., & Hardisty, A. (2022). Cross-validation of a semantic segmentation network for natural history collection specimens. *Machine Vision and Applications*, *33*(3), 39. <https://doi.org/10.1007/s00138-022-01276-z>
- de Medeiros, B.A.S., Cai, L., Flynn, P.J. et al. (2025). A composite universal DNA signature for the tree of life. *Nat Ecol Evol*. <https://doi.org/10.1038/s41559-025-02752-1>
- Duran, D. P., & Gough, H. M. (2020). Validation of tiger beetles as distinct family (Coleoptera: Cicindelidae), review and reclassification of tribal relationships. *Systematic Entomology*, *45*(4), 723–729. <https://doi.org/10.1111/syen.12440>
- French, R., Bell, A., Calladine, K., Acorn, J., & Sperling, F. (2021). Genomic distinctness despite shared color patterns among threatened populations of a tiger beetle. *Conservation Genetics*, *22*, 1–16. <https://doi.org/10.1007/s10592-021-01370-1>

- Gaumer, G. C. (1977). The variation and taxonomy of *Cicindela formosa* Say (Coleoptera: Cicindelidae) [Doctoral dissertation, Texas A&M University]. ProQuest Dissertations & Theses Global (7720372).
- Hancock, G. R. A., Cuthill, I. C., & Troscianko, J. (2025). Shining a light on camouflage evolution: Using genetic algorithms to determine the effects of geometry and lighting on optimal camouflage. *bioRxiv*. <https://doi.org/10.1101/2025.03.01.640995>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Heron, M., Hanson, V. L., & Ricketts, I. (2013). Open source and accessibility: Advantages and limitations. *Journal of Interaction Science*, *1*(1), 2. <https://doi.org/10.1186/2194-0827-1-2>
- Herbst, R., Stille, D., Gwilliam, G. F., III, et al. (2025). Unlocking the past: The potential of large language models to revolutionize transcription of natural history collections. *Data Intelligence*, *7*, 237–264. <https://doi.org/10.3724/2096-7004.di.2025.0011>
- Holmes, M. W., Hammond, T. T., Wogan, G. O. U., Walsh, R. E., LaBarbera, K., Wommack, E. A., Martins, F. M., Crawford, J. C., Mack, K. L., Bloch, L. M., & Nachman, M. W. (2016). Natural history collections as windows on evolutionary processes. *Molecular Ecology*, *25*(4), <https://doi.org/10.1111/mec.13529>
- Holovachov, O., Zatushevsky, A., & Shydlovsky, I. (2014). Whole-Drawer Imaging of Entomological Collections: Benefits, Limitations and Alternative Applications. *Journal of Conservation and Museum Studies*, *12*(1). <https://doi.org/10.5334/jcms.1021218>
- Howard, J., & Gugger, S. (2020). fastai: A Layered API for Deep Learning. *Information*, *11*(2), 108. <https://doi.org/10.3390/info11020108>
- Hudson, L. N., Blagoderov, V., Heaton, A., Holtzhausen, P., Livermore, L., Price, B. W., Walt, S. van der, & Smith, V. S. (2015). Inselect: Automating the Digitization of Natural History Collections. *PLOS ONE*, *10*(11), e0143402. <https://doi.org/10.1371/journal.pone.0143402>
- Lister, A. M. (2011). Natural history collections as sources of long-term datasets. *Trends in Ecology & Evolution*, *26*(4), 153–154. <https://doi.org/10.1016/j.tree.2010.12.009>
- Losey, J. E., & Vaughan, M. (2006). The Economic Value of Ecological Services Provided by Insects. *BioScience*, *56*(4), 311–323. [https://doi.org/10.1641/0006-3568\(2006\)56\[311:TEVOES\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2006)56[311:TEVOES]2.0.CO;2)
- Lürig, M. D. (2022). phenotype: A phenotyping pipeline for Python. *Methods in Ecology and Evolution*, *13*(3), 569–576. <https://doi.org/10.1111/2041-210X.13771>
- Mantle, B., LaSalle, J., & Fisher, N. (2012). Whole-drawer imaging for digital management and curation of a large entomological collection. *ZooKeys*, *209*, 147–163. <https://doi.org/10.3897/zookeys.209.3169>
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2024). *A Comprehensive Overview of Large Language Models*

- (arXiv:2307.06435). arXiv. <https://doi.org/10.48550/arXiv.2307.06435>
- Nj, S., N, C., Ar, D. R., & Dj, G. (2023). Leveraging natural history collections to understand the impacts of global change. *The Journal of Animal Ecology*, *92*(2), Article 2. <https://doi.org/10.1111/1365-2656.13882>
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019, December 3). *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. arXiv.Org. <https://arxiv.org/abs/1912.01703v1>
- Pearson, D. L., Knisley, C. B., & Kazilek, C. J. (2006). A field guide to the tiger beetles of the United States and Canada. Oxford University Press.
- Pearson, D. L., & Vogler, A. P. (2001). *Tiger Beetles: The Evolution, Ecology, and Diversity of the Cicindelids*. Cornell University Press.
- Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in Ecology and Evolution*, *14*(4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>
- Postema, E. G., Lippey, M. K., & Armstrong-Ingram, T. (2023). Color under pressure: How multiple factors shape defensive coloration. *Behavioral Ecology*, *34*(1), 1–13. <https://doi.org/10.1093/beheco/arac056>
- Stenhouse, A., Fisher, N., Lepschi, B., Schmidt-Lebuhn, A., Rodriguez, J., Turco, F., Reeson, A., Paris, C., & Thrall, P. H. (2025). A vision of human–AI collaboration for enhanced biological collection curation and research. *BioScience*, biaf021. <https://doi.org/10.1093/biosci/biaf021>
- Schultz, T. D., & Bernard, G. D. (1989). Pointillistic mixing of interference colours in cryptic tiger beetles. *Nature*, *337*(6202), 72–73. <https://doi.org/10.1038/337072a0>
- Schultz, T. D., & Hadley, N. F. (1987). Structural Colors of Tiger Beetles and Their Role in Heat Transfer through the Integument. *Physiological Zoology*, *60*(6), 737–745. <https://doi.org/10.1086/physzool.60.6.30159990>
- Scudder, G. G. E. (2017). The Importance of Insects. In *Insect Biodiversity* (pp. 9–43). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118945568.ch2>
- Spiesman, B. J., Gratton, C., Hatfield, R. G., Hsu, W. H., Jepsen, S., McCornack, B., Patel, K., & Wang, G. (2021). Assessing the potential for deep learning and computer vision to identify bumble bee species from images. *Scientific Reports*, *11*(1), 7580. <https://doi.org/10.1038/s41598-021-87210-1>
- Steinke, D., McKeown, J. T. A., Zyba, A., McLeod, J., Feng, C., & Hebert, P. D. N. (2024). Low-cost, high-volume imaging for entomological digitization. *ZooKeys*, *1206*, 315–326. <https://doi.org/10.3897/zookeys.1206.123670>
- Sun, J., Futahashi, R., & Yamanaka, T. (2021). Improving the Accuracy of Species Identification by Combining Deep Learning With Field Occurrence Records. *Frontiers in Ecology and Evolution*, *9*. <https://doi.org/10.3389/fevo.2021.762173>

- Sweeney, P. W., Starly, B., Morris, P. J., Xu, Y., Jones, A., Radhakrishnan, S., Grassa, C. J., & Davis, C. C. (2018). Large-scale digitization of herbarium specimens: Development and usage of an automated, high-throughput conveyor system. *TAXON*, *67*(1), Article 1. <https://doi.org/10.12705/671.10>
- Sys, S., Weißbach, S., Jakob, L., Gerber, S., & Schneider, C. (2022). CollemboAI, a macrophotography and computer vision workflow to digitize and characterize samples of soil invertebrate communities preserved in fluid. *Methods in Ecology and Evolution*, *13*(12), 2729–2742. <https://doi.org/10.1111/2041-210X.14001>
- Trautner, J., & Geigenmüller, K. (1987). Tiger beetles ground beetles: Illustrated key to the Cicindelidae and Carabidae of Europe. Margraf.
- Truong, M.-X. A., & Van der Wal, R. (2024). Exploring the landscape of automated species identification apps: Development, promise, and user appraisal. *Bioscience*, *74*(9), 601–613. <https://doi.org/10.1093/biosci/biae077>
- Van Belleghem, S. M., Papa, R., Ortiz-Zuazaga, H., Hendrickx, F., Jiggins, C. D., Owen McMillan, W., & Counterman, B. A. (2018). patternize: An R package for quantifying colour pattern variation. *Methods in Ecology and Evolution*, *9*(2), 390–398. <https://doi.org/10.1111/2041-210X.12853>
- van den Berg, C. P., Condon, N. D., Conradsen, C., White, T. E., & Cheney, K. L. (2024). Automated workflows using Quantitative Colour Pattern Analysis (QCPA): A guide to batch processing and downstream data analysis. *Evolutionary Ecology*, *38*, 387–397. <https://doi.org/10.1007/s10682-024-10291-7>
- van den Berg, C. P., Troscianko, J., Endler, J. A., Marshall, N. J., & Cheney, K. L. (2020). Quantitative Colour Pattern Analysis (QCPA): A comprehensive framework for the analysis of colour patterns in nature. *Methods in Ecology and Evolution*, *11*(2), 316–332. <https://doi.org/10.1111/2041-210X.13328>
- Wagner, D. L. (2020). Insect Declines in the Anthropocene. *Annual Review of Entomology*, *65*, 457–480. <https://doi.org/10.1146/annurev-ento-011019-025151>
- Watanabe, A. (2018). How many landmarks are enough to characterize shape and size variation? *PLOS ONE*, *13*(6), e0198341. <https://doi.org/10.1371/journal.pone.0198341>
- Weaver, W. N., Ng, J., & Laport, R. G. (2020). LeafMachine: Using machine learning to automate leaf trait extraction from digitized herbarium specimens. *Applications in Plant Sciences*, *8*(6), Article 6. <https://doi.org/10.1002/aps3.11367>
- Weaver, W. N., & Smith, S. A. (2023). From leaves to labels: Building modular machine learning networks for rapid herbarium specimen analysis with LeafMachine2. *Applications in Plant Sciences*, *11*(5), Article 5. <https://doi.org/10.1002/aps3.11548>
- Weeks, B. C., Zhou, Z., O'Brien, B. K., Darling, R., Dean, M., Dias, T., Hassena, G., Zhang, M., & Fouhey, D. F. (2023). A deep neural network for high-throughput measurement of functional traits on museum skeletal specimens. *Methods in Ecology and Evolution*, *14*(2), 347–359. <https://doi.org/10.1111/2041-210X.13864>
- Weinstein, B. G. (2018). A computer vision for animal ecology. *Journal of Animal Ecology*,

- 87(3), 533–545. <https://doi.org/10.1111/1365-2656.12780>
- Welch, K. D., & Lundgren, J. G. (2024). Introducing BugBox: A Platform for AI-Assisted Bioinventories of Arthropods. *American Entomologist*, 70(3), 31–33. <https://doi.org/10.1093/ae/tmae047>
- Weller, H. I., Hiller, A. E., Lord, N. P., & Van Belleghem, S. M. (2024). recolorize: An R package for flexible colour segmentation of biological images. *Ecology Letters*, 27(2), e14378. <https://doi.org/10.1111/ele.14378>
- Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLOS ONE*, 7(1), e29715. <https://doi.org/10.1371/journal.pone.0029715>
- Wiken, E., Jiménez Nava, F., & Griffith, G. (2011). North American terrestrial ecoregions — Level III. Commission for Environmental Cooperation.
- Wilson, R. J., de Siqueira, A. F., Brooks, S. J., Price, B. W., Simon, L. M., van der Walt, S. J., & Fenberg, P. B. (2023). Applying computer vision to digitised natural history collections for climate change research: Temperature-size responses in British butterflies. *Methods in Ecology and Evolution*, 14(2), 372–384. <https://doi.org/10.1111/2041-210X.13844>
- Wu, S., Chang, C.-M., Mai, G.-S., Rubenstein, D. R., Yang, C.-M., Huang, Y.-T., Lin, H.-H., Shih, L.-C., Chen, S.-W., & Shen, S.-F. (2019). Artificial intelligence reveals environmental constraints on colour diversity in insects. *Nature Communications*, 10(1), Article 1. <https://doi.org/10.1038/s41467-019-12500-2>
- Wührl, L., Pylatiuk, C., Giersch, M., Lapp, F., von Rintelen, T., Balke, M., Schmidt, S., Cerretti, P., & Meier, R. (2022). DiversityScanner: Robotic handling of small invertebrates with machine learning methods. *Molecular Ecology Resources*, 22(4), 1626–1638. <https://doi.org/10.1111/1755-0998.13567>
- Yamamoto, N., & Sota, T. (2020). Evolutionary fine-tuning of background-matching camouflage among geographical populations in the sandy beach tiger beetle. *Proceedings of the Royal Society B: Biological Sciences*, 287(1941), 20202315. <https://doi.org/10.1098/rspb.2020.2315>
- Yang, L. H., Postema, E. G., Hayes, T. E., Lippey, M. K., & MacArthur-Waltz, D. J. (2021). The complexity of global change and its effects on insects. *Current Opinion in Insect Science*, 47, 90–102. <https://doi.org/10.1016/j.cois.2021.05.001>
- Zelditch, M., Swiderski, D., & Sheets, H. D. (2012). *Geometric Morphometrics for Biologists: A Primer*. Academic Press.
- Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57(4), 99. <https://doi.org/10.1007/s10462-024-10721-6>