

# Novelty Assessment Report

**Paper:** Automatic Image-Level Morphological Trait Annotation for Organismal Images

**PDF URL:** <https://openreview.net/pdf?id=oFRbiaib5Q>

**Venue:** ICLR 2026 Conference Submission

**Year:** 2026

**Report Generated:** 2025-12-29

## Abstract

Morphological traits are physical characteristics of biological organisms that provide vital clues on how organisms interact with their environment. Yet extracting these traits remains a slow, expert-driven process, limiting their use in large-scale ecological studies. A major bottleneck is the absence of high-quality datasets linking biological images to trait-level annotations. In this work, we demonstrate that sparse autoencoders trained on foundation-model features yield monosemantic, spatially grounded neurons that consistently activate on meaningful morphological parts. Leveraging this property, we introduce a trait annotation pipeline that localizes salient regions and uses vision-language prompting to generate interpretable trait descriptions. Using this approach, we construct Bioscan-Traits, a dataset of 80K trait annotations spanning 19K insect images from BIOSCAN-5M. Human evaluation confirms the biological plausibility of the generated morphological descriptions. When used to fine-tune BioCLIP, a biologically grounded vision-language model, Bioscan-Traits improves zero-shot species classification on the in-the-wild Insects benchmark, underscoring the value of trait-level supervision for enhancing model generalization.

### Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

## Core Task Landscape

This paper addresses: **Automatic Morphological Trait Annotation from Biological Images**

A total of **50 papers** were analyzed and organized into a taxonomy with **14 categories**.

### Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Deep Learning-Based Trait Extraction and Annotation**
- **Geometric Morphometrics and Shape Analysis**
- **Software Toolkits and Platforms**
- **Imaging Modalities and Preprocessing**
- **Domain-Specific Applications**
- **Methodological Foundations and Reviews**

### Complete Taxonomy Tree

- Automatic Morphological Trait Annotation from Biological Images Survey Taxonomy
- Deep Learning-Based Trait Extraction and Annotation
  - Foundation Model and Vision-Language Approaches ★ (2 papers)
  - [0] Automatic Image-Level Morphological Trait Annotation for Organismal Images (Anon et al., 2026) [View paper](#)
  - [30] CellFlow: Simulating Cellular Morphology Changes via Flow Matching (Zhang Yu-hui, 2025) [View paper](#)
  - Supervised Segmentation and Classification (8 papers)
  - [3] Prediction of black soldier fly larval sex and morphological traits using computer vision and deep learning (Sarah Nawoya, 2025) [View paper](#)
  - [6] A deep learning pipeline for morphological and viability assessment of 3D cancer cell spheroids (Ajay K. Mali, 2025) [View paper](#)
  - [7] Unlocking the power of AI for phenotyping fruit morphology in Arabidopsis (Kieran Atkins, 2025) [View paper](#)
  - [12] Machine learning identification of Pseudomonas aeruginosa strains from colony image data (Jennifer B. Rattray, 2023) [View paper](#)
  - [20] Field-robotic leaf angle detection and characterization of maize plants using stereo vision and deep convolutional neural networks (Lirong Xiang, 2023) [View paper](#)
  - [29] Tellu: an object-detector algorithm for automatic classification of intestinal organoids (Eva Domínguez-Moreno, 2023) [View paper](#)
  - [41] Toward automatic phenotyping of developing embryos from videos (F. Ning, 2005) [View paper](#)
  - [49] Benchmarking and Automating the Image Recognition Capability of an In Situ Plankton Imaging System (Kevin T. Le, 2022) [View paper](#)
  - Interactive and Semi-Supervised Learning (3 papers)
  - [1] Stomata morphology measurement with interactive machine learning: accuracy, speed, and biological relevance? (T. Wacker, 2025) [View paper](#)
  - [18] Ilastik: interactive machine learning for (bio) image analysis (Stuart Berg, 2019) [View paper](#)
  - [27] CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging (Michael Held, 2010) [View paper](#)
  - Generative Modeling for Morphology Synthesis (2 papers)
  - [16] Adaptive Elastic GAN for High-Fidelity Blood Cell Image Hallucination and Classification (Issac Neha Margret, 2025) [View paper](#)
  - [45] MorphoDiff: Cellular Morphology Painting with Diffusion Models (Zeinab Navidi, 2024) [View paper](#)
- Geometric Morphometrics and Shape Analysis
  - Automated Landmark Detection and Shape Correspondence (3 papers)

- [8] Automated morphological phenotyping using learned shape descriptors and functional maps: A novel approach to geometric morphometrics (Oshane O. Thomas, 2023) [View paper](#)
- [10] 1/4Match: 3D Shape Correspondence for Biological Image Data (James Klatzow, 2022) [View paper](#)
- [25] ML4morph: A fast, accurate and general approach for automated detection and landmarking of biological structures in images (Arthur Porto, 2020) [View paper](#)
- Graph and Topological Representations (2 papers)
- [35] Shape-to-graph mapping method for efficient characterization and classification of complex geometries in biological images (William Pilcher, 2020) [View paper](#)
- [38] Image-based Morphological Characterization of Filamentous Biological Structures with Non-constant Curvature Shape Feature (Jie Fan, 2025) [View paper](#)
- Software Toolkits and Platforms (4 papers)
  - [2] PhenoLearn: a user-friendly toolkit for image annotation and deep learning-based phenotyping for biological datasets (Yichen He, 2025) [View paper](#)
  - [19] V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets (Hanchuan Peng, 2010) [View paper](#)
  - [22] Quantitative Image Analysis of Tissue Properties: A MATLAB Tool for Measuring Morphology and Co-localization in 2D Images (Shira Landau, 2024) [View paper](#)
  - [28] Pattern recognition software and techniques for biological image analysis (Lior Shamir, 2010) [View paper](#)
- Imaging Modalities and Preprocessing
  - Advanced Microscopy and Tomography (4 papers)
    - [5] Polarization microwave-induced thermoacoustic imaging for quantitative characterization of deep biological tissue microstructures (Yujing Li, 2022) [View paper](#)
    - [13] Characterization of Deformation and Failure of Fomes fomentarius Hymenium by Synchrotron Phase-Contrast Enhanced Micro-Computed Tomography, In situ Compression Testing and Three-Dimensional Optical Flow (Sophie Klemm, 2025) [View paper](#)
    - [32] Nanoscale characterization of the biomolecular corona by cryo-electron microscopy, cryo-electron tomography, and image simulation (S. Sheibani, 2021) [View paper](#)
    - [39] Non-destructive Characterization of Skeletal Muscle Extracellular Matrix Morphology by Combining Optical Coherence Tomography (OCT) Imaging with Tissue Clearing (M. Maillat, 2023) [View paper](#)
  - Specimen Preparation and Image Enhancement (4 papers)
    - [9] Specimen preparation optimization for size and morphology characterization of nanocellulose by TEM (Laura C. E. da Silva, 2020) [View paper](#)
    - [11] Mathematical morphology for quantification in biological & medical image analysis (Carl J Nelson, 2017) [View paper](#)
    - [17] Oil-immersion flow imaging microscopy for quantification and morphological characterization of submicron particles in biopharmaceuticals (Nils Krause, 2021) [View paper](#)
    - [40] Accurate Morphology Characterization Using Atomic Force Microscopy via Vertical Drift Correction and Illusory Slope Elimination (Yinan Wu, 2021) [View paper](#)
- Domain-Specific Applications
  - Plant Phenotyping (1 papers)
    - [14] Machine learning techniques to characterize functional traits of plankton from image data (Orenstein, 2022) [View paper](#)
  - Medical and Pathological Imaging (7 papers)
    - [15] Image Analysis in Histopathology and Cytopathology: From Early Days to Current Perspectives (Tibor Mezei, 2024) [View paper](#)
    - [24] In vitro to in vivo extrapolation and high-content imaging for simultaneous characterization of chemically induced liver steatosis and markers of hepatotoxicity (F. Moller, 2023) [View paper](#)
    - [31] Abstract 4141107: Genome-wide association study of aortic stiffness derived from deep learning in CMR images of 45,789 individuals identifies loci linked with cell-matrix structure (Nikhil Paliwal, 2024) [View paper](#)
    - [34] Optimizing morphology through blood cell image analysis (Anna Merino, 2018) [View paper](#)
    - [36] Label-Free Cell Detection of Acute Leukemia Using High-Content Morphological Profiling in Flow (Yoko Kawamura, 2023) [View paper](#)
    - [47] Morphological characterization of colorectal cancers in The Cancer Genome Atlas reveals distinct morphology-molecular associations: clinical and biological implications (J. Shia, 2017) [View paper](#)
    - [50] Three-dimensional morphological characterization of colorectal pits from label-free microscopy images. (L. Erbes, 2022) [View paper](#)
  - Cellular and Subcellular Morphology (4 papers)
    - [4] Morphological characterization of antibiotic combinations (MA Coram, 2021) [View paper](#)
    - [21] Cellular and morphological characterization of blastoderms from freshly laid broiler eggs (N. Pokhrel, 2017) [View paper](#)
    - [44] Activated sludge morphology characterization through an image analysis procedure (Y. G. Perez, 2006) [View paper](#)
    - [46] Automated on-chip rapid microscopy, phenotyping and sorting of *C. elegans* (Kwanghun Chung, 2008) [View paper](#)
  - Organismal and Ecological Phenotyping (3 papers)
    - [33] Morphological Characterization of the Antenna and Scent Patch of Three Danaus Species (Papilionoidea: Nymphalidae, Danainae) (Yaqi Yang, 2024) [View paper](#)
    - [42] Computer vision, machine learning, and the promise of phenomics in ecology and evolutionary biology (Moritz Lorig, 2021) [View paper](#)
    - [48] Imaging morphogenesis: technological advances and biological insights (Philipp J. Keller, 2013) [View paper](#)
- Methodological Foundations and Reviews (4 papers)
  - [23] Two-point Stokes vector diagnostic approach for characterization of optically anisotropic biological tissues (Motahareh Peyvasteh, 2020) [View paper](#)
  - [26] An Investigation of Spectroscopic Characterization on Biological Tissue (Mirza Muntasir Nishat, 2018) [View paper](#)
  - [37] Mathematical morphology and its applications to image and signal processing (John Goutsias, 2012) [View paper](#)
  - [43] Computational imaging in cell biology (Roland Eils, 2003) [View paper](#)

## Narrative

Core task: automatic morphological trait annotation from biological images. The field encompasses a diverse set of approaches organized around six main branches. Deep Learning-Based Trait Extraction and Annotation includes both specialized architectures for segmentation and measurement (e.g., Stomata Morphology Measurement[1], Maize Leaf Angle Detection[20]) and emerging foundation model or vision-language strategies that leverage large-scale pretraining. Geometric Morphometrics and Shape Analysis focuses on

landmark-based methods and shape correspondence techniques (Automated Geometric Morphometrics[8], 3D Shape Correspondence[10]), often applied to evolutionary or comparative studies. Software Toolkits and Platforms provide interactive or modular frameworks (Ilastik Interactive Learning[18], PhenoLearn Toolkit[2]) that enable biologists to annotate and analyze images without deep technical expertise. Imaging Modalities and Preprocessing addresses the acquisition and preparation of data from microscopy, flow imaging, or advanced optical systems (Flow Imaging Microscopy[17], Polarization Thermoacoustic Imaging[5]). Domain-Specific Applications target particular organisms or tissues—ranging from plankton (Plankton Trait Characterization[14]) to cancer spheroids (Cancer Spheroid Assessment[6]) and insect morphology (Black Soldier Fly Prediction[3])—while Methodological Foundations and Reviews consolidate theoretical perspectives and survey the state of the art (Phenomics Promise[42], Computational Cell Biology[43]).

Several active lines of work highlight contrasting emphases and open questions. Traditional geometric morphometrics remains valuable for hypothesis-driven shape analysis, yet deep learning methods increasingly dominate when large annotated datasets are available, raising questions about interpretability and generalization across imaging conditions. Meanwhile, foundation model and vision-language approaches—exemplified by Automatic Morphological Trait Annotation[0] and its neighbor CellFlow Morphology[30]—seek to unify diverse annotation tasks under a single pretrained framework, potentially reducing the need for task-specific labeled data. These works differ in scope: CellFlow Morphology[30] emphasizes cellular-scale phenotyping with flow-based representations, whereas Automatic Morphological Trait Annotation[0] targets broader biological imaging contexts by integrating vision-language alignment. Both contrast with more narrowly scoped deep learning pipelines (e.g., Fruit Phenotyping AI[7]) that optimize for a single organism or trait. The central trade-off lies between the flexibility and data efficiency of foundation models versus the precision and domain knowledge embedded in specialized tools, a tension that continues to shape the trajectory of automated morphological analysis.

---

## Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

### 1. CellFlow: Simulating Cellular Morphology Changes via Flow Matching

**Authors:** Zhang Yu-hui, Su YuChang, Yuhui Zhang, Wang Chen-yu, Yuchang Su, et al. (21 authors total) | **Year/Venue:** 2025 • International Conference on Machine Learning | **URL:** [View paper](#)

#### Abstract

Building a virtual cell capable of accurately simulating cellular behaviors in silico has long been a dream in computational biology. We introduce CellFlux, an image-generative model that simulates cellular morphology changes induced by chemical and genetic perturbations using flow matching. Unlike prior methods, CellFlux models distribution-wise transformations from unperturbed to perturbed cell states, effectively distinguishing actual perturbation effects from experimental artifacts such as b...

#### Relationship Analysis

Both papers belong to the Foundation Model and Vision-Language Approaches category, leveraging pre-trained models for biological image analysis. While the original paper uses sparse autoencoders on foundation model features to extract and annotate morphological traits from static organismal images via vision-language prompting, CellFlow employs flow matching to simulate dynamic cellular morphology changes induced by perturbations. The key difference is that the original paper focuses on trait discovery and annotation from existing images, whereas CellFlow generates synthetic cell images to model perturbation-induced transformations.

---

## Contributions Analysis

**Overall novelty summary.** The paper introduces a trait annotation pipeline combining sparse autoencoders on foundation-model features with vision-language prompting to generate morphological descriptions from insect images. It resides in the Foundation Model and Vision-Language Approaches leaf, which contains only two papers within the broader Deep Learning-Based Trait Extraction and Annotation branch. This leaf represents an emerging research direction, contrasting with the more populated Supervised Segmentation and Classification leaf (eight papers) that focuses on task-specific architectures. The sparse population suggests the work enters a relatively nascent area where foundation models are being adapted for biological trait discovery.

The taxonomy reveals that neighboring leaves pursue distinct strategies: Supervised Segmentation and Classification emphasizes domain-tailored networks trained on annotated datasets, while Interactive and Semi-Supervised Learning (three papers) enables non-expert annotation through corrective feedback. The paper's approach diverges by leveraging pretrained representations to bypass extensive manual labeling, aligning more closely with the vision-language paradigm than with classical supervised pipelines. Its sibling paper in the same leaf, CellFlow Morphology, targets cellular-scale phenotyping with flow-based representations, whereas this work addresses organism-level insect morphology, indicating complementary scopes within the foundation model category.

Among the twenty candidates examined across three contributions, none were flagged as clearly refuting the proposed methods. The trait annotation pipeline (ten candidates examined, zero refutable) and the BIOSCAN-TRAITS dataset (ten candidates examined, zero refutable) both appear to lack direct prior work within this limited search scope. The species-contrastive ranking method was not evaluated against any candidates. This absence of overlapping prior work, combined with the sparse leaf population, suggests the contributions occupy a relatively unexplored intersection of sparse autoencoders, vision-language models, and morphological trait extraction, though the search examined only top-twenty semantic matches rather than an exhaustive survey.

Given the limited search scope and the nascent state of the Foundation Model and Vision-Language Approaches leaf, the work appears to introduce novel technical components—particularly the use of sparse autoencoders for monosemantic neuron discovery in biological imaging—that have not been directly addressed in the examined candidates. However, the analysis reflects top-twenty semantic matches and does not cover the full breadth of foundation model or vision-language research outside this specific biological context. The dataset contribution also appears distinct within the examined scope, though broader ecological or entomological datasets may exist beyond the search perimeter.

---

This paper presents **3 main contributions**, each analyzed against relevant prior work:

### Contribution 1: Trait annotation pipeline using sparse autoencoders and vision-language prompting

**Description:** The authors propose a three-step pipeline that uses sparse autoencoders trained on foundation-model features to identify monosemantic, spatially grounded neurons corresponding to morphological parts, then localizes these regions and prompts a multimodal language model to generate trait descriptions. This approach addresses the bottleneck of extracting morphological traits from biological images without requiring manual expert annotation.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

#### 1. Sparse attention vectors: Generative multimodal model features are discriminative vision-language classifiers

**URL:** [View paper](#)

#### Brief Assessment

Sparse Attention Vectors[64] focuses on extracting discriminative features from generative multimodal models for vision-language classification tasks, not on trait annotation pipelines for biological images using sparse autoencoders to localize morphological parts.

---

## 2. VL-SAE: Interpreting and Enhancing Vision-Language Alignment with a Unified Concept Set

URL: [View paper](#)

### Brief Assessment

VL-SAE Vision-Language[63] focuses on interpreting vision-language alignment in VLMs through sparse autoencoders that map multimodal representations to unified concepts. The original paper uses SAEs for morphological trait extraction from biological images via spatial localization and MLLM prompting, which is a distinct application domain and methodology.

---

## 3. Interpreting CLIP with Hierarchical Sparse Autoencoders

URL: [View paper](#)

### Brief Assessment

CLIP Hierarchical Autoencoders[62] focuses on interpreting CLIP vision-language models through hierarchical sparse autoencoders for concept extraction and similarity search, not on morphological trait annotation from biological images. The candidate addresses multimodal representation interpretability in general computer vision, while the original work specifically targets biological trait extraction from organismal specimens.

---

## 4. SAUCE: Selective Concept Unlearning in Vision-Language Models with Sparse Autoencoders

URL: [View paper](#)

### Brief Assessment

SAUCE Concept Unlearning[69] focuses on concept unlearning in vision-language models using sparse autoencoders to suppress specific concepts, not on trait annotation or morphological feature extraction from biological images.

---

## 5. Sparse Autoencoders Learn Monosemantic Features in Vision-Language Models

URL: [View paper](#)

### Brief Assessment

Monosemantic Features SAE[66] focuses on evaluating monosemanticity in vision-language models (e.g., CLIP) and steering multimodal LLMs, not on morphological trait extraction from biological images. The candidate does not address trait annotation pipelines for organismal biology.

---

## 6. Steering LVLMs via Sparse Autoencoder for Hallucination Mitigation

URL: [View paper](#)

### Brief Assessment

Steering LVLMs SAE[68] uses sparse autoencoders to identify semantic directions for hallucination mitigation in vision-language models, not for morphological trait extraction from biological images. The candidate focuses on steering model behavior through latent directions, while the original develops a pipeline for automated trait annotation in organismal biology.

---

## 7. debiaSAE: Benchmarking and Mitigating Vision-Language Model Bias

URL: [View paper](#)

### Brief Assessment

debiaSAE Bias Mitigation[70] applies sparse autoencoders to debias vision-language models for demographic fairness, not for morphological trait extraction from biological images. The technical application domains are fundamentally different.

---

## 8. Patch-level phenotype identification via weakly supervised neuron selection in sparse autoencoders for CLIP-derived pathology embeddings

URL: [View paper](#)

### Brief Assessment

Patch-level Phenotype Identification[65] focuses on pathology image analysis for tumor/normal tissue classification using sparse autoencoders on CLIP embeddings, not on morphological trait extraction from biological organism images using vision-language prompting pipelines.

---

## 9. A novel multimodal framework for automatic recognition of individual cattle based on hybrid features using sparse stacked denoising autoencoder and group sparse $\hat{\alpha}$

URL: [View paper](#)

### Brief Assessment

Cattle Recognition Sparse[67] focuses on cattle identification using sparse stacked denoising autoencoders for feature extraction in a biometric recognition system, not on morphological trait annotation or vision-language model prompting for biological specimens.

---

## 10. Avltrack: Dynamic sparse learning for aerial vision-language tracking

URL: [View paper](#)

### Brief Assessment

Avltrack Aerial Tracking[61] focuses on aerial vision-language tracking for UAV scenarios using dynamic sparse attention for multi-modal alignment, not on morphological trait extraction from biological images using sparse autoencoders trained on foundation-model features.

---

## Contribution 2: BIOSCAN-TRAITS dataset

**Description:** The authors create a large-scale dataset containing 80,000 morphological trait annotations across 19,000 insect images by applying their pipeline to the BIOSCAN-5M corpus. This dataset provides structured, interpretable trait-level supervision at scale for training and evaluating biological foundation models.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

## 1. Zero-shot insect detection via weak language supervision

URL: [View paper](#)

### Brief Assessment

Zero-shot Insect Detection[56] focuses on zero-shot object detection using vision-language models for insect localization in images, not on creating morphological trait annotations. The candidate curates an insect image dataset for detection benchmarking, whereas the original creates trait-level descriptions of morphological features.

---

## **2. Formalizing invertebrate morphological data: A descriptive model for cuticle-based skeleto-muscular systems, an ontology for insect anatomy, and their potential**

URL: [View paper](#)

### **Brief Assessment**

Invertebrate Morphological Data[52] focuses on formalizing morphological terminology and ontology development for insect anatomy, not on creating large-scale trait annotation datasets from images using automated pipelines.

---

## **3. A multi-modal dataset for insect biodiversity with imagery and DNA at the trap and individual level**

URL: [View paper](#)

### **Brief Assessment**

Insect Biodiversity Dataset[53] focuses on multi-modal data (DNA barcodes and bulk/individual images) for taxonomic classification and abundance estimation, not on morphological trait annotation. The datasets serve fundamentally different purposes: BIOSCAN-TRAITS provides trait-level descriptions for individual specimens, while this candidate provides segmentation masks and taxonomic labels for bulk samples.

---

## **4. Worldwide revision of synanthropic silverfish (Insecta: Zygentoma: Lepismatidae) combining morphological and molecular data**

URL: [View paper](#)

### **Brief Assessment**

Silverfish Revision[57] focuses on taxonomic revision of synanthropic silverfish using morphological and molecular data for species identification, not on creating large-scale morphological trait annotation datasets for insect images.

---

## **5. Utilizing CNNs for classification and uncertainty quantification for 15 families of European fly pollinators**

URL: [View paper](#)

### **Brief Assessment**

Fly Pollinator CNNs[54] focuses on CNN-based classification of fly families using morphological features, not on creating large-scale trait annotation datasets for insect images.

---

## **6. Identification of species by combining molecular and morphological data using convolutional neural networks**

URL: [View paper](#)

### **Brief Assessment**

Species Identification Multimodal[58] focuses on integrating morphological images with DNA barcoding data for species classification tasks, not on creating large-scale morphological trait annotation datasets. The candidate does not describe any dataset containing structured trait-level annotations.

---

## **7. Artificial intelligence correctly classifies developmental stages of monarch caterpillars enabling better conservation through the use of community science photographs**

URL: [View paper](#)

### **Brief Assessment**

Monarch Caterpillar Classification[51] focuses on developmental stage classification of monarch caterpillars using YOLOv5, not on creating large-scale morphological trait annotation datasets for diverse insect images. The candidate creates a dataset of annotated monarch caterpillar developmental stages (instars), while the original creates trait-level morphological descriptions across 19,000 diverse insect images from BIOSCAN-5M.

---

## **8. STARdbi: A pipeline and database for insect monitoring based on automated image analysis**

URL: [View paper](#)

### **Brief Assessment**

STARdbi Insect Monitoring[60] focuses on a pipeline for capturing, scanning, and storing insect trap images with machine learning for ecological monitoring, not on creating large-scale morphological trait annotation datasets for training biological foundation models.

---

## **9. MAPHIS: Measuring arthropod phenotypes using hierarchical image segmentations**

URL: [View paper](#)

### **Brief Assessment**

MAPHIS Arthropod Phenotypes[59] focuses on segmenting arthropod body parts and calculating morphological measures from images, but does not create a large-scale trait annotation dataset. The candidate is a tool for measuring traits, not a dataset of 80,000 trait annotations like BIOSCAN-TRAITS.

---

## **10. Classification and morphological analysis of vector mosquitoes using deep convolutional neural networks**

URL: [View paper](#)

### **Brief Assessment**

Mosquito Classification[55] focuses on building a mosquito image dataset for species classification using deep learning, not on creating large-scale morphological trait annotations across diverse insect taxa as in the original paper's BIOSCAN-TRAITS dataset.

---

## **Contribution 3: Species-contrastive ranking method for trait selection**

**Description:** The authors develop a ranking approach that identifies SAE units by comparing their activation strength within a focal species against closely related species (congeners). This method isolates taxonomically diagnostic features that correspond to the fine-scale morphological structures recorded by taxonomists as traits.

This contribution was assessed against **0 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

## **Appendix: Text Similarity Detection**

No high-similarity text segments were detected across any compared papers.

---

## References

---

- [0] Automatic Image-Level Morphological Trait Annotation for Organismal Images [View paper](#)
- [1] Stomata morphology measurement with interactive machine learning: accuracy, speed, and biological relevance? [View paper](#)
- [2] PhenoLearn: a user-friendly toolkit for image annotation and deep learning-based phenotyping for biological datasets [View paper](#)
- [3] Prediction of black soldier fly larval sex and morphological traits using computer vision and deep learning [View paper](#)
- [4] Morphological characterization of antibiotic combinations [View paper](#)
- [5] Polarization microwave-induced thermoacoustic imaging for quantitative characterization of deep biological tissue microstructures [View paper](#)
- [6] A deep learning pipeline for morphological and viability assessment of 3D cancer cell spheroids [View paper](#)
- [7] Unlocking the power of AI for phenotyping fruit morphology in Arabidopsis [View paper](#)
- [8] Automated morphological phenotyping using learned shape descriptors and functional maps: A novel approach to geometric morphometrics [View paper](#)
- [9] Specimen preparation optimization for size and morphology characterization of nanocellulose by TEM [View paper](#)
- [10] 1/4Match: 3D Shape Correspondence for Biological Image Data [View paper](#)
- [11] Mathematical morphology for quantification in biological & medical image analysis [View paper](#)
- [12] Machine learning identification of Pseudomonas aeruginosa strains from colony image data [View paper](#)
- [13] Characterization of Deformation and Failure of Fomes fomentarius Hymenium by Synchrotron Phase-Contrast Enhanced Micro-Computed Tomography, In situ Compression Testing and Three-Dimensional Optical Flow [View paper](#)
- [14] Machine learning techniques to characterize functional traits of plankton from image data [View paper](#)
- [15] Image Analysis in Histopathology and Cytopathology: From Early Days to Current Perspectives [View paper](#)
- [16] Adaptive Elastic GAN for High-Fidelity Blood Cell Image Hallucination and Classification [View paper](#)
- [17] Oil-immersion flow imaging microscopy for quantification and morphological characterization of submicron particles in biopharmaceuticals [View paper](#)
- [18] Ilastik: interactive machine learning for (bio) image analysis [View paper](#)
- [19] V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets [View paper](#)
- [20] Field-based robotic leaf angle detection and characterization of maize plants using stereo vision and deep convolutional neural networks [View paper](#)
- [21] Cellular and morphological characterization of blastoderms from freshly laid broiler eggs [View paper](#)
- [22] Quantitative Image Analysis of Tissue Properties: A MATLAB Tool for Measuring Morphology and Co-localization in 2D Images [View paper](#)
- [23] Two-point Stokes vector diagnostic approach for characterization of optically anisotropic biological tissues [View paper](#)
- [24] In vitro to in vivo extrapolation and high-content imaging for simultaneous characterization of chemically induced liver steatosis and markers of hepatotoxicity [View paper](#)
- [25] ML-morph: A fast, accurate and general approach for automated detection and landmarking of biological structures in images [View paper](#)
- [26] An Investigation of Spectroscopic Characterization on Biological Tissue [View paper](#)
- [27] CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging [View paper](#)
- [28] Pattern recognition software and techniques for biological image analysis [View paper](#)
- [29] Tellu an object-detector algorithm for automatic classification of intestinal organoids [View paper](#)
- [30] CellFlow: Simulating Cellular Morphology Changes via Flow Matching [View paper](#)
- [31] Abstract 4141107: Genome-wide association study of aortic stiffness derived from deep learning in CMR images of 45,789 individuals identifies loci linked with cell-matrix structure [View paper](#)
- [32] Nanoscale characterization of the biomolecular corona by cryo-electron microscopy, cryo-electron tomography, and image simulation [View paper](#)
- [33] Morphological Characterization of the Antenna and Scent Patch of Three Danaus Species (Papilionoidea: Nymphalidae, Danainae) [View paper](#)
- [34] Optimizing morphology through blood cell image analysis [View paper](#)
- [35] Shape-to-graph mapping method for efficient characterization and classification of complex geometries in biological images [View paper](#)
- [36] Label-Free Cell Detection of Acute Leukemia Using High-Content Morphological Profiling in Flow [View paper](#)
- [37] Mathematical morphology and its applications to image and signal processing [View paper](#)
- [38] Image-based Morphological Characterization of Filamentous Biological Structures with Non-constant Curvature Shape Feature [View paper](#)
- [39] Non-destructive Characterization of Skeletal Muscle Extracellular Matrix Morphology by Combining Optical Coherence Tomography (OCT) Imaging with Tissue Clearing [View paper](#)
- [40] Accurate Morphology Characterization Using Atomic Force Microscopy via Vertical Drift Correction and Illusory Slope Elimination [View paper](#)
- [41] Toward automatic phenotyping of developing embryos from videos [View paper](#)
- [42] Computer vision, machine learning, and the promise of phenomics in ecology and evolutionary biology [View paper](#)
- [43] Computational imaging in cell biology [View paper](#)
- [44] Activated sludge morphology characterization through an image analysis procedure [View paper](#)
- [45] MorphoDiff: Cellular Morphology Painting with Diffusion Models [View paper](#)
- [46] Automated on-chip rapid microscopy, phenotyping and sorting of C. elegans [View paper](#)
- [47] Morphological characterization of colorectal cancers in The Cancer Genome Atlas reveals distinct morphology-molecular associations: clinical and biological implications [View paper](#)
- [48] Imaging morphogenesis: technological advances and biological insights [View paper](#)
- [49] Benchmarking and Automating the Image Recognition Capability of an In Situ Plankton Imaging System [View paper](#)
- [50] Three-dimensional morphological characterization of colorectal pits from label-free microscopy images. [View paper](#)
- [51] Artificial intelligence correctly classifies developmental stages of monarch caterpillars enabling better conservation through the use of community science photographs [View paper](#)
- [52] Formalizing invertebrate morphological data: A descriptive model for cuticle-based skeleto-muscular systems, an ontology for insect anatomy, and their potential [View paper](#)
- [53] A multi-modal dataset for insect biodiversity with imagery and DNA at the trap and individual level [View paper](#)
- [54] Utilizing CNNs for classification and uncertainty quantification for 15 families of European fly pollinators [View paper](#)

- [55] Classification and morphological analysis of vector mosquitoes using deep convolutional neural networks [View paper](#)
- [56] Zero-shot insect detection via weak language supervision [View paper](#)
- [57] Worldwide revision of synanthropic silverfish (Insecta: Zygentoma: Lepismatidae) combining morphological and molecular data [View paper](#)
- [58] Identification of species by combining molecular and morphological data using convolutional neural networks [View paper](#)
- [59] MAPHIS: Measuring arthropod phenotypes using hierarchical image segmentations [View paper](#)
- [60] STARdbi: A pipeline and database for insect monitoring based on automated image analysis [View paper](#)
- [61] Avltrack: Dynamic sparse learning for aerial vision-language tracking [View paper](#)
- [62] Interpreting CLIP with Hierarchical Sparse Autoencoders [View paper](#)
- [63] VL-SAE: Interpreting and Enhancing Vision-Language Alignment with a Unified Concept Set [View paper](#)
- [64] Sparse attention vectors: Generative multimodal model features are discriminative vision-language classifiers [View paper](#)
- [65] Patch-level phenotype identification via weakly supervised neuron selection in sparse autoencoders for CLIP-derived pathology embeddings [View paper](#)
- [66] Sparse Autoencoders Learn Monosemantic Features in Vision-Language Models [View paper](#)
- [67] A novel multimodal framework for automatic recognition of individual cattle based on hybrid features using sparse stacked denoising autoencoder and group sparse [View paper](#)
- [68] Steering LVLMs via Sparse Autoencoder for Hallucination Mitigation [View paper](#)
- [69] SAUCE: Selective Concept Unlearning in Vision-Language Models with Sparse Autoencoders [View paper](#)
- [70] debiaSAE: Benchmarking and Mitigating Vision-Language Model Bias [View paper](#)