

Descriptor: *Parasitoid Wasps and Associated Hymenoptera Dataset (DAPWH)*

JOÃO MANOEL HERRERA PINHEIRO¹ (STUDENT MEMBER, IEEE), GABRIELA DO NASCIMENTO HERRERA², LUCIANA BUENO DOS REIS FERNANDES², ALVARO DORIA DOS SANTOS³, RICARDO V. GODOY¹, EDUARDO A. B. ALMEIDA⁴, HELENA CAROLINA ONODY⁵, MARCELO ANDRADE DA COSTA VIEIRA¹, ANGÉLICA MARIA PENTEADO-DIAS² AND MARCELO BECKER¹ (MEMBER, IEEE)

¹São Carlos School of Engineering, University of São Paulo, São Carlos 13566590, SP, Brazil

²Department of Ecology and Evolutionary Biology, Federal University of São Carlos, São Carlos 13565905, Brazil

³Federal University of Tocantins, Porto Nacional, 77500000, Brazil

⁴Department of Biology (FFCLRP), University of São Paulo, Ribeirão Preto 14040901, Brazil

⁵State University of Piauí, Deputado Jesualdo Cavalcanti Campus, Corrente, 49800000, Brazil

CORRESPONDING AUTHOR: João Manoel Herrera Pinheiro (e-mail: joao.manoel.pinheiro@usp.br), Gabriela do Nascimento Herrera (e-mail: gabriela.herrera@estudante.ufscar.br), Angélica Maria Pentado-Dias (e-mail: angelica@ufscar.br) and Marcelo Becker (e-mail: becker@sc.usp.br).

This article has been accepted for publication in IEEE Data Descriptions. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/IEEEDATA.2026.3683381

ABSTRACT Accurate taxonomic identification is the cornerstone of biodiversity monitoring and agricultural management, particularly for the hyper-diverse superfamily Ichneumonoidea. Comprising the families Ichneumonidae and Braconidae, these parasitoid wasps are ecologically critical for regulating insect populations, yet they remain one of the most taxonomically challenging groups due to their cryptic morphology and vast number of undescribed species. To address the scarcity of robust digital resources for these key groups, we present a curated image dataset designed to advance automated identification systems. The dataset contains 3,556 high-resolution images, primarily focused on Neotropical Ichneumonidae and Braconidae, while also including supplementary families such as Andrenidae, Apidae, Bethyridae, Chrysididae, Colletidae, Halictidae, Megachilidae, Pompilidae, and Vespidae to improve model robustness. Crucially, a subset of 1,739 images is annotated in COCO format, featuring multi-class bounding boxes for the full insect body, wing venation, and scale bars. This resource provides a foundation for developing computer vision models capable of identifying these families.

IEEE SOCIETY/COUNCIL IEEE Computational Intelligence Society (CIS)

DATA DOI/PID 10.5281/zenodo.18501018

DATA TYPE/LOCATION Images, Brazil

INDEX TERMS arthropod, biodiversity, convolutional neural network, taxonomy

BACKGROUND

The order Hymenoptera, which includes bees, wasps, and ants [1], is among the most speciose groups within Hexapoda [2], [3]. Although Coleoptera and Lepidoptera presently account for the greatest number of described species, these four orders are widely recognized as the most species-rich overall, with accumulating evidence suggesting that

Hymenoptera may in fact represent the most speciose order [3]–[5]. The bees are the most famous insect family in the Hymenoptera order, they are recognized for their essential pollination services [6], while the population status of many other Hymenopterans, such as ants, wasps, and parasitoids, which also provide important ecosystem services, largely remains unknown [7]. Many species within this order serve

as pollen vectors, contributing to the pollination of various plants and acting as effective pollinators [8], [9].

Parasitoid wasps are likely the most diverse group within the order Hymenoptera, potentially representing the largest group in terms of species numbers [10]–[14]. The Ichneumonidea superfamily, represents a vast group of parasitoid wasps within the order Hymenoptera, these insects play a critical role in regulating the populations of other insect species [15]–[17]. As parasitoid Hymenoptera, they primarily attack the larvae and pupae of Lepidoptera and Coleoptera, many of which are important agricultural pests [17]. Beyond pest regulation, they also act as pollinators and hold significant importance in maintaining ecological diversity [18], [19]. With over 45,000 validly described species, a significant number of species within these groups remain undescribed or are challenging to identify [1], [11], [14], [20], [21], highlighting a critical need for robust resources to aid their taxonomic.

Traditional insect identification relies on expert taxonomists using microscopes and morphological keys, a slow and highly specialized process in a time of global taxonomist shortage [22]. In response to the critical need for robust identification resources, we introduce a comprehensive dataset of 3,556 high-quality images centered on the parasitoid families Ichneumonidae and Braconidae. To ensure taxonomic robustness, the dataset includes supplementary images from diverse Hymenoptera groups, including Apidae, Vespidae, and Pompilidae, among others. A key contribution of this work is the inclusion of a fully annotated subset in COCO format; these annotations delineate the full body, wing structures, and scale bars, explicitly designed to train object detection architectures. Figure 1 shows the schematic representation of the dataset organization.

It offers a valuable resource for researchers across disciplines, including entomologists, ecologists, and computer vision scientists, enabling the training and validation of machine learning and deep learning models for automated insect identification and fine-grained classification, particularly within taxonomically complex groups like Ichneumonidea.

COLLECTION METHODS AND DESIGN

The dataset comprises high-resolution images of specimens from the superfamily Ichneumonoidea as well as nine additional hymenopteran families: Andrenidae, Apidae, Bethyidae, Chrysididae, Colletidae, Halictidae, Megachilidae, Pompilidae, and Vespidae. Specimens were obtained from four primary university affiliated repositories, the specimens curated in these repositories represent the culmination of years of research projects.

Originally sampled using Malaise traps during extensive field campaigns, the material was subsequently transferred to university laboratories. There, expert taxonomists identified the specimens using morphological dichotomous keys before permanently depositing them in the collections. All images are licensed under CC BY 4.0 or CC BY-NC 4.0.

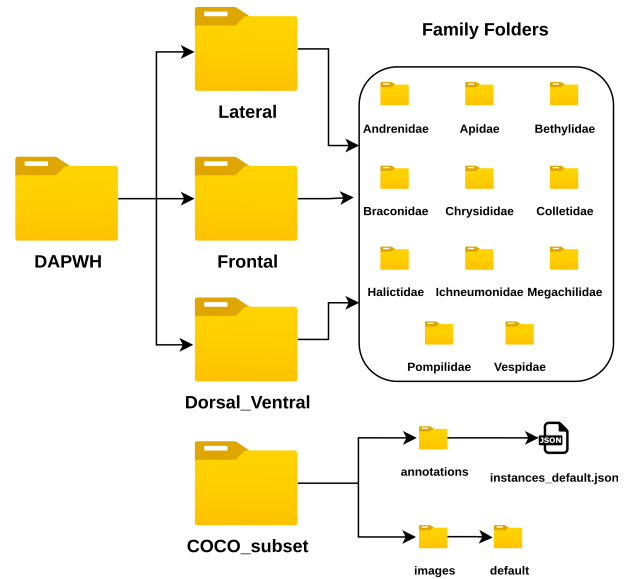


FIGURE 1. Schematic representation of the dataset file organization. The primary repository (DAPWH) [23] is structured hierarchically by anatomical view (*Lateral, Frontal, Dorsal_Ventral*), with each view containing specific families' subdirectories. Complementing this structure, the COCO_subset provides a subset version of the dataset, including the `instances_default.json` file and corresponding images.

Specimens representing the families Braconidae and Ichneumonidae were primarily acquired from the Coleção Taxonômica do Departamento de Ecologia e Biologia Evolutiva da UFSCar (DCBU) [24]. The process consisted of retrieving specimens from the collection, mounting them under a Leica M205C stereomicroscope equipped with a Leica K5C camera was used for image acquisition, with LAS X software controlling capture. Image stacking was performed using Helicon Focus software. Figure 2 illustrates the procedure adopted for specimen imaging.

Figure 3 maps the collection localities of samples. The spatial distribution highlights that the vast majority of specimens originate from the Neotropical region, specifically Brazil.

The Apoidea families (Andrenidae, Apidae, Colletidae, Halictidae, Megachilidae) and Chrysididae were primarily sourced from the Coleção Entomológica Prof. J.M.F. Camargo (RPSP) [25] provided by Prof. Dr. Eduardo A.B. Almeida, in addition to contributions from Museu de Zoologia da Universidade de São Paulo (MZUSP) and the Spencer Collection [26]. Finally, specimens for Bethyidae, Pompilidae, and Vespidae were acquired from MZUSP [27] and the Spencer Entomological Collection [26]. Figure 4 and Figure 5 shows example images from the dataset.

The taxonomic and photographic distribution of the DAPWH dataset is detailed in Table 1. The dataset comprises a total of 3,556 high-resolution images across 11 Hymenoptera families, organized into three primary morphological perspectives: dorsal/ventral, frontal, and lateral. The



FIGURE 2. Specimens were retrieved from DCBU collection, mounted under a Leica M205C stereomicroscope equipped with a Leica K5C camera.

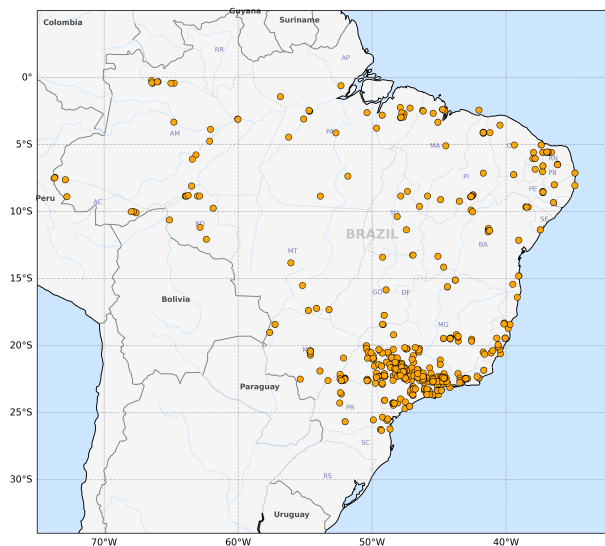


FIGURE 3. Illustrative geographic distribution of specimen sources from the DCBU, MZUSP, and RPSP collections, generated using Python, Pandas, and Cartopy.

images have an average resolution of 1892×1478 pixels, a standard deviation of 1085×858 , and a maximum resolution of 8050×5372 pixels.

The family Ichneumonidae represents the largest portion of the dataset with 786 images, followed closely by Braconidae with 648 images. These two families are of particular interest for automated identification of parasitoid wasps. From a viewpoint perspective, the Lateral view is the most prevalent, containing 1,616 images, providing a substantial foundation for profiling lateral morphological traits.

Annotation Process

To facilitate the training of detection models, a subset of 1,739 images was selected for detailed labeling and processed using the Computer Vision Annotation Tool (CVAT) [28] and the Segment Anything Model (SAM) [29]. The annotation protocol involved the precise delineation of segment masks and bounding boxes for three distinct classes: the full insect body, wings, and scale bar.

The initial annotations were performed collaboratively by computer vision specialists and taxonomists, followed by a rigorous final review by a lead taxonomist. Upon validation, the final bounding boxes and segmentation masks were generated, and the dataset was exported in the standard format Common Objects in Context (COCO) [30]. The complete pipeline for dataset construction and annotation is depicted in Figure 6.

The structural complexity of the dataset is further detailed by the annotation properties summarized in Table 2. Beyond image-level taxonomic labels, the dataset contains 1,794 high-precision segmentation masks for wings and 1,700 bounding boxes for scale bars. Within the COCO subset, Ichneumonidae and Braconidae remain the most densely annotated taxa, with 521 and 406 mask-level annotations, respectively.

TABLE 1. Distribution of images per family and view.

Family	Dorsal/Ventral	Frontal	Lateral	Total
Ichneumonidae	205	128	453	786
Braconidae	175	66	407	648
Apidae	153	155	158	466
Vespidae	230	87	143	460
Megachilidae	183	19	96	298
Chrysididae	137	31	76	244
Andrenidae	154	11	79	244
Pompilidae	57	48	85	190
Bethylidae	22	0	72	94
Halictidae	26	18	31	75
Colletidae	26	9	16	51
Total	1,368	572	1,616	3,556

VALIDATION AND QUALITY

To establish a baseline for the DAPWH dataset, we conducted comprehensive image-level classification and object detection benchmarks. Following established methodological conventions, we adopted a stratified split strategy to ensure taxonomic representation across all subsets. The dataset was partitioned into a training set (70%), with the remaining 30% equally distributed between validation and testing sets (15% each). While no single splitting ratio is universally prescriptive, this configuration was selected to balance model generalization against the bias–variance trade-off, considering the high-resolution nature of the imagery and computational constraints [31].



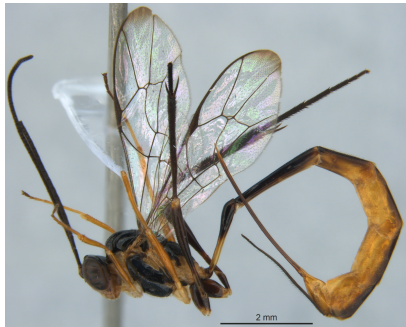
(a)



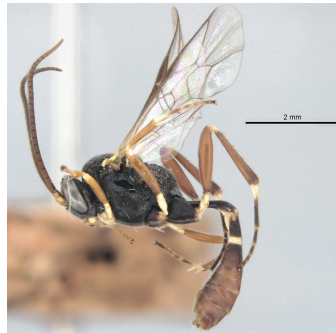
(b)



(c)



(d)



(e)



(f)



(g)



(h)



(i)



(j)



(k)



(m)

FIGURE 4. Examples of morphological variation in the DAPWH dataset. (a), (b), (c) Braconidae; (d), (e), (f) Ichneumonidae; (g), (h) Apidae; (i) Vespidae; (j) Bethyidae; (k) Pompilidae; (m) Colletidae.

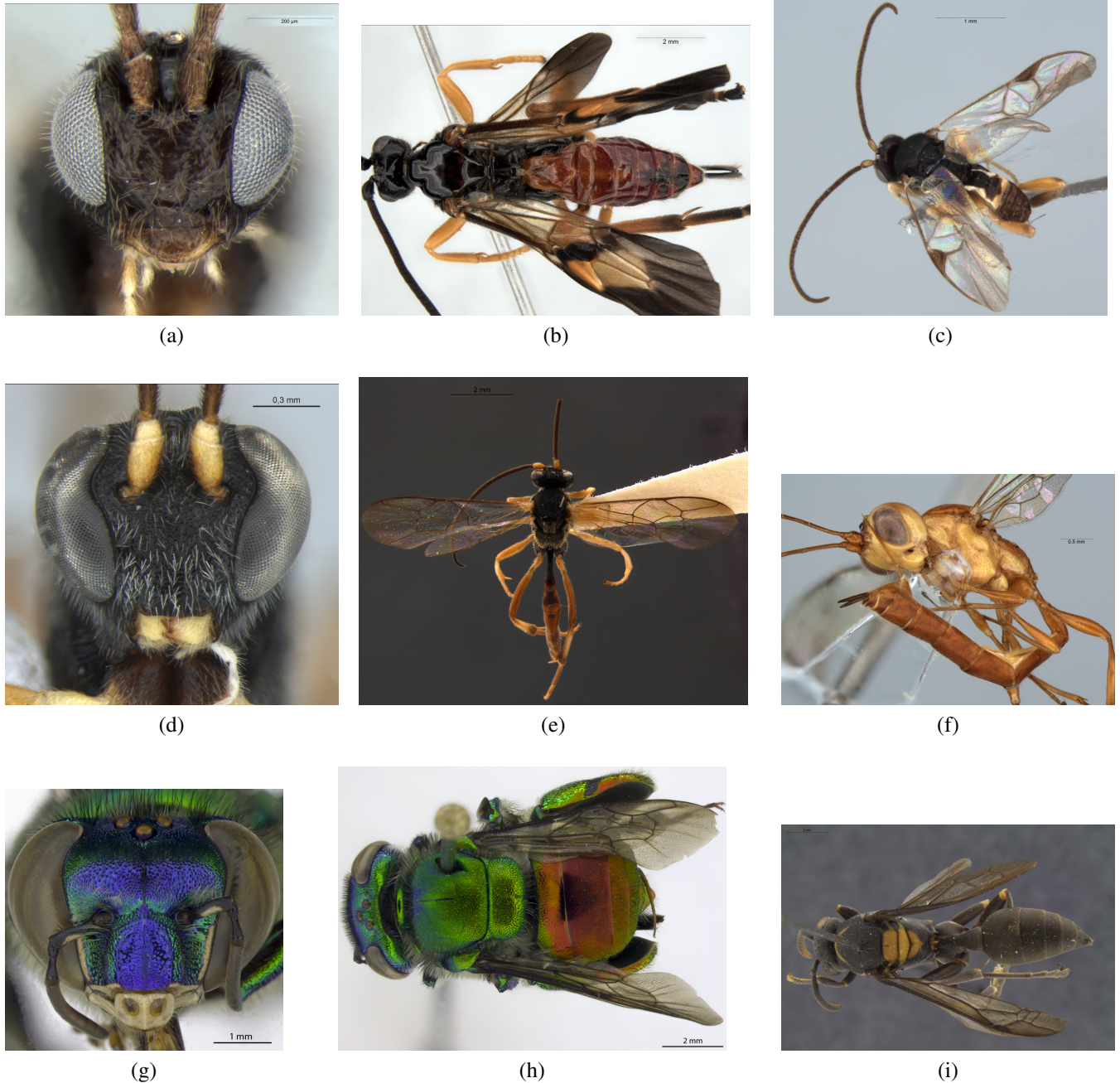


FIGURE 5. Examples of morphological variation in the DAPWH dataset. (a), (b), (c) Braconidae; (d), (e), (f) Ichneumonidae; (g), (h) Apidae; (i) Vespidae.

All computational experiments were executed using the PyTorch framework on a high-performance workstation equipped with an NVIDIA GeForce RTX 4090 GPU with 24 GB. To preserve the fine-grained morphological details inherent in the macro-photography dataset, a standardized input resolution of 224×224 and 1280×1280 pixels was maintained across the classification and detection models, respectively.

Image-Level Identification

To evaluate the discriminative power of different feature extractors on the Hymenoptera dataset, we conducted a com-

parative analysis of six distinct architectures: EfficientNetV2 [32], ConvNeXt [33], ViT-B/16 [34], VGG16 [35], ResNet50 [36], and YOLOv12 [37].

The quantitative performance metrics for the image-level classification task are consolidated in Table 3. The evaluated architectures achieved high efficacy, with Top-1 test-set accuracy ranging from 89.34% to 92.28%. Specifically, EfficientNetV2 and ConvNeXt reached a peak accuracy of 92.28%, while YOLOv12 demonstrated the highest overall robustness with an F1-score of 95.59%. Furthermore, the convergence behavior of these models is illustrated in Figure

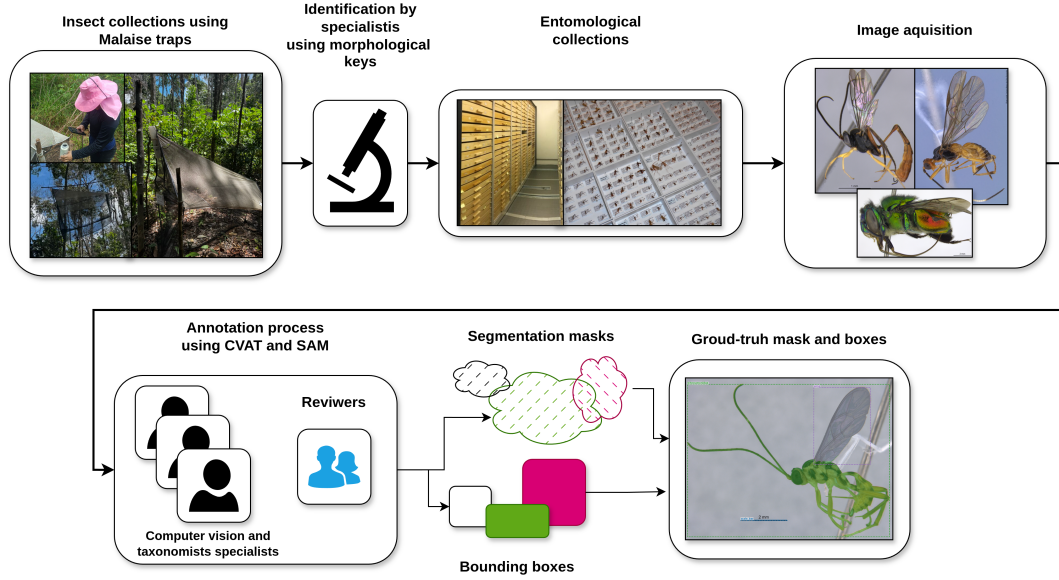


FIGURE 6. Illustration of the dataset creation and annotation workflow for the subset COCO annotated dataset. Initial segmentation masks and bounding boxes were generated by computer vision experts in CVAT, assisted by the SAM, and subsequently reviewed by taxonomists.

TABLE 2. Summary of COCO dataset and Image Properties

Parameter	Value	Notes
<i>Taxonomic Classes</i>		
Ichneumonidae	521	Family level (Masks)
Braconidae	406	Family level (Masks)
Apidae	158	Family level (Masks)
Vespidae	142	Family level (Masks)
Megachilidae	115	Family level (Masks)
Andrenidae	90	Family level (Masks)
Pompilidae	85	Family level (Masks)
Chrysididae	76	Family level (Masks)
Bethylidae	72	Family level (Masks)
Halictidae	49	Family level (Masks)
Colletidae	25	Family level (Masks)
<i>Annotated Features</i>		
Wing	1,794	Segmentation Masks
Scale bar	1,700	Bounding boxes

TABLE 3. Comparison of validation metrics for the evaluated classification models on the test dataset.

Model	Acc.	F1	Precision	Recall
EfficientNetV2	92.2%	84.07%	93.69%	82.54%
ConvNeXt Tiny	92.2%	87.00%	92.38%	86.24%
ViTB16	89.3%	80.45%	90.16%	78.48%
VGG16	89.5%	86.39%	88.88%	84.60%
ResNet50	90.1%	84.51%	88.71%	82.10%
YOLOv12	91.7%	95.59%	85.73%	88.88%

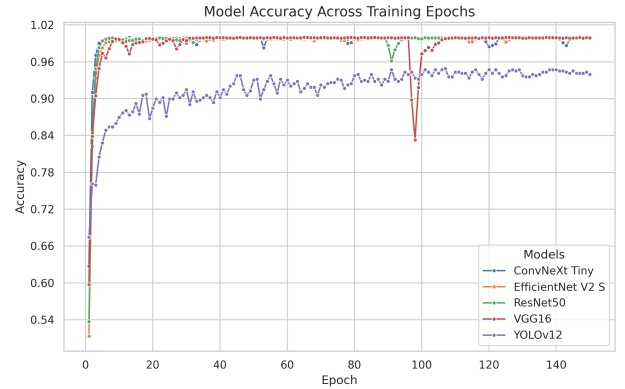


FIGURE 7. Validation accuracy per epoch across 150 epochs.

7, which shows the validation accuracy over 150 epochs of training.

Baseline Detector

The baseline detector model served as the baseline for object detection, utilizing the standard hyperparameters recommended by the YOLOv8 [38] and YOLOv12 [37] literature to establish a performance benchmark.

The object detection model was trained for 150 epochs, with performance metrics monitored across the validation

set. Figure 8 and Figure 9 illustrate the training dynamics, including loss convergence and metric evolution. Both training and validation losses demonstrated a consistent downward trend, indicating that the model effectively learned to localize bounding boxes and classify insect families without significant overfitting. The final performance metrics for the

object detection task, evaluated on the COCO subset test dataset, are summarized in Table 4.

TABLE 4. Comparative performance metrics for object detection architectures on the DAPWH test dataset.

Model	mAP50	Precision	Recall
YOLOv8	86.26%	79.08%	85.83%
YOLOv12	90.53%	90.23%	82.54%

The normalized confusion matrix for the YOLOv12s model is illustrated in Figure 10. The diagonal elements represent the recall for each category, demonstrating the model's strong ability to identify key taxonomic families and morphological features. Notably, the model achieved exceptional classification accuracy for Ichneumonidae, Apidae, Vespidae, and Bethyilidae.

The YOLOv12 model demonstrated high spatial and taxonomic precision, particularly for the identification of morphological features such as wings, which were detected with an accuracy of 96%. However, localized classification errors were observed between the families Andrenidae and Colletidae. This discrepancy is largely due to the significant class imbalance in the COCO subset; while the model achieved high recall for well-represented taxa, the Colletidae family contains only 25 mask-level annotations, representing the smallest sample size in the dataset. Such limited taxonomic representation constrains the model's ability to learn the distinct morphological boundaries of this family, especially when compared to the 927 images available for the superfamily Ichneumonoidea. Additionally, the scale_bar class shows a 0.61 background false negative rate, suggesting that while the object is detected, its localization boundaries occasionally overlap with background noise in high-resolution frames.

Figures 11 present representative inference detection examples from the test dataset. As illustrated, the object detection model shows strong spatial alignment, with the predicted bounding boxes for both the insect and wing structures closely matching the ground truth annotations.

RECORDS AND STORAGE

The DAPWH [23] dataset, comprising images of Hymenoptera families, is archived and publicly accessible via Zenodo (10.5281/zenodo.18501018). The data is released under the CC BY-NC 4.0 license. The repository includes the raw image files organized by taxonomic family and comprehensive metadata provided in the COCO (Common Objects in Context) annotation format via JSON files, facilitating its immediate use in object detection and classification pipelines, as shown in Figure 1. The dataset does not include structured environmental metadata or detailed locality-level attributes, as its primary focus is on computer vision and object detection tasks.

INSIGHTS, LIMITATIONS AND USAGE NOTES

The DAPWH dataset provides a multi-perspective visual record of Hymenoptera, featuring specimens captured in

lateral, frontal, and dorsal views to ensure comprehensive morphological coverage. Although the current annotations and benchmarks are restricted to the family level, the dataset preserves higher-resolution taxonomic information in the image filenames. This enables straightforward extension to subfamily or genus level classification as additional annotations become available.

A significant subset of these images includes high-fidelity annotations of the insect body, wings, and a reference scale bar. These annotations facilitate the automated extraction of phenotypic features and biometric measurements, such as wing venation patterns. While current annotations focus on wing and body segmentation, the dataset's structure invites the scientific community to contribute additional labels for structures such as antennae or ovipositors, supporting future research in evolutionary biology.

The classification and detection benchmarks performed on this dataset demonstrate its robust utility for automated entomological research. The image-level classification models achieved high performance, with Top-1 accuracy reaching 92.28% and YOLOv12 attaining an F1-score of 95.59%. Similarly, the object detection benchmarks yielded strong results, particularly with YOLOv12, which achieved an mAP@50 of 90.53%. These metrics validate the dataset's high quality and its effectiveness in supporting state-of-the-art computer vision architectures for the precise identification of Hymenoptera.

While the detection performance for the scale bar was lower than that of other classes, these objects remain high-value assets within the dataset. Each scale bar is accompanied by a precise metric annotation in millimeters (mm), ensuring that, even when automated detection requires refinement, the physical reference remains available for manual or semi-automated calibration.

SOURCE CODE AND SCRIPTS

The complete computational pipeline, including scripts for data preprocessing, dataset splitting, model training, and learning curve generation, is available in the GitHub repository: <https://github.com/joamh/DAPWH-2026>.

ACKNOWLEDGEMENTS

J.M.H.P., and G.N.H. contributed equally to this article, developed the dataset, annotated the images, drafted the manuscript, and performed technical validation. L.B.R.F., A.D.S., and H.C.O. reviewed the taxonomy of the insects, provided images of Ichneumonoidea and Vespidae and revised the manuscript. E.A.B.A. provided images of Apoidea families revised the manuscript and approved the final version. R.V.G., and M.A.C.V. performed technical validation, reviewed, revised, and approved the final version of the manuscript. A.M.P.D., and M.B. led the project and supervised the research, and reviewed, revised, and approved the final version of the manuscript.

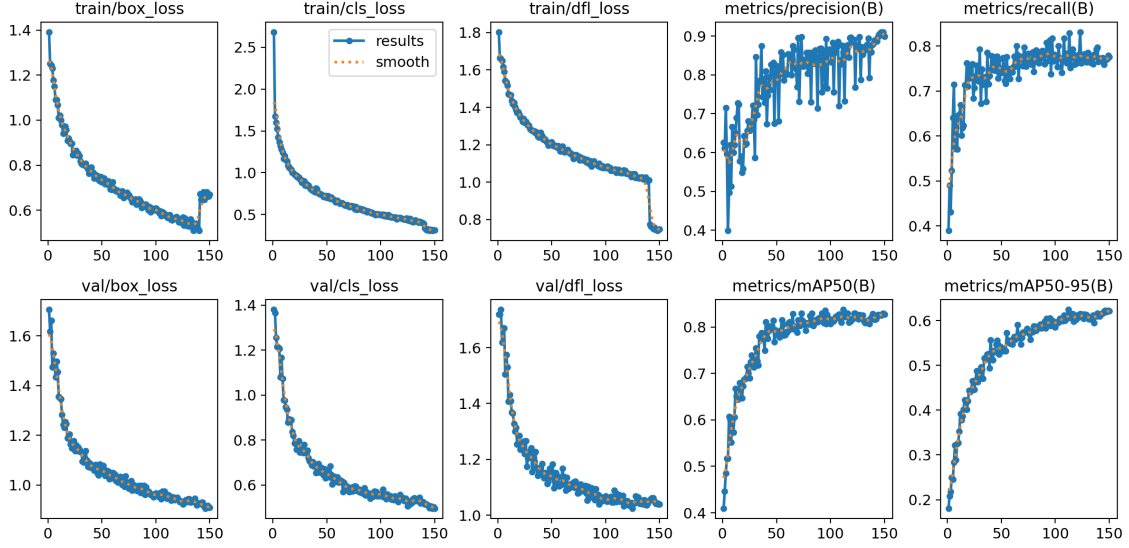


FIGURE 8. Training and validation metrics over 150 epochs for the YOLOv8 detection model.

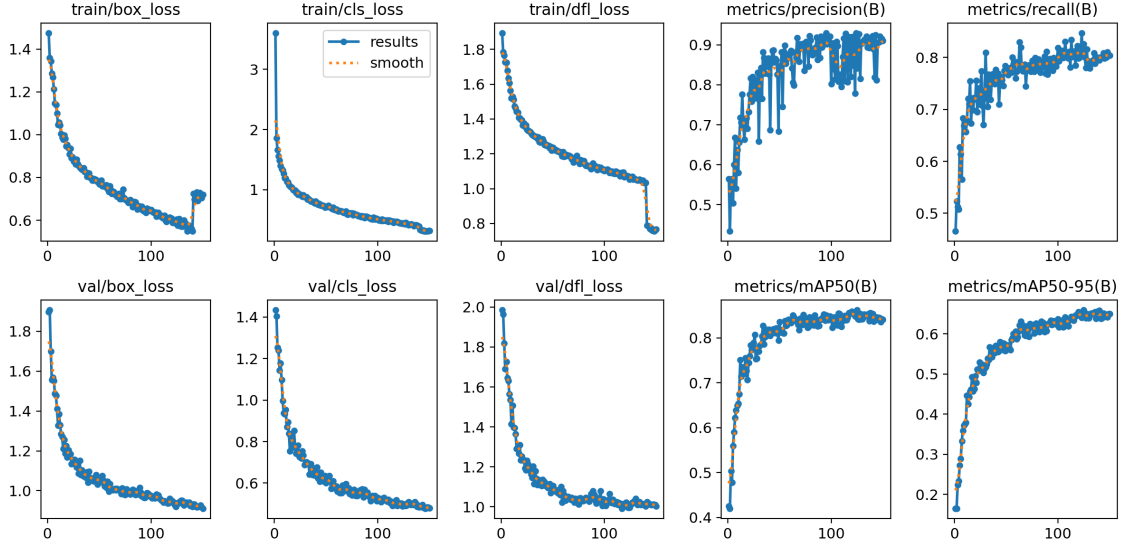


FIGURE 9. Training and validation metrics over 150 epochs for the YOLOv12 detection model.

This work was supported by the Petróleo Brasileiro S.A. - Petrobras, using resources from the R&D clause of the ANP, in partnership with the Universidade de São Paulo (USP) and the Fundação de Apoio à Física e à Química (FAFQ), under Cooperation Agreement No. 2023/00016-6 and 2023/00013-7, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) grant n°88887.002221/2024-00 and n°88887.975224/2024-00, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grant n°2014/50940-2, 2019/09215-6 and 2022/11451-2, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant n°465562/2014-0, Instituto Nacional de Ciência e Tecnologia dos Hymenoptera Parasitoides (INCT-HYMPAR).

The authors have declared no conflicts of interest.

REFERENCES

- [1] J. T. Huber, *Biodiversity of Hymenoptera*. John Wiley & Sons, Ltd, 2017, ch. 12, pp. 419–461. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118945568.ch12>
- [2] N. E. Stork, J. McBroom, C. Gely, and A. J. Hamilton, “New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 24, pp. 7519–7523, 2015. [Online]. Available: <https://www.pnas.org/doi/abs/10.1073/pnas.1502408112>
- [3] N. E. Stork, “How many species of insects and other terrestrial arthropods are there on earth?” *Annual Review of Entomology*, vol. 63, no. Volume 63, 2018, pp. 31–45, 2018. [Online]. Available: <https://www.annualreviews.org/content/journals/10.1146/annurev-ento-020117-043348>
- [4] A. A. Forbes, R. K. Bagley, M. A. Beer, A. C. Hippee, and H. A. Widmayer, “Quantifying the unquantifiable: why hymenoptera, not coleoptera, is the most speciose animal order,”

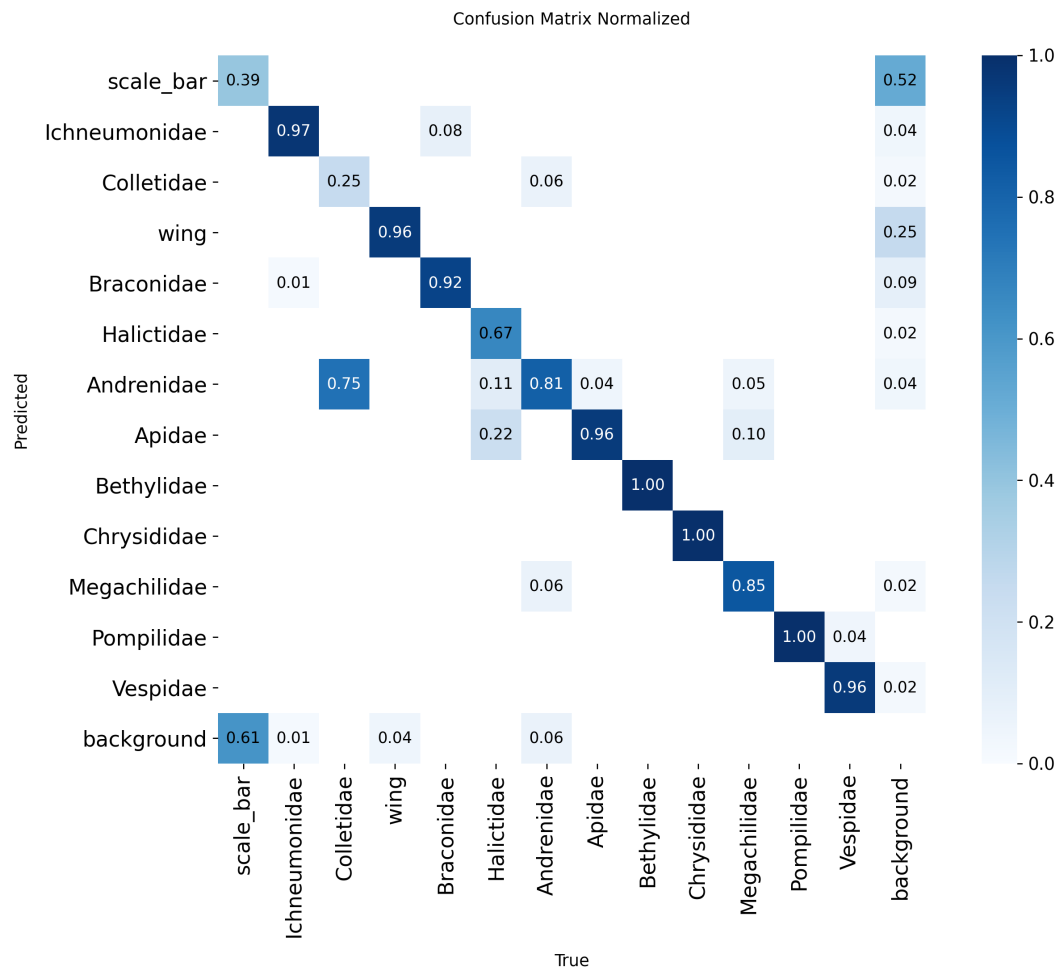


FIGURE 10. Confusion matrix normalized for YOLOv12.

- BMC Ecology*, vol. 18, no. 1, p. 21, 2018. [Online]. Available: <https://doi.org/10.1186/s12898-018-0176-x>
- [5] P. Eggleton, "The state of the world's insects," *Annual Review of Environment and Resources*, vol. 45, no. Volume 45, 2020, pp. 61–82, 2020. [Online]. Available: <https://www.annualreviews.org/content/journals/10.1146/annurev-environ-012420-050035>
 - [6] J. Ollerton, R. Winfree, and S. Tarrant, "How many flowering plants are pollinated by animals?" *Oikos*, vol. 120, no. 3, pp. 321–326, 2011. [Online]. Available: <https://nsojournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2010.18644.x>
 - [7] F. Sánchez-Bayo and K. A. Wyckhuys, "Worldwide decline of the entomofauna: A review of its drivers," *Biological Conservation*, vol. 232, pp. 8–27, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0006320718313636>
 - [8] R. Barbizan Sühs, A. Somavilla, A. Köhler, and J. Putzke, "Vespídeos (Hymenoptera, Vespidae) vetores de pólen de *Schinus terebinthifolius* Raddi (Anacardiaceae), Santa Cruz do Sul, RS, Brasil," *Brazilian Journal of Biosciences*, vol. 7, pp. 138–143, 04 2009.
 - [9] J. R. Beggs, E. G. Brockerhoff, J. C. Corley, M. Kenis, M. Masciocchi, F. Muller, Q. Rome, and C. Villemant, "Ecological effects and management of invasive alien vespidae," *BioControl*, vol. 56, no. 4, pp. 505–526, 2011. [Online]. Available: <https://doi.org/10.1007/s10526-011-9389-z>
 - [10] I. Gauld and B. Bolton, *The Hymenoptera*. British Museum (Natural History), 1988.
 - [11] D. Quicke, *The Braconid and Ichneumonid Parasitoid Wasps: Biology, Systematics, Evolution and Ecology*. Wiley, 2015.
 - [12] G. Broad, M. Shaw, and M. Fitton, *Ichneumonid Wasps (Hymenoptera: Ichneumonidae): Their Classification and Biology*, ser. Entomology Series. CABI, 2018.
 - [13] B. Butcher and D. Quicke, *The Parasitoid Wasps of South East Asia*. CAB International, 2023.
 - [14] A. Polaszek and L. Vilhemsen, "Biodiversity of hymenopteran parasitoids," *Current Opinion in Insect Science*, vol. 56, p. 101026, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2214574523000238>
 - [15] D. Scatolini and A. M. Pentead-Dias, "Análise faunística de Braconidae (Hymenoptera) em três áreas de mata nativa do Estado do Paraná, Brasil," *Revista Brasileira de Entomologia*, vol. 47, no. 2, pp. 187–195, 2003.
 - [16] K. Schoeninger, J. L. P. Souza, C. Krug, and M. L. Oliveira, "Diversity of parasitoid wasps in conventional and organic guarana (*Paullinia cupana* var. *sorbilis*) cultivation areas in the Brazilian Amazon," *Acta Amazonica*, vol. 49, no. 4, pp. 283–293, 2019.
 - [17] N. T. B. Antunes and D. R. R. Fernandes, "Faunistic analysis of Ichneumonidae (Hymenoptera) in guarana (*Paullinia cupana*) crop, with new records of genera for the Brazilian Amazon," *Arquivos do Instituto Biológico*, vol. 87, 2020.
 - [18] M. J. Sharkey, "Family Braconidae," in *Hymenoptera of the world: An identification guide to families*, H. Goulet and J. T. Huber, Eds. Ottawa, Ontario: Center for Land and Biological Resources Research, 1993, pp. 362–395.
 - [19] M. J. Sharkey and D. B. Wahl, "Superfamilia Ichneumonoidea," in *Introducción a los Hymenoptera de la Región Neotropical*, F. Fernández and M. J. Sharkey, Eds. Bogotá: Sociedad Colombiana de Entomología: Universidad Nacional de Colombia, 2006, pp. 287–292.

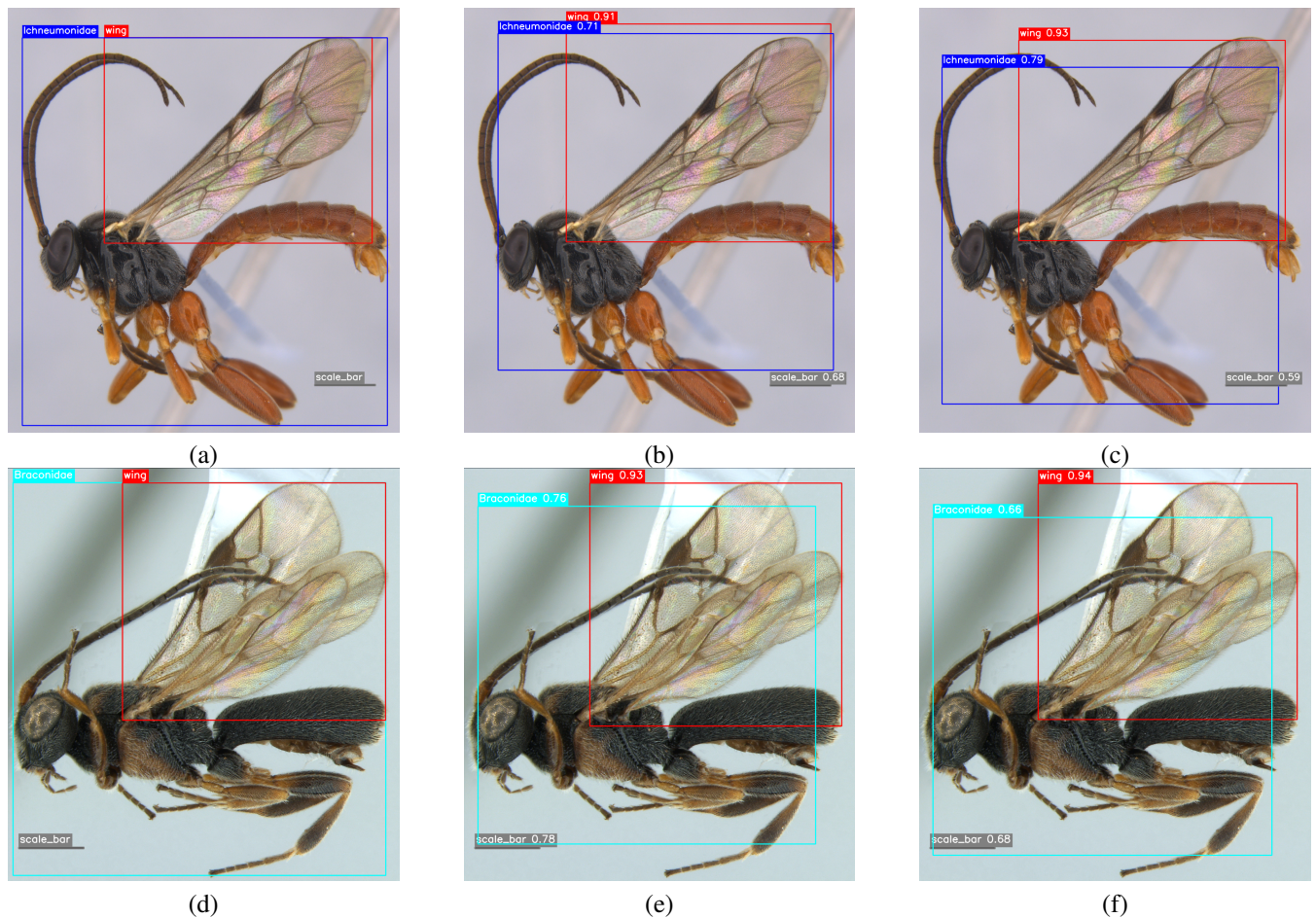


FIGURE 11. Visual comparison of detection results. (a,d) The corresponding ground truth annotations; (b,e) The predicted bounding boxes generated by the YOLOv8 during inference; (c,f) The predicted bounding boxes generated by the YOLOv12 during inference.

- [20] X.-x. Chen and C. van Achterberg, "Systematics, phylogeny, and evolution of braconid wasps: 30 years of progress," *Annual Review of Entomology*, vol. 64, no. Volume 64, 2019, pp. 335–358, 2019. [Online]. Available: <https://www.annualreviews.org/content/journals/10.1146/annurev-ento-011118-111856>
- [21] A. Petrović, "Sizing the knowledge gap in taxonomy: The last dozen years of aphidiinae research," *Insects*, vol. 13, no. 2, 2022. [Online]. Available: <https://www.mdpi.com/2075-4450/13/2/170>
- [22] S. H. Luke, H. E. Roy, C. D. Thomas, L. A. N. Tilley, S. Ward, A. Watt, M. Carnaghi, C. C. Jaworski, M. P. T. G. Tercel, C. Woodrow, S. Aown, J. A. Banfield-Zanin, S. L. Barnsley, I. Berger, M. J. F. Brown, J. C. Bull, H. Campbell, R. A. B. Carter, M. Charalambous, L. J. Cole, M. J. Ebejer, R. A. Farrow, R. S. Fartyal, M. Grace, F. Highet, J. K. Hill, A. S. C. Hood, E. S. Kent, F.-T. Krell, S. R. Leather, D. J. Leybourne, N. A. Littlewood, A. Lyons, G. Matthews, L. Mc Namara, R. Menéndez, P. Merrett, S. Mohammed, A. K. Murchie, M. Noble, M.-R. Paiva, M. J. Pannell, C.-K. Phon, G. Port, C. Powell, S. Rosell, F. Sconce, C. R. Shortall, E. M. Slade, J. P. Sutherland, J. C. Weir, C. D. Williams, N. B. Zielonka, and L. V. Dicks, "Grand challenges in entomology: Priorities for action in the coming decades," *Insect Conservation and Diversity*, vol. 16, no. 2, pp. 173–189, 2023. [Online]. Available: <https://resjournals.onlinelibrary.wiley.com/doi/abs/10.1111/icad.12637>
- [23] J. M. Herrera Pinheiro, G. do Nascimento Herrera, L. Bueno dos Reis Fernandes, A. Doria dos Santos, R. Vilela de Godoy, E. Andrade Botelho de Almeida, H. Carolina Onody, M. Andrade da Costa Vieira, A. Maria Pentead-Dias, and M. Becker, "Dataset of parasitoid wasps and associated hymenoptera (dapwh)," Feb. 2026. [Online]. Available: <https://doi.org/10.5281/zenodo.18501018>
- [24] A. M. Pentead-Dias and L. B. d. R. Fernandes, "Dcub - coleção taxonômica do departamento de ecologia e biologia evolutiva da ufscar," 2025. [Online]. Available: <https://doi.org/10.15468/xzkz3y>
- [25] E. A. B. Almeida, A. M. Costa, J. A. T. Filho, M. M. P. Zichinelli, and F. B. Quinteiro, "Illustrated catalogue of type specimens of insects (Hexapoda) at Coleção Entomológica "Prof. J.M.F.Camargo" (RPSP), Universidade de São Paulo, Brazil," *Zootaxa*, vol. 4842, no. 1, p. 1–204, Aug. 2020. [Online]. Available: <https://www.mapress.com/zt/article/view/zootaxa.4842.1.1>
- [26] K. Needham, "University of British Columbia - Spencer Entomological Collection (UBCZ). Version 13.12," 2025, accessed via GBIF.org on 2025-12-10.
- [27] T. d. O. Andrade, K. S. Ramos, H. C. Onody, A. D. d. Santos, and C. R. F. Brandão, "Type specimens of Pompiloidea, Thynnoidea and Vespoidea (Hymenoptera) deposited in the Museu de Zoologia da Universidade de São Paulo, Brazil," *Papéis Avulsos de Zoologia*, vol. 58, p. e20185839, Aug. 2018. [Online]. Available: <https://revistas.usp.br/paz/article/view/144141>
- [28] CVAT.ai Corporation, "Computer Vision Annotation Tool (CVAT)," Nov. 2023. [Online]. Available: <https://github.com/cvat-ai/cvat>
- [29] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment anything," 2023. [Online]. Available: <https://arxiv.org/abs/2304.02643>
- [30] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, "Microsoft coco: Common objects in context," 2015. [Online]. Available: <https://arxiv.org/abs/1405.0312>
- [31] S. Raschka, "Model evaluation, model selection, and algorithm selection in machine learning," 2020. [Online]. Available: <https://arxiv.org/abs/1811.12808>

- [32] M. Tan and Q. V. Le, “Efficientnetv2: Smaller models and faster training,” 2021. [Online]. Available: <https://arxiv.org/abs/2104.00298>
- [33] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “A convnet for the 2020s,” 2022. [Online]. Available: <https://arxiv.org/abs/2201.03545>
- [34] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, “An image is worth 16x16 words: Transformers for image recognition at scale,” 2021. [Online]. Available: <https://arxiv.org/abs/2010.11929>
- [35] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2015. [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [36] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [37] Y. Tian, Q. Ye, and D. Doermann, “Yolov12: Attention-centric real-time object detectors,” 2025. [Online]. Available: <https://arxiv.org/abs/2502.12524>
- [38] R. Varghese and S. M., “Yolov8: A novel object detection algorithm with enhanced performance and robustness,” in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, 2024, pp. 1–6.