Title	Dataset name and URL	Dataset description	Methods name	Accuracy of the model	Pros	Cons	Citation
Microscopic Image Dataset of Plant-Parasitic Nematodes	https://data.mendeley.com/datasets/cck8yxj3xw/2	1,016 images, 11 classes. No predefined train/Val/test split	ResNet101V2 CoAtNet-0 EfficientNetV2B0 EfficientNetV2M	97.94% test (EfficientNetV2B0/M + Respro	**1**.High-resolution (1280×1024) **2**.curated by nematologists **3.** publicly available, **4**. includes tropical species from Indonesia, **5**. supports AI development.	**1.**Imbalanced class distribution **2.** geographic limitation **3**. no predefined data splits.	Indarti, S., Shabrina, N. H., & Maharani, R. (2025). Microscopic image dataset of plant-parasitic nematode. *Data in Brief*, 61, 111687. https://doi.org/10.1016/j.dib.2025.111687
valuation of Selected Methods in the Control of Plant Parasitic Nematodes Infecting Carnation	Not Applicable (Field Experiment)	**Samples:** Soil and root samples from experimental plots. **Classes/Treatments:** 8 (Sugarcane bagasse, Tea compost, Flower compost, Fenamiphos, Neem, Molasses, *P. lilacinus*, Control). **Replicates:** 6 per treatment. Measurements taken at 90 and 180 days after treatment.	1. Sugarcane Bagasse 2. Tea Compost 3. Flower Compost 4. Fenamiphos (Nemacur®) 5. Neem (Achook®) 6. Molasses 7. *Paecilomyces lilacinus* (PL plus®) 8. Control (Untreated)	**Primary Efficacy Metric (% Reduction in Root Galling Index):** 1. Bagasse: 53% 2. Tea compost: 62% 3. Flower compost: 60% 4. Fenamiphos: 67% 5. Neem: 69% 6. Molasses: 58% 7. *P. lilacinus*: 58% 8. Control: 0%	**1**. Environmentally friendly, **2**. sustainable, **3.** improve soil health,, **4**target multiplespecies from Indonesia, **5**. Highly effective, fast-acting..	**1.** Generally less effective than the chemical standard, distribution **2**.slower to act, **3**. Harmful to non-target organisms and the environment.	Kimenju, J. W., Wachira, P. M., Lang'at, J. K., Otieno, W., & Mutua, G. K. (2014). Evaluation of Selected Methods in the Control of Plant Parasitic Nematodes Infecting Carnation. *Journal of Agricultural Science, 6*(3), 31-38. http://dx.doi.org/10.5539/jas.v6n3p31
Plant parasitic nematode management using antagonistic plants as a potential substitute to hazardous	https://doi.org/10.1080/09670874.2024.2399069	Literature review covering multiple studies; no primary dataset; focuses on antagonistic plants (Tagetes, Crotalaria, Brassica, Azadirachta indica, Lantana camara)	Review and synthesis of existing research; no model training or evaluation	Not applicable (review article)	Comprehensive overview of eco-friendly nematode management; identifies key plant species and mechanisms	No original dataset or model; limited to qualitative synthesis	Mushtaq et al. (2024). *International Journal of Pest Management*. DOI: 10.1080/09670874.2024

chemical control – a review							.2399069
A Semi-supervised approach to cluster symptomatic and asymptomatic leaves in root lesion nematode infected walnut trees	https://doi.org/10.1016/j.compag.2022.106761	Hyperspectral leaf scans (360–1700 nm) from 3 walnut genotypes; 8 measurements per tree; control vs. infected treatments; used for clustering symptomatic leaves	Semi-supervised agglomerative clustering with SAM distance, N-FINDR endmember selection, MDPA ranking	Not a classification model; genotype tolerance ranking validated against nematological ground truth	Handles within-canopy spectral variability; identifies asymptomatic leaves; provides validated genotype ranking	Requires per-genotype parameter tuning; not directly comparable to standard accuracy metrics	Omidi et al. (2022). *Computers and Electronics in Agriculture*, 194, 106761.
Dataset on the diversity of plant-parasitic nematodes in cultivated olive trees in southern Spain	https://doi.org/10.1016/j.dib.2019.104658	376 soil samples from olive orchards; 128 species of plant-parasitic nematodes across 38 genera and 13 families; abundance and biomass data per species and site.	Integrative taxonomy (morphological + molecular identification), beta diversity partitioning (LCBD/SCBD)	Not a machine learning model; taxonomic identification validated via integrative methods.	Comprehensive sampling; high taxonomic resolution; includes ecological indices (LCBD/SCBD); open access.	No imagery or genomic sequences; data is ecological/taxonomic; not suitable for image-based ML.	Archidona-Yuste et al. (2019). *Data in brief*, 27, 104658.
Advancements in Machine Vision and Deep Learning for Automated Nematode Cyst Identification in Agricultural Soil	Not publicly named/URL: Not provided	**Total Images:** 10,000+ high-resolution cyst images **Classes:** Nematode cysts vs. other soil particles (specific species not detailed) **Data Type:** Hyperspectral and high-resolution digital microscope images of isolated cysts	1. Convolutional Neural Networks (CNNs - e.g., ResNet-50) 2. Transformer-Based Models 3. Traditional Machine Learning (Baseline) **Supporting Tech:** Hyperspectral imaging, IoT, Edge Computing	**CNN:** 94.5% **Transformer-Based Model:** 96.2% **Traditional ML:** 88.3%	**- High Accuracy & Speed:** Significantly outperforms manual methods; near real-time detection (~0.8 seconds/image). **- Automation & Scalability:** Enables high-throughput analysis, reducing labor.	**- Dataset Scarcity:** Lack of large, diverse public datasets limits model generalization. **- High Resource Demand:** Requires GPUs and significant computational power.	Oluwagbade, E. (2024). Advancements in Machine Vision and Deep Learning for Automated Nematode Cyst Identification in Agricultural .

Title	Dataset name and URL	Dataset description	Methods name	Accuracy of the model	Pros	Cons	Citation
A Deep Learning-Based Decision Support Tool for Plant-Parasitic Nematode Management	**URL:** Not publicly specified; dataset was self-collected for this study.	**Samples:** 415 images of RKN eggs. **Classes:** 1 class (RKN eggs). **Split:** 80% training (332 images), 20% validation (83 images).	**YOLOv5, YOLOv6, YOLOv7 (tested at input resolutions 224×224, 480×480, 640×640)**	**Best Model (YOLOv5-640):** Precision: 0.992 Recall: 0.959 F1-score: 0.975 mAP@0.5: 0.979 Inference Time: 3.9 ms	1. Very high precision and recall. 2. Extremely fast inference suitable for real-time use. 3. Integrated into a user-friendly web tool (NemDST). 4. Provides population counts, not just detection. 5. Helps in decision-making for nematode management..	1. Dataset is not publicly available. 2. Limited to one nematode type (RKN eggs). 3. Performance may drop with low-quality images or different microscopes. 4. May confuse spoiled eggs or soil particles of similar color.	Pun, T.B.; Neupane, A.; Koech, R. A Deep Learning-Based Decision Support Tool for Plant-Parasitic Nematode Management. J. Imaging 2023, 9, 240. https://doi.org/10.3390/jimaging9110240
Machine learning algorithms accurately identify free-living marine nematode species	**URL:** https://zenodo.org/record/7278042	Samples: 145 individuals (*Acantholaimus*), 260 individuals (*Sabatieria*). Classes: 40 species (*Acantholaimus*), 60 species (*Sabatieria*). Data Type: Not images, but a data matrix of morphometric measurements and categorical morphological characters. Split: *Acantholaimus*: 131 training, 14 testing. *Sabatieria*: 227 training, 33 testing.	Random Forest (RF), Stochastic Gradient Boosting (SGBoost), Support Vector Machine (SVM) with linear and radial kernels, K-Nearest Neighbors (KNN)	**Random Forest:** *Acantholaimus*: 93% *Sabatieria*: 100% SVM (Linear): *Acantholaimus*: 92% *Sabatieria*: 100% SVM (Radial): *Acantholaimus*: 92% *Sabatieria*: 97**%**	1. Automates and speeds up species identification. 2. Makes identification accessible to non-taxonomists. 3. High accuracy achieved with the best models. 4. Based on traditional, well-established morphological data. 5. Identifies the most important taxonomic characters.	1. Performance depends on the quality and completeness of the original species descriptions. 2. Requires prior manual measurement and character categorization. 3. Less effective for species described from a single or few individuals. 4. Misclassification can occur (e.g., 1 *Acantholaimus* individual)..	Brito de Jesus S, Vieira D, Gheller P, Cunha BP, Gallucci F, Fonseca G. 2023. Machine learning algorithms accurately identify free-living marine nematode species. *PeerJ* 11:e16216 DOI 10.7717/peerj.16216
							…………………… ……………………

							………..

Related-work summary :

This summary surveys recent research on plant-parasitic nematode detection and management, highlighting a shift toward AI-driven and eco-friendly approaches.

In image-based detection, several studies applied deep learning to automate nematode identification. For instance, a microscopic image dataset of 1,016 images across 11 classes achieved 97.94% accuracy using EfficientNetV2 and ResNet models, though it suffers from class imbalance. Another study used YOLO models to detect root-knot nematode eggs with high precision (0.992) and integrated the model into a user-friendly web tool. Similarly, a cyst identification system using CNNs and transformers reached 96.2% accuracy, enabling near real-time analysis. However, these methods often face challenges such as dataset scarcity, computational demands, and limited generalization.

Beyond imagery, morphometric and spectral data have also been leveraged. One study used machine learning (Random Forest, SVM) on morphological measurements to identify marine nematode species with up to 100% accuracy. Another employed semi-supervised clustering on hyperspectral leaf data to distinguish symptomatic from asymptomatic walnut trees, supporting genotype tolerance ranking without standard accuracy metrics.

In ecological and control-oriented studies, non-AI methods were prominent. A Spanish olive cultivation study provided detailed taxonomic and diversity data using integrative taxonomy, while field experiments evaluated organic amendments like compost and neem-based products, showing up to 69% reduction in root galling. A 2024 review also highlighted antagonistic plants as sustainable alternatives to chemical nematicides.

Overall, the field is advancing through high-accuracy AI models for diagnostics and a growing emphasis on sustainable management. Key limitations include data availability, geographic bias, and the trade-off between model performance and ecological applicability. Future work may benefit from integrating multimodal data and expanding public datasets to improve model robustness and support global agricultural needs.