

Date of publication, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

# NemaSSL++:Supervised and Self-Supervised Representation Learning for Microscopic Classification of Plant-Parasitic Nematodes

**RIPON HASAN<sup>1</sup>, SHARIFUZZAMAN MAZUMDER<sup>2</sup>, MOUMITA AKTER<sup>3</sup>, MARJUK IBNA BELATED<sup>4</sup>**

<sup>1</sup>Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh (e-mail: 2022-1-60-054@std.ewubd.edu)

<sup>2</sup>Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh (e-mail: 2022-1-60-003@std.ewubd.edu)

<sup>3</sup>Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh (e-mail: 2022-1-60-105@std.ewubd.edu)

<sup>4</sup>Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh (e-mail: 2022-1-60-383@std.ewubd.edu)

This work was supported in part by the Department of Computer Science and Engineering, East West University

**ABSTRACT** Plant-parasitic nematodes (PPNs) represent a major threat to agricultural productivity worldwide, yet accurate microscopic identification remains challenging, time-consuming, and highly dependent on expert knowledge. This paper proposes NemaSSL++, a unified framework that integrates supervised and self-supervised representation learning for robust microscopic classification of PPNs. Experiments are conducted on a publicly available dataset containing 1,016 microscopic images across nine nematode genera. Multiple supervised convolutional and transformer-based backbones are evaluated alongside three self-supervised learning (SSL) methods: SimCLR, BYOL, and MoCo. Among all configurations, ResNet-50 with SimCLR pretraining achieves the highest test accuracy of 80.88% after fine-tuning, while BYOL demonstrates superior Comprehensive evaluations using linear probing, shallow classifiers, full fine-tuning nematode classification tasks.

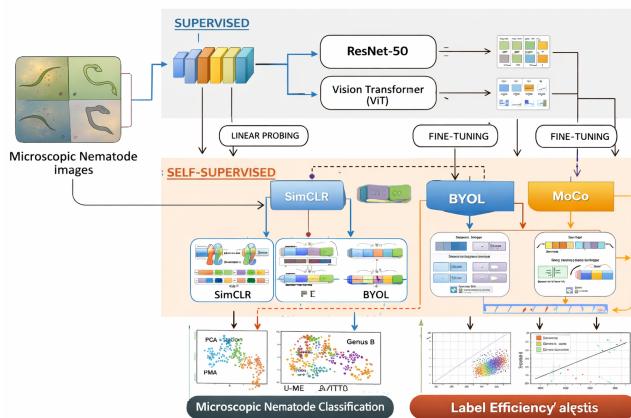
**INDEX TERMS** Plant-Parasitic Nematodes, Self-Supervised Learning, Microscopic Image Classification, Deep Learning, Label Efficiency.

## I. INTRODUCTION

**(PPNs)** Plant-parasitic nematodes are among the most destructive agricultural pests, causing substantial economic losses worldwide by damaging plant root systems and impairing nutrient absorption. These microscopic organisms affect a wide range of crops, including cereals, vegetables, and fruits, thereby posing a serious threat to global food security. Accurate identification of nematode genera is crucial for effective pest management, disease control, and sustainable agricultural practices. However, traditional nematode identification methods rely heavily on manual microscopic inspection and expert knowledge of morphological traits, which are labor-intensive, time-consuming, and prone to subjective interpretation. With the rapid advancement of computer vision and machine learning, automated image-based nematode classification has gained significant attention. Early studies focused on classical machine learning approaches such as Support

Vector Machines (SVM), k-Nearest Neighbors (kNN), and Random Forests (RF), utilizing handcrafted morphological features extracted from microscopic images. While these methods demonstrated moderate success, their performance was limited by the quality of manually engineered features and their inability to generalize across varying imaging conditions. Moreover, such approaches struggle to capture subtle intra-class variations and complex structural patterns inherent in nematode morphology. The emergence of deep learning, particularly convolutional neural networks (CNNs), marked a major shift in microscopic image analysis. Supervised CNN-based models such as VGG, ResNet, Inception, and EfficientNet have been widely adopted for biomedical and agricultural image classification tasks. Several prior works reported high accuracy using fully supervised CNN architectures for nematode detection and classification. Despite their strong performance, these methods require large-scale annotated datasets, which are difficult and costly to obtain in agricul-

tural To overcome the limitations of supervised learning,



**FIGURE 1.** Overall Workflow of the Study

self-supervised learning (SSL) has emerged as a powerful paradigm that enables models to learn meaningful feature representations from unlabeled data. SSL methods leverage pretext tasks or contrastive objectives to capture semantic and structural information without explicit annotations. Recent SSL frameworks such as SimCLR, Bootstrap Your Own Latent (BYOL), and Momentum Contrast (MoCo) have demonstrated remarkable success across various visual recognition tasks. These approaches are particularly attractive for nematode classification, where unlabeled microscopic images are relatively abundant compared to labeled ones. SimCLR employs contrastive learning by maximizing agreement between different augmented views of the same image while pushing apart representations of different images. BYOL, on the other hand, removes the need for negative samples by using an online-target network architecture, making it more stable and label-efficient. MoCo introduces a momentum-updated encoder and a dynamic memory queue to improve representation consistency across large datasets. Despite their success in general computer vision, systematic comparisons of these SSL techniques for plant-parasitic nematode classification remain scarce. In addition to CNN-based backbones, Vision Transformers (ViTs) have recently gained popularity for image representation learning by modeling long-range dependencies using self-attention mechanisms. ViTs have shown strong performance on large-scale datasets; however, their effectiveness in low-data and microscopic imaging scenarios is still an open research question. Combining SSL frameworks with both CNN and transformer-based backbones provides an opportunity to better understand architectural trade-offs in nematode classification tasks. Motivated by these gaps, this work presents a comprehensive analysis of supervised and self-supervised representation learning for microscopic classification of plant-parasitic nematodes. We evaluate multiple pretrained supervised backbones alongside three SSL frameworks—SimCLR, BYOL, and MoCo—using ResNet-50 and Vision Transformer architectures. A unified experimental setup is employed to ensure fair comparison across

methods. The learned representations are evaluated through linear probing, shallow classifiers, and full fine-tuning to assess their discriminative power, generalization capability, and computational efficiency. Extensive experiments are conducted on a multi-class microscopic nematode dataset comprising 1,016 images across several including accuracy, precision, recall, F1-score, and ROC-AUC. Furthermore, label efficiency analysis is performed by training models with varying proportions of labeled data, highlighting the practical advantages of SSL in low-annotation regimes. Embedding quality is analyzed using PCA, t-SNE, and UMAP visualizations, providing qualitative The results demonstrate that self-supervised pretraining significantly enhances performance compared to training from scratch. SimCLR achieves highest fine-tuned test accuracy, supported by well-separated embedding clusters. BYOL exhibits superior label efficiency, maintaining competitive accuracy even with a small fraction of labeled data. MoCo combined shows weaker frozen-feature performance but benefits substantial healthcare professionals reduce misdiagnosis risk, and improve patient Embedding quality is analyzed using PCA, t-SNE, and UMAP visualizations, providing qualitative insights

## II. RELATED WORKS

Research on plant-parasitic nematodes has evolved significantly over the past decade, encompassing traditional field-based control methods, classical machine learning techniques, and recent deep learning–driven computer vision approaches. Existing works can be broadly categorized into: (i) image-based nematode detection and classification, (ii) machine learning–based species identification, (iii) hyperspectral and semi-supervised disease detection, and (iv) non-image-based ecological and management studies.

### A. DEEP LEARNINGBASED IMAGE DETECTION OF NEMATODES

Pun et al. (2023) proposed a deep learning–based decision support tool for plant-parasitic nematode management using microscopic images of root-knot nematode (RKN) eggs. The study employed YOLOv5, YOLOv6, and YOLOv7 architectures on a self-collected dataset consisting of 415 images. The YOLOv5-640 model achieved a precision of 0.992 and an mAP@0.5 of 0.979, demonstrating strong real-time detection capability. However, the dataset was not publicly available and focused only on a single nematode class, limiting reproducibility and generalization. Oluwagbade (2024) investigated advancements in machine vision and deep learning for automated nematode cyst identification in agricultural soil . Approximately 10,000 high-resolution cyst images were used to compare CNN-based models (ResNet-50), transformer-based models, and traditional machine learning approaches. Transformer architectures achieved the highest accuracy of 96.2%, indicating superior feature representation capability. Despite strong performance, the lack of a public dataset and high GPU requirements remain significant limitations.

**TABLE 1.** Summary of Related Works on Nematode Detection and Management

Paper (Year)	Dataset Name	Methods	Accuracy / Outcome
Pun et al. (2023)	Self-collected RKN egg images	YOLOv5, YOLOv6, YOLOv7	Precision: 0.992, mAP@0.5: 0.979
Oluwagbade (2024)	Soil cyst image dataset	ResNet-50, Transformer	Transformer accuracy: 96.2%
Brilio de Jesus et al. (2022)	Marine nematode morphometrics	RF, SVM, KNN, SCBoost	RF accuracy: 93–100%
Indarti et al. (2025)	Microscopic nematode images (public)	EfficientNetV2, ResNet101V2	Accuracy: 97.94%
Omidi et al. (2022)	Hyperspectral leaf scans	SAM, N-FINDR, MDPA	Early detection (qualitative)
Kimenju et al. (2014)	Field experiment	Organic vs. chemical control	53–69% root gall reduction
Mushtaq et al. (2024)	Review study	Literature synthesis	N/A
Archidona-Yuste et al. (2019)	Olive soil nematode dataset	Integrative taxonomy	N/A

### B. MACHINE LEARNING FOR SPECIES IDENTIFICATION

Brilio de Jesus et al. (2022) explored the application of classical machine learning algorithms for identifying free-living marine nematode species. The dataset consisted of morphometric measurements from 40 species, including *Leptolaimus* and *Sabatieria*. Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and SC-Boost classifiers were evaluated. Random Forest achieved accuracies ranging from 93% to 100%, while SVM and KNN also showed competitive performance. Although the results were promising, the study relied on limited morphometric features and relatively small datasets.

### C. PUBLIC IMAGE DATASETS AND CNN-BASED CLASSIFICATION

Indarti et al. (2025) introduced a publicly available microscopic image dataset of plant-parasitic nematodes comprising 1,016 images across 11 classes. Several deep learning architectures, including ResNet101V2, CoAtNet-0, and EfficientNetV2, were evaluated. The highest accuracy of 97.94% was achieved using EfficientNetV2 combined with residual propagation. This dataset significantly improves reproducibility; however, challenges such as class imbalance and geographic bias persist.

### D. HYPERSPECTRAL AND SEMI-SUPERVISED APPROACHES

Omidi et al. (2022) proposed a semi-supervised learning framework for detecting root lesion nematode infection in walnut trees using hyperspectral imaging data. The study employed clustering techniques such as Spectral Angle Mapper (SAM), N-FINDR, and Minimum Distance Pixel Analysis (MDPA) to separate symptomatic and asymptomatic leaves. The approach demonstrated effectiveness in early infection detection, although extensive genotype-specific tuning was required.

### E. TRADITIONAL AND ECO-FRIENDLY NEMATODE MANAGEMENT STUDIES

Kimenju et al. (2014) conducted a field-based evaluation of organic and chemical control methods for managing plant-parasitic nematodes infecting carnation crops. The results showed a 53–69% reduction in root galling using organic

treatments, highlighting their environmental benefits despite lower effectiveness compared to chemical methods. Mushtaq et al. (2024) reviewed the use of antagonistic plants such as *Tagetes*, *Crotalaria*, and *Brassica* as eco-friendly alternatives to chemical nematicides. While the review provided broad insights into sustainable nematode management, it lacked quantitative performance evaluation.

### F. NON-IMAGE-BASED ECOLOGICAL AND TAXONOMIC STUDIES

Archidona-Yuste et al. (2019) released a large-scale dataset on the diversity of plant-parasitic nematodes in cultivated olive trees in southern Spain. The dataset included 376 soil samples covering 128 species across 38 genera. Although this work provides extensive taxonomic and ecological coverage, it is not suitable for image-based automated detection tasks.

## III. METHODOLOGY

This section describes the complete experimental pipeline adopted in this study, covering exploratory data analysis, supervised baselines, self-supervised pretraining, and downstream evaluation. The methodology is structured to ensure reproducibility, prevent data leakage, and enable fair comparison between supervised and self-supervised learning approaches.

### A. DATA DESCRIPTION

The experiments are conducted on the Microscopic Image Dataset of Plant-Parasitic Nematodes, which consists of microscopic images captured under controlled laboratory conditions. The dataset contains a total of 1,016 images belonging to 9 nematode genera (classes), including *Criconemoides*, *Helicotylenchus*, *Hirschmanniella*, *Hoplolaimus*, *Meloidogyne*, *Pratylenchus*, *Radopholus*, *Trichodorus*, and *Xiphinema*. All images are preprocessed by resizing to a fixed spatial resolution of 224 × 224 pixels with three RGB channels, ensuring compatibility with standard CNN backbones such as ResNet. Pixel values are normalized using ImageNet mean and standard deviation statistics. The dataset exhibits class imbalance, with certain genera (e.g., *Meloidogyne*) having substantially more samples than rare classes such as *Radopholus*. Additionally, images



FIGURE 2. NemaSSL++ Declaration Dataset images

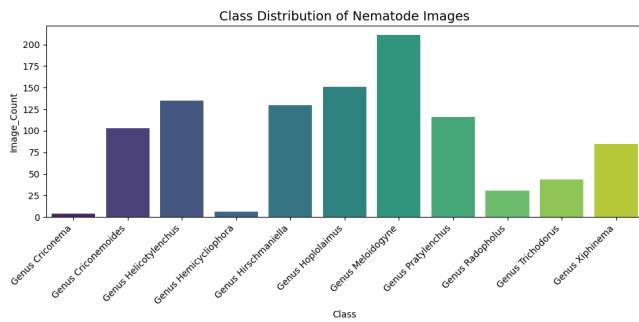


FIGURE 3. Calss Distrbution of Nematode images

## B. IMAGE-FOCUSED EXPLORATORY DATA ANALYSIS

### 1) Color Distribution and Illumination Analysis

RGB and HSV histograms are computed for each image and aggregated per class. Per-image and per-class mean and standard deviation values are analyzed to study brightness, contrast spread, and saturation clipping. This analysis helps identify illumination bias and color dominance across classes.

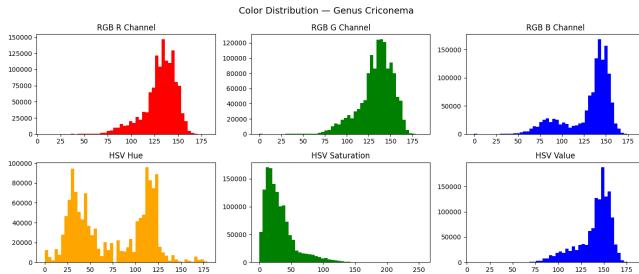


FIGURE 4. Color Distribution - Genus Criconema

## C. MODEL ARCHITECTURE

### 1) Supervised Model

**ResNet-50:** ResNet-50 is a deep convolutional neural network that employs residual connections to enable efficient training of very deep architectures. Unlike traditional convolutional networks, ResNet-50 introduces skip (identity) connections, which allow feature information to bypass one or

more convolutional layers. This design effectively mitigates the vanishing gradient problem and facilitates stable gradient propagation during training. The network consists of 50 layers organized into convolutional blocks and identity blocks, followed by global average pooling and a fully connected layer for classification. Through this hierarchical structure, ResNet-50 learns increasingly abstract representations, ranging from low-level texture and edge features to high-level morphological structures.

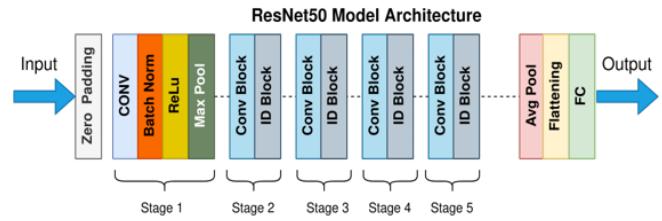


FIGURE 5. Resnet-50 Model Arhitecture

**Vision Transformer (ViT):** Vision Transformer (ViT) is a deep learning architecture that replaces convolution operations with self-attention mechanisms to model global relationships within an image. ViT splits an input image into fixed-size patches, embeds them into tokens, and processes them using transformer encoders. This design enables effective learning of long-range dependencies and contextual features. For plant-parasitic nematode image classification, ViT is valuable as it captures global shape, spatial relationships, and structural patterns that complement convolutional neural network features.

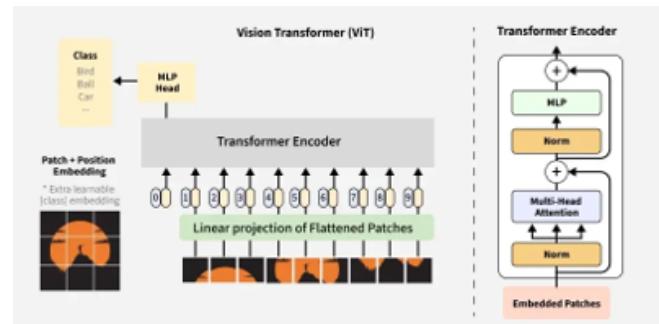
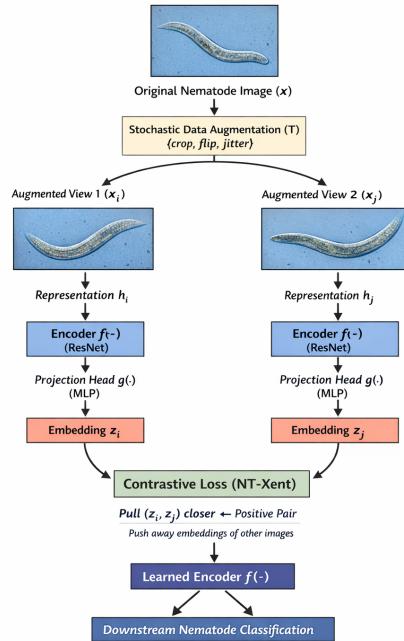


FIGURE 6. ViT Model Architecture

### 2) Self-Supervised Learning Models

**SimCLR:** (Simple Framework for Contrastive Learning of Representations) is a self-supervised learning architecture designed to learn robust visual representations without using labeled data. In SimCLR, each input nematode image is transformed into two different augmented views using a stochastic augmentation pipeline that includes operations such as random cropping, flipping, color jittering, and blurring. Both augmented views are passed through a shared base encoder (e.g., ResNet-50), producing feature representations. These features are further mapped by a projection

head, typically a small multilayer perceptron (MLP), into a latent embedding space. A contrastive loss function is then applied to maximize the similarity between embeddings of the same image (positive pairs) while minimizing similarity with embeddings from other images in the batch (negative pairs).



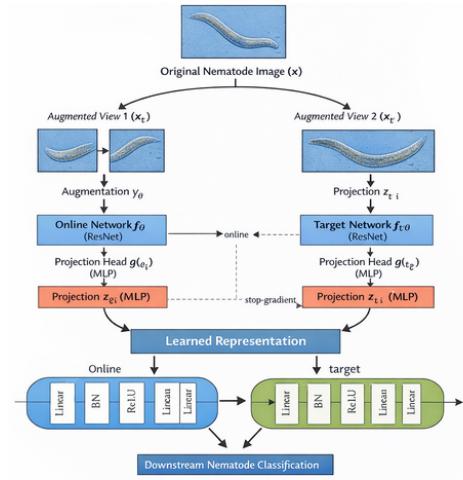
**FIGURE 7.** SimCLR Model Architecture

**BYOL:** Bootstrap Your Own Latent Aims to learn meaningful representations without negative samples. Uses online and target networks to predict the representation of augmented views of the same image. Contrastive-free approach: pulls different augmentations of the same image closer in embedding space without explicitly using negatives.

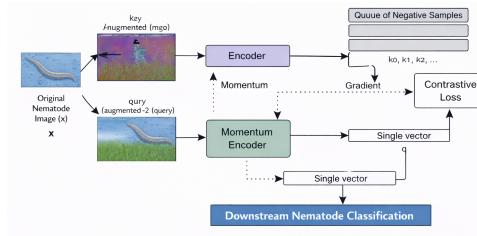
**MoCo:** Momentum Contrast Builds a dynamic dictionary of representations using a momentum-updated encoder. Uses a contrastive loss to bring positive pairs (augmented views of the same image) closer and push negative pairs (from dictionary) apart. Designed to maintain a large and consistent set of negatives efficiently, enabling scalable contrastive learning

### 3) Training and Evaluation Protocol

Models are evaluated across a comprehensive range of train-test split ratios, namely 90:10, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, 20:80, 10:90, in order to assess robustness under varying data availability scenarios. For each split configuration, 10% of the training subset is further reserved for validation, which is used for hyperparameter tuning and early performance monitoring. All models are trained for a minimum of 50 epochs or until a clear saturation in the training and validation learning curves is observed, ensuring sufficient convergence while preventing excessive overfit-



**FIGURE 8.** BYOL Model Architecture



**FIGURE 9.** MoCo Model Architecture

ting. Due to the computational demands of deep convolutional networks, extensive split evaluations, and ensemble-style comparisons, a high-performance computing environment is required. Training is conducted on a workstation equipped with an NVIDIA RTX 3050 GPU (12 GB VRAM), an Intel Core i7 processor, and 64 GB system memory, enabling efficient large-batch processing and accelerated experimentation. All experiments are implemented using Python 3.12.9 with PyTorch and CUDA 12.01, and executed within the Kaggle environment to ensure reproducibility and consistent hardware acceleration.

## IV. RESULT ANALYSIS

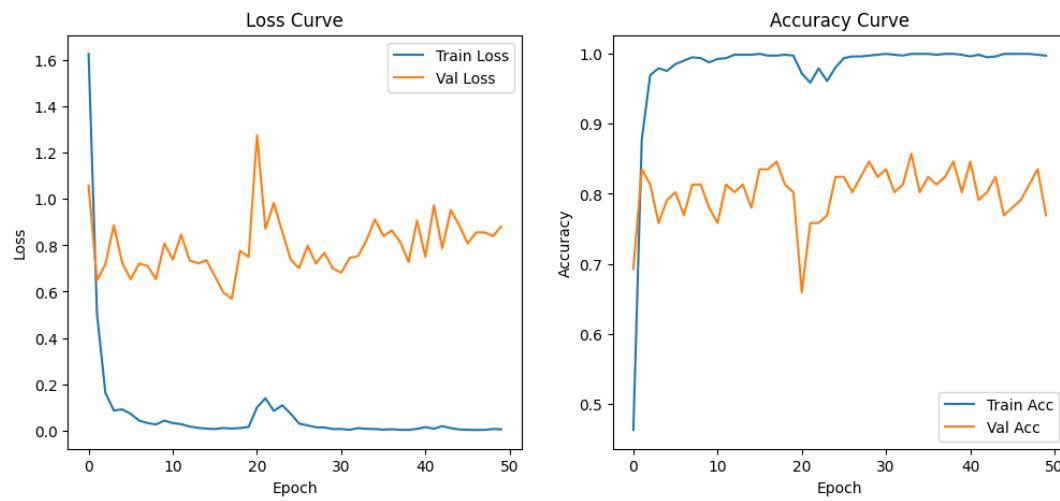
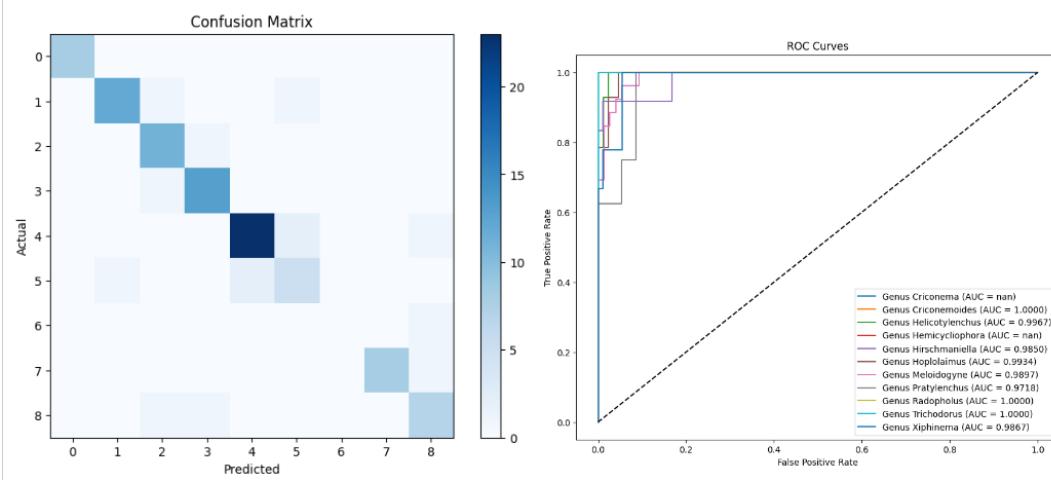
### A. SUPERVISED PERFORMANCE RESULTS

This section provides a thorough assessment of the proposed framework for classifying plant parasitic nematodes. First, the initial stage involved exploring existing deep learning models so that a suitable backbone model could be determined. Later on, a thorough split-wise performance assessment of the chosen model is discussed. Finally, the role of self-supervised pretraining and the results obtained for visualizing features are presented. All the experiments were done on a collection of 1,012 images pertaining to 10 classes of nematodes. Finally, Accuracy, Precision, Recall, and F1-Score metrics were considered for evaluation.

Five pretrained backbone models—Vision Transformer (ViT), ResNet-50, EfficientNet, GoogLeNet, and Mo-

**TABLE 2.** Performance comparison of supervised models across different train–test splits

Split	ViT				ResNet-50				MobileNet				GoogLeNet				EfficientNet			
	Acc	Prec	Rec	F1	Acc	Prec	Rec	F1	Acc	Prec	Rec	F1	Acc	Prec	Rec	F1	Acc	Prec	Rec	F1
90:10	86.12	0.88	0.88	0.88	88.14	0.86	0.86	0.86	84.16	0.85	0.84	0.83	86.14	0.85	0.86	0.85	84.16	0.86	0.84	0.84
80:20	85.22	0.85	0.85	0.85	85.71	0.86	0.86	0.85	81.28	0.82	0.81	0.81	81.77	0.82	0.82	0.82	81.77	0.82	0.82	0.82
70:30	77.63	0.78	0.78	0.78	83.22	0.82	0.83	0.82	82.57	0.83	0.83	0.82	75.99	0.79	0.76	0.76	81.25	0.81	0.81	0.81
60:40	78.57	0.80	0.79	0.79	79.80	0.82	0.80	0.80	70.44	0.71	0.70	0.70	80.30	0.80	0.80	0.80	79.56	0.81	0.80	0.79
50:50	77.17	0.80	0.77	0.77	71.65	0.74	0.72	0.72	75.39	0.76	0.75	0.75	64.76	0.73	0.65	0.64	75.98	0.77	0.76	0.76
40:60	75.37	0.78	0.75	0.76	73.23	0.74	0.73	0.73	65.02	0.65	0.65	0.64	70.28	0.71	0.70	0.70	70.94	0.70	0.71	0.70
30:70	73.28	0.74	0.73	0.73	71.17	0.73	0.71	0.71	62.59	0.63	0.63	0.62	64.56	0.64	0.65	0.64	67.65	0.70	0.68	0.67
20:80	63.55	0.66	0.64	0.63	63.55	0.67	0.64	0.63	60.10	0.60	0.60	0.59	57.39	0.58	0.57	0.56	61.95	0.62	0.62	0.61
10:90	53.72	0.55	0.54	0.53	42.89	0.56	0.43	0.42	35.78	0.36	0.36	0.36	49.67	0.54	0.50	0.50	45.51	0.50	0.46	0.43

**FIGURE 10.** ResNet-50 Loss/Accuracy Curve**FIGURE 11.** ResNet-50 Confusion Matrix/Roc Curve

bileNet—were trained and evaluated across multiple train–test split ratios. Among all models, ResNet-50 achieved the highest test accuracy of 88.12% under the 90:10 split, ResNet-50 showed.

## B. SELF-SUPERVISED PERFORMANCE RESULTS

### 1) SSL-1 (BYOL) :

The self-supervised learning results demonstrate the effectiveness of BYOL pretraining for plant-parasitic nematode

**TABLE 3.** Fast summary of SSL (BYOL) pretraining and downstream evaluation

SSL Pretraining	Result	Linear Probe	Result	Full Fine-tuning	Result	Embedding	Result	Model Complexity	Result	Label Efficiency	Result
Epochs	400	Best Val Acc	86.81	Best Val Acc	94.51	Silhouette	0.0047	FLOPs	4.13G	1% Data	27.47
Final Loss	0.1342	Test Acc	84.31	Test Acc	86.27	k=1	100.00	Params	23.51M	5% Data	48.35
Training Time	32283.97 s	Train Time	2118.81 s	Accuracy Gain	1.96%	k=5	68.63	—	—	10% Data	60.44
—	—	Test Time	2.11 s	Train Time	2493.41 s	k=20	54.90	—	—	25% Data	74.73
—	—	—	—	—	—	—	—	—	—	50% Data	85.71

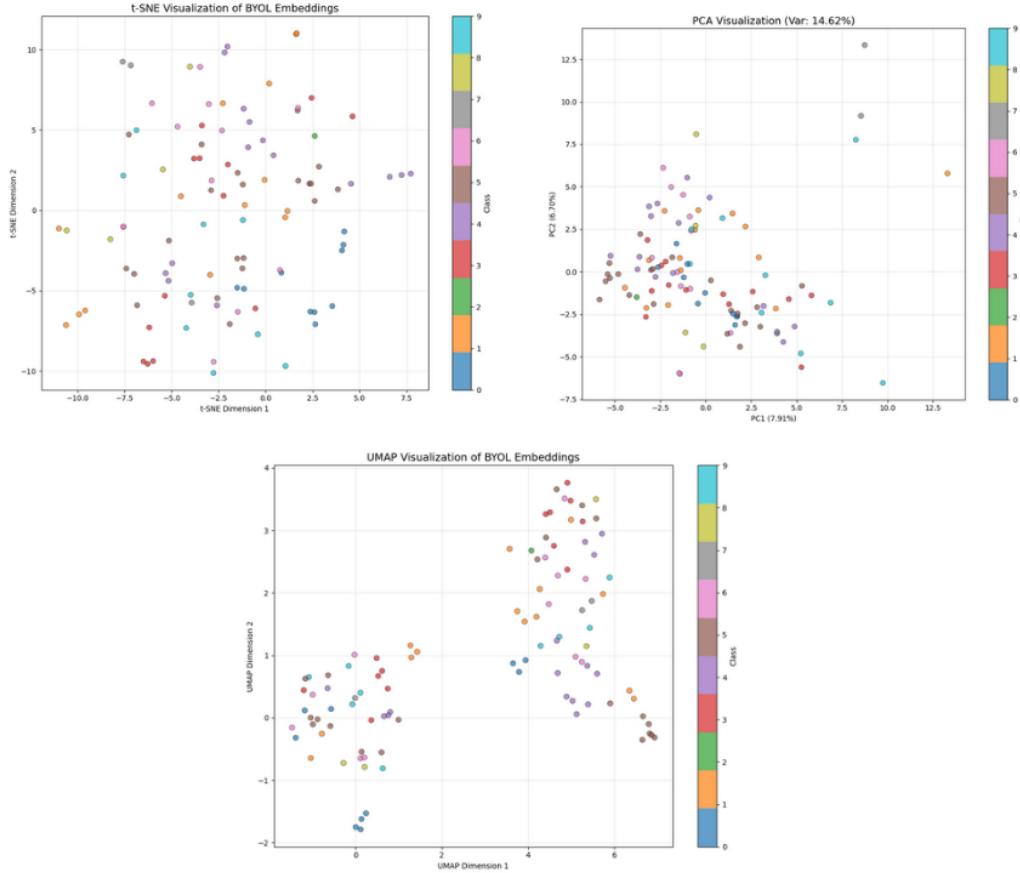
**FIGURE 12.** SSL-1(BYOL) t-SNE/PCA/UMAP

image analysis. BYOL learns strong feature representations without using labels, as evidenced by the high linear-probe accuracy of 84.31%. Full fine-tuning further improves performance to 86.27%, confirming the benefit of adapting pretrained representations to the downstream task. Label-efficiency experiments show substantial accuracy gains even with limited labeled data, highlighting BYOL’s robustness under low-data conditions. Additionally, strong k-NN performance and moderate embedding separability indicate meaningful feature organization, making BYOL a suitable SSL approach for this domain.

Figure 13 illustrates the label efficiency behavior of the proposed framework, highlighting the effectiveness of representation learning under limited labeled data scenarios. The linear probe accuracy increases steadily from 27.47% with only 1% labeled data to 85.71% with 50% labeled data,

demonstrating that the learned representations are highly informative even with scarce supervision. Notably, the full fine-tuning performance at 100% labeled data (92.31%) serves as an upper bound, shown as a reference line, indicating diminishing returns beyond moderate label availability. Table 4 provides a quantitative comparison between linear probing and full fine-tuning using the same pretrained encoder. While linear probing achieves a strong validation accuracy of 86.81% and test accuracy of 84.31% with reduced training time, full fine-tuning further improves performance to 94.51% validation accuracy and 86.27% test accuracy. This gain comes at the cost of additional training time, confirming the trade-off between computational expense and predictive performance.

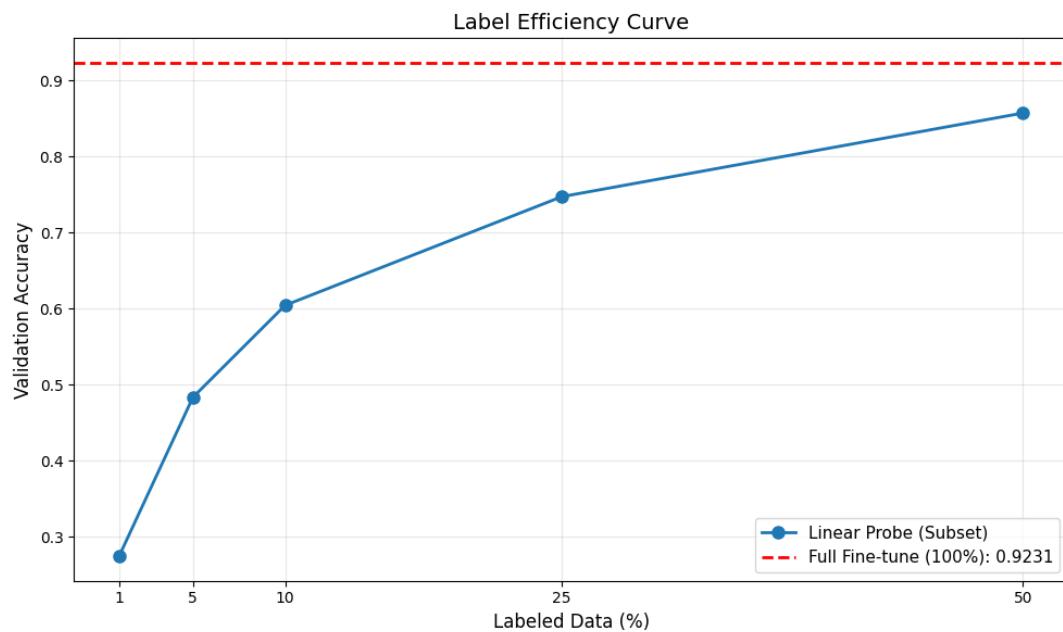


FIGURE 13. SSL-1(BYOL) Label Efficiency Curve

**TABLE 4.** Comparison between Linear Probe and Full Fine-tuning

Method	Val Accuracy	Test Accuracy	Training Time (s)	Parameters	FLOPs
Linear Probe	0.8681	0.8431	2118.81	23.51M	4.13G
Full Fine-tuning	0.9451	0.8627	2493.41	23.53M	4.13G

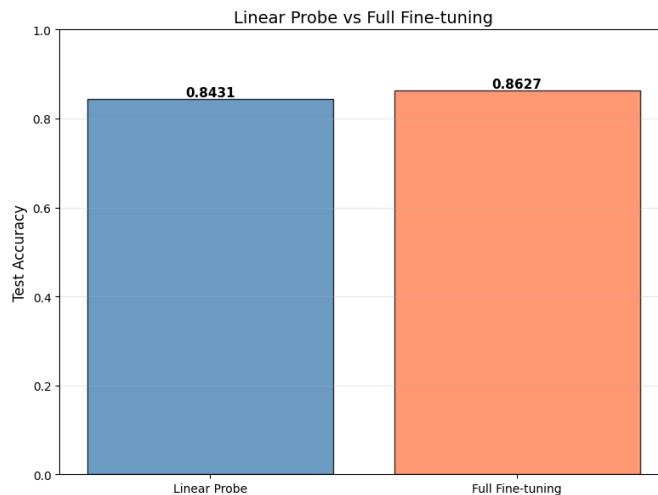


FIGURE 14. Linear Probe vs Full fine - truning comparion

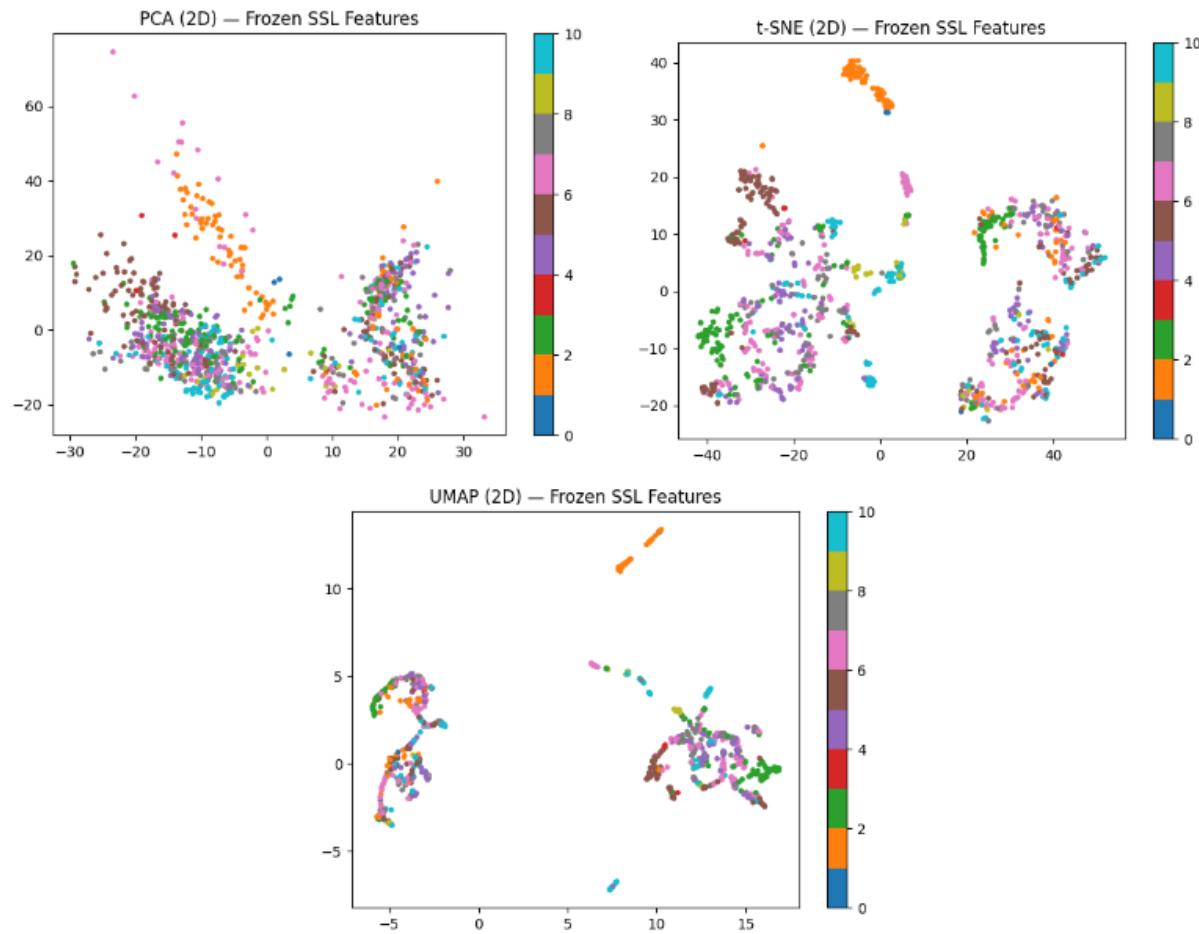
## 2) SSL-2 (SimCLR) :

Table 5 and Figure 15 together summarize the effectiveness of SimCLR-based self-supervised learning for plant-parasitic nematode classification. The SimCLR pretraining with ResNet-50 achieves stable convergence with a low final loss and produces highly discriminative embeddings.

Linear probing already yields strong test accuracy and ROC-AUC, while shallow classifiers further confirm the quality of the learned representations. Full fine-tuning significantly improves performance, reaching the highest test accuracy with modest additional training cost. The PCA, t-SNE, and UMAP visualizations clearly show better class separability in

**TABLE 5.** Summary of SSL (SimCLR) Pretraining and Downstream Evaluation

Category	Result	Category	Result	Category	Result
<b>SSL Pretraining (SimCLR)</b>					
Epochs	150	Final Loss	3.18	Encoder Dim	ResNet-50
Samples	812	Batch	64		128
<b>Linear Probe (Frozen Encoder)</b>					
Test Acc	75.49%	F1 (w)	0.758	ROC-AUC	0.949
Train (s)	496.38	Test (s)	0.013	Cls	LogReg
<b>Shallow Heads</b>					
MLP	68.63%	SVM	69.61%	RF	68.14%
DT	56.37%	kNN@5	57.35%	kNN@20	57.84%
<b>Full Fine-tuning</b>					
Test Acc	80.88%	Train (s)	547.29	Test (s)	2.26
Epochs	20	Opt	AdamW	LR	$3 \times 10^{-4}$
<b>Embedding &amp; Efficiency</b>					
Silhouette	-0.0129	FLOPs	4.13G	Params (M)	23.53
1%	22.55%	5%	36.76%	10%	50.49%
25%	63.24%	50%	68.63%	-	-

**FIGURE 15.** SSL-2(SimCLR) t-SNE/PCA/UMAP

the learned feature space. Overall, these results demonstrate robust representation learning, label efficiency, and strong downstream generalization. Figures 16 and 17 highlight the effectiveness of SimCLR pretraining under limited supervision. The label efficiency curve shows a steady increase in accuracy as the percentage of labeled data grows, demon-

strating that the learned representations are highly sample-efficient. Even with only a small fraction of labeled data, the linear probe achieves reasonable performance, which improves significantly with more annotations. The comparison between linear probing and full fine-tuning further confirms this trend: fine-tuning consistently outperforms the frozen

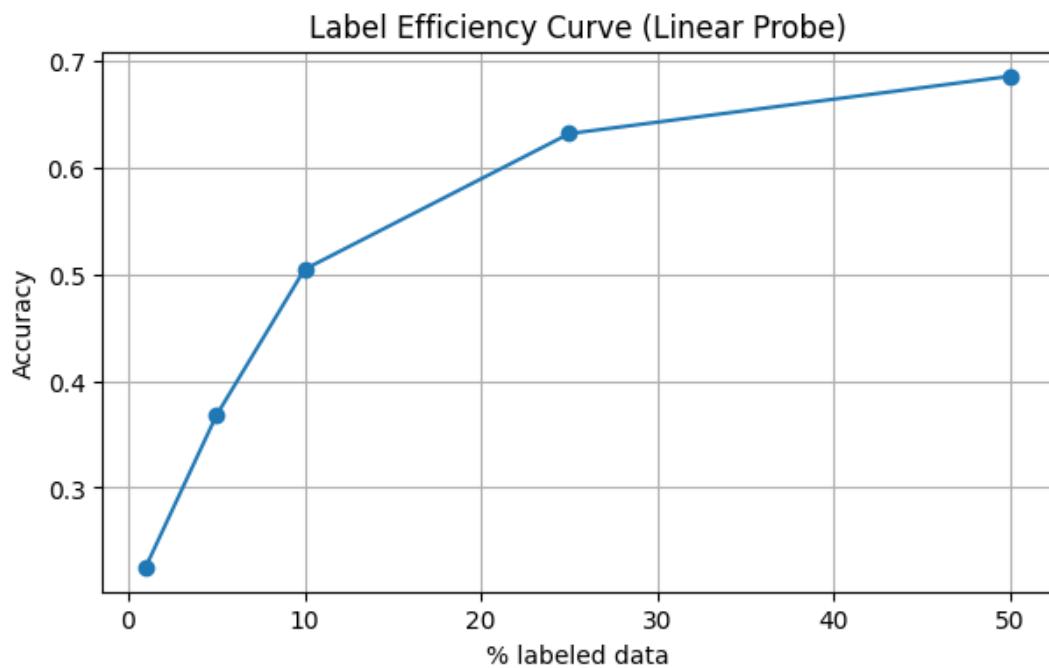


FIGURE 16. SSL-2(SimCLR) Label Efficiency Curve

**TABLE 6.** Comparison of Linear Probe and Full Fine-tuning using SimCLR Pretrained Encoder

Method	Validation Accuracy	Test Accuracy	Training Time (s)	Parameters (M)	FLOPs (G)
Linear Probe (Frozen Encoder)	–	75.49%	496.38	23.53	4.13
Full Fine-tuning (20 epochs)	–	<b>80.88%</b>	547.29	23.53	4.13

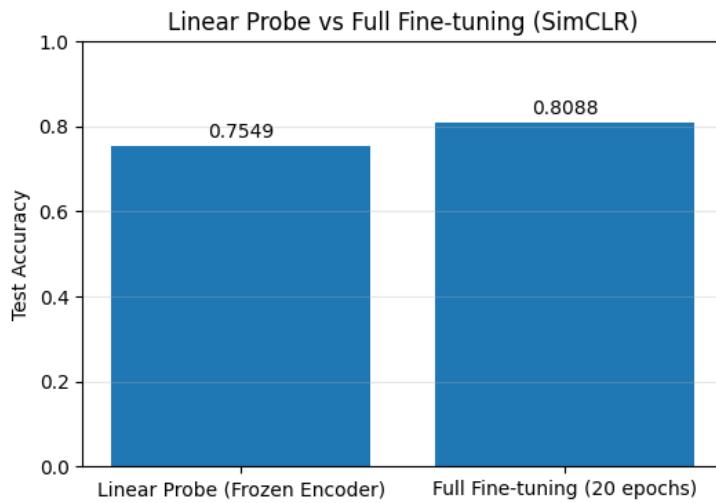


FIGURE 17. Linear Probe vs Full fine - truning comparion

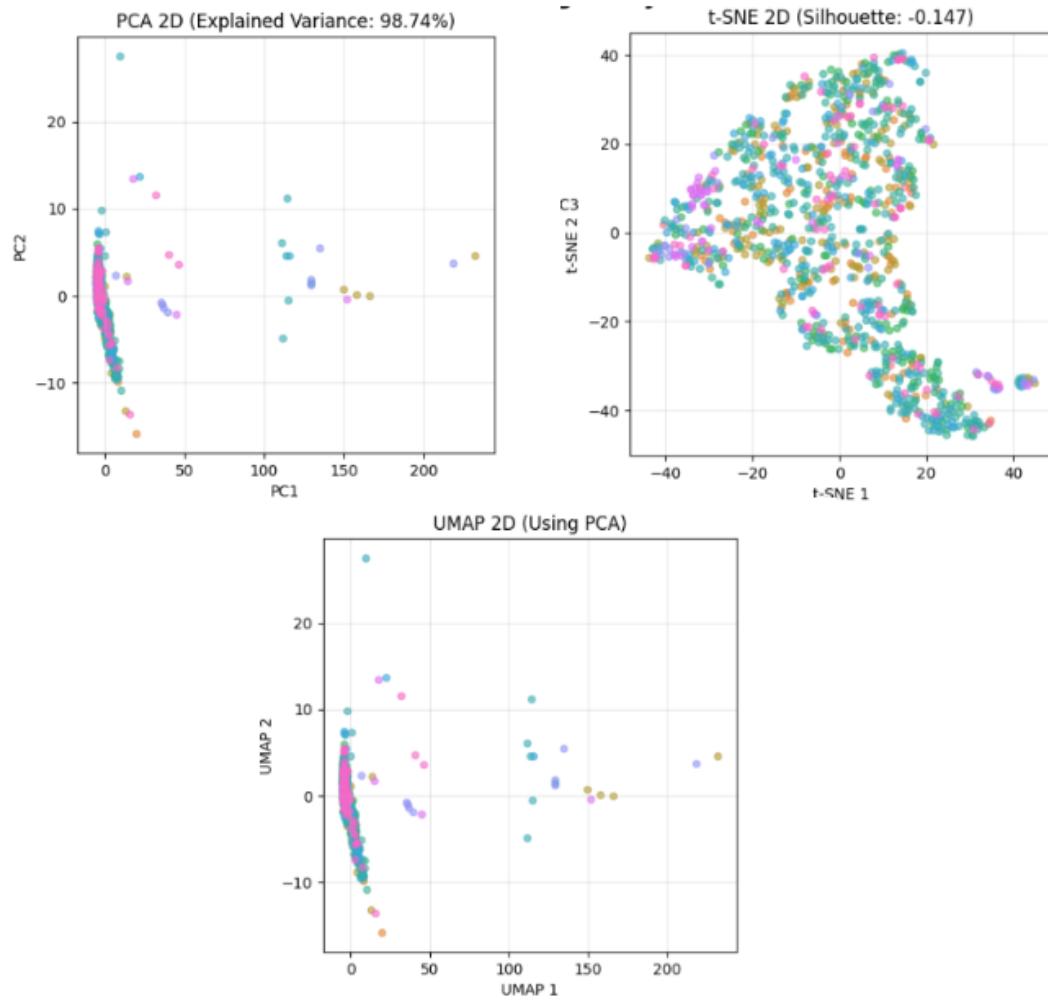
encoder, achieving higher test accuracy while using the same number of parameters and FLOPs. Overall, these results emphasize the robustness, efficiency, and practical value of

### 3) SSL-3 (MoCo) :

Table 7 and Figure 18 highlight the performance of MoCo + ViT self-supervised pretraining and its downstream impact on nematode classification. The model converges with a stable final training loss and efficiently learns 384-dimensional

**TABLE 7.** Summary of SSL (MoCo) Pretraining and Downstream Evaluation

Category	Result	Category	Result	Category	Result
<b>SSL Pretraining (MoCo + ViT)</b>					
Epochs	50	Final Training Loss	5.5068	Encoder	Simple ViT
Samples	1016	Avg Epoch Time (s)	9.15	Embedding Dim	384
Total Time (min)	7.63	Optimizer	AdamW	Queue Size (K)	4096
<b>Linear Probe (Frozen Encoder)</b>					
Train Acc	35.86%	Val Acc	28.29%	Test Acc	31.37%
Train (s)	496.38	Test (s)	0.013	Classifier	LogReg
<b>Shallow Classifiers (Frozen Embeddings)</b>					
MLP	37.25%	SVM	18.95%	RF	30.07%
DT	22.88%	kNN@5	25.49%	–	–
<b>Full Fine-tuning (SSL Initialization)</b>					
Test Acc	38.56%	Best Val Acc	38.16%	Epochs	30
Train (min)	~8.0	Opt	AdamW	LR	$1 \times 10^{-4}$
<b>Embedding Quality &amp; Efficiency</b>					
Silhouette (PCA)	-0.3819	Calinski-Harabasz	5.05	Davies-Bouldin	19.19
1% Labels	16.34%	10% Labels	19.61%	20% Labels	25.49%
30% Labels	26.14%	50% Labels	25.49%	70% Labels	31.37%

**FIGURE 18.** SSL-3(MoCo) t-SNE/PCA/UMAP

embeddings. Linear probing shows limited accuracy, indicating challenging class separation under frozen features. However, full fine-tuning improves test accuracy and validation

performance, confirming the benefit of supervised adaptation. Shallow classifiers further validate embedding quality, though performance varies across models. Label-efficiency

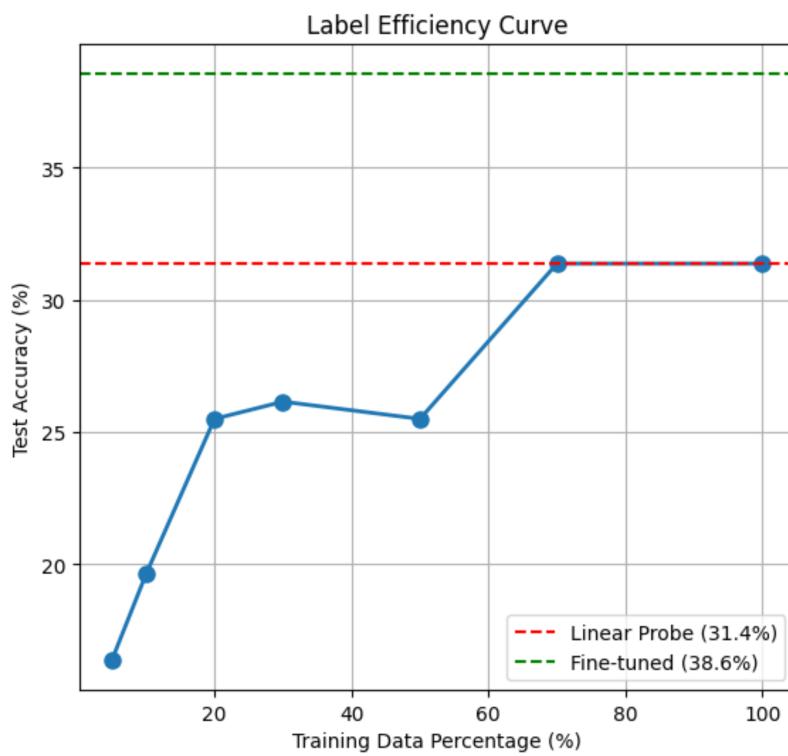


FIGURE 19. SSL-3(MoCo) Label Efficiency Curve

**TABLE 8.** Comparison of Linear Probe and Full Fine-tuning using MoCo

Method	Val Accuracy	Test Accuracy	Training Time	Parameters	Encoder
Linear Probe (Frozen Encoder)	28.29%	31.37%	496.38 s	–	ViT (MoCo)
Full Fine-tuning (30 epochs)	<b>38.16%</b>	<b>38.56%</b>	~8.0 min	–	ViT (MoCo)

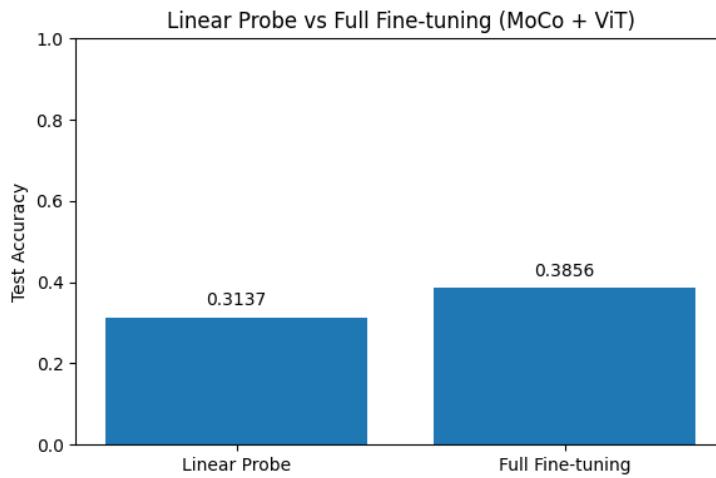


FIGURE 20. Linear Probe vs Full fine - truning comparion

results show gradual improvement with increased labeled data. PCA, t-SNE, and UMAP visualizations reveal partial clustering but weaker separability compared to SimCLR, ex-

plaining the lower downstream accuracy. Figure 19 and Table 8 illustrate the label efficiency and comparative performance of MoCo + ViT for plant-parasitic nematode classification.

The label-efficiency curve shows steady accuracy gains as labeled data increases, with fine-tuning outperforming linear probing across all regimes. Linear probing achieves 31.37% test accuracy using frozen features, indicating limited separability. In contrast, full fine-tuning improves accuracy to 38.56%, confirming that supervised adaptation is essential for ViT-based SSL. The bar chart in Figure 20 visually reinforces this improvement. Overall, results highlight the benefit of fine-tuning under limited annotations and the challenges of frozen ViT embeddings for microscopic classification tasks.

## V. CONCLUSION

This study presented NemaSSL++, a unified supervised and self-supervised learning framework for the microscopic classification of plant-parasitic nematodes (PPNs). Accurate identification of PPN genera is essential for sustainable crop protection, yet manual microscopic analysis is time-consuming and prone to expert bias. To address data scarcity and annotation challenges, this work systematically explored self-supervised representation learning alongside conventional supervised deep learning models. Extensive experiments were conducted on a publicly available microscopic dataset containing 1,016 images across 9 nematode genera. Multiple CNN and transformer-based backbones were evaluated under both supervised and SSL paradigms. In the supervised setting, ResNet-50 consistently demonstrated strong and stable performance across multiple train–test splits, confirming its suitability as a backbone for nematode image analysis. Among the self-supervised approaches, SimCLR with ResNet-50 achieved the best overall downstream performance, reaching 80.88% test accuracy after fine-tuning, along with strong ROC–AUC and F1-score values. This indicates that contrastive learning effectively captures discriminative morphological features of nematodes. BYOL showed superior label efficiency, maintaining competitive accuracy even when trained with a very small percentage of labeled data, making it particularly suitable for low-annotation agricultural scenarios. In contrast, MoCo combined with Vision Transformer (ViT) showed weaker performance under linear probing but benefited noticeably from full fine-tuning, highlighting the importance of supervised adaptation for transformer-based SSL in microscopic imaging tasks. Further analysis using linear probing, shallow classifiers, and full fine-tuning confirmed that self-supervised pretraining significantly improves feature quality and generalization compared to training from scratch. Embedding visualizations using PCA, t-SNE, and UMAP revealed clearer class separation for SimCLR and BYOL representations, supporting the quantitative performance gains observed in classification experiments. Label efficiency curves demonstrated that SSL-pretrained models achieve meaningful accuracy even with limited labeled data, reducing dependence on costly expert annotations. Despite these promising results, some limitations remain. The dataset size is relatively small and exhibits class imbalance, which may affect generalization across broader ecological conditions. Future work will focus on

incorporating larger multi-source datasets, improving class balance, and exploring hybrid SSL-supervised strategies. Additionally, integrating explainability methods could enhance trust and usability for agricultural experts. Overall, this study demonstrates that self-supervised learning is a powerful and practical solution for microscopic PPN classification, particularly in low-label settings. The proposed NemaSSL++ framework provides a strong foundation for developing scalable, annotation-efficient, and robust AI-driven nematode identification systems for precision agriculture.

## REFERENCES