TWFSL-MM: Few-Shot Learning using Meta-Learning and Metric-Learning for Disease Detection in Azadirachta Indica

H. A. Vidya
Department of Computer Science and Engineering, Kalpataru Institute of Technology, Visvesvaraya Technological University, Belagavi-590018, India
vidya.ha@gmail.com (corresponding author)

M. S. Narasimha Murthy
Department of Information Science and Engineering, BMS Institute of Technology and Management, Autonomous Institution Under Visvesvaraya Technological University, Belagavi-590018, India
narasimhamurthyms@bmsit.in

Received: 9 December 2024 | Revised: 3 January 2025 and 22 January 2025 | Accepted: 27 January 2025

 DOI: https://doi.org/10.48084/etasr.9886

ABSTRACT

Few-Shot Learning (FSL) is one of the emerging and promising approaches used in machine learning for image classification and prediction. This work proposes a Two-Way Five-Shot Learning with Meta-learning and Metric-learning (TWFSL-MM) model that can detect plant diseases with limited data, reducing the cost of implementation and improving the quality of Azadirachta Indica. The proposed method addresses the drawbacks of FSL by employing meta-learning and metric-learning approaches. Experimental results showed that the proposed model achieved an accuracy of 92.09%, an average loss of 0.18, an average precision of 0.94, a recall of 0.93, and an F1 score of 0.93. FSL is a promising strategy for plant disease detection, achieving higher accuracy with a limited dataset. The TWFSL-MM model outperforms other state-of-the-art models, demonstrating its potential to improve crop yields and quality.

Keywords-deep learning; few-shot learning; metric-learning; meta-learning

I. INTRODUCTION

Humans grow plants for food, clothing, medicine, shelter, etc. However, plants are affected by diseases that occur due to various factors, such as biological or environmental factors. Biologically, plant diseases are caused by fungi, bacteria, and viruses. Environmental conditions, such as temperature, humidity, rainfall, and soil, affect plants. Plant diseases reduce crop yields and quality or may even cause plants to die [1]. They also cause agricultural and economic hindrances [2]. Hence, plant disease detection is one of the prominent fields in which researchers are working today. Azadirachta indica is the scientific name of the neem tree, which is one of the vital plant species with high medicinal values used in the treatment of various diseases in Ayurveda. It is used in the preparation of drugs for skin problems, dental problems, and diabetes. Neem extracts are also used in the preparation of oil and cosmetics [3]. It is also used in pesticides and insecticides to treat plant diseases [4]. Artificial intelligence and deep learning play a significant role in almost all fields to solve problems [5]. Nowadays, the use of DL techniques is promising and reduces the cost of labor, improving the quality and yield of crops vital in agriculture.

Deep learning models require large amounts of data to learn features to identify and classify diseases. In the current state of the art, few paradigms in artificial intelligence work with small amounts of data, one of them being Few-Shot Learning (FSL) [6]. FSL is one of the emerging and promising approaches used in machine learning for image classification and prediction [7]. FSL takes labeled data as input and produces the classification output using probabilities of the classes used, dividing the data into two sets, namely a support set and a query set. The support set contains a few samples per new class of data. The query set contains existing and new classes of data on which the model generalizes using the knowledge gained from the support set [8]. FSL is an N-way K-shot learning approach where N-way indicates the number of new classes used in the support set on which any pre-trained model generalizes. K-shot specifies the number of data samples for each class in the support set. As the value of N increases, complexity also increases, and a lower value of K results in lower accuracy. Typical values for K range from 1 to 5. If K is zero, it is a zero-shot learning, an unsupervised learning technique [9]. One-shot learning [10] is a variation in FSL.

FSL is important for several reasons. First, it can use a few labeled samples with satisfactory results. Second, it saves time and power by using a pre-trained model. Third, endangered plant and animal species can be studied with limited samples. However, FSL may have limited applicability and overfitting problems. Researchers are working on these limitations and exploring various techniques, such as data augmentation, meta-learning, and regularization [11]. FSL uses the similarity function to identify similarities between the support and query set images [12]. It uses probability scores as output, where the class with the highest similarity score is the one predicted by the model.

In [13], FSL was employed to address the issue of large dataset requirements using traditional deep learning techniques. The study used a Siamese network with a triplet loss function to identify new plant leaves with the reduced dataset, achieving an accuracy of 90% with 80 images per new class. In [14], a novel semi-supervised FSL approach was proposed to improve the accuracy of plant leaf disease classification. This method leveraged unlabeled data to enhance the learning process, making it more efficient. The average improvement was 2.8% using a single semi-supervised method and 4.6% using an iterative method. In [15], Pre-training, Meta-learning, and Fine-tuning with Feature Attention (PMF+FA) were used to improve the accuracy of FSL methods. This approach achieved 90% accuracy on the PlantDoc dataset. Future work implied exploring different FSL methods and more diverse datasets. In [16], a novel Background-Filtering Feature-Enhanced Graph Network (BFFE-GNN) was proposed, using two publicly available datasets, namely MiniImageNet and TiredImagenet. Experiments were carried out for 1-shot and 5-shot learning, achieving 61.83% and 79.01% accuracy, respectively. In [17], a Biased-Reduction Alternative Network (BRAVE) was proposed, which had a vector quantized variational autoencoder. This study investigated 1-shot and 5-shot scenarios using the ResNet-12 pre-trained model, achieving 68.55% accuracy on the MiniImagenet dataset for 1-shot and 88.93% for 5-shot learning, and 73.79% for 1-shot and 89.05% for 5-shot learning on the TiredImageNet dataset.

Previous works on FSL have used public datasets, and existing state-of-the-art FSL models show accuracy up to 90%. Most FSL models are based on the Siamese network [18]. Deep learning techniques require a large amount of data to train the model for classification. An FSL model can accomplish the task of classification and prediction of diseases with limited data. The contributions of this work include:

- Collection of data samples in the region of Tiptur, Tumkur, and Mysuru, Karnataka, India along with images in public datasets.
- Preprocessing and augmenting data to increase the number of samples, thus addressing the overfitting problem.
- Addressing the drawbacks of existing FSL methods by employing meta-learning and metric-learning approaches.

Normally, FSL uses a Siamese network and the triplet loss function to find the similarity. There are four different approaches in FSL, namely, data level, parameter level, metric level, and gradient-based meta-learning. Combining meta-learning and metric-learning approaches is a new paradigm that provides improved generalization, faster adaptation, and improved performance for real-time data. The objective of this work is to develop an FSL model that can detect plant diseases with limited data, reducing the cost and improving the quality of crops by combining meta-learning and metric-learning.

II. PROPOSED WORK

The proposed TWFSL-MM model was designed for a dataset consisting of two classes: healthy and diseased. It utilizes support and query set images by randomly selecting five images from each class, resulting in 10 images (5 healthy and 5 diseased) for both the support and query sets, totaling 20 images. These support and query images are randomly chosen from the dataset and processed in each batch during the epochs, ensuring that all images are utilized over time. The model effectively leverages the entire dataset across multiple iterations. The dataset used is a mixture of publicly available data from [19, 20] and real-time images, consisting of healthy and diseased classes. As data for diseased neem leaves are not publicly available, real-time data were collected, including both leaves and trees. Healthy and diseased images were captured using a high-resolution camera in the local areas of Tiptur, the forest area of Tumkur, and Mysuru, Karnataka, India. The dataset contains two classes with 2000 healthy and 1800 diseased images. The data collected was preprocessed by applying image resizing and image denoising techniques to enhance the quality of images and augment the dataset. The total number of images used for training and testing was 3800. 80% of the data were considered for training and 20% for testing and validation.

Figure 1 shows a flow diagram of the proposed approach. In the initial step, the model takes the neem leaf image dataset as input. The dataset is divided into training, testing, and validation. Feature extraction is performed by the ResNet-50 pre-trained model to extract features from the images. The training dataset is split into a support set and a query set in the meta-learning process. The support set is used to train the model, and the query set is used for classification. Training the model includes meta-learning, where it learns to adapt to new tasks with limited data, and metric learning, where it calculates the similarity between features of support and query images. The model then classifies the query images as healthy or diseased based on learned features and distances. The model was evaluated by comparing the predicted with the actual labels of the query images. If the classification is correct, the process ends, and if the classification is incorrect, the model is further trained to improve its accuracy. With these steps, the model can detect diseases in neem trees even with limited training data. Figure 2 shows the architecture of the proposed TWFSL-MM model. Its key components are input data, feature extractor, flatten layer, fully connected layer, meta-learning module, metric learning module, and final classifier layer. The input data is a set of images with dimensions 224×224×3, where the resolution is 224×224 with three RGB channels. The feature extractor used is the pre-trained ResNet-50 model with the last layer removed to make it compatible with the meta-learning process. The final convolutional layer of ResNet-50 produces a feature map of 7×7×2048, which means 7×7 spatial size and 2048 channels.

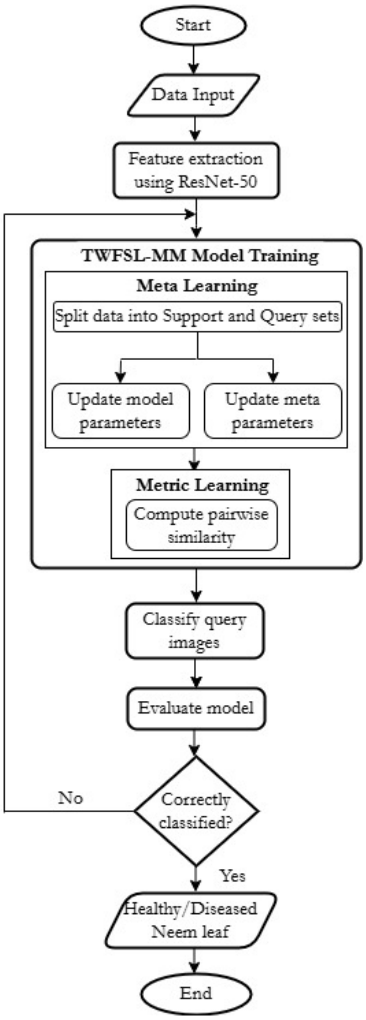

Fig. 1. Flowchart of the proposed TWFSL-MM model.

The flattened layer takes the 3D feature map and converts it to a 1D feature vector of size 2048. The fully connected layer reduces the dimensionality of the feature vector from 2048 to 256 by applying a linear transformation. This is a form of regularization that helps to prevent overfitting and improve the generalization of the proposed model. This also reduces computational complexity and improves the overall performance of the meta-learning process. Training data is split into support and query sets in the meta-learning process of the model. The meta-learning process includes an inner loop and an outer loop. The support set contains images used to update the model parameters during the inner loop of the meta-learning process, and the query set contains images that are used to update the meta-parameters during the outer loop. The model parameters are the weights and biases of the neural network used in ResNet-50, fully connected, and classification layers. Meta-parameters are the learning rate of the optimizer. The inner loop uses the Stochastic Gradient Descent (SGD) optimizer, and the outer loop uses the Adam optimizer.

The metric-learning module in the architecture is the process of learning suitable similarity metrics or distance functions to measure the relationship between input samples. Similarity is found using the Euclidean distance function [21]. Metric-learning is often guided by a contrastive cross-entropy loss function that helps the model learn a discriminative feature representation. The final classification layer is responsible for classifying the images into the target classes: healthy and diseased. The reduction in dimensionality from 256 to 2 aligns the output of the model with the number of target classes in the problem. Algorithm 1 describes in pseudocode the entire process of the TWFSL-MM model.

The model starts by initializing the ResNet-50 model. Loss is calculated in the inner and outer loops using the cross entropy loss function:

$$Loss = -\sum\left(y_i * log(p_i) + (1-y_i) * log(1-p_i)\right) \quad (1)$$

where y_i is the true label for the i^{th} class (0 or 1) and p_i is the predicted probability for the i^{th} class.

Initialization of the optimizer for the inner loop is performed using the SGD optimizer (2). The update rule for a parameter θ using SGD is:

$$\theta = \theta - n * \nabla\theta L \quad (2)$$

where θ is the parameter to be updated, n is the learning rate controlling each step size of the update, and $\nabla\theta L$ is the gradient of the loss function L for the parameter θ.

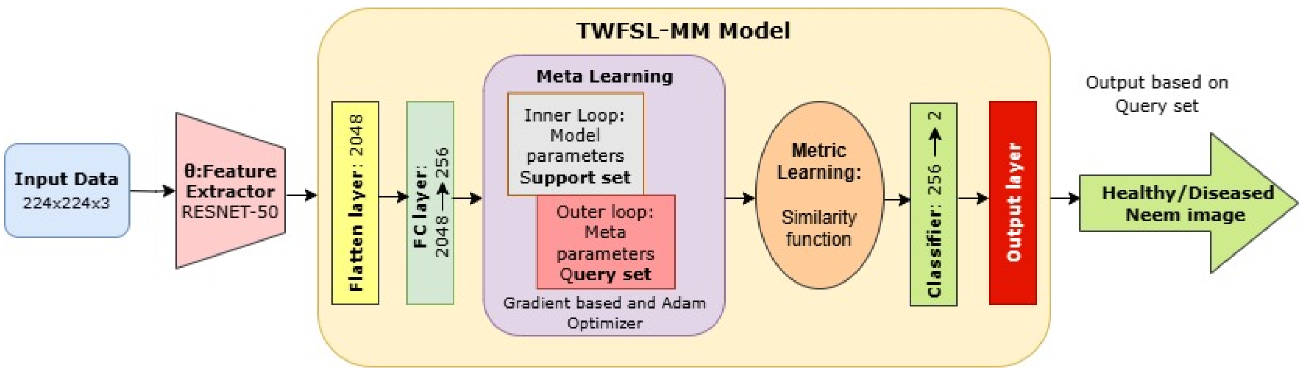

Fig. 2. The architecture of the proposed TWFSL-MM model.

Algorithm 1: Proposed TWFSL-MM model
Input: Training data (a dataset of images and their corresponding labels), Number of epochs, Batch size, Inner loop iterations, Learning rate for inner and outer optimizers, Model architecture ResNet-50
Output: Trained Model
Model:
  Initialize model ResNet-50
  Initialize Adam optimizer
  For each epoch:
    For each batch:
      Split batch into support set S and query set Q
      Inner Loop:
      Initialize inner optimizer
      For each inner loop iteration:
        Forward pass on support set
        $Logits_S = model(S)$
        Calculate loss:
        $Loss_S = CrossEntropyLoss(logits_S, labels_S)$
        Backward pass:
        Update model parameters using gradients of $loss_S$ and inner optimizer
      Outer Loop:
      Forward pass on query set:
      $Logits_Q = model(Q)$
      Calculate loss:
      $Loss_Q = CrossEntropyLoss(logits_Q, labels_Q)$
      Backward pass:
      Update meta-parameters using gradients of $loss_Q$ and meta-optimizer
      Similarity-based classification:
      Extract feature embeddings for support and query sets
      Calculate pairwise Euclidean distances
      Assign class of nearest support image to each query image
    Evaluation of model on validation set:
    Calculate and print accuracy, precision, recall, and F1-score

Initialization of the optimizer for the outer loop is performed using the Adam optimizer given. The update rule for a parameter θ using Adam is:

$$m_t = \beta_1 * m_{t-1} + (1 - \beta_1) * \nabla\theta L \quad (3)$$

$$v_t = \beta_2 * v_{t-1} + (1 - \beta_2) * (\nabla\theta L)^2 \quad (4)$$

$$m_{t_{hat}} = m_t / (1 - \beta_1^t) \quad (5)$$

$$v_{t_{hat}} = v_t / (1 - \beta_2^t) \quad (6)$$

$$\theta = \theta - n * m_{t_{hat}} / \sqrt{v_{t_{hat}} + \epsilon} \quad (7)$$

where θ is the parameter to be updated, n is the learning rate, β_1 and β_2 are the hyperparameters controlling the decay rates of the first- and second-moment estimates, m_t and v_t are the first- and second-moment estimates of the gradients, $m_{t_{hat}}$ and $v_{t_{hat}}$ are the biases correct first- and second-moment estimates, and ϵ is a small constant to prevent division by zero.

Metric learning is carried out using the similarity-based Euclidean distance function:

$$distance(x_i, x_j) = \sqrt{\sum_{i=0}^{d} (x_{i_k} - x_{j_k})^2} \quad (8)$$

where x_i is the feature embedding of the i^{th} query image, x_j is the feature embedding of j^{th} support image, d is the dimensionality of the feature embeddings, x_{i_k} is the k^{th} element of the feature vector x_i, and x_{j_k} is the k^{th} element of the feature vector x_j. The output metrics considered are accuracy, precision, recall, and F1-score, calculated as follows.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

$$Precision = \frac{TP}{TP+FP} \quad (10)$$

$$Recall = \frac{TP}{TP+FN} \quad (11)$$

$$F1 - Score = 2 * \frac{(Precision * Recall)}{Precision + Recall} \quad (12)$$

where TP denotes true positives, TN denotes true negatives, FP denotes false positives, and FN denotes false negatives.

III. RESULTS

The proposed model achieved promising results on the hybrid dataset, outperforming other state-of-the-art models by a significant margin. The model's innovative architecture, combining both meta-learning and metric-learning techniques, enabled it to effectively capture complex patterns within the data. The support set and query set images are displayed as shown in Figures 3 and 4.

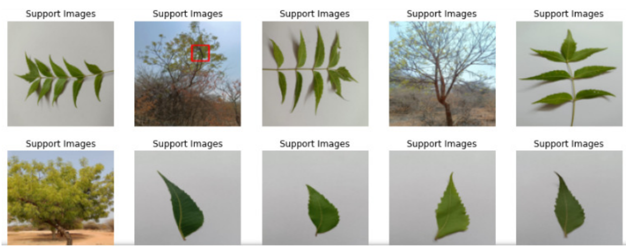

Fig. 3. Support set images

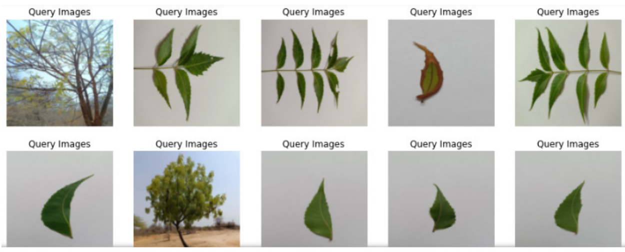

Fig. 4. Query set images.

```
Support Labels: [0, 0, 0, 0, 0, 1, 1, 1, 1, 1]
Query Labels: [0, 0, 0, 0, 0, 1, 1, 1, 1, 1]
Ground Truth / Predicted
diseased / healthy
diseased / diseased
diseased / diseased
diseased / diseased
diseased / diseased
healthy / healthy
healthy / healthy
healthy / healthy
healthy / healthy
healthy / healthy
```

Fig. 5. Ground truth and predicted values of two classes.

(a) (b) (c) (d)

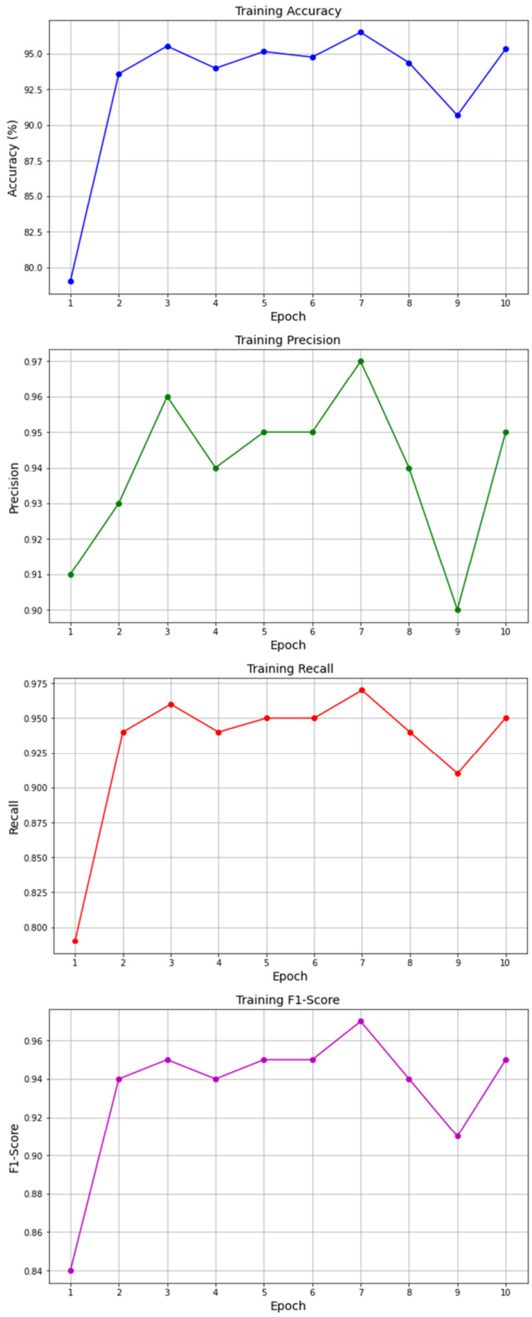

Fig. 6. Training results for the proposed TWFSL-MM model: (a) Accuracy, (b) Precision, (c) F1-score, (d). Recall.

The model ran for 10 epochs in different splits with a batch size of 64 and a learning rate of 0.001, demonstrating 92.09% accuracy, 0.18 average loss, 0.94 average precision, 0.93 recall, and 0.93 F1-score. Figure 5 shows the ground truth and the predicted values of the two neem classes (healthy and diseased). The training results are shown in Figure 6 and the validation results are shown in Figure 7.

(a) (b) (c) (d)

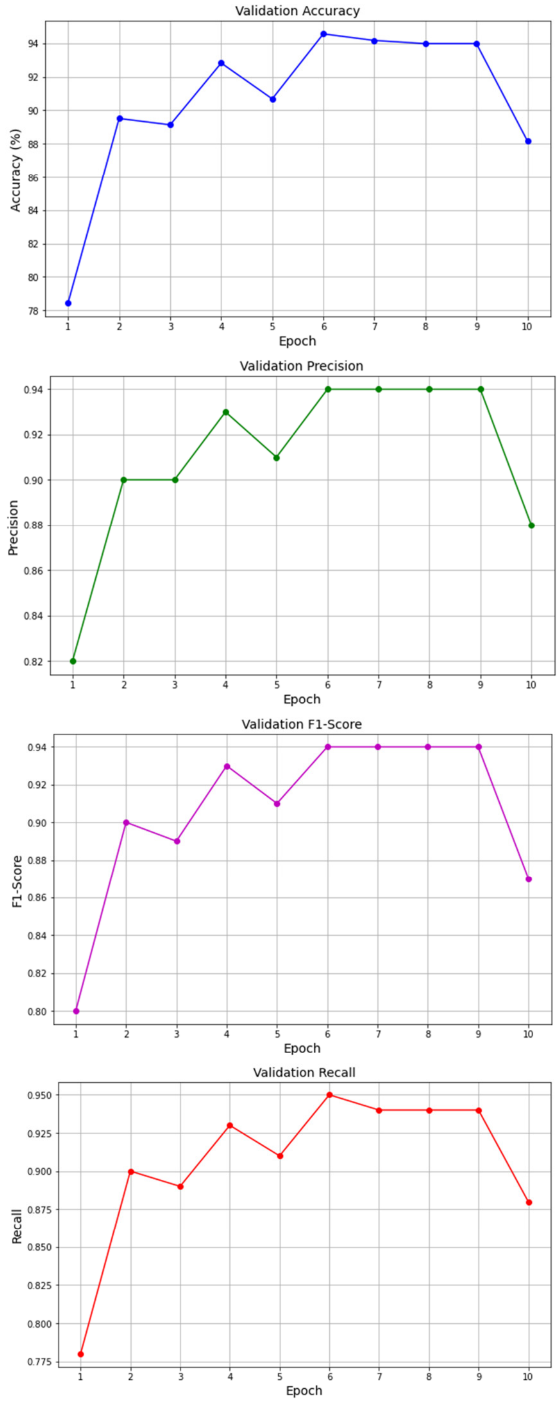

Fig. 7. Validation results for the proposed TWFSL-MM model: (a) Accuracy, (b) Precision, (c) F1-score, (d) Recall.

The proposed model was compared with other related works, as seen in Table I, showing its superiority.

TABLE I. COMPARATIVE ANALYSIS OF THE PROPOSED WITH EXISTING MODELS

Study	Dataset	Model	#Shots	Accuracy
[7]	Plant Village Dataset	Meta-learning	5	83.25%
[15]	Plant Village, Plant Doc dataset	PMF+FA	40	90.12%
[16]	MiniImageNet, TiredImageNet, Tool dataset	Background-Filtering Feature-Enhanced GNN for FSL	5	58.68%
[21]	Plant Village Dataset	Semi-supervised FSL	5	90%
[22]	Plant Village Dataset, FGVC8 dataset	MAFDE-DN4	1 5	57.5% 81.41%
[23]	Not mentioned	AlexNet GoogleNet	5 5	42.53% 44.56%
[24]	Tomato weeds dataset	ResNet50	0	77.8%
This study	Hybrid dataset	TWFSL-MM	5	92.09%

Table I shows that most studies used 5 shots, achieving accuracies up to 90%. The proposed model had superior performance compared to the other few-shot learning models. Figure 9 shows an analysis chart. As the ideal values for *K* range from 1 to 5, 5 shots were considered in the proposed model to achieve improved accuracy.

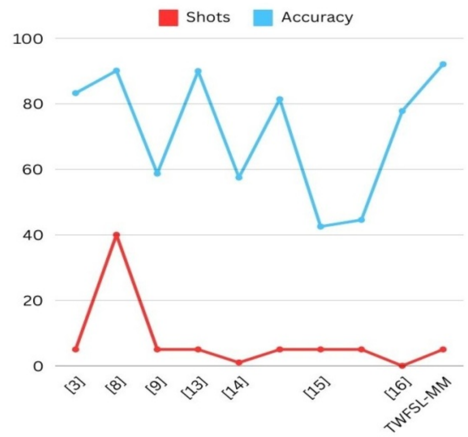

Fig. 8. Comparative analysis of the proposed model with related works.

IV. CONCLUSION

Deep learning models require large amounts of data for effective processing. However, collecting data for species of plants and animals, particularly those in danger of extinction, poses significant challenges. In such cases, where data collection becomes a bottleneck, innovative techniques are required. FSL is a promising approach that works for limited data. The proposed TWFSL-MM model outperformed other state-of-the-art FSL models, demonstrating its potential to improve crop yields and quality. The proposed model achieved 92.09% accuracy, 0.18 average loss, 0.94 average precision, 0.93 recall, and 0.93 F1 score. The novelty of this work is the combination of meta-learning and metric-learning on hybrid data that helps improve generalization, adaptation, and performance by leveraging the strengths of both techniques. This model shows an improvement in accuracy compared to other existing FSL methods. In the future, the model can be implemented to detect different plant diseases, and other approaches can be combined to further improve the accuracy of FSL models.

REFERENCES

[1] Y. Gai and H. Wang, "Plant Disease: A Growing Threat to Global Food Security," *Agronomy*, vol. 14, no. 8, Aug. 2024, Art. no. 1615, https://doi.org/10.3390/agronomy14081615.

[2] E. Agliardi, R. Agliardi, and W. Spanjers, "The Economic Value of Biodiversity Preservation," *Environmental and Resource Economics*, vol. 87, no. 6, pp. 1593–1610, Jun. 2024, https://doi.org/10.1007/s10640-024-00855-0.

[3] J. F. Islas *et al.*, "An overview of Neem (*Azadirachta indica*) and its potential impact on health," *Journal of Functional Foods*, vol. 74, Nov. 2020, Art. no. 104171, https://doi.org/10.1016/j.jff.2020.104171.

[4] S. K. Tulashie, F. Adjei, J. Abraham, and E. Addo, "Potential of neem extracts as natural insecticide against fall armyworm (*Spodoptera frugiperda* (J. E. Smith) (Lepidoptera: Noctuidae)," *Case Studies in Chemical and Environmental Engineering*, vol. 4, Dec. 2021, Art. no. 100130, https://doi.org/10.1016/j.cscee.2021.100130.

[5] A. Jafar, N. Bibi, R. A. Naqvi, A. Sadeghi-Niaraki, and D. Jeong, "Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations," *Frontiers in Plant Science*, vol. 15, Mar. 2024, https://doi.org/10.3389/fpls.2024.1356260.

[6] J. Sun, W. Cao, X. Fu, S. Ochi, and T. Yamanaka, "Few-shot learning for plant disease recognition: A review," *Agronomy Journal*, vol. 116, no. 3, pp. 1204–1216, 2024, https://doi.org/10.1002/agj2.21285.

[7] H. Lin, R. Tse, S. K. Tang, Z. Qiang, and G. Pau, "Few-Shot Learning for Plant-Disease Recognition in the Frequency Domain," *Plants*, vol. 11, no. 21, Jan. 2022, Art. no. 2814, https://doi.org/10.3390/plants11212814.

[8] R. Duan, D. Li, Q. Tong, T. Yang, X. Liu, and X. Liu, "A Survey of Few-Shot Learning: An Effective Method for Intrusion Detection," *Security and Communication Networks*, vol. 2021, no. 1, 2021, Art. no. 4259629, https://doi.org/10.1155/2021/4259629.

[9] J. Yang, X. Guo, Y. Li, F. Marinello, S. Ercisli, and Z. Zhang, "A survey of few-shot learning in smart agriculture: developments, applications, and challenges," *Plant Methods*, vol. 18, no. 1, Mar. 2022, Art. no. 28, https://doi.org/10.1186/s13007-022-00866-2.

[10] M. H. Saad and A. E. Salman, "A plant disease classification using one-shot learning technique with field images," *Multimedia Tools and Applications*, vol. 83, no. 20, pp. 58935–58960, Jun. 2024, https://doi.org/10.1007/s11042-023-17830-4.

[11] Y. Fu *et al.*, "Long-tailed visual recognition with deep models: A methodological survey and evaluation," *Neurocomputing*, vol. 509, pp. 290–309, Oct. 2022, https://doi.org/10.1016/j.neucom.2022.08.031.

[12] X. Li, X. Yang, Z. Ma, and J.-H. Xue, "Deep metric learning for few-shot image classification: A Review of recent developments," *Pattern Recognition*, vol. 138, Jun. 2023, Art. no. 109381, https://doi.org/10.1016/j.patcog.2023.109381.

[13] D. Argüeso *et al.*, "Few-Shot Learning approach for plant disease classification using images taken in the field," *Computers and Electronics in Agriculture*, vol. 175, Aug. 2020, Art. no. 105542, https://doi.org/10.1016/j.compag.2020.105542.

[14] Y. Li and X. Chao, "Semi-supervised few-shot learning approach for plant diseases recognition," *Plant Methods*, vol. 17, no. 1, Jun. 2021, Art. no. 68, https://doi.org/10.1186/s13007-021-00770-1.

[15] M. Rezaei, D. Diepeveen, H. Laga, M. G. K. Jones, and F. Sohel, "Plant disease recognition in a low data scenario using few-shot learning," *Computers and Electronics in Agriculture*, vol. 219, Apr. 2024, Art. no. 108812, https://doi.org/10.1016/j.compag.2024.108812.

[16] B. Wang, Y. Wang, and Y. Xu, "Background-Filtering Feature-Enhanced Graph Neural Networks for Few-Shot Learning," *Applied*

Sciences, vol. 14, no. 15, Jan. 2024, Art. no. 6571, https://doi.org/10.3390/app14156571.

[17] H. Ji, L. Luo, and H. Peng, "BRAVE: A cascaded generative model with sample attention for robust few shot image classification," *Neurocomputing*, vol. 610, Dec. 2024, Art. no. 128585, https://doi.org/10.1016/j.neucom.2024.128585.

[18] B. Wang and D. Wang, "Plant Leaves Classification: A Few-Shot Learning Method Based on Siamese Network," *IEEE Access*, vol. 7, pp. 151754–151763, 2019, https://doi.org/10.1109/ACCESS.2019.2947510.

[19] G. Pushpa, "Indian Medicinal Leaves Image Datasets." Mendeley, May 05, 2023, https://doi.org/10.17632/748F8JKPHB.3.

[20] P. Sarma, "MED117_Medicinal Plant Leaf Dataset & Name Table." Mendeley, Jan. 18, 2023, https://doi.org/10.17632/DTVBWRHZNZ.3.

[21] Y. Li and J. Yang, "Meta-learning baselines and database for few-shot classification in agriculture," *Computers and Electronics in Agriculture*, vol. 182, Mar. 2021, Art. no. 106055, https://doi.org/10.1016/j.compag.2021.106055.

[22] Y. Zhao, Z. Zhang, N. Wu, Z. Zhang, and X. Xu, "MAFDE-DN4: Improved Few-shot plant disease classification method based on Deep Nearest Neighbor Neural Network," *Computers and Electronics in Agriculture*, vol. 226, Nov. 2024, Art. no. 109373, https://doi.org/10.1016/j.compag.2024.109373.

[23] X. Liu and C. Aldrich, "Recognition of flotation froth conditions with k-shot learning and convolutional neural networks," *Journal of Process Control*, vol. 128, Aug. 2023, Art. no. 103004, https://doi.org/10.1016/j.jprocont.2023.103004.

[24] N. Belissent, J. M. Peña, G. A. Mesías-Ruiz, J. Shawe-Taylor, and M. Pérez-Ortiz, "Transfer and zero-shot learning for scalable weed detection and classification in UAV images," *Knowledge-Based Systems*, vol. 292, May 2024, Art. no. 111586, https://doi.org/10.1016/j.knosys.2024.111586.