Interdisciplinary Journal of Information, Knowledge, and Management

An Official Publication of the Informing Science Institute InformingScience.org

IJIKM.org

Volume 19, 2024

EMPHASIZING DATA QUALITY FOR THE IDENTIFICATION OF CHILI VARIETIES IN THE CONTEXT OF SMART AGRICULTURE

Wiwin Suwarningsih*	Research Center for Information and Data Sciences, National Research and Innovation Agency, Bandung, Indonesia	wiwi005@brin.go.id
Rinda Kirana	Research Center of Horticulture and Plantation, National Research and Innovation Agency, Bandung, Indonesia	rind005@brin.go.id
Purnomo Husnul Khotimah	Research Center for Information and Data Sciences, National Research and Innovation Agency, Bandung, Indonesia	purn005@brin.go.id
Andri Fachrur Rozie	Research Center for Information and Data Sciences, National Research and Innovation Agency, Bandung, Indonesia	andr035@brin.go.id
Dianadewi Riswantini	Research Center for Information and Data Sciences, National Research and Innovation Agency, Bandung, Indonesia	dian003@brin.go.id
Ekasari Nugraheni	Research Center for Information and Data Sciences, National Research and Innovation Agency, Bandung, Indonesia	ekas003@brin.go.id
Devi Munandar	Research Center for Information and Data Sciences, National Research and Innovation Agency, Bandung, Indonesia	devi002@brin.go.id
Andria Arisal	Research Center for Information and Data Sciences, National Research and Innovation Agency, Bandung, Indonesia	andr015@brin.go.id
Noor Roufiq Ahmadi	Center for Standard Testing of Vegetable Plant Instruments, Ministry of Agriculture, Bandung, Indonesia	noorroufiqa@gmail.com

* Corresponding author

Accepting Editor Dimitar Grozdanov Christozov │ Received: November 30, 2023 │ Revised: February 15, 2024 │ Accepted: February 23, 2024.

Cite as: Suwarningsih, W., Kirana, R., Khotimah, P. H., Rozie, A. F., Riswantini, D., Nugraheni, E., Munandar, D., Arisal, A., & Ahmadi, N. R. (2024). Emphasizing data quality for the identification of chili varieties. *Interdisciplinary Journal of Information, Knowledge, and Management*, *19,* Article 5. https://doi.org/10.28945/5257

ABSTRACT

Aim/Purpose	This research aims to evaluate models from meta-learning techniques, such as Riemannian Model Agnostic Meta-Learning (RMAML), Model-Agnostic Meta-Learning (MAML), and Reptile meta-learning, to obtain high-quality metadata. The goal is to utilize this metadata to increase accuracy and efficiency in identifying chili varieties in smart agriculture.
Background	The identification of chili varieties in smart agriculture is a complex process that requires a multi-faceted approach. One challenge in chili variety identification is the lack of a large and diverse dataset. This can be addressed using meta-learning techniques, which allow the model to leverage knowledge learned from other related tasks or artificially expand the dataset by applying transformations to existing data. Another challenge is the variation in growing conditions, which can affect the appearance of chili varieties. Meta-learning techniques can help address this challenge by allowing the model to adapt to variations in growing conditions with task-specific embeddings and optimizations. With the help of meta-learning techniques, such as data augmentation, data characterization, selection of datasets, and performance estimation, quality metadata for accurate identification of chili varieties can be achieved even in the presence of limited data and variations in growing conditions. Furthermore, the use of meta-learning techniques in chili variety identification can also assist in addressing challenges related to the computational complexity of the task.
Methodology	The research approach employed is quantitative, specifically comparing three models from meta-learning techniques to determine which model is most suitable for our dataset. Data was collected from the variety assembly garden in the form of images of chili leaves using a mobile device. The research successfully gathered 1,974 images of chili leaves, with 697 images of large red chilies, 649 images of curly red chilies, and 628 images of cayenne peppers. These chili leaf images were then processed using augmentation techniques. The results of image data augmentation were categorized based on leaf characteristics (such as oval, lancet, elliptical, serrated leaf edges, and flat leaf edges). Subsequently, training and validation utilized three models from meta-learning techniques. The final stage involved model evaluation using 2-way and 3-way classification, as well as 5-shot and 10-shot learning scenarios to select the dataset with the best performance.
Contribution	Improving classification accuracy, with a focus on ensuring high-quality data, allows for more precise identification and classification of chili varieties. Enhancing model training through an emphasis on data quality ensures that the models receive reliable and representative input, leading to improved generalization and performance in identifying chili varieties.
Findings	With small collections of datasets, the authors have used data augmentation and meta-learning techniques to overcome the challenges of limited data and variations in growing conditions.
Recommendations for Practitioners	By leveraging the knowledge and adaptability gained from meta-learning, accurate identification of chili varieties can be achieved even with limited data and variations in growing conditions. The use of meta-learning techniques in chili variety identification can greatly improve the accuracy and reliability of the identification process.

Recommendations for Researchers	Using meta-learning techniques, such as transfer learning and parameter optimization, researchers can overcome challenges related to limited data and variations in growing conditions in chili variety identification.
Impact on Society	The findings from this research can help identify superior chili seeds, thereby motivating farmers to cultivate high-quality chilies and achieve bountiful harvests.
Future Research	We intend to verify our approach on a more extensive array of datasets and explore the implementation of more resilient regularization techniques, going beyond image augmentation, within the meta-learning techniques. Furthermore, our goal is to expand our research to encompass the automatic learning of parameters during training and tackle issues associated with noisy labels. Building on the insights gained from our observed outcomes, a future objective is to enhance the refinement of model-agnostic meta-learning techniques that can effectively adapt to intricate task distributions with substantial domain gaps between tasks. To realize this aim, our proposal involves devising model-agnostic meta-learning techniques specifically designed for multi-modal scenarios.

Keywords	chili variety identification, meta-learning, 2-3 way classification, 5-10 shot classification

INTRODUCTION

The application of deep learning technology in image-based plant identification has experienced significant growth across diverse domains, encompassing various plant varieties (Hati & G, 2013; Lee & Chen, 2006), plant diseases (Rumpf et al., 2010; V. Singh & Misra, 2017), and general plant identification (Hati & G, 2013; Liming & Yanchao, 2010; Too et al., 2019). The integration of deep learning intelligence presents both a substantial challenge and an opportunity to enhance the economic potential and efficiency of agricultural production (Geetharamani & Arun Pandian, 2019; Krishnasamy & Nanjundappan, 2016). Advanced deep learning technology serves as a fundamental tool for developing identification models through image processing (Hussain et al., 2020; Oktaria et al., 2019).

Deep learning aims to replicate the functioning of the human brain by combining input data, weights, and biases. These components collaborate to accurately recognize, classify, and describe objects within existing data (Atban et al., 2023; Hussin et al., 2012). However, a notable limitation of deep learning is its voracious appetite for extensive datasets (Atban et al., 2023; Li et al., 2023; Rumpf et al., 2010). Neural networks require diverse and substantial data to effectively learn intricate patterns (Noh et al., 2016; Park et al., 2004), posing a challenge in fields with limited data availability, such as agriculture (Sangeetha & Govindarajan, 2023; Too et al., 2019), and environmental health (Karaaslan et al., 2023; Lan et al., 2023)

Recent studies have showcased the potential of deep learning models in agriculture, particularly in forecasting weather data for precision agriculture (Jin et al., 2020) and predicting environmental parameters for early warning systems (Shuchang Chen et al., 2018). Notably, these models leveraging LSTM and GRU networks have demonstrated high accuracy and effectiveness. Deep learning has found successful applications in various agricultural challenges, including disease detection and fruit classification, yielding impressive accuracy results (Santos et al., 2020). However, the application of deep learning in agriculture also presents challenges, such as optimizing planting strategies and enhancing yields (Alibabaei et al., 2022).

Several studies have explored traditional and machine-learning methods for chili variety identification. Notably, Ahmad Loti et al. (2020) found that deep learning feature-based approaches outperformed traditional methods in identifying chili pests and diseases, achieving an accuracy of 92.10%.

Islam et al. (2020) also employed machine-learning techniques to detect weeds in chili fields with an accuracy of 96%. Saad et al. (2020) applied a deep learning approach to detect chili and its flower in plant images, achieving high accuracy in classification and detection. Aminuddin et al. (2022) provided a comprehensive review of computational approaches for automated disease identification in chili leaf images, emphasizing their potential. In Suwarningsih et al.'s (2022) study, a dataset comprising images of chili leaves with 12 distinct classes of variety showed classification accuracy ranging from 70.18% to 78.37%. However, the achieved accuracy fell short of expectations, mainly due to imbalanced data distribution among classes and the relatively small dataset.

Despite advancements in deep learning and machine learning models designed to address these challenges, inherent limitations persist (Liu et al., 2022). Deep learning algorithms requiring extensive data face challenges in the context of variety identification and classification (Xu et al., 2023). Training models on limited examples can impact generalization capability and introduce biases (Gaikwad & Tidke, 2022). To overcome these limitations and create high-performing models with restricted data, a meta-learning approach proves valuable (Gama et al., 2023). Meta-learning streamlines the trial-and-error process, leading to improved predictions in a shorter timeframe (Saeed et al., 2022). This approach utilizes the output and metadata from machine learning algorithms as input for prediction.

Numerous approaches in the realm of few-shot learning leverage meta-learning, preparing a network for learning from minimal training data, common in few-shot learning scenarios (Zhu et al., 2022). Meta-learning extends to other computer vision tasks, such as rapid adaptation for video tracking (Liu et al., 2022). Mechanisms in meta-learning approaches fall into two broad categories, deploying specialized network architectures to encode information gathered during the meta-learning phase. This includes techniques like fast weights (R. Singh et al., 2021), neural plasticity values (Saeed et al., 2022), custom update rules (Tabealhojeh et al., 2023), temporal convolutions (Khodadadeh et al., 2019), or the use of LSTM memory (Solgi et al., 2021). While offering the advantage of fine-tuning the architecture for efficient encoding of meta-learning information, a drawback is that it confines the selection of network architectures, limiting the seamless translation of innovations in network design to the meta-learning approach (Gama et al., 2023). In a custom network architecture meta-learning model, the target learning phase diverges from conventional network learning, requiring effective leveraging of the custom encoding (Verma et al., 2023).

Researchers have embraced strategies such as transfer learning (Kandel & Castelli, 2020) and have explored various fine-tuning techniques specific to chili leaf variety datasets. It is crucial to note that fine-tuning requires a prerequisite – the dataset used to pre-train the model should exhibit a similar data distribution to the dataset intended for the fine-tuning process. In the realm of chili leaf variety image classification, the absence of an extensive database like ImageNet for pre-training poses a challenge (Lin et al., 2021). To address the scarcity of data, efforts have been directed towards generating synthetic data using techniques such as data augmentation (Zhao et al., 2019) and Generative Adversarial Network (GAN) (Sandfort et al., 2019). Training a GAN involves fine-tuning multiple parameters to produce high-quality virtual images. Consequently, novel strategies are imperative to combat the issue of data scarcity. Meta-learning emerges as a promising paradigm, offering versatile solutions for various challenges, particularly in the agricultural domain.

Meta-learning is a subfield of machine learning that focuses on the development of algorithms and models capable of learning to learn. These algorithms and models aim to improve the efficiency and effectiveness of the learning process by leveraging past experiences and knowledge gained from solving similar tasks. Meta-learning techniques have found direct applications in various fields, such as supervised learning and reinforcement learning (Durall et al., 2019). The popularity of meta-learning has surged in recent years due to its ability to quickly adapt to new and unseen tasks (Gupta et al., 2021).

In Lemke et al.'s (2015) study, the authors explored the architectures and frameworks of meta-learning, highlighting future trends such as extracting meta-features, selecting base-learners, and dynamically updating base-learners. However, recognizing that the categorization in Vanschoren's (2018) work may not fully capture all aspects of meta-learning and the connections between diverse meta-learning frameworks, Hospedales et al. (2022) categorized the meta-learning problem into meta-representation, meta-objective, and meta-optimizer. In their comprehensive review, they positioned the definition of meta-learning in relation to other fields, such as transfer learning, multi-task learning, and hyperparameter optimization. Additionally, they addressed the few-shot learning problem as a primary application of meta-learning in real-world scenarios. Vanschoren (2018) introduced the principles of meta-learning and categorized meta-learning techniques based on learning tasks, including learning from model evaluations, learning from task properties, and learning from prior models. This survey also addressed few-shot learning problems.

A comparative analysis of various definitions of meta-learning provided by different research teams, with a particular focus on classification problems (Vilalta & Drissi, 2002). Their review centered on summarizing methods primarily applied in classification contexts, offering valuable insights into the utilization of meta-knowledge. Vilalta and Drissi (2002) concluded that the key objective in defining meta-learning is to explore the interplay between learning mechanisms and specific application contexts.

In Y. Wang et al.'s (2020) exhaustive survey, the few-shot learning problem was distinctly differentiated from conventional machine learning problems. They proposed a novel categorization of few-shot learning based on how prior knowledge can be leveraged to minimize empirical risk. Meta-learning (J. X. Wang, 2021) is a well-established field that enhances learning strategies to tackle the challenges of Artificial General Intelligence (AGI). While AGI remains a distant goal, meta-learning has made significant advances in domains such as image recognition, reinforcement learning, and regression. The inspiration for this new paradigm, meta-learning, is drawn from the quick adaptation of human intelligence to new skills based on prior experiences (Si Chen et al., 2022).

The potential to significantly enhance the accuracy and reliability of agricultural processes lies in focusing on data quality. The following contributions can be identified:

(i) *Improved Classification Accuracy*. Ensuring high-quality data allows for more accurate identification and classification of chili varieties. This contributes to the reliability of the smart agriculture system in distinguishing between different types of chilies, aiding farmers in better decision-making.

(ii) *Enhanced Model Training*. Emphasizing data quality ensures that the model receives reliable and representative input, leading to better generalization and performance when identifying chili varieties. By placing emphasis on data quality, the likelihood of misclassification and errors in chili variety identification is minimized. This is crucial for preventing incorrect recommendations or actions based on flawed data, ultimately improving the overall efficiency of smart agricultural systems.

This paper is organized as follows. The proposed method is outlined in the next section, followed by the details of the experimental validation and discussion. Finally, the findings and potential future directions are highlighted.

MATERIAL AND METHOD

Meta-learning is an approach that focuses on developing algorithms and models capable of learning and adapting to different tasks or domains more efficiently. The primary goal of meta-learning is to enable a system to leverage past learning experiences, enhancing its performance on new, unseen tasks. The following materials and methods can be considered.

Materials

The data we collected originated from 1,974 variety assembly gardens in digital image form. These chili leaf images were captured using various types of mobile devices with different resolutions and pixels. The images showcase various varieties of chili plants, encompassing large red chilies, curly red chilies, and cayenne peppers. Large red chilies comprise six varieties with 697 leaf images, while curly red chilies consist of five varieties, totaling 649 leaf images. Cayenne peppers encompass four varieties, amounting to 628 leaf images (examples of leaves are illustrated in Figure 1). The dataset is categorized into three groups: large red chilies, curly red chilies, and cayenne peppers, each containing specific classes. The red chili cluster encompasses six classes (*Tanjung-2, Branang, Ciko, Lingga, Inata Agrohorti, Carvi Agrihorti*), the curly red chili cluster comprises five classes (*Ateng Maninjau, Kampung Manangah, Randah, Sijunjung, Paijan*), and the cayenne pepper group includes four classes (*Hiyung, Rawita, Sigantung, Patra*).

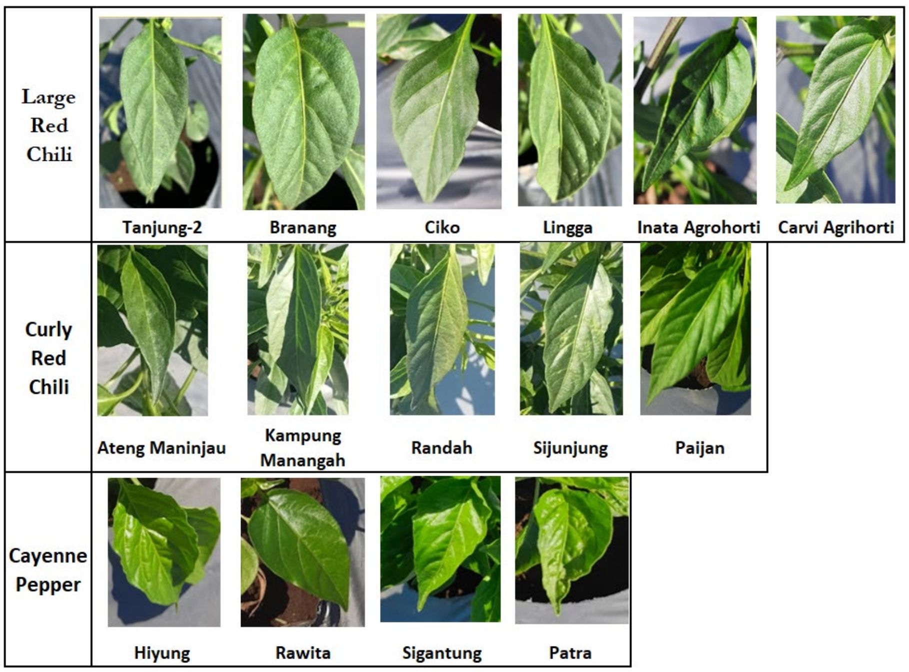

Figure 1. Example of an image of chili leaves-based varieties

Methods

The method employed in this research involves a quantitative model, specifically evaluating the model using meta-learning techniques. The analyzed models include RMAML, MAML, and Reptile ML. This stage aims to determine the most suitable model for our dataset. A flow diagram illustrating the augmentation process and the meta-learning techniques is presented in Figure 2. This diagram visually represents the key steps involved in the implementation, encompassing data augmentation techniques and meta-learning techniques.

Data augmentation

In this research, we conducted data augmentation to increase the number and variety of datasets we have. The augmentation stages we carried out include glow edge, salient edge map, de-texturized, and flip/rotate. This process provides a larger dataset for training and enables the model to discover a wider range of features.

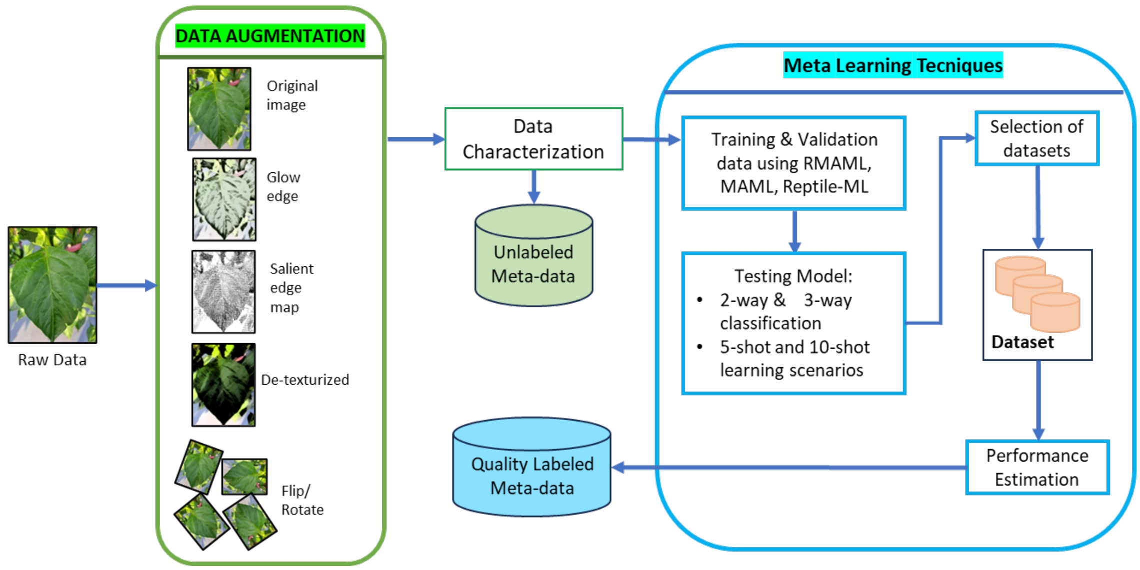

Figure 2. Flow diagram for proposed method

Data characterization

Characterization is the process of identifying specific traits in chilies, which serves to differentiate between varieties and individuals. The objective is to provide a detailed description of the three types of chili varieties (an example of a data characterization-based morphology grouping combination can be seen in Figure 3). To maintain diversity within the dataset, we adjusted the number of leaves in each group during combination creation. Table 1 details the unique leaf groupings and their associated quantities of images. Additionally, we divided the images within each combination into training and testing sets using a 75:25 split ratio to evaluate the performance of pre-trained models.

Combinations labeled as M1 to M10 are employed in generating metadata and training the meta-learner, while combinations labeled as M11 to M14 are reserved for testing the proposed framework. The selection of combinations for testing considers specific factors that showcase the effectiveness and practicality of the proposed framework:

M11 = Two leaves with a serrated lanceolate and elongated serrated.
M12 = Two leaves with serrated elliptical and serrated ovate.
M13 = Three leaves with a flat elliptical, serrated elliptical, and serrated ovate.
M14 = Three leaves with a flat oval, serrated elliptical, and serrated ovate.

Table 1. Leaf morphology grouping combinations

ID#	Combination	Images#
M1	flat lancet, elongated evenly and serrated lanceolate	193
M2	flat lancet, elongated evenly and serrated elongated	122
M3	flat elliptical, flat oval, and serrated elliptical	157
M4	flat elliptical, flat oval, and serrated ovate	87
M5	lancet and elongated	151
M6	elliptical and ovate	166
M7	flat lancet and elongated evenly	264
M8	flat elliptical and flat oval	232
M9	flat lancet, serrated lanceolate, and elongated serrated	75

ID#	Combination	Images#
M10	elongated evenly, serrated lanceolate, and elongated serrated	140
M11	serrated lanceolate and elongated serrated	40
M12	serrated elliptical and serrated ovate	94
M13	flat elliptical, serrated elliptical, and serrated ovate	105
M14	flat oval, serrated elliptical, and serrated ovate	148

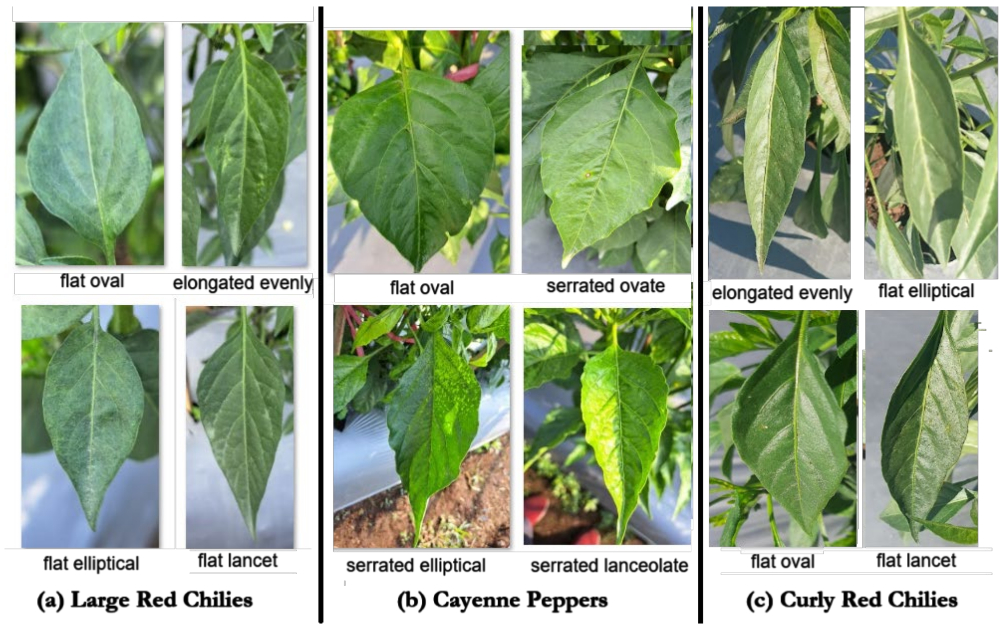

Figure 3. Example of a characterization data-based morphology grouping combinations

Training and validation using Riemannian Model-Agnostic Meta-Learning (RMAML) with orthogonality constraints

Gradient-based techniques represent crucial optimization methods in both Euclidean and Riemannian spaces. In the Euclidean space, the standard Gradient Descent (GD) method updates the learnable parameters by moving along the negative gradient direction. Its stochastic variant, namely Stochastic Gradient Descent (SGD), updates parameters using mini-batches of data, leading to faster convergence and a reduced likelihood of getting stuck in local minima in the parameter space. However, the computed gradients during SGD may exhibit significant variances due to the random generation of mini-batches, resulting in undesirable oscillations of model parameters around optimal values (Roy et al., 2018).

To address this issue, Momentum-SGD (M-SGD) is a widely employed technique that incorporates a momentum term to mitigate these oscillations over time. More recently, several methods have focused on improving the convergence rate through automatic learning-rate adaptation.

Training and validation using Model-Agnostic Meta-Learning (MAML)

MAML, introduced by the research team led by Finn et al. (2017), initializes the model parameters and undergoes a few gradient update steps to adapt to new tasks. The core concept behind MAML is to optimize the initial parameters of the model to facilitate efficient learning on a new task with minimal gradient steps. Given that slight parameter adjustments can lead to significant improvements in task learning, MAML emphasizes acquiring transferable representations rather than focusing on specific learning rules. The strengths of MAML include its ability to mitigate overfitting issues commonly encountered with limited data and its versatility across various application domains, such as regression, classification, and reinforcement learning.

However, Antoniou et al. (2019) highlighted certain drawbacks of MAML: the training process of MAML is often unstable, relying heavily on the network architecture and parameter setup, and the necessity of sharing the learning rate for all parameters in MAML requires repeated hyperparameter searches to determine the correct learning rate for a specific task, thereby increasing the computational burden.

Training and validation using Reptile Meta-Learning

Reptile, introduced with the goal of quickly learning the initial parameters of a model for efficient adaptation to new tasks, is proposed by the research community (Zhang et al., 2022). However, it is important to note that Reptile is more apt for optimization problems that require a substantial number of updated steps. In Reptile, the stochastic gradient descent (SGD) algorithm is employed to perform gradient updates on each task in the shortest possible descent.

The advantages of Reptile include the following: (1) it requires less computation and memory than MAML since it does not unroll a computation graph or calculate second derivatives, and (2) Reptile enhances the model's convergence speed due to lower variance. However, it is worth noting that existing research has primarily demonstrated the effectiveness of Reptile in few-shot classification. There is currently no empirical evidence supporting its performance in other few-shot learning tasks, such as regression or reinforcement learning.

RESULT AND DISCUSSION

Meta-learning can play a crucial role in chili variety identification by enabling the model to learn how to adapt to different types of macronutrients found in chili varieties. By using meta-learning, the model can learn from previous experiences and apply that knowledge to new chili varieties, improving its accuracy and efficiency in identification. Meta-learning can also help optimize the hyperspectral imaging process by learning which wavelengths and spectral bands are most informative for chili variety identification. Moreover, meta-learning can facilitate the transfer of knowledge from other crops, such as sunflowers and papaya, where deep learning models have shown high accuracy in identifying a variety of species. Using meta-learning techniques, the chili variety identification model can learn to recognize patterns and features that are common across different varieties, allowing it to make accurate predictions for new, unseen chili varieties.

EXPERIMENTAL SETTING

We assessed the effectiveness of our approach using three datasets of chili images, as detailed in Table 2. These datasets amalgamate leaf morphology data derived from the chili leaf dataset. Our evaluations covered 2-way and 3-way classification tasks involving 5-shot and 10-shot learning scenarios. We utilized both traditional and advanced augmentation techniques. Additionally, we explored a transfer-learning approach, where the model underwent initial training on all meta-train classes and subsequent fine-tuning on a few shots from the meta-test set.

In the realm of meta-learning, conventional augmentation methods like flipping and rotating contribute minimal additional information for model learning. This limitation arises because, during meta-training, we sample not only images but also classes. Hence, there is a need to create novel classes with new images.

EXPERIMENT RESULT

In all our experiments, we utilized accuracy (%) as the primary evaluation metric, aligning with the standard practice in classification tasks. It is important to highlight that the accuracy metric maintains its robustness during meta-testing due to the even distribution of images across all classes, mitigating the impact of class imbalance. The reported accuracy reflects the average performance across 100 randomly sampled few-shot tasks from the meta-testing dataset.

Our evaluation encompasses three distinct leaf image datasets, focusing on both 2-way and 3-way classification tasks and covering 5-shot and 10-shot learning scenarios. We employed a variety of augmentation techniques, ranging from traditional to advanced methods, in these assessments. Moreover, we explored a transfer learning approach, wherein the model underwent initial training on all meta-train classes and subsequently underwent fine-tuning using multiple shots from the meta-test set.

For each of the k-shot, n-way experiments, including 3-shot, 5-shot, and 10-shot learning scenarios in both 2-way and 3-way classification tasks, we repeated the process using three different augmentation techniques: RMAML, MAML, and Reptile. Similarly, we conducted evaluations for the transfer learning approach.

Transfer learning stands out as a widely acknowledged strategy for constructing deep learning classification models in situations characterized by limited data availability. Hence, we adopted transfer learning as our baseline method, as detailed in Table 2, for a comprehensive comparison with meta-learning approaches. During the transfer learning process, we employed the entire meta-training dataset for supervised learning over 1000 epochs. Following model selection, the chosen model underwent fine-tuning using the same procedure applied in our meta-learning experiments. To ensure a fair comparison, we maintained a consistent network architecture and hyperparameters across all experiments.

Table 2 presents the results of the transfer learning method, shown in the third column for 2-way classification and the seventh column for 3-way classification. However, for 5-shot and 10-shot learning challenges, the transfer learning approach did not yield the highest test accuracy in both 2-way and 3-way classification scenarios. The identification accuracy for chili varieties, averaging 83%, represents a noteworthy advancement compared to earlier research. Yet, it is crucial to note that this accuracy does not straightforwardly translate to the diverse and intricate conditions in which chili peppers grow, given their vast numbers. Precise recognition of chili types holds great importance in applications like quality control, breeding, and seed production. This paper introduces meta-learning methods – RMAML, MAML, and Reptile – to augment the effectiveness of models trained through gradient descent.

Table 2. Experiment results of evaluation metric

Few-shot task	2-way classification meta-learning				3-way classification meta-learning			
	Normal augmentation	RMAML	MAML	Reptile	Normal augmentation	RMAML	MAML	Reptile
5-SHOT	73.88%	83.65%	82.52%	82.12%	60.12%	79.89%	79.45%	79.10%
10-SHOT	81.23%	83.98%	83.01%	83.12%	65.89%	78.45%	78.21%	78.12%

RMAML incorporates the concept of task relations in its meta-learning process, allowing it to grasp the similarities and differences among various chili pepper varieties. This inclusion of additional information enhances RMAML's capacity to accurately discern between these varieties. In contrast, MAML stands out as a widely utilized meta-learning algorithm that can be integrated with any learning algorithm, facilitating rapid adaptation to new tasks.

Reptile, another meta-learning algorithm, presents a simplified and computationally efficient alternative to MAML. Unlike MAML, Reptile performs stochastic gradient descent or Adam on each task just once, foregoing the repetitive nature of MAML. Consequently, the Reptile algorithm conserves training time while achieving performance comparable to MAML.

In the realm of chili variety identification, RMAML, MAML, and Reptile meta-learning techniques offer promising strategies. These approaches have the potential to elevate the accuracy and efficiency of chili variety identification through the application of meta-learning methods. To validate the efficacy of our approach, we continuously evaluate and contrast three techniques against several more recent few-shot learning algorithms. Table 3 showcases the outcomes for the datasets of large red chilies, curly red chilies, and cayenne peppers, respectively. Our findings reveal that, compared to recent methodologies, our three algorithms notably enhance few-shot recognition performance, particularly on the cayenne peppers dataset. As illustrated in Table 3, when applied to the detailed dataset of chilies, our proposed method is pitted against various state-of-the-art methods utilizing MAML. The results demonstrate that RMAML outperforms all considered methods in both 5-shot and 10-shot scenarios. Moreover, RMAML surpasses all prior state-of-the-art algorithms in the 10-shot classification of the curly red chilies dataset. In 5-shot experiments, RMAML exhibits superior accuracy compared to methods employing the same MAML and Reptile approach.

Table 3. Additional comparative analysis of few-shot classification results on the Large Red Chile (LRC), Curly Red Chilies (CRC) and Cayenne Peppers (CP) dataset

Method	5-shot			10-shot		
	LRC	CRC	CP	LRC	CRC	CP
RMAML	**80.38%**	**81.37%**	**81.87%**	**84.25 %**	**80.62%**	**82.12%**
MAML	80.08%	81.01%	80.21%	82.12%	80.12%	81.98%
Reptile	80.12%	81.12%	80,01%	82.48%	79.98%	81.01%
Matching-Net	80.01%	80.98%	79.23%	81.75%	79.12%	80.11%
Proto-Net	79.89%	80.56%	79.45%	81.25%	79.25%	80.24%
ML-LSTM	79.99%	80.23%	79.30%	80.56%	79.36%	80.07%

Based on Table 3, we observed that across all datasets and in both 5-shot and 10-shot learning scenarios, most of the prediction scores generated by the (RMAML, MAML, & Reptile) meta-learning approach were close to 0.8-0.9. In contrast, another set of meta-learning models (Matching-Net, Proto-Net, ML-LSTM) performed poorly, with most predictions falling within the confidence score range of 0.7–0.8. This pattern held true for both 5-shot and 10-shot classifications. In the smart agriculture domain, deep learning algorithms are used to assist experts rather than making decisions independently. Therefore, having a confident model will help experts make correct and confident decisions.

We conducted standard 5-shot and 10-shot experiments for single-domain few-shot classification. Table 3 presents the results of these experiments, focusing on different chili datasets such as large red chilies, curly red chilies, and cayenne peppers. Across all experimental setups, RMAML demonstrates a noteworthy improvement over Matching-Net, Proto-Net, and ML-LSTM in terms of classification accuracy (e.g., 80.38%, 81.37%, 81.87% for the 5-shot experiment on the cayenne peppers dataset). Furthermore, RMAML consistently outperforms or achieves comparable results with other meta-learning techniques considered for both 5-shot and 10-shot experiments. As depicted in Table 3, we observe that RMAML enhances few-shot recognition performance on the large red chilies, curly red chilies, and cayenne peppers datasets compared to its counterparts.

DISCUSSION

We observed that the implementation of advanced augmentation techniques effectively mitigates the issue of overfitting in both 5-shot and 10-shot learning scenarios. The plots in Tables 2 and 3 illustrate the outcomes for 5-shot and 10-shot learning, respectively, focusing on the large red chilies, curly red chilies, and cayenne peppers dataset in a 2-way classification task.

Our observations reveal instances where the training accuracy is lower than the test accuracy for some iterations. This discrepancy arises because the data split employed during the meta-training stage is relatively challenging, and the augmentations introduced during this stage generate more intricate virtual samples. Consequently, the model initially struggles with the meta-training dataset. We speculate that this behavior is intrinsic to the dataset itself and the specific split between meta-train and meta-test sets. The impact of augmentation is particularly prominent in 5-shot and 10-shot learning scenarios, as the disparity between training and testing values significantly diminishes in these instances. We contend that 5-shot learning represents an extreme case, and consequently, the use of augmentation has a lesser impact on model performance.

Utilizing meta-learning for the purpose of identifying chili varieties entails harnessing acquired knowledge and adjusting to novel tasks. This process involves training a meta-learning model on a diverse dataset containing various chili varieties, enabling it to grasp the shared patterns and distinctive characteristics among different varieties. The meta-learning model, such as RMAML, MAML, and Reptile, can be applied to rapidly adapt and categorize new chili images based on their acquired knowledge. This strategy proves particularly beneficial when facing limited data for new chili varieties, as the meta-learning model leverages its past knowledge to make precise predictions for novel varieties. Integrating meta-learning techniques empowers authors to enhance the accuracy and efficiency of chili variety identification.

RMAML, MAML, and Reptile methods offer valuable guidance for developing effective techniques in chili variety identification. Overall, by employing meta-learning techniques, authors can elevate the precision and efficiency of chili variety identification through the extraction of pertinent features from chili leaf images. Achieving this involves training the meta-learning model on an extensive dataset of chili leaf images, enabling it to discern the distinctive features that set different varieties apart. However, the field of chili variety identification faces several challenges that warrant attention. One such challenge involves navigating difficult illumination conditions, such as dim light or total darkness, which can compromise the quality of visual information crucial for accurate identification. Additionally, the intricate and ever-changing chili growing environment presents a hurdle for precise variety identification.

CONCLUSION AND FUTURE WORK

In this study, we proposed a meta-learning approach by framing the classification of leaf images in low-data scenarios as a few-shot learning challenge. Moreover, we incorporated advanced augmentation techniques such as 2-way and 3-way classification to improve the model's regularization. We evaluated the effectiveness of this approach on three complex leaf image datasets, and after thorough analysis, we concluded that meta-learning outperforms transfer learning across all dataset experiments.

The transfer learning model exhibited a lack of confidence in its predictions, posing potential risks, particularly in the plant domain. Additionally, our observations highlight that the comparison of RMAML, MAML, and Reptile methods serves to regularize the model, resulting in a substantial improvement in test accuracy across all datasets. As expected, there was a noticeable enhancement in model performance as we transitioned from 5-shot to 10-shot experiments. Throughout our experimentation, we noted that the choice of the optimizer, along with its hyperparameters, significantly influences the model's performance, emphasizing the importance of careful consideration in selecting these elements. We are optimistic that our work will be valuable for the community employing meta-learning approaches within the plant domain.

In future research, we plan to validate our methodology across a broader range of datasets and explore the application of more robust regularization strategies beyond image augmentation within the meta-learning framework. Additionally, we aim to extend our work to include the automated learning of parameters during training and address challenges related to noisy labels. Another future objective, inspired by our observed results, is to advance the development of more effective model-agnostic meta-learning techniques capable of handling complex task distributions with significant domain gaps between tasks. To achieve this, our proposal involves the creation of model-agnostic meta-learning techniques tailored for multi-modal scenarios.

REFERENCES

Ahmad Loti, N. N., Mohd Noor, M. R., & Chang, S. (2020). Integrated analysis of machine learning and deep learning in chili pest and disease identification. *Journal of the Science of Food and Agriculture*, *101*(2), 3582-3594 https://doi.org/10.1002/jsfa.10987

Alibabaei, K., Gaspar, P. D., Lima, T. M., Campos, R. M., Girão, I., Monteiro, J., & Lopes, C. M. (2022). A review of the challenges of using deep learning algorithms to support decision-making in agricultural activities. *Remote Sensing*, *14*(3), 638. https://doi.org/10.3390/rs14030638

Aminuddin, N. F., Joret, A. B., Zulkifli, S. A., Abdul Kadir, H., Morsin, M., & Tukiran, Z. (2022). Computational approaches based on image processing for automated disease identification on chili leaf images: A review. *Emerging Advances in Integrated Technology*, *3*(2), 11-26. https://doi.org/10.30880/emait.2022.03.02.002

Antoniou, A., Edwards, H., & Storkey, A. (2019, May). How to train your MAML. *Proceedings of the 7th International Conference on Learning Representations, New Orleans, USA.* https://openreview.net/pdf?id=HJGven05Y7

Atban, F., Ekinci, E., & Garip, Z. (2023). Traditional machine learning algorithms for breast cancer image classification with optimized deep features. *Biomedical Signal Processing and Control*, *81*, 104534. https://doi.org/10.1016/j.bspc.2022.104534

Chen, S., Li, B., Cao, J., & Mao, B. (2018). Research on agricultural environment prediction based on deep learning. *Procedia Computer Science*, *139*, 33-40. https://doi.org/10.1016/j.procs.2018.10.214

Chen, S., Wang, L., Wang, Z., Yan, Y., Wang, D.-H., & Zhu, S. (2022). Learning meta-adversarial features via multi-stage adaptation network for robust visual object tracking. *Neurocomputing*, *491*, 365–381. https://doi.org/10.1016/j.neucom.2022.03.031

Durall, R., Pfreundt, F.-J., & Keuper, J. (2019, December). Semi few-shot attribute translation. *Proceedings of the International Conference on Image and Vision Computing, Dunedin, New Zealand.* https://doi.org/10.1109/IVCNZ48456.2019.8961034

Finn, C., Abbeel, P., & Levine, S. (2017). *Model-agnostic meta-learning for fast adaptation of deep networks.* arXiv:1703.03400. https://doi.org/10.48550/arXiv.1703.03400

Gaikwad, S., & Tidke, S. (2022). Multi-spectral imaging for fruits and vegetables. *International Journal of Advanced Computer Science and Applications*, *13*(2), 743-760. https://doi.org/10.14569/IJACSA.2022.0130287

Gama, P. H. T., Oliveira, H., dos Santos, J. A., & Cesar, R. M. (2023). An overview on meta-learning approaches for few-shot weakly-supervised segmentation. *Computers & Graphics*, *113*, 77–88. https://doi.org/10.1016/j.cag.2023.05.009

Geetharamani, G., & Arun Pandian, J. (2019). Corrigendum to: 'Identification of plant leaf diseases using a nine-layer deep convolutional neural network', Computers & Electrical Engineering Journal at Volume 76, June 2019, Pages 323-338. *Computers and Electrical Engineering*, *78*, 536. https://doi.org/10.1016/j.compeleceng.2019.08.010

Gupta, A., Jonnalagedda, P., Bhanu, B., & Roy-Chowdhury, A. K. (2021). Ada-VSR: Adaptive video super-resolution with meta-learning. *Proceedings of the 29th ACM International Conference on Multimedia* (pp. 327-336). ACM. https://doi.org/10.1145/3474085.3475320

Hati, S., & G, S. (2013). Plant recognition from leaf image through artificial neural network. *International Journal of Computer Applications*, *62*(17), 15–18. https://doi.org/10.5120/10172-4897

Hospedales, T., Antoniou, A., Micaelli, P., & Storkey, A. (2022). Meta-learning in neural networks: A survey. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 44(9), 5149-5169. https://doi.org/10.1109/TPAMI.2021.3079209

Hussain, E., Mahanta, L. B., Das, C. R., & Talukdar, R. K. (2020). A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network. *Tissue and Cell*, *65*, 101347. https://doi.org/10.1016/j.tice.2020.101347

Hussin, R., Juhari, M. R., Kang, N. W., Ismail, R. C., & Kamarudin, A. (2012). Digital image processing techniques for object detection from complex background image. *Procedia Engineering*, *41*, 340–344. https://doi.org/10.1016/j.proeng.2012.07.182

Islam, N., Rashid, M. M., Wibowo, S., Wasimi, S., Morshed, A., Xu, C., & Moore, S. (2020). Machine learning based approach for weed detection in chilli field using RGB images. In H. Meng, T. Lei, M. Li, K. Li, N. Xiong, & L. Wang (Eds.), Advances in natural computation, fuzzy systems and knowledge discovery (pp. 1097-1105). Springer. https://doi.org/10.1007/978-3-030-70665-4_119

Jin, X., Yu, X., Wang, X., Bai, Y., Su, T., & Kong, J. (2020). Deep learning predictor for sustainable precision agriculture based on Internet of Things system. *Sustainability*, *12*(4), 1433. https://doi.org/10.3390/su12041433

Kandel, I., & Castelli, M. (2020). How deeply to fine-tune a convolutional neural network: A case study using a histopathology dataset. *Applied Sciences, 10*(10), 3359. https://doi.org/10.3390/app10103359

Karaaslan, Ş., Kobat, S. G., & Gedikpınar, M. (2023). A new method based on deep learning and image processing for detection of strabismus with the Hirschberg test. *Photodiagnosis and Photodynamic Therapy*, *44*, 103805. https://doi.org/10.1016/j.pdpdt.2023.103805

Khodadadeh, S., Bölöni, L., & Shah, M. (2019). Unsupervised meta-learning for few-shot image classification. *Proceedings of the 33rd Conference on Neural Information Processing System, Vancouver, Canada*, Article 909.

Krishnasamy, U., & Nanjundappan, D. (2016). Hybrid weighted probabilistic neural network and biogeography based optimization for dynamic economic dispatch of integrated multiple-fuel and wind power plants. *International Journal of Electrical Power and Energy Systems, 77*, 385–394. https://doi.org/10.1016/j.ijepes.2015.11.022

Lan, P., Guo, L., Sun, H., Zhang, Y., & Jiang, Y. (2023). Modeling stream baseflow nitrate concentration in an agricultural watershed using neural network and bootstrap method. *Ecological Indicators, 156*. https://doi.org/10.1016/j.ecolind.2023.111097

Lee, C. L., & Chen, S. Y. (2006). Classification of leaf images. *International Journal of Imaging Systems and Technology, 16*(1), 15–23. https://doi.org/10.1002/ima.20063

Lemke, C., Budka, M., & Gabrys, B. (2015). Metalearning: A survey of trends and technologies. *Artificial Intelligence Review*, *44*, 117–130. https://doi.org/10.1007/s10462-013-9406-y

Li, D., Wang, X., Zhang, J., & Ji, Z. (2023). Automated deep learning system for power line inspection image analysis and processing: Architecture and design issues. *Global Energy Interconnection, 6*(5), 614–633. https://doi.org/10.1016/j.gloei.2023.10.008

Liming, X., & Yanchao, Z. (2010). Automated strawberry grading system based on image processing. *Computers and Electronics in Agriculture, 71*(SUPPL. 1), 32–39. https://doi.org/10.1016/j.compag.2009.09.013

Lin, Y., Guo, X., & Lu, Y. (2021). Self-supervised video representation learning with meta-contrastive network. *Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada,* 8219-8229. https://doi.org/10.1109/ICCV48922.2021.00813

Liu, X., Zhang, X., Peng, W., Zhou, W., & Yao, W. (2022). A novel meta-learning initialization method for physics-informed neural networks. *Neural Computing and Applications*, *34*, 14511–14534. https://doi.org/10.1007/s00521-022-07294-2

Noh, H., Seo, P. H., & Han, B. (2016). Image question answering using convolutional neural network with dynamic parameter prediction. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, 30–38. https://doi.org/10.1109/CVPR.2016.11

Oktaria, A. S., Prakasa, E., & Suhartono, E. (2019). Wood species identification using Convolutional Neural Network (CNN) architectures on macroscopic images. *Journal of Information Technology and Computer Science, 4*(3), 274–283. https://doi.org/10.25126/jitecs.201943155

Park, S. B., Lee, J. W., & Kim, S. K. (2004). Content-based image classification using a neural network. *Pattern Recognition Letters, 25*(3), 287–300. https://doi.org/10.1016/j.patrec.2003.10.015

Roy, S. K., Mhammedi, M., & Harandi, M. (2018). Geometry aware constrained optimization techniques for deep learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA,* 4460–4469. https://doi.org/10.1109/CVPR.2018.00469

Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., & Plümer, L. (2010). Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture, 74*(1), 91–99. https://doi.org/10.1016/j.compag.2010.06.009

Saad, W. H. M., Karim, S. A. A., Razak, M. S. J. A., Radzi, S. A., & Yussof, Z. M. (2020). Classification and detection of chili and its flower using deep learning approach. *Journal of Physics: Conference Series*, *1502*, 012055. https://doi.org/10.1088/1742-6596/1502/1/012055

Saeed, S. U., Fu, Y., Stavrinides, V., Baum, Z. M. C., Yang, Q., Rusu, M., Fan, R. E., Sonn, G. A., Noble, J. A., Barratt, D. C., & Hu, Y. (2022). Image quality assessment for machine learning tasks using meta-reinforcement learning. *Medical Image Analysis, 78*. https://doi.org/10.1016/j.media.2022.102427

Sandfort, V., Yan, K., Pickhardt, P. J., & Summers, R. M. (2019). Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Scientific Reports*, *9*, Article 16884. https://doi.org/10.1038/s41598-019-52737-x

Sangeetha, J., & Govindarajan, P. (2023). Prediction of agricultural waste compost maturity using fast regions with convolutional neural network(R-CNN). *Materials Today: Proceedings*. https://doi.org/10.1016/j.matpr.2023.01.112

Santos, L., Santos, F. N., Oliveira, P. M., & Shinde, P. (2020). Deep learning applications in Agriculture: A short review. In M. Silva, J. Luís Lima, L. Reis, A. Sanfeliu, D. Tardioli (Eds.), *Robot 2019: Fourth Iberian Robotics Conference* (pp. 139-151). Springer. https://doi.org/10.1007/978-3-030-35990-4_12

Singh, R., Bharti, V., Purohit, V., Kumar, A., Singh, A. K., & Singh, S. K. (2021). MetaMed: Few-shot medical image classification using gradient-based meta-learning. *Pattern Recognition, 120*. https://doi.org/10.1016/j.patcog.2021.108111

Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information Processing in Agriculture, 4*(1), 41-49. https://doi.org/10.1016/j.inpa.2016.10.005

Solgi, R., Loáiciga, H. A., & Kram, M. (2021). Long short-term memory neural network (LSTM-NN) for aquifer level time series forecasting using in-situ piezometric observations. *Journal of Hydrology*, *601*, 126800. https://doi.org/10.1016/j.jhydrol.2021.126800

Suwarningsih, W., Khotimah, P. H., Rozie, A. F., Arisal, A., Riswantini, D., Nugraheni, E., Munandar, D., & Kirana, R. (2022). Ide-cabe: Chili varieties identification and classification system based leaf. *Bulletin of Electrical Engineering and Informatics, 11*(1), 445-453. https://doi.org/10.11591/eei.v11i1.3276.

Tabealhojeh, H., Adibi, P., Karshenas, H., Roy, S. K., & Harandi, M. (2023). RMAML: Riemannian meta-learning with orthogonality constraints. *Pattern Recognition, 140*. https://doi.org/10.1016/j.patcog.2023.109563

Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture, 161*(March), 272–279. https://doi.org/10.1016/j.compag.2018.03.032

Vanschoren, J. (2018). *Meta-learning: A survey.* arXiv:1810.03548v1. https://arxiv.org/abs/1810.03548

Verma, S., Kumar, P., & Singh, J. P. (2023). A meta-learning framework for recommending CNN models for plant disease identification tasks. *Computers and Electronics in Agriculture, 207*. https://doi.org/10.1016/j.compag.2023.107708

Vilalta, R., & Drissi, Y. A. (2002). Perspective view and survey of meta-learning. *Artificial Intelligence Review, 18*, 77–95. https://doi.org/10.1023/A:1019956318069

Wang, J. X. (2021). Meta-learning in natural and artificial intelligence. *Current Opinion in Behavioral Sciences, 38*, 90-95. https://doi.org/10.1016/j.cobeha.2021.01.002

Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2020). Generalizing from a few examples: A survey on Few-shot Learning. ACM Computing Surveys, *53*(3), Article 63. https://doi.org/10.1145/3386252

Xu, X., Bao, X., Lu, X., Zhang, R., Chen, X., & Lu, G. (2023). An end-to-end deep generative approach with meta-learning optimization for zero-shot object classification. *Information Processing and Management, 60*(2). https://doi.org/10.1016/j.ipm.2022.103233

Zhang, X., Hu, C., He, B., & Han, Z. (2022). Distributed reptile algorithm for meta-learning over multi-agent systems. *IEEE Transactions on Signal Processing, 70*, 5443-5456. https://doi.org/10.1109/TSP.2022.3222734

Zhao, A., Balakrishnan, G., Durand, F., Guttag, J. V., & Dalca, A. V. (2019, June). Data augmentation using learned transformations for one-shot medical image segmentation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA*, 8535-8545. https://doi.org/10.1109/CVPR.2019.00874

Zhu, J., Jang-Jaccard, J., Singh, A., Welch, I., AI-Sahaf, H., & Camtepe, S. (2022). A few-shot meta-learning based Siamese neural network using entropy features for ransomware classification. *Computers and Security, 117*. https://doi.org/10.1016/j.cose.2022.102691

AUTHORS

Wiwin Suwarningsih graduated with a degree in Informatics in 2000 and obtained her doctoral degree in 2017 from the School of Electrical and Informatics Engineering at the Bandung Institute of Technology. Currently, she is affiliated with the Research Center for Data and Information Sciences at the National Agency of Research and Innovation (BRIN, Indonesia). Her research interests encompass artificial intelligence, computational linguistics, image processing, data mining, information retrieval, and visual question answering systems. Since 2020, she has been leading a project focused on using deep learning for identifying chili varieties, with the aim of maintaining seed purity and supporting the certification of high-quality chili seeds. One of the study's objectives is to provide an early detection system for chili varieties. For further inquiries, she can be contacted at wiwi005@brin.go.id

Rinda Kirana obtained her bachelor's and master's degrees at Padjadjaran University. She obtained the doctoral degree from the Institut Teknologi Bandung. She worked at the Indonesian Vegetable Research Institute (IVEGRI), Ministry of Agriculture, for 20 years as a researcher/breeder. During her career, she has released 18 varieties of chili peppers. Since 2022, she has been working at the Research Center for Horticultural and Estate Crops, the Research Organization for Agriculture and Food, National Research, and Innovation Agency (BRIN). Currently, her research focuses on chili pepper breeding using deep learning with artificial intelligence (AI) features as a plant selection tool and on uncovering potential metabolites of local chili genetic resources as nutritious food and raw materials for medicinal plants. She can be contacted at rind005@brin.go.id.

Purnomo Husnul Khotimah received her bachelor's degree from Universitas Indonesia in 2005 and her master's degree from Institut Teknologi Bandung, Indonesia, in 2009. Later, she obtained her Ph.D. degree in Informatics from Kyoto University, Japan in 2018. She attained a postdoctoral position in Kyoto University Hospital, Japan, from 2018 to 2020. Currently, she is a researcher at the Research Center for Data and Information Sciences of the National Research and Innovation Agency (BRIN). Her research interests are data mining, information retrieval, data integration, information systems, distributed information systems, web-based implementations, and open source. Currently, she works on the topics of, but not limited to, health informatics, smart and precision farming, autonomous vehicles, and small and media enterprises. She can be contacted at purn005@brin.go.id

Andri Fachrur Rozie received his undergraduate degree, majoring in Informatics Engineering, and a master's degree from the Department of Computer and Radio Communications Engineering at Korea University. His research areas are data science, machine learning, natural language processing, and autonomous vehicles. Currently, he is working at the Research Center for Data and Information Sciences, National Agency of Research and Innovation (BRIN, Indonesia). He was previously involved in conducting research for the development of data and information center systems for weather and air quality to support natural resource management and in NLP research for text classification. Since 2021, he has been involved in the research of autonomous vehicles. His responsibility includes creating environment simulation for autonomous vehicle testing using RoadRunner and Carla and the teleoperation dashboard. He can be contacted via email at andri.fachrur.rozie@brin.go.id

Dianadewi Riswantini received a scholarship from the Overseas Fellowship Program (a collaboration between the Indonesian Government and World Bank) for a bachelor's and master's degree in computer science from the Delft University of Technology, the Netherlands, completed in 1994. She has her second master's degree in business administration and is currently a Ph.D. candidate in the School of Business and Management, Bandung Institute Technology, Indonesia. She engaged in big data analytics for business and management. She is a member of the Information Retrieval Research Group at the Research Center for Data and Information Sciences, National Agency of Research and Innovation (BRIN, Indonesia). Her research interests include data analytics, text mining, natural language processing, and machine learning in the fields of social and medical informatics. She can be contacted at dianadewi.riswantini@brin.go.id

Ekasari Nugraheni obtained a bachelor's degree in information management from the AKRIND Institute of Technology Yogyakarta in 1996. She received a scholarship from the Indonesian Ministry of Research and Technology for her master's degree in informatics from the Bandung Institute of Technology, Indonesia, in 2016. Currently, she is a researcher of the Information Retrieval Research Group at the Research Center for Data and Information Sciences, National Research and Innovation Agency (BRIN, Indonesia). Her research interests include data analysis, data mining, deep learning, and natural language processing. She has been involved in many research areas, such as asset management, tide monitoring, semantic data warehouse, emotion recognition, human activities recognition, and others. Her latest research project is mining social media for public perception toward the halal food industry. She can be contacted at ekasari.nugraheni@brin.go.id

Devi Munandar received his undergraduate degree in Informatics Engineering and obtained a master's degree from the Department of Mathematics, Universitas Padjadjaran, Bandung, Indonesia. His research is in the fields of data science, machine learning, natural language processing, multivariate statistics, and prediction analysis. Currently, he is working at the Research Center for Data and Information Sciences, National Agency of Research and Innovation (BRIN, Indonesia). He was previously involved in conducting research for the development of data and information center systems for weather and air quality to support natural resource management and in NLP research for text classification and POS-Tagging on Indonesian-language tweets. From 2020 to 2022, he was involved in statistical modeling research and stochastic processes, using principal component analysis (PCA) integrated with vector autoregressive (VAR) and a data mining approach for forecasting climate effects. He can be contacted at devi.munandar@brin.go.id

Andria Arisal (Member, IEEE) received his Bachelor of Engineering in Informatics Engineering from Institut Teknologi Bandung and Master of Engineering in Distributed Computing from the University of Melbourne. Currently, he is working as a researcher at the Research Center for Data and Information Sciences of the National Research and Innovation Agency (BRIN). He is involved in various multi-modal data management and analytics. He also operates institutions' high-performance computing infrastructure and supports other researchers in using the infrastructure. His research interests are data management and analytics in distributed and parallel computing environments. He can be contacted at: andria.arisal@brin.go.id

Noor Roufiq Ahmadi received his Bachelor of Agricultural Technology from Stiper Agricultural Institute in 1998, completed the master's program in the Food Science and Technology Study Program at Gadjah Mada University in 2004, and earned his Doctorate in the Agricultural Industrial Technology Study Program from Bogor Agricultural Institute in 2012. Currently, he serves as the head of the Center for Standard Testing of Vegetable Crop Instruments. His field of expertise is Middle Expert Standardization Analyst. The research he conducts includes agricultural products and perishables after harvesting, as well as the development of superior commodity-based agrotourism areas. He can be contacted at noorroufiqa@gmail.com