Chili leaf segmentation using meta-learning for improved model accuracy

Wiwin Suwarningsih[1], Rinda Kirana[2], Purnomo Husnul Khotimah[1], Dianadewi Riswantini[1], Andri Fachrur Rozie[1], Ekasari Nugraheni[1], Devi Munandar[1], Andria Arisal[1], Noor Roufiq Ahmadi[3]
[1]Research Center for Data and Information Sciences, National Research and Innovation Agency, Bandung, Indonesia
[2]Research Center of Horticulture, National Research and Innovation Agency, Bandung, Indonesia
[3]Center for Standard Testing of Vegetable Plant Instruments, Ministry of Agriculture, Bandung, Indonesia

Article Info

Article history:

Received Nov 24, 2023
Revised Dec 2, 2024
Accepted Dec 25, 2024

Keywords:

Chili leaf image
K-shot classification
Meta-learning
Model agnostic
N-way classification

ABSTRACT

Recognizing chili plant varieties through chili leaf image samples automatically at low costs represents an intriguing area of study. While maintaining and protecting the quality of chili plants is a priority, classifying leaf images captured randomly requires considerable effort. The quality of the captured leaf images significantly impacts the development of the model. This study applies a meta-learning approach to chili leaf image data, creating a dataset and classifying leaf images captured using mobile devices with varying camera specifications. The images were organized into 14 experimental groups to assess accuracy. The approach included 2-way and 3-way classification tasks, with 3-shot, 5-shot, and 10-shot learning scenarios, to analyze the influence of various chili leaf image factors and optimize the classification and segmentation model's accuracy. The findings demonstrate that a minimum of 10 shots from the meta-test dataset is sufficient to achieve an accuracy of 84.87% using 2-way classification meta-learning combined with the mix-up augmentation technique.

Corresponding Author:

Wiwin Suwarningsih
Research Center for Data and Information Sciences, National Research and Innovation Agency (BRIN)
Tower 2 Lt. 5 Jl. Sangkuriang 21 Bandung, Jawa Barat, Indonesia
Email: wiwin.suwarningsih@brin.go.id

1. INTRODUCTION

Deep learning has emerged as a critical technology in image-based plant identification, finding applications in areas like plant variety recognition [1], [2], plant disease detection [3], [4], and species classification [1], [5], [6]. While these methods offer transformative potential for improving agricultural productivity and efficiency [7], [8], they also present challenges. Cutting-edge deep learning techniques enable the development of image-processing models [9], [10], simulating human cognitive processes by processing input data with weighted connections and biases. This approach facilitates accurate categorization and detailed object descriptions in datasets [11], [12]. However, deep learning models heavily depend on the availability of extensive datasets [4], [11], [13]. With limited data, achieving consistent and reliable results becomes increasingly difficult. Neural networks require substantial and diverse datasets to grasp complex patterns effectively [14], [15], making it challenging in domains like agriculture [6], [16] and environmental health [17], [18], where data scarcity is common. Achieving optimal performance often necessitates iterative experimentation. Notably, deep learning accuracy tends to improve with larger datasets [3], [5], [13]. To mitigate challenges posed by limited data, researchers have utilized meta-learning [13], a strategy described by [19], as "learning to learn." This approach systematically evaluates the performance of various machine-learning algorithms across diverse tasks, reducing the number of trials needed to achieve better predictions in less time. By leveraging the outputs and metadata from existing models, meta-learning refines predictions efficiently.

Several few-shot learning techniques are grounded in meta-learning, which emphasizes preparing systems to operate effectively with limited training data a characteristic feature in these scenarios. Interestingly, meta-learning has also been applied in other domains, such as adapting swiftly to variations in video tracking tasks [20]. The methods utilized in meta-learning can generally be grouped into two distinct categories. The first group leverages specialized network architectures to encapsulate knowledge obtained during the meta-learning phase. Examples of such techniques include fast weights [21], neural plasticity values [22], custom update rules [23], temporal convolutions [24], and long short-term memory (LSTM) memory modules [25], [26]. This approach excels in refining architectures to efficiently encode meta-learning information. Nonetheless, a key limitation lies in its dependency on specific architectural designs, which hampers the seamless integration of novel network innovations with meta-learning frameworks. In systems employing these custom architectures, the learning process deviates from traditional workflows, necessitating a robust strategy to optimize the tailored encoding.

The use of transfer learning techniques has shown promising results in enhancing the detection of plant diseases. Nevertheless, the challenge of identifying the most appropriate pre-trained model for a particular task often demands significant time and computational resources [27]. Meta-learning methods provide a more efficient alternative by drawing on insights from prior evaluations of models across various datasets to simplify the selection process for new tasks. For example, studies [19], [28] utilized meta-learning approaches to identify the optimal image segmentation algorithm tailored to specific image characteristics. This approach involves deriving meta-features, such as the mean value of a channel, entropy within intensity channels, and the Spearman correlation coefficient between different channels.

Another study, [29] explored meta-learning to utilize knowledge from comparable few-shot tasks, aiming to enable fine-tuning of a base-learner for novel tasks using only limited samples. Their approach involved consolidating available datasets and training a model capable of managing diverse examples across various categories. Subsequently, the feature extraction module of the model was retained intact, while the classifier portion was removed. The retained weights were then applied to initialize the meta-training phase for a new task.

An alternative strategy proposed by [30] utilized meta-learning to address few-shot classification problems in agriculture. This approach prioritized adapting convolutional neural network (CNN) models for novel tasks by incorporating insights drawn from prior tasks. Similarly, Zhai and Wibowo [31], implemented meta-learning to develop CNN models tailored for plant disease detection, emphasizing robust performance against noisy data. Their innovative method featured a rectification module aimed at mitigating the influence of biased image samples. While both strategies were centered on plant disease classification, their primary focus lay in refining training methodologies for novel models. Unlike these approaches, the current study highlights the creation of ranked recommendations for benchmark models applied to previously unseen plant disease classification tasks.

The core idea of meta-learning is to accumulate knowledge from experience [19], or meta-data, enabling the rapid acquisition of new tasks at a much faster pace than would be achievable otherwise. The formalization of meta-learning can be expressed through mathematical frameworks, employing conventional supervised machine learning concepts. Meta-learning effectively leverages knowledge from a limited set of acquired examples, bridging gaps when dealing with small training datasets. In meta-learning, the approach involves training multiple models rather than maintaining a single model, with the objective of optimizing performance across various shot setups. One approach that stands out for its ability to achieve this without the need to assemble numerous models and establish a network of relationships is the model agnostic meta-learning (MAML) method. MAML aims to provide an adaptable foundation for meta-learning. The meta-learning initialization presented in this paper is designed to generate structured output for the segmentation of chili leaf images.

The study delves into generating tasks that incorporate synthetic reward functions, eliminating the need for supervision. By using this approach, the policy network undergoes meta-training on these synthetic tasks, which equips it to learn real-world tasks defined by manual reward functions with greater efficiency and fewer data samples:

a. How can one investigate the impact of N-way and K-shot variations on the performance of few-shot classification through an extensive series of experiments?
b. How can samples of a well-balanced dataset suitable for few-shot classification be generated?
c. Can the MAML algorithm be applied to the chili image dataset?

The main contributions of this study are twofold. Firstly, it demonstrates the high potential of meta-learning to enhance the accuracy of chili variety identification, even with limited datasets. Secondly, it introduces augmentation techniques for generating samples to balance the chili image dataset.

The paper is organized as follows: section 2 describes the proposed method, while section 3 introduces the experimental validation and discussion. Finally, section 4 concludes by summarizing the findings and highlighting potential future directions.

2. METHOD

In this research, an existing model, namely MAML [27], was utilized as the foundation for evaluating the compatibility of the dataset with the MAML model. The objective of this approach was to initiate the analysis process based on existing models. This decision was motivated by the limited dataset available. Consequently, a meta-learning process was conducted to enhance the model's capability in accurately identifying chili varieties. The process began with a dataset in which each sample was labeled with its corresponding class. The training dataset was constructed by randomly selecting N distinct classes, denoted as $C_{L\text{-}i}$, and then selecting a random sample, Xi, from each of these chosen classes. Subsequently, for the validation dataset, a distinct sample, X'_i, was chosen from the same class as the corresponding training sample (refer to Figure 1). In this study, the meta-learner was an automated reweighting algorithm designed to transform imbalanced data into a more balanced distribution by adjusting sample weights [19], [28].

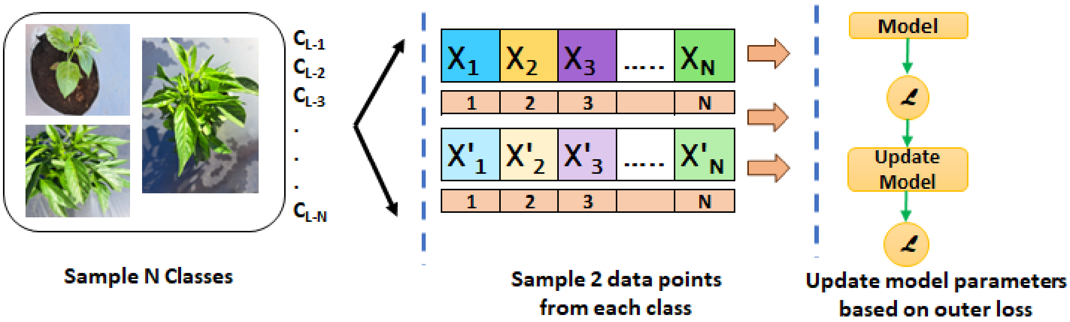


Figure 1. Chili leaf segmentation research method

2.1. Dataset

The dataset used in this study was obtained from an assembly garden of various chili plant varieties, including large red chilies, curly red chilies, and cayenne peppers. The large red chilies consist of 9 varieties, the curly red chilies consist of 7 varieties, and the cayenne peppers consist of 11 varieties. The data can be accessed through the following link: https://data.brin.go.id/dataset.xhtml?persistentId=hdl:20.500.12690/RIN/QSKTET. Table 1 provides details on the grouping of leaves and the corresponding number of images. Additionally, the images in each combination were divided into training and testing datasets using a 75:25 split ratio to evaluate the performance of pre-trained models. Combinations labeled L1 to L10 were used to generate metadata and train the meta-learner, while combinations labeled L11 to L14 were reserved for testing the proposed framework. The selection of combinations for testing was based on specific factors that highlight the effectiveness and practicality of the proposed framework.
L11: three morphological types characterized by a flat lancet shape, an evenly elongated form, and a serrated lanceolate structure.
L12: three morphological variations defined by a flat lancet shape, an evenly elongated design, and a serrated elongated structure.
L13: three morphological categories identified as flat elliptical, flat oval, and serrated elliptical forms.
L14: three morphological groups distinguished as flat elliptical, flat oval, and serrated ovate shapes.

2.2. Model agnostic meta-learning data training

In this research, a meta-task is defined as a training task involving sample data. When dealing with a task with a limited amount of raw data, meta-learning can be categorized as either a few-shot learning challenge or a zero-shot learning challenge.

For the MAML algorithm [27], the data needs to be in the form of (x, y), requiring a total of N-way * K-shots of these pairs. In the context of MAML algorithm, the data is typically structured as pairs of (x, y), where: ‘x’ represents the input data or features of a given example. In the context of meta-learning, 'x' is often an input image, sequence, or any form of data relevant to the task at hand. For example, if you are working on an image classification task, 'x' would be the pixel values of an image. ‘y’ represents the corresponding label or output associated with the input 'x'. In classification tasks, 'y' usually denotes the class or category to which the input belongs. It could be a one-hot encoded vector for multi-class classification or a scalar value for binary classification. So, for an N-way K-shot meta-learning task, you would need a total of N classes with K examples each, resulting in N * K pairs of (x, y). Each pair corresponds to a single task in the meta-learning setup, where the algorithm seeks to train a model capable of rapidly adapting to new tasks based on a limited number of examples (K-shots) per class (N-way).

Table 1. Leaves grouping based on morphology combination

Combination	Leaves morphology	#Images
L1	Lance-shaped and extended in length.	193
L2	Oval-shaped and egg-like.	232
L3	Lanceolate with serrated edges and elongated with serrated features.	105
L4	Elliptical shape with serrated edges and ovate shape with serrated edges.	94
L5	Lancet-shaped with a flat surface and evenly elongated shape.	264
L6	Elliptical shape with a flat surface and oval shape with a flat surface.	166
L7	Lancet-shaped with a flat surface, lanceolate with serrated edges, and elongated with serrated features.	75
L8	Evenly elongated shape, lanceolate with serrated edges, and elongated shape with serrated features.	140
L9	Elliptical shape with a flat surface, elliptical shape with serrated edges, and ovate shape with serrated edges.	40
L10	Oval shape with a flat surface, elliptical shape with serrated edges, and ovate shape with serrated edges.	148
L11	Lancet-shaped with a flat surface, evenly elongated shape, and lanceolate shape with serrated edges.	151
L12	Lancet-shaped with a flat surface, evenly elongated shape, and elongated shape with serrated edges.	122
L13	Elliptical shape with a flat surface, oval shape with a flat surface, and elliptical shape with serrated edges.	157
L14	Elliptical shape with a flat surface, oval shape with a flat surface, and ovate shape with serrated edges.	87

During the meta-training phase, the specific labels used are not significant since they are discarded. They can be substituted with artificial labels, which can be assigned from the set {1, 2, ..., N}. However, it is crucial to ensure that these artificial labels maintain the distinctions between classes. In other words, simply put, data points with the same label were assigned an identical artificial label, while those with different labels were given distinct artificial labels.

2.3. Model agnostic meta-learning data validation

Within the framework of the MAML approach [27], it is important to highlight that the verification data used in tasks designed for meta-training serves as input for the outer loop of training. Consequently, it is necessary to create a validation dataset consisting of pairs such as $(x'_1, 1)$, $(x'_2, 2)$, ..., (x'_N, N) for every task. Therefore, appropriate validation data must be generated for the artificial task.

One fundamental prerequisite for this verification data is that it must be precisely annotated within the specified context. In practical terms, this implies that the synthetic numerical label must correspond to the identical category in the unlabeled dataset in both instances. In other words, for a given class C, it should be the case that both the training sample (x_i) and its corresponding validation sample (x'_i) belong to this same class, represented as follows: for a class C, x_i, and x'_i should both fall within C.

2.4. Pre-training agnostic model

The study encompasses all major categories of pre-trained model-agnostic techniques, including transfer learning (Matching-NET) and MAML with Cut-Out, Mix-Up, and Cut-Mix augmentation techniques, to assess the effectiveness of the proposed framework for chili leaf variety classification tasks. Furthermore, the meta-learner selects the top-performing models from these categories for each new task. Table 2 provides details about the four model-agnostic techniques that have been trained.

Table 2. Trained model agnostic

Model	Depth
Transfer learning	17
MAML-Cut-Out	20
MAML-Mix-Up	59
MAML-Cut-Mix	72

The augmentation techniques (Cut-Out, Mix-Up, and Cut-Mix) are employed to augment training data and improve the performance, generalization, and robustness of deep learning models, particularly in the context of image classification tasks. Each technique introduces unique variations to the training process, encouraging the model to become more adaptable and resistant to overfitting.

Cut-Out is an image augmentation technique commonly used in computer vision tasks, including deep learning models for image classification [4]. During the training process, a random rectangular portion of the image is removed or "cut out," creating a black region in that specific area. This aims to force the model to focus on other parts of the image, promoting robust feature learning and preventing over-reliance on specific details. The cut-out region is usually replaced with either zeros or the mean pixel values.

Mix-Up is a data augmentation technique that involves blending two or more images and their corresponding labels during the training process [10]. This blending is done by taking a weighted sum of the pixel values and labels of two randomly selected images. This method helps enabling the model to generalize more effectively by introducing a combination of different patterns and features from multiple images. Mix-Up is especially beneficial for avoiding overfitting and boosting the model's ability to handle variations in the input data.

Cut-Mix is a combination of Cut-Out and Mix-Up techniques [10], [15]. It involves randomly selecting a region (like Cut-Out) from one image and replacing it with the same region from another image while mixing their labels. This process effectively combines the features of two different images, encouraging the model to learn from the information present in both. Cut-Mix is designed to strengthen the robustness and generalization abilities of the model, like Mix-Up, while also introducing local modifications through the cut-out operation.

Similar to how the activations of a meta-model trained on a visual recognition task act as a rich representation of input images, the weights acquired during the model's task-specific training provide valuable representations of that task. Thus, in this study, these learned weights are used as meta-features to uniquely represent the dataset.

2.5. Updating agnostic model

The primary objective of this research is to recommend suitable models for chili variety identification, aiming to streamline the process and save time and resources. This involves reducing the number of models under consideration while maintaining high predictive performance. To accomplish this, rank-biased overlap (RBO) [29] is employed as a method to assess how well the predicted ranking of models aligns with the original ranking generated by evaluating each model on the dataset. The RBO score varies from 0 to 1, with 1 indicating that the two rankings are identical and 0 indicating no similarity between them. In (1) defines the RBO metric for comparing two infinite ranked lists, denoted as $L1$ and $L2$:

$$RBO\ (L_1, L_2, p) = (1-p)\sum p^{d-1} A_d \tag{1}$$

In (1), where d ranges from 1 to infinity (representing the depth of the ranking under consideration), X_d signifies the size of the overlap between the rankings L_1 and L_2 up to a depth of d. A_d is calculated as the ratio of X_d to d, and the parameter p is a tunable parameter that falls within the range (0,1).

3. RESULTS AND DISCUSSION

3.1. Experimental setting

The performance of our method was evaluated across three chili image datasets, as depicted in Table 2, which integrates leaf morphology information from the chili leaf dataset. Our evaluations encompassed 2-way and 3-way classification tasks, spanning 3-shot, 5-shot, and 10-shot learning scenarios, and employed a combination of conventional and sophisticated augmentation techniques. Furthermore, a transfer-learning strategy was evaluated, where the model was initially trained on all meta-train classes and subsequently fine-tuned using a few samples from the meta-test set. In meta-learning, standard augmentation techniques like flipping and rotating offer limited extra information for model training. This is due to the fact that, during meta-training, both images and classes are sampled. Therefore, it is necessary to create novel classes with new images. It also involved an evaluation of augmentation techniques such as Cut-Out, Mix-Up, and Cut-Mix for medical data (example images is illustrated in Figure 2).

As evident from Figure 2, these augmentation strategies are exclusively utilized during the meta-training stage to mitigate the risk of overfitting. A brief explanation of the three augmentation techniques employed - Cut-Out, Mix-Up, and Cut-Mix - will now be provided. Figure 2(a) displays the original chili leaf dataset without any augmentation, serving as a baseline representation of the input data. The Cut-Out technique, as described in reference [30], involves the random creation of a square mask, with pixel values within this mask set to zero. An example of a batch of chili leaf images with Cut-Out augmentation is illustrated in Figure 2(b). With the application of Cut-Out, features are removed at the initial input stage, ensuring that no feature map includes features related to the masked area.

Mix-Up is an augmentation technique aimed at enhancing the generalization capabilities of deep learning models by generating virtual samples from the existing data distribution. It has been observed that Mix-Up augmentation introduces subtle alterations to the images, which might go unnoticed without close examination. Figure 2(c) displays images that have been augmented using Mix-Up, where pixel values from two images are blended, leading to an intermediate representation of both samples.

Another augmentation process is Cut-Mix. The fundamental concept behind Cut-Mix is the generation of a new sample by cutting out a portion of an image and incorporating it into a different image from the training set. Additionally, the labels corresponding to the ground truth are combined in proportion to the area of the cut sections. Figure 2(d) illustrates the application of Cut-Mix, where distinct portions of different leaf images are merged, creating hybrid samples that enrich the dataset and improve model robustness.

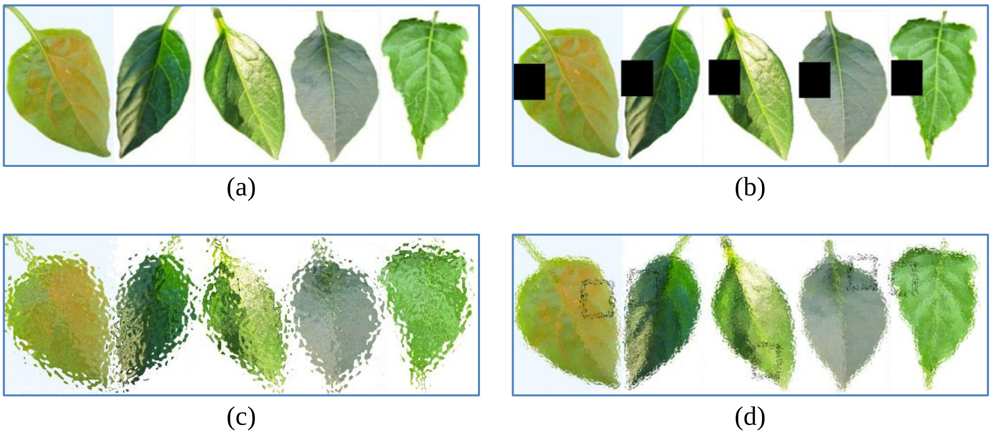


Figure 2. Example chili leaf dataset with; (a) no augmentation (original), (b) Cut-Out, (c) Mix-Up, and (d) Cut-Mix

3.2. Experimental result

3.2.1. Result of evaluation metric

Throughout all experiments, accuracy (%) was utilized as the primary evaluation metric, consistent with standard practices in classification tasks. Importantly, the accuracy metric demonstrated robustness during meta-testing, attributed to the balanced distribution of images across all classes, which minimized the effect of class imbalance. The reported accuracy reflected the average performance across few-shot tasks that were randomly sampled from the meta-testing dataset.

The model's performance was assessed in 3-shot, 5-shot, and 10-shot learning scenarios for both 2-way and 3-way classification tasks. For each of these k-shot, n-way scenarios, the evaluation was repeated using three distinct augmentation techniques: Cut-Out, Mix-Up, and Cut-Mix. Additionally, transfer learning evaluations were also conducted.

The evaluation focused on three separate leaf image datasets. Assessments were carried out for both 2-way and 3-way classification tasks, encompassing 3-shot, 5-shot, and 10-shot learning scenarios. Both conventional and advanced augmentation techniques were applied during these evaluations. Furthermore, a transfer learning approach was explored, wherein the model was initially trained on all meta-training classes and subsequently fine-tuned using multiple samples from the meta-test set

Transfer learning is a widely recognized approach for developing deep learning classification models in scenarios with limited data [16]. Consequently, transfer learning was chosen as the baseline method (refer to Table 3 for a comparison between meta-learning and transfer learning). During the transfer learning process, the entire meta-training dataset was employed for supervised learning across 500 epochs. After selecting the best model, it was fine-tuned following the same approach followed during our meta-learning trials. Consistent network architecture and hyperparameters were maintained across all experiments to ensure a fair comparison.

In Table 3, you can find the outcomes regarding the transfer learning technique displayed in the third column for 2-way classification and the seventh column for 3-way classification. However, in the case of 3, 5, and 10-shot learning problems, the transfer learning strategy failed to achieve the top test accuracy in either the 2-way or 3-way classification tasks.

Table 3. Experiment result of evaluation metric

Few-shot task	2-way classification meta learning					3-way classification meta learning				
	MAML-traditional (%)	MAML-Cut-Out (%)	MAML-Mix-Up (%)	MAML-Cut-Mix (%)	Transfer learning (%)	MAML-traditional (%)	MAML-Cut-Out (%)	MAML-Mix-Up (%)	MAML-Cut-Mix (%)	Transfer learning (%)
3-shot	75.75	77.25	82.88	84.12	75.62	63.08	67.33	68.83	70.58	61.25
5-shot	77.62	79.75	83.25	80.88	77.85	68.00	63.41	66.50	68.25	64.78
10-shot	79.75	83.37	84.87	81.75	79.12	74.16	73.91	75.12	74.24	69.98

3.2.2. The experiment investigated the impacts of Cut-Out, Mix-Up, and Cut-Mix

The results across various experiments on the Combination Leaves Morphology Grouping dataset, spanning all magnification levels, suggest that the utilization of these advanced augmentation methods generally contributes to improved accuracy. However, a few experiments did not demonstrate the same positive impact. In both the two-combination leaves morphology and three-combination leaves morphology datasets, incorporating augmentation markedly enhances the model's accuracy. While a conclusive winner is not evident in the comparison among the three advanced augmentation techniques—Cut-Out, Mix-Up, and Cut-Mix—Cut-Mix tends to outperform the others in most of the experiments. The application of advanced augmentation introduces challenging samples to the model during meta-training, facilitating the learning of more effective representations. A comparable impact of augmentation, specifically in mitigating the overfitting issue observed in the chili leaves dataset, was noted.

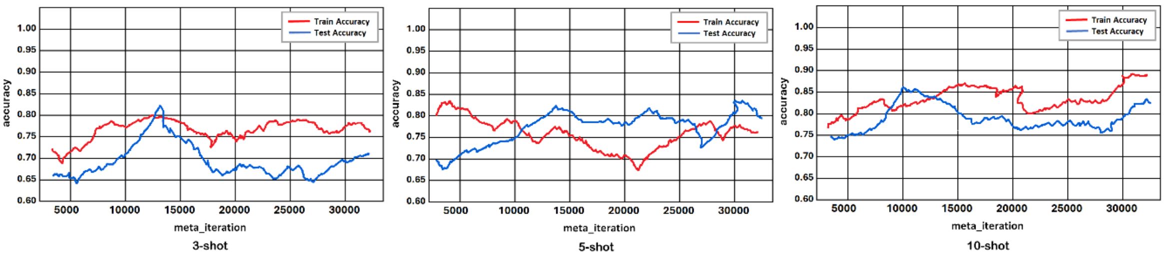
Figure 3 illustrates the accuracy versus meta-iterations plot for 3, 5, and 10-shot learning in the context of a 2-way classification task. The model was trained over the range of 5,000 to 30,000 epochs to obtain optimal parameters and enhance accuracy. Figure 3(a) presents the results using the Cut-Out augmentation technique, where certain features are randomly removed, forcing the model to learn robust representations despite missing information. Figure 3(b) shows the accuracy progression under the Mix-Up augmentation, which blends two samples to generate virtual data points, improving generalization. Lastly, Figure 3(c) depicts the results using Cut-Mix augmentation, where portions of different images are combined, leading to stronger feature representations and enhanced model performance.

In Figure 3, the graph illustrates the dynamic changes in a model's accuracy as it undergoes multiple meta-iterations, specifically in the context of various scenarios related to few-shot learning within a 2-way classification task. The plot serves to visually showcase the model's performance evolution over time and iterations during the meta-learning process. It provides a clear depiction of how the accuracy of the model develops and adapts across different stages of meta-iterations. The graph not only captures the fluctuations in accuracy but also highlights the correlation between the number of meta-iterations and the achieved accuracy of the model. Essentially, meta learning iterations are conducted within the range of 5,000 to 30,000 with the objective of enhancing accuracy, facilitating the acquisition of optimal parameter configurations. Meta learning encompasses the adjustment of model or algorithm parameters. Through iterative processes within this range, this research identifies optimal parameter values to augment accuracy. The iteration involves the continuation of the learning process by incorporating additional data or training batches. The aim is to enhance the model's capabilities by gaining a deeper understanding of a broader spectrum of data. In instances where the model needs to adapt to dynamic changes in the data or environment, iteration proves beneficial, enabling the model to continually update and enhance accuracy over time.

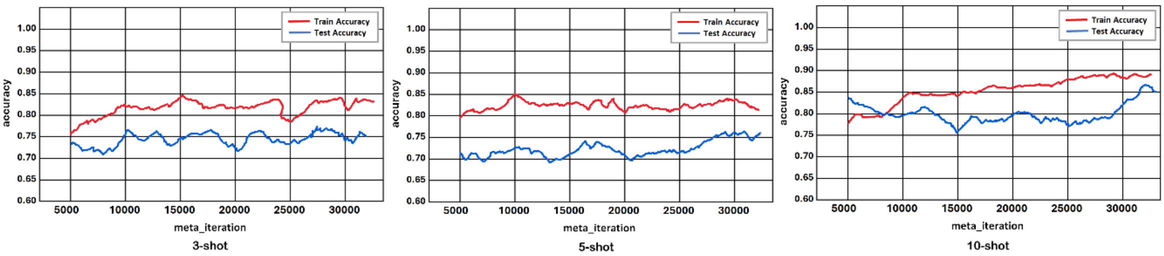
The depicted plot or graph is focused on scenarios related to 3, 5, and 10-shot learning. In these scenarios, the plot provides information for three specific instances of few-shot learning, each involving a different number of shots: 3-shot, 5-shot, and 10-shot. The term "few-shot learning" refers to a machine learning paradigm where a model is trained and evaluated on a small number of examples per class. The crucial aspect highlighted is that, within these scenarios, the model is subject to evaluation and adaptation based on a limited number of examples, referred to as "shots," for each class. The mention of "N-way" signifies the number of classes involved in the classification task. In this specific context, N is specified as 2, indicating that the model is engaged in a 2-way classification task. In a 2-way classification task, the model categorizes input data into one of two classes or categories. Therefore, the plot analyzes the model's performance under conditions where it is evaluated and adapted with a restricted number of examples per class, focusing on scenarios involving 3, 5, and 10 shots in the context of a 2-way classification task.

The analysis is conducted within the framework of a 2-way classification task. In this specific type of classification task, the primary objective is to train the model to categorize input data into two distinct categories or classes. The decision to adopt a 2-way classification task implies that there are only two possible outcomes or labels for each example the model encounters. This choice of a 2-way classification task simplifies the prediction task for the model, as it only needs to assign input data to one of the two predefined categories. In contrast to tasks with multiple classes, where each example could belong to one of several categories, a 2-way classification task reduces the complexity of the learning problem.

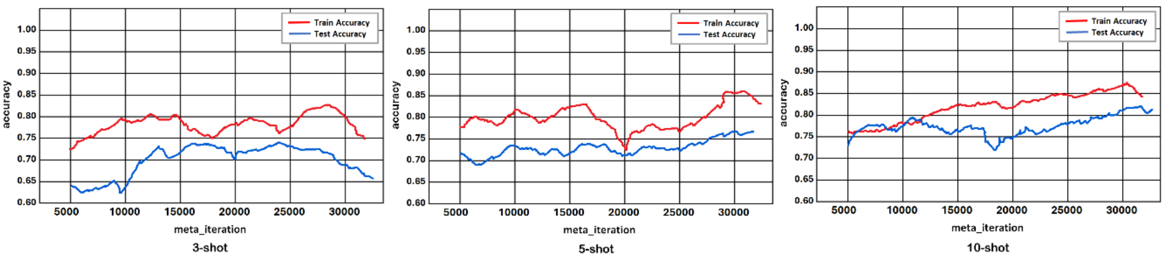
The outcomes across many experiments at various magnification levels suggest that employing advanced augmentation is generally advantageous for enhancing accuracy, with only a few exceptions. The substantial improvement in model accuracy is attributed to the use of augmentation. While there is no definitive winner among the three advanced augmentation techniques Cut-Out, Mix-Up, and Cut-Mix it is observed that Cut-Mix tends to outperform the others in most experiments. The application of advanced augmentation introduces challenging samples during meta-training, enabling the model to acquire more robust representations.



(a)



(b)



(c)

Figure 3. Meta-iteration versus accuracy in 2-way classification for; (a) cut-out, (b) mix-up, and (c) cut-mix

3.3. Discussion

Based on the outcomes of the conducted experiments (refer to Table 3), it is evident that the 2-way classification task consistently exhibits superior performance compared to the 3-way classification task across all variations and scenarios of the few-shot task, encompassing 3-shot, 5-shot, and 10-shot setups. The inherent complexity of the 3-way classification task contributes to its difficulty, as reflected in consistently lower performance across different variations of the few-shot task. This suggests that the model faces increased challenges in accurately classifying data into three distinct categories, making the 3-way classification inherently more demanding than its 2-way counterpart.

Analyzing the performance of MAML compared to transfer learning reveals that, in all scenarios of multiple-shot tasks (3-shot, 5-shot, 10-shot), 2-way classification with MAML-Cut-Mix consistently delivers the most optimal results. Additionally, for 3-shot and 5-shot tasks, MAML-Cut-Out and MAML-Mix-Up exhibit competitive performance. While transfer learning contributes positively, it does not consistently surpass the effectiveness of MAML-Cut-Mix. On the other hand, in the context of 3-way classification, MAML-Cut-Mix maintains its superior performance across all scenarios of few-shot tasks. Although transfer learning shows positive contributions, particularly in the 10-shot scenario, its performance remains lower compared to MAML-Cut-Mix.

In the analysis of Few-shot Task Performance, across various few-shot tasks such as 3-shot, 5-shot, and 10-shot scenarios, MAML-Cut-Mix consistently stands out by delivering the most optimal results. This consistency highlights the effectiveness of the MAML-Cut-Mix technique in adapting to new tasks, particularly those with a limited number of training samples. Furthermore, MAML-Mix-Up demonstrates commendable performance, especially in the context of 2-way classification tasks, particularly in the challenging 10-shot scenario. While transfer learning contributes significantly to the performance, it falls short compared to MAML-Cut-Mix, especially evident in 3-way classification tasks where MAML-Cut-Mix maintains its superior standing.

Therefore, the conclusion drawn is that MAML-Cut-Mix emerges as the most successful technique across all variations and scenarios in the context of the few-shot task, demonstrating its superiority in the context of this experiment. While transfer learning contributes positively, its effectiveness is contingent upon the number of shots and the nature of the classification task. Specifically, in 3-way classification tasks, where models often face challenges in learning intricate patterns, MAML-Cut-Mix continues to be the superior choice.

The exploration of this study entails navigating through constraints associated with limited datasets. The author has strategically utilized methods such as data augmentation combined with meta-learning strategies to address challenges arising from a scarcity of variations in data and cultivation conditions. The societal implications of this research carry substantial weight, given that the outcomes can play a pivotal role in identifying superior chili seeds. Consequently, this identification has the potential to inspire farmers to engage in the cultivation of high-quality chilies, ultimately resulting in more plentiful harvests.

4. CONCLUSION

This research demonstrated promising results, with the meta-learning approach achieving an accuracy of 84.87% using a minimum of 10 shots from the meta-test dataset and leveraging the mix-up augmentation method. A novel approach to meta-learning was introduced, framing leaf image classification as a few-shot learning problem in scenarios with limited data. Advanced augmentation methods including CutOut, MixUp, and CutMix were employed to enhance the model's ability to generalize. The success of this method was evaluated on three intricate leaf image datasets. Comprehensive evaluation revealed that meta-learning consistently outperformed leveraging transfer learning across all experimental datasets.

The transfer learning framework exhibited lower prediction reliability, raising potential concerns in the plant domain. However, integrating advanced augmentation techniques improved test accuracy and model consistency across datasets. Notably, transitioning from 3-shot to 5-shot, and subsequently to 10-shot experiments, significantly enhanced model performance. The choice of optimizers, hyperparameters, and their specific values played a crucial role in performance, underscoring the importance of these factors. This work is anticipated to benefit researchers employing meta-learning techniques in plant science. The findings of this study demonstrate the potential of leveraging modest datasets combined with combining data augmentation and meta-learning to tackle challenges posed by restricted datasets and diverse growth conditions. Insights from meta-learning enable accurate recognition of chili varieties even under these constraints, enhancing the process's reliability and precision.

Future research will aim to validate the methodology across broader datasets and investigate more robust regularization strategies beyond traditional image augmentation. Plans include extending this work to address challenges such as noisy labels and automating parameter optimization during training. Building on these results, further advancements in MAML techniques will focus on handling complex task distributions with significant domain gaps. Additionally, developing tailored MAML approaches for multimodal scenarios remains a key area of interest.

ACKNOWLEDGEMENTS

The authors extend gratitude to the Research Center for Data and Information Sciences BRIN, as well as the Research Center of Horticulture and Plantation BRIN, for their provision of facilities during the manuscript development process. Special thanks are also extended to the Center for Standard Testing of Vegetable Plant Instruments in Lembang, Ministry of Agriculture, for their technical support and the provision of a chili planting area for data collection.

FUNDING INFORMATION

The authors declare that no funding was received for the conduct of this research.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Wiwin Suwarningsih	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Rinda Kirana		✓				✓		✓	✓	✓	✓	✓		
Purnomo Husnul Khotimah	✓		✓	✓		✓			✓		✓		✓	
Dianadewi Riswantini					✓		✓		✓		✓		✓	
Andri Fachrur Rozie	✓	✓		✓	✓				✓	✓				
Ekasari Nugraheni		✓	✓			✓		✓		✓			✓	✓
Devi Munandar	✓		✓	✓		✓			✓		✓		✓	
Andria Arisal				✓	✓		✓		✓	✓			✓	
Noor Roufiq Ahmadi					✓		✓		✓		✓		✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

The authors also acknowledge individual contributions to this work: W.S. conceived the research idea and supervised the study, R.K. provided plant biology insights, P.H.K. implemented data augmentation techniques, D.R. prepared and managed the datasets, A.F.R. optimized machine learning frameworks, E.N. developed the meta-learning model, D.M. conducted the literature review, A.A. supported model testing and analysis, and N.R.A. refined the manuscript and ensured its scientific rigor. The authors express their gratitude to everyone who provided support and reviews, significantly contributing to the accomplishment of this work.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this research paper.

INFORMED CONSENT

This study does not involve human participants, personal data, or any activities requiring informed consent. Therefore, informed consent was not applicable to this research.

ETHICAL APPROVAL

This study does not involve human participants, animals, or sensitive data requiring ethical clearance. All research activities were conducted in accordance with ethical research standards and guidelines.

DATA AVAILABILITY

The data that support the findings of this study are openly available in the BRIN Data Repository at https://data.brin.go.id/dataset.xhtml?persistentId=hdl:20.500.12690/RIN/ND3BJI.

REFERENCES

[1] S. Hati and S. G, "Plant Recognition from Leaf Image through Artificial Neural Network," *International Journal of Computer Applications*, vol. 62, no. 17, pp. 15–18, Jan. 2013, doi: 10.5120/10172-4897.

[2] C. L. Lee and S. Y. Chen, "Classification of leaf images," *International Journal of Imaging Systems and Technology*, vol. 16, no. 1, pp. 15–23, Jan. 2006, doi: 10.1002/ima.20063.

[3] A. N. Rathod, B. Tanawal, and V. Shah, "Image Processing Techniques for Detection of Leaf Disease," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 11, pp. 1–3, 2013.

[4] T. Rumpf, A. K. Mahlein, U. Steiner, E. C. Oerke, H. W. Dehne, and L. Plümer, "Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance," *Computers and Electronics in Agriculture*, vol. 74, no. 1, pp. 91–99, Oct. 2010, doi: 10.1016/j.compag.2010.06.009.

[5] X. Liming and Z. Yanchao, "Automated strawberry grading system based on image processing," *Computers and Electronics in Agriculture*, vol. 71, no. SUPPL. 1, pp. S32–S39, Apr. 2010, doi: 10.1016/j.compag.2009.09.013.
[6] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, Jun. 2019, doi: 10.1016/j.compag.2018.03.032.
[7] U. Krishnasamy and D. Nanjundappan, "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*, vol. 77, pp. 385–394, May 2016, doi: 10.1016/j.ijepes.2015.11.022.
[8] G. Geetharamani and J. A. Pandian, "Corrigendum to: 'Identification of plant leaf diseases using a nine-layer deep convolutional neural network," *Computers and Electrical Engineering*, vol. 76, pp. 323-338, Jun. 2019, doi: 10.1016/j.compeleceng.2019.08.010.
[9] A. S. Oktaria, E. Prakasa, and E. Suhartono, "Wood Species Identification using Convolutional Neural Network (CNN) Architectures on Macroscopic Images," *Journal of Information Technology and Computer Science*, vol. 4, no. 3, pp. 274–283, Dec. 2019, doi: 10.25126/jitecs.201943155.
[10] E. Hussain, L. B. Mahanta, C. R. Das, and R. K. Talukdar, "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*, vol. 65, pp. 1–8, Aug. 2020, doi: 10.1016/j.tice.2020.101347.
[11] F. Atban, E. Ekinci, and Z. Garip, "Traditional machine learning algorithms for breast cancer image classification with optimized deep features," *Biomedical Signal Processing and Control*, vol. 81, Mar. 2023, doi: 10.1016/j.bspc.2022.104534.
[12] R. Hussin, M. R. Juhari, N. W. Kang, R. C. Ismail, and A. Kamarudin, "Digital image processing techniques for object detection from complex background image," *Procedia Engineering*, vol. 41, pp. 340–344, 2012, doi: 10.1016/j.proeng.2012.07.182.
[13] D. Li, X. Wang, J. Zhang, and Z. Ji, "Automated deep learning system for power line inspection image analysis and processing: Architecture and design issues," *Global Energy Interconnection*, vol. 6, no. 5, pp. 614–633, Oct. 2023, doi: 10.1016/j.gloei.2023.10.008.
[14] H. Noh, P. H. Seo, and B. Han, "Image question answering using convolutional neural network with dynamic parameter prediction," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2016, pp. 30–38. doi: 10.1109/CVPR.2016.11.
[15] S. B. Park, J. W. Lee, and S. K. Kim, "Content-based image classification using a neural network," *Pattern Recognition Letters*, vol. 25, no. 3, pp. 287–300, Feb. 2004, doi: 10.1016/j.patrec.2003.10.015.
[16] J. Sangeetha and P. Govindarajan, "Prediction of agricultural waste compost maturity using fast regions with convolutional neural network(R-CNN)," *Materials Today: Proceedings*, Jan. 2023, doi: 10.1016/j.matpr.2023.01.112.
[17] P. Lan, L. Guo, H. Sun, Y. Zhang, and Y. Jiang, "Modeling stream baseflow nitrate concentration in an agricultural watershed using neural network and bootstrap method," *Ecological Indicators*, vol. 156, pp. 1–7, Dec. 2023, doi: 10.1016/j.ecolind.2023.111097.
[18] Ş. Karaaslan, S. G. Kobat, and M. Gedikpınar, "A new method based on deep learning and image processing for detection of strabismus with the Hirschberg test," *Photodiagnosis and Photodynamic Therapy*, vol. 44, pp. 1–12, Dec. 2023, doi: 10.1016/j.pdpdt.2023.103805.
[19] X. Xu, X. Bao, X. Lu, R. Zhang, X. Chen, and G. Lu, "An end-to-end deep generative approach with meta-learning optimization for zero-shot object classification," *Information Processing and Management*, vol. 60, no. 2, Mar. 2023, doi: 10.1016/j.ipm.2022.103233.
[20] Z. Liu, W. Gao, J. Zhu, Z. Yu, and Y. Fu, "Surface deformation tracking in monocular laparoscopic video," *Medical Image Analysis*, vol. 86, May 2023, doi: 10.1016/j.media.2023.102775.
[21] R. Singh, V. Bharti, V. Purohit, A. Kumar, A. K. Singh, and S. K. Singh, "MetaMed: Few-shot medical image classification using gradient-based meta-learning," *Pattern Recognition*, vol. 120, pp. 1–13, Dec. 2021, doi: 10.1016/j.patcog.2021.108111.
[22] S. U. Saeed *et al.*, "Image quality assessment for machine learning tasks using meta-reinforcement learning," *Medical Image Analysis*, vol. 78, pp. 1–15, May 2022, doi: 10.1016/j.media.2022.102427.
[23] H. Tabealhojeh, P. Adibi, H. Karshenas, S. K. Roy, and M. Harandi, "RMAML: Riemannian meta-learning with orthogonality constraints," *Pattern Recognition*, vol. 140, Aug. 2023, doi: 10.1016/j.patcog.2023.109563.
[24] S. Khodadadeh, L. Bölöni, and M. Shah, "Unsupervised meta-learning for few-shot image classification," *Advances in Neural Information Processing Systems*, vol. 32, pp. 10132–10142, 2019.
[25] M. Tan, C. dos Santos, B. Xiang, and B. Zhou, "LSTM-based Deep Learning Models for Non-factoid Answer Selection," *arXiv preprint*, 2015, doi: 10.48550/arXiv.1511.04108.
[26] X. Shao, H. Wang, X. Zhu, F. Xiong, T. Mu, and Y. Zhang, "EFFECT: Explainable framework for meta-learning in automatic classification algorithm selection," *Information Sciences*, vol. 622, pp. 211–234, Apr. 2023, doi: 10.1016/j.ins.2022.11.144.
[27] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," *34th International Conference on Machine Learning, ICML 2017*, vol. 3, pp. 1856–1868, 2017.
[28] J. Zhu, J. Jang-Jaccard, A. Singh, I. Welch, H. AL-Sahaf, and S. Camtepe, "A few-shot meta-learning based siamese neural network using entropy features for ransomware classification," *Computers and Security*, vol. 117, Jun. 2022, doi: 10.1016/j.cose.2022.102691.
[29] P. H. T. Gama, H. Oliveira, J. A. dos Santos, and R. M. Cesar, "An overview on Meta-learning approaches for Few-shot Weakly-supervised Segmentation," *Computers and Graphics (Pergamon)*, vol. 113, pp. 77–88, Jun. 2023, doi: 10.1016/j.cag.2023.05.009.
[30] S. Verma, P. Kumar, and J. P. Singh, "A meta-learning framework for recommending CNN models for plant disease identification tasks," *Computers and Electronics in Agriculture*, vol. 207, Apr. 2023, doi: 10.1016/j.compag.2023.107708.
[31] C. Zhai and S. Wibowo, "A systematic review on cross-culture, humor and empathy dimensions in conversational chatbots: the case of second language acquisition," *Heliyon*, vol. 8, no. 12, pp. 1–13, 2022, doi: 10.1016/j.heliyon.2022.e12056.

BIOGRAPHIES OF AUTHORS

Wiwin Suwarningsih graduated from her bachelor's degree at Informatics Program, Adityawarman Institute of Technology Bandung in 1996. She got her master's degree in 2000 at the Informatics Study Program and doctoral degree in 2017 at the School of Electrical and Informatics Engineering, Bandung Institute of Technology. Currently, she works at the Research Center for Data and Information Sciences, National Agency of Research and Innovation (BRIN, Indonesia). Her research interests are artificial intelligence, computational linguistics, Indonesian natural language processing and text mining, information retrieval and question answering systems. Since 2020, she has led a project about the identification of chili varieties using deep learning to maintain seed purity and supporting the certification of high-quality chili seeds. One of the study's objectives is to provide early detection system for chili varieties. She can be contacted at email: wiwin.suwarningsih@brin.go.id.

Rinda Kirana earned her bachelor's and master's degrees at Padjadjaran University. She obtained the Doctor from Institut Teknologi Bandung. She worked at Indonesian Vegetable Research Institute (IVEGRI), Ministry of Agriculture for 20 years as a researcher/breeder. During her career he has released 18 varieties of chili pepper. Since 2022, she has been working at Research Center for Horticultural and Estate Crops, 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 plant selection tool and research on uncovering potential metabolites of local chili genetic resources as nutritious food and raw materials for medicinal plants. She can be contacted at email: rind005@brin.go.id.

Purnomo Husnul Khotimah received her bachelor's degree (S.T.) from Universitas Indonesia in 2005 and her master's degree (M.T.) 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 working as a researcher at the Research Center for Data and Information Sciences of National Research and Innovation Agency (BRIN). Her research interests are in 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 enterprise. She can be contacted at email: purn005@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 that are 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 email: dianadewi.riswantini@brin.go.id.

Andri Fachrur Rozie received his undergraduate majoring in Informatics Engineering and obtained master's degree from Department of Computer and Radio Communications Engineering from Korea University. His research area is in the field of data science, machine learning, natural language processing, and autonomous vehicles. Currently, he is working at 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 at email: andri.fachrur.rozie@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, completed 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 halal food industry. She can be contacted at email: ekasari.nugraheni@brin.go.id.

Devi Munandar received his undergraduate degree majoring in Informatics Engineering. Obtained master's degree from 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 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. In 2020-2022, he has been involved in statistical modeling research and stochastic processes, using principal component analysis (PCA) integrated with vector autoregressive (VAR) using a datamining approach for forecasting climate effects. He can be contacted at email: 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 University of Melbourne. Currently he is working as a researcher at the Research Center for Data and Information Sciences of 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 in data management and analytics in distributed and parallel computing environments. He can be contacted at email: 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 as a 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 email: noorroufiqa@gmail.com.