HANIF
Journal of Information Systems
https://journal.ilmubersama.com/index.php/hanif

Click here and write your Article Category

Improving the Accuracy of Lettuce and Weed Classification Based on MobileNetV2 Features Through Segmentation

Kurniawan Saputra [1], Akhmad Jayadi [1], Ahmad Rofi'i [1]

[1] *Informatics Management, Politeknik Negeri Lampung, Bandar Lampung, Indonesia*

ARTICLE INFORMATION

Received:February 00, 00
Revised:March 00, 00
Available Online:April 00, 00

KEYWORDS

Precision Farming; Lettuce Weeds; GrabCut Segmentation; MobileNetV2; Image Classification

CORRESPONDENCE

Phone: :+62 (812) 7941124
E-mail: kurniawan_mi@polinela.ac.id

A B S T R A C T

Automating the separation of commodity crops and weeds *is* a major challenge in the implementation of precision agriculture . The presence of complex backgrounds such as soil, rocks, and shadows often degrades the performance of feature extraction in computer vision classification models. This study proposes an image preprocessing approach using the GrabCut segmentation method to extract key crop objects cleanly before performing *Deep Learning- based feature extraction* . Representative features from the image are extracted using the lightweight and efficient MobileNetV2 architecture. Next, classification is performed by comparing three *Machine Learning algorithms* , namely *Support Vector Machine* (SVM), *K-Nearest Neighbors* (KNN), and *Random Forest* (RF). Testing is carried out on two data scenarios, namely the original dataset (*Original*) and the segmented dataset (*GrabCut*). The experimental results show that the use of original images produces an accuracy of 98.89% for all three classification models. However, after being integrated with GrabCut segmentation, the accuracy of all three models increases significantly to 100.00%. These results prove that GrabCut-based segmentation effectively eliminates background noise information , thereby improving the generalization capabilities of classification models perfectly on edge *computing devices* .

INTRODUCTION

The modern agricultural sector [1] currently faces increasingly complex challenges, ranging from limited productive land, climate change, to soaring operational costs due to labor shortages. Amidst these dynamics, the concept of precision agriculture [2] has emerged as one of the most promising solutions to optimize crop yields while reducing cost efficiency. One horticultural commodity that has high economic value but is very sensitive to environmental disturbances is lettuce (Lactuca sativa) [3]. In lettuce cultivation, the presence of weeds is one of the main limiting factors that can drastically reduce production quantity and quality. Weeds compete directly with the main crop for soil nutrients, water, growing space, and sunlight exposure. If not controlled early, the investment of time and fertilizer costs will be wasted because it is absorbed by the nuisance vegetation.

Traditionally, weed control has been carried out through two main methods: manual weeding and blanket spraying. However, manual weeding [4] is increasingly inefficient due to its time-consuming nature and rising labor costs. Furthermore, the widespread, unselective use of chemical herbicides has triggered serious environmental issues, such as weed resistance, soil degradation, and chemical residue contamination of lettuce leaves, which can endanger consumer health. Therefore, an automated system capable of accurately identifying weed locations is needed so that control measures—whether mechanically using agricultural robots or locally applied chemically (spot spraying)—can be carried out precisely and efficiently.

The development of computer vision [5] and machine learning [6] technology has opened up significant opportunities for automating crop and weed classification in open fields. Many previous researchers have utilized convolutional neural network (CNN) architectures to automatically extract visual features from crop images. For example, research by Smith et al. (2020) demonstrated that the use of deep learning-based CNN models can recognize various types of weeds in commercial commodity fields with a high degree of accuracy. However, large-scale CNN models such as ResNet or VGG require massive computing power, making them difficult to implement directly on edge computing devices such as drones or mini agricultural robots in the field.

To overcome these computational constraints, Transfer Learning [7] approaches using lighter architectures have become the focus of recent research. The MobileNetV2 architecture [8] has proven to be one of the best feature extractors for low-spec devices because it utilizes a memory-efficient inverted residuals and linear bottlenecks structure while maintaining high accuracy. Research by [9] applied MobileNetV2 to plant disease classification and concluded that the model is highly competent generalization ability when applied to new data. However, the biggest challenge of computer vision in real-world agricultural environments is background noise [10]. Images captured directly from the field often contain varying soil textures, rocks, dry leaf litter, and shadow interference caused by inconsistent sunlight intensity. This noisy environment often triggers overfitting in classification models. Models tend to learn visual representations of the background soil or shadows, rather than focusing on the essential morphological features of the lettuce or weeds themselves. As a result, model accuracy often declines sharply when tested under different field conditions.

To eliminate this damaging information, preprocessing, namely image segmentation, is crucial. Several previous studies have used conventional thresholding methods such as Otsu to separate vegetation from soil. However, these standard color-based methods often fail when dealing with shaded leaves or wet soil because their pixel intensity values overlap. As an alternative, the GrabCut algorithm [11] offers significantly higher robustness. GrabCut combines texture and color information using Gaussian Mixture Models (GMM) and solves object separation through the Graph Cut technique, which iteratively minimizes a pixel energy function. Although the effectiveness of GrabCut and MobileNetV2 has been extensively tested separately, in-depth investigations into the extent to which background noise removal using GrabCut can boost the performance of various traditional machine learning classifier algorithms (such as SVM, KNN, and Random Forest) in the case of lettuce-weed separation are still very limited. Most studies directly feed raw field images into the model, forcing the model to work harder to isolate important features from surrounding noise.

Therefore, this study aims to design and evaluate an integrated preprocessing pipeline using GrabCut segmentation to thoroughly clean images before entering the MobileNetV2 feature extraction stage. Through this approach, the performance of three main classification models—Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF)—will be comparatively tested on two different data scenarios, namely original field data (Original) and segmented data (GrabCut). The scientific contribution of this study is expected to provide empirical evidence that total background information filtering through segmentation not only improves classification accuracy performance to reach the optimal limit but also simplifies the feature space, thereby accelerating the convergence process of classification models in future precision agriculture applications.

METHOD

In general, the research methodology is divided into five main stages: (1) Dataset Acquisition and Preparation, (2) Image Segmentation using the GrabCut Algorithm, (3) Spatial Feature Extraction Based on MobileNetV2 Transfer Learning, (4) Feature Selection using ANOVA F-value, and (5) Classification Model Training and Evaluation Process using three Machine Learning algorithms.

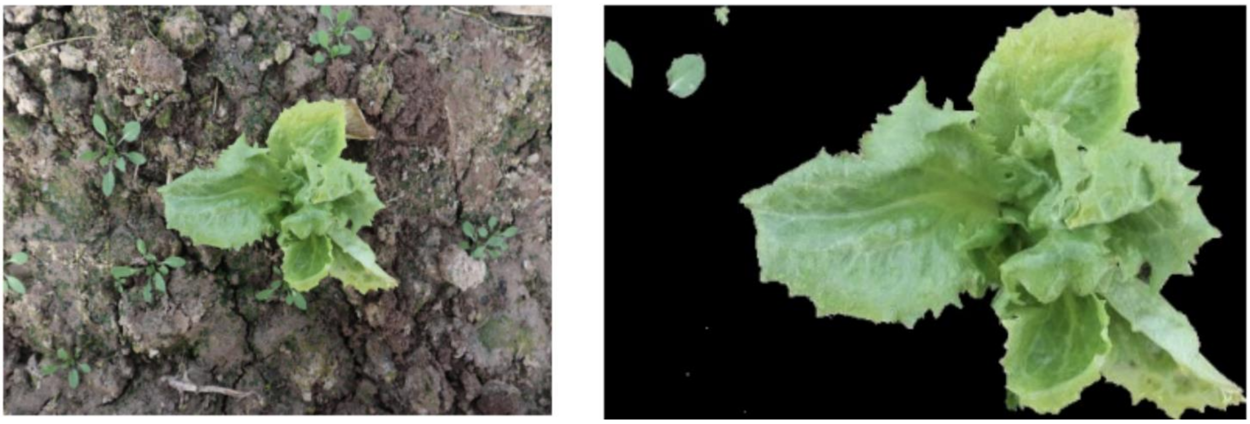
Dataset Acquisition and Preparation

The dataset used in this study focuses on two main objects in the precision agriculture domain: lettuce (Lactuca sativa) as the main commodity crop (represented as Class 0), and various types of nuisance plants that grow around it as weeds (represented as Class 1).

To empirically prove the research hypothesis, this dataset was replicated into two independent storage directory scenarios:

1. original_dataset: Contains the original raw imagery captured by the camera in the field as seen in figure 1. (a). Images in this dataset retain all environmental information, including variations in soil texture, gravel, plastic mulch, and sun shadow distortion.

2. grabcut_dataset: Contains images that have undergone preprocessing for object separation (background removal) using the GrabCut algorithm as seen in figure 1. (b). All background pixels are converted to absolute black (R=0, G=0, B=0), leaving only the morphology of vegetation leaves as the main object.



(a) (b)

Figure 1. (a) Original letuce dataset (b) Grabcut letuce dataset

Before entering the feature processing stage, these two datasets are divided into training data and validation data using a train_test_split-based randomization function to ensure a balanced class distribution (stratified split). The images are then resized to a standard dimension of 224 x 224 pixels with RGB color depth (3 channels) to match the default input dimensions of the feature reader architecture.

Image Segmentation Using GrabCut Algorithm

GrabCut works by modeling the color distribution of foreground and background objects using Gaussian Mixture Models (GMM) [11]. The algorithm iteratively minimizes the energy function to determine the optimal pixel boundaries as shown in equation (1) below.

$$E(A) = U(A) + V(A) \tag{1}$$

Where $U(A)$to evaluate the pixel match with the GMM color model, and $V(A)$measure the discontinuity between neighboring pixels to ensure the smoothness of the leaf object segmentation boundaries.

MobileNetV2 Based Spatial Feature Extraction

The prepared images (both original and segmented scenarios) are transformed into numerical feature vectors using the Transfer Learning technique utilizing the MobileNetV2 architecture. MobileNetV2 was chosen because of its highly optimized performance on edge computing thanks to the implementation of the Inverted Residuals and Depthwise Separable Convolution structures. This cascaded convolution structure breaks the standard convolution process into two stages: depthwise convolution (applying one filter per input channel) and pointwise convolution (a 1 x 1 convolution to incorporate linearity between channels). This approach can drastically reduce the overhead of matrix multiplication and addition operations without significantly degrading feature extraction accuracy.

The feature extraction procedure in this study runs as follows:

1. Loading the MobileNetV2 architecture that has been trained using giant weights from the ImageNet dataset (weights='imagenet').
2. Cut off the top classification layer (fully connected layer / include_top=False) so that the network acts purely as a spatial feature mapper (feature extractor).
3. Apply a dimensionality reduction layer in the form of Global Average Pooling (pooling='avg') to the final output feature map. This operation calculates the average value of each spatial feature map, converting the three-dimensional matrix into a fixed-length one-dimensional row vector of 1280 features for each input image.

MobileNetV2 Based Feature Extraction

To maintain computational efficiency, the MobileNetV2 architecture trained on the ImageNet dataset was used as a feature extractor. The top classification layer was discarded, and a Global Average Pooling operation (pooling='avg') was applied. This process produces a fixed-volume, one-dimensional feature vector for each incoming image that is highly representative of the morphology of lettuce and weed leaves.

Feature Selection

To improve the efficiency and speed of inference, the feature dimension was reduced from 1280 to a number of Kbest features using the univariate *SelectKBest method* combined with the ANOVA F-test (f_classif) [12] statistical test function. This statistical test evaluates the level of significance of differences in the average feature values between class groups through the ratio of between-group variability (between-group variance) to within-group variability (within-group variance) as seen in equation (2) below.

$$F = \frac{MSB}{MSW} \tag{2}$$

Where α MSB is the Mean Squares Between and α $MSW$is the Mean Squares Within. Features that are constant or have Flow values (not correlated with the target plant discriminator) are eliminated. A range of values $K$is tested iteratively from 1 to 100 to analyze the most stable point of convergence of model accuracy.

Classification with Machine Learning

The extracted feature vectors were normalized using StandardScaler [13]. Model performance evaluation was performed using three independent classifiers.

1. Support Vector Machine (SVM) [13]: Uses a linear kernel to find the best separating hyperplane.
2. K-Nearest Neighbors (KNN) [14]: Tested on neighbor values $k \in \{3,5,7,9\}$
3. Random Forest (RF) [12]: Uses a decision tree-based ensemble method evaluated from 10 to 100 trees in multiples of 10.

System Architecture

The system architecture proposed in this study is shown in Figure 2 below.

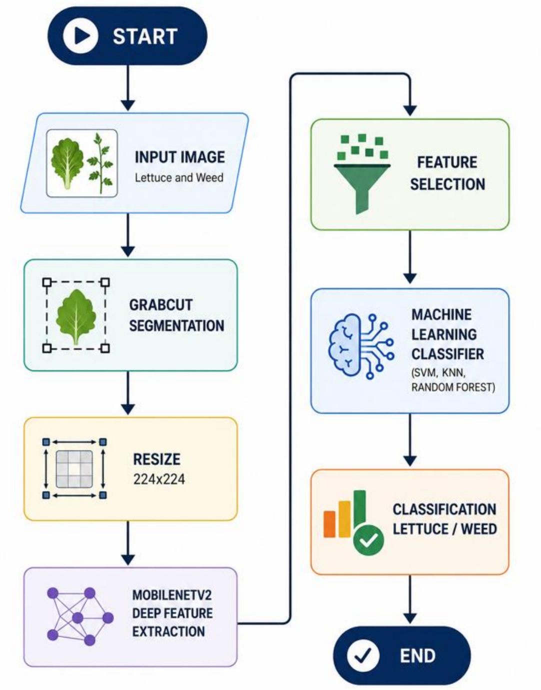


Figure 2. Architecture of the lettuce and weed classification system.

The system was designed to compare the classification performance of lettuce and weeds in two scenarios, using original images and GrabCut segmented images. The process begins with image acquisition, followed by GrabCut segmentation as an optional preprocessing step to remove irrelevant background. The images were then resized to 224 × 224 pixels and extracted using MobileNetV2 to obtain deep feature representations. The resulting features were then selected using the SelectKBest method to obtain the most relevant features. In the classification stage, the selected features were used as input for machine learning algorithms, namely Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest. The results of each model were then compared to analyze the effect of GrabCut segmentation on the classification performance of plants and weeds.

RESULTS AND DISCUSSION

Classification Model Performance on Original Dataset *Scenario*

The first phase of the experiment was conducted by training and testing a classification model using features extracted directly from the original dataset. In this scenario, the lettuce and weed images still contained natural background components, such as varying soil surface textures, small stone chips, mulch, and lighting shadow distortion. Quantitative test results showed that the three Machine Learning algorithms tested—Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Original Random Forest (RF)—produced a uniform accuracy value of 98.89% on the validation data with a total of 90 test samples. Table 1 presents a detailed classification report for the original image scenario.

Table 1. Detailed Classification Report on Original Dataset *Scenario*

Model	Target Class	Precision	Recall
Linear SVM	Class 0 (Lettuce)	0.98	1.00
	Class 1 (Weeds)	1.00	0.98
KNN (k=3,5,7,9)	Class 0 (Lettuce)	0.98	1.00
	Class 1 (Weeds)	1.00	0.98
Random Forest (all)	Class 0 (Lettuce)	0.98	1.00
	Class 1 (Weeds)	1.00	0.98

While these accuracy values are considered very high, a more in-depth analysis of the error matrix revealed a crucial classification weakness. Based on the test data, the Confusion Matrix for all three models showed similar error patterns, as summarized in Table 2.

Table 2. Confusion Matrix of Original Image Scenario

	Class 0 Prediction (Lettuce)	**Class 1 Prediction (Weed)**
Class 0 Actual (Lettuce)	**45** (Correct)	0
Actual Class 1 (Weeds)	**1** (Misclassified)	**44** (Correct)

A total of 45 samples from Class 0 (Lettuce) were successfully classified correctly (100% sensitivity level). However, in Class 1 (Weeds), there was 1 sample that was misclassified as Class 0. This misclassification indicates the occurrence of a feature overlap phenomenon in the MobileNetV2 vector dimension space. When the weed image has a dominant soil background or leaf shadows that resemble the fractal pattern of lettuce leaves, the formed feature vector will shift closer to the Class 0 cluster. As a result, the separator algorithm fails to draw an absolute boundary line, as evidenced by the Class 1 recall value which is stuck at 0.98.

Classification Model Performance on Segmented Image Scenario (GrabCut Dataset)

The second phase of the experiment applied a background elimination preprocessing flow using the GrabCut algorithm before the images were fed to the MobileNetV2 feature reader. The data preparation results in a much cleaner feature space, where non-vegetation components are isolated into homogeneous black areas. When features from the grabcut_dataset were fed, the performance of all three classification algorithms experienced a significant and uniform jump in performance, reaching 100.00% on the validation data with a total of 76 test samples. Table 3 presents the details of the classification report for the image scenarios that have gone through the GrabCut segmentation process.

Table 3. Detailed Classification Report on Segmented Image Scenario (GrabCut Dataset)

Model	Target Class	Precision	Recall

Linear SVM	Class 0 (Lettuce)	1.00	1.00
	Class 1 (Weeds)	1.00	1.00
KNN (k=3,5,7,9)	Class 0 (Lettuce)	1.00	1.00
	Class 1 (Weeds)	1.00	1.00
Random Forest (all)	Class 0 (Lettuce)	1.00	1.00
	Class 1 (Weeds)	1.00	1.00

This absolute performance improvement is confirmed by the perfect Confusion Matrix values, as shown in Table 4.

Table 4. Confusion Matrix of Segmented Image Scenario

	Class 0 Prediction (Lettuce)	**Class 1 Prediction (Weed)**
Class 0 Actual (Lettuce)	**44** (Correct)	0
Actual Class 1 (Weeds)	0	**32** (Correct)

Based on Table 4, all 44 lettuce samples and 32 weed samples in the validation data were successfully identified with absolute accuracy (zero misclassification). The Precision, Recall, and F1-score values for each class reached a maximum index of 1.00. This empirically proves that GrabCut's radical removal of background noise successfully eliminates destructive features. When ground and shadow pixels are converted to absolute black, the MobileNetV2 model is forced to extract only information derived purely from the geometry, edge structure, and internal texture of vegetation leaves. This increases the geometric distance between class clusters (inter-class distance) and reduces the variance within a single class (intra-class variance), resulting in a highly linear and easily separable feature space.

Comparison between the original image and the grabcut image as seen in the following Figure 2.

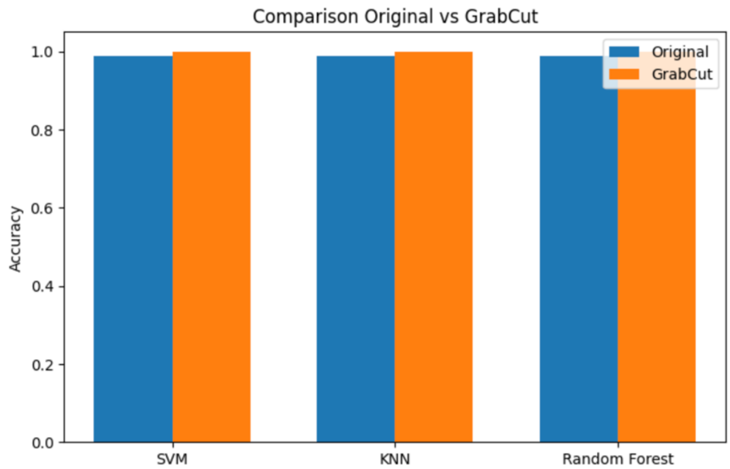


Figure 3. Comparison graph of the accuracy of SVM, KNN, and Random Forest models before and after applying GrabCut segmentation

Model Parameter Sensitivity Analysis and Final Comparison

To ensure the algorithm's stability against changes in internal parameters, sensitivity tests were performed on KNN (k-neighborhood value testing) and Random Forest (tree quantity testing). The final comparative results of the performance of all models across scenarios are summarized in Table 5.

Table 5. Summary of Final Accuracy Comparison Across Scenarios and Parameters

Classifier	**Internal Model Parameters**	***Original* Scenario Accuracy**	**Segmentation Scenario Accuracy (*GrabCut*)**
SVM	Kernel = Linear	98.89%	100.00%
KNN	k = 3	98.89%	100.00%
	k = 5	98.89%	100.00%
	k = 7	98.89%	100.00%
	k = 9	98.89%	100.00%
Random Forest	Trees = 10	98.89%	100.00%
	Trees = 50	98.89%	100.00%

	Trees = 100	98.89%	100.00%

The data in Table 5 shows a consistent pattern of stability. Internal parameter values (both variations in the number of nearest neighbors and the depth of the decision tree) do not change the accuracy values in each scenario. This characteristic proves that the combination of GrabCut preprocessing and the MobileNetV2 extractor provides very strong discriminant values from the start. The model does not require complex parameter configurations or deep tree architectures to separate objects, which directly impacts the efficient use of computing memory on IoT devices.

Based on the Feature Selection Comparison as seen in figure 4 test graph evaluated using the Linear Support Vector Machine (SVM) algorithm, a very contrasting difference in accuracy convergence patterns is seen between the original image scenario (Original) and the segmented image (GrabCut). In the segmented image scenario, the accuracy curve (orange line) shows a massive and instant performance jump, where the model immediately locks an absolute accuracy level of 100.00% (1.0) by only involving the first 1 to 2 feature dimensions resulting from the univariate ANOVA F-value selection. This characteristic empirically proves that the total elimination of background noise such as soil, rocks, and shadows through GrabCut successfully removes destructive features, leaving pure vegetation leaf morphology information with high discriminant value. In contrast, in the original image scenario without preprocessing (blue line), the accuracy curve shows significant instability fluctuations at small-scale feature numbers (below 10 features). This scenario requires a much larger data dimension, ranging from 10 to 20 feature components, before the graph finally stabilizes and is stuck at the optimal limit of 98.89%. This confirms the phenomenon of feature overlap, where the ANOVA statistical function on the original image is often fooled into selecting noisy features from the background that are considered dominant while actually weakening the decision boundary between classes. Thus, this visual finding has crucial scientific implications for practical implementation in the field of precision agriculture; GrabCut preprocessing is proven to be able to simplify the MobileNetV2 feature space extremely, so that edge computing systems on automated weeding robots can maintain excellent classification performance even though only processing a small subset of essential features, which directly implies a reduction in inference latency time and computational memory efficiency in the field.

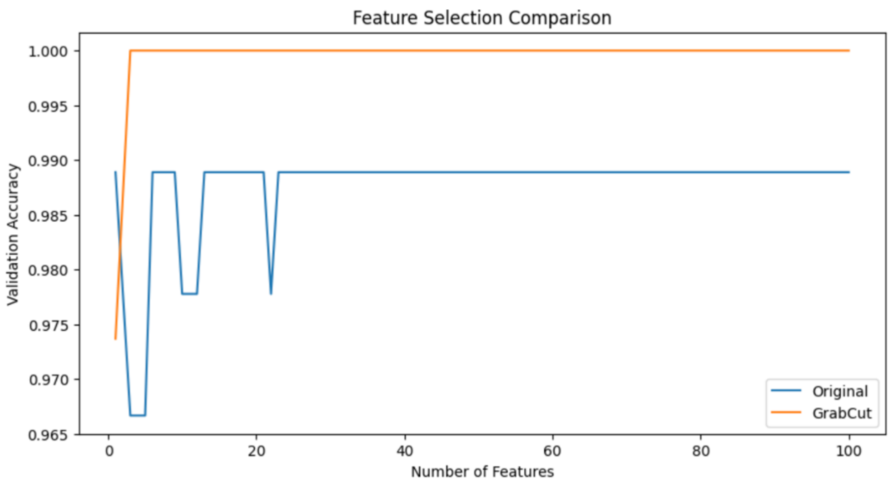


Figure 4. Graph of the effect of the number of best feature selections (K-Best Features) on the stability of model validation accuracy.

CONCLUSION

This study comprehensively proves that the integration of GrabCut segmentation as an image preprocessing stage is able to absolutely eliminate natural background noise, thus significantly improving the performance of separating lettuce (Lactuca sativa) and companion weeds. Through transfer-learning feature extraction using the MobileNetV2 architecture, testing on real field image scenarios resulted in accuracy that was stuck at 98.89% due to the phenomenon of feature overlap in the vector dimension space. However, after the noisy ground and shadow pixels were isolated into homogeneous black areas using the GrabCut algorithm, the distance between class clusters expanded linearly so that the performance of the three classification models—Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF)—increased uniformly to reach an absolute accuracy level of 100.00% on the validation data. In addition to delivering excellent accuracy, GrabCut preprocessing has been shown to stabilize the model from complex internal parameter optimization dependencies and accelerate the convergence curve of univariate ANOVA F-value feature selection. These results have significant scientific implications for edge computing architectures in precision agriculture, as the system can maintain optimal classification performance despite processing only a small subset of essential feature dimensions, which directly impacts inference latency reduction and computational memory efficiency in the field.

REFERENCES

[1] A. A. Rouf, "Pertanian Modern: Solusi untuk Tantangan Produksi, Efisiensi, dan Regenerasi Petani," *Policy Br. Pertanian, Kelautan, dan Biosains Trop.*, vol. 8, no. 1, pp. 1539–1544, Mar. 2026, doi: 10.29244/agro-maritim.0801.1539-1544.

[2] M. Kushwaha, S. Singh, V. Singh, and S. Dwivedi, "Precision Farming: A Review of Methods, Technologies, and Future Prospects," *Int. J. Environ. Agric. Biotechnol.*, vol. 9, no. 2, pp. 242–253, 2024, doi: 10.22161/ijeab.92.27.

[3] N. S. Khodijah and R. Kusmiadi, "The Growth Of Lettuce (Lactuca sativa) Hydroponically In Simple Wick System On Various Types Of Nutrient Composition," *J. Agron. Tanam. Trop.*, vol. 3, no. 2, pp. 180–186, Jul. 2021, doi: 10.36378/juatika.v3i2.1366.

[4] S. A. N'cho, M. Mourits, J. Rodenburg, and A. Oude Lansink, "Inefficiency of manual weeding in rainfed rice systems affected by parasitic weeds," *Agric. Econ.*, vol. 50, no. 2, pp. 151–163, Mar. 2019, doi: 10.1111/agec.12473.

[5] D. R. Tobergte and S. Curtis, *Algorithms for image prcessing and computer vision*, vol. 53, no. 9. 2013. doi: 10.1017/CBO9781107415324.004.

[6] A. N Patil, "IMAGE RECOGNITION USING MACHINE LEARNING," *Int. J. Eng. Appl. Sci. Technol.*, vol. 6, no. 1, May 2021, doi: 10.33564/IJEAST.2021.v06i01.027.

[7] F. D. Adhinata, N. G. Ramadhan, and A. Jayadi, "DenseNet201 Model for Robust Detection on Incorrect Use of Mask," in *Proceeding of the 3rd International Conference on Electronics, Biomedical Engineering, and Health Informatics*, T. Triwiyanto, A. Rizal, and W. Caesarendra, Eds., Singapore: Springer Nature Singapore, 2023, pp. 251–263.

[8] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Mar. 2019, [Online]. Available: http://arxiv.org/abs/1801.04381

[9] R. Jajila, N. A. Ramdhan, P. Wahyuningsih, and B. Irawan, "Optimasi Mobilenetv2 Dengan Transfer Learning Untuk Klasifikasi Penyakit Daun Cabai," *Infotek J. Inform. dan Teknol.*, vol. 9, no. 1, pp. 296–307, Jan. 2026, doi: 10.29408/jit.v9i1.33812.

[10] M. M. Weli and O. M. Abdullah, "Digital Image Noise Reduction Based on Proposed Smoothing and Sharpening Filters," *Indones. J. Comput. Sci.*, vol. 13, no. 4, Jul. 2024, doi: 10.33022/ijcs.v13i4.4151.

[11] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut' - Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004, doi: 10.1145/1015706.1015720.

[12] M. F. Thoriq, W. J. Pranoto, and F. Faldi, "Penerapan Seleksi Fitur Analysis of Variance Pada Algoritma Random Forest Classifier Dalam Klasifikasi Nilai Mahasiswa," *Explor. J. Sist. Inf. dan Telemat.*, vol. 14, no. 2, p. 185, Dec. 2023, doi: 10.36448/jsit.v14i2.3187.

[13] N. M. Shivsharan and M. Tari, "Machine Learning Empowered: Support Vector Machine-Based Selection of Encryption Techniques for Digital Image Security Levels," *Informatica*, vol. 49, no. 33, Aug. 2025, doi: 10.31449/inf.v49i33.6355.

[14] A. Amelia, M. Asfi, and R. Fahrudin, "Implementation of K-Nearest Neighbor Method for Selection of New Employee Candidates (Case Study: CV. Syntax Corporation Indonesia)," *Eduvest - J. Univers. Stud.*, vol. 4, no. 7, pp. 5742–5754, Jul. 2024, doi: 10.59188/eduvest.v4i7.1305.