Evaluating Deep Learning Architectures for Potato Pest Identification: A Comparative Study of NasNetMobile, DenseNet, and Inception Models

Sri Hadianti[1,*], Dwiza Riana[2], Daning Nur Sulistyowati[3]

[1,2,3]Faculty of Technology Information, Universitas Nusa Mandiri, Jalan Raya Jatiwaringin No 2, Cipinag Melayu, Makasar, Jakarta, Indonesia

(Received: July 8, 2024; Revised: September 1, 2024; Accepted: October 17, 2024; Available online: December 30, 2024)

Abstract

Manual potato pest identification that is still applied today is often time-consuming and highly dependent on farmer skills in the field. This causes delays in taking action and inaccurate reporting, especially in pest emergencies. In addition, these limitations slow down the response to pest control which ultimately risks reducing crop yields and farmer income. This study aims to develop a more accurate, fast, and consistent deep learning-based approach to identify potato pests, in order to support practical solutions that farmers can implement independently. This study contributes by comparing three deep learning architecture models, namely NasNetMobile, DenseNet, and Inception which are designed to identify pest images. The potato pest image dataset used was collected from various sources equipped with an augmentation process to increase data diversity. The model was drilled using transfer learning techniques to utilize previously learned features on a large dataset. The evaluation model was carried out comprehensively based on accuracy, precision, and inference time efficiency. The results showed that the DenseNet model achieved the highest accuracy of 97% with an inference time of 11 seconds, and this model maintained a relatively stable performance and was superior several times compared to other models. Based on these results, DenseNet was chosen as the most effective and reliable model to be developed for practical applications in the field. This study provides practical implications in the form of providing a model that can be integrated into a mobile-based application that is easy to use by farmers, including in remote areas. This allows farmers to identify pests independently without requiring in-depth technical expertise. In addition, this study is a new benchmark for the development of artificial intelligence-based pest identification systems in other crops and opens up opportunities for integration with IoT-based technologies to support sustainable agricultural practices.

Keywords: Comparative, potato pests, NasNetMobile, DenseNet, Inception

1. Introduction

Early identification of potato pests is a challenge in increasing the effectiveness of overall pest control. Manual identification methods that are currently widely used often take a long time [1]. This is because the identification process is highly dependent on the skills and expertise of officers in the field. This dependence not only makes it difficult in emergency situations that require a quick response but also increases the risk of errors in pest identification [2]. As a result, the effectiveness of pest control can decrease, which has an impact on decreased production results and potential economic losses for farmers [3]. With these limitations, a new, more efficient and accurate approach is urgently needed to support a sustainable agricultural system and increase food security [4].

Artificial intelligence technology especially deep learning shows great potential as an innovative solution in potato pest identification [5]. Through the deep learning approach, the identification process can be carried out quickly and accurately and reduces dependence on time-consuming manual methods [6]. The advantage of deep learning in recognizing complex patterns from pest images allows for more consistent and reliable detection [7]. In addition this deep learning model can be integrated into applications that are easily accessible to farmers and field extension workers, so that they can independently identify pests, even in remote locations with results approaching expert accuracy [8].

Various deep learning architectures have been developed and successfully achieved high accuracy in image classification tasks including in plant pest identification [9]. Among these architectures, the NasNetMobile [10],

*Corresponding author: Sri Hadianti (sri.shv@nusamandiri.ac.id)
DOI: https://doi.org/10.47738/jads.v6i1.545

ISSN 2723-6471

DenseNet [11], [12], and Inception [13], [14], models have been known to be able to handle image complexity with a high level of accuracy and relatively fast inference time. However, there has been no comprehensive study comparing the performance of these three architectures specifically for potato plant pest identification. Comparative analysis is needed to understand which model is more optimal, both in terms of accuracy and computational efficiency so that it can be used practically in the field.

This study aims to evaluate and compare the performance of NasNetMobile, DenseNet, and Inception models in potato plant pest identification. The results of this study are expected to provide recommendations for the most appropriate deep learning model for pest identification applications that can help farmers in pest management effectively and efficiently. This study is also expected to be a reference for further development in the application of deep learning technology in the agricultural sector.

The main contributions of this study are: Comparative evaluation of NasNetMobile, DenseNet, and Inception models for potato pest identification provides a new benchmark in agricultural pest detection, demonstrating higher accuracy and efficiency than existing methods on the potato pest dataset. Performance issues caused by high variations in pest images, such as varying backgrounds and lighting conditions, which often affect model accuracy, are successfully addressed by using customized image preprocessing techniques for each model. Superior results are achieved with NasNetMobile in terms of computational efficiency, while DenseNet shows excellent feature extraction capabilities. These findings are significant when compared to previous studies of each model for agricultural pest identification.

2. Literature Review

Table 1 is a related study that is relevant to this study. The studies in table 1 have used deep learning techniques to identify plant pests. The level of accuracy validation in these studies varies, some are perfect at 100%, but some are still low at 73%. Research [15] used the deep learning method to identify pests in tomato plants. The results showed that the DenseNet169 model was the model with the highest accuracy value. Research [16] used deep learning techniques to recognize ten types of pests found in rice plants. The deep learning approach was applied to the pest dataset by readjusting the model layers and hyperparameters. The results showed that the adjusted ResNet-50 model produced a better accuracy of 95.012% compared to other trained models. The results obtained showed the effective performance of the model in identifying rice plant pests. Research [17] used a traditional feature-based approach and a deep learning feature-based approach in extracting chili plant pest features, then the features were trained using three machine learning classifiers. The results of the study showed that the InceptionV3 + DenseNet201 model performed better than other models. Research [18] developed an improved YOLOv4 model in identifying tomato plant pests. The results showed that the proposed model can effectively increase the accuracy value in identifying tomato plant pests compared to other models. Research [19] proposed a new model called DeepPestNet in identifying plant pests. The proposed model consists of 11 learning layers containing eight convolutional layers and three fully connected (FC) layers. The proposed model is able to identify 10 types of pests effectively. Research [20] proposed a new model called PestDetector which can identify 4 types of jute plant pests with high accuracy. The accuracy of the proposed model is 99.18%, indicating that the model can identify very well. Research [21] used the VGG16, ResNet50, and Xception models to identify cassava plant pests. The Xception model is the best proposed model with an accuracy value of 94.56%. Research [22] proposed a new model named JutePestDetect in identifying jute plant pests. The models used to develop the model were DenseNet201, InceptionV3, MobileNetV2, VGG19, and ResNet50. These models were modified by adding a dropout layer for regularization and changing the classification layer with a global average pooling layer. The JutePestDetect model, which is a development of DenseNet201, showed better performance than other models. So that the proposed model can identify jute plant pests well. Research [23] proposed a DCNN model in identifying jute plant pests. The proposed model showed an accuracy value of 95.86% in four classes of jute plant pests. Research [10] proposed a MobileNet SSD model for identifying corn plant pests. The results showed that the proposed MobileNet SSD model had good accuracy values with limited datasets. These results validate the potential use of neural networks in real-time corn plant pest identification. Research [24] proposed the YOLOv5 model that adjusts the variation of patience as an initial stopping parameter of 100, 200, and 300. The results of the study showed that the proposed model can identify pests with small and large sizes well. From these studies, it can be seen that the proposed deep learning model is good at identifying plant pests.

ISSN 2723-6471

However, research on potato pest identification is still limited, including research [25] using five deep learning models CMobileNetV2, CNASLargeNet, CXception, CDenseNet201, and CInceptionV3, to develop a model named PotatoPestNet to identify potato pests. Among the trained models, the CTInceptionV3 model achieved the highest accuracy of 91%, indicating that the proposed model is able to identify potato pests accurately. Our previous study [25] used the VGG16 model to identify potato pests. This study showed that the VGG16 CNN can identify pest types with an accuracy rate of 73%. Research [26] used the MobileNet model to identify corn pests. This study shows that MobileNet can identify pest types with an accuracy of 90%. From previous studies that have existed, this study will compare the NasNetMobile, DenseNet and Inception models that are already good at identifying plant pests, and will be trained on potato pests by adding augmentation techniques.

Table 1. Literature Review

Researchers	Domain	Method	Accuracy (%)
Pattnaik G, dkk [19]	Identification of Tomato Pests	DenseNet16	88.83
Malathi V, dkk [20]	Identification of Paddy Pests	ResNet 50	90.01
Ahmad Loti NN [21]	Identification of Chili Pests	InceptionV3+ DenseNet201	92.00
Liu J, dkk [22]	Identification of Tomato Pests	YOLOv4	93.40
Ullah N, dkk [23]	Identification of Plant Pests	DeepPestNet	100.00
Karim DZ, dkk [24]	Identification of Jute Pests	CNN	99.18
Gómez-Pupo SM, dkk [25]	Identification of Casava Pests	Xception	94.55
Talukder MS, dkk [26]	Identification of Jute Pests	Deep learning	99.00
Sourav MS, dkk [27]	Identification of Jute Pests	Deep CNN	93.00
Agustian I, dkk [28]	Identification of Red Chili Pepper Pest	YOLOv5	90.00
Talukder MS, dkk [29]	Identification of Potato Pests	CTInceptionV3-RS	93.00
Hadianti S, dkk [30]	Identification of Potato Pests	VGG16	73.00
Maican E, dkk [31]	Identification of Corn Pest	MobileNet-SSD	90.00

3. Methodology

This research was conducted by following the stages that have been planned systematically, which include several important steps starting from planning, data collection, analysis, to compiling research results. These stages are explained in detail in figure 1 which describes the overall research flow starting from data collection, data processing, to evaluation and conclusions obtained.

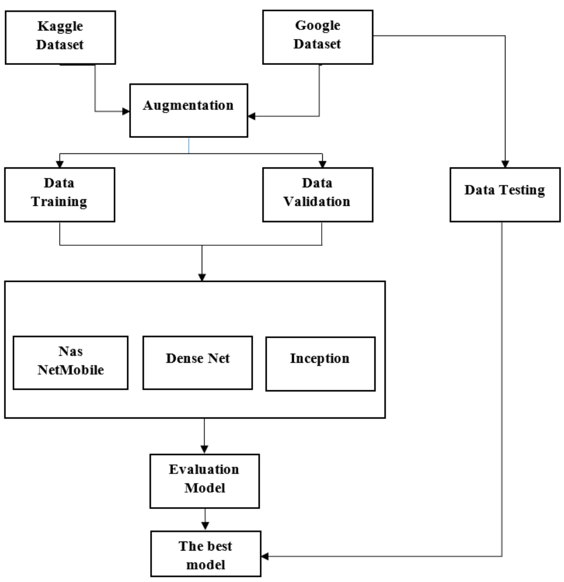


Figure 1. Research Flow

The first stage is collecting potato plant pest image data. This study uses potato plant pest images consisting of 8 types of pests consisting of Phthorimaea operculella (Potato Borer Moth), Amrasca devastans (Cotton Planthopper), Aphis gossypii Glover (Aphids), Agrotis ipsilon (Black Armyworm), Brachytrypes portentosus Lichtenstein (Necklace Cricket), Bemisia tabaci (Whitefly), Epilachna vigintioctopunctata (28-Spot Beetle), and Myzus persicae (Green Aphids) [25]. The dataset used in this study consists of 1,040 images, with 459 images sourced from Kaggle and an additional 581 images collected from the Google search engine. To address the limited dataset size, we employed data augmentation techniques such as rescale, shear, zoom, horizontal flip, rotation, width shift, height shift, and brightness, which helped increase the effective number of training images. We also acknowledge that a larger dataset would contribute to better generalizability and plan to explore additional data collection efforts in the future. The dataset contains images from several regions across bangladesh, but seasonal and other environmental variations were not specifically tracked in the current dataset. We recognize that this could influence the appearance of pests, and we suggest that future work may consider incorporating data across different seasons and regions for a more comprehensive analysis. Data descriptions can be seen in table 2.

Table 2. Potato Pests Dataset

No	Class name	Kaggle Dataset	Additional Datasets	Sub Total
1	Amrasca devastans (Cotton Planthopper)	62	68	130
2	Agrotis ipsilon (Black gobbler)	103	27	130
3	Aphis gossypii Glover (Aphids)	37	93	130
4	Bemisia tabaci (Whitefly)	35	95	130
5	Brachytrypes portentosus Lichtenstein (Necklace Cricket)	35	95	130
6	Epilachna vigintioctopunctata (28 Spot Beetle)	70	60	130
7	Phthorimaea operculella (Potato Borer Moth)	42	88	130
8	Myzus persicae (Green Aphid)	75	55	130
	Total	**459**	**581**	**1040**

After obtaining the dataset collection, the next step is to augment the data to address the limitation of its small size. The augmentation techniques aim to improve the model's robustness, particularly in handling image disturbances or variations in lighting. These techniques include:

The images are rescaled using the formula shown in the formula (1).

$$rescale = x\ \frac{1}{255} \tag{1}$$

where x represents the original pixel value. This step normalizes the pixel values to a range between 0 and 1, ensuring consistency in the input data and facilitating better convergence during training. Next the Shear Range is carried out where this technique applies a shear transformation and images are tilted or rotated within a maximum shear limit. It improves the model's ability to handle distortions and varying perspectives in real-world scenarios. Next Zoom Range is performed by randomly enlarging or reducing the image size by a factor between 0.8 and 1.2. This technique helps the model adapt to variations in object scale, making it more robust to objects that appear closer or further away in real images. Next, Horizontal Flip is performed where random horizontal flips along the vertical axis ensure the model learns to recognize objects regardless of their orientation, such as left-to-right or right-to-left variations. Then, Rotation Range is performed where images are randomly rotated during training. This technique allows the model to generalize well to images with varying orientations, reducing sensitivity to specific rotation alignments. Next, Width Shift Range is performed which introduces variations in the horizontal position of objects in the image, allowing the model to recognize objects even when their horizontal position changes slightly. Then, Height Shift Range which is similar to width shift, this technique introduces variations in vertical position, helping the model recognize objects with changes in their vertical position. Then, Brightness Range is performed by varying the brightness level, this addition increases the diversity of the dataset related to lighting conditions, making the model robust to variations in lighting images, such as shadows or overexposure. The augmentation techniques employed, such as rescaling and varying the brightness, are particularly designed to address issues like lighting distortions and noise in real-world images. For example, adjusting the brightness range ensures the model is more robust to variations in illumination, including shadows and overexposure. Similarly, random zoom and rotation allow the model to handle object distortions caused by perspective changes or camera noise.

After the dataset is augmented, it is then divided into three, namely training data, validation data and testing data with a division of 80%:10%:10% [27]. After the data is divided, the training data and validation data are then used to train the compared models. The compared models consist of the NasNetMobile, Inception, and DenseNet models. The architecture of the proposed model can be seen in figure 2.

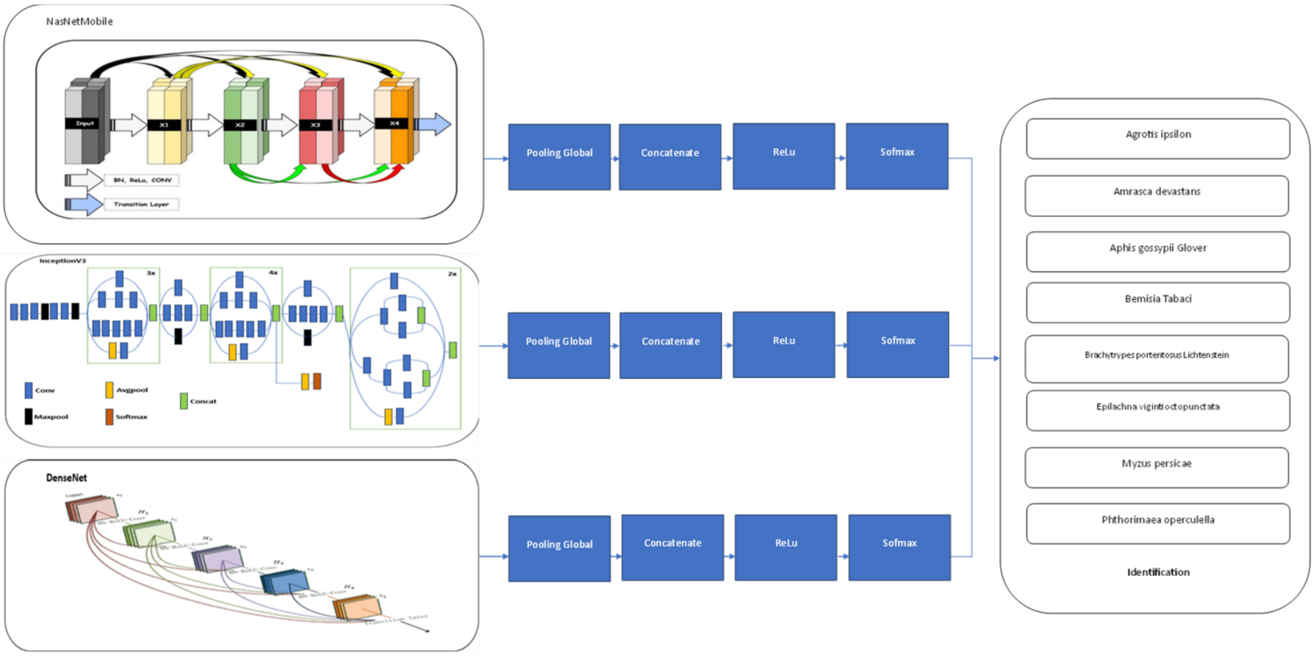


Figure 2. Model Architecture

After being trained the three compared models, evaluation was conducted using a confusion matrix with the calculation of accuracy, recall, F1-Score, and precision. In addition, to ensure the certainty and generalization of the model, k-fold cross-validation with K=5 was performed. K=5 was performed during the training process [28]. This method divides the dataset into five subsets (folds), with each fold used as the validation set while the remaining four serve as the training set. This process is repeated five times, and the evaluation metrics are averaged across all folds to obtain stable and unbiased performance metrics. Cross-validation ensures minimizing the risk of overfitting and the model is robust to unseen data. In this study True Positives (TP) refers to the number of positive examples that are predicted correctly. True Negatives (TN) is the number of negative examples that are predicted correctly. False Positives (FP) is the number of negative examples that are incorrectly predicted as positive. False Negatives (FN) is the number of positive examples that are incorrectly predicted as negative. The selection of accuracy, recall, precision, and F1-score as evaluation metrics is motivated by the critical requirements of agricultural pest identification systems. Accuracy provides a general measure of the overall correctness of the model, which is essential to ensure the reliability of the system. Recall is particularly relevant in this context because it reflects the model's ability to identify all instances of pest presence, minimizing the risk of missing pests that could cause significant damage. Precision is equally important, as it quantifies the model's capability to avoid false positives, ensuring that identified pests are indeed present and avoiding unnecessary interventions. The F1-score is included as a harmonic mean of precision and recall, providing a balanced measure that is particularly useful in situations where the dataset might have imbalanced classes, as some pests may appear less frequently than others. Together, these metrics provide a comprehensive evaluation of the model's effectiveness in addressing the challenges of agricultural pest identification.

ISSN 2723-6471

Accuracy provides a general idea of how well a classification model predicts the correct labels overall. The accuracy formula is given in the equation (2).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \text{x}10 \quad (2)$$

Recall aims to assess the extent to which the model is able to find all correct positive examples. The recall formula can be seen in the equation (3).

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \text{x}100 \quad (3)$$

F1-score aims to provide an overview of the performance of a classification model by balancing precision and recall. The F1-score formula can be seen in the equation (4)

$$\text{F1} - \text{score} = 2\text{x}\frac{\text{Precision} + \text{Recall}}{\text{Precision} + \text{Recall}} \text{x}100 \quad (4)$$

The precision formula can be seen in formula (5).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \text{x}100 \quad (5)$$

Furthermore, testing data is used to test the best model with new data so that it is known whether the best model can really recognize objects in the image identification well or not.

4. Results and Discussion

4.1. Results

Implementation and evaluation of the comparison of three models using Google Collaboratory with GPU support. This study uses the Python programming language with libraries consisting of pandas, numpy, matplotlib, keras, and tensorflow. These libraries are used in the training and testing process of deep learning models.

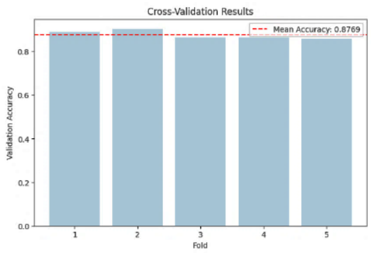
This study aims to determine which model is most appropriate for identifying pests in potato plants effectively. To achieve the desired goal, training was carried out on three deep learning models, namely NasNetMobile, Inception, and DenseNet. The three models were trained using the ADAM optimizer with a learning rate of 0.00001 and a dropout rate of 0.4. The ADAM Optimizer was chosen because of its adaptive ability to adjust the learning rate for each parameter based on the first and second gradients. This makes the training process more stable and faster than other optimizers such as SGD. A Learning Rate of 0.00001 is used to avoid too large a parameter update step, which can cause the model to not converge or pass the optimal solution. This learning rate is suitable for models with high complexity, such as NasNetMobile, Inception, and DenseNet. A Dropout Rate of 0.4 was chosen to reduce the risk of overfitting. This level is moderate enough to allow the model to learn a strong representation without losing important information from the training data. These hyperparameters are applied uniformly to all three models to maintain consistency in the initial evaluation. However, additional optimization processes are performed through grid search or random search experiments to determine whether a combination of learning rate, dropout rate, or other parameters (such as batch size) is more suitable for each specific architecture. This optimization aims to ensure the best performance of each model, considering that each architecture has different sensitivities to hyperparameter configurations. With this approach, training can improve generalization and prevent overfitting, as well as ensure that the latest weights produced are stable and effective for the fine-tuning process.

To ensure the reliability and generalization of the models, 5-fold cross-validation was applied during the training process. Each model was trained and evaluated across 5 different subsets of the dataset, allowing for a more robust estimation of their performance. Table 3 shows the evaluation results of each fold for the three models: NasNetMobile, Inception, and DenseNet. Each model was tested on five different subsets of the dataset, with varying accuracy results across folds. From table 3, it can be seen that NasNetMobile shows greater accuracy in terms of. This model starts with an accuracy of 88% on Fold 1 but tends to decrease in subsequent folds, with lower accuracies (86% on Folds 3, 4, and 5). This shows that although this model is relatively stable at the start, it is less able to handle variations in the data in subsequent folds. Although Inception has slight fluctuations between 87% and 91%, it shows more stable performance than NasNetMobile. This model is able to maintain fairly consistent accuracy across folds, although the highest accuracy is achieved on Fold 4 (91%). DenseNet shows fairly good performance overall, with accuracies ranging from 88% to 92%. Despite a slight decrease in accuracy on Fold 5 (88%), the model maintains relatively stable performance and outperforms on several folds, especially on Fold 3 and Fold 2, with the highest accuracies of 92% and 91%, respectively.

Table 3. Comparison Cross Validation

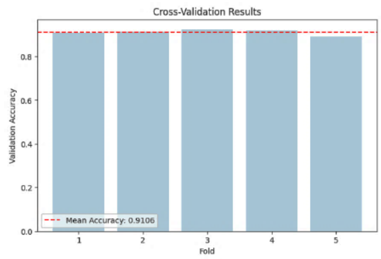
Model	Fold 1 Accuracy (%)	Fold 2 Accuracy (%)	Fold 3 Accuracy (%)	Fold 4 Accuracy (%)	Fold 5 Accuracy (%)
NasNetMobile	88	90	86	86	86
Inception	87	90	87	91	87
DenseNet	90	91	92	91	88

The graph of the cross-validation comparison between models can be seen in figure 3.



(a)

(b)

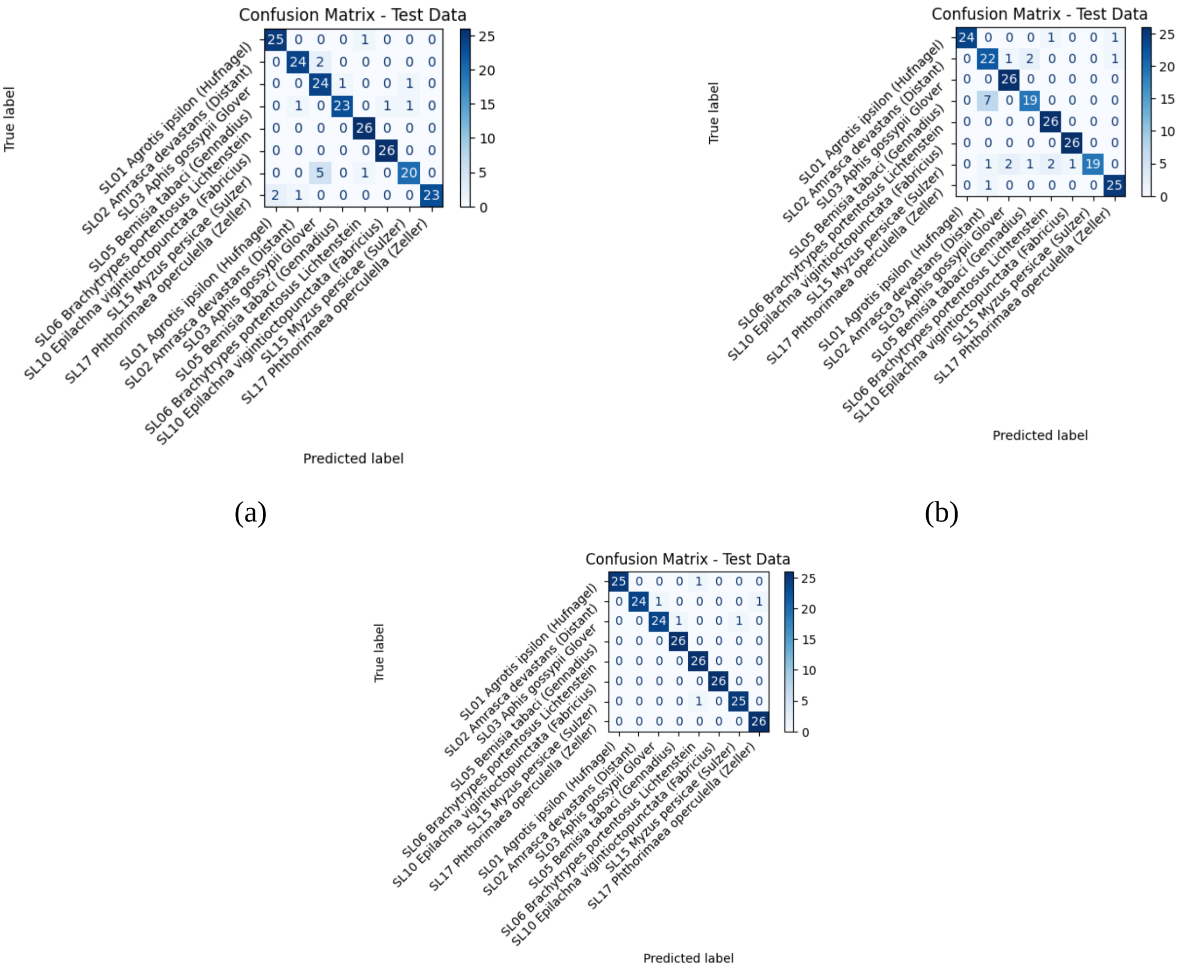


(c)

Figure 3. Graph of Cross Validation: (a) NasNetMobile, (b) Inception, (c) DenseNet

The evaluation metrics include accuracy, recall, F1-score, precision, and inference time as shown in figure 4. The matrix provides an overview of the correct classification and misclassification of the model. The results of the model evaluation can be seen in table 4, which shows a comparison of the performance of the three models tested. The model with the lowest accuracy is Inception, with accuracy, recall, F1-score, and precision values of 89% with inference time 6 second. Furthermore, the NasNetMobile model shows an increase with an accuracy, recall, F1-score, and precision value of 92% with inference time 5 second. Meanwhile the DenseNet model provides the best performance with an accuracy, recall, F1-Score and precision of 97%, with inference time 11 second which shows its superiority in predicting classes consistently. The differences in the performance of the models can be attributed to several factors. Inception, although a powerful architecture, struggles with consistently identifying certain features within the dataset, resulting in the lowest performance metrics. Its 6 second inference time, while relatively fast, indicates that the model might be over-complicating its predictions, leading to lower accuracy and precision. On the other hand, the NasNetMobile model benefits from its optimized architecture for mobile devices, making it more efficient with a 5 second inference time. Despite the efficiency, it does show a slight increase in performance metrics (accuracy, recall, precision, F1-score) compared to Inception, as it strikes a better balance between complexity and speed. This makes NasNetMobile a good choice for real-time applications where speed is critical.

ISSN 2723-6471



(a) (b)

(c)

Figure 4. Confusion Matrix: (a) NasNetMobile, (b) Inception, (c)DenseNet

However, the DenseNet model stands out as the best performer, with significantly higher accuracy, recall, F1-score, and precision (see table 4). Its architecture, which utilizes densely connected layers, enables the model to capture more intricate patterns within the data, leading to more accurate and consistent predictions. The trade-off is a longer inference time (11 seconds), which reflects the increased computational effort required for its deeper connections and more detailed analysis of features. Despite the longer inference time, the DenseNet model's superior performance in all metrics demonstrates its robustness and suitability for tasks where high accuracy is paramount.

Table 4. Comparison Model

Models	Accuracy (%)	Recall (%)	F1-Score (%)	Precision (%)	Inference Time (second)
NasNetMobile	91	91	91	92	5
Inception	89	89	89	90	6
DenseNet	97	97	97	97	11

To facilitate analysis, a comprehensive comparison of the three models is also presented in visual form in figure 5 so that the performance differences between the models can be easily identified and understood.

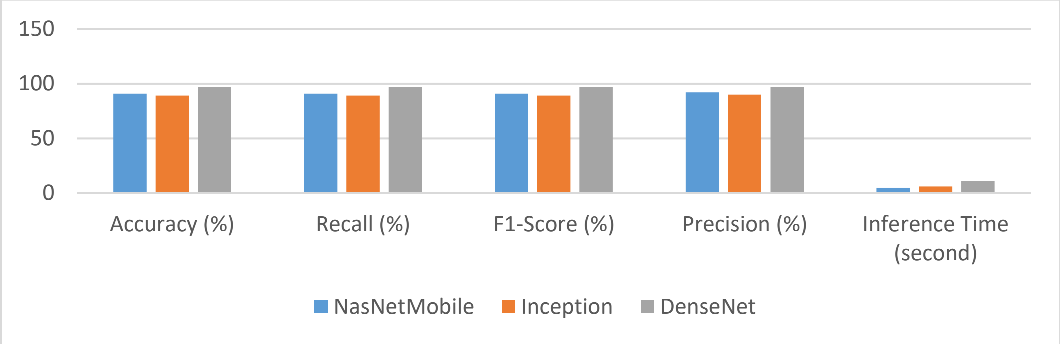


Figure 5. Comparison Among of Accuracy, Recall, F1 Score, Precision, Inference Time

Figure 6 shows an example of the results of pest image identification on potato plants carried out using the DenseNet model. The image shows that the model is able to make accurate predictions of various types of pests in the input image. The model successfully identified the types of pests on potato plants with fairly high accuracy and is more reliable than other models. This shows a good ability to distinguish the visual characteristics of each pest. These results show that the DenseNet model is not only able to identify pest types accurately but also has resistance to various variations in input images. These accurate prediction results are important to support early detection systems in maintaining the health of potato plants automatically.

Figure 6. Image Identification

4.2. Discussion

The results of the analysis carried out from the three transfer learning models used can be seen that the accuracy value of the Inception model has a low accuracy value. The Inception model showed relatively low accuracy in our study compared to DenseNet and NasNetMobile, with results similar to those observed in [29] where an Inception model achieved an accuracy of 89%. One potential reason for this underperformance could be differences in learning rate selection and data augmentation techniques used. The optimal learning rate for Inception might differ from that of DenseNet or NasNetMobile, leading to slower convergence or suboptimal training. Furthermore, the data augmentation strategy employed in this study might have been more beneficial for the other models, resulting in better generalization and performance. This can occur due to differences in the selection of learning rates and augmentation techniques. For the accuracy value of the NasNetMobile model, it has an accuracy value of 91%. These results are lower than those of the study [10] which has an accuracy of 95%. This may be due to different modifications of the NasNetMobile model used. The highest accuracy value obtained in this study was by using the DenseNet model which has an accuracy value of 97%. The accuracy results are greater than those of the study [30] which has an accuracy value of 92%. However, these accuracy results from all three models were obtained through 5-fold cross-validation, ensuring that the performance metrics provided are more reliable and stable. Cross-validation helped mitigate the potential biases caused by random splitting of the dataset, providing a more generalizable evaluation of model performance. The cross-validation process also highlighted how each model responded to different subsets of data, ensuring that the reported accuracy is not due to overfitting on a particular set. The DenseNet model showed the highest accuracy across folds, but it did come at the cost of longer inference time.

Interpretability of deep learning models remains a critical challenge. Although the models used in this study do not provide built-in interpretability tools techniques like gradient-weighted class activation mapping (Grad-CAM) can be utilized to generate heatmaps that highlight the areas of the image that the model focused on when making a prediction

[31]. Future work could explore the integration of such tools to increase model transparency and aid users in understanding why certain pests were identified.

Of the three models tested, DenseNet can be said to be the more effective and reliable model in identifying the potato pest dataset that has been trained. However, to further improve performance, future research will explore the potential use of noise reduction and image sharpening techniques to improve image quality, especially in the context of pest identification in agricultural imagery.

5. Conclusion

This study compares three deep learning architectures using NasNetMobile, DenseNet, and Inception models in identifying potato pests. The model evaluation results show that the DenseNet model has an effective and reliable model with an accuracy and precision value of 97% with an inference time of 11 seconds and this model maintains relatively stable performance and is several times superior to other models, indicating the excellent ability of this model in establishing relationships with various characteristics of potato pests. The deep learning approach in this study offers a practical solution to overcome the challenges in recognizing potato pests that have so far relied on manual methods. With deep learning, farmers can perform independent detection that approaches expert performance levels even in remote locations. However, in addition to the model, several policy challenges must be considered, such as ensuring stable internet connectivity and integrating the system with mobile devices that may not always be accessible or functioning properly in rural areas. Thus, accelerating the response to pest control and potentially increasing yields and income. The results of this study provide an important contribution to the development of artificial intelligence-based pest identification systems and support more efficient and sustainable agriculture. In addition, this study also opens up opportunities for further research in the application of the best deep learning models to low-resolution pest images.

6. Declarations

6.1. Author Contributions

Conceptualization: S.H., D.R., and D.N.S.; Methodology: S.H.; Software: D.R.; Validation: D.R., S.H., and D.N.S.; Formal Analysis: D.R., S.H., and D.N.S.; Investigation: D.R.; Resources: S.H.; Data Curation: S.H.; Writing—Original Draft Preparation: D.R., S.H., and D.N.S.; Writing—Review and Editing: S.H., D.R., and D.N.S.; Visualization: D.R. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The author would like to thank the Directorate General of Higher Education, Ministry of Education, Culture, Research and Technology for supporting this research through the National Competitive Applied Research Grant "MaTangDetect: Development of a Deep Learning Model to Support Early Identification of Potato Plant Pest Varieties Based on Artificial Intelligence ", 2024.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M Shoaib, B Shah, S Ei-Sappagh, A Ali, A Ullah, F Alenezi, T Gechev, T Hussain, F Ali., "An advanced deep learning models-based plant disease detection: A review of recent research," *Front. Plant Sci*., vol. 14, no. March, pp. 1–22, 2023, doi: 10.3389/fpls.2023.1158933.

[2] A. Sing, N Shraogi, R Verma, J Saji, AK Kar, S Tehlan, D Ghosh, S Patnaik., "Challenges in current pest management practices: Navigating problems and a way forward by integrating controlled release system approach," Chem. Eng. J., vol. 498, no. 15, pp. 154989, 2024

[3] T. C. Durham and T. Mizik, "Comparative economics of conventional, organic, and alternative agricultural production systems," *Economies*, vol. 9, no. 2, pp. 1–22, 2021, doi: 10.3390/economies9020064.

[4] R. Çakmakçı, M. A. Salık, and S. Çakmakçı, "Assessment and Principles of Environmentally Sustainable Food and Agriculture Systems," *Agric.*, vol. 13, no. 5, pp. 1–27, 2023, doi: 10.3390/agriculture13051073.

[5] Olabimpe Banke Akintuyi, "Adaptive AI in precision agriculture: A review: Investigating the use of self-learning algorithms in optimizing farm operations based on real-time data," *Open Access Res. J. Multidiscip. Stud.*, vol. 7, no. 2, pp. 016–030, 2024, doi: 10.53022/oarjms.2024.7.2.0023.

[6] A. Sharifi, H. Mahdipour, E. Moradi, and A. Tariq, "Agricultural Field Extraction with Deep Learning Algorithm and Satellite Imagery," *J. Indian Soc. Remote Sens.*, vol. 50, no. 2, pp. 417–423, 2022, doi: 10.1007/s12524-021-01475-7.

[7] M. Chithambarathanu and M. K. Jeyakumar, *Survey on crop pest detection using deep learning and machine learning approaches*, vol. 82, no. 27. pp. 42277–42310, Springer US, 2023.

[8] B. Kariyanna and M. Sowjanya, "Unravelling the use of artificial intelligence in management of insect pests," *Smart Agric. Technol.*, vol. 8, no. June, pp. 100517, 2024, doi: 10.1016/j.atech.2024.100517.

[9] A. Ahmad, D. Saraswat, and A. El Gamal, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," *Smart Agric. Technol.*, vol. 3, no. May 2022, pp. 2772-3755, 2023, doi: 10.1016/j.atech.2022.100083.

[10] P. Lanjewar, Madhusudan , G Morajkar and P. P, "Modified transfer learning frameworks to identify potato leaf diseases," *Multimed. Tools Appl.*, vol. 83, no. 17, pp. 50401–50423, 2024.

[11] Hongxing Peng, Huiming Xu1, Zongmei Gao, Zhiyan Zhou , Xingguo Tian, Qianting Deng, Huijun He andChunlong Xian, "Crop pest image classification based on improved densely connected convolutional network," *Front. Plant Sci.*, vol. 14, no. April, pp. 1–12, 2023, doi: 10.3389/fpls.2023.1133060.

[12] H Gong, T Liu, T Luo, J Guo, R Feng, J Li, X Ma, Y Mu, T Hu, Y Sun, S Li, Q Wang, Y Guo Agronomy, "Based on FCN and DenseNet Framework for the Research of Rice Pest Identification Methods," *Agronomy*, vol. 13, no. 2, pp. 1-14, 2023, doi: 10.3390/agronomy13020410.

[13] S. Yu, L. Xie, and Q. Huang, "Inception convolutional vision transformers for plant disease identification," *Internet of Things*, vol. 1, no. 21, pp. 100650, 2023.

[14] Zhuoxin Li, Cong Li, Linfan Deng, Yanzhou Fan, Xianyin Xiao, Huiying Ma, Juan Qin, Liangliang Zhu, "Improved AlexNet with Inception-V4 for Plant Disease Diagnosis," *Comput. Intell. Neurosci.*, vol. 2022, no. 5862600, pp. 1-12, 2022, doi: 10.1155/2022/5862600.

[15] G. Pattnaik, V. K. Shrivastava, and K. Parvathi, "Transfer Learning-Based Framework for Classification of Pest in Tomato Plants," *Appl. Artif. Intell.*, vol. 34, no. 13, pp. 981–993, 2020, doi: 10.1080/08839514.2020.1792034.

[16] V. Malathi and M. P. Gopinath, "Classification of pest detection in paddy crop based on transfer learning approach," *Acta Agric. Scand. Sect. B Soil Plant Sci.*, vol. 71, no. 7, pp. 552–559, 2021, doi: 10.1080/09064710.2021.1874045.

[17] N. N. Ahmad Loti, M. R. Mohd Noor, and S. W. Chang, "Integrated analysis of machine learning and deep learning in chili pest and disease identification," *J. Sci. Food Agric.*, vol. 101, no. 9, pp. 3582–3594, 2021, doi: 10.1002/jsfa.10987.

[18] J. Liu, X. Wang, W. Miao, and G. Liu, "Tomato Pest Recognition Algorithm Based on Improved YOLOv4," *Front. Plant Sci.*, vol. 13, no. July, pp. 1–10, 2022, doi: 10.3389/fpls.2022.814681.

[19] N. Ullah, J. A. Khan, L. A. Alharbi, A. Raza, W. Khan, and I. Ahmad, "An Efficient Approach for Crops Pests Recognition and Classification Based on Novel DeepPestNet Deep Learning Model," *IEEE Access*, vol. 10, no. July, pp. 73019–73032, 2022, doi: 10.1109/ACCESS.2022.3189676.

[20] D. Z. Karim, T. A. Bushra, and M. M. Saif, "PestDetector: A Deep Convolutional Neural Network to Detect Jute Pests," *2022 4th Int. Conf. Sustain. Technol. Ind. 4.0, STI 2022*, December 2022, vol. 12, no. 17, pp. 1–6, 2022, doi: 10.1109/STI56238.2022.10103326.

[21] S. M. Gómez-Pupo, M. A. Fennix Agudelo, A. Patiño-Vanegas, A. Patiño-Saucedo, and E. C. Mesa, "Convolutional neural networks for the recognition of diseases and pests in Cassava leaves (Manihot esculenta)," *Proc. LACCEI Int. Multi-conference Eng. Educ. Technol.*, vol. 2022-July, no. October, pp. 1-5, 2022, doi: 10.18687/LACCEI2022.1.1.759.

[22] Md. Simul Hasan Talukder, Mohammad Raziuddin Chowdhury, Md Sakib Ullah Sourav, Abdullah Al Rakin, Shabbir Ahmed Shuvo, Rejwan Bin Sulaiman, Musarrat Saberin Nipun, Muntarin Islam, Mst Rumpa Islam, Md Aminul Islam, Zubaer Haque, "JutePestDetect: An intelligent approach for jute pest identification using fine-tuned transfer learning," *Smart Agric. Technol.*, vol. 5, no. July, pp. 1-13, 2023, doi: 10.1016/j.atech.2023.100279.

[23] M. S. U. Sourav and H. Wang, "Intelligent Identification of Jute Pests Based on Transfer Learning and Deep Convolutional Neural Networks," *Neural Process. Lett.*, vol. 55, no. 3, pp. 2193–2210, 2023, doi: 10.1007/s11063-022-10978-4. LINK: https://link.springer.com/article/10.1007/s11063-022-10978-4

[24] I. Agustian, R. Faurina, S. I. Ishak, F. P. Utama, K. Dinata, and N. Daratha, "Deep learning pest detection on Indonesian red chili pepper plant based on fine-tuned YOLOv5," *Int. J. Adv. Intell. Informatics*, vol. 9, no. 3, pp. 383–401, 2023, doi: 10.26555/ijain.v9i3.864.

[25] M. S. H. Talukder, R. Bin Sulaiman, M. R. Chowdhury, M. S. Nipun, and T. Islam, "PotatoPestNet: A CTInceptionV3-RS-based neural network for accurate identification of potato pests," *Smart Agric. Technol.*, vol. 5, no. July, pp. 1-13, 2023, doi: 10.1016/j.atech.2023.100297.

[26] E. Maican, A. Iosif, and S. Maican, "Precision Corn Pest Detection: Two-Step Transfer Learning for Beetles (Coleoptera) with MobileNet-SSD," *Agric.*, vol. 13, no. 12, pp. 1-24, 2023, doi: 10.3390/agriculture13122287.

[27] S. D. Daphal and S. M. Koli, "Enhancing sugarcane disease classification with ensemble deep learning: A comparative study with transfer learning techniques," *Heliyon*, vol. 9, no. 8, pp. 1-19, 2023, doi: 10.1016/j.heliyon.2023.e18261.

[28] R. Lapcharoensuk, C. Fhaykamta, W. Anurak, W. Chadwut, and A. Sitorus, "Nondestructive Detection of Pesticide Residue (Chlorpyrifos) on Bok Choi (Brassica rapa subsp. Chinensis) Using a Portable NIR Spectrometer Coupled with a Machine Learning Approach," *Foods*, vol. 12, no. 5, pp. 1-13, 2023, doi: 10.3390/foods12050955.

[29] A. Sohel, S. Shakil, S. T. Siddiquee, and A. Al, "Enhanced Potato Pest Identification : A Deep learning approach for identifying potato pests," *IEEE Access*, vol. 12, no. 2024, pp. 172149-172151, 2024, doi: 10.1109/ACCESS.2024.3488730.

[30] S Banothu, K Madhavi, KMVM Kumar, R Gajula, C Mallikarjuna Rao, S Dixit, A Chhetri, "Plant disease identification and pesticides recommendation using Dense Net," *Cogent Eng.*, vol. 11, no. 1, pp. 1-11, 2024, doi: 10.1080/23311916.2024.2353080.

[31] M. J. Karim, M. O. F. Goni, M. Nahiduzzaman, M. Ahsan, J. Haider, and M. Kowalski, "Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad-CAM," *Sci. Rep.*, vol. 14, no. 1, pp. 1–23, 2024, doi: 10.1038/s41598-024-66989-9.