JOURNAL MEDICAL INFORMATICS TECHNOLOGY
medinftech.org ISSN 2988-7003 (online)

Identification of Potato Plant Pests Using the Convolutional Neural Network VGG16 Method

Sri Hadianti[1], Faruq Aziz[2], Daning Nur Sulistyowati[3], Dwiza Riana[4], Ridwan Saputra[5], and Kurniawantoro[6]*

[1,2,3,4,5,6]*Universitas Nusa Mandiri, Indonesia*

ARTICLE HISTORY

Received: 20 June 24
Final Revision: 28 June 24
Accepted: 29 June 24
Online Publication: 30 June 24

KEYWORDS

Convolutional Neural Network, Identification, Plant Pests, Potato, VGG16

CORRESPONDING AUTHOR

sri.shv@nusamandiri.ac.id

DOI

10.37034/medinftech.v2i2.37

A B S T R A C T

Pests are one of the main challenges in potato cultivation that can significantly reduce crop yields. Therefore, quick and accurate pest identification is crucial for effective pest control. This research aims to develop a pest identification system for potato plants using the Convolutional Neural Network (CNN) method with the VGG16 architecture. The dataset used consists of images of pests commonly found on potato plants. After the labeling process, these images were used to train the CNN VGG16 model. The research results show that the CNN VGG16 method can identify types of pests with an accuracy rate of 73%. The results serve as a reference to help farmers and agricultural practitioners detect the presence of pests earlier and take the necessary actions to reduce crop losses.

1. Introduction

Potato (Solanum tuberosum) is one of the major agricultural commodities worldwide that significantly contributes to global food security [1]. The consistent production of potatoes faces serious challenges due to pest attacks and diseases, which can cause significant losses in yields if not effectively managed [2]. Pest attacks such as flea beetles (Phyllotreta), cutworms (Agrotis), and diseases like root rot (Rhizoctonia solani) and bacterial wilt (Ralstonia solanacearum) are some of the main threats that commonly affect potato crops across various regions of the world [3].

Traditional approaches to identify pests and diseases in potato plants often rely on visual inspection by farmers or agricultural experts. While these methods can provide valuable initial information, they often have limitations in detecting pest attacks promptly, especially in large agricultural areas or when symptoms of pest infestation are difficult to recognize visually [4]. This can lead to delays in implementing effective preventive or curative actions, which in turn can reduce productivity and the quality of the harvest.

In the current digital era, advancements in digital image processing and artificial intelligence have opened new opportunities to enhance early detection and management of pest and disease attacks in plants. One prominent technique is Convolutional Neural Networks (CNN), an approach in deep learning that has proven effective in pattern recognition in image [5]. CNN allows systems to automatically learn important features from plant images, including symptoms of pest attacks that may be difficult for the human eye to recognize.

One popular variant of CNN is VGG16, renowned for its capability to classify images with high accuracy [6]. VGG16 has a deep architecture consisting of 16 convolutional layers and associated layers, enabling it to recognize complex visual patterns in images. This capability makes VGG16 an attractive choice for implementing systems to identify pests and diseases in potato plants.

The application of CNN VGG16 in agricultural contexts, particularly in identifying pests in potato plants, promises to provide more accurate and efficient solutions compared to conventional methods. This technology allows for in-depth analysis of potato plant images from various perspectives, detecting subtle symptoms that may be difficult to recognize manually [7]. Therefore, systems utilizing CNN VGG16 have the potential to become valuable tools for farmers and researchers in monitoring and managing pest attacks more effectively. They enable faster and more accurate decision-making processes, enhancing the ability to respond promptly to pest outbreaks.

Previous research has demonstrated the success of CNN VGG16 in various object recognition applications, including the detection and classification of pests and diseases in plants [8], [9], [10]. However to implement this technology widely in field practice, further validation is needed across various agricultural conditions, as well as integration with existing agricultural management systems [11].

This research aims to explore the full potential of CNN VGG16 in supporting comprehensive and practical management of pests and diseases in potato plants. By integrating artificial intelligence and advanced image processing technology, the goal is to develop systems that enhance agricultural production efficiency, strengthen global food security, and support sustainable and environmentally friendly farming practices.

2. Research Method

This research aims to identify pests in potato plants. The research methodology involves several main stages: data collection, data preprocessing, CNN model development, model training and evaluation, and result analysis. Figure 1 provides detailed information about each stage.

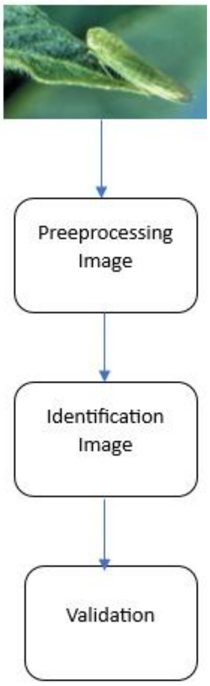

Figure 1. Research Method

2.1. Dataset

The dataset obtained from Kaggle regarding potato crop pests consists of 459 images categorized into 8 main pest classes. Each class represents a different type of pest such as Amrasca devastans, Aphis gossypii, Brachytrypes portentosus, Bemisia tabaci, Epilachna vigintioctopunctata, Agrotis ipsilon, Myzus persicae, and Phthorimaea operculella. The purpose of collecting this data is to support the development of an automated recognition and classification system that can assist farmers in early detection and management of pests in potato crops [12]. The available data may also include additional information such as image capture locations and plant conditions, which can provide further value for effective pest control analysis. By utilizing this dataset, researchers can train models to accurately identify various types of potato crop pests based on the available visual images. Figure 3 is an example of an image used in the dataset.

Figure 3. Potato Pets

2.2. Preprocessing Data

The preprocessing of the dataset involves several critical steps to ensure the images are ready for training a machine learning model, with a significant focus on image augmentation. Initially, each image is resized to a consistent dimension, typically 224x224 pixels, to standardize the input size. Normalization follows, scaling pixel values to a range of 0 to 1 to facilitate faster convergence during training. Various image augmentation techniques are then applied to increase the diversity of the dataset. These techniques include random rotations between -30 to 30 degrees, horizontal and vertical flipping, zooming in and out, horizontal and vertical shifting, shearing to introduce skewness, and adjustments to brightness and contrast to simulate different lighting conditions. The augmented dataset is subsequently split into training, validation, and test sets, typically allocating 70-80% for training, 10-15% for validation, and the remaining 10-15% for testing [13]. By implementing these preprocessing and augmentation strategies, the dataset becomes more robust, enhancing the performance and generalization capabilities of the trained model.

2.3. Identification Image

Image identification is performed using the VGG16 model. VGG16 is a well-known deep learning model, consisting of 16 layers, including convolutional and fully connected layers, designed to identify and classify images with high precision [14]. In the context of this research, the VGG16 model will be applied through several crucial stages. Initially, the pre-trained VGG16 model, trained using the ImageNet dataset, will be initialized. ImageNet is a large dataset containing millions of images classified into thousands of categories, providing the VGG16 model with a strong foundation for recognizing various visual features. The architecture of the VGG16 model can be seen in Figure 3.

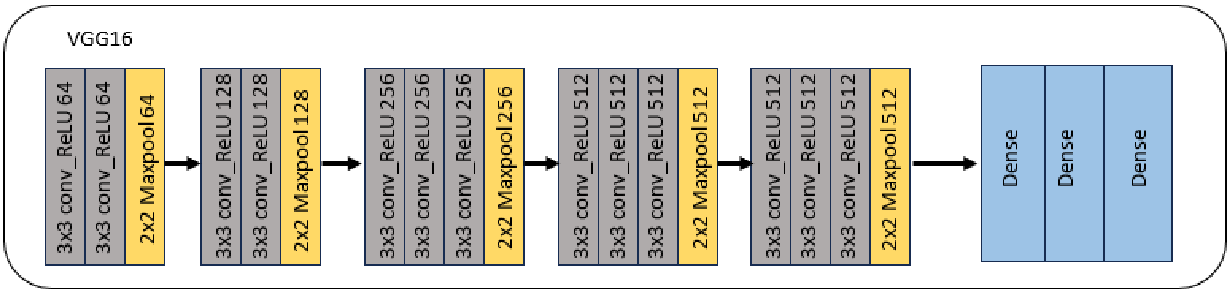

Figure 3. Architecture of VGG16

The next stage is fine-tuning, where the final layers of the VGG16 model will be modified and adjusted to identify specific classes of potato crop pests [15]. This involves replacing the last fully connected layer with a new layer that corresponds to the desired number of pest classes. After modification, several of the final layers of the model will be retrained using the pre-prepared potato crop pest image dataset. This process allows the model to adapt its knowledge from ImageNet to the specific task of identifying potato crop pests, thereby improving the accuracy and relevance of its identification results.

2.4. Validation

The model validation process is a crucial stage to ensure the model's capability in identifying and classifying potato plant pest images with high accuracy. This involves using a separate validation dataset from the training dataset to objectively measure the model's performance. Evaluation metrics such as accuracy, precision, recall, and F1-score play a role in assessing how well the CNN model can recognize various types of pests on potato plants [16]. These metrics help evaluate the reliability and accuracy of the model in specific classification tasks.

Additionally cross-validation is performed if possible to test the model's stability and consistency across different data variations. This step supports model generalization its ability to recognize pest images not seen during training, and ensures that evaluation results are trustworthy and reproducible in different scenarios. The primary goal of validation is to ensure that the developed CNN model can be relied upon for accurate pest identification on potato plants. This evaluation aims to assess the model's generalization ability in a broader context and to ensure that all training and evaluation procedures are conducted with valid methodologies that can be replicated to validate the findings of this research.

3. Result and Discussion

In this study we tested 459 images of potato plant pests using the CNN VGG16 model, with the following results:

3.1. Preprocessing Data

The data which originally amounted to 459 images had a different number for each class, as seen in Figure 4. To avoid class imbalance, an augmentation technique was used.

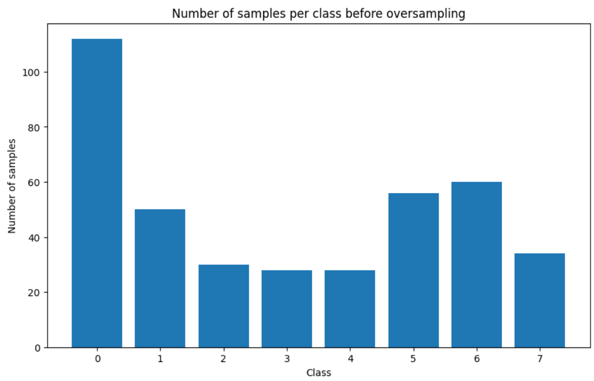

Figure 4. Samples Class Before Augmentation and Oversampling

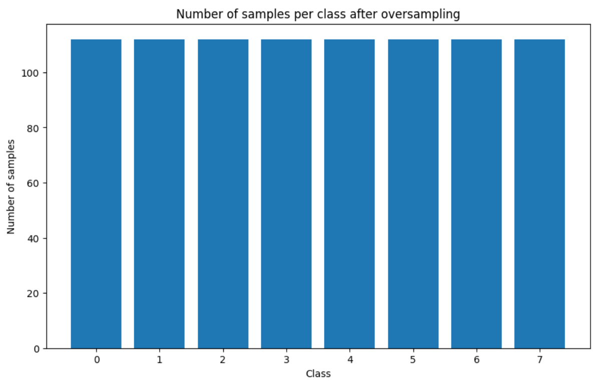

Figure 5. Samples Class After Augmentation and Oversampling

Figure 4 and Figure 5 can be seen that this indicates that the dataset initially had 459 samples before applying data augmentation and oversampling techniques. The classes in the dataset were not balanced, meaning they had different numbers of samples. After applying these techniques, the dataset now contains 896 samples. Furthermore, each class in the dataset now has 112 samples, which suggests that the oversampling technique was used to balance the classes. This increased size and balanced distribution of the data can potentially improve the performance of machine learning models by reducing the impact of class imbalance and increasing the robustness of the models to overfitting.
The augmented dataset is subsequently split into training, validation, and test sets typically allocating 80% for training, 10% for validation, and the remaining 10% for testing. The distribution of the dataset can be seen in Table 1.

Table 1. Distribution Dataset

Class	Training	Validation	Testing
Amrasca devastans	89	11	11
Aphis gossypii	89	11	11
Brachytrypes portentosus	89	11	11
Bemisia tabaci	89	11	11
Epilachna vigintioctopunctata	89	11	11
Agrotis ipsilon,	89	11	11
Myzus persicae	89	11	11
Phthorimaea operculella	89	11	11

The table provides the distribution of samples across different classes in the training, validation, and testing sets. Each class represents a specific insect species. The numbers indicate the number of samples for each class in each set. This suggests that the dataset is split into three parts: training, validation, and testing. Each class has the same number of samples in the training set (89), and the same number of samples in the validation and testing sets (11). This balanced distribution can be beneficial for training machine learning models, as it helps to reduce the impact of class imbalance and ensures that each class is adequately represented in each set.

3.1. Model Summary

Table 1 provides information from the VGG16 model summary model used in this research.

Referring to Table 2, it is evident that the VGG16 architecture begins with an input layer that accepts images sized 224x224 with three color channels (RGB). The first convolutional layer generates 64 feature maps with 1,792 parameters. Subsequently, the second convolutional layer also produces 64 feature maps, but with an increased parameter count of 36,928.

Table 2. Model Summary

Layer (type)	Pixel	Parameter
InputLayer	224, 224, 3	0
Conv2D	224, 224, 64	1792
Conv2D	224, 224, 64	36928
MaxPooling2D	112, 112, 64	0
Conv2D	112, 112, 128	73856
Conv2D	112, 112, 128	147584
MaxPooling2D	56, 56, 128	0
Conv2D	56, 56, 256	295168
Conv2D	56, 56, 256	590080
Conv2D	56, 56, 256	590080
MaxPooling2D	28, 28, 256	0
Conv2D	28, 28, 512	1180160
Conv2D	28, 28, 512	2359808
Conv2D	28, 28, 512	2359808
MaxPooling2D	14,14,512	0
Conv2D	14,14,512	2359808
Conv2D	14,14,512	2359808
Conv2D	14,14,512	2359808
MaxPooling2D	7, 7, 512	0
Flatten	25088	0
Dense	4096	102764544
Dense	4096	16781312
Dense	1000	4097000

The first pooling layer then reduces the resolution of the feature maps to 112x112x64 without adding any parameters. This process repeats with convolutional layers generating 128, 256, and 512 feature maps at progressively lower resolutions, interspersed with pooling layers that further reduce the resolution. These convolutional layers are followed by a flatten layer that converts the 3D tensor into a 1D vector of size 25,088. The model concludes with three fully connected layers: the first two layers each have 4,096 units with parameters totaling 102,764,544 and 16,781,312 respectively, while the final layer has 1,000 units for classification with parameters totaling 4,097,000.

3.3. Evaluation Model

Below are visual representations depicting the accuracy and loss values obtained from experiments conducted using the optimal epoch, which is epoch 10.

Figures 6 and 7 display the progress of accuracy and loss values for each epoch during training and validation. The yellow line represents the changes in accuracy and loss values for the validation data, while the blue line indicates these changes for the training data. The graphs illustrate a gradual decrease in loss values up to epoch 10, alongside a continuous increase in accuracy values until epoch 10. By the end of the epochs, the training data achieved an accuracy of 0.73, and the validation data reached an accuracy of 0.68.

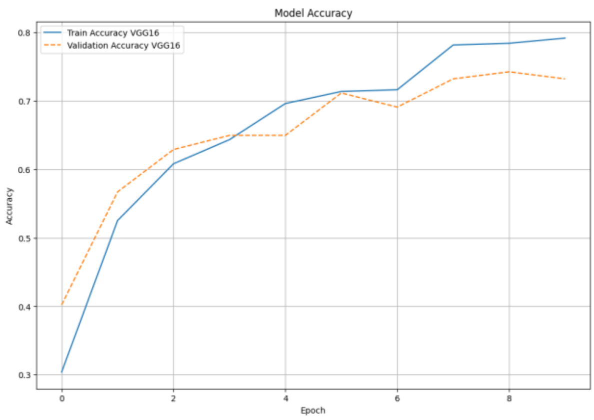

Figure 6. Accuracy of VGG16

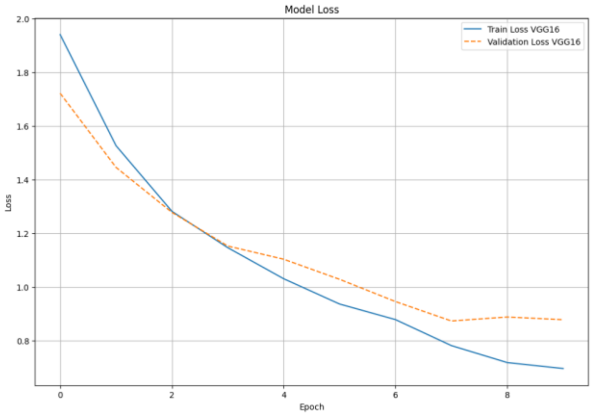

Figure 7. Loss of VGG16

Table 3. Evaluation Result

	Precision	Recall	F1 – Score	Support
0	0.96	0.81	0.88	27
1	0.42	0.92	0.58	12
2	1.00	0.29	0.44	7
3	1.00	0.29	0.44	7
4	0.75	0.86	0.80	7
5	0.88	1.00	0.93	14
6	0.62	0.53	0.57	15
7	0.86	0.75	0.80	8
Accuracy			0.73	97
Macro Avg	0.81	0.68	0.68	97
Weighted Avg	0.81	0.73	0.73	97

The evaluation results of the CNN model using the VGG16 architecture for potato pest identification demonstrate promising performance can be seen in Table 3. The model achieved an overall accuracy of 73% across all classes, as indicated in the precision, recall, and F1-score metrics for each pest class. Precision values ranged from 0.42 to 1.00, with class 0 showing the highest precision at 0.96. Recall values varied from 0.29 to 1.00, with class 1 exhibiting the highest recall at 0.92. The F1-scores also varied, with class 5 achieving the highest F1-score of 0.93, indicating a balanced performance between precision and recall. The macro average scores for precision, recall, and F1-score were 0.81, 0.68, and 0.68 respectively, while the weighted averages were 0.81, 0.73, and 0.73. These results suggest that while the model performs well in general, there is room for improvement, particularly in enhancing the recall for classes with fewer samples. Overall, this study provides a solid foundation for developing more efficient and accurate pest detection systems, which can aid farmers in better managing their crops and reducing losses due to pest infestations.

4. Conclusion

The results of the conducted research show promising outcomes in the application of Convolutional Neural Network (CNN) technology for potato pest identification. The use of the VGG16 architecture in this model achieved an accuracy of 73% after 10 epochs of training. This indicates that the VGG16 model has good capabilities in recognizing visual patterns associated with potato pests. However, there is still room for improvement in terms of both accuracy and model efficiency. This research provides a strong foundation for the development of more efficient and accurate pest detection systems in the future, which in turn can help farmers manage their crops better and reduce losses due to pest infestations. Further implementation with larger datasets, greater data variety, and model parameter optimization can enhance the performance of this model. Thus, this technology has the potential to be widely applied in modern agriculture to increase productivity and sustainability.

Acknowledgements

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: Pengembangan Model Deep Learning dalam Mendukung Identifikasi Awal Varietas Hama Tanaman Kentang Berbasis Kecerdasan Buatan", 2024.

References

[1] J. M. Al-Khayri, S. M. Jain, and D. V. Johnson, *Advances in Plant Breeding Strategies: Vegetable Crops: Volume 8: Bulbs, Roots and Tubers*, vol. 8. 2021.

[2] J. F. Kreuze, J. A. C. Souza-Dias, A. Jeevalatha, A. R. Figueira, J. P. T. Valkonen, and R. A. C. Jones, “Viral Diseases in Potato,” The Potato Crop, pp. 389–430, 2020, doi: 10.1007/978-3-030-28683-5_11.

[3] A. Devaux, J. Goffart, and P. Kromann, “The Potato of the Future : Opportunities and Challenges in Sustainable Agri-food Systems Introduction : The Current Situation of Global Food Security,” *Potato Res.*, vol. 64, pp. 681–720, 2021.

[4] T. Domingues, T. Brandão, and J. C. Ferreira, “Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey,” *Agric.*, vol. 12, no. 9, pp. 1–23, 2022, doi: 10.3390/agriculture12091350.

[5] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, “A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, 2022, doi: 10.1109/TNNLS.2021.3084827.

[6] R. A. Sholihati, I. A. Sulistijono, A. Risnumawan, and E. Kusumawati, “Potato Leaf Disease Classification Using Deep Learning Approach,” *IES 2020 - Int. Electron. Symp. Role Auton. Intell. Syst. Hum. Life Comf.*, pp. 392–397, 2020, doi: 10.1109/IES50839.2020.9231784.

[7] H. Tian, T. Wang, Y. Liu, X. Qiao, and Y. Li, “Computer vision technology in agricultural automation —A review,” *Inf. Process. Agric.*, vol. 7, no. 1, pp. 1–19, 2020, doi: 10.1016/j.inpa.2019.09.006.

[8] C. R. Rahman *et al.*, “Identification and recognition of rice diseases and pests using convolutional neural networks,” *Biosyst. Eng.*, vol. 194, pp. 112–120, 2020, doi: 10.1016/j.biosystemseng.2020.03.020.

[9] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, “Recent

advances in image processing techniques for automated leaf pest and disease recognition – A review," *Inf. Process. Agric.*, vol. 8, no. 1, pp. 27–51, 2021, doi: 10.1016/j.inpa.2020.04.004.

[10] L. Li, S. Zhang, and B. Wang, "Plant Disease Detection and Classification by Deep Learning - A Review," *IEEE Access*, vol. 9, no. Ccv, pp. 56683–56698, 2021, doi: 10.1109/ACCESS.2021.3069646.

[11] A. Villa-Henriksen, G. T. C. Edwards, L. A. Pesonen, O. Green, and C. A. G. Sørensen, "Internet of Things in arable farming: Implementation, applications, challenges and potential," *Biosyst. Eng.*, vol. 191, pp. 60–84, 2020, doi: 10.1016/j.biosystemseng.2019.12.013.

[12] 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, p. 100297, 2023, doi: 10.1016/j.atech.2023.100297.

[13] Z. Yang, R. O. Sinnott, J. Bailey, and Q. Ke, *A survey of automated data augmentation algorithms for deep learning-based image classification tasks*, vol. 65, no. 7. Springer London, 2023.

[14] Z. P. Jiang, Y. Y. Liu, Z. E. Shao, and K. W. Huang, "An improved VGG16 model for pneumonia image classification," *Appl. Sci.*, vol. 11, no. 23, 2021, doi: 10.3390/app112311185.

[15] N. Tilahun, "Potato'S Leaves Disease Detection Model Using Convolutional Neural Network," no. November, 2020.

[16] S. Aladhadh, S. Habib, M. Islam, M. Aloraini, M. Aladhadh, and H. S. Al-Rawashdeh, "An Efficient Pest Detection Framework with a Medium-Scale Benchmark to Increase the Agricultural Productivity," *Sensors*, vol. 22, no. 24, pp. 1–17, 2022, doi: 10.3390/s22249749.