Benchmarking CNN-Based Systems for Corn Leaf Pest Detection using Fine-Tuning

Risco Cosavalente, Mariana Lucia*
Universidad Nacional de Trujillo
Trujillo, Perú
mrisco@unitru.edu.pe

Chang Román, Sulei Jackeline †
Universidad Nacional de Trujillo
Trujillo, Perú
schang@unitru.edu.pe

Cortegana Silva, Carlos Arturo ‡
Universidad Nacional de Trujillo
Trujillo, Perú
ccortegana@unitru.edu.pe

Abstract

This research presents a computer vision system for the detection of diseases in maize leaves using convolutional neural networks. The Peruvian valley of Chicama was the focus of our study, where images were collected and subsequently added to the Plant Village dataset. Image preprocessing techniques, including GrabCut and data augmentation, were employed to enhance the quality of the images. We compared a number of fine-tuned architectures, including DenseNet121, DenseNet201, ResNet50, ResNet101, VGG16 and VGG19, to identify the most accurate model for maize leaf diseases. The results demonstrated that VGG16 achieved the highest accuracy of 93.16%. DenseNet121 followed closely with an accuracy of 93.03%, indicating its strong performance. In contrast, ResNet50 showed the lowest accuracy at 87.94%. The complete implementation can be accessed at **Github**

1 Introduction

In recent years, developed countries have made notable advancements in the implementation of farming techniques. These innovations have emerged primarily in response to the necessity of detecting and controlling plant pests and diseases, which represent a significant phytosanitary challenge in agriculture. The term "pest" is used to describe any organism that interferes with the functioning or structure of plants, thereby affecting their health and productivity[14]. These phytosanitary problems result in considerable losses, amounting to 20-40% of global agricultural production on an annual basis [8]. Despite its reputation as a resilient crop, maize is frequently afflicted by diseases and pests. The resulting losses are reflected in both the quantity and quality of the harvested crops[14]. As agricultural products constitute a fundamental component of the human diet and represent a significant source of income for many countries, the impact of pests has not only social repercussions but also a considerable economic impact.

Maize is susceptible to a number of significant foliar diseases, including Common Rust (Puccinia sorghi), which presents as brown pustules and is exacerbated in cool, humid climates; Grey Spot (Cercospora zeae-maydis), identifiable by greyish leaf spots, which thrives in high humidity and warm

*`https://www.linkedin.com/in/mariana-risco-cosavalente/`
†`https://www.linkedin.com/in/sulei-jackeline-chang-roman-33358b283/`
‡`https://www.linkedin.com/in/arturocsdev/`

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

temperatures; and Northern Leaf Blight (Exserohilum turcicum), which causes elongated lesions and can severely affect susceptible hybrids [14]. The aforementioned diseases are illustrated in Figure 1, which presents images captured by our research team in the Chicama Valley. In addition to affecting yield, these diseases will be examined as cases in the neural network, with reference to a healthy leaf to identify optimal conditions.

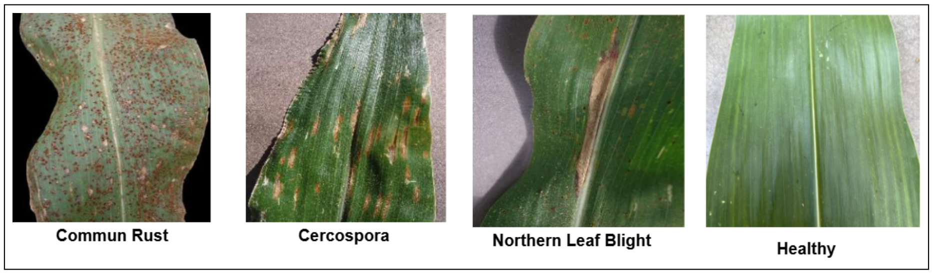

Figure 1: Pest on Corn Leaf “Commun Rust” (reddish-brown pustules), “Cercospora” (reddish-brown borders), “Northern Leaf Blight” (elliptical grayish-green lesions) and “Healthy".

Maize plays an important role in Peru and globally that extends well beyond its function as a foodstuff. Globally, maize is essential for a number of reasons, including its significant production volume and its role in human and animal feed. Additionally, it serves as a raw material in the manufacture of a range of industrial products, including oils, flours, sweeteners, ethanol, and even plastics and fibers. This grain is fundamental in various industries, including the automotive, pharmaceutical and cosmetic industries. It is also used in the manufacture of organic fertilizers and other industrial compounds that drive the world economy [4].

In Peru, maize is cultivated across the entire country, from the coastal regions to the highlands and jungle, reflecting its adaptability to diverse altitudes and geographical conditions. Hard yellow maize is the predominant variety grown on the Peruvian coast and in the jungle, whereas starchy maize is more common in the Andes [4]. Given the extensive geographical distribution and economic importance of maize, the implementation of systems for the early detection of maize leaf disease is of vital importance. Such systems will ensure the continued productivity and sustainability of this crop, thus protecting the income of Peruvian farmers and ensuring the supply of this strategic resource.

This study employs computer vision techniques to identify the presence of diseases in maize leaves in the Chicama Valley, Peru. A number of different Convolutional Neural Networks (CNN) architectures are evaluated, including DenseNet121, DenseNet201, ResNet50, ResNet101, VGG16 and VGG19 [7, 5] . Given the limited size of the dataset, we employed fine-tuning transfer learning. The weights of the pre-trained models were derived from the Imagenet dataset [1]. The objective of this research is to identify the optimal model for the classification of plant diseases.

1.1 Related Works

In their research, Saputra et al. [12] compared the performance of three convolutional neural network models (DenseNet121, DenseNet169 and DenseNet201) for the classification of rice leaf diseases. In their methodology, the researchers employ convolutional neural networks, specifically the DenseNet family of models, to train an algorithm capable of accurately detecting diseases in rice leaves. The results obtained, such as 91.67% accuracy with DenseNet121, illustrate the potential of these models for agricultural applications. This approach is closely related to my current work, in which I am implementing the DenseNet201 and DenseNet101 models for disease detection in plants, specifically in maize leaves.

Ganatra and Patel [3], explores the potential of convolutional neural network (CNN) models, including VGG16, Inception V4, ResNet50 and ResNet101, for the early detection of plant leaf diseases. A dataset comprising 38 classes and 87,000 images was used to implement transfer learning for training the models. The results demonstrate that ResNet50 and ResNet101 achieve accuracies of 99.70% and 99.73%, respectively, thereby highlighting their effectiveness in plant disease classification.

Hu et al. [6] this study addresses the detection of diseases in corn leaves using convolutional neural networks, data augmentation techniques and transfer learning. The researchers put forth an optimized model for the classification of diseases, including Corn Gray Leaf Spot, Corn Common Rust and Corn Northern Leaf Blight, as well as healthy leaves. The process included data expansion and parameter fitting in pre-trained models, including GoogLeNet, ResNet18, VGG16 and VGG19, resulting in an average accuracy of 97.6%. The two works are directly pertinent to the present research, as the objective is also to detect the same diseases in maize leaves using models such as DenseNet201, DenseNet101, ResNet50, ResNet101, VGG16 and VGG19. Furthermore, as in the aforementioned studies, transfer learning and data augmentation techniques were employed for the purpose of comparing the performance of different networks and optimizing disease classification accuracy.

2 Materials and Methods

Figure 2 illustrates the methodology employed for the comparative analysis of the models, as described in the following section.

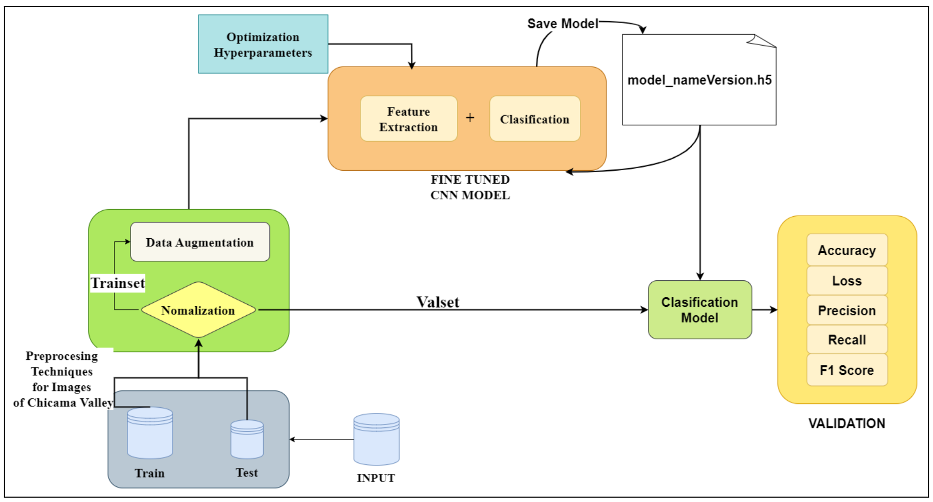

Figure 2: Methodology block diagram, we used data augmentation and fine tuning models

2.1 Dataset Collection

This project employs two distinct image sources. The Plant Village Dataset and images captured in the Chicama Valley.

Plant Village Dataset The dataset, sourced from Mohanty [10], comprises pre-processed RGB images, with colour correction and without background, which proved advantageous for the purposes of neural network training.

Images of Chicama Valley The images were collected by taking photographs of corn leaves in the Chicama Valley. The presence of Cercospora Grey Spot and Northern Leaf Blight diseases was identified, as well as healthy corn leaves. Following this, labelling was carried out. In order to integrate the newly captured images into the existing Plant Village dataset, it was necessary to apply some image pre-processing techniques. The images constituted part of the training set.

The dataset comprises a total of, , 256 x 256 images, including those that have been added with the aforementioned pre-processing. Of these, 6,256 are included in the training set and 804 in the validation set.

2.2 Preprocessing Techniques

2.2.1 GrabCut

The GrabCut tool was utilised to remove the background from the images, as illustrated in Figure 3. This technique was employed on images of Cercospora Grey Spot and Northern Leaf Blight diseases, utilising a rectangle for cropping [11]. Pixels situated outside the rectangle were designated as known, and pixels were assigned to the most probable component of the GMM models (foreground or background). Thereafter, a new GMM was trained and graph pruning was performed to update the pixel classifications. This process was repeated until convergence [11].

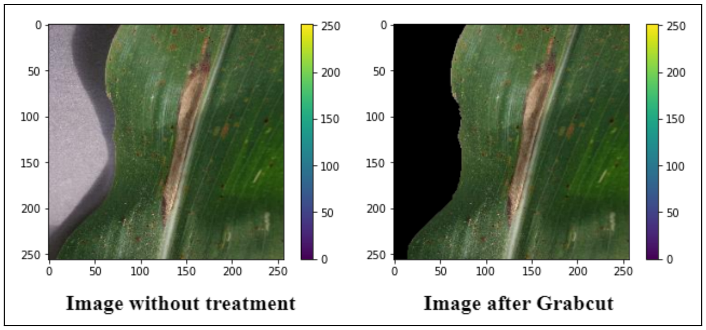

Figure 3: Grabcut preprocessing of corn leaves in the Chicama Valley.

2.2.2 Zoom Center

The zoom technique was developed by our research team in the analysis of images of healthy maize leaves obtained in the Chicama Valley (see Figure 4). This pre-processing technique enables the user to focus attention on specific areas of the image by zooming in on the centre, which is particularly useful for highlighting relevant features that may be crucial for the identification and classification of plant diseases.

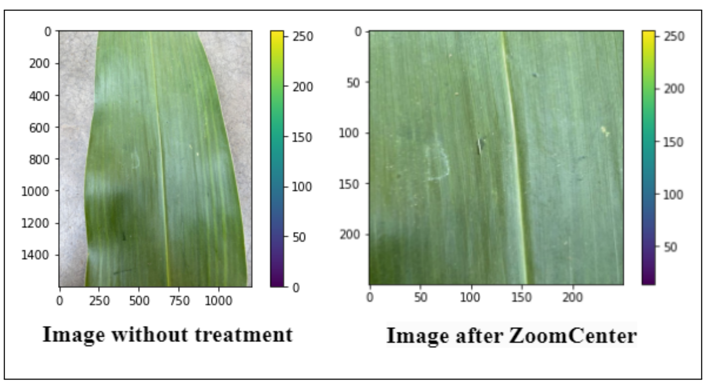

Figure 4: Zoom Center preprocessing of healthy corn leaves in the Chicama Valley.

2.2.3 Data Augmentation

The purpose of this stage is to enhance the robustness of the dataset, as this improves the efficacy of the learning model. Augmentation is conducted on the existing images, employing data augmentation techniques such as rotation, scaling, image cropping, modifying the width and height of the image, zooming, and so forth. To enable CNN training, it is essential to provide the model with the requisite data to facilitate learning and information extraction. The techniques applied are detailed in Table 1.

Table 1: Data Augmentation Techniques

Technique	Range/Description
Scaling	1/255.0
Random Rotation	0° to 45°
Random Zoom	0.3
Random Shift	Width and height shift (0.2)
Shear	0.4
Horizontal Flip	True

2.3 Convolutional Neural Networks Models for Image Classification

This research employs deep learning models as a methodology for the classification of corn leaf diseases. The investigation encompasses three architectural models and their respective variations. Convolutional neural networks have been developed and refined over time. The following architectures will be evaluated using fine-tuning.

2.3.1 VGG

Simonyan and Zisserman [13] focused on investigating how increasing network depth affects performance in large-scale image recognition tasks. Specifically, they explored using small convolutional filters (3x3) in deeper networks, with configurations that ranged from 16 to 19 layers. The results showed a significant improvement in image classification accuracy compared to previous models, particularly in tasks like localization and classification within the ImageNet dataset [1]. One of the most significant contributions of VGG was to illustrate that augmenting the depth of the network while preserving the simplicity of the convolution filters can result in more precise models for image recognition. This proved an effective solution to the problem of improving performance in this area without significantly increasing model complexity. As a result, VGG has become a benchmark in computer vision and has influenced the development of subsequent deep network architectures.

2.3.2 ResNet

He et al. [5] propose ResNet architecture to address the challenges of training deep neural networks, particularly the degradation problem, where adding more layers to a network results in worse performance. To address this issue, the researchers introduced 'shortcut' connections that enable layers to learn residual functions in place of the original functions, thereby facilitating the training of networks comprising hundreds of layers. This approach has been shown to achieve superior performance in image recognition tasks, as evidenced by the results of the ImageNet challenge. Models such as ResNet-50, ResNet-101, and ResNet-152, which have 50, 101, and 152 layers, respectively, have demonstrated this effectiveness despite their depth. Their design allows for efficient training with far more layers than previous architectures like VGG.

2.3.3 DenseNet

In their research, Huang et al. [7] propose DenseNet, a convolutional network architecture that seeks to increase the depth of the network. This is achieved through the use of dense blocks, where each layer receives input from previous layers to preserve the accuracy of the classifier (feedback). In contrast to traditional convolutional neural networks (CNNs), which have L layers and L connections, DenseNet has L(L+1)/2 direct connections. The author put forth this neural network as it has fewer parameters than other architectures, resulting in a shorter computation time. Additionally, it addresses the leakage gradient issue and necessitates less memory for optimal performance.

2.4 Hyperparameters

The Adam optimizer was selected for its capacity to adapt the learning rate, which facilitates convergence in neural network training. The Adam optimizer combines optimization techniques such as gradient descent with moments and RMSProp[9]. Additionally, as illustrated in Table 2, we have implemented the ReduceLROnPlateau parameter to adjust the learning rate by the validation metric,

Table 2: Hyperparameters used by each model.

Model	Learning Rate	Factor	Patience	Min LR
DenseNet121	0.0001	0.8	10	1e-5
DenseNet201	0.0001	0.7	10	1e-5
ResNet50	0.0001	0.3	10	1e-6
ResNet101	0.0001	0.3	10	1e-6
VGG16	0.0001	0.3	10	1e-6
VGG19	0.0001	0.5	10	1e-6

thereby preventing overfitting and enhancing generalization. These parameters have been adjusted for each model.

3 Results

3.1 Architecture models Fine-Tuning

Fine-tuning with transfer learning will be employed due to the limited size of the dataset. This method exploits pre-trained models that have acquired pertinent features from extensive data sets, thereby enhancing the precision of specific tasks with limited data availability [16, 2]. The final (CNN) architectures that were implemented are illustrated in Figure 5.

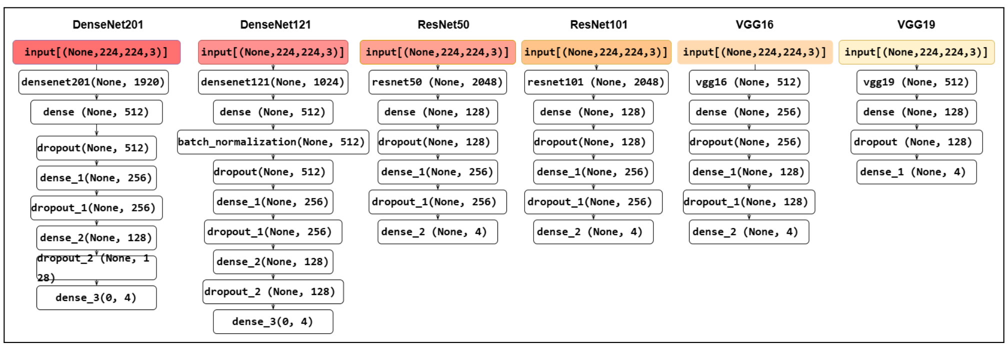

Figure 5: Architecture models Fine-Tuning of DenseNet121, DenseNet201, ResNet50, ResNet101, VGG16 and VGG19

3.2 Training and Validation Losses

The models were trained for 10 epochs with ImageNet dataset weights [1]. In the DenseNet and VGG architectures, the final 50 training layers were unfrozen, whereas in ResNet only the final 30 layers were unfrozen.

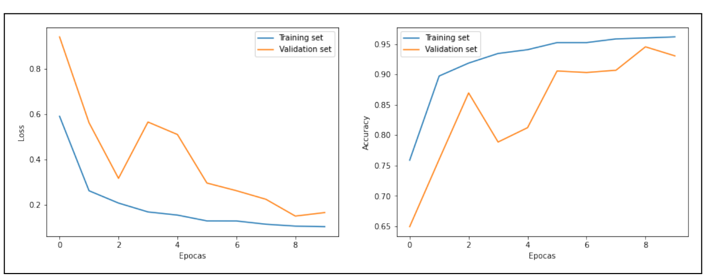

Figure 6: Accuracy and Loss DenseNet121.

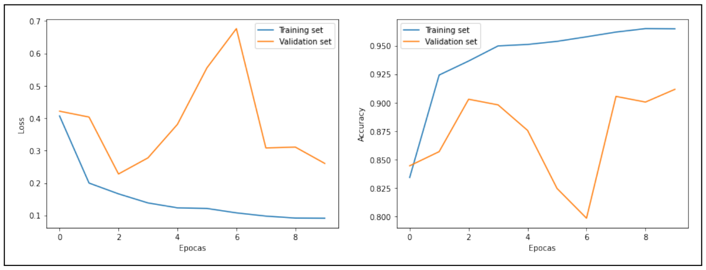

Figure 7: Accuracy and Loss DenseNet201.

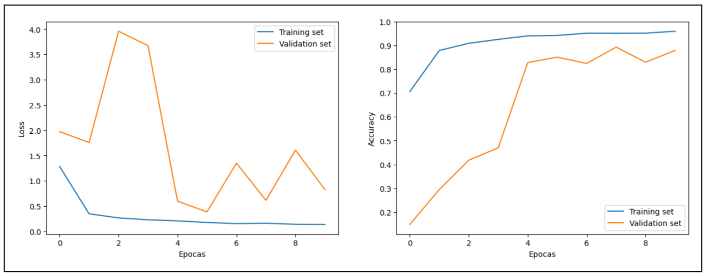

Figure 8: Accuracy and Loss ResNet50.

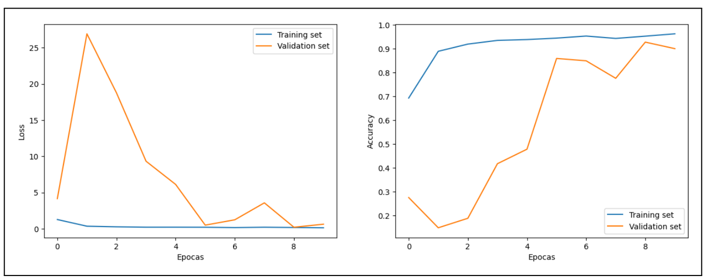

Figure 9: Accuracy and Loss ResNet101.

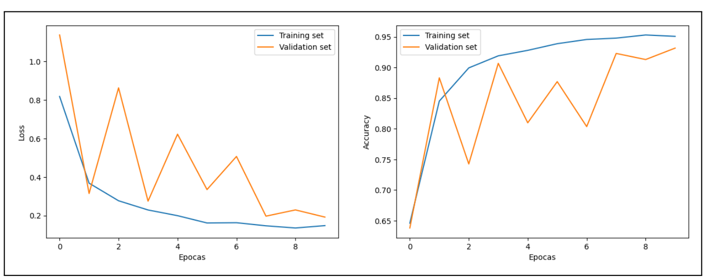

Figure 10: Accuracy and Loss VGG16.

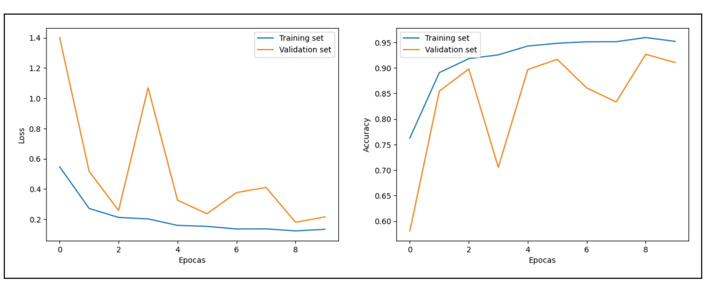

Figure 11: Accuracy and Loss VGG19.

4 Discussion

We evaluated the models using accuracy, precision, recall, and F1-score metrics. Integrated Gradients Sundararajan et al. [15] were used to highlight key features in the images. Figure 11 shows that DenseNet201 provides the clearest focus on leaf details. Models with better performance (DenseNet121 and DenseNet201) tend to highlight the affected areas more accurately.

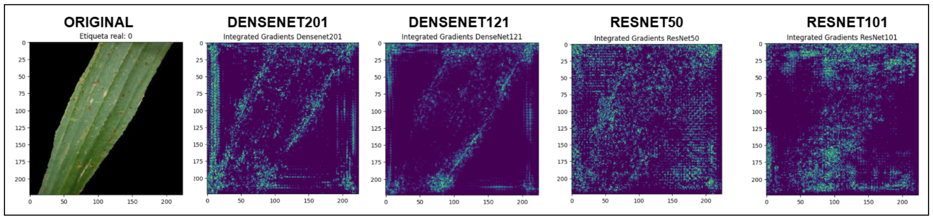

Figure 12: Visualization of important features identified using Integrated Gradients in DenseNet121, DenseNet201, ResNet50, and ResNet101. Adapted from Sundararajan et al. [15].

The results are summarised in Table 3. Among the models evaluated, VGG16 stood out with the highest accuracy of 93.16%, accompanied by a precision of 93.68%. This demonstrates its effectiveness in classifying maize leaf diseases.

DenseNet121 also showed a solid performance with an accuracy of 93.03%, highlighting that a smaller number of parameters (7,729,092) can still deliver competitive results compared to more complex models such as ResNet101 (42,954,500).

In contrast, ResNet50 achieved an accuracy of 87.94%, reflecting difficulties in identifying certain classes, lower than the 99.70% and 99.73% reported by Ganatra and Patel [3] for ResNet50 and ResNet101, respectively, in the early detection of plant leaf diseases. DenseNet201 and VGG19 showed accuracies of 91.17% and 91.04%, respectively, but did not outperform DenseNet121. These results are comparable to those reported by Saputra et al. [12], who achieved an accuracy of 91.67% with DenseNet121 in classifying rice leaf diseases. This indicates that models belonging to the DenseNet family are effective in agricultural applications.

Table 3: Result Evaluation metrics

Model	**Metrics**				**Total Params**
	Accuracy	**Precision**	**Recall**	**F1-Score**	
DenseNet121	**0.9303**	**0.9305**	**0.9303**	**0.9294**	7,729,092
DenseNet201	0.9117	0.9184	0.9117	0.9112	19,470,276
ResNet50	0.8794	0.9002	0.8794	0.8794	23,884,036
ResNet101	0.9005	0.9119	0.9005	0.9002	42,954,500
VGG16	**0.9316**	**0.9368**	**0.9316**	**0.9311**	14,879,428
VGG19	0.9104	0.9155	0.9104	0.9096	20,090,564

5 Conclusion

The results of this study demonstrate the efficacy of convolutional neural network (CNN) models in the detection of diseases in maize leaves. VGG16 demonstrates an accuracy of 93.16%, while DenseNet121 exhibits an accuracy of 93.03%. Conversely, DenseNet201 and ResNet101 demonstrate comparatively diminished performance, with accuracy levels of 91.17% and 90.05%, respectively. ResNet50 exhibits the least favorable performance, with an accuracy level of 87.94%. These findings are consistent with those of previous studies highlighting the effectiveness of CNN architectures in agricultural applications. The selection of an appropriate model is paramount, and the application of transfer learning and data augmentation techniques is essential to enhance the accuracy of plant disease classification. In conclusion, the findings indicate that VGG16 and DenseNet121 are the optimal choices for disease classification on corn leaves in the Chicama Valley. This indicates a promising direction for future research in this field, including the classification of plant diseases using video data.

References

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. doi: 10.1109/CVPR.2009.5206848. URL `https://doi.org/10.1109/CVPR.2009.5206848`.

[2] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. Decaf: A deep convolutional activation feature for generic visual recognition, 2013. URL `https://arxiv.org/abs/1310.1531`.

[3] Nilay Ganatra and Atul Patel. Performance analysis of fine-tuned convolutional neural network models for plant disease classification. *International Journal of Control and Automation*, 13(3):293–305, 2020.

[4] Pedro José García Mendoza. El cultivo del maíz en el mundo y en perú. *Revista De Investigaciones De La Universidad Le Cordon Bleu*, 2017. doi: https://doi.org/10.36955/RIULCB.2017v4n2.005. URL `https://revistas.ulcb.edu.pe/index.php/REVISTAULCB/article/view/65`.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. URL `https://arxiv.org/abs/1512.03385`.

[6] Rongjie Hu, Shan Zhang, Peng Wang, Guoming Xu, Daoyong Wang, and Yuqi Qian. The identification of corn leaf diseases based on transfer learning and data augmentation. In *Proceedings of the 3rd International Conference on Computer Science and Software Engineering*, CSSE '20, page 58–65, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450375528. doi: 10.1145/3403746.3403905. URL `https://doi.org/10.1145/3403746.3403905`.

[7] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks, 2018. URL `https://arxiv.org/abs/1608.06993`.

[8] Mohamed Esmail Karar, Fahad Alsunaydi, Sultan Albusaymi, and Sultan Alotaibi. A new mobile application of agricultural pests recognition using deep learning in cloud computing system. *Alexandria Engineering Journal*, 60(5):4423–4432, 2021. ISSN 1110-0168. doi: https://doi.org/10.1016/j.aej.2021.03.009. URL `https://www.sciencedirect.com/science/article/pii/S1110016821001642`.

[9] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017. URL `https://arxiv.org/abs/1412.6980`.

[10] SP Mohanty. Plantvillage-dataset: Dataset of diseased plant leaf images and corresponding labels, 2016. URL `https://github.com/spMohanty/PlantVillage-Dataset`.

[11] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "grabcut": interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314, aug 2004. ISSN 0730-0301. doi: 10.1145/1015706.1015720. URL `https://doi.org/10.1145/1015706.1015720`.

[12] Adi Dwifana Saputra, Djarot Hindarto, and Handri Santoso. Disease classification on rice leaves using densenet121, densenet169, densenet201. *Sinkron: jurnal dan penelitian teknik informatika*, 7(1):48–55, 2023. doi: 10.33395/sinkron.v8i1.11906. URL `https://doi.org/10.33395/sinkron.v8i1.11906`.

[13] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015. URL `https://arxiv.org/abs/1409.1556`.

[14] D. R. Smith and D. G. White. *Diseases of Corn*, chapter 12, pages 687–766. John Wiley Sons, Ltd, 1988. ISBN 9780891182122. doi: https://doi.org10.2134/agronmonogr18.3ed.c12. URL `https://acsess.onlinelibrary.wiley.com/doi/abs/10.2134/agronmonogr18.3ed.c12`.

[15] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks, 2017. URL `https://arxiv.org/abs/1703.01365`.

[16] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks?, 2014. URL `https://arxiv.org/abs/1411.1792`.