JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 06 Issue 01
2023
Research Article
https://doi.org/10.56979/601/2023

Olive Leaves Disease Detection and Classification using Deep Learning

Nida Anwar[1], Sohail Akhtar[2], Muhammad Munwar Iqbal[2]*, Noureen Hameed[1], and Ammad Noor[2]

[1]Department of Computer Science and IT, Virtual University of Pakistan, Lahore, Pakistan.
[2]Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan.
Corresponding Author: Muhammad Munwar Iqbal. Email: munwariq@gmail.com.

Received: November 07, 2023 **Accepted:** December 01, 2023 **Published:** December 05, 2023

Abstract: Plants play a vital role in our environment and food substantially. Many plants are part of our food chain. The olive plant is one of the most significant plants, and we are achieving many benefits from it. The disease that affects these plants harms the quality of the olive and reduces its productivity, which also negatively impacts the economy. The Early detection and classification of these diseases is a big Challenge for the farmers. The early identification and targeted interventions can reduce crop management costs, minimize pesticide use, and prevent losses. A standard dataset called Zeytin_224x224_Augmented repository is utilized in this research. The dataset is openly available on Kaggle. Timely and accurate disease identification is essential for growers' income and olive production quality. The methodology utilized in this study for evaluating and classifying the effectiveness of InceptionV3 in detecting olive leaf disease and classification. Inception V3, deep learning enables early disease detection, allowing for timely treatment and ensuring the long-term sustainability of olive farming. We proposed a deep-learning model, Inception V3 that achieves the highest accuracy of 98.8%.

Keywords: Plants leaves disease detection; Olive leaves disease detection; classification; Deep Learning; Inception V3.

1. Introduction

Plant diseases are the major source of agricultural losses, making early detection important. Farmers use the best fertilizers and develop ideal growing conditions to reduce losses, minimize operating costs, and boost profits. Infections can arise in a variety of crop organs, and farmers frequently do visual testing. Early detection needs Automated approaches like computer vision and deep learning are essential for early detection, which meet the issues faced by farmers in big agricultural areas (Metre 2021). Many kinds of crops are traded on a global scale. Olive is one of the products with an increasing market (Chakraborty, Kishor et al. 2022). Saudi Arabia is the major producer of olives in the Southern Mediterranean region, ranking second in the country only to the European Union, with 6% of global production. Olive oil contributes to approximately 5.5% of all Saudi exports and half of all exported items (Jellali, Hachicha et al. 2021). The olive oil trade is the Kingdom of Saudi Arabia's fifth largest source of foreign cash, making it a critical industry for the country's economy. To improve the quality of olive oil and expand land, the government has spent heavily in restructuring and modernizing the olive sector (Ksibi, Ayadi et al. 2022). In Saudi Arabia, olive producers are dealing with significant illnesses such as olive leaf spot, Aculus

Olearius, and olive scab. These illnesses cause early loss of leaves and lower output in olive trees. The condition can be detected visually, but expert aid is required. The issue is lengthy, expensive, and frequently errors made by people. As the initial step in resolving the problem, a more quickly, programmed, and precise technique for detecting such conditions is required. Additionally, there is at present a need for an automatic, low-cost, and accurate technique for combating these diseases (Gupta and Nahar 2023).

In recent years, deep learning has been recognized as a disruptive force in agriculture, providing significant prospects for crop management and disease detection. Its outstanding image recognition and pattern analysis capabilities brought about a new area of precise farming. Deep learning models, such as the Inception V3 architecture employed in this study, can potentially transform how we monitor and control olive leaf diseases in the context of olive farming. Other models, including AlexNet, Google Net, Shuffle Net, and Squeeze Net, have shown considerable potential in this sector in addition to Inception V3(Balafas, Karantoumanis et al. 2023).

The model aids in early detection of diseases, pests, or nutrient deficiencies in olive leaves, enabling timely intervention. It also aids farmers in implementing precision agriculture techniques, efficiently applying resources like water, fertilizers, and pesticides. This early identification and targeted interventions can reduce crop management costs, minimize pesticide use, and prevent losses. The following are the particular objectives of this work:

- The first objective of this study is to create a reliable and accurate system for detecting and classifying diseases in olive leaves using deep learning techniques.
- The second objective is to categorize various diseases affecting olive leaves according to their features and symptoms using an image dataset.
- The third objective is to create a practical system that can help identify and diagnose diseases in olive leaves early on, thereby limiting crop losses and halting the spread of the diseases.

We not only improve the accuracy and efficiency of disease diagnosis by leveraging the power of this cutting-edge technology, but we also open the path for sustainable olive farming practices. As we go deeper into the heart of this research (Slimani, El Mhamdi et al. 2023), we examine the practical impact of various deep-learning techniques and methods in the world of olives, demonstrating the enormous potential for evolving agriculture and ensuring the future of this vital crop.

2. Literature Review

This study (Uğuz and Uysal 2021) gives a data set of 3400 olive leaf samples, including healthy leaves, to detect Aculus olearius and olive peacock spot illnesses, both of which are frequent in Turkish olive plants. Transfer learning methods were applied in this experimental work on VGG16 and VGG19 architectures and our suggested CNN architecture. One goal of this study was to examine the effects of data augmentation on performance. . In the experimental investigations utilizing data augmentation, the highest success value in trained models was 95%, whereas the highest value in the experiments without data augmentation was 88%. The effect of the Adam, AdaGrad, RMS Prop, and stochastic gradient descent optimization techniques on network performance is another area of interest for this research. Adeel et al. (Adeel, Khan et al. 2019) organized grape leaf photos sourced from the widely used Plant village dataset in their study. This dataset comprises a vast collection of plant photographs and includes four distinct class labels: black rot, black measles, leaf blight, and healthy. The researchers partitioned the dataset into training and test sets utilizing the K-fold cross-validation technique. The experimental results yielded accuracy scores ranging from 90% to 94% when employing a Support Vector Machine (SVM), K-Nearest Neighbor, and other machine learning algorithms derived from the previous techniques.

The primary objective of this study (El Akhal, Yahya et al. 2023) is to employ machine learning and deep learning methodologies to identify and classify illnesses that affect olive plants. The researchers gathered a total of 4138 photographs of olive leaves from various origins. They employed six pre-trained convolutional neural network architectures and five machine learning classifiers to construct a set of 30 distinct deep hybrid models (DHMs). The performance of the DHMs was assessed using established performance evaluation criteria and rigorous cross-validation processes. The highest performing DHM was achieved by the integration of the EfficientNetB0 model and a logistic regression classifier, yielding an exceptional accuracy rate of 96.14%. The utilization of this approach has significant promise in assisting olive growers in promptly and precisely detecting diseases, hence mitigating financial losses. This study highlights the efficacy of machine learning (ML) and deep learning (DL) methodologies in the identification and categorization of illnesses that pose a threat to olive trees.

This paper (Alshammari, Gasmi et al. 2022) introduces a novel approach to illness detection and classification in olive leaves, utilizing a sophisticated ensemble learning strategy that combines convolutional neural network and vision transformer models. The objective is to ascertain the optimal characteristic and ensure enhanced accuracy. As a result of the studies, the Adam and SGD optimization algorithms were found to produce superior outcomes (Dogo, Afolabi et al. 2018, Soydaner 2020). In the context of applications involving machine learning, researchers are required to do manual feature extraction on images using a range of machine learning approaches for addressing specific problems. The performance of the hybrid deep learning model, which incorporated three distinct models together with a modified vision transformer model, surpassed that of alternative models in both binary and multiclass classification tasks. The highest-performing model, which integrates the ViT model, and the VGG-16 model, attained a binary classification accuracy of 97% (Rezaei, Rahmani et al. 2023, Suha and Islam 2023). The researchers intend to modify this approach for application to more plant collections, while expanding their collection of photographs depicting olive illnesses. The findings of the study provide evidence of the efficacy of the deep-learning methodology in the identification and categorization of olive leaf diseases. The highest success value in trained models in the experimental investigations that used data augmentation was 95%, while the highest value in the experiments that did not use data augmentation was 88%. Another focus of this study is the impact of the Adam, AdaGrad, Stochastic gradient descent, and RMS Prop optimization techniques on network performance (Lachgar, Hrimech et al. 2022). The literature review summary is presented in Table 1.

Table 1. Literature Review Summary

References	Year	Plant	Dataset	Technique	Accuracy
El Akhal et al. (El Akhal, Yahya et al. 2023)	2023	Olive	4138 Olive images	Efficient Net	96.14
Ksibi et al. (Ksibi, Ayadi et al. 2022)	2022	Olive	5400 olive leaf images were collected from Olive groves from vehicle	CNN+DFC+Mobile Resnet	97.08
Alshammr Et al. (Alshammari,	2022	Olive	Olive images	VGG-16, VGG-19, Vision	97

Gasmi et al. 2022)				Transform	
Uğuz et al. (Uğuz and Uysal 2021)	2021	Olive	3400 olive leaf images were collected from Denzil city of Turkey	CNN, VGG16, VGG19	95
Adeel et al. (Adeel, Khan et al. 2019)	2019	Plant disease	Plant Village	KNN, SVM	92

3. Proposed Model

This section describes in detail the approach used in this study to evaluate and assess the effectiveness of InceptionV3 in identifying and classifying olive leaf disease. This section contains thorough details on the picture dataset creation and preprocessing procedure, as well as the settings and hyper parameter values used for the validation, testing, and training phases of the InceptionV3 architecture.

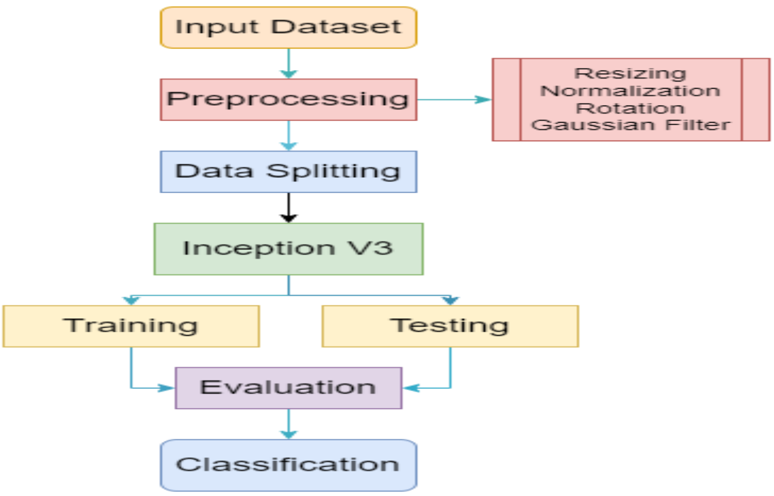


Figure 1. Research methodology of COVID-19 sentiments analysis

3.1 Dataset

We check the detection and classification ability of our framework. A standard dataset called Zeytin_224x224_Augmented repository is utilized in this work. The dataset is openly available on Kaggle. Olive leaf photos were used to train and test the model. This dataset includes a total of 6,961 images that contain healthy, Aculus olearius-affected, and olive peacock spot-affected leaves. The dataset is split into 70% for training and 30% for testing.

Figure 2. Sample images from dataset

3.2 Preprocessing of Dataset

We performed various important steps during preprocessing to prepare the dataset before input to the model. Every image was first scaled to 165x165 pixels. This scaling ensured image standardization and reduced model training, which is the computational cost of model training. Normalizing image intensities to a common scale is essential for model convergence. Normalization prevents large or very small data values, allowing the model to learn and adapt..

Additionally, a Gaussian filter was added to increase image quality and uniformity. This filter reduced image noise and enhanced key features. Image-based classification tasks require noise reduction to focus the model on olive leaf disease patterns and features. Data augmentation was used to improve the model's robustness and dataset generalization. Random rotations within a range were included. This method subjected the model to leaf orientation variations, a common real-world phenomenon. Rotated images helped the model spot diseases from different angles, enhancing its performance. These preprocessing techniques primed the dataset, allowing the deep learning network to successfully detect and classify olive leaf diseases detect and classify olive leaf diseases successfully. This rigorous data preparation was vital to the research's success and the model's ability to contribute to agriculture and disease control.

Data normalization and image resizing are crucial preprocessing operations in computer vision tasks. Standardizing input image size and aspect ratio facilitates processing and increases comparability. Inception v3 uses 165x165 pixels for input images, reducing computational costs and improving speed and efficiency. Data normalization modifies input data values, facilitating faster convergence and generalization to novel data and improving performance and accuracy.

3.3 Data splitting

Testing and training datasets to partition the dataset, 30% is used for testing and 70% for training. model. We do this to maximize availability. Training data would produce an accurate model. The dataset is composed of three parts: the train set, test set, and the validation set. Three categories have been established for the dataset: 20% for testing, 10% for validation, and 70% for training. The learning process of deep convolution neural networks depends on an optimizer. Before splitting, randomize the whole dataset needed.

3.4 Proposed Model

Our proposed model Inception v3, exhibits performance that is at the forefront of the field as shown in Fig 3. To assess performance on unobserved data and prevent overfitting, the model is evaluated on the validation and testing sets after being implemented with Kera's, a high-level deep learning library. The traditional 7 x 7 convolution has been factorized into three 3 x 3 convolutions. The Inception part of the network has three traditional modules with 288 filters each, which are reduced to a 17 x 17 grid with 768 filters using grid reduction techniques. The factorized modules are then reduced to an 8 x 8 x 1280 grid, with two Inception modules at the coarsest 8 x 8 level. Hyperparameters including batch size, number of epochs, and number of classes are established throughout the training process. In this instance, sample sizes of 25 are utilized, and the model is trained with three classes for epochs 10. Our model achieves optimal performance with a group size of 25 and 10 epochs. A validation set evaluation prevents overfitting during training by monitoring the performance of the model. Following the conclusion of the training phase, the model is assessed on the testing set to determine its performance on data that it has not been exposed to thus far. The Inception network comprises multiple modules with four parallel operations: 1x1 conv layer, 3x3conv layer, and 5x5conv layer, with max pooling.

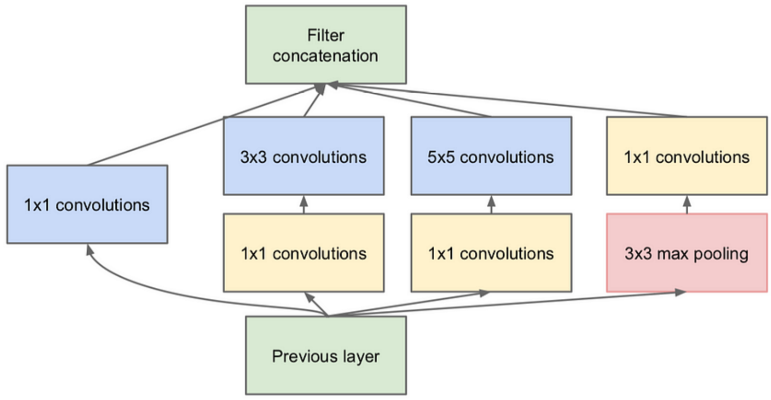


Figure 3. Inception V3 Model

The Adam optimizer was used to achieve optimal accuracy in batch sizes of 25, 32, 10, 20, 50, and 60, with a learning rate of 0.001 and input size of 165x165x3. The model achieved the highest training accuracy of 98.950 and the highest validation accuracy of 98.8, demonstrating good generalization without overfitting.

The convolution layer transforms the input into a feature map, which is subsequently transmitted to the activation layer. The feature map is sorted by shape number, height, width, and channels. Intending to reduce the quantity of network statistics and parameters, the pooling layer reduces the magnitude of the representation area. The average pooling rate is the process by which the mean value of each element in a distinct map is aggregated publicly. The aim of this procedure iis procedure aims to sample the in rectangular circuits by dividing the input into regions and calculating the average value for each.

4. Performance Evaluation Metrics

This section presents a discussion of the performance that the proposed work has achieved. The experimental configuration utilized in this investigation is described initially. Following this, we establish our performance metrics. The evaluation of each class's accuracy, precision, recall, and F1 score was calculated by Equations 1,2,3, and 4 Respectively.

$$\text{Accuracy}=(TP+TN)/(TP+TN+FP+FN) \tag{1}$$

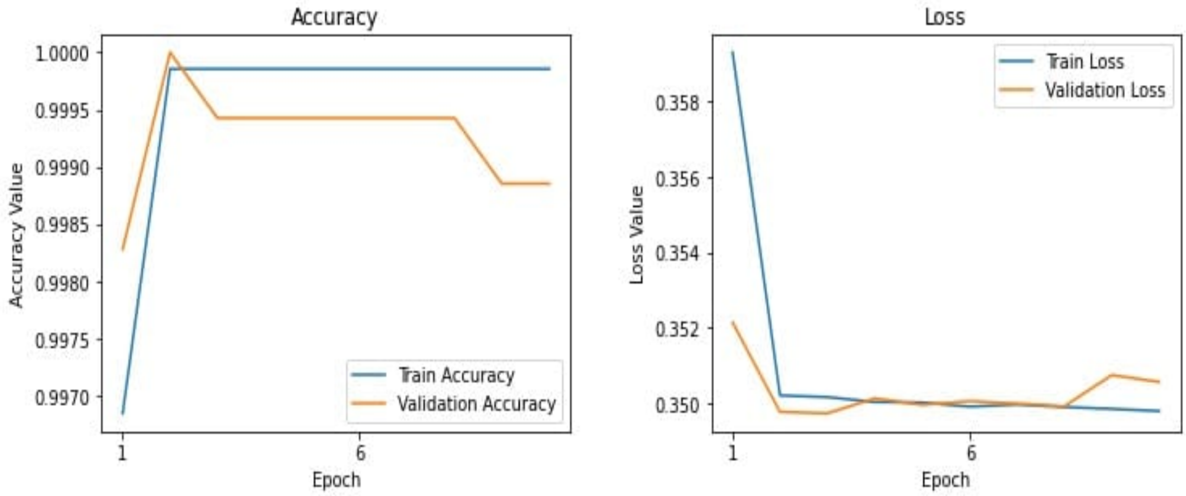
$$\text{Percision} =TP/(TP+FP) \tag{2}$$

$$\text{Recall}=TP/(TP+FN) \tag{3}$$

$$\text{F1 Score}=(2\times(\text{Percision}\times\text{recall}))/(\text{percision}+\text{recall}) \tag{4}$$

We used the Adam optimizer with a batch size of 25 to achieve maximum accuracy and trained for 10 epochs. To be compatible with our Inception V3 model, we resized images to 165x165 pixels with a depth of 3. We strategically froze some layers while unfreezing the most important layers, namely 'conv2d_367', 'conv2d_369', 'conv2d_375', and 'conv2d_86', using a fine-tuned pre-trained model. We used the early-stopping technique selectively to prevent overfitting.

Our proposed model performed brilliantly throughout the training process, achieving an outstanding maximum training accuracy of 97.75%. The model's validation accuracy was also remarkable, reaching 88.6039%. The test score and test accuracy are the final measures of our model's efficiency in olive leaves image classification. Our model achieved a remarkable test score of 0.9619700908660889 and a remarkable peak test accuracy of 98.8%, which is presented in the accuracy vs loss graph which is shown in Fig. 4. This remarkable accuracy highlights the model's ability to accurately classify various types of diseases accurately.



.

Figure 4. Accuracy and Loss graph

The Fig 5. shows the evaluation matrices in which the classification report of the classification report of proposed work provides accuracy, precision, recall, and F1-Score for all classes. A thorough examination of the classification results, as provided in the classification report, indicates that our model showed exceptional precision, recall, and F1 scores across all disease classes. Particularly, the healthy and infected classes were the easiest to distinguish, as evidenced by their high precision, recall, and F1 scores. This performance shows the model's ability to identify and classify distinct diseases with high accuracy and accurately identify and classify them.

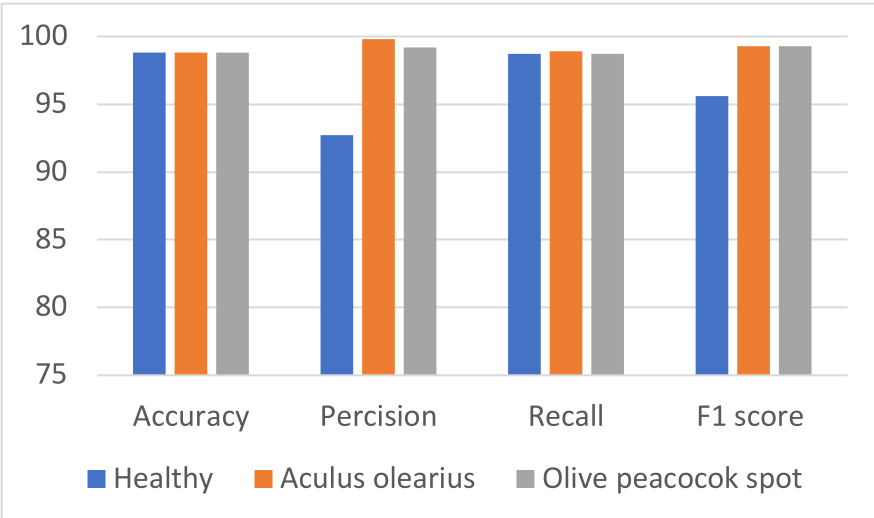


Figure 5. Evaluation Metrices

Our model ROC AUC Scores were exceptional, with a macro score of 0.999792 and a prevalence-weighted score of 0.999833, indicating that the model excelled in classifying the various classes. Similarly, the One-vs-Rest ROC AUC Scores received excellent results, with a macro score of 0.999792 and a prevalence-weighted score of 0.999833, demonstrating the model's top-tier performance in the classification of the classes. The ROC curves for all three classes exhibited a perfect area of 1.00, as illustrated in Fig 6, signifying that the model achieved an ideal balance between sensitivity (true positive rate) and specificity (true negative rate) in classifying the various classes. Hence, the InceptionV3 model proved to be a dependable model dependable for classifying images.

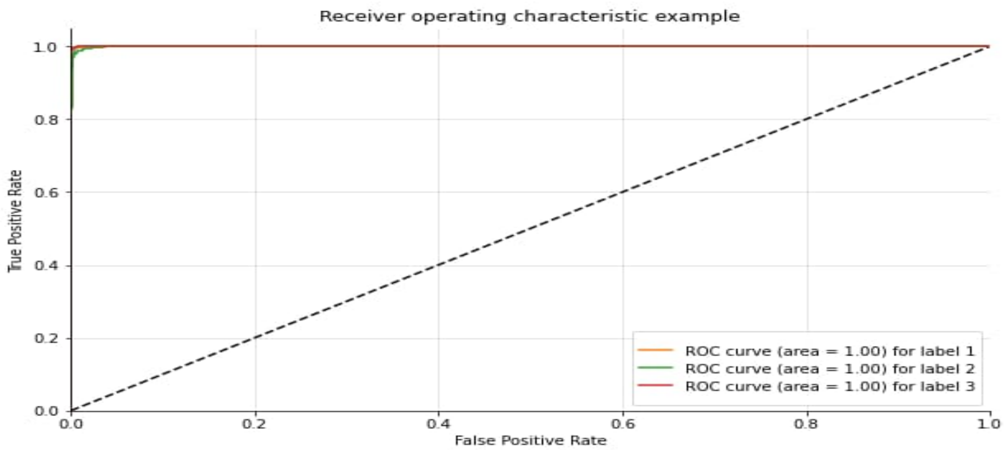


Figure 6. ROC Curve

The evaluation of the InceptionV3 model using the confusion matrix revealed that the model achieved a high level of accuracy in classifying most of the test data. The "Olive peacock spot " class had the highest number of correct predictions, with a value of 734, followed by the "Healthy " class, with a value of 416. However, the model struggled to predict the "Aculus olearius" class, with a value of 340, as shown in Fig. 7.

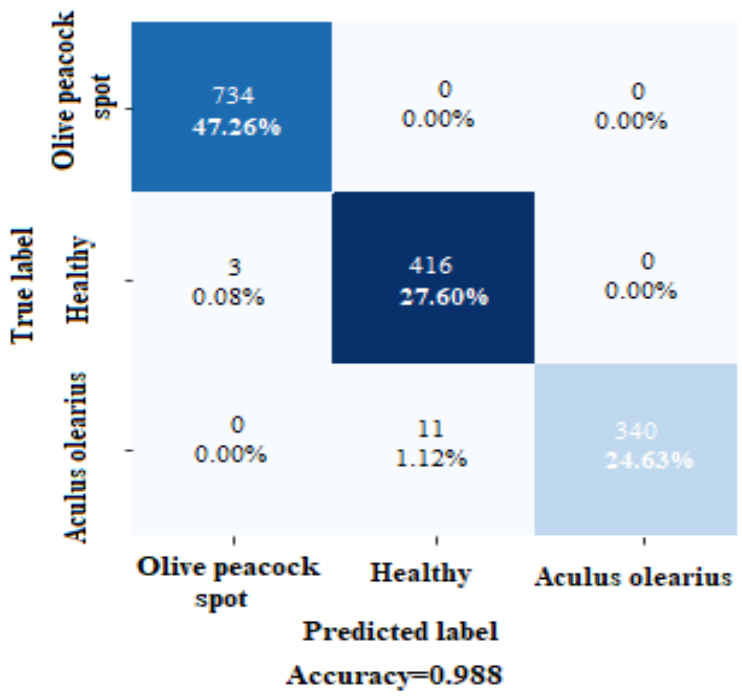


Figure 7. Confusion Matrix

Table 2. Comparison with Existing Techniques

References	Year	Plant	Technique	Accuracy
El Akhalet al. (El Akhal, Yahya et al. 2023)	2023	Olive	Efficient Net	96.14
Ksibi et al. (Ksibi, Ayadi et al. 2022)	2022	Olive	CNN + DFC + Mobile Resnet	97.08

Alshammr Et al. (Alshammari, Gasmi et al. 2022)	2022	Olive	VGG-16, VGG-19, Vision Transform	97
Uğuz et al. (Uğuz and Uysal 2021)	2021	Olive	CNN, VGG16, VGG19	95
Adeel et al. (Adeel, Khan et al. 2019)	2019	Plant disease	KNN, SVM	92
Proposed Model	2023	Olive	Inception V3	98.8

Table 02 supports the abstract claims the proposed model achieves the highest accuracy of 98.8%. It presents enough comparative analysis with existing methods or models. It contains a comparison to demonstrate the superiority of the Inception V3 model in detecting and classifying olive leaf diseases.

5. Conclusion

This research investigates that specifically the InceptionV3 model can detect and classify olive leaf diseases with 98.85% accuracy across three classes: healthy leaves, Aculeus olearius-affected leaves, and olive peacock spot-affected leaves, Aculeus olearius-affected, and olive peacock spot-affected. This achievement shows the transformative effects of advanced deep-learning techniques in agriculture. These fast and efficient disease detection capabilities impact olive growers' income and olive production quality. Deep learning identifies diseases early, enabling quick remediation and ensuring the success of olive cultivation. This research advances plant disease detection and classification by promoting the adoption of cutting-edge agricultural technology. In the future, the integration of drone technology equipped with cameras holds promise for real-time disease detection, as images captured in the field can be processed by the classifier, enabling rapid and precise disease identification and treatment.

Acknowldgment: I owe my appreciation to my teacher Dr Muhammad Munwar Iqbal, who made this work attainable and whose contribution to my graduate experience I will cherish to the end.

References

1. Adeel, A., et al. (2019). "Diagnosis and recognition of grape leaf diseases: An automated system based on a novel saliency approach and canonical correlation analysis based multiple features fusion." Sustainable Computing: Informatics and Systems 24: 100349.
2. Alshammari, H., et al. (2022). "Olive disease classification based on vision transformer and CNN models." Computational Intelligence and Neuroscience 2022.
3. Balafas, V., et al. (2023). "Machine learning and deep learning for plant disease classification and detection." IEEE Access.
4. Chakraborty, C., et al. (2022). "Novel Enhanced-Grey Wolf Optimization hybrid machine learning technique for biomedical data computation." Computers and Electrical Engineering 99: 107778.
5. Dogo, E. M., et al. (2018). A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. 2018 international conference on computational techniques, electronics and mechanical systems (CTEMS), IEEE.
6. El Akhal, H., et al. (2023). "A novel approach for image-based olive leaf diseases classification using a deep hybrid model." Ecological Informatics 77: 102276.
7. Gupta, A. and P. Nahar (2023). "Classification and yield prediction in smart agriculture system using IoT." Journal of Ambient Intelligence and Humanized Computing 14(8): 10235-10244.
8. Jellali, A., et al. (2021). Sustainable Configuration of the Tunisian Olive Oil Supply Chain Using a Fuzzy TOPSIS-Based Approach. Sustainability 2021, 13, 722, s.
9. Ksibi, A., et al. (2022). "MobiRes-net: a hybrid deep learning model for detecting and classifying olive leaf diseases." Applied Sciences 12(20): 10278.
10. Metre, V. A. (2021). "Research review on plant leaf disease detection utilizing swarm intelligence." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12(10): 177-185.
11. Rezaei, M., et al. (2023). "Role of artificial intelligence in the diagnosis and treatment of diseases." Kindle 3(1): 1-160.
12. Slimani, H., et al. (2023). "Drone-Assisted Plant Disease Identification Using Artificial Intelligence: A Critical Review." International Journal of Computing and Digital Systems 14(1): 10433-10446.
13. Soydaner, D. (2020). "A comparison of optimization algorithms for deep learning." International Journal of Pattern Recognition and Artificial Intelligence 34(13): 2052013.
14. Suha, S. A. and M. N. Islam (2023). "Exploring the dominant features and data-driven detection of polycystic ovary syndrome through modified stacking ensemble machine learning technique." Heliyon 9(3).
15. Lachgar, M., et al. (2022). "Optimization techniques in deep convolutional neuronal networks applied to olive diseases classification." Artificial Intelligence in Agriculture 6: 77-89.
16. Uğuz, S. and N. Uysal (2021). "Classification of olive leaf diseases using deep convolutional neural networks." Neural Computing and Applications 33(9): 4133-4149.