International Journal of Advanced Trends in Engineering and Management
(ISSN 2583-7052)

Vol. V, No. 03, pp. 20-26, March

Automated Olive Leaf Disease Identification using K-Means segmentation and VGG-19 –based Classification

J N Rajesh Kumar[3]

[3]Assistant Professor, Department of Computer Science and Engineering,
Sree Sastha Institute of Engineering and Technology, Chembarrambakkam,
Chennai, Tamil Nadu, India.

rajeshjnin@yahoo.com

ABSTRACT: Early and accurate detection of Olive Leaf Diseases (OLD) is essential for effective agricultural management and significantly affect crop productivity and quality. This paper proposed a VGG-19-based classification a Deep Learning (DL) model, for automated OLD identification. Preprocessing based on image resizing and Gaussian Filtering (GF) to enhance image quality and reduce noise. Segmentation based on K-Means segmentation is used to accurately isolate the diseased regions from the leaf background. Feature extraction is then performed by the Histogram of Oriented Gradients (HOG) technique to capture discriminative texture and shape characteristics associated with different disease patterns. VGG-19 is proposed which effectively classifies the complex visual representations for disease recognition. Experimental results demonstrate that the proposed framework attains high classification accuracy of 94.8% and precision, recall and F1-score of 95% and robust performance in identifying various OLD. The novelty of the paper enhances disease recognition accuracy while reducing the influence of background noise.

Keywords: Olive Leaf Diseases VGG-19, Gaussian Filtering, K-Means, Histogram of Oriented Gradients

1. Introduction

Olive cultivation is an important agricultural activity in many countries, contributing significantly to economic growth and food production [1]. Consequently, Saudi Arabia is home to some of the world's highest-quality olive plantations, covering approximately 6% of the global olive cultivation area of 10.6 million hectares [2]. Furthermore, the identification of most olive cultivars is primarily based on the characteristics of their fruits and endocarps [3]. However, it negatively affects both the quality and quantity of agricultural production [4]. Moreover, OliVaR classifies olive varieties using the morphological features of olive endocarps [5], Olive plant diseases are often identified through visible changes in leaf color and appearance [6]. To overcome this issue several DL techniques are utilized for OLD and CLD [7].

In [8] Ksibi et al (2022) have proposed MobiRes-Net model technique for OLD detection. Classify multiple OLD even under varying lighting and environmental conditions. However, accurate disease recognition requires a substantial number of labeled OL images. Furthermore, in [9] Alshammari et al (2023) have presented Artificial Neural Network (ANN) to analyses the OLD as leaf features and disease classes. Nevertheless, perform poorly on unseen samples if not properly regularized. Moreover, in [10] Huaquipaco et al (2024) have invented Xception and VGG16 to detect peacock spot in OLD. Nonetheless, performance degrades when trained on imbalanced datasets. Furthermore, in [11] Pacal et al (2025) have developed MetaFormer to detect the OLD. Perform well across diverse image conditions and disease categories. Nevertheless, it is difficult to deploy on edge devices agricultural monitoring systems. Moreover, in [12] Gunduz et al (2026) have proposed ResNet101 with LGBM for OLD detection. Identify subtle disease symptoms such as spots, lesions, and discoloration. However, performs best when trained on a large OL images.

1.1 Problem statement

Existing OLD detection methods suffers from sensitivity to image noise, variations in illumination, complex leaf backgrounds, and the inability to effectively capture discriminative disease characteristics. To overcome this issue this paper proposes a VGG-19 for OLD detection.

1.2 Research Contribution

- Preprocessing based on image resizing and GF is used to improve image quality, suppress noise, segmentation based on K-Means segmentation is to isolate diseased regions from the leaf background.
- Feature Extraction based on HOG is utilized to extract texture and shape features associated with different OLD
- Proposed VGG-19 classification effectively learns complex visual patterns and accurately distinguishes among various olive leaf disease categories.

2. Proposed Methodology

Figure 1 displays the proposed architecture for OLD detection.

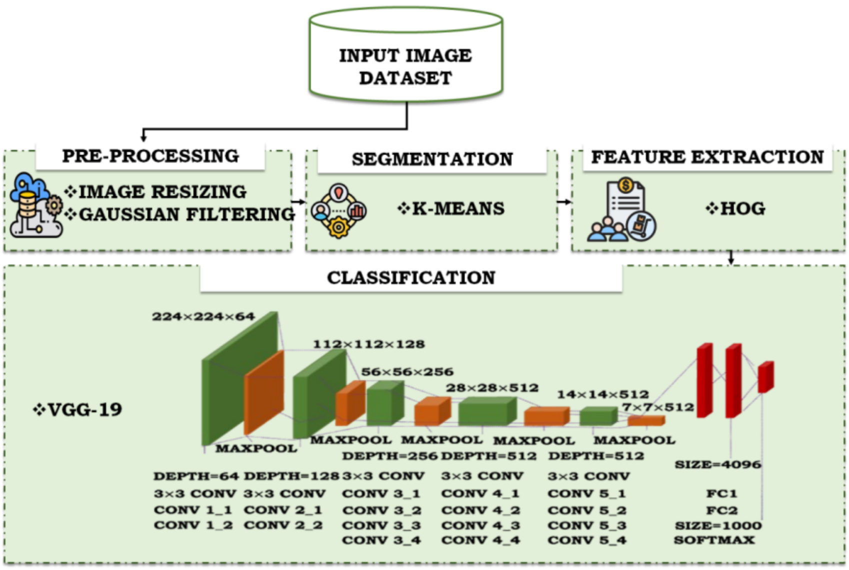

Figure 1: *Proposed Architecture*

The input dataset is fed to pre-processing based on image resizing standardizes the input dimensions and GF removes noise. Next, K-Means segmentation is applied to accurately separate the diseased regions from the leaf background. The segmented image is fed to feature extraction stage to extract the feature using HOG which is used to extract discriminative texture and shape features that effectively represent disease characteristics. The VGG-19-based classifier, which learns complex visual patterns and classifies the OL images into their respective disease categories.

2.1 Pre-Processing for OLD detection

For OLD detection the input dataset is resized using image resizing technique that standardizes the input dimension. GF preserving important disease features such as spots, lesions, and texture patterns

$$G(x,y) = \frac{1}{2\pi\sigma^2}\exp(-\frac{x^2+y^2}{2\sigma^2}) \quad (1)$$

Where, $G(x,y)$ denotes Gaussian kernel value at position (x,y), σ represents standard deviation controlling the degree of smoothing x,y signifies spatial coordinates relative to the kernel center, e indicates Euler's constant. Following to this the proceed image fed to segmentation stage.

2.2 Segmentation by K-Means

For OLD detection the K-Means segmentation is to segment the diseased region from the background. A data point distance from these cluster centers is compared, and it is part of the closest cluster.

$$D(x,y) = \|x-y\|^2 \quad (2)$$

$$l_k(x_k) = argmin_i D(x_k - C_i) = argmin_i \|x_k - C_i\|^2 \quad (3)$$

Where $D(x,y)$ represents data points, l_k denotes data pointx_k label. Following to this the segmented image is fed to feature extraction.

2.3 Feature Extraction by HOG

For OLD detection the HOG feature is used to extract the diseased characterises and it is a type of descriptor that detects object attributes. These features capture local structural variations and spatial relationships that are critical for distinguishing the diseased region.

$$I(x,y) = I(x,y)^{gamma} \quad (4)$$

Here, $I(x,y)$ takes value based on the effect. $Gamma$ denotes the feature. Feature vector by combining all of the components at last to find OLD. Next, the featured OL image is given as the input to the proposed VGG-19.

2.4 Classification by proposed VGG-19

In OLD detection, the proposed VGG-19 automatically extracts discriminative features from leaf images and classifies them into healthy or diseased categories. VGG-19 contains 19 layers in which 16 convolutional layers and 3 fully connected layers. Input layer receives the pre-processed OL image, convolutional layer remove low-level then high-level features like spots, lesions, discoloration, max-pooling reduce feature map dimensions, fully connected integrate the extracted features for classification, softmax layer produces the probability of each disease class and output layer predicts the OLD category. Table 1 shows the hyperparameter for VGG-19. Fig.2 shows the VGG-19 architecture.

Table 1: *Hyper parameter for VGG-19*

Hyperparameter	Values
Optimizer	Adam
Batch size	32
Hidden layer, output layer activation	ReLU, Softmax
Training time	58-83min
Epochs	100
Dropout rate	0.5

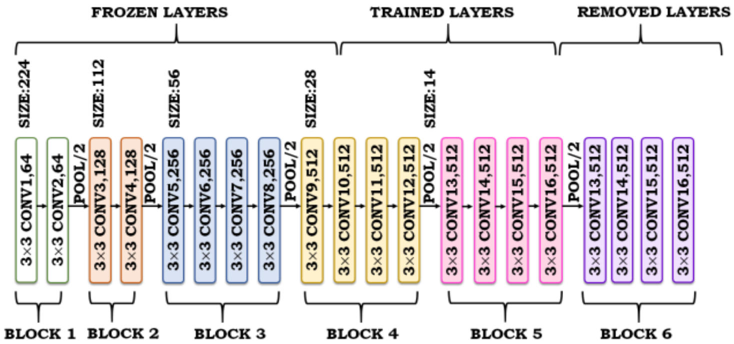

Figure 2: *VGG-19 architecture*

Each layer classify its OL by its category and it is mathematically expressed as

$$F(i,j) = \sum_m \sum_n I(i-m, j-n).K(m,n) \quad (5)$$

Where, I denotes input image, K signifies convolution kernel and Frepresents feature map. ReLU activation function. The max-pooling operation is

$$P(i,j) = {}_{(m,n)\in R}^{max} F(m,n) \quad (6)$$

Where R signifies the pooling region, $P(yi)$ denotes probability of class, j . Thus by this the proposed VGG-19 classify the OLD clearly.

3. Result and Discussion

For OLD the proposed VGG-19 is implemented in python software to classify the disease. Using Olive leaf image dataset the work is done and the input images have the size of 224,224, and the total number of train set is 2720, test set is 680 and total is 3400 images the class names are aculus_olearius have 690 train set, 200 test set and total 890 images, Healthy have 830 train set, 220 test set and total 1050 images, olive_peacock_spot have 1200 train set, 260 test set and total 14600 images.

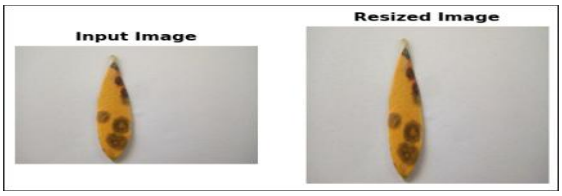

Figure 3: *Input to resize image*

Figure 3 displays the input to resize image. The input image is chosen from olive leaf image dataset here the input image is resized using image resizing technique.

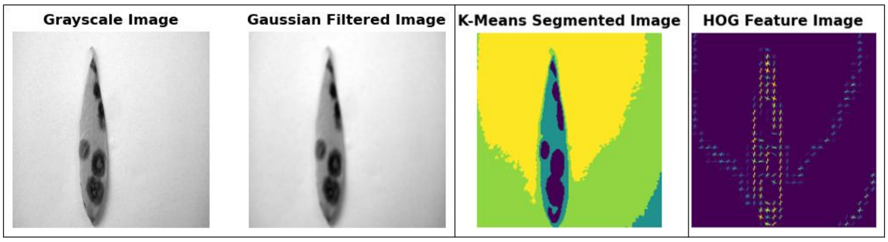

Figure 4: *Grayscale, GF image, K-means, HOG*

Figure 4 shows the grayscale, GF image, K-means, HOG. The resized image to grayscale here the color image is changed into a single-channel intensity image. The noises are removed using GF image. The K-Means segmented image separates the leaf and diseased regions into distinct clusters. In HOG the edge orientations and texture patterns of the leaf lesions, extracting discriminative features that the HOG feature vector length is 26244.

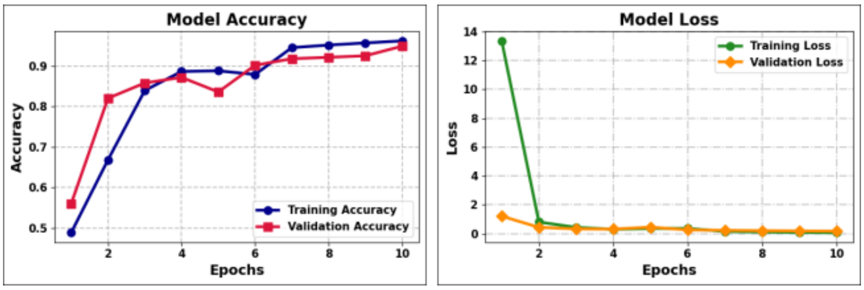

Figure 5: *Model accuracy and loss*

Figure 5 displays model accuracy and loss. The validation accuracy improves from 0.56 to 0.95, effective model learning and strong generalization. The validation loss drops from 1.15 to 0.16, improved prediction performance with minimal overfitting.

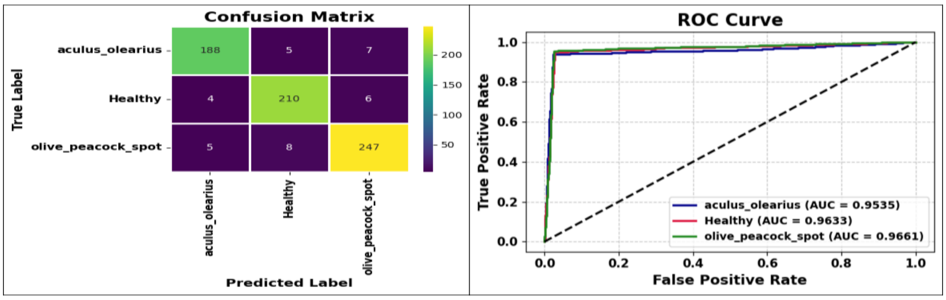

Figure 6: *Confusion matrix and ROC curve for proposed VGG-19*

Figure 6 displays the confusion matrix for proposed VGG-19 for OLD detection. Here the classes aculus_olearius of 188 with precision, F1-score of 95%, recall of 94%,healthy of 210 with precision of 96%, recall, F1-score of 95%, and olive_peacock_spot of 247 with precision, F1-score of 94% and recall of 95%,. The AUC value for aculus_olearius is 0.95, healthy and olive_peacock_spot is 0.96 correspondingly.

Ablation study

In the ablation study the proposed VGG-19 is compared with the existing techniques as shown in table 2.

Table 2: *Comparison for performance metrics*

Study	Accuracy	Precision	Recall	F1-Score
YOLOv8 [1]	90.3%	84%	77%	82%
ResNet101 [2]	91.8%	73.08%	93.7%	94.9%
MobileNet [8]	94.6%	93%	93%	94%
Proposed VGG-19	**94.8%**	**95%**	**95%**	**95%**

4. Conclusion

In this paper OLD detection is detected using proposed VGG-19 for accurate and reliable disease classification. The OLD is resized and removed noised using image resizing and GF, K-Means for precisely isolated the diseased regions, and HOG extract the OLD image. The proposed VGG-19 architecture identifies different OLD. Experimental results demonstrate the effectiveness of the proposed framework, achieving a classification accuracy of 94.8% with a precision, recall, and F1-score of 95%. Future work focus on integrating advanced attention mechanisms, lightweight DL architectures, and real-time deployment on edge devices to further improve detection accuracy and practical applicability in smart agricultural environments

References

1. Konstantinos Prousalidis; Stavroula Bourou; Terpsichori-Helen Velivassaki; Artemis Voulkidis; Aikaterini Zachariadi; Vassilios Zachariadis, Year: 2024, “Olive tree segmentation from UAV imagery”, Drones, Vol: 8, No: 8, pp. 408.
2. Antonio Fazari; Oscar J. Pellicer-Valero; Juan Gómez-Sanchıs; Bruno Bernardi; Sergio Cubero; Souraya Benalia, Giuseppe Zimbalatti, and Jose Blasco, Year: 2021, “Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS/NIR hyperspectral images”, Computers and Electronics in Agriculture, Vol: 187, pp. 106252.
3. Konstantinos N. Blazakis; Danil Stupichev; Maria Kosma; Mohamad Ali Hassan El Chami; Anastasia Apodiakou; George Kostelenos; Panagiotis Kalaitzis, Year: 2024, “Discrimination of 14 olive cultivars using morphological analysis and machine learning algorithms”, Frontiers in Plant Science, Vol: 15, pp. 1441737.
4. Mohamed Lachgar; Hamid Hrimech; Ali Kartit, Year: 2022, “Optimization techniques in deep convolutional neuronal networks applied to olive diseases classification”, Artificial Intelligence in Agriculture, Vol: 6, pp. 77-89.
5. Hristofor Miho; Giulio Pagnotta; Dorjan Hitaj; Fabio De Gaspari; Luigi Vincenzo Mancini; Georgios Koubouris; Gianluca Godino; Mehmet Hakan; Concepción Muñoz Diez, Year: 2024, “OliVaR: Improving olive variety recognition using deep neural networks”, Computers and Electronics in Agriculture, Vol: 216, pp. 108530.
6. Ali Hakem Alsaeedi; Ali Mohsin Al-Juboori; Haider Hameed R; Al-Mahmood; Suha Mohammed Hadi; Husam Jasim Mohammed; Mohammad R. Aziz; Mayas Aljibawi; Riyadh Rahef Nuiaa, Year: 2023, “Dynamic clustering strategies boosting

deep learning in olive leaf disease diagnosis", Sustainability, Vol: 15, No: 18, pp. 13723.

7. Sivakumar Rajendran, Year: 2025, "Enhancing Cotton Leaf Disease Detection using Deep Learning Approach: A Study and Comparison", In 2025 International Conference on Electronics and Renewable Systems (ICEARS), pp. 1450-1454. IEEE.
8. Amel Ksibi; Manel Ayadi; Ben Othman Soufiene; Mona M. Jamjoom; Zahid Ullah, Year: 2022, "MobiRes-net: a hybrid deep learning model for detecting and classifying olive leaf diseases", Applied Sciences, Vol. 12, No. 20, pp. 10278.
9. Hamoud H.Alshammari; Ahmed I. Taloba; Osama R. Shahin, Year: 2023, "Identification of olive leaf disease through optimized deep learning approach", Alexandria Engineering Journal, Vol. 72, pp. 213-224.
10. Saul Huaquipaco; Oscar Vera; Victor Yana-Mamani; Wilson Mamani; Helarf Calsina; Flavio Puma; Eli Morales-Rojas; Norman Beltran; Jose Cruz, Year: 2024, "Peacock spot detection in Olive leaves using self supervised learning in an assembly Meta-Architecture", IEEE Access, Vol: 12, pp. 192828-192839.
11. Ishak Pacal; Serhat Kilicarslan; Burhanettin Ozdemir; Muhammet Deveci; Seifedine Kadry, Year: 2025, "Efficient and autonomous detection of olive leaf diseases using AI-enhanced MetaFormer", Artificial Intelligence Review, Vol: 58, No: 10, pp. 303.
12. Hakan Gunduz, Year: 2026, "Efficient Olive Leaf Disease Detection Using Composite Feature Selection and Ensemble Learning", Agronomy, Vol: 16, No: 11, pp. 1057.