FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 9, September, 2025, pp 207 – 217
DOI: https://doi.org/10.33003/fjs-2025-0909-4032

COMPARATIVE ANALYSIS OF BINARY AND MULTICLASS POTATO LEAF DISEASE CLASSIFICATION USING VGG19 MODEL

*[1]Evwiekpaefe Abraham Eseoghene, [1]Chinyio Darius Tienhua and [2]Nwanna Ndubuisi Peter

[1]Department of Computer Science, Nigerian Defence Academy, Kaduna, Kaduna State, Nigeria.
[2]Directorate of Information and Communication Technology, Nigerian Defence Academy, Kaduna, Kaduna State, Nigeria.

*Corresponding authors' email: aeevwiekpaefe@nda.edu.ng

ABSTRACT

Agriculture in Nigeria has been the source of livelihood yielding sustainable development across the country. However, potato farming in Nigeria faces numerous challenges such as unknown diseases and challenges in potato leaf disease classification. This study discovered a problem in potato leaf disease classification using VGG19 model in which binary class of potato leaf (potato early blight and potato late blight diseases) was not enough for dataset generalization. Therefore, this study aimed to conduct a comparative analysis of binary and multiclass potato leaf disease classification using VGG19 model. The research used comparative analysis tools to compare the result of the binary class (early blight and late blight leaves) and multiclass (early blight, late blight, virus disease and healthy potato leaves) in which VGG19 model with binary class obtained at epoch 40, training accuracy of 93.25% and loss of 0.2513, validation accuracy of 90.00% and loss of 0.2970 and testing accuracy of 91.67% and loss of 0.2735 and VGG19 model with multiclass obtained at epoch 40, training accuracy of 91.28% and loss of 0.3794, validation accuracy of 87.50% and loss of 0.3893 and testing accuracy of 91.67% and loss of 0.4956. The result showed that the higher the number of data classes in VGG19 model, the lower the training accuracy in VGG19 model. Finally, this work has achieved its aim and objective; and it can be evaluated for future study.

Keywords: Binary Class, Multiclass, Potato leaf disease, VGG19 model

INTRODUCTION

The agro-ecological zones of Nigeria is endowed with diverse and a rich agricultural heritage, which stands at the threshold of a transformative journey in the realm of agriculture. In recent years, there has been a growing realization of the pivotal role that plant classification can play in improving agricultural practices and ensuring food security (Sibhatu and Qaim, 2024).

Potato, whose botanical name is Solanum tuberosum, also an annual plant in the nightshade family (Solanaceae), grown for its starchy edible tubers that are highly digestible and supply vitamin C, protein, thiamin, and niacin (Ghosh et al., 2023). Potato is one of the most important staple crops globally, providing a significant source of food and income for millions of people. However, potato plants are vulnerable to various diseases that can severely affect crop yield and quality. These diseases often manifest on the leaves such as potato early blight, potato late blight, potato mosaic virus and making early and accurate detection are crucial for effective disease management and prevention of crop losses.

Nevertheless, potato leaf disease classification involves identifying and categorizing different types of diseases that affect potato leaf. Traditionally, this process relies heavily on expert knowledge, where trained agronomists would visually inspect the plants for symptoms. However, this method is time-consuming, prone to human error, and not scalable for large-scale farming (Rashid et al., 2021). With advancements in technology, especially in the fields of computer vision and machine learning, automated systems for disease classification have become more feasible. Machine learning models, particularly those based on deep learning, have shown great promise in this area. Convolutional Neural Networks (CNNs), for instance, can learn to recognize patterns and features in leaf images that are indicative of specific diseases. Moreover, Convolutional Neural Networks have demonstrated exceptional capability in recognizing complex patterns and features in leaf images that correspond to specific diseases (Singh and Yogi, 2023). By training these models on large datasets of labeled images, they can achieve high levels of accuracy in disease classification, often matching or exceeding the performance of human experts which ultimately improve crop yield and sustainability thereby making leaf disease classification an essential task in agricultural management.

Ghosh et al. (2023) was evaluation and contrast of three state-of-art CNN models: VGG19, DenseNet121 and ResNet50 to identify and forecast two potato leaf classes (potato late blight and potato early blight). Their study's result showed that VGG19 emerged as the top performer with an accuracy of 92.71% followed closely by DenseNet121 and ResNet50. However, this study discovered a problem in potato leaf disease classification using VGG19 model in which binary class of potato leaf (potato early blight and potato late blight diseases) was not enough for dataset generalization.

The aim of this study is to conduct a comparative analysis of binary and multiclass potato leaf disease classification using VGG19 model. The objective of this study is used comparative analysis tools to compare the result of the binary class (early blight and late blight leaves) and multiclass (early blight, late blight, virus disease and healthy potato leaves).

Motivation of this study arise from Ghosh et al. (2023)'s work in which binary class of potato leaf (potato early blight and potato late blight diseases) was not enough for dataset generalization. And, this problem made this study to conduct a comparative analysis of binary and multiclass potato leaf disease classification using VGG19 model. More so, this study is motivated by the available methodologies required to conduct a comparative analysis of binary and multiclass potato leaf disease classification using VGG19 model.

Potato Leaves

A potato leaf is a unique leaf type found primarily on potato plants and it has its own characteristics, differences from regular leaves, and its significance in gardening. These characteristics are broad and smooth leaflets, simple leaf structure which are larger leaflets that appear more unified, flat texture and appearance with few slight ridges. Potatoes are good source of carbohydrates, potassium, and vitamin C; they are also a source of fiber, vitamin B6, and folate and they are low in fat, calories, and sodium (Britannica, 2024). Having seen potato leaf characteristic which can help to know when the potato leaf has deficiency and there are numerous diseases affecting potato, it's mostly detected from potato leaves, these potato leaf diseases includes potato late blight (P*hytophthora infestans*), potato early blight leaf (*Alternaria solani*), potato Bacterial Wilt leaf (*Ralstonia solanacearum*), potato powdery mildew leaf (*Erysiphe cichoracearum*) etc., for the purpose of this study, the researcher considers dealing with healthy potato leaf and three major dominant potato leaf diseases at the time of data capture in Northern part of Nigeria, these are potato early blight leaf, potato late blight leaf and potato leaf virus.

The healthy potato leaf has green colour and is not edible. They are poisonous, because they are solanaceous plants of Solanaceae, as well as pepper, eggplant, tomato and other plants (Britannica, 2024). Healthy potato is shown in figure 1.

Figure 1: Healthy Potato Leaf

Potato early blight leaf disease is caused by the fungus, Alternaria solani, which can cause disease in potato, tomato, other members of the potato family, and some mustards (Plant-village, 2024). This disease, also known as target spot, rarely affects young, vigorously growing plants. Potato early blight leaf is shown in figure 2.

Figure 2: Potato Early Bacterial Blight Leaf

Potato late blight leaf disease caused by the fungus phytophthora infestans is the most important disease of potato that can result into crop failures in a short period if appropriate control measures are not adopted (Plant-village, 2024). Losses in potato yield can go as high as 80% in epidemic years. It is shown in figure 3.

Figure 3: Potato Late Bacterial Blight Leaf

Potato virus leaf disease is caused by pathogen, a viral infection transmitted by aphids. It can remain latent in tubers, spreading to the next crop through infected seed potatoes.
Symptoms of potato virus leaf disease includes mosaic or mottled yellowing of leaves, yellowing of veins and areas between veins (chlorosis), leaves may curl or wrinkle stunted growth and reduced yield (Plant-village, 2024). Potato virus leaf disease is shown in figure 4.


Figure 4: Potato Virus Leaf

Review of Related Works

Ghosh et al. (2023) expounded that potato crops are vital to global food security and economy, yet they are vulnerable to a wide range of leaf diseases that can significantly impact yield and quality. Rapid diagnosis and accurate identification of these disorders are critical for effective disease control and prevention. The research offered an extensive evaluation and contrast of three state-of-art CNN models- VGG19, DenseNet121 and ResNet50 in order to identify and forecast potato leaf diseases. Our study employed a sizable dataset of potato leaf images, containing diverse healthy and afflicted specimens, to train and assess the performance of the chosen CNN models. Extensive data augmentation techniques were employed to enhance the dataset's diversity and generalization capabilities. The research evaluated the models considering their accuracy, precision, recall, F1-score and computational efficiency to determine the most fitting model for real-life applications. The results demonstrate that all three CNN models achieved high performance in identifying and predicting potato leaf diseases, with VGG19 emerging as the top performer followed closely by DenseNet121 and ResNet50.

Rashid et al. (2021) highlighted the difficulty of detecting potato leaf diseases at an early stage due to variations in crop types, disease symptoms, and environmental influences. These factors hinder early diagnosis. Although several machine learning methods exist, they often underperform on broader datasets because they are typically trained on region-specific images. To address this, a two-level deep learning model was developed. The first stage used YOLOv5 for segmenting potato leaves from plant images, while the second stage employed a CNN to identify early and late blight diseases. The model was trained on 4,062 leaf images from Central Punjab, Pakistan, achieving a 99.75% accuracy. Validation on the Plant Village dataset confirmed its robustness, outperforming several existing models in both accuracy and efficiency.

Madhumitaa et al. (2024) addressed the challenges in soybean farming, especially under abiotic stress conditions like drought and salinity prevalent in tropical regions. The study bridged research gaps by integrating morphological traits into machine learning models for genotype classification. Using six morphological features, various algorithms including SVM, RFC, GB, kNN, and Naive Bayes were tested. Plants were classified into Control and Stress groups. Among the models, SVM stood out with a 96.79% accuracy, proving that morphological data significantly improves classification and supports sustainable soybean cultivation.

Kumar and Patel (2023) emphasized the importance of agriculture in national economies and stressed the role of smart farming in managing crop diseases. They proposed a Hierarchical Deep Learning Convolutional Neural Network (HDLCNN) for detecting leaf diseases. The model involved image denoising via Median Filtering, feature extraction using Intuitionistic Fuzzy Local Binary Pattern (IFLBP), and disease classification through HDLCNN. It outperformed existing methods like VGG-INCEP, Deep CNN, and Spiking Neural Networks in accuracy, precision, recall, and other metrics. The system helps farmers take timely action, enhancing potato crop yield and reducing disease impact.

Ajoh et al. (2024) introduced a novel CNN model for yam disease classification using a hybrid activation function that combines ReLU and ELU. While ReLU is fast and simple, ELU tackles issues like the dying ReLU effect. Their combined use improved the CNN's learning capacity and generalization. Testing on healthy and diseased yam images showed superior performance compared to using individual activations. The model captured highly distinctive features of yam diseases, improving accuracy and robustness. The research underscores the critical role of activation functions in boosting CNN performance for agricultural image classification tasks.

Lee et al. (2021) described potato as the world's fourth most significant food crop, often vulnerable to early and late blight diseases. Early detection is crucial to mitigate losses. This study proposed an optimized CNN architecture tailored for potato disease detection, leveraging image processing techniques for dataset creation. Adam optimizer and cross-entropy loss were used for training, with Softmax as the final classification layer. The model achieved 99.53% accuracy and significantly reduced parameter usage by 99.39%, supporting real-time disease monitoring in smart farming.

Emuoyibofarhe et al. (2019) developed machine learning models to identify two major cassava diseases in Nigeria—Cassava Mosaic Disease (CMD) and Cassava Bacterial Blight Disease (CBBD). From over 18,000 images collected at various symptom stages, 46 models were trained. One model assessed leaf health while another diagnosed specific diseases in infected leaves. Using 5-fold cross-validation, the best-performing models were Cubic SVM (83.9% accuracy for healthy/unhealthy detection) and Coarse Gaussian SVM (61.6% for disease classification).

Owomugisha et al. (2021) proposed using spectral data instead of traditional image data to detect cassava diseases, especially before symptoms appear. Visible and near-infrared spectra were gathered from leaves affected by cassava mosaic and brown streak diseases. Since the spectral data contained high-dimensional features, functional decomposition and matrix relevance learning were applied for dimensionality reduction and classification. The Generalized Matrix Relevance Learning Vector Quantization method performed well in identifying the most informative wavelengths for accurate disease diagnosis.

Metlek (2023) worked on detecting cassava leaf diseases using the "Cassava-Leaf-Disease-Classification" dataset with over 21,000 images. Images underwent Chan-Vese segmentation to isolate areas of interest. Features were extracted using ResNet50 and MobileNetV2, then classified via SVM and kNN. With 5-fold cross-validation, ResNet50 paired with SVM achieved the highest accuracy—85.4% on training data and 84.4% on testing. The trained models were deployed through a web interface using MATLAB Builder NE, offering farmers a user-friendly decision support system.

Jagadish et al. (2023) focused on simplifying cassava disease detection by automating the process with machine learning. Manual methods are often inconsistent and time-consuming. The automated system trained on a labeled dataset provided reliable classification of plant conditions. It aimed to assist greenhouse farmers in early intervention and crop management, showing how technology can replace traditional visual inspection for improved accuracy and efficiency in plant health monitoring.

Many researchers have conducted research to address the problems related to analysis of potato leaf disease classification using different analysis tools and different potato leaf diseases. Meanwhile, among the related studies on analysis of potato leaf disease classification, it is recommended that more research are needed because there are some limitations in some models. Nevertheless, Ghosh et al. (2023)'s was evaluation and contrast of three state-of-art CNN models: VGG19, DenseNet121 and ResNet50 to identify and forecast potato leaf classes (potato late blight and potato early blight). Their study's result showed that VGG19 emerged as the top performer with an accuracy of 92.71% followed closely by DenseNet121 and ResNet50. However, this study discovered a problem in potato leaf disease classification using VGG19 model in which binary class of potato leaf (potato early blight and potato late blight diseases) was used and it was not enough for dataset generalization. Therefore, this research consider this challenge in Ghosh et al. (2023)'s as research gap and aims to conduct comparative analysis of binary and multiclass potato leaf disease classification using VGG19 model to solve this research gap.

MATERIALS AND METHODS

This research adopted "Experimental Research Design Methodology" to investigate the effect of class granularity on the performance of the VGG19 deep learning model in classifying potato leaf diseases. The study compares two classification scenarios: binary classification (two classes) and multiclass classification (four classes). In both cases, the same model architecture, training parameters, and dataset structure are used to ensure consistency and reliability of results. This comparative experimental approach allows for objective evaluation of model performance across both classification tasks using quantitative metrics. The methodology consists of seven key phases, which includes:

i. Dataset Collection
ii. Dataset Preprocessing and Augmentation
iii. Architecture of VGG19 Model
iv. Feature Extraction, Training, Validation and Testing of Dataset
v. Model Performance Evaluation Metrics
vi. Comparative Analysis of Results
vii. Visualization and Statistical Interpretation

Dataset Collection

This study identified four classes of potato leaf from local farms in Kaduna and was confirmed by a local farmer that the potato leaves were healthy potato leaf, potato virus leaf, potato early blight leaf and potato late blight leaf. Then, after the confirmation of the potato leaf classes, the potato leaf image was captured using Sony W610 Camera at a distance of 0.3m and these potato leaf images was gathered to form a single dataset. More so, these locally gathered potato leaf images will serves as dataset bank for future study.

Dataset Preprocessing and Augmentation

These images collected were preprocessed and resized to 224x224 pixels to make the suitable input dimension for the VGG 19 model. The dataset was split into 80% for training, 10% for testing and 10% for validating respectively. The dataset contains 1,200 images which includes four potato leaf classes, these include one control class for healthy potato leaf (300 images) and three classes of potato leaf diseases which are potato early blight Leaf (300 images), potato late blight leaf (300 images) and potato virus leaf (300 images). The reason for equal amount of data class size was to avoid data imbalance.

Architecture of VGG19 Model

The VGG model, or VGGNet, that supports 19 layers is also referred to as VGG19, which is a convolutional neural network model proposed by Zisserman, .A and Simonyan, K., from the University of Oxford. These researchers published their model in the research paper titled, "Very Deep Convolutional Networks for Large-Scale Image Recognition". ImageNet is a dataset consisting of more than 14 million images belonging to nearly 1000 classes. Moreover, it was one of the most popular models submitted to ILSVRC-2014. It replaces the large kernel-sized filters with several 3×3 kernel-sized filters one after the other, thereby making significant improvements over AlexNet. The VGG16 model was trained using Nvidia Titan Black GPUs for multiple weeks. As mentioned above, the VGGNet-19 supports 19 layers and can classify images into 1000 object categories, including keyboard, animals, plants, pencil, mouse, etc. Additionally, the model has an image input size of 224-by-224. Final, VGG-19 consists of 16 convolutional layers and three fully connected layers. Figure 5 shows VGG19 network structures.

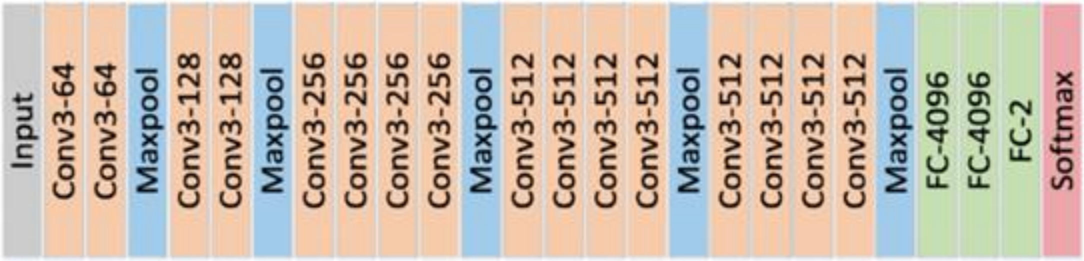

Figure 5: VGG19 Network Structure (Wang et al., 2020)

Feature Extraction, Training, Validation and Testing of Dataset

The feature extraction was done by convolutional and pooling layers (feature extraction layers) of VGG19 model. After the last convolutional layers extract spatial feature maps from the input image, it passed to the fully connected layers (FCL). The first FC layer aggregates spatial and channel-wise features into a high-dimensional feature vector, learning global patterns across the image and passed it to second FC layer. The second FC layer refines the feature representation by learning higher-level combinations of features from the first FC layer while the third FC layer maps the high-dimensional feature vector (4,096 dimensions) to the number of output classes and passes it to a softmax layer to select the highest probability as the final predicted label for image classification. At each epoch, the model performance is evaluate on unseen data after train epoch to check its ability to generalize. Validation dataset are passed through the model at the end of each epoch and metrics like accuracy and loss are computed to monitor model's performance but the model's parameters are not changed or updated. Likewise, testing dataset are passed through the model after training and metrics like accuracy precision, recall, F1score, or confusion matric are used to calculated the model performance.

Model Performance Evaluation Metrics

The VGG 19 model performance evaluation for the potato leaf disease classification was done using these performance evaluation metrics: Accuracy, Precision, Recall, F_1-Score, Macro F_1 Score, Confusion Matrix and ROC Curve. Their equation and computations for performance metrics are below (Evwiekpaefe & Amrevuawho, 2023).

Accuracy

This is the sum of all TP divided by the number of instances in test dataset as expressed in equation (1).

$$Accuracy = \frac{\Sigma \text{TP}}{\text{Number of test data}} \quad (1)$$

Precision

Precision is the ratio of true positives and total positives predicted. It is expressed mathematically in equation (2)

$$Precision\,TP = \frac{TP}{\text{TP + EP}} \quad (2)$$

Recall

The recall (sensitivity) is the measure of your true positive over the count of actual positive outcomes which is expressed mathematically in equation (3).

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

F_1-Score

F_1 score is the harmonic mean between precision and recall which is expressed mathematically in equation (4).

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

Macro F_1 – Score

This is the average over all F_1-Score for a multiclass task and it is expressed mathematically in equation (5) where n is the number of classes in the target class.

$$Macro\,F1 - Score = \frac{\Sigma \text{ni = F1 – Score}}{n} \quad (5)$$

Confusion Matrix

The ground-truth labels and model predictions are shown in a table as a confusion matrix. In the confusion matrix, each row represents an instance in a predicted class, and each column represents an instance in an actual class (Evwiekpaefe & Lawi, 2024).

Table 1: Confusion Matrix Values

		Actual Values	
Predicted Values		**Positive (1)**	**Negative (0)**
	Positive (1)	TP	FP
	Negative (0)	FN	TN

Where TP, FP. FN and TN are interpreted as:
True Positive (TP): Predicted as True and it is True in reality.
True Negative (TN): Predicted as False and it is False in reality.
False Positive (FP): Predicted as True and it is False in reality.
False Negative (FN): Predicted as False and it is True in reality.

ROC Curve

The ROC Curve (Receiver Operating Characteristic curve) is a graphical representation used to evaluate the performance of a binary classification model. It illustrates the trade-off between True Positive Rate (TPR) and False Positive Rate (FPR) at various threshold levels. The curve helps assess how well a model distinguishes between two classes. It is expressed mathematically in equation (6) and equation (7).

$$True\ Positive\ Rate\ (TPR) = \frac{\text{True Positive (TP)}}{\text{True Positive (TP) + False Negative (FN)}} \quad (6)$$

$$False\ Positive\ Rate\ (FPR) = \frac{\text{False Positive (FP)}}{\text{False Positive (FP) + True Negative}} \quad (7)$$

Comparative Analysis of Results

Comparative analysis of model results was done using performance evaluation metrics. After training both models, their performance was compared in terms of their total epoch metrics (training, validation and testing accuracy and loss).

Visualization and Statistical Interpretation

Line plot was used for visualization and paired t-test was used statistical interpretation of the binary and multiclass potato leaf classification using VGG19 model results.

Tools and Environment

Python Programming Language was used for training and analysis of binary class and multiclass of potato leaf disease classification using VGG19 model in Jupyter Notebook Environment of Spyder Anaconda IDE.

RESULTS AND DISCUSSION

Results and Evaluation of Multiclass Potato Leaf Disease Classification using VGG19 Model

From figure 6 to figure 10 were the performance and evaluation report obtained from "Multiclass Potato Leaf Disease Classification using VGG19 Model". These includes: the VGG19 Model training, validation, testing result and evaluation report.

Training, Validation and Testing Result of Multiclass Potato Leaf Disease Classification using VGG19 Model

The training, validation and testing result of the trained Multiclass VGG 19 Model were shown in figure 6 and figure 7 respectively. From the result, at epoch 40, VGG 19 Model achieved training accuracy of 91.28% and loss of 0.3794, validation accuracy of 87.50% and loss of 0.3893 and testing accuracy of 91.67% and loss of 0.4956.

```
Epoch 40/40
30/30 ━━━━━━━━━━━━━━━━━━━━ 190s 6s/step - accuracy: 0.9128 - loss: 0.3794 - val_accuracy: 0.8750 - val_loss: 0.4248
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format :
nd using instead the native Keras format, e.g. `model.save('my_model.keras')` or `keras.saving.save_model(model, 'my_model.keras
4/4 ━━━━━━━━━━━━━━━━━━━━ 22s 5s/step - accuracy: 0.8625 - loss: 0.3893
Validation Accuracy: 87.50%
4/4 ━━━━━━━━━━━━━━━━━━━━ 21s 5s/step - accuracy: 0.8677 - loss: 0.4956
Test Accuracy: 91.67%
```

Figure 6: Training, Validation and Testing Report of the Trained Multiclass Potato Leaf Disease Classification using VGG19 Model

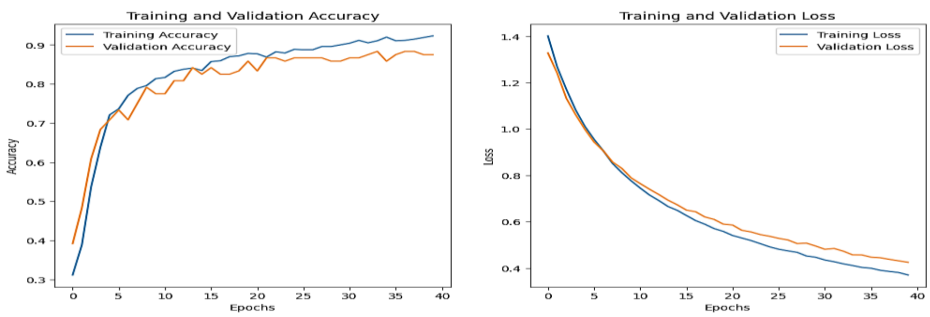

Figure 7: Training and Validation Accuracy and Training and Validation Loss of the Trained Multiclass Potato Leaf Disease Classification using VGG19 Model

The figure 7 shows the graphical representation of training and validation accuracy and training and validation loss across the epochs. Also, it shows that, as the training and validation accuracy increases, the training and validation loss decreases across the epochs.

Evaluation Report of the Trained the trained Multiclass Potato Leaf Disease Classification using VGG19 Model

The evaluation report of trained Multiclass VGG 19 Model which includes the confusion matrix, ROC curve and classification report (precision, recall, f1-score and support) were shown in figure 8, figure 9 and figure 10 respectively.

Classification Report:

	precision	recall	f1-score	support
Potato leaf has Early Blight Disease	1.00	0.73	0.85	30
Potato leaf has Late Blight Disease	0.85	0.97	0.91	30
Potato leaf has Virus Disease	0.97	0.97	0.97	30
Potato leaf is Healthy	0.88	1.00	0.94	30
accuracy			0.92	120
macro avg	0.93	0.92	0.91	120
weighted avg	0.93	0.92	0.91	120

Figure 8: Classification Report of Multi Class

The figure 8 shows the values obtained from the different classes of potato leaf and its corresponding precision, recall, f1-score and support.

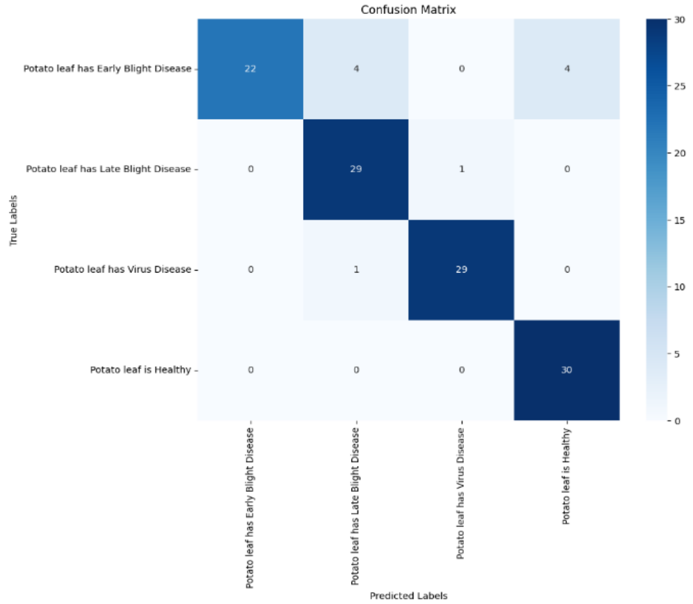

Figure 9: Confusion Matrix of Multiclass

The figure 9 shows the confusion matrix of the trained VGG 19 model for potato leaf disease classification. It also shows the actual values given in the test data and the predicted values by the VGG 19 model. However, the actual values were placed accordingly to the potato leaf classes respectively. More so, thirty healthy potato leaves, twenty-nine potato virus leaves, twenty-nine potato late blight leaves and twenty-two potato early blight leaves were correctly classified while one potato virus leaf was misclassified as potato late blight leaf, one potato late blight leaf was misclassified as potato virus leaf, four potato early blight leaves were misclassified as potato late blight leaves and four early blight leaves were misclassified as healthy potato leaves.

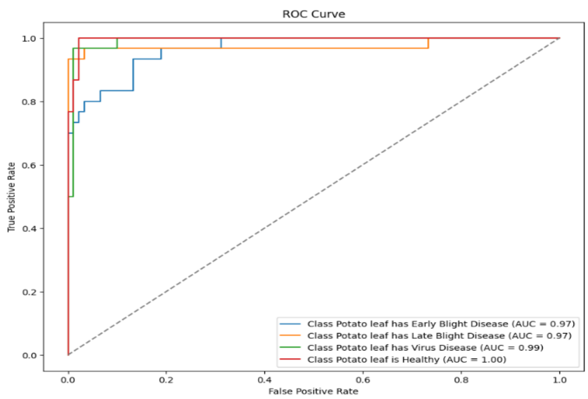

Figure 10: ROC Curve of Multiclass

Figure 10 illustrates the trade-off between True Positive Rate (TPR) and False Positive Rate (FPR) at various threshold levels of the potato leaf classification.

Results and Evaluation of Binary Class Potato Leaf Disease Classification using VGG19 Model

From figure 11 to figure 15 were the performance and evaluation report obtained from "Binary Class Potato Leaf Disease Classification using VGG19 Model". These includes: the VGG19 Model training, validation, testing result and evaluation report.

Training, Validation and Testing Result of Binary Class Potato Leaf Disease Classification using VGG19 Model

The same potato leaf dataset containing binary class (potato early blight and potato late blight) were trained using VGG 19 model as shown in figure 4.6. At epoch 40, VGG 19 model achieved training accuracy of 93.25% and loss of 0.2513, validation accuracy of 90.00% and loss of 0.2970 and testing accuracy of 91.67% and loss of 0.2735.

```
Epoch 40/40
15/15 ━━━━━━━━━━━━━━━━━━━━ 87s 6s/step - accuracy: 0.9325 - loss: 0.2513 - val_accuracy: 0.9000 - val_loss: 0.2702
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format is
nd using instead the native Keras format, e.g. `model.save('my_model.keras')` or `keras.saving.save_model(model, 'my_model.keras')`
2/2 ━━━━━━━━━━━━━━━━━━━━ 10s 5s/step - accuracy: 0.8917 - loss: 0.2970
Validation Accuracy: 90.00%
2/2 ━━━━━━━━━━━━━━━━━━━━ 10s 5s/step - accuracy: 0.9028 - loss: 0.2735
Test Accuracy: 91.67%
```

Figure 11: Result of the Training, Validation and Testing Report of the Trained Binary Class Potato Leaf Disease Classification using VGG19 Model

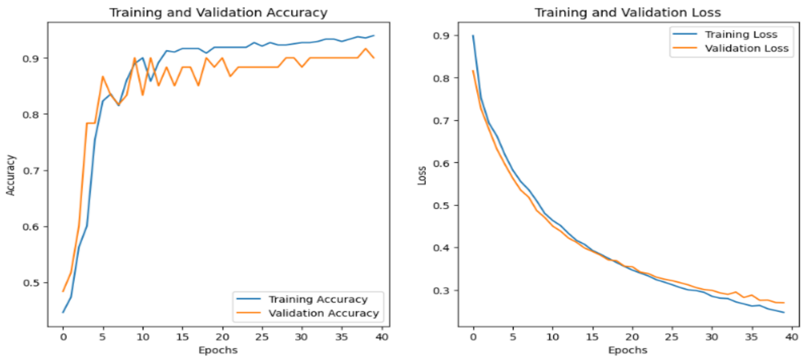

Figure 12: Training and Validation Accuracy and Training and Validation Loss of the trained Binary Class Potato Leaf Disease Classification using VGG19 Model

The figure 12 shows the graphical representation of training and validation accuracy as it increases across the epoch. Also, it shows the graphical representation of training and validation loss as it decreases across the epoch.

Evaluation Report of the Trained the Trained Binary Class Potato Leaf Disease Classification using VGG19 Model

The evaluation report of trained Binary Class VGG 19 Model which includes the confusion matrix, ROC curve and classification report (precision, recall, f1-score and support) were shown in figure 13, figure 14 and figure 15 respectively.

```
Classification Report:
                                      precision    recall  f1-score   support

Potato leaf has Early Blight Disease       0.96      0.87      0.91        30
 Potato leaf has Late Blight Disease       0.88      0.97      0.92        30

                            accuracy                           0.92        60
                           macro avg       0.92      0.92      0.92        60
                        weighted avg       0.92      0.92      0.92        60
```

Figure 13: Classification Report of Binary Class

The figure 13 shows the values obtained from the different classes of potato leaf and its corresponding precision, recall, f1-score and support.

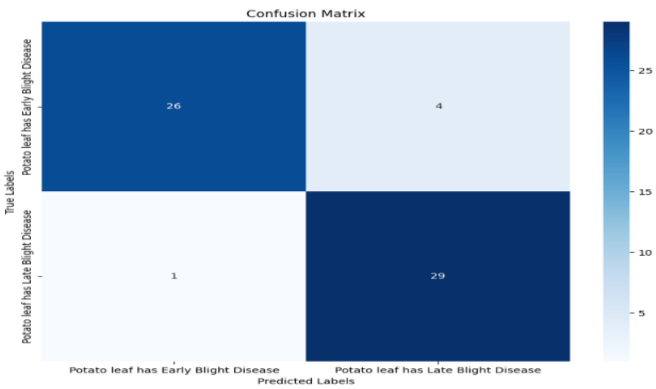

Figure 14: Confusion Matrix of Binary Class

The figure 14 shows the confusion matrix of the trained VGG 19 model for potato leaf disease classification. It also shows the actual values given in the test data and the predicted values by the VGG 19 model. However, the actual values were placed accordingly to the potato leaf classes respectively. More so, twenty-nine potato late blight leaves and twenty-six potato early blight leaves were correctly classified while one potato late blight leaf was misclassified as potato early blight leaf and four potato early blight leaf was misclassified as potato late blight leaves.

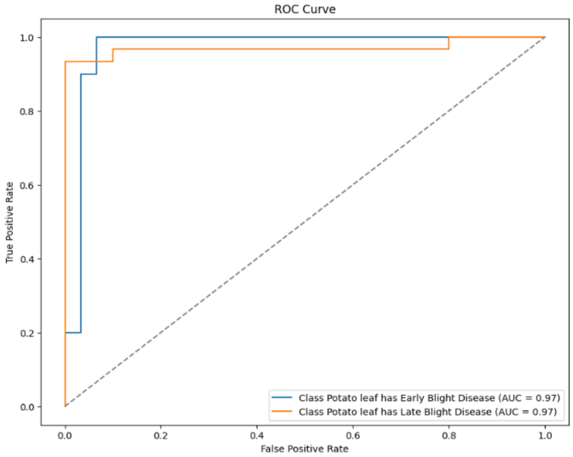

Figure 15: ROC Curve of Binary Class

The figure 15 illustrates the trade-off between True Positive Rate (TPR) and False Positive Rate (FPR) at various threshold levels of the potato leaf classification.

Comparison between Binary and Multiclass Potato Leaf Disease Classification using VGG19 Model

From the result of this study, it shows that both data classes has same test accuracy but other results varies. However, Table 2 shows the constructive comparison between the Binary and Multiclass Potato Leaf Disease Classification using VGG19 Model.

Table 2: Comparison Between Binary and Multiclass Potato Leaf Disease Classification using VGG19 Mode

S/N	Metrics	Binary Class (%)	Multiclass (%)
1.	Training Accuracy	93.25	91.28
2.	Training Loss	0.2513	0.3794
3.	Validation Accuracy	90.00	87.50
4.	Validation Loss	0.2970	0.3893
5.	Testing Accuracy	91.67	91.67
6.	Testing Loss	0.2735	0.4956

Visualization and Statistical Interpretation of Binary and Multiclass Potato Leaf Disease Classification using VGG19 Model

Paired t-test was used statistical interpretation of the binary and multiclass potato leaf classification using VGG19 model results across the 40 epochs which gave t-statistic of 7.4054 and p-value of 4.08e-05. Line plot was used for visualization of the binary and multiclass potato leaf classification using VGG19 model results across the 40 epochs which gave t-statistic of 8.1434 and p-value of 0.0.

```
print("Paired t-test result:")
print("t-statistic =", round(t_stat, 4))
print("p-value =", round(p_value, 6))

# Interpretation
if p_value < 0.05:
    print("🔍 The difference in accuracy is statistically significant.")
else:
    print("✅ No significant difference in accuracy between binary and multiclass.")
```

```
Paired t-test result:
t-statistic = 8.1434
p-value = 0.0
🔍 The difference in accuracy is statistically significant.
```

Figure 16: Paired t-Test of the Binary and Multiclass Potato Leaf Classification Using VGG19 Model

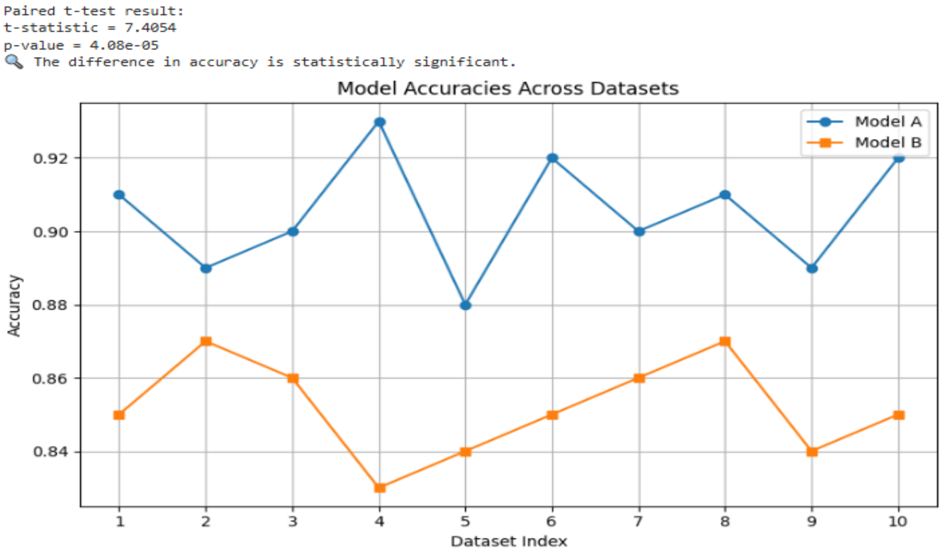

Figure 17: Line plot of the Binary and Multiclass Potato Leaf Classification using VGG19 Model

The testing accuracy obtained from both binary class (two class) and multiclass (four class) potato leaf disease using VGG 19 model was 91.67% accuracy, the training accuracy of binary class and multiclass was 93.25% and 91.28% respectively while the validation accuracy of binary class and multiclass was 90.00% and 87.50% accuracy respectively. This indicates that VGG19 model trained on two disease classes showed slightly better performance in training and validation metrics, achieving higher accuracy and lower loss. However, both models reached the same testing accuracy of 91.67%. The higher testing loss in the multiclass (four-class) model may indicate greater difficulty in generalizing to more classes, suggesting that model complexity or dataset size may need to be adjusted for multiclass classification. Therefore, this indicates that the VGG 19 model's testing accuracy was good but the training and validation accuracy reduces as the dataset class increases, which means that there is likelihood of overfitting as dataset class increases. Finally, the higher the number of data class in VGG19 model, the lower the training and validation accuracy.

CONCLUSION

This research explored the comparative analysis of binary and multiclass potato leaf disease classification using VGG19 Model. This study used model evaluation metrics and comparative tools to comparatively analyze binary and multiclass potato leaf disease classification using VGG19 Model. Testing accuracy of both binary class (two class) and multiclass (four class) potato leaf disease using VGG 19 model was 91.67% accuracy, the training accuracy of binary class and multiclass was 93.25% and 91.28% respectively while the validation accuracy of binary class and multiclass was 90.00% and 87.50% accuracy respectively. This indicates that the VGG 19 model's testing accuracy was good but the training and validation accuracy reduces as the dataset class increases, which means that there is likelihood of overfitting as dataset class increases. Thus, the higher the number of data class in VGG19 model, the lower the training and validation accuracy.

While there are many works undertaken in the field of deep learning and plant classification, this study used model evaluation metrics and comparative tools to comparatively analyze binary and multiclass potato leaf disease classification using VGG19 Model. This study achievement and contribution to knowledge shows that the higher the number of data class in VGG19 model, the lower the training and validation accuracy.

This work has explored many research papers, identified gap, and tried to cover the identified gap by proffering solution classes of potato leaf disease classification using VGG 19 model. This work improved on the previous work by using model evaluation metrics and comparative tools to comparatively analyze binary and multiclass potato leaf disease classification using VGG19 Model which shows that shows that the higher the number of data class in VGG19 model, the lower the training and validation accuracy.

It is recommended that the "Comparative Analysis of Binary and Multiclass Potato Leaf Disease Classification Using VGG19 Model" be adopted and used for potato leaf disease classification thereby improving potato crop yield. More so, it is recommended that the future work should include improving VGG19 model. Moreover, work can further be extended to include the classification of other plant diseases.

REFERENCES

Ajoh, S. L, Ogala, E., Samuel, J. T., Akpa, J., & Daniel, A. M. (2024). A Combined Activation Function for Learning Performance Improvement of Convolution Neural Network for Yam Disease Classification. *Journal of Software Engineering and Applications, 1* (3), 110 -127, April 2024.

Albelwi, S. & Mahmood, A. (2017). A Framework for Designing the Architectures of Deep Convolutional Neural Networks. *Entropy 2017, 19*(6), 1-20. https://doi.org/10.3390/e19060242

Alex, P. (2022). *How to Train an Ensemble of Convolutional Neural Networks for Image Classification.* Retrieved on 12th July, 2024 from https://medium.com/@alexppppp/how-to-train-an-ensemble-of-convolutional-neural-networks-for-image-classification-8fc69b087d3

Anbarjafari, G. (2022). *Digital Image Processing*. Retrieved on 12th July, 2024 from https://sisu.ut.ee/imageprocessing/book

Beheshti, N. (2022). *Guide to Confusion Matrices and Classification PerformanceMetrics*. Retrieved on 21st June, 2024 from https://towardsdatascience.com/guide-to-confusion-matrices-classification-performance-metrics-a0ebfc08408e

Britannica (2024). *Potato*. Encyclopedia Britanncia Retrieved on 21st August, 2024 from https://www.britannica.com/science/starch

Chugh, G., Sharma, A., Choudhary, P., & Khanna, R. (2020). Potato Leaf Disease Detection Using Inception V3. *International Research Journal of Engineering and Technology (IRJET), 07* (11), 1363-1366, e-ISSN: 2395-0056, p-ISSN:2395-0072, www.irjet.net

Emuoyibofarhe, O., Emuoyibofarhe, J.O., Adebayo, S., Ayandiji, A., Demeji, O., & James, O. (2019). Detection and Classification of Cassava Diseases Using Machine Learning. *International Journal of Computer Science and Software Engineering (IJCSSE), 8:7, July 2019* Page: 166-176. ISSN (Online): 2409-4285 www.IJCSSE.org

Evwiekpaefe, A. E, & Amrevuawho, O. F. (2023). Classification of Dermatologic Manifestations of Cardiovascular Disease Using Efficient NetV2 CNN Model. *International Journal of Intelligent Computing and Information Sciences, IJICIS, 23* (1), 115-127. https://doi.org/10.21608/ijicis.2023.184311.1242

Evwiekpaefe, A. E., & Lawi, V. P. (2024). An Artificial Neural Network Model for Predicting Children at Risk of Defaulting from Routine Immunization in Nigeria. *Journal of Information Systems and Informatics, 6* (2), 810-829. https://doi.org/10.51519/journalisi.v6i2.689

Ghosh, H., Rahat I. S., Shaik, K., Khasim, S., & Yesubabu, M. (2023). Potato Leaf Disease Recognition and Prediction using Convolutional Neural Networks. *Journal of EAI Endorsed Transactions on Scalable Information Systems 2023, 10* (6), 1-8. https://publications.eai.eu/index.php/sis/issue/view/409

IBM (2020a). *Deep learning*. International Business Machine. Retrieved on 21st June, 2024 from https://www.ibm.com/topics/machine-learning

IBM (2020b) *Machine learning*. International Business Machine. Retrieved on 21st June, 2024 from https://www.ibm.com/topics/machine-learning

Jadon, S. & Jasim, M. (2019). *Video Summarization using Keyframe Extraction and Video Skimming.* Retrieved on 12th July, 2024 from https://www.researchgate.net/publication/336510662

Jagadish, K., Jagadhiswar, Sai, P. K., Vishnu, V. B., Jahnavi, P., Janaki, M., & Rao, R.K. (2023) Detection and Classification of Cassava Diseases using Machine Learning. *International Journal for Research in Applied Science & Engineering Technology (IJRASET) 11* (12), 803-807. https//:www.ijraset.com

Kumar, A. & Patel, V. K. (2023). Classification and identification of disease in potato leaf using hierarchical based deep learning convolutional neural network. *Multimed Tools Appl 82,* 31101–31127. https://doi.org/10.1007/s11042-02314663-z

Lee, T. Y., Lin, A., & Yu, J. Y., (2021). High Efficiency Disease Detection for Potato Leaf with Convolutional Neural Network. *SN COMPUT. SCI. 2* (297), 691-699.

Madhumitaa, P. S., Ragav, I. C., Kiranmayi C., Prabhavathy, V. M & Prabhavathy P. (2024). Drought and Salinity Stress Classification in Soya bean Crops: Comparative Analysis of Machine Learning Models. *2024 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), Karaikal, India, 2024,* 1-5. https://doi.org/10.1109/IConSCEPT61884.2024.10627854

Malinowski, E., Zimányi, E., Joseph, S. K., Warehouse, D., Inmon, B., Analytical, O., Olap, P., Gatziu, S., Vavouras, A., Nilsson, A. A., & Merkle, D. (2019). *Data Vault* 2.0, *1* (1), 1–15; https://doi.org/10.1007/978-3-322-94873-1

Metlek, S. (2021). Disease Detection from Cassava Leaf Images With Deep Learning Methods in Web Environment. *Int. Journal of 3D Printing Tech. Dig. Ind., 5*(3): 625-644.

Owomugisha, G., Melchert, F., Mwebaze, E., Quinn, J.A., & Biehl, M. (2021) Matrix Relevance Learning From Spectral Data for Diagnosing Cassava Diseases. *IEEE Access Journal. 9,* (83355-83363) https://doi.org/10.1109/ACCESS.2021.3087231

Plant-village (2024). *Potato*. Plant Village. Retrieved on 21st August, 2024 from https://plantvillage.psu.edu/topics/potato/infos

Rashid, J., Khan, I., Ali, G., Almotiri, S.H., AlGhamdi, M. A., & Masood K., (2021). Multi-Level Deep Learning Model for Potato Leaf Disease Recognition. *Electronics 2021, 10(17), 2064; https://doi.org/10.3390/electronics10172064*

Sibhatu, K.T., & Qaim, M. (2024). *Rural food security, subsistence agriculture, and seasonality.* Retrieved on 21st August, 2024 from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0186406

Simske, S. (2019). *Meta-analytics: consensus approaches and system patterns for data analysis.* Morgan Kaufmann.

Wang, W., Tian, J., Zhang, C., Luo, Y., Wang, X., & Li, J. (2020). An improved deep learning approach and its applications on colonic polyp images detection. *BMC Medical Imaging 20 (*83), 1-14. https://doi.org/10.1186/s12880-020-00482-3

Zhang, C., & Zhang, Z. (2024). Improving multiview face detection with multi-task deep convolutional neural networks. Retrieved on 21st August, 2024 from 2014 *IEEE Winter Conference on Applications of Computer Vision, WACV 2014,* Colorado, *USA.* 1036–1041. https://doi.org/10.1109/WACV.2014. 6835990