

Article

Not peer-reviewed version

---

# Revealing GLCM Metric Variations across Plant Disease Dataset: A Comprehensive Examination and Future Prospects for Enhanced Deep Learning Applications

---

[Masud Kabir](#) , [Fatih Unal](#) <sup>\*</sup> , [Tahir Cetin AKINCI](#) <sup>\*</sup> , Alfredo A. Martinez-Morales , [Sami Ekici](#) <sup>\*</sup>

Posted Date: 25 April 2024

doi: 10.20944/preprints202404.1566.v1

Keywords: GLCM metrics; deep learning; darkNet19; plant diseases; open datasets; criteria



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Revealing GLCM Metric Variations Across Plant Disease Dataset: A Comprehensive Examination and Future Prospects for Enhanced Deep Learning Applications

Masud Kabir<sup>1</sup>, Fatih Unal<sup>1</sup>, Tahir Cetin Akinci<sup>2,\*</sup>, Alfredo A. Martinez Morales<sup>3</sup> and Sami Ekici<sup>1,\*</sup>

<sup>1</sup> Department of Energy System Engineering, Firat University, Turkiye; sekici@firat.edu.tr

<sup>2</sup> WCGEC, University of California Riverside, 92507, Riverside, Ca, USA. tahircetin.akinci@ucr.edu

<sup>3</sup> WCGEC, University of California Riverside, 92507, Riverside, Ca, USA.

\* Correspondence: sekici@firat.edu.tr; tahircetin.akinci@ucr.edu

**Abstract:** The intricate relationship between Gray-Level Co-occurrence Matrix (GLCM) metrics and machine learning model performance underscores the need for rigorous dataset evaluation and selection protocols to ensure the reliability and generalizability of classification outcomes. This study involved a thorough examination of selected publicly available plant diseases datasets, with an emphasis on how well they performed as measured by GLCM metrics. After first classifying the datasets according to their GLCM metrics, dataset\_2 (D2) and dataset\_5 (D5) were found, respectively, to be the best-performing dataset in all GLCM analyses. The same datasets were then used to train deep learning models, and their classification performances were assessed. A noteworthy association was observed between the results of training deep learning models and the performance ratings derived from GLCM studies. More specifically, dataset\_2 (D2) performed best in both GLCM analysis and deep learning model performance, indicating a strong correlation between the accuracy of classification and the textural qualities that GLCM captured. In the context of plant disease identification, in particular, these results highlight the significance of clearly defined dataset selection criteria in deep learning applications. Scholars can improve the accuracy and dependability of deep learning models for diagnosing plant diseases by giving preference to datasets with favorable GLCM metrics. The research also emphasizes the importance of texture features being taken into account in addition to conventional image features, highlighting the necessity of transparency and rigor in dataset selection procedures.

**Keywords:** GLCM metrics; deep learning; darkNet19; plant diseases; open datasets; criteria

## 1. Introduction

In recent times, the world is facing an increase in population that in turn calls for a red alarm on food security. According to the Food and Agricultural Organization of the United Nations (FOA), 70% increase of recent food production is globally needed by the year 2050 to meet up with the bouncing population [1]. This is posing a serious demand on agriculture to meet up with the foreseen population rapidly and sustainably. However, Plants being the primary source of food are faced with serious challenges in growth. Abnormalities in its growth mainly caused by diseases, pests and environmental factors are yet to reach an end.

Plant diseases can cause anything from modest symptoms to severe damage to large fields of crops, incurring significant expenditures and having a significant negative influence on the agricultural economy. Plant pests and diseases accounts for about 40% loss of global food crops and trade loss of over USD 220 billion every year [2]. These diseases promoted by invisible pathogens poses the most serious effect on the plant's health [3] thus causing a significant threat to food security and affecting the environment. In addition to that, it has been noticed lately that the recent climate

change promotes plant's health problems by early appearance of pests and their dissemination to areas they have never been noticed [4].

Since ancient times, human beings have been struggling to counter these plant health problems. They largely rely on visible changes in plant parts, such as the leaves and stems, due to the absence of enough technologies to pre-detect the appearances hence causing a huge loss that may reach to famine in some regions [5]. This method is time and resource consuming, full of uncertainties, and labor intensive. Moreover, traditional methods were mainly focusing on particular locations thus limiting its efficiency across geographies. In contrast, plant diseases can spread across geographies unknowingly thus the need for a more generalized system of its detection [6].

With the advancements of technology, the world of today has been developing several techniques to tackle these problems. Artificial intelligence brought tremendous novelty in the name of machine learning and deep learning which are currently relied on in these regards. Machine learning techniques were widely used in the research community since the last two decades, when its applications were addressed and studies for agriculture and plant diseases were widely evaluated [7]. Conventional machine learning methods, however, still require time-consuming feature extraction for training models [8]. Moreover, it has limited success under some circumstances such as its inability to process natural data in their raw form [5,9]. As a result, deep learning algorithms that automatically extract features as part of their natural operation have been incorporated into research for about a decade [10–12]. According to [13], deep learning has emerged as a more preferred method for plant disease detection and management due to its increased computing power, storage capabilities, and accessibility of big datasets.

Deep learning simply learns from datasets of the past disease images to detect and identify the plants' health problem(s) using some sets of complex algorithms. As a data-hungry technology, researches in use of deep learning battle with datasets preparation for individual applications. This made the scientific research communities to provide datasets for the public use, simplifying the path for model developments, learning and knowledge sharing. PlantVillage Datasets, developed by [14] is considered one of the most significant datasets that fueled the application of deep learning in plant disease detection and management [15–18]. Maize plant disease dataset [19], Digipathos dataset [20], multiple plant disease datasets [21], PlantDoc dataset [22], cassava disease image dataset [23] and RoCoLe coffee disease dataset [24] are typical examples of available plant disease datasets. Datasets preparation for machine-learning or deep learning-based applications can also be done through the capturing of images on live plants over a period, using remote sensing and hyperspectral imagery [25]. Thermal imagery has also recorded success in capturing such images [26].

However, after a critical review and pre-tests undergone for the purpose of this study, most of the open datasets for plant diseases detection are observed to possess a number of technical problems worthy of noting for deep learning applications. For example, many of the public available datasets such as the PlantVillage dataset are developed in controlled laboratory contexts. This controlled environment provides for exact control over variables, assuring standardized data gathering conditions. Though Authors of [27] trained a deep convolutional neural network (DCNN) with the PlantVillage dataset using AlexNet and GoogleNet architectures and obtain a performance efficiency of 99.35%, Conversely, the controlled setting may not entirely mirror real-world situations encountered by farmers. In lab settings, field factors such as fluctuating light intensities, various backgrounds, and unpredictable weather are not faithfully duplicated. As a result, models trained on such datasets may perform poorly when applied to real-world agricultural contexts [28]. While these datasets give useful insights and initial model training, there is an urgent need to supplement them with field-based data in order to improve the resilience and practical application of deep learning models for plant disease detection. The works of [29,30] tangibly proved high efficiency of model trained with field-based dataset.

Moreover, the size and resolution of the images offered in the open plant disease datasets frequently varies. Some datasets contain very high-resolution images, like the RoCoLe coffee disease dataset [24] and Paddy Doctor dataset [31], while others may contain images of inferior quality. For model training, images with high resolution can be computationally intensive and may necessitate

large processing resources. Low-resolution images, on the other hand, may lack fine-grained features necessary for effective disease detection. It is thus critical to strike a balance between image resolution and processing efficiency. To standardize image sizes in this case, preprocessing procedures are necessary to ensure that models can handle the data successfully while preserving the key visual information [30,32].

Plant diseases are extremely diverse, with hundreds of distinct pathogens affecting thousands of different plant species [6]. Each disease has unique symptoms, making it difficult to create a single universal dataset that includes all possible disease variations. Diseases can present differently even among the same plant species depending on factors such as pathogen strain, climatic circumstances, and host genetics [33]. Because of this complication, dedicated datasets for individual diseases or groups of linked disorders were necessary.

Plant disease databases have been created in certain cases by directly collecting photos from search engine results, notably from platforms such as Google [34]. While this strategy may appear to be convenient, it creates numerous serious obstacles to the dataset's integrity and trustworthiness. These photos lack consistent quality and metadata, and thus may be in violation of copyright laws. Because of search engine biases, their diversity and representativeness are jeopardized. Stringent data gathering techniques from trustworthy and authorized sources are thus required to ensure valid and reliable plant disease datasets. This not only maintains ethical norms, but also improves the dataset's quality and usability for relevant scientific research and deep learning applications in plant pathology.

Researchers are, nevertheless, expected to be mindful of the constraints imposed by specialized datasets. While they are useful for some disease categories, they may not be appropriate for broader applications without further data augmentation or transfer learning approaches, especially for open-field generated datasets for mobile application development.

Using texture analysis methods, including Gray-Level Co-occurrence Matrix (GLCM) metrics, which identify complex texture qualities in images, is one way to tackle these problems [35,36]. These metrics offer a quantitative depiction of texture features and offer insightful information about the spatial correlations between pixel intensities [37,38]. Nevertheless, little is known about how machine learning model performance is impacted by differences in GLCM metrics between datasets.

Progress in the field of plant disease identification depends on the scientific community's ability to comprehend how differences in GLCM variables affect machine learning model performance. Along with providing insight into which datasets are appropriate for training models, it also helps to build reliable, broadly applicable models that are able to diagnose plant diseases in a variety of agricultural contexts. Consequently, it is imperative to conduct a thorough analysis of the differences in GLCM metrics among diverse plant disease datasets and evaluate their consequences for improved deep learning applications in plant disease identification. This research is therefore an attempt to define novel criteria, using some publicly available plant disease datasets in the eyes of deep learning application, through the exploration of technical and textual features of plant disease datasets for enhancement of machine learning and deep learning applications, as the case may be.

## 2. Methodology

### 2.1. Data Collection

A set of five distinct plant disease datasets that are publicly available were acquired online via:

- "plant\_village dataset" - [<https://github.com/spMohanty/PlantVillage-Dataset>]
- "a\_database\_of\_leaf\_images" - [<https://data.mendeley.com/datasets/hb74ynkjc/4>]
- "RoCoLe dataset" - [<https://data.mendeley.com/datasets/c5yvn32dzc/2>]
- "FGVCx\_cassava dataset" - [<https://github.com/icassava/fgvcx-icassava/tree/master#fgvcx-cassava-disease-diagnosis>] and
- "paddy\_doctor dataset" [<https://ieee-dataport.org/documents/paddy-doctor-visual-image-dataset-automated-paddy-disease-classification-and-benchmarking>]

Each of the mentioned datasets comprises a set of diverse plant species, diseases, and image resolutions.

## 2.2. Image Preprocessing & GLCM Metrics

Standardized image sizes and formats across datasets were captured and summarized using the Image Text analysis feature in Python programming language interface.

Gray Level Co-occurrence Matrix (GLCM) was used in further processing of these datasets to obtain the relevant GLCM metrics. These are texture features derived from images that describe spatial relationships between pixel intensities. A total of Ten (10) GLCM metrics including Energy, Contrast, Correlation, Homogeneity, Total Variance, Difference Variance, Maximum Probability, Joint Entropy, Difference Entropy and Angular Second Moment were extracted using predefined parameters and utilized for further analysis.

## 2.3. Statistical Analysis

Descriptive statistics was employed to compute and derive insights into the datasets. These include computed mean, variance, and sums of GLCM metrics for each dataset. Normality and homogeneity of data checks were conducted in an attempt to assess normal distribution assumptions using Shapiro-Wilk and Kolmogorov-Smirnov tests using GraphPad Prism 10.1.2 (324). Applied Kruskal-Wallis test for non-normally distributed data was used for Variability Assessment. Several charts as figures were generated to display correlations and distributions in GLCM metrics. Relationships between these metrics within the datasets were also presented and discussed in the results sections.

## 2.4. Deep Learning Model Training and Evaluation

From each dataset, 77 images were randomly collected (minimum number of classes available from the overall datasets properties used herein) to ensure uniformity as well as avoiding bias in the representation of datasets. A total of 6160 images were summed, from the different classes across the datasets, and used as training data.

According to a conventional procedure, a deep learning model was trained independently with regards to each dataset. The following procedures were part of the training process:

- i. **Preprocessing of the Data:** To improve model performance, better feature representation and guarantee uniformity before processing, the images were resized to a uniform resolution (227x227) and pixel values.
- ii. **Model Selection:** The DarkNet19 architecture was chosen as the foundational model for training since it has a track record of success in image classification challenges. DarkNet19 is a deep convolutional neural network (CNN) model.
- iii. **Hyperparameter Tuning:** To maximize classification accuracy, the DarkNet19 model's hyperparameters—such as initial learning rate (0.0001), mini-batch size (128), maximum epoch (10), weight and bias learning rate factors (10) were adjusted as same values for all training trials.

To guarantee a fair representation of classes, a validation dataset, equal to 20% of the overall dataset, was randomly selected from each dataset. This 80:20 split ratio was aimed at reducing the possibility of overfitting while guaranteeing a sufficient quantity of data for both training and evaluation. Based on the trained models' predictions on the validation dataset, performance metrics like accuracy, precision, recall, and F1-score were calculated. These metrics shed light on how well the model classified plant images that belonged to various disease classifications.

## 3. Results and Discussions

### 3.1. Ideals of Image Properties and Distribution across the Datasets

Images sizes and formats within the relevant plant disease datasets used in this research portrayed some versatile properties that has a potential of inflicting deep learning applications. In a more detailed form, these will be digested in the following sub-headings:

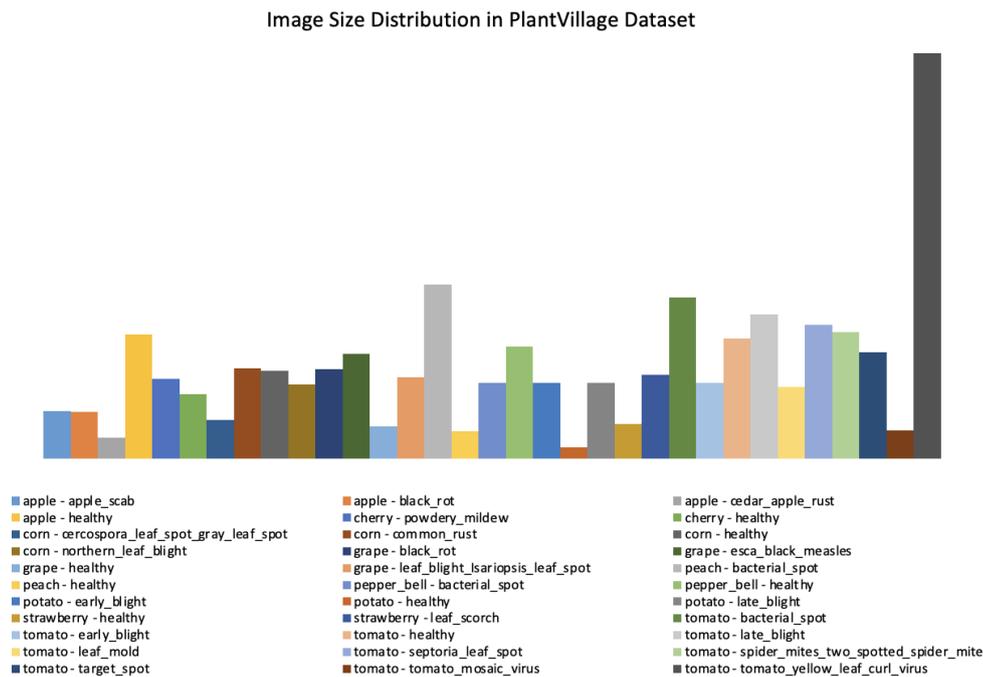
### 3.1.1. Data Imbalance

In order to create high-quality datasets for the detection of plant diseases, data balancing is crucial. It comprises making certain that a roughly equal fraction of each data class or category is represented. This holds particular importance in scenarios where certain classes, like particular plant diseases, might be less prevalent than others. Data balance is important because it allows the model to learn equally and fairly from every class, making the detection system more dependable and strong [39].

The performance of the plant disease identification model may suffer from an imbalanced dataset in several ways, including where certain classes contain much fewer samples than others [40]. Biased training is one such outcome. A presentation of the data imbalance across the datasets, with the respective annotations assigned in this study, could be outlined in a general form and summary through **Error! Reference source not found.** A more specific analysis will be also presented for each dataset.

**Table 1.** General dataset.

Annotation	Dataset name	Total Images	Image Resolutions	Setting	Disease Classes	Plants involved
dataset_1	plant_village dataset	54303	[256x256] throughout	Lab	38	Apple, Cherry, Corn, grape, Peach, Pepper, Potato, Strawberry, and Tomato Mango, Arjun, Alstonia Scholaris, Guava, Bael,
dataset_2	a_database_of_leaf_images	4503	[6000x4000] throughout	Lab	22	Jamun, Jatropha, Pongamia Pinnata, Basil, Pomegranate, Lemon, and Chinar
dataset_3	RoCoLe dataset	1560	[2048x1152], 768 images, [1280x720], 479 images, [4128x2322], 313 images	Field	5	Coffee
dataset_4	FGVCx_cassava dataset	537	Variable between [213x231] to [960x540]	Field	5	Cassava
dataset_5	paddy_doctor dataset	16,225	[1080x1440], 16219 images, and [1440x1080], 6 images	Field	13	Paddy



**Figure 1.** Data size disparity in PlantVillage Dataset (dataset\_1).

Across different classes, the dataset\_1 appears to be significantly unbalanced, with different image sizes and numbers. For example: 'tomato\_yellow\_leaf\_curl\_virus' has a greater number of images in the 'tomato' subfolders than other classes such as 'bacterial\_spot' or 'leaf\_mold'. In addition, 'Healthy' subfolders—like 'cedar\_apple\_rust' and 'background\_without\_leaves'—occasionally included less photos than subfolders dedicated to certain diseases.

A class imbalance is also realized in dataset\_2 between different plant species and their health statuses. For example, there are substantially different numbers of diseased and healthy pictures in classes like "arjun" and "alstonia\_scholaris." Moreover, the inscribed 'Diseased' subfolders typically contain more photos in nearly every class than their 'healthy' counterparts.

The 'coffee' dataset in dataset\_3 includes pictures of coffee plants in various states of health (healthy, red spider mite, rust levels 1–4). There is a notable disparity in class between the subfolders: 'healthy' has the greatest number of images, followed by 'rust\_level\_1' and 'red\_spider\_mite'. As rust gets less severe, there are increasingly fewer images for "rust\_level\_2," "rust\_level\_3," and "rust\_level\_4".

The Image data imbalance in dataset\_4 is found to be significant where 'Healthy' is the class with the most images, followed by 'Brown\_streak\_disease', 'Green\_mite', 'Mosaic\_disease', and 'bacterial\_blight', which each contain fewer images. Within a few disease classes, even certain resolutions appear to be more prevalent (Figure 3).

As dataset\_5 also follow the cue, there are a total of 2351 images in "blast" and 450 images in "bacterial\_panicle\_blight" respectively. In contrast to certain disease classes, the 'healthy' or 'healty' class, which represents healthy paddy plants, as annotated in the dataset's nomenclature style, contains a significant number of images. Further variations in the class data could be observed in Figure 4.

Due to variations and the smaller sample sizes, classes with fewer images may produce GLCM-based features that are less reliable, which could have an impact on model performance. Moreover, if unaddressed during model training, imbalanced classes may result in biased models that perform well on majority classes and badly on minority classes. Other factors that may hinder similar performances relevant to datasets' feature extraction are defined in literature (**Error! Reference source not found.**). In the coming sections, an in-depth observation into these GLCM features is presented.

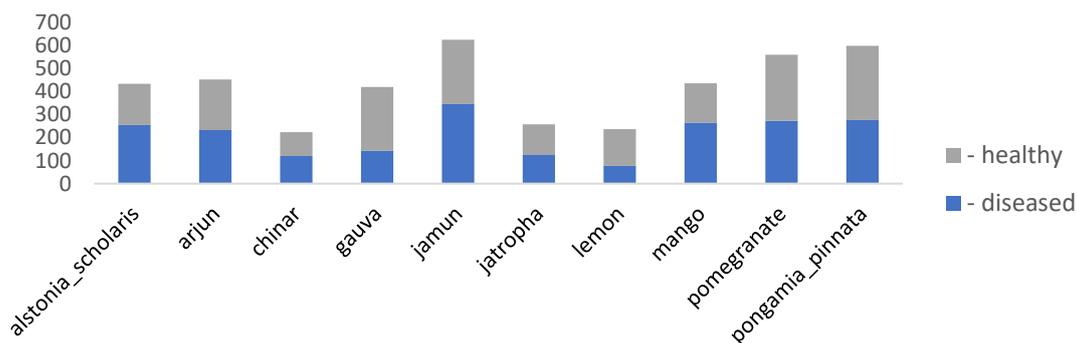


Figure 2. Image Size Distribution in dataset\_2.

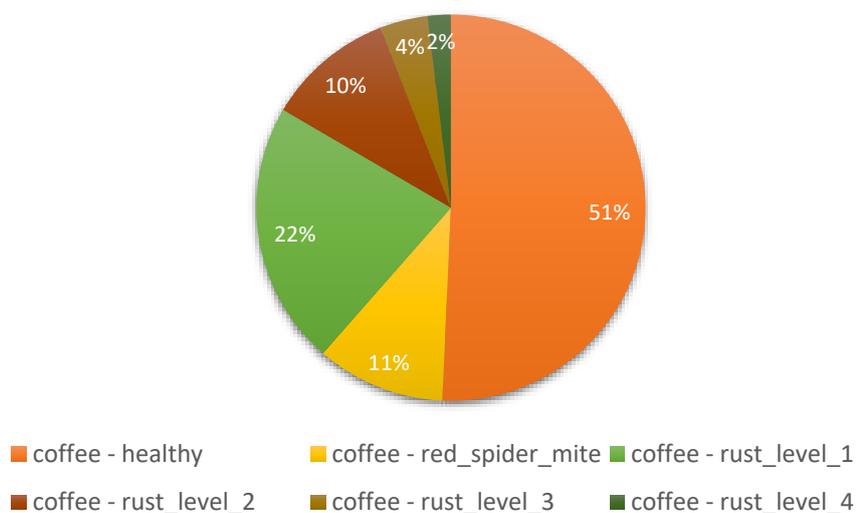


Figure 3. Data Imbalance in FGVCx\_cassava dataset (dataset\_4).

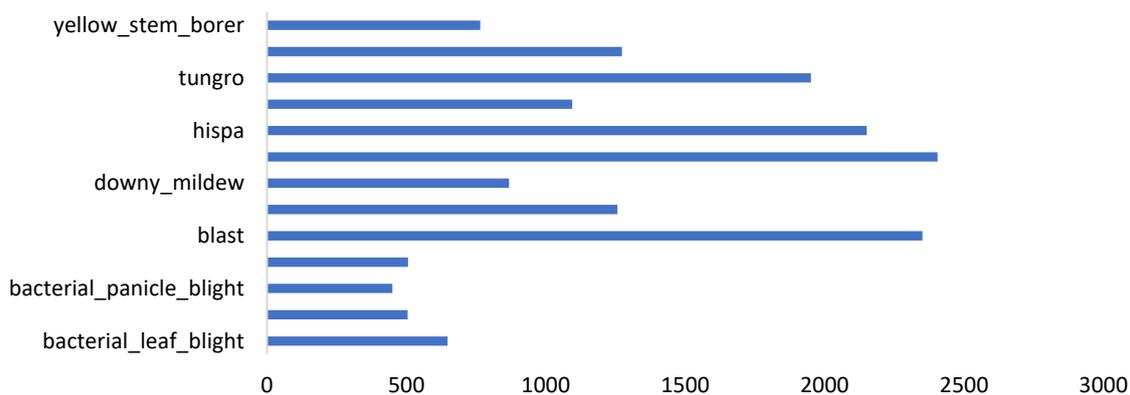


Figure 4. Image data Imbalance in paddy\_doctor dataset (dataset\_5).

Table 2. Factors fostering fine-grained extraction style in plant disease detection research.

Factors	Effects	Source
External factors such as uneven lighting, extensive occlusion, and fuzzy details	Variations in the visual characteristics of affected plants.	[41]

Variations in the presence of illness and the growth of a pest	Subtle differences in the characterization of the same diseases and pests in different regions, resulting in “intra-class distinctions”.	[42]
Similarities in the biological morphology and lifestyles of subclasses of diseases and pests	Problem of “inter-class resemblance”	[39]
Background disturbances	Makes it harder to detect plant pests and diseases In actual agricultural settings	[43]

### 3.1.2. Image Resolutions

For texture analysis techniques like Gray-Level Co-occurrence Matrix (GLCM) to ensure consistent feature extraction, resolutional consistency in pictures within a dataset is essential. More precision and dependability may be achieved in the extraction of texture information when photos have consistent resolutions. The resolutional consistency of datasets 1 and 2, for example, is demonstrated by the images, which are usually scaled at 256x256 or 256x192 pixels for dataset 1 and 6000x4000 pixels for dataset 2. With all of the images in these datasets, this consistency improves the dependability of GLCM analysis.

Nevertheless, dataset\_3 exhibits a significant range of dimensions, from 2048x1152 to 1280x720 pixels. Likewise, dataset\_4 and dataset\_5 have different resolutions of 213x213 to 960x540 pixels and 1080x1440 and 1440x1080 pixels, respectively. These variations in image resolutions between datasets could lead to unpredictability in GLCM measurements, which could affect texture analysis outcomes.

The differences in image resolutions found in datasets 3, 4, and 5 may result in inconsistent texture feature extraction, which could compromise the precision and dependability of GLCM analysis. Furthermore, because the models may find it difficult to generalize across images of different resolutions, these inconsistencies could provide problems for machine learning or deep learning models trained on these datasets. In order to ensure the robustness and generalizability of texture analysis results and ensuing machine learning models in the context of plant disease detection, resolutional discrepancies must be addressed.

### 3.2. Distribution of GLCM Metrics

Images are used to generate texture characteristics called GLCM (Gray-Level Co-occurrence Matrix) metrics, which show the spatial correlations between pixel intensities. Particularly when applied to plant disease image datasets, these metrics can be quite relevant for describing images since they capture various facets of texture. The 10 GLCM metrics distribution between the 5 different datasets were captured for the purpose of this research and presented in **Error! Reference source not found.** and **Error! Reference source not found.**. A general summary of these metrics is provided:

#### Energy

Energy represents the total of the GLCM's squared elements. Greater texture complexity or a wider range of pixel pairings in the image are indicated by higher energy ratings. It is computed using equation (1).

$$\sum_{i,j}^{K=1} P(i,j)^2 \quad (1)$$

Where  $P(i,j)$  is the  $(i,j)$ th entry in GLCM. While images in dataset\_2 possess the highest energy level, dataset\_5 has the lowest energy level (Figure 5). As energy changes proportionately, there is an indication of the homogeneity in the image texture.

#### Contrast

It measures the local variations existing in an image. High contrast values suggest a great difference between pixel intensities, hence indicating a more textured surface. It is computed using equation (2).

$$\sum_{i,j}^{K-1} |i-j|^2 P(i,j) \quad (2)$$

Where  $P(i,j)$  is the  $(i,j)$ th entry in GLCM.

Contrast distribution presented **Error! Reference source not found.** (Figure 5) indicates the lowest index for dataset\_2, while dataset\_1 and dataset\_5 possess equal and highest levels. A linear variation of contrast could be further observed between dataset\_3 to dataset\_5 in a row. This could stand treasured in understanding how contrast-related features change across different datasets, potentially impelling the interpretation of plant disease images or the performance of image analysis algorithms.

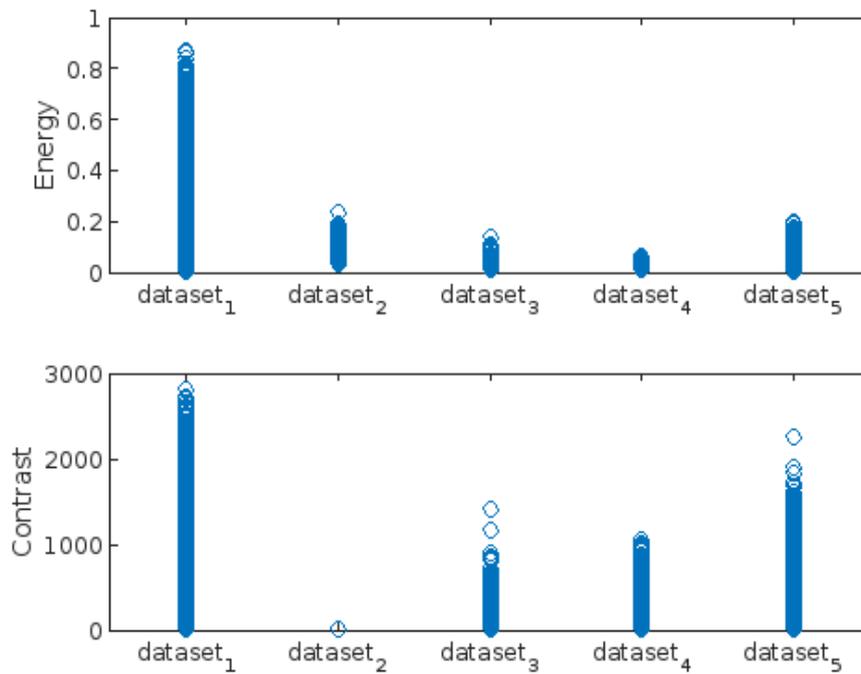


Figure 5. GLCM Metrics: Energy and Contrast.

### Correlation

It Indicates the linear dependency of gray levels in an image. High correlation values suggest a more linear association between pixel pairs. Correlation feature is computer using equation (3)

$$\sum_{i,j}^{K-1} \frac{(i,j)P(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \quad (3)$$

where  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$  and  $\sigma_y$  are the means and standard deviations of Px and Py

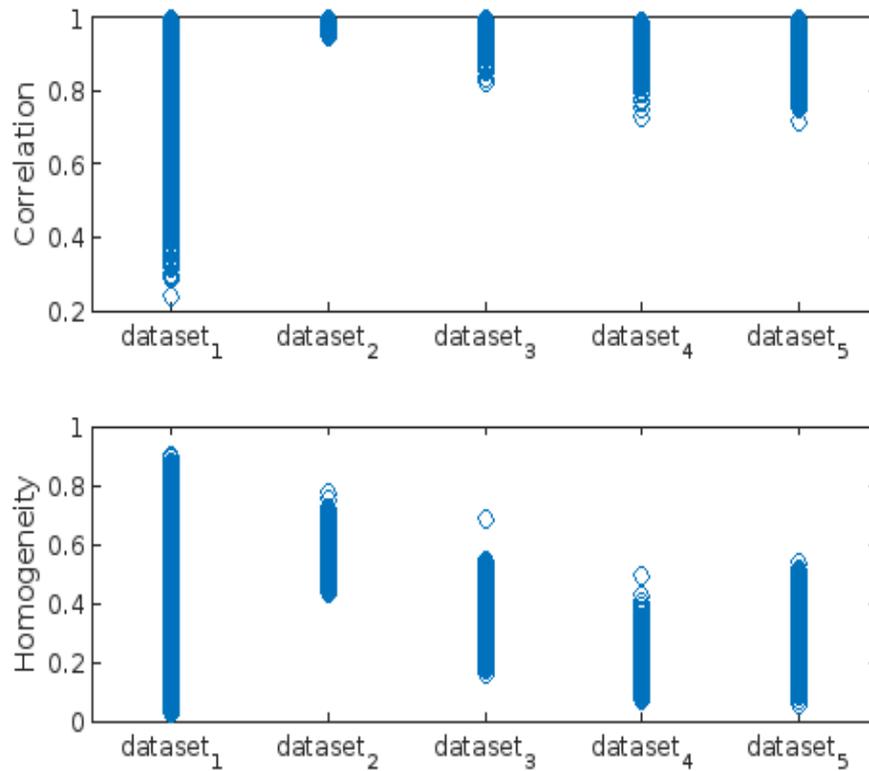
From the GLCM results distribution (Figure 6), dataset\_1 has the lowest correlation while dataset\_2 and dataset\_3 possesses the highest.

### Homogeneity:

This is an indicator that reflects the closeness of the distribution of elements in the GLCM to the GLCM diagonal. High homogeneity highlights the uniformity in the image. It is given using the equation (4) below.

$$\sum_{i,j=0}^{K-1} \frac{P(i,j)}{(1 + |i - j|^2)} \quad (4)$$

Dataset\_2 has the highest correlation while dataset\_4 shows the lowest (Figure 6).



**Figure 6.** GLCM Metrics: Correlation and Homogeneity.

Angular Second Moment (ASM):

This represents the consistency or smoothness of an image. Higher ASM values indicate a more homogeneous texture. ASM is computed using equation (5).

$$\sum_{i,j=0}^{K-1} P_{i,j}^2 \quad (5)$$

In terms of meaning, energy is somehow similar to ASM. However, dataset<sub>1</sub> and dataset<sub>2</sub> has the highest ASM, while dataset<sub>4</sub> and dataset<sub>5</sub> shows the lowest (Figure 7).

Total Variance:

This represents the variance of the GLCM as it provides an overall view of the variance in the image texture. It's related to texture complexity; higher values indicate more complexity. It is computed using equation (6) below.

$$\sum_{i,j} |i - \mu|^2 P(i,j) \quad (6)$$

GLCM distribution results (Figure 7) indicates that dataset<sub>1</sub> and dataset<sub>2</sub> possesses the highest and lowest total variance respectively.

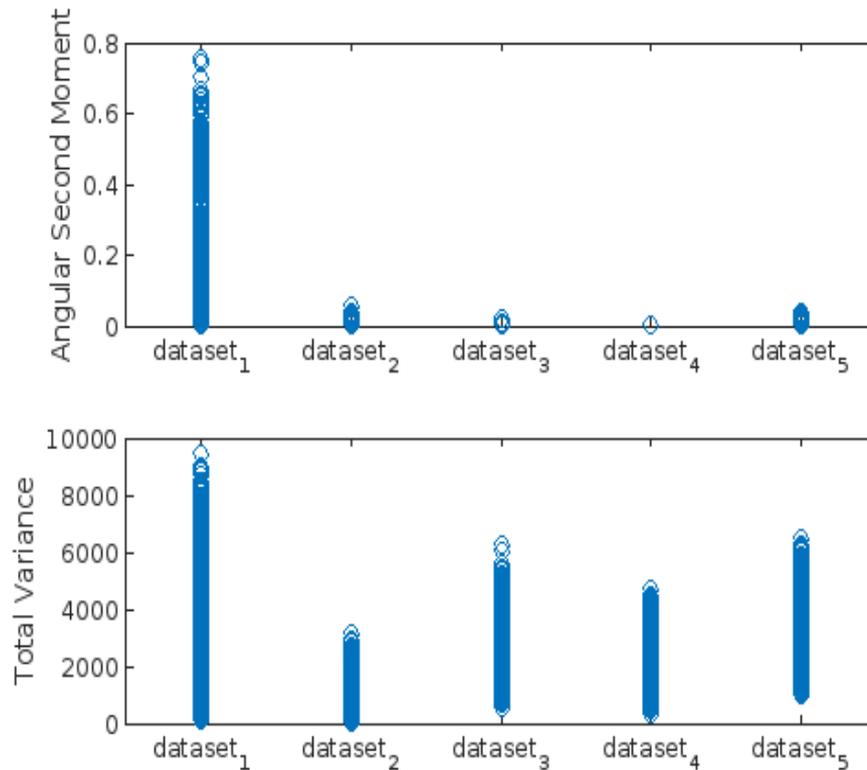


Figure 7. GLCM Metrics: ASM and Total Variance.

Maximum Probability:

This represents the most frequently occurring intensity pair in the image, as in equation (7). Higher values of maximum probability indicate a dominant texture pattern within the image sets.

$$\text{Maximum Probability} = \max(P) \quad (7)$$

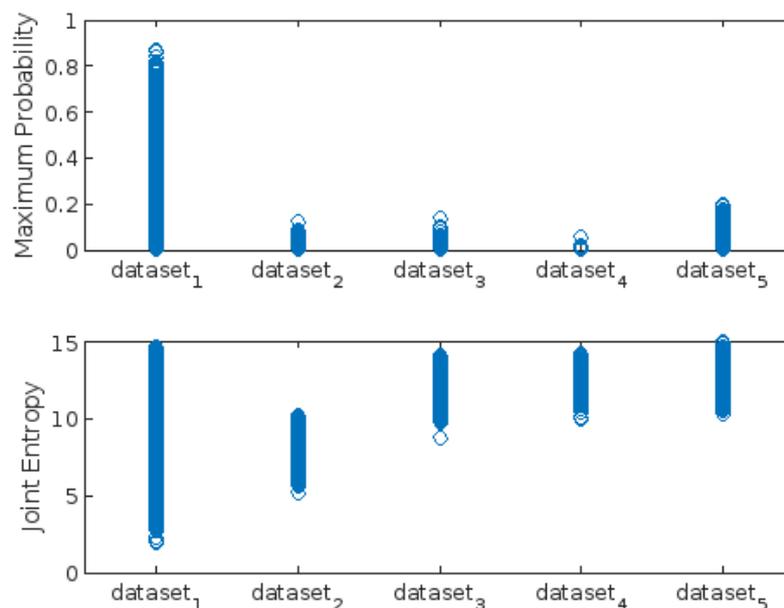
Difference Variance:

This measures the variance of the differences between adjacent pixel sets. It reveals alterations in intensity between neighboring pixels. Equation (8) shows the formula of difference variance.

$$\text{Difference Variance} = \text{variance of } P_{x-y} \quad (8)$$

Joint Entropy:

This reflects the amount of info or ambiguity present in the image. Higher joint entropy values indicate more randomness or less predictability in the image's texture (Figure 9).



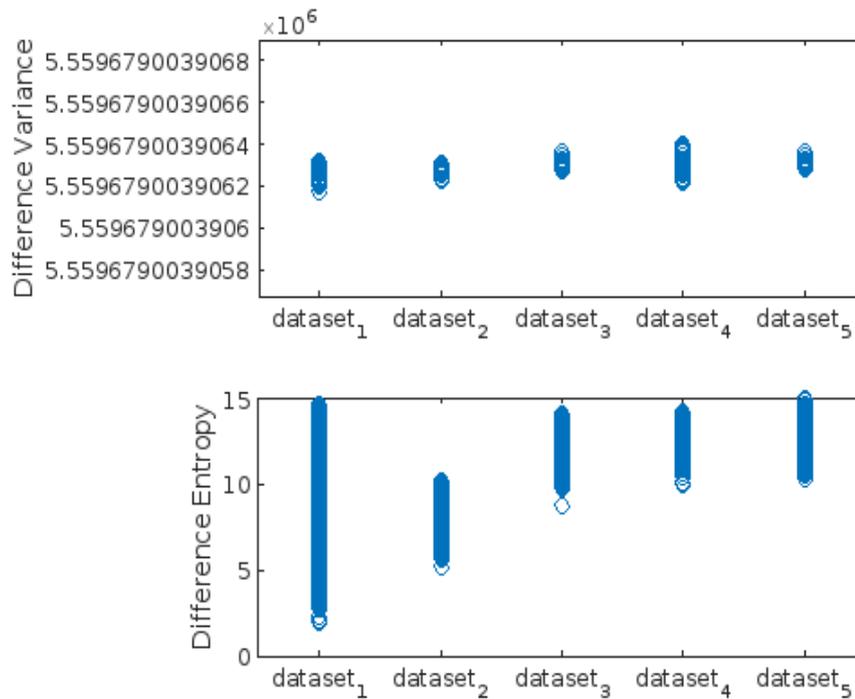


Figure 9. Joint entropy and Difference Variance.

Difference Entropy:

This reflects the randomness or unpredictability of the differences between adjacent pixel pairs. It is computed using equation (9) below.

$$\text{Difference Entropy} = \sum_{i=0}^{K-1} P_{x-y}(i) \log\{P_{x-y}(i)\} \quad (9)$$

### 3.3. Highest-Lowest GLCM Metric's Scorecard

Texture analysis for plant disease diagnosis relies heavily on image quality control and dataset selection criteria, as seen by the observed differences in GLCM metric scores between datasets. Although GLCM metrics were uniform and constant in certain datasets, they varied significantly in others, which can affect the validity of texture-based characteristics that were derived from the images. To distinguish between the performances of individual datasets, with regards to being the highest or lowest per metrics generated, the scorecard presented in **Error! Reference source not found.** was developed.

The dataset\_3 images include high-resolution variations spanning from 2048x1152 to 1280x720 pixels, indicating that image resolution may have an impact on texture analysis results. While lower resolution images may result in information loss and poorer texture differentiation, higher resolution pictures may capture finer texture features, leading to more nuanced GLCM measurements.

Table 3. GLCM Metric's scorecard.

GLCM Metrics	Highest in GLCM Metrics (dataset_x) / Lowest in GLCM Metrics (dataset_x)									
	Highest in GLCM Metrics (dataset_x)					Lowest in GLCM Metrics (dataset_x)				
	1	2	3	4	5	1	2	3	4	5
Energy		X								X
Contrast	X				X		X			
Correlation		X					X			
Homogeneity		X					X			

Angular_Second_Moment	X	X						X	X	
Total_Variance				X	X					
Maximum_Probability	X							X		
Joint_Entropy				X	X					
Difference_Variance			X		X				X	
Difference_Entropy				X	X					
<b>Sum of scores</b>	<b>3</b>	<b>4</b>	<b>0</b>	<b>1</b>	<b>4</b>	<b>3</b>	<b>4</b>	<b>0</b>	<b>2</b>	<b>3</b>

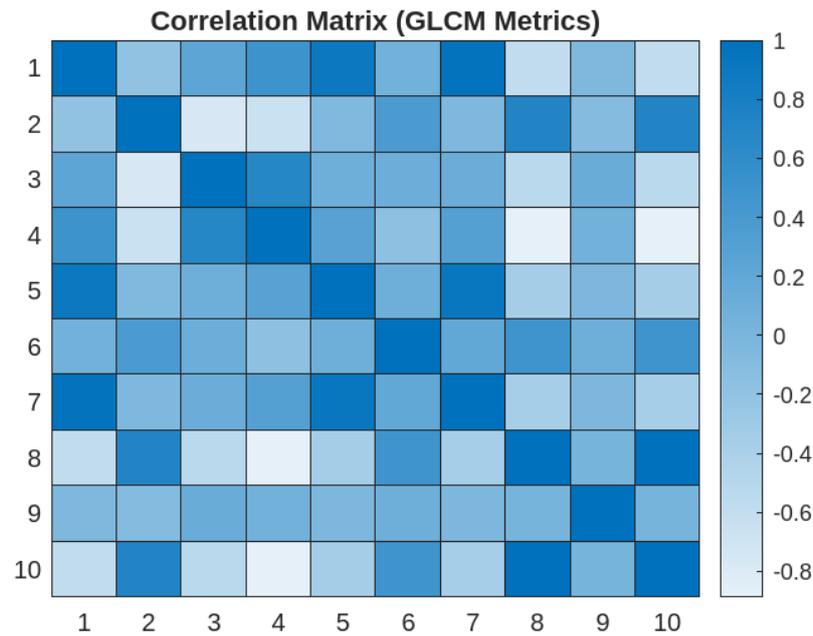
Dataset\_2 and dataset\_5 scored the highest overall performance while dataset\_2 showed the lowest overall scores per GLCM metrics.

Differences in image acquisition settings, environmental factors, and disease severity may be the cause of variations in GLCM metrics such as energy, contrast, and homogeneity between the respective datasets. For example, texture properties in images from lab-based datasets 1 and 2 might be more consistent because the images were taken under controlled conditions, while field-based datasets 3, 4, and 5 (referring to **Error! Reference source not found.**) presents more variability because of natural variations in plant physiology and environmental factors. This in turn calls for the need to conduct a more comprehensive analysis into the reasons why field-based datasets are lower in GLCM metrics score and whether this has an impact on machine/deep learning applications.

The observed discrepancies in GLCM measure scores among datasets bear significant consequences for the robustness and generalizability of machine learning and deep learning models that are trained on these datasets. When compared to models trained on datasets with significant variability in texture qualities, models trained on datasets with consistent GLCM metrics might perform better and be more generalizable. The capacity of the model to identify discriminative patterns linked to various disease classes may be hampered by bias and noise introduced into the feature space by inconsistent GLCM metrics between datasets. In practical applications, this can result in less-than-ideal performance and decreased dependability, especially when used in varied and ever-changing agricultural settings.

#### 3.4. Correlation Matrix of GLCM Metrics

To ascertain the relationships between individual GLCM metrics of these datasets, a correlation matrix was generated and presented as a heatmap for easy visualization (Figure 10). The GLCM metrics were ordered in the following series: Energy (1), Contrast (2), Correlation (3), Homogeneity (4), Angular Second Moment (5), Total Variance (6), Maximum Probability (7), Joint Entropy (8), Difference Variance (9) and Difference Entropy (10) respectively.



**Figure 10.** Correlation Matrix of the derived GLCM Metrics.

With the use of the correlation matrix, the link between variables by calculating values between -1 and 1 can be ascertained. Prior to developing any machine learning model, most data scientists believe that this is a decisive step to do as it helps determine which variables are most relevant for their model.

Strongly correlated metrics appeared darker in color. Results suggested pairwise correlation between the following metrics:

1) Strong Correlations:

- *Energy and Angular Second Moment*

Higher values of Energy are generally correlated with higher values of Angular Second Moment, according to a strong positive correlation observed. In accordance with this, images with more uniform pixel values also tend to be more homogeneous. This has the potential of helping machine learning algorithms to identify patterns or homogeneity in textures. Moreover, understanding this correlation is key for feature selection in image processing. Consequently, supposing either of the two metrics is highly representative, using both may not provide additional information.

- *Energy and Maximum Probability*

The maximum probability of pixel pairs tends to grow as the uniformity (Energy) of the image increases, due to a strong association observed. This indicates that more uniform plant disease images are likely to have a pixel pair that occurs more regularly than others. For machine learning and deep learning applications, this could impact tasks where the existence of specific pixel pairs is essential, such as in identifying unique texture patterns in certain diseases.

- *Maximum Probability and Angular Second Moment*

There appears to be a considerable association between the likelihood of a certain pixel pair recurring frequently and the homogeneity of the image. This could mean that certain patterns or textures appear frequently and consistently across the image. For applications requiring a given texture pattern to occur frequently and to be uniform, such as diagnosing diseases based on recurring patterns will be worthwhile.

- *Joint Entropy and Difference Entropy*

A substantial positive correlation suggests that as the information content (Joint Entropy) of an image increases, the randomness in intensity differences (Difference Entropy) also increases. More information-rich images may show a wider range of intensity variations. More information-rich images may show a wider range of intensity variations. This implies that the information content (Joint Entropy) of an image is related to the distribution of pixel intensities (Maximum Probability) classified within itself. Images with different distributions of pixel intensity could have more entropy. In terms of deep learning, this is cherishable for tasks where understanding both the randomness in intensity differences and the overall information content is decisive, e.g., in tasks requiring diverse texture patterns.

## 2) Moderate Correlations:

- *Contrast and Joint Entropy*

According to a reasonable correlation between Joint Entropy and Contrast, there tends to be a corresponding increase in the intensity difference between adjacent pixels (Contrast) as the individual plant disease image's information content rises (greater Joint Entropy). This correlation further implies that images with higher entropy (more varied pixel pair intensities) also seem to have more noticeable contrasts in intensity between neighboring pixels.

This correlation may be useful for activities where it is important to comprehend both the overall information content and the fluctuations in local intensity. For example, it could be helpful to capture different texture patterns at the global and local levels in disease identification.

- *Contrast and Difference Entropy*

A moderate link has been found between Difference Entropy and Contrast, indicating that a rise in the unpredictability of intensity differences between pixels (higher Difference Entropy) is accompanied by an increase in the intensity difference between a pixel and its neighbors (Contrast). According to this correlation, images with more diverse intensity differences between individual pixels also typically exhibit more pronounced intensity differences between neighboring pixels.

The Effect on Machine Learning/Deep Learning Applications could be linked to tasks where it is necessary to capture both the local changes in intensity and the global randomness in intensity differences. This correlation may aid in the identification of various texture patterns with unique local properties in the context of plant disease image analysis.

- *Correlation and Homogeneity*

A moderately positive correlation between these measures indicates that there is a tendency for stronger correlation to be associated with higher homogeneity. This possibly will suggest, in the context of machine/deep learning, that textures with more homogeneity (homogeneity) shows a relationship with higher correlations between pixel values at various spatial distances, thus signifying a texture that is more predictable.

## 3) Weak Correlations

- *Difference Entropy and Difference Variance*

The low correlation seen between these metrics may suggest that the information contained in these differences (Difference Entropy) is not highly correlated with fluctuations in pixel differences across spatial distances (Difference Variance). This could imply, in terms of machine learning, that although pixel disparities vary, they may not significantly add to the image's data content.

## 4) Inverse Correlation

- *Joint Entropy and Homogeneity*

The negative correlation suggests that homogeneity and entropy are inversely proportional. That is to say as homogeneity seems to drop as Joint Entropy rises. This could imply that images tend to be less homogeneous when their entropy is larger. This may suggest that images with a wider range of pixel intensities have less homogeneity in machine learning.

#### 5) Kruskal-Wallis test of Variance

To find out whether there are statistically significant differences between the GLCM metrics, Kruskal-Wallis test for non-parametric data was utilized. This is due to the violation of the datasets to ANOVA assumptions, including normality of data and its hypotheses, as insights from the presented distribution of the parameters across the datasets portrays. In this case a statistical significance is indicated by a p-value less than 0.0001. Given the size of the datasets, the p-value is determined to be estimated rather than exact. Moreover, the shown multiple stars signify much more of the said significance. It's true that there is a significant difference ( $P < 0.05$ ) in the medians, as well, between the 10 GLCM metrics. Kress-Wallis Statistic value estimated as 619192 is given herein as the determined test statistic value.

In light of the findings as summarized in **Error! Reference source not found.**, machine/deep learning applications could be impacted owing to the significant differences in GLCM metrics across plant disease datasets, through the feature relevance of these metrics might be decisive in distinguishing plant diseases. Integrating these metrics through advanced modellings could significantly impact classification accuracy. Also, tailoring the deep learning or machine learning algorithms, as the case may be, to accommodate these differences has the potential to enhance their performance. Such algorithms could adapt their weights or learning rates based on the dataset's specific characteristics shown by its GLCM metrics. Nevertheless, it'll be crucial to understand which metrics vary significantly alongside the individual correlations between them. This will viable help allow targeted training through a fine-tuning style or separate training for each dataset, hence optimizing their analytical power for specific diseases.

**Table 4.** Kruskal-Wallis test.

Table Analyzed	GLCW 10 parameters
Kruskal-Wallis test	
P value	<0.0001
Exact or approximate P value?	Approximate
P value summary	****
Do the medians vary signif. ( $P < 0.05$ )?	Yes
Number of groups	10
Kruskal-Wallis statistic	619192
Data summary	
Number of treatments (columns)	10
Number of values (total)	637010

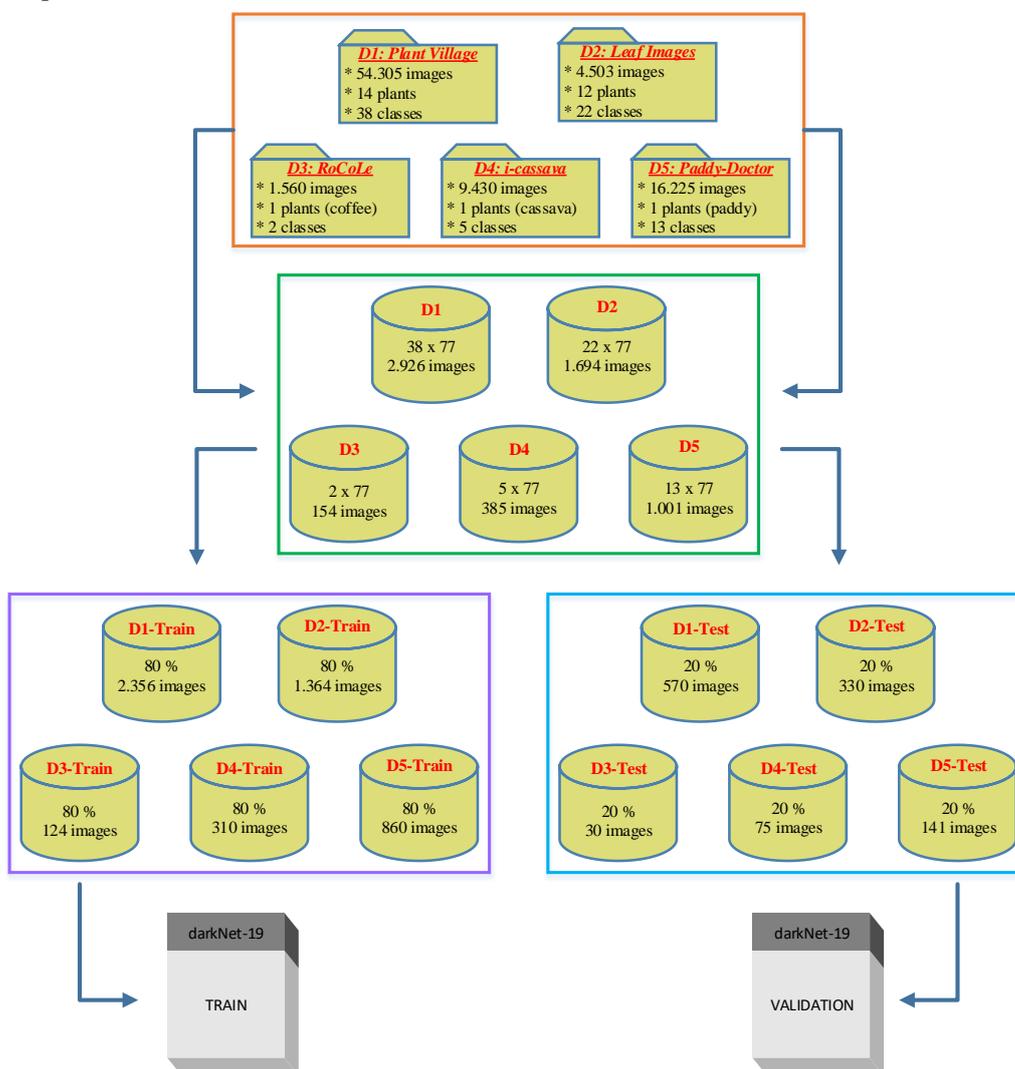
Furthermore, metrics showing substantial unevenness or variations might correlate with disease severity. The severity of plant diseases based on image features will potentially help in accurately assessing these variations if properly leveraged. This could be true for enhancing the robustness of future machine learning and deep learning model's robustness thus a better handle diverse conditions and variations observed in different diseases.

#### 3.5. Deep Learning Model's Development and Analysis

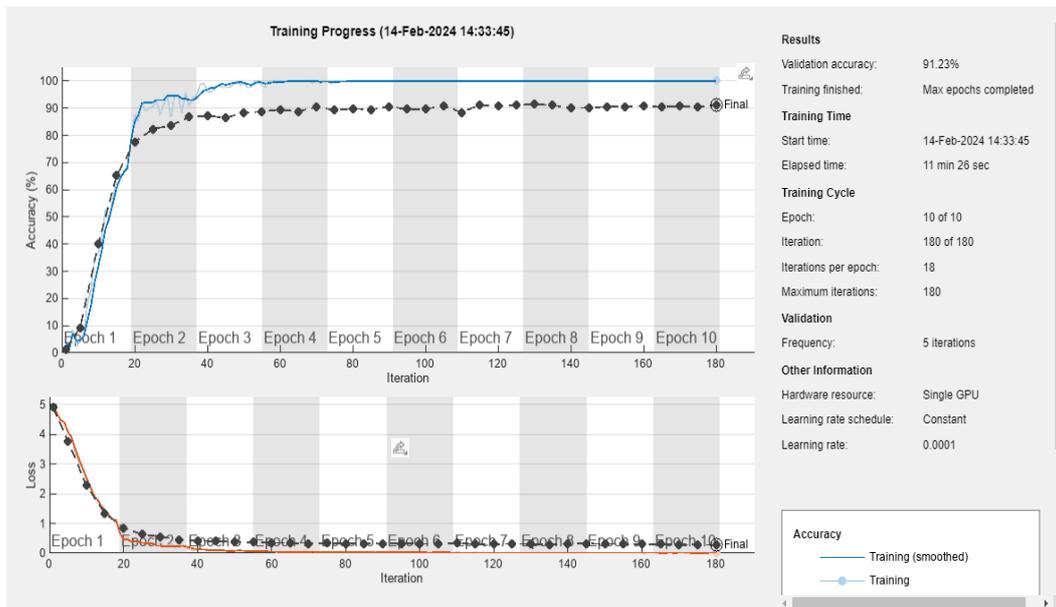
To maximize performance, the model was trained through a number of epochs (10) kept fixed for the whole datasets. The training dataset is traversed entirely in each epoch, and the model's

parameters are updated using mini-batches of data throughout each iteration. To guarantee the best results, hyperparameters including learning rate, batch size, and regularization strategies were carefully adjusted (Figure 12).

As it is known, the input data size significantly affects the model performance. Like other hyperparameters, the number of inputs should be the same for all data sets in order to make a fair evaluation. For this purpose, the class with the lowest number of samples was determined after all data sets were converted into separate image data stores with a Matlab code. Accordingly, the subclass labeled "Leamon (P10)\_diseased" of dataset D2 has the lowest number of image samples (77 images). Using this threshold value, 77 images were randomly selected from the subclasses of each dataset to obtain the final image datastores. Then, 80% of the data was randomly allocated for training and the remaining 20% for testing. The stages of preparing the datasets are shown in Figure 11. Results obtained from the deep learning training are presented in **Error! Reference source not found.** The average accuracy, precision, recall, and F1-score measures show differences in the deep learning models' performance across the five datasets (D1–D5).



**Figure 11.** The stages of preparing final datasets with equal class-image numbers.



**Figure 12.** Training Phase for D1.

Training on the D1 and D2 datasets, the deep learning model demonstrated highly encouraging results on several assessment criteria, averaging 91.22% and 90.6 % average testing accuracy, respectively. This suggests that there was a strong agreement between the models' predictions and the ground truth labels for the test samples. The models' capacity to properly classify diseased and healthy plant image samples while limiting false positives and false negatives was demonstrated by the average precision, recall, and F1-scores. The models found an effective compromise between accurately recognizing unhealthy plants (recall) and minimizing misclassifications (precision), as evidenced by the high values of average precisions, recalls, and F1-scores, respectively. Generally, when working with imbalanced datasets, the F1-score is very helpful as it provides a thorough assessment of the model's overall performance by taking into account both precision and recall [21,44].

**Table 5.** Testing Performance Metrics of the datasets in the Deep learning.

Datasets	Av. accuracy	Av. precision	Av. recall	Av. F1-score
D1	0.9122	0.9141	0.9123	0.9111
D2	0.9060	0.9116	0.9061	0.9056
D3	0.6666	0.7329	0.6667	0.6411
D4	0.5866	0.5885	0.5867	0.5867
D5	0.5897	0.5996	0.5897	0.5852

As evidenced by the confusion matrix derived for D1 (Figure 13), a 100% accuracy for 15 out of the 38 classes were recorded. The lowest was recorded as 9 accurate predictions thus indicating the suitability of both the dataset and model utilized (darkNet19) since it has minimal errors. The average accuracy of D1 and D2 being higher than that of D3, D4, and D5, respectively, implies that in terms of accurately classifying disease classifications, the models trained on datasets D1 and D2 performed better overall.



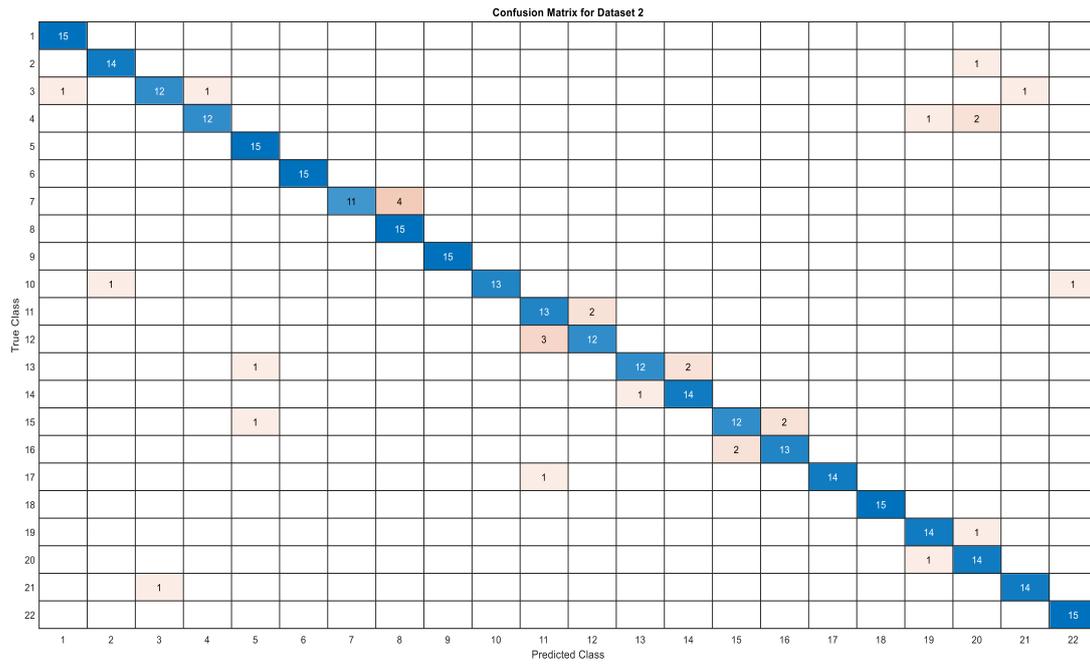


Figure 14. Confusion Matrix for D2.

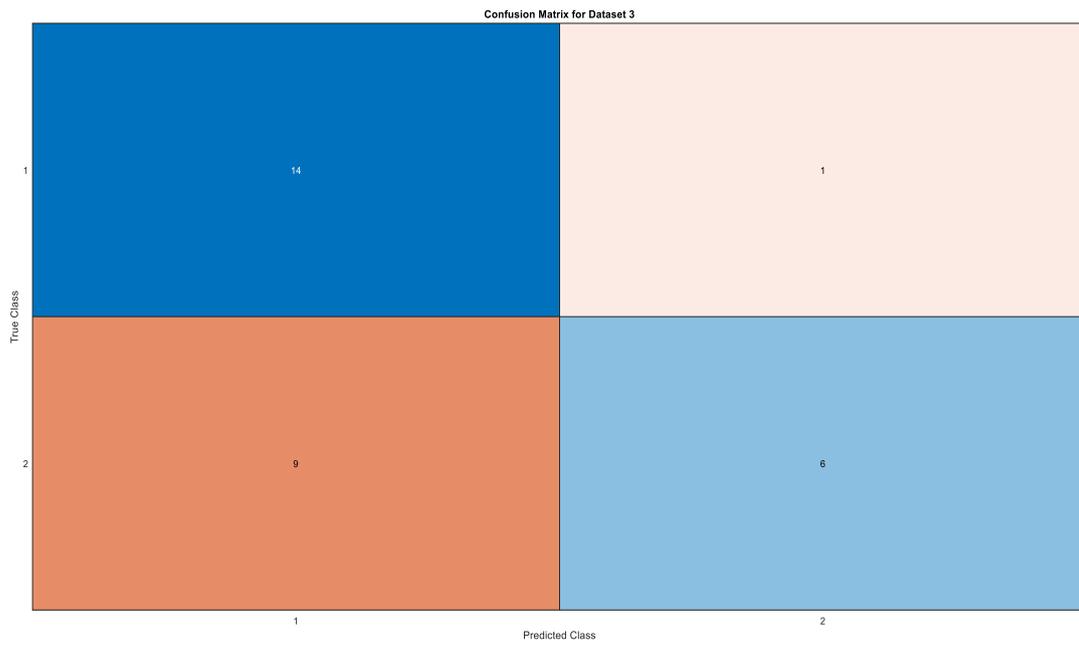


Figure 15. Confusion Matrix for D3.

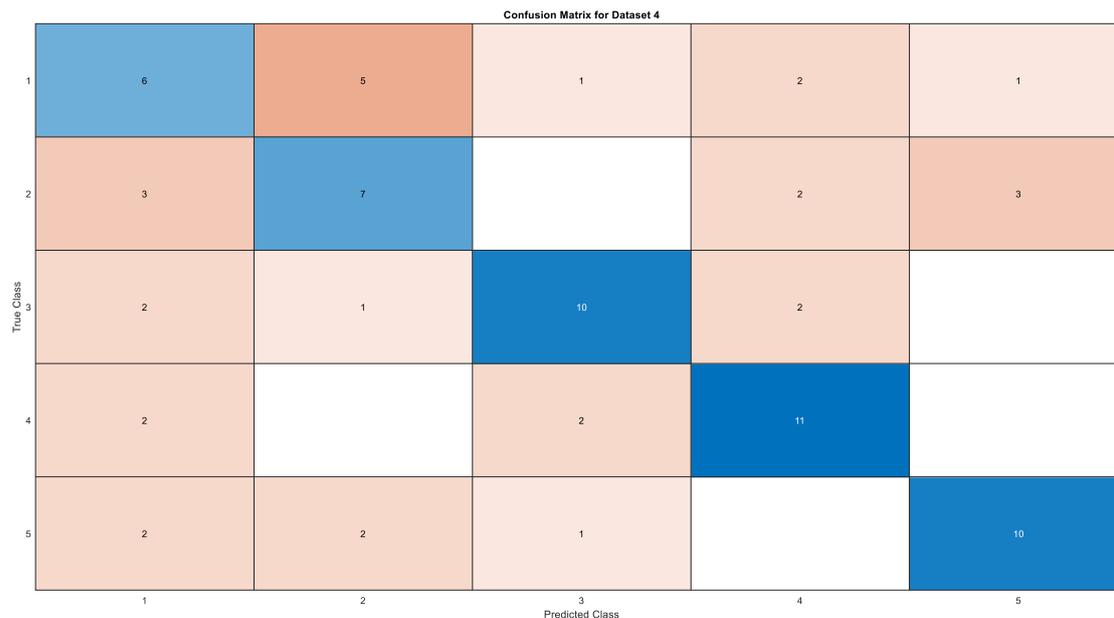


Figure 16. Confusion Matrix D4.

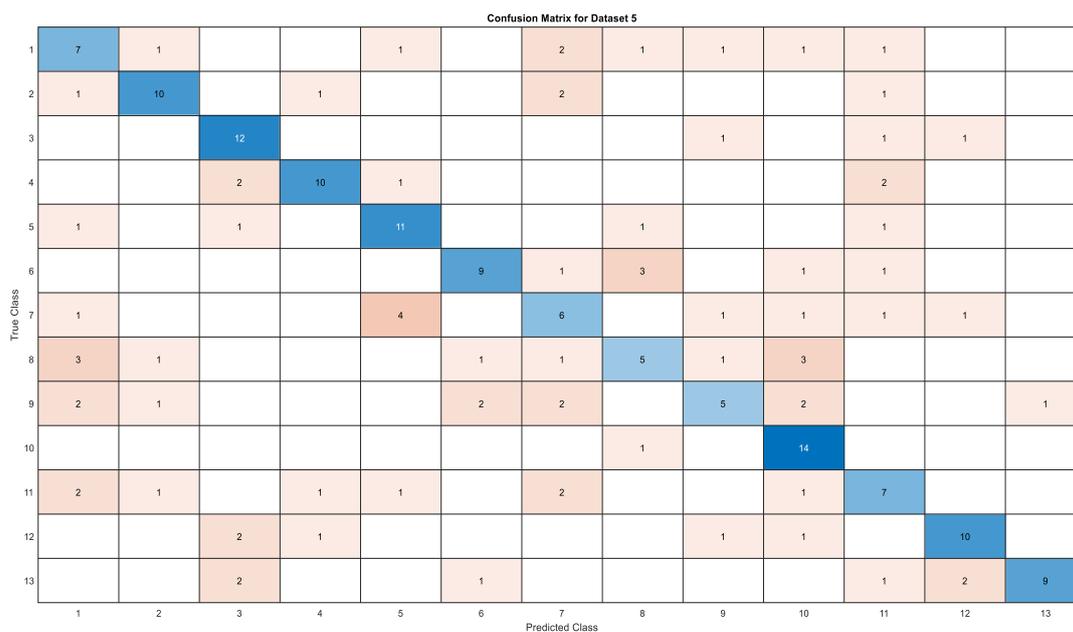


Figure 17. Confusion Matrix D5.

Comparing models trained on D3, D4, and D5 to those trained on D1 and D2, the former show greater average recall values, suggesting fewer false negatives (missed detections). The lower recall scores for D3, D4, and D5 indicate a higher proportion of false negatives, implying that the models may have failed to identify incidences of disease classes in the dataset.

The average F1-scores, which incorporate recall and precision, are greater for D1 and D2. This shows that the models trained on these datasets have a superior balance between precision and recall. The F1-score, in particular, suggests a good trade-off between precision and recall, is considered so essential for accurate disease detection [44]. It is possible that the class imbalances across these datasets are to blame for the lower F1-scores for D3, D4, and D5, which show a less than ideal balance between precision and recall.

#### 4. Conclusion

In conclusion, this study explored the complex terrain of GLCM (Gray-Level Co-occurrence Matrix) metric changes in various plant disease datasets and have concluded with a thorough investigation of their implications for deep learning applications in plant disease diagnosis. To shed light on the textural variations present, GLCM metrics were initially generated by carefully compiling a collection of plant images from five different disease datasets. Afterwards, a number of meticulous statistical analyses were carried out to identify trends and variances in GLCM measures between the datasets, providing the foundation for wise choices in later phases of this research. Specifically, these analyses, were used to assign scores to the datasets according on how well they performed on the GLCM measure, creating a quantitative standard for evaluating the quality of datasets.

Within the field of deep learning model building, our methodology stands out by a careful synchronization with the understandings obtained from the GLCM analyses. Using the DarkNet19 architecture, each dataset was subjected to rigorous model training, with particular attention dedicated to validation and hyperparameter tuning. The trained models were thoroughly evaluated using performance indicators as benchmarks for the models' ability to identify plant diseases.

Most importantly, our results revealed a strong convergence between deep learning model performance and dataset quality as measured by GLCM measures. Of all the datasets analyzed, dataset\_2 (D2) was identified as the best as it had the highest GLCM scores and the best model performance metrics. This correspondence highlights how important dataset properties are, especially texture-based features that are captured by GLCM metrics, on the efficacy of deep learning models in plant disease detection. Significantly, this study emphasizes how important it is to have a discriminating criterion when choosing datasets for deep learning applications in plant disease detection in the future. The study further emphasizes the need for a comprehensive approach to dataset curation by comparing dataset quality indicators obtained from GLCM analysis with the results of subsequent deep learning model training. To assure dataset compatibility and improve a model's performance, this means going beyond conventional visual characteristics and including texture-based descriptors like GLCM metrics.

Overall, this work provides important new understandings of the complex interactions among dataset properties, GLCM measures, and deep learning model performance in plant disease identification. By promoting a thorough and discriminating approach to dataset selection, a door is hereby opened to better decision-making and increased effectiveness in future deep learning applications, pushing the boundaries of agricultural research and supporting initiatives for sustainable food production and global food security.

## 5. Recommendations and Future Works

1. Reference to the GLCM Metrics distribution across the datasets, the results emphasize how crucial it is to carry out a more thorough investigation in order to comprehend the fundamental causes of the variations in GLCM metric scores between datasets that are field- and lab-based. The possible effects of these variations on the effectiveness of deep learning or machine learning applications in the identification of plant diseases are also called into question. To improve texture analysis techniques' resilience and dependability as well as their use in practical situations, these issues might need to be resolved.
2. Owing to the limitations of this study in terms of generalization, the findings may be constrained by the specific plant disease datasets and deep learning methodologies employed. Generalizing the results to broader contexts, diverse plant classes or other datasets deserves caution and further authentication.
3. Deep learning models trained in this work may not be as robust or as generalizable due to the inherent biases and constraints of the chosen plant disease datasets, which include differences in image quality, disease severity, and class diversity. Acquiring datasets on plant diseases from multiple sources and repositories in order to encompass a wider range of plant species, diseases, and environmental factors will be great development. Furthermore, to maintain consistency and comparability across various datasets, standardizing the procedures for acquiring such images, annotating them, and curating the data is recommended. This entails developing standards for camera settings, illumination, image resolution, and disease severity rating.

4. The incorporation of multispectral imaging techniques may allow for the acquisition of spectral information beyond the visible spectrum, thereby improving disease detection capacities and resilience to environmental variability.
5. Future research could investigate the integration of other texture descriptors, such as Gabor filters or Local Binary Patterns (LBP), in addition to the GLCM metrics that were the focus of our study on texture analysis. This multi-modal method could lead to improved model performance by offering a more thorough analysis of texture features in plant disease images.
6. Improving model generalizability and guaranteeing dependable performance across various imaging conditions and illness scenarios require addressing dataset biases and variability in texture features. To improve the accuracy and dependability of automated plant disease detection systems, future research should concentrate on creating strong feature extraction techniques and model architectures that can adjust to changes in textural characteristics and environmental factors.

**Author Contributions:** Conceptualization, M.K., F.U., and S.E.; methodology, S.E.; M.K., F.U., S.E. T.C.A., and A.A.M.M., validation, S.E.; M.K., F.U., S.E. T.C.A., and A.A.M.M.; formal analysis, X.X.; investigation, X.X.; resources M.K., F.U., and S.E.; data curation, M.K., F.U., and S.E.; writing—original draft preparation, S.E.; M.K., F.U., S.E., T.C.A., and A.A.M.M.

**Funding:** This research received no external funding.

**Data Availability Statement:** We encourage all authors of articles published in MDPI journals to share their research data. In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Where no new data were created, or where data is unavailable due to privacy or ethical restrictions, a statement is still required. Suggested Data Availability Statements are available in section “MDPI Research Data Policies” at <https://www.mdpi.com/ethics>.

**Acknowledgments:** This study was supported by Firat University, ADEP Project No. 23.23.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. FAO. The State of Food and Agriculture, 1974, *Lancet* 1975, 306, 313–314. [https://doi.org/10.1016/S0140-6736\(75\)92740-3](https://doi.org/10.1016/S0140-6736(75)92740-3).
2. FAO. FAO - News Article: New standards to curb the global spread of plant pests and diseases. *Fao*, (2018), 2020. <https://www.fao.org/news/story/en/item/1187738/icode/> (accessed August 14, 2022).
3. Horst, R.K. Plant Diseases and Their Pathogens. *Westcott’s Plant Dis. Handb.* 2001, 65–530. [https://doi.org/10.1007/978-1-4757-3376-1\\_3](https://doi.org/10.1007/978-1-4757-3376-1_3).
4. FAO. International Year of Plant Health – Final report. FAO 2021. <https://doi.org/10.4060/cb7056en>.
5. Buja, I.; Sabella, E.; Monteduro, A.G.; Chiriaco, M.S.; De Bellis, L.; Luvisi, A.; Maruccio, G. Advances in plant disease detection and monitoring: From traditional assays to in-field diagnostics. *Sensors* 2021, 21, 1–22. <https://doi.org/10.3390/s21062129>.
6. Strange, R.N.; Scott, P.R. Plant Disease: A Threat to Global Food Security. *Annu. Rev. Phytopathol.* 2005, 43, 83–116. <https://doi.org/10.1146/annurev.phyto.43.113004.133839>.
7. Witten, I.H.; Cunningham, S.; Holmes, G.; McQueen, R.J.; Smith, L.A. Practical Machine Learning and its Potential Application to Problems in Agriculture. In *Proc New Zeal. Comput. Conf., Department of Computer Science, University of Waikato, 1993*, pp. 308–325. <https://researchcommons.waikato.ac.nz/handle/10289/9915> (accessed July 6, 2023).
8. Shrivastava, V.K.; Pradhan, M.K. Rice plant disease classification using color features: a machine learning paradigm. *J. Plant Pathol.* 2021, 103, 17–26. <https://doi.org/10.1007/s42161-020-00683-3>.
9. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* 2015, 521, 436–444. <https://doi.org/10.1038/nature14539>.
10. Tibdewal, M.N.; Kulthe, Y.M.; Bharambe, A.; Farkade, A.; Dongre, A. Deep Learning Models for Classification of Cotton Crop Disease Detection. *Zeichen J.* 2022.
11. Prakash, N.; Udayakumar, E.; Kumareshan, N. Design and development of Android based Plant disease detection using Arduino. In *2020 7th Int. Conf. Smart Struct. Syst. ICSSS 2020*, Institute of Electrical and Electronics Engineers Inc., 2020. <https://doi.org/10.1109/ICSSS49621.2020.9202278>.
12. A, P.; S, B.K.; Murugan, D. Paddy Leaf diseases identification on Infrared Images based on Convolutional Neural Networks. 2022. <https://doi.org/10.48550/arxiv.2208.00031>.

13. Ahmad, A.; Saraswat, D.; El Gamal, A. A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. *Smart Agric. Technol.* 2023, 3, 100083. <https://doi.org/10.1016/j.atech.2022.100083>.
14. Hughes, D.P.; Salathe, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. 2015. <http://arxiv.org/abs/1511.08060> (accessed July 24, 2022).
15. Pardede, H.F.; Suryawati, E.; Zilvan, V.; Ramdan, A.; Kusumo, R.B.S.; Heryana, A.; Yuwana, R.S.; Krisnandi, D.; Subekti, A.; Fauziah, F.; Rahadi, V.P. Plant diseases detection with low resolution data using nested skip connections. *J. Big Data.* 2020. <https://doi.org/10.1186/s40537-020-00332-7>.
16. Thakur, P.S.; Sheorey, T.; Ojha, A. VGG-ICNN: A Lightweight CNN model for crop disease identification. *Multimed. Tools Appl.* 2023, 82, 497–520. <https://doi.org/10.1007/s11042-022-13144-z>.
17. Nagi, R.; Tripathy, S.S. Plant disease identification using fuzzy feature extraction and PNN. *Signal, Image Video Process.* 2023, 1–7. <https://doi.org/10.1007/s11760-023-02499-x>.
18. Hanh, B.T.; Manh, H.V.; Nguyen, N.V. Enhancing the performance of transferred efficientnet models in leaf image-based plant disease classification. *J. Plant Dis. Prot.* 2022, 129, 623–634. <https://doi.org/10.1007/s41348-022-00601-y>.
19. Wiesner-Hanks, T.; Stewart, E.L.; Kaczmar, N.; Dechant, C.; Wu, H.; Nelson, R.J.; Lipson, H.; Gore, M.A. Image set for deep learning: Field images of maize annotated with disease symptoms. *BMC Res. Notes.* 2018, 11, 1–3. <https://doi.org/10.1186/s13104-018-3548-6>.
20. Barbedo, J.G.A.; Koenigkan, L.V.; Halfeld-Vieira, B.A.; Costa, R.V.; Nechet, K.L.; Godoy, C.V.; Junior, M.L.; Patricio, F.R.A.; Talamini, V.; Chitarra, L.G.; Oliveira, S.A.S.; Ishida, A.K.N.; Fernandes, J.M.C.; Santos, T.T.; Cavalcanti, F.R.; Terao, D.; Angelotti, F. Annotated plant pathology databases for image-based detection and recognition of diseases. 2018. <https://doi.org/10.1109/TLA.2018.8444395>.
21. Pérez-Enciso, M.; Zingaretti, L.M. A Guide for Using Deep Learning for Complex Trait Genomic Prediction. *Genes (Basel).* 2019, 10, 553. <https://doi.org/10.3390/genes10070553>.
22. Singh, D.; Jain, N.; Jain, P.; Kayal, P.; Kumawat, S.; Batra, N. PlantDoc: A dataset for visual plant disease detection. In *ACM Int. Conf. Proceeding Ser.*; Association for Computing Machinery, 2020, pp. 249–253. <https://doi.org/10.1145/3371158.3371196>.
23. Oyewola, D.O.; Dada, E.G.; Misra, S.; Damaševičius, R. Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing. *PeerJ Comput. Sci.* 2021, 7, 1–15. <https://doi.org/10.7717/peerj-cs.352>.
24. Parraga-Alava, J.; Cusme, K.; Loor, A.; Santander, E. RoCoLe: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. *Data Br.* 2019, 25, 104414. <https://doi.org/10.1016/j.dib.2019.104414>.
25. Zhang, X.; Han, L.; Dong, Y.; Shi, Y.; Huang, W.; Han, L.; González-Moreno, P.; Ma, H.; Ye, H.; Sobeih, T. A Deep Learning-Based Approach for Automated Yellow Rust Disease Detection from High-Resolution Hyperspectral UAV Images. *Remote Sens.* 2019, 11, 1554. <https://doi.org/10.3390/rs11131554>.
26. Bhakta, I.; Phadikar, S.; Majumder, K. Thermal Image Augmentation with Generative Adversarial Network for Agricultural Disease Prediction. In *Lect. Notes Networks Syst.*; Springer Science and Business Media Deutschland GmbH, 2022, pp. 345–354. [https://doi.org/10.1007/978-981-19-3089-8\\_33](https://doi.org/10.1007/978-981-19-3089-8_33).
27. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 2016. <https://doi.org/10.3389/fpls.2016.01419>.
28. Chen, M.; French, A.P.; Gao, L.; Ramcharan, A.; Hughes, D.P.; McCloskey, P.; Baranowski, K.; Mbilinyi, N.; Mrisho, L.; Ndalaha, M.; Legg, J.; Hughes, D.P. A Mobile-Based Deep Learning Model for Cassava Disease Diagnosis. *Front. Plant Sci.* 2019, 10, 272. <https://doi.org/10.3389/fpls.2019.00272>.
29. Johannes, A.; Picon, A.; Alvarez-Gila, A.; Echazarra, J.; Rodriguez-Vaamonde, S.; Navajas, A.D.; Ortiz-Barredo, A. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Comput. Electron. Agric.* 2017, 138, 200–209. <https://doi.org/10.1016/j.compag.2017.04.013>.
30. Ahmad, J.; Jan, B.; Farman, H.; Ahmad, W.; Ullah, A. Disease detection in plum using convolutional neural network under true field conditions. *Sensors* 2020, 20, 1–18. <https://doi.org/10.3390/s20195569>.
31. Petchiammal; Kiruba, B.; Murugan, P. Arjunan; Paddy Doctor: A Visual Image Dataset for Automated Paddy Disease Classification and Benchmarking. In *\*Proc. 6th Jt. Int. Conf. Data Sci. Manag. Data (10th ACM IKDD CODS 28th COMAD)*, ACM, New York, NY, USA, 2023, pp. 203–207. <https://doi.org/10.1145/3570991.3570994>.
32. Mikołajczyk, A.; Grochowski, M. Data augmentation for improving deep learning in image classification problem. In *2018 Int. Interdiscip. PhD Work. IIPHDW 2018*; Institute of Electrical and Electronics Engineers Inc., 2018, pp. 117–122. <https://doi.org/10.1109/IIPHDW.2018.8388338>.
33. Velásquez, A.C.; Castroverde, C.D.M.; He, S.Y. Plant-Pathogen Warfare under Changing Climate Conditions. *Curr. Biol.* 2018, 28, R619–R634. <https://doi.org/10.1016/j.cub.2018.03.054>.
34. Wang, Q.; He, G.; Li, F.; Zhang, H. A novel database for plant diseases and pests classification. *ICSPCC 2020 - IEEE Int. Conf. Signal Process. Commun. Comput. Proc.* 2020. <https://doi.org/10.1109/ICSPCC50002.2020.9259502>.

35. Gadkari, D. Image Quality Analysis Using GLCM. *Electron. Theses Diss.* **2004**, 1–120. <https://stars.library.ucf.edu/etd/187>.
36. Hall-Beyer, M. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *Int. J. Remote Sens.* **2017**, *38*, 1312–1338. <https://doi.org/10.1080/01431161.2016.1278314>.
37. Mall, P.K.; Singh, P.K.; Yadav, D. GLCM based feature extraction and medical X-RAY image classification using machine learning techniques. *2019 IEEE Conf. Inf. Commun. Technol. CICT 2019.* **2019**. <https://doi.org/10.1109/CICT48419.2019.9066263>.
38. Kadir, A. A Model of Plant Identification System Using GLCM, Lacunarity And Shen Features. *Publ. Res. J. Pharm. Biol. Chem. Sci.* **2014**, *5*, 1–10. <https://arxiv.org/abs/1410.0969v1>.
39. Shoaib, M.; Shah, B.; El-Sappagh, S.; Ali, A.; Ullah, A.; Alenezi, F.; Gechev, T.; Hussain, T.; Ali, F. An advanced deep learning models-based plant disease detection: A review of recent research. *Front. Plant Sci.* **2023**, *14*, 1158933. <https://doi.org/10.3389/fpls.2023.1158933>.
40. Gupta, H.P.; Chopade, S.; Dutta, T. Computational Intelligence in Agriculture. In *Emerg. Comput. Paradig.*; John Wiley & Sons, Ltd, **2022**, pp. 125–142. <https://doi.org/10.1002/9781119813439.ch7>.
41. Wang, G.; Sun, Y.; Wang, J. Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. *Comput. Intell. Neurosci.* **2017**, **2017**. <https://doi.org/10.1155/2017/2917536>.
42. Barbedo, J.G.A. Factors influencing the use of deep learning for plant disease recognition. *Biosyst. Eng.* **2018**, *172*, 84–91. <https://doi.org/10.1016/j.biosystemseng.2018.05.013>.
43. Barbedo, J.G.A. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput. Electron. Agric.* **2018**, *153*, 46–53. <https://doi.org/10.1016/j.compag.2018.08.013>.
44. Theodoridis, S. *Machine Learning: A Bayesian and Optimization Perspective*, Second Edition, **2020**. <https://doi.org/10.1016/C2019-0-03772-7>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.