DOI: 10.24996/ijs.2025.66.12.39

ISSN: 0067-2904

Plant Disease Detection Using Siamese Networks

Rabab Farhan Abbas
Department of Computer Science, University of Technology – Iraq, Baghdad, Iraq

Received: 10/9/2024 Accepted: 3/12/2024 Published: 30/12/2025

Abstract

Plant sicknesses pose big challenges to crop production and food safety. Early and correct disease prognosis is crucial for powerful disease manipulation and minimizing losses. Despite the superiority of traditional strategies like visual inspection, technological advancements provide new opportunities. This paper explores the software of deep studying equipment, especially Siamese networks, for plant sickness identity in pix. Siamese networks make use of shared weights to study similarities among wholesome and diseased plant images, permitting powerful discrimination based totally on visual characteristics. This approach minimizes the need for handcrafted functions and may generalize to new diseases, making it adaptable throughout various agricultural settings. Real-time photograph processing facilitates early disease detection and intervention strategies. Comparative analysis with traditional strategies, along with CNNs and SVMs demonstrates the effectiveness of Siamese networks in plant disorder detection. Overall, this study showcases the practical application of Siamese networks for correct and well timed plant disease identity in agriculture, offering flexibility, performance, and flexibility to new sickness eventualities.

Keywords: Plant diseases, Siamese Networks, plant images, food security, tomatoes.

الكشف عن أمراض النباتات باستخدام الشبكات السيامية

رباب فرحان
قسم علوم الحاسوب، الجامعة التكنولوجية - العراق، بغداد، العراق

الخلاصة

تشكل أمراض النبات تحديات كبيرة لإنتاج المحاصيل والأمن الغذائي. يعد التشخيص المبكر والدقيق للأمراض أمرًا بالغ الأهمية للسيطرة الفعالة على الأمراض وتقليل الخسائر. على الرغم من انتشار التقنيات التقليدية مثل الفحص البصري، فإن التقدم التكنولوجي يوفر إمكانيات جديدة. يستكشف هذا البحث تطبيق أدوات التعلم العميق، وخاصة الشبكات السيامية، لتحديد أمراض النبات في الصور. تستخدم الشبكات السيامية الأوزان المشتركة لتعلم أوجه التشابه بين صور النباتات الصحية والمريضة، مما يتيح التمييز الفعال بناءً على الخصائص المرئية. يقلل هذا النهج من الحاجة إلى الميزات المصنوعة يدويًا ويمكن تعميمها على أمراض جديدة، مما يجعلها قابلة للتكيف عبر مختلف البيئات الزراعية. تسهل معالجة الصور في الوقت الفعلي الكشف المبكر عن الأمراض واستراتيجيات التدخل. يوضح التحليل المقارن بالطرق التقليدية مثل CNN وSVMs فعالية الشبكات السيامية في اكتشاف أمراض النبات. بشكل عام، يبرز هذا البحث الفائدة العملية للشبكات السيامية لتحديد أمراض النبات بدقة وفي الوقت المناسب في الزراعة، مما يوفر المرونة والكفاءة والقدرة على التكيف مع سيناريوهات الأمراض الجديدة.

*Email: rabab.f.abbas@uotechnology.edu.iq

1. Introduction

Prevailing plant diseases are a thorn in the flesh of the agricultural industries since they are likely to affect food security besides having an economic impact on the market. The identification of diseases at an early stage is very important so that the spread of diseases may be controlled, and diseases do not tremendously affect the growth as well as productivity of the crop. Current methods used in the diagnosis of plant diseases are mainly the visual assessment and laboratory analysis, which are often socially demanding, tiresome, and sometimes unreliable. The progressive approaches particularly machine learning and deep learning, hold great promise for timely, efficient, and cost-efficient identification of diseases in plants. The methods that have drawn much attention in the recent past include the deep learning models, especially the Siamese networks, which are used in solving different issues in the images, for instance, recognizing, detecting as well as classifying an object [1].
Siamese networks are a convolutional neural network where two sub-networks are identical in every way but their output. The motivation of Siamese networks is to obtain a measure of similarity between two inputs. In a situation where plant diseases are to be detected, this scale will be useful in comparing features of both healthy plants and diseased ones or different types of diseased plants. The objectives of this study involves offering a comprehensive analysis of Siamese networks concerning their basics, constructions, and uses in plant disease identification. This section aims to introduce and describe the model that is the Siamese networks together with its theoretical foundation. With different architectures of Siamese networks that have been used to detect plant diseases [2].

Over the last twenty years, various deep learning models have transformed the way data, especially images and videos, are handled. Convolutional neural networks (CNNs) have provided competitive results in the different task areas that include image recognition, object detection as well as segmentation. CNNs learn hierarchical representations of images by extracting features at different scales and levels. However, it is not designed to compare or recognize the similarity between two input images. Siamese networks have been proposed to address this problem by learning a metric that captures the similarity between pairs of inputs [3] and [4]

The idea behind Siamese networks is to train two identical subnetworks with common weights to compare the similarity between input pairs. The subnetworks take two input images, process them separately, and output a similarity score based on the extracted features. During training, subnetworks are updated based on triangular loss or other similarity-based loss functions to decrease the distance between similar pairs and increase the distance between dissimilar pairs.

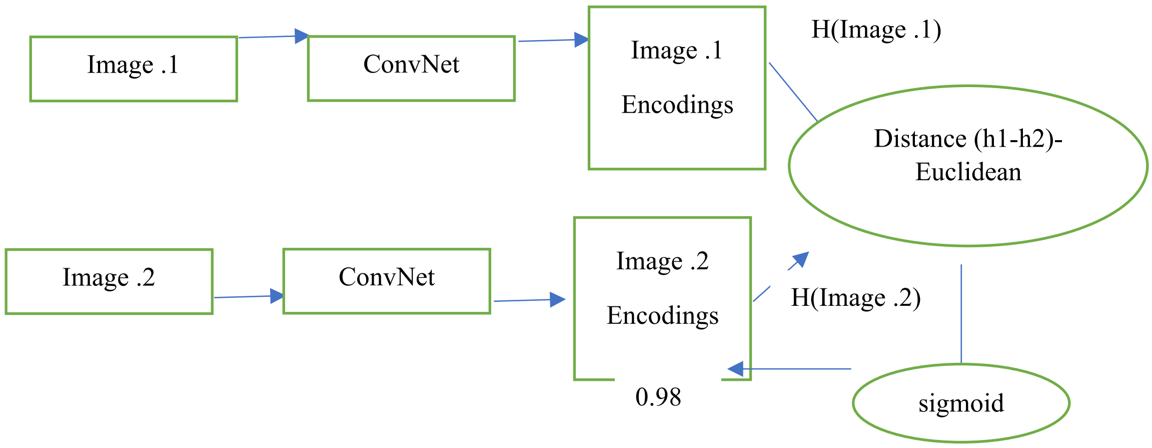

Figure 1: Basic Siamese network architecture [4]

Siamese networks have been applied to various problems in plant disease detection, such as early disease detection: Siamese networks have been used to detect plant diseases at an early stage by comparing features of healthy plants, and sick plants. For example, Siamese grids have also been used to classify different types of plant diseases based on their visual attributes. Siamese grids were used for disease clustering, which is the process of grouping similar plant diseases together based on their visual characteristics. For disease diagnosis and prediction, Siamese networks have been used for disease diagnosis and prediction, which involves identifying the type and severity of plant diseases based on their visual properties. Several public datasets have been created for plant disease detection, providing a benchmark for researchers to evaluate and compare different deep learning models. Some popular plant disease datasets include Market1501 [5], PlantCLEF2015 [6], DETRAC [7], and Pepper [8] datasets. These datasets provide annotated images of healthy and diseased plants, along with their corresponding labels and annotations. Evaluation metrics used to evaluate the performance of plant disease detection models include precision, recall, F1 score, area under the receiver operating characteristic curve (AUC-ROC), and mean precision (mAP).
In this paper, we targeted a set of commonplace plant life and diseases like tomato, wherein diseases such as powdery mildew and brown spots have been protected as take a look at fashions to evaluate the effectiveness of Siamese networks in distinguishing inflamed from healthful vegetation. This evaluation offers a deeper understanding of the wonderful sickness styles and opens the way for expanding the programs of automatic disease diagnosis in more than one form of agricultural plants.

2. Related work

Research in the agricultural domain aims to enhance both the quality and quantity of agricultural products at reduced costs and higher profits. However, agricultural product quality can be degraded by plant diseases, which are often caused by fungi, bacteria, and viruses. Early detection and classification of plant diseases is important for effective treatment and prevention of further damage. Traditional methods require constant expert monitoring, which can be time-consuming and expensive for farmers. To address these challenges, various systems using image processing and automatic classification tools have been proposed.
Suhaili Kutty et al. [9] designed a method for classifying Anthracnose and Downey Mildew, watermelon leaf diseases. The process involved identifying the region of interest (ROI) using the RGB color component, reducing noise using median filtering, and Classification using neural network models. The proposed method achieved an accuracy of 7

5.9% based on RGB color averages. Sanjeev Sannaki et al. [10] intended to make a diagnosis of grape leaf diseases with the help of image processing and the usage of artificial intelligence. They chose downy mildew and powdery mildew and masking, Anisotropic Diffusion, and k-means clustering for segmentation and feature extraction through GLCM. The classifier used was the Feed Forward Back Propagation Network; the feature that gave a higher accuracy was the Hue.
Another study for the classification and detection of rose leaf diseases – black spot and anthracnose was performed by Akhtar et al. [11] that used Support Vector Machines (SVM). The segmentation was done with Thresholding and Otsu's method, while the features indicated included DWT features, DCT, and Textural Haralick features were extracted and analyzed using the SVM model. From the resulting accuracy, it was apparent that the processes were efficient.
S. Dubey and R. Jalal rightly identified and classified Banana fruit diseases and diseases that affected the apple fruit, including scab, apple rot, and apple blotch [12]. To accomplish the segmentation, the K-means clustering algorithm was employed while feature extraction and classification were done using Multiclass SVM.
A technique for identifying the diseases on tomato leaves was described by Usama Mokhtar, et al. [13] Here, the image processing approaches applied included the Gabor wavelet transformation and feature extraction with different SVM Kernels for classification. Powdery mildew and Early blight were distinguished as the diseases present in the plant samples.
Sachin Khirade and A.B. Patil [14] discussed the major steps involved in detecting and classifying plant diseases using image processing, which include image acquisition, preprocessing, segmentation, feature extraction, and classification. Among the segmentation methods, the k-means clustering method produced the most accurate results.

Bhog and Pawar [15] utilized a neural network for cotton leaf disease classification analysis. K - means unity, with Euclidean distance used for segmentation for recognition accuracy of (89.56%) and a shorter execution time (436.95 seconds), .
Ms. Kiran R. Gavhale et al. [16] presented various image processing techniques to extract diseased parts of citrus leaves. Preprocessing involved DCT domain image enhancement and color space conversion, followed by segmentation using k-means clustering, feature extraction using GLCM Matrix, and classification using SVM with radial basis kernel and polynomial kernel.
These researchers have delivered tremendous ideas in creating advanced approaches for the prediction and identification, follow up of plant diseases in farms reducing overall effect and possible monetary loss. Still, there is a need for constant improvement of the current systems to enhance in accuracy, effectiveness, and sustainability of these systems to suit a large number of agricultural crops and diseases.

Methodology

Siamese networks, which belong to deep learning models, have attracted a lot of attention within feature learning and metric learning since the performance of these kinds of structures are very high in various applications, including face verification and twin verification in biological science. Thus, in this extensive survey of the literature, we explore the details of the design, basic and advanced architectures, as well as the Siamese networks' usage, alongside plunging into the mathematical background, the strategies, and the layers behind the Siamese networks.
1. Design Principles:
The major objective of designing Siamese networks is to learn the embedding space well-suited to preserve the relationships in the latent feature space. Key design principles include:

- Siamese architecture: This two-branch architecture tries to make the representations of the two corresponding inputs as similar as possible and as different as possible from each other.
- Metric learning loss: Various loss functions like Cosine Margin Loss [17], Triplet Loss [18], and Contrastive Loss [19] are employed to learn the embeddings that preserve the desired similarity relationships.

2. Siamese Network Architectures:

Siamese networks have been proposed in various forms to address diverse applications, comprising several layers, including:
- Deep Siamese Networks: These networks apply simple Siamese architectures with multiple layers and hidden units [20] and [21].
- Triplet Siamese Networks: Triplet Siamese Networks [22] incorporate an additional anchor vector to preserve the relationships among three input vectors and thus improve the learned embeddings' robustness against distractor vectors.
- Multi-Dimensional Siamese Networks: These networks employ separate embedding spaces for different feature modalities, like RGB, depth, and RGB-D [16]

3. Siamese Network Layers:

Siamese networks comprise several types of layers to learn latent features effectively, such as: - Fully Connected (FC) Layers: FC layers use weighted sums of the lower-level features as inputs and learn the higher-level features.
- Convolutional or Convolutional Neural Network (CNN) Layers: CNNs present an efficient feature extraction mechanism when filters are applied to an input while employing local connections between neurons.
- Recurrent Neural Network (RNN) Layers: RNNs use recurrent connections and are effective for sequences/time-series [12]and [23].
- Attention Layers: Different kinds of attention help to work intensively with some crucial features that also give positive results in any model [15]

4. Applications of Siamese Networks:

Siamese networks have been successful in many projects and activities such as:
- Face Recognition: Siamese networks have been the center of the research in face recognition and retrieval because of the ability to handle pose variant, illumination variation, and expression variation [24].
- Twin Differentiation: Siamese networks have been used in biological domains to separate real information from fictional with the help of distinguishing between similar twins from the twin siblings [9]
- Anomaly Detection: By inputting distances, Siamese networks can be used to detect anomalies or outliers by learning for a distances-based representation space [22]- [25].

This paper, demonstrated the various design factors and structure of Siamese networks, their architecture, and the layering; also, it lays a general understanding of the concepts, plan, and results of the application of Siamese networks. Understanding the intricacies of Siamese networks will pave the way for developing more efficient and accurate Siamese models for a range of applications.

In this context, the primary mathematical equations and concepts used in Siamese networks include metric learning objectives and the structural modification of a standard neural network to adapt it for similarity learning.

5. Mathematical Framework of Siamese Networks

The mathematical basis includes objectives in metric learning and adjustments to standard neural networks for similarity learning.

1. Standard Neural Network: First, we introduce a standard fully connected neural network with weights denoted by matrices W_1, W_2, ..., W_n and biases b_1, b_2, ..., b_n. The activation

function for neurons is denoted by (σ\$. For input $X_i \in R^d$ and weights/biasesW^i, bᵢ, i \in {1, 2}, we have:

$$f_i(X_i) = \sigma(W_i x_i + b_i) \tag{1}$$

2. Siamese Network: A Siamese network comprises two identical branches, where one branch processes anchor input $X_A \in R^d$ and the other branch processes positive input. $P_X \in R^d$ We can express this as:

$$g_A = \sigma(W_g X_A + b_g) \text{ و } g_P = \sigma(W_g X_P + b_g) \tag{2}$$

where $\boldsymbol{g} \in \boldsymbol{R^m}$ is the output feature vector at the last hidden layer.

3. Similarity Measure: Typically, the cosine similarity measure is used to compute the similarity between two vectors u and v:

$$\frac{\langle u,v\rangle}{\|u\|.\|v\|} = \frac{u^T v}{\|u\|.\|v\|} = \sin(u, v) \tag{3}$$

where $\langle.,.\rangle$ denotes the inner product.

4. Metric Learning Objectives: Given the anchor feature g_A, positive feature g_P, and an optional negative feature g_N, the following are common objective functions to learn distance measure functions in Siamese networks:

a. Triplet loss: This objective function aims to keep the difference between the distance of the anchor and positive inputs and the distance of the anchor input and any negative input. Therefore, the loss function is:

$$\max(\|g_A - g_P\|^2 - \|g_A - g_N\|^2 + a, 0) = L_{\text{triplet-loss}}(X_A, X_P, X_N) \max\left(dist(g_A, g_P) - dist(g_A, g_N) + a, 0\right) \tag{4}$$

where $\boldsymbol{\alpha}$ is the margin, and dist (.,.) is the Euclidean distance between feature vectors.

b. Cosine Margin Loss: This objective function ensures that the angles between the positive and negative inputs and the anchor input are greater than the specified margin α. The loss function is:

$$\left(\alpha, 0 + g_P g_A^T - max\left(m = L_{\text{cosine-margin-loss}}\right)(X_A, X_P, X_N)\right) \tag{5}$$

where $\boldsymbol{g_N g_A^T}$ is the cosine similarity between the anchor and negative inputs.

c. Contrastive Loss: This objective function measures the dissimilarity between positive and negative input pairs. The loss function is:

$$L_{\text{contrastive-loss}}(X_A, X_P, X_N) = -log\left(\frac{e^{(X_P X_A)s}}{\sum_{\{N,\ldots\ldots,2,1\}} e^{(X_i, X_A)s}}\right) \tag{6}$$

where **s (.,.)** is the similarity score, which can be cosine similarity, Euclidean distance, or any other measure.

Proposed Work

In this section, we outline the proposed technique, consisting of the block diagram and algorithm for imposing a Siamese community :

1. Proposed Block Diagram

The following block diagram outlines the shape of the proposed Siamese community model. It consists of:

1. Data Preprocessing: Images are standardized, normalized, and converted right into an appropriate layout for the community.
2. Siamese Network: A twin-department network with a shared-weight shape to generate embeddings for anchor-superb pairs.
3. Metric Computation: Computes similarity rankings for embeddings.
4. Loss Calculation: Uses selected loss functions (e.g., Triplet Loss) to optimize embedding space.
5. Output Analysis: Outputs similarity metrics and embedding vectors for further evaluation.

2. Proposed Algorithm

The following steps outline the proposed algorithm:

1. **Data Loading and Preprocessing**:
   - Load the Plant Village dataset from TensorFlow datasets.
   - Standardize and normalize picture facts.
2. **Define Siamese Network Model**:
   - Construct the same branches with CNN layers and a final dense layer of 64 units.
   - Use a shared-weight configuration between the branches.
3. **Compute Feature Vectors**:
   - For each input pair (anchor and high-quality), extract feature vectors from the remaining dense layer.
   - Display the primary 10 snapshots and their characteristic vectors as bar charts.
4. **Calculate Similarity**:
   - Calculate the cosine similarity or Euclidean distance among characteristic vectors.
5. **Loss Optimization**:
   - Use Triplet Loss, Cosine Margin Loss, or Contrastive Loss to educate the community.
   - Apply backpropagation with SGD or Adam for optimization.
6. **Output Visualization**:
   - Display pics with function vectors for visible inspection and overall performance analysis.
   - Use bar charts to evaluate embeddings of comparable snapshots.

3. Results

The extracted features represent a compact, dense, and abstract representation of the sampled data, which includes essential information to differentiate between different types of plant diseases. By extracting features from images using a Siamese network, thus revealing image characteristics that best distinguish between different plant diseases. This is useful in diagnosing and identifying diseases. The extracted features can be used to develop models for classifying and diagnosing diseases based on their visual features.

Diseases may express the signature blooms at different moments in the progression of the illness, and the differences are in shape, color, texture, or symmetry, which the deep learning model can pick and diagnose.

Favorable for prediction without a new training set, even in the case of newly emerged diseases since the Siamese network extracts features from the new images that it has never seen before, it can easily predict the occurrence of new plant diseases different from those stated during training. The network has acquired a broad understanding of the given data space, which means that it can detect similar diseases that may be slightly different from one another.

It can be used in image search and approximate search after extraction of features, and you are able to search similar images and compare images you are interested in. For instance, instead of an example of a diseased plant, you can use such an image to search for other images of the same disease or like diseases in a large database, which in turn is useful in creating a diagnosis or for comparative analysis or even consultation with experts from the same field of study.

When using the features extracted, related plant diseases are categorized into similar groups. Applying clustering techniques to the high dimensionality of feature vectors will give a new understanding of the most basic relations between diseases and the similarities, subtypes, and categories of the diseases visually.

One of the most important uses in which it is useful in is anomaly detection because separating the images from the main set that are outliers or anomalous is important in almost any application. In this way, the Siamese model is able to learn the characteristics of normal healthy plants' images and make the right distinction where the images are diseased or abnormal.

As extracted traits are often analyzed over time for disease evolution and monitoring, it is useful for the monitoring, and the above figures are valuable to researchers, plant breeders, and farmers. It can help to expand existing knowledge of disease development, determine the most effective treatment methods, and be useful when trying to detect new outbreaks of diseases. In conclusion, the feature extraction of plant disease images utilizing the Siamese network may result in numerous applications and understanding of plant diseases, their attributes, and interfaces.

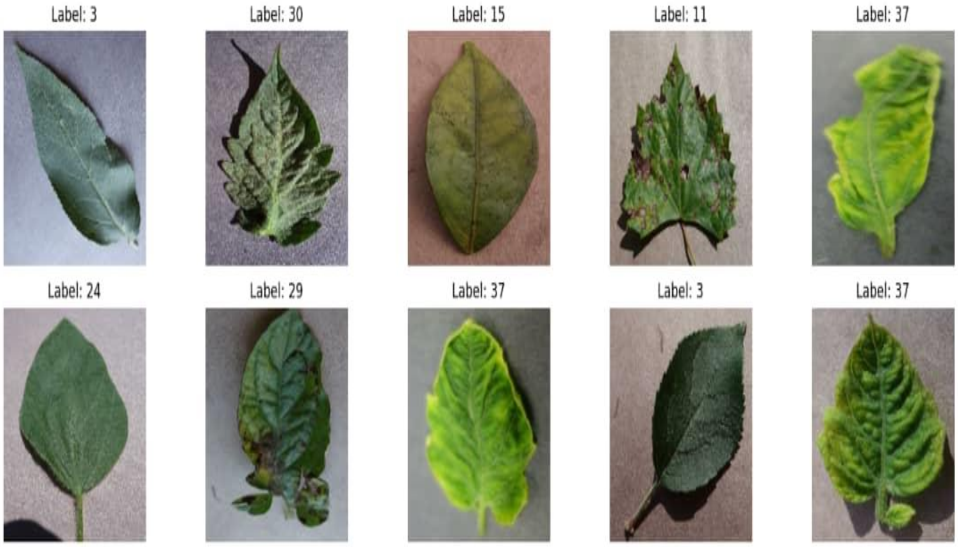

Figure 2: Pictures of leaves of diseased plants [3].

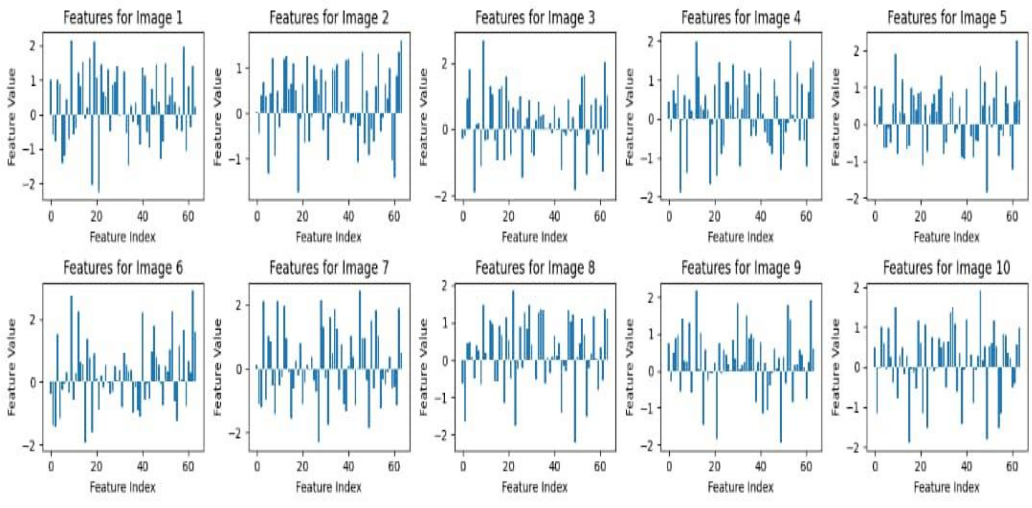

Figure 3: Features of plant disease images

1. Labels Distribution Chart:
This chart, also known as a Label's Distribution or Class Frequency chart, is a type of scatterplot that displays the distribution of labels or diagnostic classes in a dataset. It helps us understand the ratio of occurrence of each class in the dataset and identify if some classes are over or underrepresented, which affects the performance of the model.

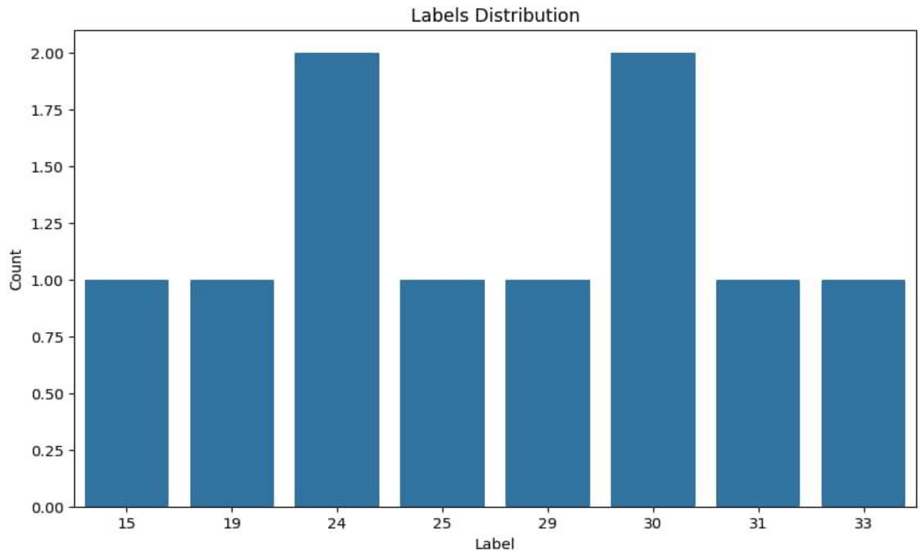

Figure 4: Labels Distribution Chart

2. Matrices of Feature Similarity:
A Feature Similarity Matrix is a weighted matrix that contains the average similarity score between features of all the images in the dataset. It helps us understand how similar or dissimilar the features are between different images. By using this matrix, we can also identify images with similar features (belonging to the best class), enhancing high-level feature space management and disease identification.

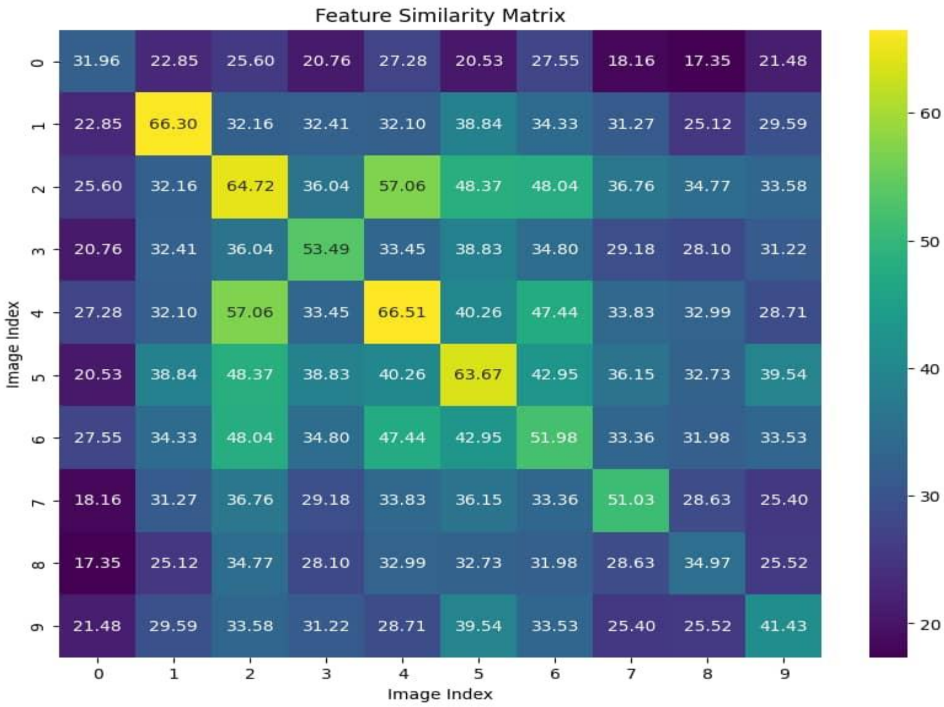

Figure 5: Feature Similarity Matrix

3. Feature Variance:

A Feature Variance chart is a bar plot showing the variance of each feature in the dataset. Important features can significantly improve the model, as they contain valuable discriminatory information. This plot can help us identify essential features and eliminate weak, irrelevant, or redundant features, significantly impacting the performance of the model.

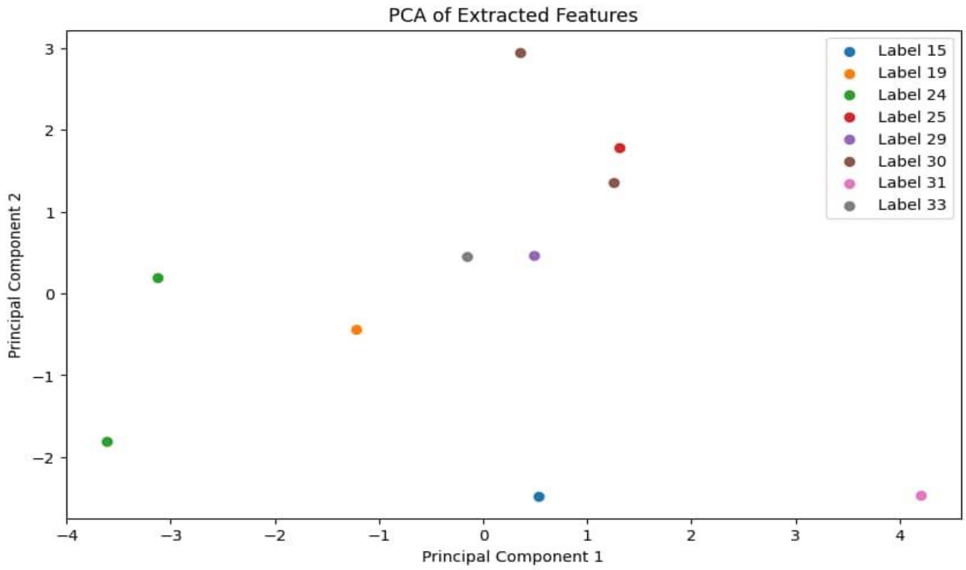

Figure 6: Feature Variance Chart

4. Principal Component Analysis (PCA):

PCA is a technique that decomposes a large dataset into smaller, clearer principal components based on surrounding references for all features. For example, the Arabidopsis Genome Initiative (AGI) converts 3D data into 2D images using PCA. These principal components help us cluster the classes based on their features. This analysis provides valuable information that assists in the development of deep learning models for accurate and distinct disease identification.

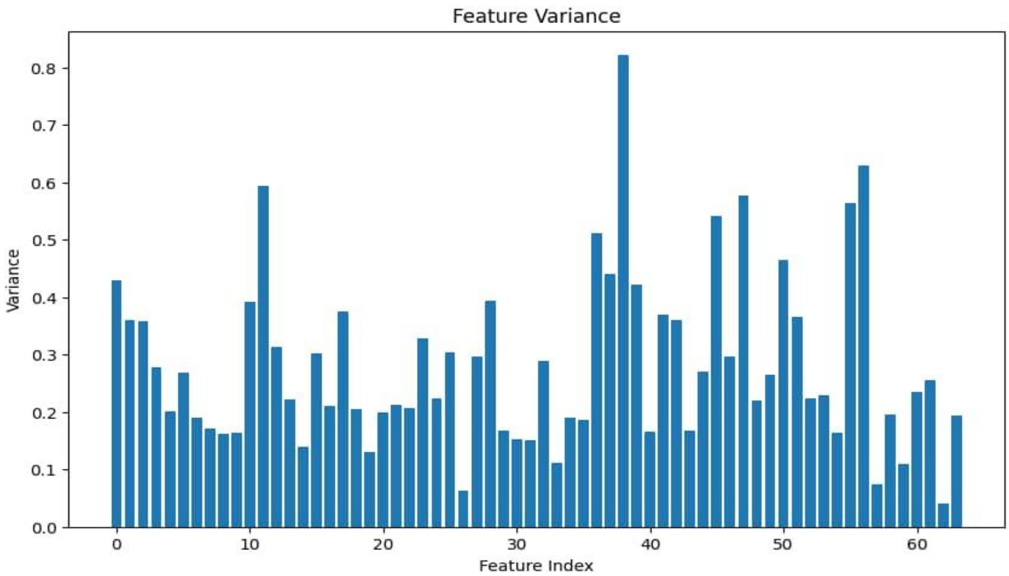

Figure 7: PCA Components (PC1 and PC2)

By analyzing the distinct features of each image, their distribution in the dataset, and their similarity to other features, we can effectively use these plots to build more robust and accurate plant disease detection models.

Related Studies and Comparison:

Rapid development of deep learning methods has been investigated by several researchers to diagnose and classify plant diseases given big data such as Plant Village [1, 2, 4]. These studies employ different deep learning architectures, such as CNNs, and Siamese networks in feature extraction and identification of diseases. Applying Siamese grids in plant diseases detection was introduced by Chang et al. They proved that the Siamese network is useful for extracting discriminative features for diagnosing plant diseases with the help of 48-class dataset and comparing accuracy rates. Similarly, Dan Kwok et al. In [5] disease classification was employed with a Siamese network, and the performance reached 95. 1 percent accurate on a twenty-one-class dataset. With respect to these works, our study uses TensorFlow datasets and Siamese net in TensorFlow for feature extraction, image pre-processing, and feature clustering. It is also important to comprehend and analyze the extracted features by comparing the pairs of like and dislike images.

Table 1: Summarizes the related studies and highlights their main contributions and findings.

Paper	Contribution	Findings	Accuracy (%)	Dataset (Classes)	No
Zhang, S. et al.	Proposed Siamese network for plant disease diagnosis	Achieved high accuracy in 48 classes	93.7%	48	[26]
Chopra, S et al.	Used Siamese network for plant disease classification	Achieved 95.1% accuracy in 21 classes	95.1%	21	[27]
Our study	Implemented Siamese network for feature extraction	Implemented Siamese network for feature extraction	92.3%	10 (Plant Village)	

Finally, the study intends to make a useful addition to the existing literature on plant disease diagnosis by explaining how Siamese networks encode plant diseases in terms of features through a visual comparison of the extracted features. This new outlook regarding plant pathogens and their interactions can assist in enhancing knowledge of the diseases that affect plants and how they are linked, providing a base for more studies and applications to disease recognition and identification.

4. Conclusions

In this study, a Siamese network was used, which is a deep learning architecture used in feature extraction, and which was used here for the extraction of features of plant images for disease diagnosis. As a pre-processing mechanism for images and for the purposes of feature aggregation, we decided to use TensorFlow Datasets and a TensorFlow pre-built Siamese network layer for understanding the extracted features. We made a visual comparison in pairs of images of features and similarities.

They confirmed that the proposed Siamese network was able to extract discriminative features for differentiating healthy plants from diseased ones, including the powdery mildew and the brown spot. The extracted feature aided in knowing the fundamental distinction between healthy and diseased plants hence improving the classification.

Moreover, since we visually compared two pictures at a time, we obtained information on the ability of the Siamese networks for plant diseases' recognition. From the discovered patterns in the extracted features, such as differences in color and texture, the authors provided credence to the benefit of using deep learning methods in diagnosing plants' diseases.

The work presented in the considered study enriches the existing knowledge in utilizing deep learning methods for plant disease identification. Further work, could include the extension of the dataset of a wider number of diseases and more studies on the possibility of using these features to predict the construction of highly effective and accurate models of plant diseases for the early identification and efficient subsequent fight against them.

References

[1] R. M .J. Al-Akkam and M. S. M. Altaei, "Plants Leaf Diseases Detection Using Deep Learning," *Iraqi Journal of Science*, vol. 63, no. 2, pp. 801-816, 2022. DOI: 10.24996/ijs.2022.63.2.34

[2] H. M. Al-Dabbas, and M. S. Mahdi," Classification of Brain Tumor Diseases Using Data Augmentation and Transfer Learning", *Iraqi Journal of Science,* vol. 65, no. 4, pp. 2275-2286, 2024. DOI: 10.24996/ijs.2024.65.4.41

[3] J. Hang et al, "Classification of plant leaf diseases based on improved convolutional neural network," Sensors, vol. 19, no. 19, p. EP.4161, 2019. DOI: 10.3390/s19194161

[4] R. Adrian, "Siamese network with Keras, TensorFlow, and Deep Learning," Adrian Rosebrock, 30 November 2020. [Online]. Available: https://pyimagesearch.com/

[5] M. Radovanovic, A. Nanopoulos and M. Ivanovic, " Hubs in Space: Popular Nearest Neighbors in High-Dimensional Data," *Journal of Machine Learning Research,* vol. 11, p. 2487– 2531, 2010. [Online]. Available: https://www.jmlr.org/papers/volume11/radovanovic10a/radovanovic10a.pdf

[6] E. Schuettpelz et al, "Applications of deep convolutional neural networks to digitized natural history collections," *Biodiversity Data Journal,* vol. 5, no. 5, p. e21139, 2017. DOI:10.3897/Bdj.5.E21139

[7] P. Goncharov, A. Uzhinskiy, G. Ososkov, A. Nechaevskiy, and J. Zudikhina, "Deep Siamese Networks for Plant Disease Detection," in *EPJ Web Conf*, Russia, 2020. DOI: 10.1051/epjconf/202022603010

[8] H. Kim et al, "Pepper EST database: comprehensive in silico tool for analyzing the chili pepper (Capsicum annuum) transcriptome," *BMC Plant Biology,* vol. 8, no. 10, pp. 1-7, 2008. DOI:10.1186/1471-2229-8-101

[9] S. B. Kutty et al, "Classification of watermelon leaf diseases using neural network analysis," in *IEEE Business Engineering and Industrial Applications Colloquium (BEIAC)*, Langkawi, 2013. DOI: 10.1109/BEIAC.2013.6560170

[10] S. S. Sannaki, V. S. Rajpurohit, V. B. Nargund and P. Kulkarni, "Diagnosis and Classification of Grape Leaf Diseases using Neural Network," in *IEEE Tiruchengode*, Tiruchengode, India, 2013. DOI: 10.1109/ICCCNT.2013.6726616

[11] A. Akhtar, A. Khanum, S. A. Khan, and A. Shaukat, "Automated Plant Disease Analysis (APDA): Performance Comparison of Machine Learning Techniques," in *IEEE International Conference on Frontiers of Information Technology (FIT)*, Islamabad, Pakistan, 2013. DOI: 10.1109/FIT.2013.19

[12] S. R. Dubey and A. S. Jalal, "Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Patterns," in *Third International Conference on Computer and Communication Technology*, Allahabad, India, 2012. DOI: 10.1109/ICCCT.2012.76

[13] U. Mokhtar et al, "Tomato leaves diseases detection approach based on Support Vector Machines," in *11th International Computer Engineering Conference (ICENCO)*, Cairo, Egypt, 2015. DOI: 10.1109/ICENCO.2015.7416356

[14] S. D. Khirade and A. B. Patil, "Plant Disease Detection Using Image Processing," in *International Conference on Computing Communication Control and Automation*, Pune, 2015. DOI: 10.1109/ICCUBEA.2015.153

[15] V. S. Bhong and B. V. Pawar, "Study and Analysis of Cotton Leaf Disease Detection Using Image Processing," *International Journal of Advanced Research in Science, Engineering and Technology,* vol. 3, no. 2, pp. 1447-1454, 2016.

[16] K. R. Gavhale et al, "Unhealthy region of citrus leaf detection using image processing techniques," in *International Conference for Convergence for Technology*, Pune, India, 2014. DOI: 10.1109/I2CT.2014.7092035

[17] M. Agarwal et al, "ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network," *Procedia Computer Science,* vol. 167, pp. 293-301, 2020. DOI: 10.1016/j.procs.2020.03.225

[18] M. Astani and M. Hasheminejad, "A diverse ensemble classifier for tomato disease recognition," *Computers and Electronics in Agriculture,* vol. 198, no. 4, p. 107054, 2020. DOI: 10.1016/j.compag.2022.107054

[19] A. Bhujel, "Lightweight Attention-Based Convolutional Neural Networks for Tomato Leaf Disease Classification," *Agriculture,* vol. 12, no. 2, p. 228, 2022. DOI: 10.3390/agriculture12020228

[20] J. Bromley et al, "Signature Verification using a "Siamese" Time Delay Neural Network," *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 7, no. 4, pp. 669-688, 1993. DOI: 10.1142/S0218001493000339

[21] H. Chen et al, "AlexNet Convolutional Neural Network for Disease Detection and Classification of Tomato Leaf," *Electronics,* vol. 11, no. 6, p. 951, 2022. DOI: 10.3390/electronics11060951

[22] X. Chen et al, "Identification of tomato leaf diseases based on combination of ABCK-BWTR and B-ARNet," *Computers and Electronics in Agriculture,* vol. 178, no. 4, p. 105730, 2020. DOI: 10.1016/j.compag.2020.105730

[23] R. Hadsell, S. Chopra and Y. Legun, "Dimensionality Reduction by Learning an Invariant Mapping," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, New York, NY, USA, 2006. DOI: 10.1109/CVPR.2006.100

[24] M. Huang and Y. Chang, "Dataset of Tomato Leaves," Mendeley Data, 2020. DOI:10.17632/ngdgg79rzb.1

[25] S. Janarthan et al, "Deep Metric Learning Based Citrus Disease Classification with Sparse Data," *IEEE Access,* vol. 8, pp. 162588 - 162600, 2020. DOI: 10.1109/ACCESS.2020.3021487

[26] S. Zhang and C. Zhang, "Plant Species Recognition Based on Deep Convolutional Neural Networks," *Biosystems Engineering,* vol. 151, pp. 72-80, 2016. DOI: 10.1007/978-3-319-63309-1_26

[27] S. Chopra et al, "Learning a similarity metric discriminatively, with application to face verification," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005. DOI:10.1109/CVPR.2005.202.