

RESEARCH ARTICLE

Foliagen: Framework for Crop Foliage Image Generation From Individual Leaf Images

NABIN PAKKA¹, (Member, IEEE), **TERESA WILKERSON**², **LI CHEN**¹, (Senior Member, IEEE),
AND NIAN-FENG TZENG¹, (Life Fellow, IEEE)

¹School of Computing and Informatics, University of Louisiana at Lafayette, Lafayette, LA 70504, USA

²Department of Agricultural Science and Plant Protection, Mississippi State University, Starkville, MS 39762, USA

Corresponding author: Nabin Pakka (nabin.pakka1@louisiana.edu)

This work was supported in part by the U.S. National Science Foundation under Grant OIA-2327452.

ABSTRACT While machine learning (ML)-based crop disease classifiers mostly targeted individual leaf images, real-world applications require disease classification on in-field foliage images. Since properly annotated disease foliage image datasets are scarce for ML model training, it is desirable to generate large, annotated foliage image datasets with primary crop disease categories to effectively train and evaluate state-of-the-art (SOTA) crop disease classifiers on foliage images. This paper introduces a structure-based framework that generates crop foliage images utilizing available datasets of individual leaf images, termed Foliagen. Foliagen follows structural characteristics and diseased leaf distributions of natural plants to generate the foliage image datasets of arbitrary sizes, instead of employing a generative artificial intelligence (AI) approach, which usually produces ill-structured foliage images and is ineffective for model training. Being annotated by design, such generated datasets are valuable for 1) augmenting often scarce diseased foliage images for effective SOTA classifier training and 2) pre-training SOTA classifiers for in-field diseased foliage image classification after fine-tuning them with a handful of in-field images. The Foliagen framework is exemplified by generating foliage image datasets for soybean and tomato. Our evaluation shows that 1) training classifiers on a combination of our generated dataset with a small portion (say, 5%) of an available natural foliage dataset achieves up to 98% accuracy when evaluated on the natural foliage dataset and 2) classifiers pre-trained on our dataset can be fine-tuned via a few natural foliage images to achieve high classification performance. Foliagen is effective in foliage image generation for in-field applications.

INDEX TERMS Classifier training, crop leaf images, disease image classification, image generation, machine learning.

I. INTRODUCTION

Crops have been indispensable to human civilization since its inception, serving as fundamental sources of food, medicine, clothing, shelter, etc. According to the Food and Agriculture Organization (FAO), however, crop diseases cost approximately \$220 billion annually in the world [1], with soybean alone accounting for a loss up to \$3.9 billion USD in the USA [2]. They usually show prominent symptoms, manifesting themselves as changes in the soybean's leaf appearance and/or shape. For example, rust in soybean exhibits small, pale green to yellow spots on the upper surface of leaves (see Figure 1(a)). Many crop disease categories

possibly exist as illustrated in Figure 1, and early disease diagnosis makes it possible to apply proper measures at the onset of diseases for curbing damage they may cause, retaining crop yields as best as possible.

Instead of relying on experienced farmers for disease identification, machine learning (ML) has been adopted [3], [4], [5], [6], [7], [8], [9] to classify soybean and tomato diseased leaf images with success. State-of-the-art (SOTA) ML models automate disease classification through training on large amounts of diseased and healthy leaf images. As exemplified in Figure 1, quality images of individual diseased soybean and tomato leaves exist in public datasets [10], [11], [12] for model training. Although SOTA ML models exhibit high classification performance on the existing datasets of individual leaf images, they lack practicality for in-field

The associate editor coordinating the review of this manuscript and approving it for publication was Ikramullah Lali¹.

applications, where sequences of foliage images are usually captured over areas of interest by the cameras of uncrewed aerial vehicles (UAVs) for disease identification. Therefore, accurate automated disease classification on foliage images is essential for practical applications. It requires a large volume of high-quality, properly annotated diseased foliage image datasets for effective training. However, obtaining such datasets requires highly experienced agronomists with substantial time and effort to annotate segmented parts of foliage images for accurate classification, considered prohibitively expensive and impractical to apply for in-field applications. To the best of our knowledge, there is just one annotated soybean foliage image dataset available to the public, MH-SoyaHealthVision (SHV) [13], and no publicly available dataset of tomato foliage images exists. Unfortunately, the SHV dataset fails to include many predominant soybean disease categories (see Figure 7 of Appendix) and is likely to be inadequate to train an effective classifier for identifying the diseases at their early stages, since its images were not captured at the disease onset. While efforts have resorted to generative artificial intelligence (GenAI) to generate single-leaf images [14], [15], [16], [17], [18], none have attempted to produce diseased foliage images. Diseased foliage image generation remains challenging for GenAI models, due to (1) a large number of leaves present within each foliage image and (2) the absence of sufficient foliage images needed for effective model training.

This paper pioneers a framework, called Foliagen (short for foliage gener-ation), for generating agronomist-approved structure-based [19], [20], [21] crop foliage images by leveraging available single-leaf image datasets. Foliagen follows plant structural characteristics and diseased leaf distributions to generate large foliage image datasets. Each generated foliage image has a specified rate of diseased leaves (denoted by γ), with the rest being healthy. Such generated datasets are annotated by design, tailored to in-field applications for foliage image classification. They are valuable for (1) augmenting often scarce natural foliage images for SOTA classifier training to get high classification performance, unachievable solely by the scarce natural foliage images alone and (2) pre-training general SOTA classifiers for target farmland applications by employing a small number of disease foliage images obtained from the farmland (say, tens of them per disease category) for fine-tuning the pre-trained ones to exhibit classification performance far higher than that achievable by the obtained farmland foliage images alone. The Foliagen framework is exemplified by generating foliage datasets for soybean and tomato, using the diseased single-leaf images from ASDID [10], PlantVillage [11], and Kaggle [12] datasets. Its inputs are (1) disease categories and (2) γ (the diseased leaves rate of each foliage image). Each generated foliage image dataset covers nine predominant disease categories for soybean and tomato (with samples depicted in Figure 4 and respectively in Figure 7 and Figure 8 of Appendix). Such datasets enable classifiers to

train for early disease identification when foliage images are generated at a small disease rate γ (say, 5%).

Our Foliagen framework aims to generate foliage datasets with specified disease distributions. When foliage images are generated following the disease distribution of the SHV dataset (the only known natural foliage dataset), SOTA classifiers are expected to exhibit similar performance results on generated foliage images as on SHV images, as demonstrated in Section IV-A. Diseased foliage datasets are inherently scarce, usually with limited numbers of images and disease categories, due to the high cost and lengthy time of skilled agronomists involved in various farmlands for dataset creation. Fortunately, generated foliage images can serve to augment often scarce natural foliage images for classifier training to get high performance, as illustrated in Section IV-B. In addition, generated foliage images can pre-train classifiers for target field applications after they are easily fine-tuned for high classification performance, as revealed in Section IV-B. Classifiers so pre-trained learn the abstract features of diseased foliage images under various disease categories, making it possible to adapt soundly for classifying the real-world foliage images accurately. Hence, Foliagen facilitates the generation of diseased foliage image datasets for any crop, enhancing the performance of crop foliage disease classifiers for in-field applications, even when only a limited amount of annotated field foliage images is available. The overall contributions of this paper are summarized as follows:

- The Foliagen framework is introduced for generating properly annotated crop foliage image datasets by utilizing public datasets of individual leaf images. Datasets so generated can be arbitrarily large, properly annotated, and aimed to cover various crop disease categories common in the field and to target early disease identification by setting a small disease rate γ (say, 5%).
- Foliagen is demonstrated by generating foliage image datasets for soybean and tomato, where the generated soybean datasets are used to (1) augment the scarce natural foliage images available (say, tens of them per disease category only) for training SOTA classifiers, resulting in significantly improved performance when compared to training on scarce natural foliage images alone.
- We demonstrate that generated foliage image datasets can pre-train a general model for crop disease classification, so that the pre-trained model can then be fine-tuned by a small number of natural foliage images obtained from the field to lift performance when applied to field applications, as a result of transfer learning.

II. RELATED WORK

A. PLANT LEAF DATASETS

Various plant leaf datasets are available to the public, with some covering a variety of plant species each [11], [22], [23],

[24] and others being plant-specific [10], [12], [13], [25] (all datasets detailed in A. Specifically, PlantVillage [11], FieldPlant [22], PlantSeg [24], and PlantDoc [23] cover a variety of plant species each. PlantVillage, FieldPlant, and PlantSeg all contain expertly curated diseased single-leaf images, with the first one captured in a laboratory-controlled environment, the second one obtained in real-world environments, and the third one being internet-sourced to suffer from heterogeneous quality and inconsistent scales. PlantDoc involves single-leaf images across 13 species downloaded from the internet without quality annotations. On the other hand, several plant-specific single-leaf image datasets are publicly available, including the Auburn Soybean Diseased Image Dataset (ASDID) [10], a soybean single-leaf diseased dataset from Kaggle [12], and the TomatoVillage dataset [25]. ASDID contains individual soybean leaf images captured in a laboratory-controlled setup and covers 7 disease categories, whereas the soybean dataset from Kaggle involves 9 disease categories. The TomatoVillage dataset includes individual tomato leaf images captured in real-world settings, covering 9 disease categories. All of the aforementioned datasets contain single-leaf images, limiting their applicability to real-world disease foliage classification.

To the best of our knowledge, only one publicly available diseased foliage image dataset exists, MH-SoyaHealthVision [13], which was collected using a UAV-mounted camera over soybean fields. However, it covers only two disease categories, whereas in practical agriculture settings, a wider range of diseases must be considered. The acquisition of high-quality foliage images with expert annotations is both costly and time-consuming, leading to the scarcity of diseased foliage datasets. Therefore, there is an urgent need for inexpensive and fast solutions, such as our Foliagen, to advance research in real-time diseased foliage image classification, whether deployed at fixed stations or on edge devices.

B. GENERATION METHODS

Simple copy & paste methods have been utilized to augment image data [26], [27] for improving object detection [28], image classification [29], and instance segmentation [30], [31], [32]. Separately, a probabilistic model of DDPM [33], and a statistical model [34] were developed to generate crop foliage images, with their generated images usually ill-structured and clearly different from those captured in a crop farm field.

Various generative AI (GenAI) models exist for leaf data generation and augmentation [14], [15], [16], [17], [18], [24], [35], [36], [37], [38], [39], [40]. For example, RDCGAN [39] combines Deep neural networks (DNNs) with Generative Adversarial Networks (GANs) for unsupervised representation learning. Similarly, PixelCNN [38] autoregressively models image distributions using masked convolutions to capture spatial dependencies. They either generate only single-leaf images or produce foliage images

that are structurally different from real ones. GenAI models like Pixel-CNN [38], RDCGAN [39], and DDPM [33] fail to generate realistic diseased foliage images due to a lot of leaves per image and insufficient training data. Their outputs differ starkly from UAV-captured field images (see Figure 10 and Figure 11 of Appendix), necessitating Foliagen for effective disease classifier training.

C. CROP LEAF DISEASE CLASSIFICATION MODELS

Automated disease classification studies have been conducted lately based on the aforementioned publicly available datasets, plus certain privately collected datasets, to exhibit impressive classification outcomes [3], [4], [5], [6], [7], [8], [9], [41], [42]. For example, standard CNN-based models were employed to classify single-leaf images [10]. Enhanced DenseNet121 [9] relied on transfer learning to classify individual soybean leaf images. A conditional generative adversarial network (C-GAN) was applied [3] to synthesize individual tomato diseased leaves, on which DenseNet121 [43] achieved good classification performance.

Attention mechanisms [6], [7], [8], [44], [45], [46], [47] have been proven effective for leaf classification, where leaf colors and shapes are key features of interest [5], [8]. Swin Transformer [46] introduces a hierarchical segmentation along with the vision transformer to implement shifted windows for capturing main features. CBAM-ConvNeXt [8] employs both channel attention and spatial attention plus ConvNeXt [48] to classify the individual soybean leaf images.

While exhibiting impressive feats, almost all known studies aim only at classifying single-leaf images, which have limited application for real-world foliage disease detection (as described in VII-D). In comparing SOTA disease classification models under foliage image datasets, we chose VGG19 [49], ResNet50 [50], DenseNet121 [43], Swin Transformer [46], and CBAM-ConvNeXt [8] as representative models for comparative evaluation, with their results in Tables 1-4.

III. METHODOLOGY

The Foliagen framework generates diseased foliage images out of available single-leaf diseased images for the classification of diseased foliage images for (1) objectively evaluating known and future crop disease classifiers when deployed for in-field applications and (2) pre-training general crop disease classifiers, which can be tailored for specific fields with high classification performance after fine-tuned by a small number of annotated foliage images gathered in those fields, as the result of transfer learning. Instead of employing a GenAI model, which usually fails to produce foliage images resembling those gathered from crop farm fields by cameras, Foliagen follows plant structural characteristics and diseased leaf distributions to generate the annotated foliage image datasets of arbitrary sizes, useful to train AI models aimed at real-world crop disease classification. It is exemplified by generating foliage images of soybean and

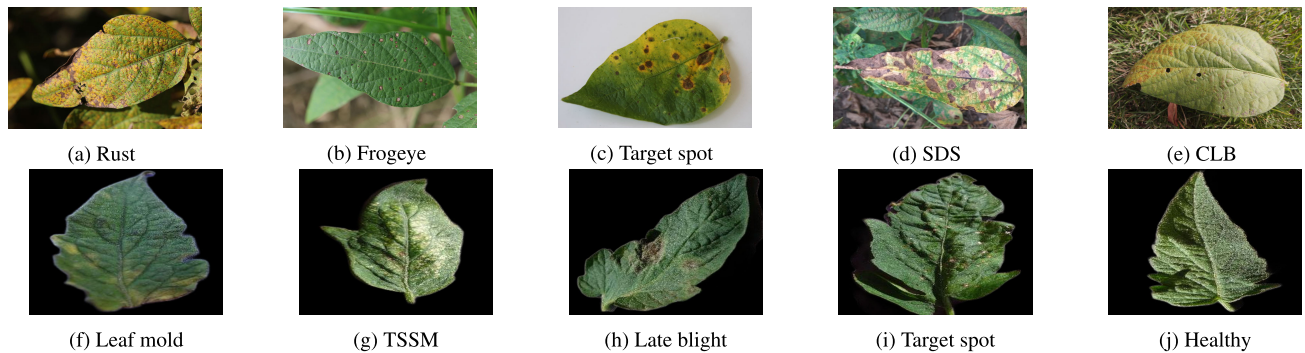


FIGURE 1. Single-leaf images of soybean (a)–(e) and tomato (f)–(j), (f)–(j), with SDS, CLB, and TSSM denoting Sudden Death Syndrome, Cercospora Leaf Blight, and Two Spotted Spider Mite, respectively. (All single-leaf disease categories for soybean and tomato are shown in Figure 5 and Figure 6 of Appendix, respectively).

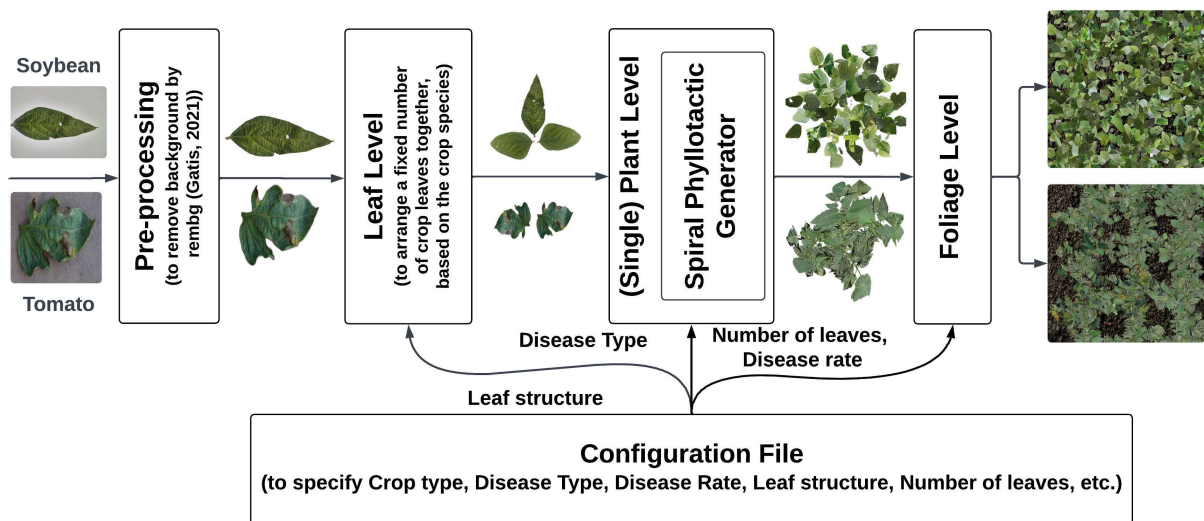


FIGURE 2. Overall foliagen.

tomato, leveraging three publicly available individual-leaf image datasets, ASDID [10], a Kaggle dataset [12], and the PlantVillage dataset [11], which are detailed in III-A. Note that the above three datasets were chosen from several considered datasets (discussed in II-A) due to their large numbers of well-annotated images, with high fidelity and clarity. Single soybean leaf images are preprocessed (as described in III-A) before being utilized by Foliagen to generate diverse sets of foliage images for evaluating known classifiers objectively. Such an image preprocessing has been proven to be effective for image segmentation and object detection data augmentation [28], [30], [31]. In addition, known classifiers, after being pre-trained by generated soybean foliage images, are fine-tuned via a small fraction (say 5%) of images from the real-world foliage dataset of MH-SoyaHealthVision (SHV) [13] to achieve improved performance when classifying soybean foliage images captured from the farm fields.

A. DATA COLLECTION AND PRE-PROCESSING

1) SOYBEAN

A total of 10,722 high-resolution single-leaf images covering 7 disease categories, namely, bacterial blight, cercospora

leaf blight (CLB), downy mildew, frogeye, soybean rust, target spot, and potassium deficiency, from ASDID [10] and 132 images (with 22 for Mosaic Virus and 110 for Sudden Death Syndrome (SDS)) from a Kaggle dataset [12] were chosen for foliage image generation. These images feature diverse and complex natural/artificial backgrounds, which are undesirable when generating foliage images. Therefore, we pre-process the images using an open-sourced and AI-enabled background remover, rembg [51], as shown in the first stage of Foliagen depicted in Figure 2. Freely available soybean field soil images are then included in the generated foliage images as their backgrounds, to emulate the natural habitat of soybean plants as best as possible.

The only publicly available and annotated dataset of diseased soybean foliage images, SHV [13], comprises two disease categories: rust and mosaic virus. The original images have a very high resolution, and the dataset is imbalanced, with rust images outnumbering healthy images by a factor of four, leading to biased predictions favoring the majority class and consequently reducing the model's generalizability. To address these issues, we crop the high-resolution foliage images into ones with a lower resolution to ensure a more balanced data distribution and to obtain a total of

3210 images, comprising 1084 rust images, 1027 healthy images, and 1099 mosaic virus images, respectively.

2) TOMATO

The single-leaf diseased images of tomato are from the PlantVillage dataset [11], with 9 primary tomato disease categories, namely, leaf curl virus, Mosaic virus, leaf mold, septoria spot, two-spotted spider mite (TSSM), bacterial spot, early blight, late blight, and target spot, and the dimension of 256×256 , some of the disease categories are shown in Figures 1(f)-(j).

B. FOLIAGEN FRAMEWORK

Foliagen takes pre-processed individual leaf images as input to generate foliage images with a customizable rate of diseased leaves. After preprocessing single leaf images for background removal via rembg [51], it then involves three generating levels, Leaf Level, Plant Level, and Foliage Level, as depicted in Figure 2. Each level incorporates plant-specific information like disease categories, target disease rate, and inherent leaf structural characteristics, as denoted by ‘Configuration File’ (detailed in VII-I) in Figure 2. Pathologically, the infected leaves of a disease foliage image are often clustered in a few small areas of the image, called disease clusters. A foliage image is expected to have more disease clusters as its disease rate (γ) rises. Under a small γ (say, 5% to indicate an early disease stage), a foliage image is likely to contain just one disease cluster, making it harder for accurate disease classification.

1) LEAF LEVEL

Plant leaves exhibit diverse inherent structural arrangements (e.g., trifoliate, bifoliate, etc.). For example, soybean leaves arrange themselves into 3 leaflets in each petiole, with a slightly bigger central leaflet and lateral leaflets on the two sides of the central leaflet, as illustrated in Figure 2. On the other hand, tomato leaves arrange themselves around a central axis, called the rachis. In practice, crop leaflets may be either healthy or diseased. Therefore, each foliage image contains a mixture of healthy and diseased leaflets, with the proportion of diseased leaflets governed by the disease rate (γ).

2) PLANT LEVEL

The number of leaves in an adult crop differs largely, based on the crop. For example, an adult soybean plant might contain 30-40 leaves, while an adult tomato plant may have 20-40 leaves. Both crops exhibit spiral phyllotaxis [19], [20], in which leaves are arranged in a spiral arrangement that makes the golden angle, i.e., 137.5° , to maximize sunlight exposure and minimize leaf overlap. The algorithm for implementing spiral phyllotaxis (Algorithm 1) is shown in Appendix F. Foliagen creates plants individually, with those within a disease cluster having diseased leaves (at the rate governed by the foliage disease rate) and all others possessing solely healthy leaves. In an individual plant, both diseased

and healthy leaves are arranged according to spiral phyllotatic coordinates at the plant level to determine leaf positions and angles, thereby maintaining natural spiral phyllotaxis [19].

As a tomato plant has branches, with a pair of leaves attached at a similar stem height and arranged in opposite directions, a sub-layer, called the branch layer, is added to create branches, each with 3-9 leaves. Such an emulated branch is then attached to the main stem in a way following the spiral phyllotactical order.

3) FOLIAGE LEVEL

A foliage image usually consists of multiple rows of crops planted in a farm field. Observation from real-world foliage images taken using UAVs [13] reveals that the major area of an image is covered by leaves, with only a small area being field soil [52]. Hence, Foliagen partitions every foliage image into a grid of (α) rows and (β) columns, resulting in a total of ($N = \alpha \cdot \beta$) positions, each corresponding to a single plant. Each disease foliage image usually involves just one or a few disease clusters, with the disease rate(s) of its disease cluster(s) summing to the image’s disease rate γ (e.g., 5% or 15%). Plants located within a disease cluster are assigned positional disease rates, denoted by γ_p , whereas plants outside disease clusters contain only healthy leaves. This clustered disease leaf distribution enables realistic modeling of disease spread across generated foliage images. Samples of generated foliage images are illustrated in Figure 4. Note that a soybean (or tomato) foliage image generated by Foliagen is dominated by leaves, with only a small fraction of its area being background (see Figure 4 and also Figure 7 and Figure 8 of Appendix), in contrast to an individual leaf image where background or noise takes up a considerable area (see Figure 5 and Figure 6 of Appendix).

The foliage framework outlined above is realized via two algorithms, Algorithm 2 and Algorithm 3, which correspond respectively to the plant level and the foliage level, besides Algorithm 1 on spiral phyllotaxis, as detailed in Appendix F.

C. DISEASE DISTRIBUTION

The distribution of disease in plant leaves is influenced by multiple interacting factors, including insect vectors, wind-mediated spore dispersal, plantation age, and environmental conditions, such as humidity, rainfall, and temperature. Although the spatial pattern of disease may vary considerably depending on these influences, a consistent phenomenon is that infections generally begin in a small area and subsequently spread outward to one (or multiple) neighboring area(s), ultimately forming a larger area, called a disease cluster [53], [54], [55]. This diseased leaf distribution is confirmed by our close examination of real-world foliage images in the SHV dataset [13] (see Figure 3(c)), leading us to devise a three-level disease distribution manner, as explained next.

1) REGION LEVEL

The whole image is divided into a $\alpha \times \beta$ grid with a total of $\alpha \times \beta$ regions. e.g., $4 \times 3 = 12$ regions. Most regions are

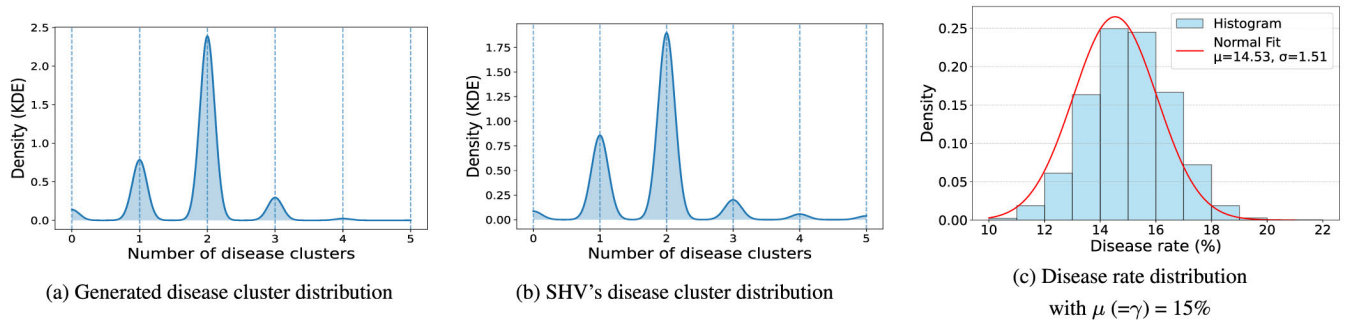


FIGURE 3. (a) Disease cluster distribution across all generated soybean foliage images for $\gamma = 15\%$, (b) the disease cluster distribution of the natural foliage images in SHV [13], and (c) the disease rate distribution of $\gamma = 15\%$.

disease-free when the disease rate (γ) is small, and the disease regions tend to fall within one or a few disease clusters. Each disease cluster is provided with a cluster disease rate, with all cluster disease rate(s) aggregating to γ . Infected leaves within a disease cluster are distributed normally across the cluster.

2) CLUSTER LEVEL

The number of disease clusters in a single foliage image varies from 1 to 5, given that our target disease rate for early detection is low (say, $\leq 20\%$). The overall disease rate (γ) of a foliage image is distributed across its disease cluster(s), represented by a vector γ_c , with its i^{th} vector element registering the disease rate of the i^{th} disease cluster. The disease rate of the i^{th} cluster, $\gamma_c[i]$, is distributed normally across all regions within the cluster, with the highest disease rate at the cluster center (see Lines 10–12 of Algorithm 3 in Appendix F). Naturally, a higher γ is expected to yield more disease clusters in a foliage image. The distribution of disease clusters among all generated foliage images follows a skewed graph as shown in Figure 3(a), which is similar to that of SHV images [13], as depicted in Figure 3(b) with the number of disease clusters per image dominated by 2 and the mean disease rate of approximately 15%.

3) DATASET LEVEL

We have analyzed SHV images fully, discovering that the disease rate (γ) of those images follows a normal distribution. Accordingly, Foliagen generates every foliage image dataset with the disease distribution of its images governed by the normal distribution characterized by a standard deviation (σ) of 1.5 and the variable mean μ equal to γ . Figure 3(c) depicts the disease rate distribution for a mean value of $\gamma = 15\%$, signifying that the disease rate varies from 10% to 20%. For early disease identification, foliage image datasets are also generated with their disease rates ranging from 1% to 9% for the mean value of $\gamma = 5\%$.

4) CROP LEVEL CUSTOMIZATION

Foliagen is a common framework, aiming to generate foliage imagery datasets for various crops, utilizing those crops' single leaf images. Given crops differ in many factors (like leaf shapes, leaf arrangements, phyllotaxis, number of leaves in single branch, numbers of leaves in single branches, etc.),

Foliagen is provisioned with a configuration file as its input to account for the crops' variability (see Figure 2), with the file listing such crop-specific customization parameters as the disease rate, the size of individual leaf, the size of individual plant, the foliage size, disease categories, etc., (as described in VII-I) to properly emulate crops' natural structures.

IV. EVALUATION AND RESULT DISCUSSION

Extensive experiments have been conducted to evaluate state-of-the-art (SOTA) crop disease classifiers under generated foliage datasets and SHV [13] on two workstations, each housing 2 NVIDIA RTX 6000 Ada GPUs (each with 48 GB of GDDR6X VRAM). Various crop foliage image datasets are generated by Foliagen for comparatively evaluating the five representative disease classifiers of VGG19 [49], ResNet50 [50], DenseNet121 [43], Swin Transformer [46], and CBAM-ConvNeXt [8] objectively without any classifier-specific data pre-processing or manipulation. Each generated soybean or tomato foliage dataset includes nine primary disease categories that exist in the original individual-leaf datasets. It contains about 1000 foliage images for each disease category, plus a similar number of healthy leaf foliage images. Foliage images with the disease rate (γ) ranging from 1% to 20% are generated for evaluation, with $\gamma = 5\%$ reflecting an early disease stage. Our full analysis of the SHV dataset indicates that the disease rate (γ) of foliage images ranges from 12% to 18%. Hence, subsequent evaluation results reported mostly for $\gamma = 15\%$.

Our evaluation adopts the Adam optimizer for model training due to its fast convergence and adaptive learning rates. Models are trained using a batch size of 8, constrained by GPU memory, and an initial learning rate of 1×10^{-4} for up to 100 epochs, with early stopping (patience = 5) applied. Accuracy and F1-score (detailed in Appendix E) are used as the primary evaluation metrics, as accuracy reflects the overall prediction correctness, while F1-score captures the harmonic mean of precision and recall, making it more suitable for classification scenarios. These comparative evaluation results shed light on the value of the generated dataset for real-world applications, where disease identification is based on in-field images captured by a camera.

Since our extensive classifier evaluation involves the real soybean foliage dataset (SHV) and/or generated foliage

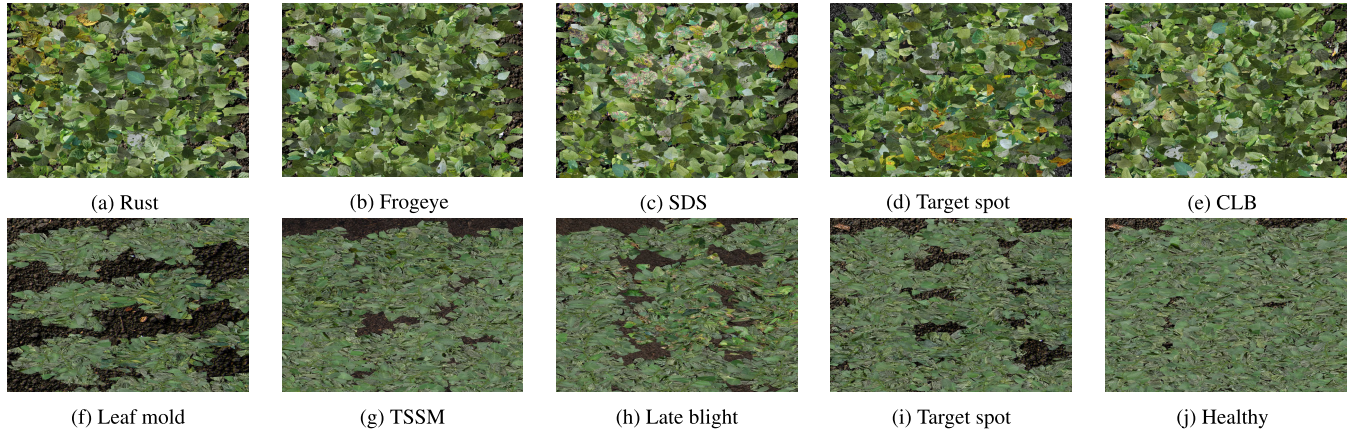


FIGURE 4. Generated foliage images of soybean (a)-(e) and tomato (f)-(j), for $\gamma = 15\%$, with SDS, CLB, and TSSM denoting Sudden Death Syndrome, Cercospora Leaf Blight, and Two Spotted Spider Mite, respectively. (All generated foliage image categories for soybean and tomato are shown in Figure 7 and Figure 8 of Appendix, respectively).

TABLE 1. Notation for denoting classifier evaluation scenarios, where I (or S) represents natural soybean individual leaf datasets (or the natural soybean foliage image dataset, i.e., SHV), and G_{So}^γ (or G_{To}^γ) represents the generated soybean (or tomato) foliage dataset with the disease rate of γ (in %).

Notation	Description
$S(p\%)$ (or $G_{So}^\gamma(p\%)$)	$p\%$ of S (or G_{So}^γ) is used in a phase (training, validation, or testing)
$\langle \text{TR}; \text{VA}; \text{TE} \rangle$	$\langle \text{TR}$: training samples, VA : validation samples, TE : testing samples $\rangle$
$\langle S(5\%); S(10\%); S(85\%) \rangle$	Scenario where 5%, 10%, and 85% of S are used for training, validation, and testing, respectively
$\langle S(5\%) + G_{So}^\gamma(80\%); G_{So}^\gamma(10\%); S(85\%) \rangle$	Scenario where 5% of S and 80% of G_{So}^γ are mixed for training, 10% of G_{So}^γ are used for validation, and 85% of S are used for testing.
$\langle G_{So}^\gamma(80\%); G_{So}^\gamma(10\%) \mid S(5\%); S(85\%) \rangle$ (or $\langle I(80\%); I(10\%) \mid S(5\%); S(85\%) \rangle$)	Scenario where a classifier is first trained on $G_{So}^\gamma(80\%)$ (or $I(80\%)$) and validated on $G_{So}^\gamma(10\%)$ (or $I(10\%)$), then fine-tuned on $S(5\%)$ and tested on $S(85\%)$

TABLE 2. Comparative performance analysis under S (the SHV dataset) and G_{So}^{15} (a generated soybean dataset with γ ranging from 10% to 20% for the mean of 15%).

$\langle S(80\%); S(10\%); S(10\%) \rangle$			$\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%); G_{So}^{15}(10\%) \rangle$	
Classifiers	Accuracy	F1-score	Accuracy	F1-score
VGG19	85.38	86.6	83.08	83.51
ResNet50	88.31	89.69	87.79	88.43
DenseNet121	87.5	87.95	90.31	90.46
Swin Transformer	91.48	91.47	94.37	94.51
CBAM-ConvNeXt	88.82	87.06	90.87	91.15

datasets, several notations are defined first and employed for subsequent explanation throughout this section, as listed in Table 1. Let image samples utilized for the three ML phases be denoted by $\langle \text{TR}; \text{VA}; \text{TE} \rangle$, where TR refers to samples for the classifier training phase and VA (or TE) denotes samples for the validation (or testing) phase. Samples can be from the SHV dataset (represented by S) or a generated soybean (or tomato) dataset with the disease rate of γ (represented by G_{So}^γ (or G_{To}^γ)). For example, the scenario of classifier evaluation using 5% of foliage images from S for training, 10% from S for validation, and 85% from S for testing is denoted concisely by $\langle S(5\%); S(10\%); S(85\%) \rangle$. In general, foliage images can be from both S and G_{So}^γ for classifier evaluation. As an instance, the scenario of $\langle S(5\%) + G_{So}^\gamma(80\%); G_{So}^\gamma(10\%); S(85\%) \rangle$ signifies that 5% of S plus 80% of G_{So}^γ are mixed for training, 10% of G_{So}^γ are for

validation, and 85% from S are for testing. In the cases when fine-tuning is involved, the notation includes fine-tuning sample information in its second element, preceded by “|”. Hence, $\langle G_{So}^\gamma(80\%); G_{So}^\gamma(10\%) \mid S(5\%); S(85\%) \rangle$ (or $\langle I(80\%); I(10\%) \mid S(5\%); S(85\%) \rangle$) denotes the scenario that classifiers are trained by 80% of G_{So}^γ (or real soybean individual leaf images, I) and then validated by 10% of G_{So}^γ (or I) before being fine-tuned by 5% (a small fraction) of real foliage images (S) for testing on 85% of S .

A. DATASET VALIDATION

The disease rate (γ) distribution of SHV images was analyzed fully according to the area covered by diseased leaves in each foliage image with respect to the whole foliage image area. Our analytic results indicate that (1) the disease rates of SHV foliage image exhibit a near normal distribution

and (2) the variable mean of $\mu = \gamma \approx 15\%$, as shown in Figure 3(c). For validating the value and usefulness of generated datasets in comparison with SHV (the natural soybean dataset), Foliagen generates a dataset with the variable mean of $\mu = \gamma = 15\%$ and including three categories (namely, rust, mosaic virus, and healthy ones) as present in the SHV dataset. All five SOTA classifiers are trained on both the SHV dataset and the generated dataset, with their performance outcomes reported in Table 2. Both datasets comprise similar numbers of images (around 1000) per category. Our validation hypothesizes that the generated dataset is representative of the natural foliage dataset (SHV) if five SOTA classifiers yield similar performance outcomes on the two datasets (treated as opaque systems). The accuracy (or F1-score) outcome of each classifier on a opaque system is viewed as one accuracy (or F1-score) feature of the opaque system, with five classifiers to obtain five accuracy (or F1-score) features for each opaque system, i.e., dataset. Table 2 lists evaluated performance outcomes in terms of accuracy and F1-score. For gauging the resemblance degree of classifier accuracy outcomes under the two datasets, cosine similarity is chosen, expressed by $\frac{A \cdot B}{\|A\| \|B\|}$, where A (or B) denotes the accuracy (or F1-score) outcomes of classifiers under one dataset (or the other dataset) as a vector, and the numerator (or denominator) is the vector dot product (or the product of the two vector Euclidean norms). A cosine similarity value closer to 1.0 signifies that the two vectors are more resemblant (with 1.0 only if A and B are identical). The accuracy values of the considered classifiers listed in Table 2 result in a cosine similarity equal to 0.9914, under the accuracy metric. Meanwhile, the F1-scores in the table give rise to the cosine similarity of 0.9997, under the F1-score metric. These cosine similarity amounts (from the two evaluation metrics) suggest that the generated dataset is closely similar to SHV.

Additionally, the geometric distances between the performance-metric vectors of the two datasets are computed in a numeric space to quantify their similarity. The average Euclidean distance per classifier based on accuracy (or F1-score) is approximately 1% (or 2%), reaffirming that the dataset generated by Foliagen is similar to the natural dataset. In general, Foliagen is expected to generate datasets that effectively capture data distributions and underlying feature representations comparable to real-world datasets, despite being constructed solely from individual-leaf images. This validates the usefulness and value of generated foliage datasets, able to serve as reliable substitutes for real data, while significantly reducing the cost and labor associated with large-scale foliage data collection and annotation.

B. EVALUATION ON REAL SOYBEAN FOLIAGE IMAGE DATASET

1) BASELINE

As baseline evaluation, the five SOTA classifiers, under various splits of SHV foliage images, were experimented with the results for $\langle S(80\%); S(10\%); S(10\%) \rangle$ shown in

Table 2, where the results are unsatisfactory. The highest achieved accuracy and F1-score values stay below 91.5% under Swin Transformer, despite the fact that the dataset set involves just two disease categories. This unsatisfactoriness is primarily attributed to insufficient training data and a lack of image heterogeneity required for high performance and robust generalization. In contrast, the Foliagen framework can generate arbitrarily large foliage datasets out of individual-leaf images available from multiple public sources, where images are captured under diverse environmental conditions to have image heterogeneity. The results under $\langle S(80\%); S(10\%); S(10\%) \rangle$ in Table 2 serve as the baseline for comparatively evaluating our generated datasets below.

2) AUGMENTING SCARCE SHV IMAGES (SAY, 5%) WITH GENERATED IMAGES

Properly annotated public natural diseased foliage images are usually scarce and imbalanced (e.g., the SHV dataset), posing a significant challenge for training classifiers to effectively classify diseased foliage images. Training solely on a small fraction of the SHV dataset (to reflect likely real-world situations where the annotated foliage image count of each disease category may be in the order of tens) yields inadequate performance. To address this inadequacy, our inherently heterogeneous foliage images generated for $\gamma = 15\%$ to train the five SOTA classifiers, which are then evaluated on 85% of the SHV dataset. Note that, in this experiment, our generated dataset includes only the same two disease categories present in the SHV dataset, along with the healthy category. From the performance results under the scenario of $\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%); S(85\%) \rangle$ as listed in Table 3, it is clear that training classifiers solely on the generated foliage images for evaluation on real-world foliage images yields comparable performance results as those under the scenario of $\langle S(5\%); S(10\%); S(85\%) \rangle$ for two classifiers, namely, ResNet50 and CBAM-ConvNeXt. Meanwhile, DenseNet121 performs better under the scenario of $\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%); S(85\%) \rangle$, where VGG19 and Swin Transformer exhibit markedly lower performance. The unsatisfactory results of the latest two are likely attributable to domain shifts between the generated and the natural datasets, caused by discrepancies in visual characteristics such as texture patterns, lighting variations, background diversity, and sensor noise, which together limit the models' ability in effective generalization for real-world data. For more accurate classification on natural foliage images, our generated dataset is combined with a small portion (say, 5%) of the SHV dataset for model training, creating the scenario of $\langle G_{So}^{15}(80\%) + S(5\%); G_{So}^{15}(10\%); S(85\%) \rangle$ for comparative evaluation, with the results listed rightmost in Table 3. The classifiers trained on the combined foliage images are hypothesized to learn the underlying disease patterns from the generated data, while the inclusion of 5% of SHV images facilitates adaptation to the structural characteristics and the distribution of the natural foliage dataset. Supporting this

TABLE 3. Comparative performance results of models trained under different data sample mixes for evaluation on 85% of S .

$\langle S(5\%); S(10\%); S(85\%) \rangle$			$\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%); S(85\%) \rangle$		$\langle G_{So}^{15}(80\%) + S(5\%); G_{So}^{15}(10\%); S(85\%) \rangle$	
Classifiers	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
VGG19	72.52	73.36	49.79	51.12	89.5	90.63
ResNet50	80.87	82.11	79.15	80.03	98.52	98.55
DenseNet121	62.85	61.37	73.14	74.42	96.43	96.48
Swin Transformer	74.3	75.59	65.35	66.49	92.17	92.41
CBAM-ConvNeXt	76.52	77.08	78.23	78.31	91.07	91.32

TABLE 4. Comparative transfer learning performance of models pre-trained on the generated dataset of G_{So}^{15} and G_{So}^5 and individual soybean leaf images (I), followed by fine-tuning on $S(5\%)$ for evaluation on $S(85\%)$.

$\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%) S(5\%); S(85\%) \rangle$			$\langle G_{So}^5(80\%); G_{So}^5(10\%) S(5\%); S(85\%) \rangle$		$\langle I(80\%); I(10\%) S(5\%); S(85\%) \rangle$	
Classifiers	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
VGG19	88.00	88.22	84.03	84.05	33.33	32.56
ResNet50	95.83	95.91	90.61	90.86	74.84	75.31
DenseNet121	94.40	94.48	90.40	90.48	35.37	34.43
Swin Transformer	92.32	91.33	87.57	88.52	69.45	69.92
CBAM-ConvNeXt	95.38	95.53	86.76	88.06	49.32	50.25

TABLE 5. Ablation study of Foliagen for Soybean and Tomato ($\gamma = 15\%$), with accuracy as the resulting measure.

$\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%); G_{So}^{15}(10\%) \rangle$					$\langle G_{To}^{15}(80\%); G_{To}^{15}(10\%); G_{To}^{15}(10\%) \rangle$			
Classifiers	Foliagen	No plant level	Random copy-paste	Unprocessed	Foliagen	No plant level	Random copy-paste	Unprocessed
VGG19	76.26	69.29	72.59	37.52	92.10	82.60	70.60	53.71
ResNet50	85.31	82.01	61.22	45.78	95.47	87.39	68.74	55.32
DenseNet121	94.47	84.66	77.97	76.66	97.85	90.32	78.32	71.63
Swin Transformer	72.36	65.94	60.75	32.58	90.18	66.53	64.97	43.84
CBAM-ConvNeXt	77.76	73.15	59.88	49.15	91.35	72.51	65.51	59.50

hypothesis, all SOTA classifiers demonstrate significantly higher performance metric values than those under the other two scenarios listed in Table 3, up to 34% improvement when compared to the scenario of $\langle S(5\%); S(10\%); S(85\%) \rangle$, mainly due to the heterogeneity of generated foliage images capable of both data augmentation and implicit regularization for improving the model's generalization to unseen foliage images. These results indicate that foliage images generated using Foliagen can mix with scarce, homogeneous, and imbalanced natural foliage images to yield a high-quality dataset for classifier training to substantially lift classification accuracy. The generated image datasets are thus useful to make scarce real-world foliage images (with just tens of them per disease category) effective and adequate for classifier training, largely curbing the often costly effort of in-field image collection and annotation.

3) TRANSFER LEARNING

Generated foliage image datasets can be leveraged for pre-training crop disease classifiers, enabling them to become generalizable classifiers able to be efficiently adapted to any real-world foliage dataset via fine-tuning with a small fraction (say, 5%) of the target dataset. To this end, the pre-trained classifiers are evaluated under the real-world SHV

dataset [13] for various scenarios, with only the evaluation results of a representative scenario, $\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%) | S(5\%); S(85\%) \rangle$, listed in the leftmost of Table 4. It can be observed from tabular results that the pre-trained classifiers with fine-tuning elevate their performance metric values noticeably, to exceed 92.3% (or 91.3%) in accuracy (or F1-score) for all but Classifier VGG19. Their performance boosts signify that pre-training under generated foliage datasets enables the classifiers to transfer the underlying features captured from generated datasets to the real-world dataset for performance improvement. It is interesting to note that Swin Transformer, which excels without transform learning (as demonstrated in Table 2), falls below ResNet50, DenseNet121, and CBAM-ConvNeXt, partly due to its limitations in (1) discriminating among disease categories without noticeable visual differences (e.g., healthy, frog-eye, and downy mildew) and (2) capturing fine-scale features (e.g., small lesions in downy mildew and frog-eye leaf spot).

When classifiers are pre-trained by a generated soybean foliage dataset with a small γ (say, 5%), they are geared for identifying soybean diseases at an early stage, especially useful for real-world field applications. To evaluate the SOTA classifiers for early disease identification, soybean foliage image datasets with $\gamma = 5\%$ were produced, with the

TABLE 6. Comparative performance evaluation results (in %) under generated soybean and tomato foliage image datasets with $\gamma = 15\%$ and 5% .

$\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%); G_{So}^{15}(10\%) \rangle$ (or $\langle G_{So}^5(80\%); G_{So}^5(10\%); G_{So}^5(10\%) \rangle$)			$\langle G_{To}^{15}(80\%); G_{To}^{15}(10\%); G_{To}^{15}(10\%) \rangle$ (or $\langle G_{To}^5(80\%); G_{To}^5(10\%); G_{To}^5(10\%) \rangle$)	
Classifiers	Accuracy	F1-score	Accuracy	F1-score
VGG19	76.26 (or 65.45)	77.17 (or 65.67)	92.10 (or 80.53)	91.83 (or 81.01)
ResNet50	85.31 (or 81.91)	85.40 (or 83.58)	95.47 (or 78.84)	95.54 (or 81.39)
DenseNet121	94.47 (or 87.56)	95.45 (or 90.81)	97.85 (or 86.93)	97.92 (or 89.47)
Swin Transformer	72.36 (or 65.38)	72.99 (or 64.50)	90.18 (or 66.53)	90.29 (or 66.71)
CBAM-ConvNeXt	77.76 (or 66.33)	79.73 (or 69.59)	91.35 (or 64.85)	91.40 (or 65.56)

outcomes of a representative scenario ($G_{So}^5(80\%); G_{So}^5(10\%) | S(5\%); S(85\%)$) listed in Table 4. As expected, the results for classifiers pre-trained on the generated dataset with $\gamma = 5\%$, show consistent performance declines, when compared with those listed on the leftmost in Table 4. This is likely due to the scarcity of disease-specific features in the dataset with such a small γ , limiting the models' ability to learn meaningful and transferable feature representations. Nonetheless, ResNet50 and DenseNet121 still exhibit impressive classification performance, with their accuracy and F1-score values all $\geq 90.4\%$.

We also pre-trained the same five classifiers on individual-leaf images, based on which Foliagen generates foliage images. The pre-trained classifiers were then fine-tuned via 5% of the SHV dataset, with their results shown in Table 4 under the scenario of ($I(80\%); I(10\%) | S(5\%); S(85\%)$). As expected, the results reveal that pre-trained classifiers using individual-leaf images perform poorly when evaluated on 85% of the SHV dataset, even after they were fine-tuned. Substantial performance gaps are observed between classifiers pre-trained on generated foliage images ($\gamma = 15\%$ or 5%) and those pre-trained on individual-leaf images. This occurs because individual-leaf images are typically zoomed-in views where disease symptoms are prominent and occupy a large portion of the image area. In contrast, foliage images contain much smaller, spatially dispersed diseased regions. Consequently, classifiers pre-trained on individual-leaf datasets fail to effectively capture the subtle and distributed disease patterns present in foliage images. Only diseased foliage datasets with proper annotations can train classifiers for effective and accurate classification of diseased foliage images.

Note that foliage image annotations are usually very expensive and time-consuming, making it infeasible to have quality-annotated foliage image datasets. As a result, our inexpensive and fast framework, Foliagen, for generating properly annotated disease foliage datasets of arbitrarily large sizes at any disease rate, is essential for (1) uplifting disease foliage image classification performance, (2) early disease detection, with lower γ value for timely notification to the

farmers for appropriate actions, and (3) the further study in diseased foliage classification.

Due to the lack of publicly available real-world diseased tomato foliage datasets, transfer learning evaluation could not be conducted. Nevertheless, similar trends to those observed for soybean are expected, as the classifiers that learn tomato underlying disease patterns from generated foliage images can also be transferred to better real-world foliage disease classification.

C. ABLATION STUDY

The Foliagen framework aims to ensure that generated foliage images closely resemble real ones. The ablation study of Foliagen is listed in Table 5 for both soybean and tomato. The three columns after "Foliagen" reflect the accuracy measures of SOTA classifiers when classifying foliage images produced by Foliagen with (1) the plant-level component excluded, with foliage images generated following the spiral phyllotactic pattern disregard to individual plants (indicated by "No plant level"), (2) randomly copy-pasting leaves without enforcing phyllotaxis (indicated by "Random copy-paste"), which resembles a simple copy and paste technique [30], and (3) background in individual leaf images retained (indicated by "Unprocessed").

The results demonstrate that the final Foliagen framework consistently outperforms all prior variants across the evaluated SOTA models for both soybean and tomato. As expected, retaining background content in leaf images degrades performance significantly, underscoring the importance of precise leaf segmentation for realistic foliage synthesis. Furthermore, experiments with random leaf arrangements confirm that preserving natural phyllotaxis is essential for accurately mimicking the crop structure. Lastly, it is clearly revealed that plant-level phyllotaxis plays a vital role in enhancing the naturalness and overall quality of the generated images.

D. EVALUATION ON GENERATED FOLIAGE IMAGE DATASETS

We generated many diseased foliage datasets of soybean and tomato based on the ASDID [10], Kaggle dataset [12], and

PlantVillage dataset [11] (with 9 primary disease categories), under a range of γ for classifier evaluation. The evaluation metric outcomes under a generated dataset with a larger γ are expected to be higher because more diseased leaves exist in each foliage image, enabling the classifiers to capture more underlying features for better accurate and effective classification. The comparative performance evaluation results are obtained for γ ranging from 1% to 50% (detailed at VII-H2), and they are found to follow similar trends. Here, the results of foliage images for $\gamma = 5\%$ and $\gamma = 15\%$ are listed in Table 6 as representative ones, denoted respectively by the scenarios of $\langle G_{So}^{15}(80\%); G_{So}^{15}(10\%); G_{So}^{15}(10\%) \rangle$ and $\langle G_{So}^5(80\%); G_{So}^5(10\%); G_{So}^5(10\%) \rangle$ for Soybean and by the scenarios of $\langle G_{To}^{15}(80\%); G_{To}^{15}(10\%); G_{To}^{15}(10\%) \rangle$ and $\langle G_{To}^5(80\%); G_{To}^5(10\%); G_{To}^5(10\%) \rangle$ for Tomato. Note that $\gamma = 5\%$ signifies an early disease stage, while $\gamma = 15\%$ represents the observed disease rate in the natural SHV dataset. As evident from the Table 6, the considered SOTA models all perform better under $\gamma = 15\%$ than under $\gamma = 5\%$, with respect to accuracy and F1-score, underscoring the fact that they tend to struggle in early disease detection (under $\gamma = 5\%$). The evaluation results exhibit that DenseNet121 outperforms the rest consistently, making it the most desirable classifier for in-field foliage image classification.

E. COMPARISON WITH OTHER GENERATION METHODS

We compared Foliagen with several state-of-the-art (SOTA) image generation counterparts for image generation and augmentation in terms of generated image quality. Specifically, we evaluated Google's Simple Copy-Paste (SCP) augmentation method [30], a probabilistic diffusion model (DDPM [33]), and generative AI models of PixelCNN [38] and RDCGAN [39]. As illustrated in Figure 10 of Appendix, DDPM, PixenCNN, and RDCGAN all failed to produce realistic diseased foliage images because of the large number of leaves present in each foliage image and the limited amount of training data available. While SCP generated visually plausible foliage images, it lacks structural awareness, resulting in disease leaves being randomly distributed throughout the foliage without preserving the natural plant architecture.

In addition, Foliagen and the RCP (Random Copy-Paste) method [30] are quantitatively compared under five established SOTA image classifiers, where RCP denotes a structure-agnostic implementation of the Google's SCP strategy. The classification results, presented in Table 5, demonstrate that foliage images generated by Foliagen consistently excels across all classifiers, indicating that biological foliage structure preservation can substantially improve the quality and utility of the generated foliage images.

V. DISCUSSION

Foliagen demonstrates that realistic crop foliage images can be generated from publicly available individual leaf datasets while maintaining biologically plausible plant structures and spatial disease distributions. Experimental results show

that classifiers trained with Foliagen-generated datasets consistently outperform those trained using conventional data generation methods, highlighting the importance of structural realism and controllable disease prevalence for improving in-field disease recognition. Moreover, the ability to generate arbitrarily large, fully annotated foliage datasets helps address the limited availability of expert-annotated field images in agricultural computer vision.

Although the proposed framework is designed to be general, its effectiveness depends on the quality and diversity of the available individual leaf images. For crops with substantially different leaf morphologies, overlapping canopies, or complex branching structures, incorporating crop-specific architectural characteristics may be necessary to maintain biological realism. Likewise, diseases whose visual symptoms vary with growth stage, environmental conditions, or interactions among neighboring leaves may require more advanced disease distribution strategies. Differences in leaf shape, venation, texture, and canopy density across crop species may also affect the realism of the generated foliage, providing opportunities for further extending Foliagen toward a more universal crop image generation framework.

VI. CONCLUDING REMARKS

This article introduces a framework (called Foliagen) to generate fully-annotated and arbitrarily-sized datasets of crop foliage images, with a given rate of diseased leaves (γ) in each generated foliage image. Following plant structural characteristics and diseased leaf distributions, it produces foliage image datasets to cover all disease categories that exist in publicly available datasets of individual leaf images. The generated foliage images closely resemble those captured by UAV-mounted cameras in crop farm fields with the disease rate of γ , making it possible to better and objectively evaluate state-of-the-art (SOTA) or future leaf disease classifiers for real-world applications. Given limited natural crop foliage images with expert annotations so far, the abundant foliage image datasets produced easily and quickly by Foliagen are valuable for training crop disease classifiers for in-field application. Generated foliage images can augment scarcely available natural foliage images so as to better train classifiers for high accuracy. In addition, when a small number of natural quality-annotated foliage images (say, only tens of them per disease category) are available from the field, a pre-trained classifier can be fine-tuned using those available images for in-field application to yield high classification performance. While Foliagen is exemplified for classifying the diseases of soybean and tomato via SOTA models in this paper, it is readily useful for the disease classification of other crops and also for objectively evaluating future disease classifiers aiming at real-world applications.

Future work will extend Foliagen to a broader range of crop species with diverse leaf morphologies and canopy structures. Since the exact locations of diseased leaves are known during image generation, the framework can also be extended to generate datasets for semantic and instance segmentation.

In addition, incorporating crop-specific structural characteristics, disease progression, and environmental variations could further improve the realism and generalization of the generated datasets, enabling broader applications in precision agriculture.

VII. DATA AND CODE AVAILABILITY

The datasets that support the findings of our study are available in hugging face (Foliagen). Our Foliagen code is available in github (Foliagen).

APPENDIX

A. AVAILABLE DATASETS

All the publicly available datasets that were inspected for the Foliagen framework are listed next. These datasets were analyzed in terms of image characteristics, disease coverage, annotation quality, and suitability for foliage-level synthesis. Particular attention was given to the availability of healthy and diseased leaf samples, image resolution, and background conditions, as these factors directly influence the quality of foliage images generated by the Foliagen framework. The selected datasets provide the foundational single-leaf imagery used to construct realistic foliage-level representations for downstream disease classification tasks.

PlantVillage dataset. PlantVillage [11] contains over 54,300 expertly curated healthy and diseased leaf images from various plants, including thirteen major crop species like soybean, tomato, etc. Its leaf classification has been attempted by GoogleNet [56] and AlexNet [57] to attain high accuracy.

PlantDoc dataset. The PlantDoc dataset [23] contains 2,598 single-leaf images of 17 disease categories across 13 plant species, including tomato and soybean. With its images gathered in a controlled laboratory environment, this dataset has limited applicability in real-world scenarios.

PlantSeg dataset. The PlantSeg dataset represents a major advance in plant disease research with 7,774 in-the-wild images across 115 disease classes and 34 plant species, featuring expert-annotated pixel-level segmentation masks. Unlike traditional lab-captured datasets like PlantVillage, PlantSeg sources its images from internet search engines, capturing real-world complexities such as occlusions, variable lighting, and complex backgrounds essential for field deployment. This addresses the domain gap limiting prior datasets, though its heterogeneity—ranging from single leaves to full foliage with occasional text overlays—poses challenges for standardized model training.

ASDID. Auburn Soybean Diseased Image Dataset (ASDID) [10] provides high-quality individual leaf images of the soybean plant, covering 9 disease categories, namely, bacterial blight, cercospora leaf blight, downy mildew, frogeye leaf spot, soybean rust, target spot, and potassium deficiency. The dataset was captured primarily at the EV Smith Agricultural Research Station in Tallassee, Alabama,

and added with 80 images per disease category from the publicly available Image Database of Plant Disease Symptoms (PDDb) [58]. Several ML-based classifiers [43], [49], [50] were employed to evaluate the dataset, with DenseNet201 [43] achieving the highest performance.

Kaggle dataset. Soybean Diseased Leaf Dataset contains individual leaf images from Kaggle [12], embracing 10 disease categories and having an artificially generated complex background added to each image. Among the 10 diseases, only mosaic virus and sudden death syndrome are considered.

FieldPlant dataset. FieldPlant [22] is a dataset of individual crop leaf images annotated by pathologists, containing 8,629 images across 27 disease categories for three crops, including tomato. It aims at practical crop disease classification with every leaf image involving a complex background.

Tomato-Village dataset. Compensating for the negative effects due to the laboratory-controlled environment setup for gathering PlantVillage's leaf images, the Tomato-Village dataset [25] contains real-world images, which belong to three groups, respectively for (1) multi-class tomato disease classification, (2) multi-label tomato disease classification, and (3) object detection-based tomato disease detection. This dataset covers seven disease categories, namely, early blight, late blight, leaf miner, magnesium deficiency, nitrogen deficiency, potassium deficiency, and spotted wilt virus, for multi-class classification applications.

MH-SoyaHealthVision dataset. As far as we know, MH-SoyaHealthVision [13] is the only public and well-annotated dataset of diseased soybean foliage images. The dataset provides both ground-level leaf images and foliage leaf images, collected using a UAV, from the soybean fields of Maharashtra, India. It comprises a total of 5,680 high-resolution images grouped into (1) single leaf images of 4 diseases categories (i.e., frogeye, mosaic virus, septoria brown spot, and rust), two types of pest attacks, and healthy leaves and (2) UAV-captured soybean foliage images, which belong to two disease categories of mosaic virus and rust, plus the healthy category.

B. SINGLE-LEAF IMAGES

Samples of single-leaf images for each disease category, along with healthy samples, are illustrated in Figure 5 for soybean and in Figure 6 for tomato. These single-leaf images are obtained from ASDID [10], Kaggle dataset [12], and Tomato village [25], which are preprocessed and subsequently used to generate foliage images for each disease category using our Foliagen framework.

C. GENERATED FOLIAGE IMAGES

Our Foliagen framework, after preprocessing the single-leaf images, creates foliage images as described in Section III in the main text. The generated diseased foliage images for both Soybean and Tomato are shown next in Figure 7 and Figure 8, respectively.

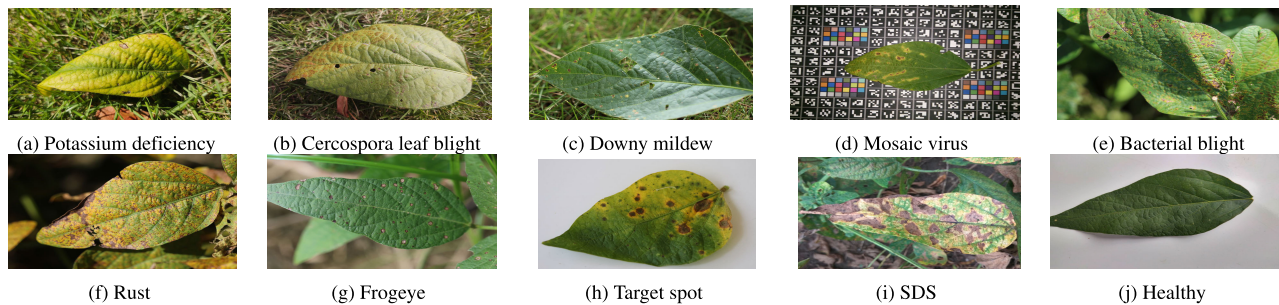


FIGURE 5. Soybean single leaf images of 9 diseases, labeled by (a) to (i), with SDS denoting Sudden Death Syndrome.

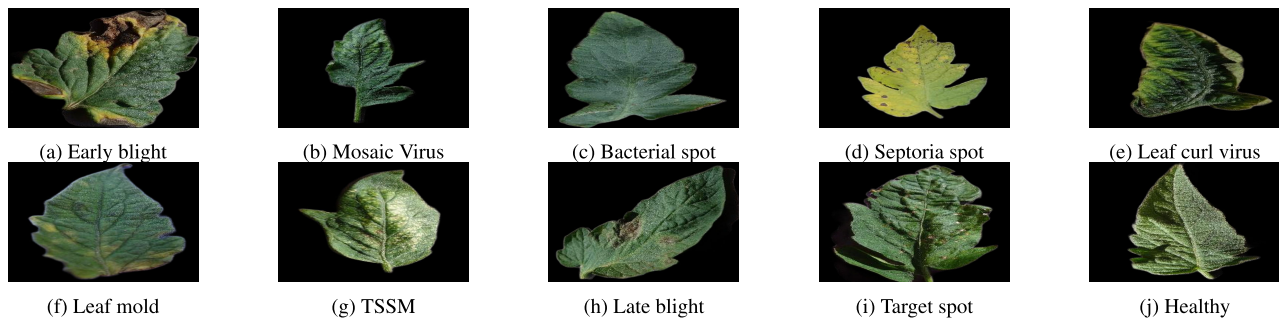


FIGURE 6. Tomato single leaf images of 9 diseases, labeled by (a) to (i), with TSSM denoting Two Spotted Spider Mite.

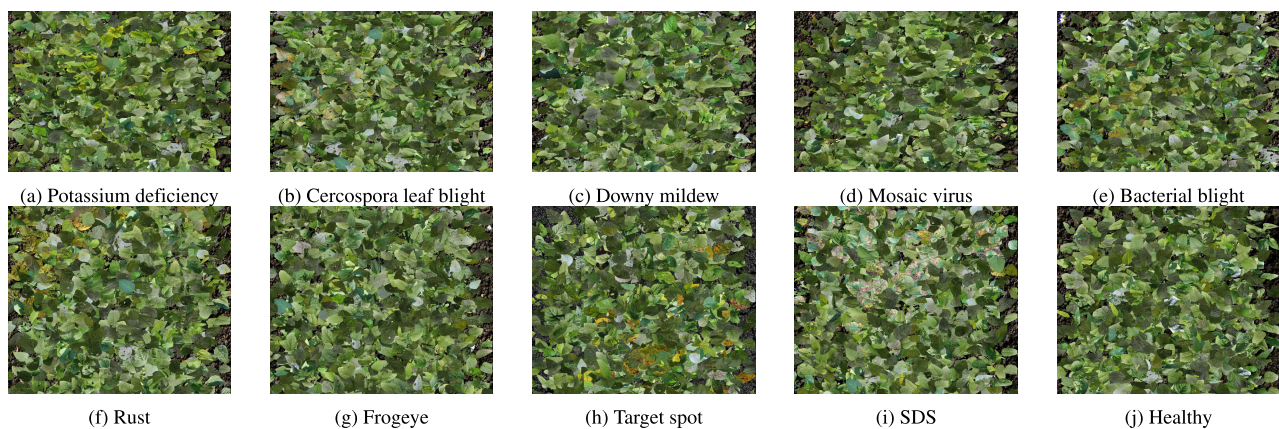


FIGURE 7. Generated soybean foliage images of 9 diseases, labeled by (a) to (i), for $\gamma = 15\%$, with SDS denoting Sudden Death Syndrome.

D. FOLIAGE GENERATION USING GENERATIVE MODEL

Recently, generative models have been used to generate almost all digital artifacts, including photos, videos, and texts. In the context of photos, the generative adversarial networks (GANs) and diffusion models have been used extensively. While these models performed very well in many aspects, there is very little evidence that they are capable of generating diseased plant foliage images. Hence, we experiment with recent state-of-the-art (SOTA) models, namely, Regularized Deep Convolutional Generative Adversarial Network (RDCGAN), Denoising Diffusion Probabilistic Model (DDPM), and PixelCNN, to generate diseased foliage images from the real-world MH-SoyaVisionHealth dataset. All generative models are trained for 100 epochs before generating images from them, with their outcomes depicted in Figure 10.

E. NATURAL FOLIAGE IMAGE

Foliagen is based on observations of natural foliage and findings reported in published studies to generate high-quality data for foliage disease classification. Our work evaluates MH-SoyaVisionHealth [13], a natural soybean foliage dataset containing two disease categories: rust and mosaic virus. Figure 11 presents real-world soybean foliage images collected using UAVs at a farm in India.

F. ALGORITHMS

Foliagen comprises three core algorithms for generating realistic foliage images of arbitrary crop species. The key steps of each algorithm are briefly described next.

Spiral Phyllotaxis. The placement of leaves in a single plant follows a spiral phyllotaxis pattern (as illustrated in Algorithm 1), which is inspired by the natural arrangement of

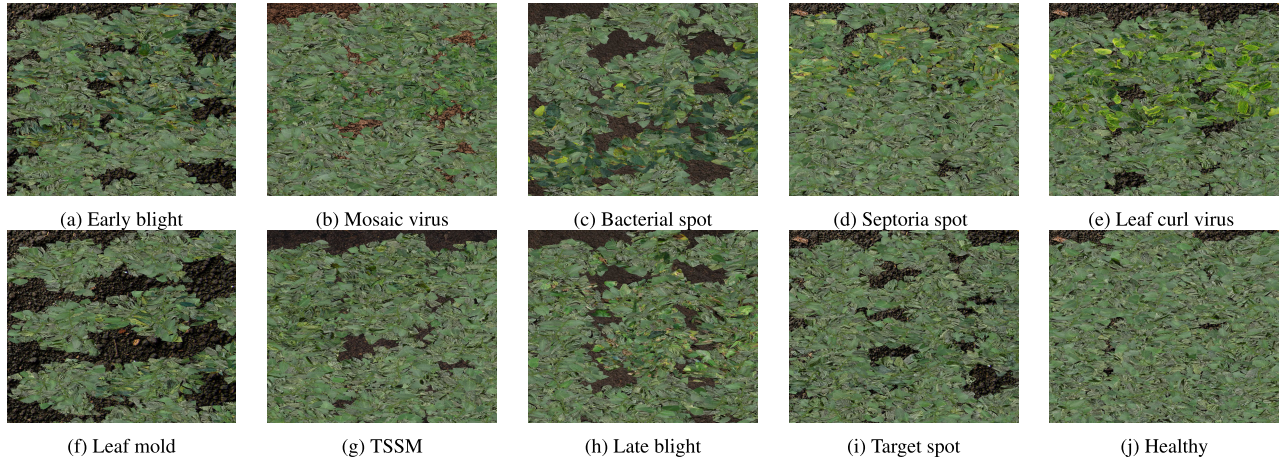


FIGURE 8. Generated tomato foliage images of 9 diseases, labeled by (a) to (i), for $\gamma = 15\%$, with TSSM denoting Two Spotted Spider Mite.

Algorithm 1 Leaf Coordinates via Spiral Phyllotaxis

Require: Center (C_o^c)(x_0, y_0), Number of leaves $\mathcal{N}_l$

Ensure: Leaf coordinates $\mathcal{C}_o$

- 1: **Initialize:** Scaling factor (s) $\leftarrow 35$, $\mathcal{C}_o \leftarrow \emptyset$
 - 2: Golden angle (θ_g^{rad}) $\leftarrow 137.5^\circ \times \frac{\pi}{180}$
 - 3: **for** $n = 1$ to $\mathcal{N}_l$ **do**
 - 4: $\theta_n \leftarrow n \cdot \theta_g$, $r_n \leftarrow s\sqrt{n}$
 - 5: $x_n \leftarrow x_0 + r_n \cos(\theta_n)$
 - 6: $y_n \leftarrow y_0 + r_n \sin(\theta_n)$
 - 7: $\mathcal{C}_o \leftarrow \mathcal{C}_o \cup \{(x_n, y_n)\}$
 - 8: **end for**
 - 9: **return** $\mathcal{C}_o$
-

leaves observed in soybean (or tomato) like plants. Starting from the center of the plant, each new leaf is positioned by rotating it by the golden angle (137.5°) relative to the previous leaf while gradually increasing its distance from the center. This approach distributes leaves evenly in all directions, reducing leaf overlap and producing a realistic canopy structure. A scaling factor controls leaf spacing (Line 1), allowing the density of the generated plant to be adjusted while preserving the natural spiral pattern (see Line 4). As a result, Algorithm 1 creates visually realistic leaf arrangements using a simple and computationally efficient procedure.

Plant Generation. Given a target disease class and the desired disease prevalence, Algorithm 2 generates a single plant by first determining the total number of leaves and the corresponding numbers of healthy and diseased leaves (Lines 2 and 3). Leaf positions and orientations are then computed (see Line 6) using Algorithm 1, resulting in a biologically plausible spatial arrangement around the plant center. Based on the assigned disease category, each leaf is subsequently selected from the corresponding leaf pool, progressively scaled to introduce depth perception (see Line 17), rotated to the appropriate orientation before pasted onto its spiral phyllotaxis coordinate (as specified in

Algorithm 2 Plant Generation

Require: Position disease rate γ_p , disease type δ , plant image size ($H \times W$)

Ensure: Single plant image $\mathcal{I}_s$

- 1: **Initialize:** Depth scale factor (s_d), depth scale offset (Δs)
 - 2: No. of leaves ($\mathcal{N}_l$) $\in [\min, \max]$
 - 3: $\mathcal{N}_{hl} \leftarrow \text{NumHealthyLeaves}(\mathcal{N}_l, \gamma_p)$
 - 4: $\mathcal{I}_s \leftarrow$ single plant image
 - 5: $\mathcal{C}_o^c \leftarrow$ center coordinates
 - 6: $\mathcal{C}_o \leftarrow \text{GetLeafCoords}(\mathcal{C}_o^c, \mathcal{N}_l)$
 - 7: $\mathcal{A} \leftarrow \text{GetLeafAngles}(\mathcal{C}_o^c, \mathcal{C}_o)$
 - 8: **for** $i = 1$ to $|\mathcal{C}_o|$ **do**
 - 9: **if** $i \bmod 2 = 0$ **and** $\mathcal{N}_{hl} > 0$ **then**
 - 10: $\delta_i \leftarrow \text{healthy}$
 - 11: $\mathcal{N}_{hl} \leftarrow \mathcal{N}_{hl} - 1$
 - 12: **else**
 - 13: $\delta_i \leftarrow \delta$
 - 14: **end if**
 - 15: $s_d \leftarrow s_d + \Delta s \cdot i$
 - 16: $\mathcal{L} \leftarrow \text{GetLeaves}(\delta_i, \mathcal{A}[i])$
 - 17: $\mathcal{L}_s \leftarrow \text{Scale}(\mathcal{L}, s_d)$
 - 18: $\mathcal{I}_s \leftarrow \text{PasteLeaf}(\mathcal{I}_s, \mathcal{L}_s, \mathcal{C}_o[i])$
 - 19: **end for**
 - 20: **return** $\mathcal{I}_s$
-

Line 18), and composited onto the soil background at its designated location. By iteratively repeating this process for all leaves, Algorithm 2 produces a realistic single-plant image with controllable disease severity while preserving the natural spatial organization and visual characteristics of the foliage. The alternating scheme listed in Lines 9 – 14 ensures a balanced distribution of healthy and diseased leaves across both depthwise and planar dimensions.

Foliage Generation. Algorithm 3 generates composite foliage images by arranging the single-plant images produced using Algorithm 2. It first determines the number of disease clusters and distributes the overall disease severity among

Algorithm 3 Foliage Generation**Require:** Disease rate γ , disease type δ , grid $G(\alpha, \beta)$ **Ensure:** Composite foliage image $\mathcal{I}$

```

1: Initialize:  $\mathcal{C} \leftarrow \emptyset$ ,  $\gamma \leftarrow []$ 
2:  $N \leftarrow \alpha \cdot \beta$ ,  $\mathcal{I} \leftarrow$  background image
3:  $n_c \leftarrow \text{GetNumClusters}(\gamma)$ 
4:  $\gamma_c \leftarrow \text{GetClusterRates}(\gamma, n_c)$ 
5: for  $i = 1$  to  $n_c$  do
6:    $h \leftarrow \text{GetClusterIndex}(1, N)$ 
7:    $\mathcal{N}(h) \leftarrow \text{GetNeighbors}(h)$ 
8:    $\mathcal{C} \leftarrow \mathcal{C} \cup \mathcal{N}(h) \cup \{h\}$ 
9:    $K \leftarrow |\mathcal{N}(h)| + 1$ 
10:   $\mathbf{r} \leftarrow \text{NormalizedSplit}(\gamma_c[i], K)$ 
11:   Assign  $\max(\mathbf{r})$  to  $h$ 
12:   Append  $\mathbf{r}$  to  $\gamma$ 
13: end for
14: for each position  $p \in [1 \dots N]$  do
15:   if  $p \in \mathcal{C}$  then
16:      $\mathcal{P}_p \leftarrow \text{GeneratePlant}(\gamma_p, \delta)$ 
17:   else
18:      $\mathcal{P}_p \leftarrow \text{GeneratePlant}(0)$ 
19:   end if
20:   Paste  $\mathcal{P}_p$  onto  $\mathcal{I}$  at  $\text{coord}(p)$ 
21: end for
22: return  $\mathcal{I}$ 

```

TABLE 7. Comparative performance evaluation results (in %) under the generated tomato foliage image dataset with 2× upscaled leaf size.

Models	Accuracy	F1-score	Precision	Recall
VGG19 [49]	65.23	65.41	65.59	65.23
ResNet50 [50]	76.73	77.61	78.51	76.73
DenseNet121 [43]	83.78	83.14	82.38	83.92
Swin Transformer [46]	67.95	68.96	70.02	67.91
CBAM-ConvNeXt [8]	71.88	71.96	72.03	71.88

these clusters. Cluster centers are randomly selected on the foliage grid, and neighboring plants are assigned related disease rates to create localized infection patterns that closely resemble those observed under natural field conditions. A single plant is then generated for each grid position (see Line 16 of Algorithm 3) according to its assigned disease rate, while plants outside the selected clusters are synthesized as healthy (see Line 18). Finally, all generated plants are composited onto a background image, producing realistic foliage images with controllable disease prevalence, spatial distribution, and natural variability.

G. EVALUATION METRICS

In classification tasks, predictions are evaluated using true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), representing correctly and incorrectly classified instances. Based on these, four metrics are calculated for the performance evaluation of classifiers on generated and natural datasets, which are detailed next.

Accuracy measures the overall proportion of correctly classified instances among all predictions. It provides a general view of model performance but can be misleading when the dataset is imbalanced. It is most useful when all classes are represented relatively equally. Its formula is shown next.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision measures how many predicted positive instances are actually positive, calculated as shown next. A high precision indicates that the model makes few false positive errors. It is important in tasks where false alarms should be minimized.

Recall measures how many actual positive instances are correctly identified by the model. A high recall means the model misses few positive cases. It is crucial when missing positives is costly. Its formula is shown next.

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}$$

F1-score is the harmonic mean of precision and recall, balancing both metrics into a single value. It is especially useful when classes are imbalanced or when both false positives and false negatives matter. A higher F1-score indicates better combined precision and recall performance. The calculation formula is presented next.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

H. ADDITIONAL EXPERIMENTAL RESULTS

We conducted extensive experimentation with multiple strategies aimed at strengthening the viability and reliability of our dataset. These efforts included (1) exploring alternative data preparation and augmentation techniques and (2) comparing the results of Foliagen with the baselines to ensure that its resulting dataset captures realistic variability while preserving essential pathological characteristics. The outcomes of these investigations are presented in the following sections.

1) VARYING LEAF SIZES

The current version of the dataset reduces the size of individual leaves to a similar size as that of natural foliage leaves. This step gave a huge performance hike in the evaluated model. Many variations in leaf sizes, maintaining their aspect ratio, were used to create the dataset and were evaluated. One of the experimental results is shown in Table 7. It is vivid from the table that upscaling the leaf size leads to degraded performance for all classifiers; this is also true when the leaves are downscaled in size.

2) SENSITIVITY ANALYSIS OF DISEASE RATE

Figure 9 illustrates the sensitivity analytic results across disease rates, revealing consistent performance patterns among the classifiers. Accuracy increases for all models as

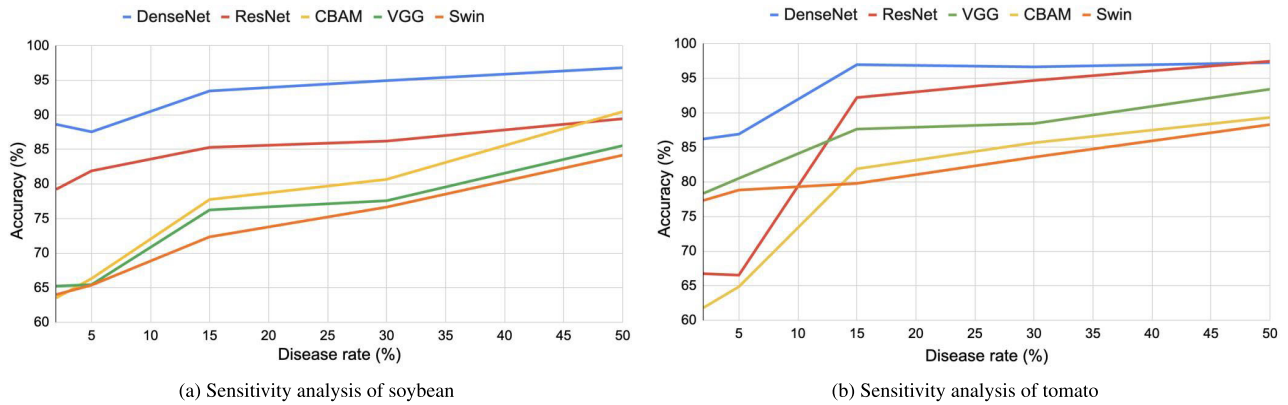


FIGURE 9. Sensitivity analysis on the disease rate for soybean and tomato.

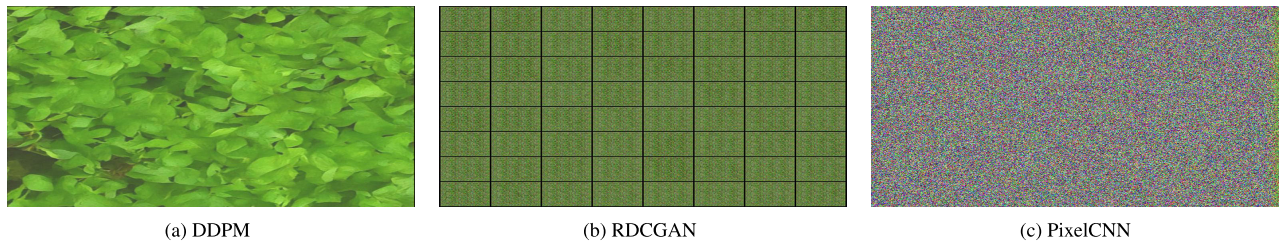


FIGURE 10. Foliage generated using generative models.



FIGURE 11. Sample soybean foliage images from the MH-SoyaVisionHealth dataset [13].

TABLE 8. Comparative performance results for models trained using 5 categories of generated images with starkly different features and with similar features (The classifiers were trained on $G_{So}^{15}(80\%)$; $G_{So}^{15}(10\%)$; $G_{So}^{15}(10\%)$ data split).

Models	Similar categories				Dissimilar categories			
	Accuracy	F1-score	Precision	Recall	Accuracy	F1-score	Precision	Recall
VGG19	45.07	45.29	45.19	45.38	64.97	63.12	62.19	64.08
ResNet50	83.16	85.80	88.45	83.31	89.21	90.07	91.11	89.05
DenseNet121	91.64	92.20	92.8	91.60	93.31	92.47	91.67	93.29
Swin Transformer	40.15	38.80	37.67	40.00	93.48	93.01	92.67	93.35
CBAM-ConvNeXt	90.78	91.15	91.54	90.77	91.78	91.66	91.54	91.77

disease prevalence intensifies, with DenseNet maintaining the highest and most stable performance levels across the full disease rate range and ResNet showing a similarly smooth upward trend. VGG improves more gradually, becoming competitive only at higher disease rates, whereas CBAM remains highly sensitive to disease severity, performing poorly at low levels but rising sharply once pathological cues become pronounced. Swin consistently yields the weakest performance with limited benefit from increased disease information.

3) EXPLAINING BEHAVIOURS OF SOTA CLASSIFIERS ON OUR DATASET

After extensive experimentation with different variations of the dataset under the five state-of-the-art (SOTA) classifiers, we observed that models performed better on different variants of the dataset. Especially, the Swin transformer performed better than other models for a dataset containing disease categories displaying starkly different (distinguishable by the human eye) features than the dataset displaying similar leaf symptoms. Two datasets with 4 disease categories

each were created with (1) diseases displaying starkly different features (including mosaic virus, rust, sudden death syndrome, and target spot) and (2) diseases displaying similar features (including downey mildew, mosaic virus, frogeye, and potassium deficiency), indicated by “Dissimilar categories” and “Similar categories”, respectively, in Table 8. Note that both datasets contain healthy foliage images as well, making the datasets have 5 categories.

I. CONFIGURATION FILE

As depicted in Figure 3 of the main text, all levels of foliage generation require plant-specific information to generate high-quality foliage images. The configuration file includes the number of leaves, disease rate, and other factors listed next.

```
{
  "num_leaves": 50,
  "diseases": "list_of_diseases",
  "foliage_size": "(1024, 1500)",
  "single_plant_size": 512,
  "single_leaf_size": 70,
  "num_plants": 16,
  "plant_offset": 100,
  "disease_rate": 15,
  "background_image_path": "<path>",
  "input_path": "<path_to_input_images>",
  "output_path": "<path_to_output_images>",
  "type": "tomato",
  "leaf_spacing": 60
}
```

REFERENCES

- [1] M. O. Triticum, “Pathogens, precipitation and produce prices,” *Nature Climate Change*, vol. 11, no. 8, p. 635, 2021.
- [2] C. A. Bradley et al., “Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2015 to 2019,” *Plant Health Prog.*, vol. 22, no. 4, pp. 483–495, 2021. [Online]. Available: <https://api.semanticscholar.org/CorpusID:240990424>
- [3] A. Abbas, S. Jain, M. Gour, and S. Vankudothu, “Tomato plant disease detection using transfer learning with C-GAN synthetic images,” *Comput. Electron. Agricult.*, vol. 187, Aug. 2021, Art. no. 106279.
- [4] A. Karlekar and A. Seal, “SoyNet: Soybean leaf diseases classification,” *Comput. Electron. Agricult.*, vol. 172, May 2020, Art. no. 105342.
- [5] R. Pan, J. Lin, J. Cai, L. Zhang, J. Liu, X. Wen, X. Chen, and X. Zhang, “A two-stage feature aggregation network for multi-category soybean leaf disease identification,” *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 35, no. 8, Sep. 2023, Art. no. 101669.
- [6] J. Sun, X. Hao, and M. Zhao, “Image recognition of soybean leaf disease based on bilateral branching network BBN,” in *Proc. 4th Int. Conf. Bioinf. Intell. Comput.*, Jan. 2024, pp. 47–53.
- [7] Y. Sun, L. Ning, B. Zhao, and J. Yan, “Tomato leaf disease classification by combining EfficientNetv2 and a Swin transformer,” *Appl. Sci.*, vol. 14, no. 17, p. 7472, Aug. 2024.
- [8] Q. Wu, X. Ma, H. Liu, C. Bi, H. Yu, M. Liang, J. Zhang, Q. Li, Y. Tang, and G. Ye, “A classification method for soybean leaf diseases based on an improved ConvNeXt model,” *Sci. Rep.*, vol. 13, no. 1, p. 19141, Nov. 2023.
- [9] V. Yogabalajee, K. Sundaram, and K. Kanagaraj, “Soybean leaf disease classification using enhanced DenseNet121,” in *Proc. 10th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Mar. 2024, pp. 1515–1520.
- [10] N. Bevers, E. J. Sikora, and N. B. Hardy, “Soybean disease identification using original field images and transfer learning with convolutional neural networks,” *Comput. Electron. Agricult.*, vol. 203, Dec. 2022, Art. no. 107449.
- [11] D. P. Hughes and M. Salathe, “An open access repository of images on plant health to enable the development of mobile disease diagnostics,” 2015, *arXiv:1511.08060*.
- [12] SIVM205. (2023). *Soybean Diseased Leaf Dataset*. [Online]. Available: <https://www.kaggle.com/datasets/sivm205/soybean-diseased-leaf-dataset>
- [13] S. Shinde and V. Attar, “An Indian UAV and leaf image dataset for integrated crop health assessment of soybean crop,” *Data Brief*, vol. 60, Jun. 2025, Art. no. 111517.
- [14] A. Benfenati, D. Bolzi, P. Causin, and R. Oberti, “A deep learning generative model approach for image synthesis of plant leaves,” *PLoS ONE*, vol. 17, no. 11, Nov. 2022, Art. no. e0276972.
- [15] D. B. Ward, P. Moghadam, and N. Hudson, “Deep leaf segmentation using synthetic data,” in *Proc. Comput. Vis. Problems Plant Phenotyping (CVPPP) Workshop BMVC*, 2018, p. 327.
- [16] O. Khare, S. Mane, H. Kulkarni, and N. Barve, “LeafNST: An improved data augmentation method for classification of plant disease using object-based neural style transfer,” *Discover Artif. Intell.*, vol. 4, no. 1, p. 50, Jul. 2024.
- [17] Q. H. Cap, H. Uga, S. Kagiwada, and H. Iyatomi, “LeafGAN: An effective data augmentation method for practical plant disease diagnosis,” *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 2, pp. 1258–1267, Apr. 2022.
- [18] Y. Yang, D. Mao, H. Santo, Y. Matsushita, and F. Okura, “NeuraLeaf: Neural parametric leaf models with shape and deformation disentanglement,” in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2025, pp. 28167–28176.
- [19] K. J. Niklas, “The role of phyllotactic pattern as a ‘development constraint’ on the interception of light by leaf surfaces,” *Evolution*, vol. 42, no. 1, pp. 1–16, 1988.
- [20] K. Kanahama, K. Iwabuchi, and M. Suda, “Phyllotaxis on the main shoot of wild tomato plants calculated by the parastichy system,” *Tohoku J. Agricult. Res.*, vol. 44, no. 1, pp. 15–22, 1994. [Online]. Available: <https://tohoku.repo.nii.ac.jp/record/65366/files/KJ00000715054.pdf>
- [21] J. Coussemment, M. Henke, P. Lootens, I. Roldán-Ruiz, K. Steppe, and T. De Swaef, “Modelling leaf spectral properties in a soybean functional-structural plant model by integrating the prospect radiative transfer model,” *Ann. Botany*, vol. 122, no. 4, pp. 669–676, Sep. 2018.
- [22] E. Moupojou, A. Tagne, F. Retraint, A. Tadonkemwa, D. Wilfried, H. Tapamo, and M. Nkenlifack, “FieldPlant: A dataset of field plant images for plant disease detection and classification with deep learning,” *IEEE Access*, vol. 11, pp. 35398–35410, 2023.
- [23] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, “PlantDoc: A dataset for visual plant disease detection,” in *Proc. 7th ACM IKDD CoDS 25th COMAD*, Jan. 2020, pp. 249–253.
- [24] T. Wei, Z. Chen, X. Yu, S. Chapman, P. Melloy, and Z. Huang, “A large-scale in-the-wild dataset for plant disease segmentation,” *Sci. Data*, vol. 13, no. 1, p. 205, Feb. 2026.
- [25] M. Gehlot, R. K. Saxena, and G. C. Gandhi, “‘Tomato-village’: A dataset for end-to-end tomato disease detection in a real-world environment,” *Multimedia Syst.*, vol. 29, no. 6, pp. 3305–3328, Dec. 2023.
- [26] Q. Guo, “Enrich the content of the image using context-aware copy paste,” 2024, *arXiv:2407.08151*.
- [27] K. Higuchi, T. Mizuhashi, F. Matulic, and T. Igarashi, “Interactive generation of image variations for copy-paste data augmentation,” in *Proc. Extended Abstr. CHI Conf. Human Factors Comput. Syst.*, Apr. 2023, pp. 1–7.
- [28] D. Dwibedi, I. Misra, and M. Hebert, “Cut, paste and learn: Surprisingly easy synthesis for instance detection,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1310–1319.
- [29] C. S. Mesnage, A. Corbett, J. Curry, S. Rowlands, A. Dutta, R. Everson, and B. I. Simmons, “Copy-paste augmentation improves automatic species identification in camera trap images,” *Ecology Evol.*, vol. 15, no. 11, p. 72357, Nov. 2025.
- [30] G. Ghiasi, Y. Cui, A. Srinivas, R. Qian, T.-Y. Lin, E. D. Cubuk, Q. V. Le, and B. I. Zoph, “Simple copy-paste is a strong data augmentation method for instance segmentation,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 2917–2927.
- [31] T. Remez, J. Huang, and M. Brown, “Learning to segment via cut-and-paste,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 39–54.
- [32] W.-H. Shen and M.-L. Li, “Copy-paste image augmentation with Poisson image editing for ultrasound instance segmentation learning,” in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2023, pp. 1–3.

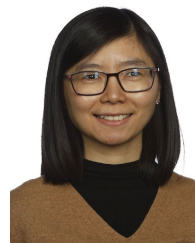
- [33] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020, pp. 6840–6851.
- [34] D. Bradley, D. Nowrouzezahrai, and P. Beardsley, "Image-based reconstruction and synthesis of dense foliage," *ACM Trans. Graph.*, vol. 32, no. 4, pp. 1–10, Jul. 2013.
- [35] T. Moon, H. Choi, D. Kim, I. Hwang, J. Kim, J. Shin, and J. E. Son, "Autonomous construction of parameterizable 3D leaf models from scanned sweet pepper leaves with deep generative networks," *Silico Plants*, vol. 4, no. 2, p. 015, Jul. 2022.
- [36] N. Asada, X. Liu, K. Xu, R. Miyakawa, Y. Yang, H. Santo, Y. Toda, and F. Okura, "LeafGen: Structure-aware leaf image generation for annotation-free leaf instance segmentation," *Plant Phenomics*, vol. 7, no. 4, Dec. 2025, Art. no. 100092.
- [37] A. Muhammad, Z. Salman, K. Lee, and D. Han, "Harnessing the power of diffusion models for plant disease image augmentation," *Frontiers Plant Sci.*, vol. 14, Nov. 2023, Art. no. 1280496.
- [38] A. Van Den Oord, N. Kalchbrenner, and K. Kavukcuoglu, "Pixel recurrent neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1747–1756.
- [39] M. Mehralian and B. Karasfi, "RDCGAN: Unsupervised representation learning with regularized deep convolutional generative adversarial networks," in *Proc. 9th Conf. Artif. Intell. Robot. 2nd Asia-Pacific Int. Symp.*, Dec. 2018, pp. 31–38.
- [40] I. Egusquiza, L. Benito-Del-Valle, A. Picón, A. Bereciartua-Pérez, L. Gómez-Zamanillo, A. Elola, E. Aramendi, R. Espejo, T. Eggers, C. Klukas, and R. Navarra-Mestre, "When synthetic plants get sick: Disease graded image datasets by novel regression-conditional diffusion models," *Comput. Electron. Agricult.*, vol. 229, Feb. 2025, Art. no. 109690.
- [41] M. Bouni, B. Hssina, K. Douzi, and S. Douzi, "Synergistic use of hand-crafted and deep learning features for tomato leaf disease classification," *Sci. Rep.*, vol. 14, no. 1, p. 26822, Nov. 2024.
- [42] X. Yu, Q. Gong, C. Chen, and L. Lu, "Automatic diagnosis of soybean leaf disease by transfer learning," *Amer. J. Biochemistry Biotechnol.*, vol. 18, no. 2, pp. 252–260, Feb. 2022.
- [43] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
- [44] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11531–11539.
- [45] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
- [46] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 9992–10002.
- [47] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.
- [48] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11976–11986.
- [49] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Proceedings 3rd Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–14.
- [50] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [51] D. Gatis. (2021). *Rembg: Remove Image Background*. Accessed: May 8, 2025. [Online]. Available: <https://github.com/danielgatis/rembg>
- [52] Freepik. (2025). *Freepik: Graphic Resources for Everyone*. [Online]. Available: <https://www.freepik.com/>
- [53] C. Chen, H. Keunecke, E. Neu, F. J. Kopisch-Obuch, B. A. McDonald, and J. Stapley, "Molecular epidemiology of cercospora leaf spot on resistant and susceptible sugar beet hybrids," *Plant Pathol.*, vol. 74, no. 1, pp. 69–83, Jan. 2025.
- [54] J. Tao, P. Cao, Y. Xiao, Z. Wang, Z. Huang, J. Jin, Y. Liu, H. Yin, T. Liu, and Z. Zhou, "Distribution of the potential pathogenic alternaria on plant leaves determines foliar fungal communities around the disease spot," *Environ. Res.*, vol. 200, Sep. 2021, Art. no. 111715.
- [55] X. B. Yang, "Patterns of rhizoctonia foliar blight on soybean and effect of aggregation on disease development," *Phytopathology*, vol. 81, no. 3, pp. 287–293, 1991.
- [56] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
- [57] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012, pp. 1097–1105.
- [58] J. G. A. Barbedo, L. V. Koenigkan, and T. T. Santos, "Identifying multiple plant diseases using digital image processing," *Biosyst. Eng.*, vol. 147, pp. 104–116, Jul. 2016.



NABIN PAKKA (Member, IEEE) received the M.S. degree in computer science from the School of Computing and Informatics, University of Louisiana at Lafayette, where he is currently pursuing the Ph.D. degree in computer science. His research interests include machine learning, computer vision, natural language processing, knowledge extraction, and distributed computing.



TERESA WILKERSON received the Ph.D. degree from the Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University, in 2021. She is currently an Assistant Research Professor with the Department of Agricultural Science and Plant Protection and the Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University. Her research interests include plant-parasitic nematodes, crop root and stem diseases, seed treatments, and grain quality concerns.



LI CHEN (Senior Member, IEEE) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Toronto, in 2018. She is currently an Associate Professor with the School of Computing and Informatics, University of Louisiana at Lafayette, USA. Her research interests include big data analytics systems, machine learning systems, cloud computing, datacenter networking, and resource allocation.



NIAN-FENG TZENG (Life Fellow, IEEE) received the Ph.D. degree in computer science from the University of Illinois at Urbana-Champaign, in 1986. He has been with the Center for Advanced Computer Studies, School of Computing and Informatics, University of Louisiana at Lafayette, since 1987. His current research interests include high-performance computer systems, computer networks, and parallel and distributed processing.

...