IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 12 March 2024, accepted 31 March 2024, date of publication 6 May 2024, date of current version 24 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3397570

RESEARCH ARTICLE

Improving Wheat Leaf Disease Classification: Evaluating Augmentation Strategies and CNN-Based Models With Limited Dataset

SYED TAHA YEASIN RAMADAN[1], TANJIM SAKIB[1], FAHMID AL FARID[2], (Member, IEEE), MD. SHOFIQUL ISLAM[1,3], JUNAIDI BIN ABDULLAH[4], (Member, IEEE), MD ROMAN BHUIYAN[5], SARINA MANSOR[2], AND HEZERUL BIN ABDUL KARIM[2], (Senior Member, IEEE)

[1]Department of Computer Science and Engineering, Military Institute of Science and Technology (MIST), Dhaka 1216, Bangladesh
[2]Faculty of Engineering, Multimedia University, Cyberjaya 63100, Malaysia
[3]Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Kuantan, Pahang 26300, Malaysia
[4]Faculty of Computing and Informatics, Multimedia University (MMU), Cyberjaya 63100, Malaysia
[5]Faculty of Computer Science and Informatics, Berlin School of Business and Innovation (BSBI), WV1 1LZ Wolverhampton, U.K.

Corresponding author: Hezerul Bin Abdul Karim (hezerul@mmu.edu.my)

This work was supported in part by Multimedia University, Malaysia, under Grant MMUI/230023.

ABSTRACT Global food security is seriously threatened by wheat leaf disease, which makes effective and precise disease detection and classification techniques necessary. For efficient disease control and the best possible crop health, timely identification and precise classification are essential. However, the limited availability of datasets for wheat leaf diseases hinders the development of effective and robust classification models. This research emphasizes the importance of precise wheat leaf disease diagnosis for global food security. The existing methods face challenges with limited data and computational demands. The research explores the potential of deep learning for automated disease detection, considering these challenges. CycleGAN proved to be the most effective among various augmentation techniques, enhancing the performance of classifiers DenseNet121, ResNet50V2, DenseNet169, Xception, ResNet152V2, and MobileNetV2. ADASYN also significantly improved classification accuracy, with MobileNetV2 consistently outperforming across different augmentation methods. This technique excels in overcoming challenges posed by limited datasets and class imbalances. Using CycleGAN for data augmentation notably enhanced classifier performance, addressing the scarcity of real-world samples. Evaluation through confusion matrix analysis revealed a minimal number of misclassified images—possibly as low as 0 to 3 images over the test dataset. The exceptional 100% accuracy achieved by the MobileNetV2 model on both CycleGAN and ADASYN augmented datasets highlights the potential of these techniques to unlock new levels of accuracy in wheat disease classification. This augmentation technique fine-tuned the classifier, reducing errors and highlighting the crucial role of CycleGAN in enhancing the accuracy and precision of wheat disease classification models. The proposed method establishes CycleGAN's effectiveness in augmenting wheat leaf disease classification and recognizes ADASYN's potential. The developed technique shows promise for automated disease detection in agriculture, enhancing global food security. Future research may optimize computational efficiency and explore integrating emerging technologies such as edge computing.

INDEX TERMS Wheat leaf disease, deep learning, augmentation techniques, CycleGAN, SMOTE, SMOTETomek, ADASYN, disease diagnosis, accuracy, agricultural automation.

I. INTRODUCTION

Wheat stands as a crucial staple crop globally, wielding substantial economic significance. Its paramount role in

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

ensuring global food security and economic stability cannot be overstated. Serving as a dietary cornerstone for a vast majority of the global population, especially in Europe, North America, and Asia, wheat cultivation and trade emerge as pivotal factors propelling the agricultural sectors and overall economies of numerous nations. The production and commerce of wheat wield a profound impact on both national and global economic landscapes. Beyond being a key revenue source for farmers, it acts as a catalyst for job creation in rural areas. The international trade of wheat, encompassing both exports and imports, contributes significantly to global trade dynamics, fostering economic growth and providing foreign exchange benefits for exporting nations. As per the Food and Agriculture Organization (FAO), 2020 witnessed a record-breaking wheat output of 773 million metric tons, with China, India, Russia, and the United States taking the lead in this agricultural milestone [1].

Nevertheless, various diseases have exerted a substantial toll on wheat production, leading to significant global reductions in yield. Pathogens, including fungi, bacteria, and viruses, pose a persistent threat to wheat harvests, contributing to diminished production and economic setbacks for farmers and the agricultural sector. In a study by Savary et al. that delved into the impact of wheat diseases on global productivity, it was revealed that ailments like wheat rusts, powdery mildew, and Fusarium head blight can precipitate yield losses ranging from 10% to 50% across different regions [2]. These losses not only affect the quantity of wheat produced but also impinge on the quality and marketability of the harvested grain.

The manual identification of diseases in wheat fields is a laborious and time-intensive task. As technology progresses, machine learning approaches have emerged as valuable tools for automating disease diagnosis in agriculture. Deep learning models, particularly convolutional neural networks (CNNs), have shown outstanding performance in image-based classification and detection [3], [4], [5]. Leveraging algorithms and models, these technologies analyze extensive databases of plant images, enabling swift and reliable disease identification [6], [7], [8]. CNNs, for instance, exhibit promise in detecting wheat diseases by being trained on labeled datasets containing images of both healthy and diseased wheat plants, allowing them to recognize patterns distinguishing between the two [9].

Despite the potential benefits, a primary challenge in implementing machine learning algorithms for wheat disease detection is the scarcity of labeled datasets. The collection and annotation of large-scale datasets for training such models can be both time-consuming and costly, especially considering the diverse kinds and variants of wheat diseases. However, addressing this limitation, the application of data augmentation techniques, including Generative Adversarial Networks (GANs), presents an effective solution. GANs, a type of deep learning model, generate synthetic data samples by learning the underlying patterns and distributions of the training dataset [8]. By employing GANs for data augmentation in wheat disease detection, researchers can create a more extensive and diverse dataset, mitigating the challenge of limited data [10].

To enhance the dataset and improve accuracy, various augmentation techniques such as CycleGAN, SMOTE, SMOTETomek, and ADASYN were employed in this study. The results from the augmented dataset were compared with those from the original dataset, revealing that the enhanced dataset yielded superior results, particularly when CycleGAN was applied. This research presents a novel contribution through a comparative analysis of classification results for wheat leaf disease detection obtained from GAN-augmented datasets and the original datasets.

1) This research introduces a groundbreaking application of Generative Adversarial Networks (GAN) in the realm of wheat leaf disease identification, a domain that has not been explored previously. This novel approach explores the potential of GANs to generate synthetic images depicting wheat leaf diseases.
2) The study tackles the formidable issue of extremely restricted dataset availability in the identification of wheat diseases. Specifically, there are 102 images in the "healthy" class, 208 images in the "stripe rust" class, and 97 images in the "septoria" class.
3) The research extensively evaluates the effectiveness of augmenting smaller datasets using Generative Adversarial Networks (GAN) to enhance both the size and quality of training data. The study suggests employing data augmentation techniques such as CycleGAN, SMOTE, SMOTETomek, and ADASYN. The evaluation highlights the utility and effectiveness of samples generated by CycleGAN in improving model performance on limited datasets.

In Figure 1, the visual representation offers insight into the entire workflow employed in the classification of wheat leaf diseases, providing succinct yet comprehensive guidance.

II. LITERATURE REVIEW

Jiang et al. ([11]) undertake a comparative analysis of seven distinct CNN architectures, namely VGG-16, Inception-v3, ResNet-50, DenseNet-121, EfficientNet-B6, ShuffleNet-v2, and MobileNetV3. The objective is to identify major wheat leaf diseases from field images. Employing a combination of transfer learning and retuning strategies, the study achieves peak accuracy, with Inception-v3 reaching an impressive 92.5% accuracy on the test dataset. Notably, lightweight CNN models such as ShuffleNet-v2 and MobileNetV3 exhibit faster processing times and lower memory requirements but trade off accuracy, delivering around 87%. In another study ([12]), 450 wheat images are collected and subjected to normalization preprocessing to enhance the training and testing accuracy of a CNN. The pre-trained model attains 89.9% accuracy in classifying powdery wheat disease and 86.5% when applied to the CGIAR dataset using transfer learning. Fang et al. ([13]) introduce a lightweight multiscale

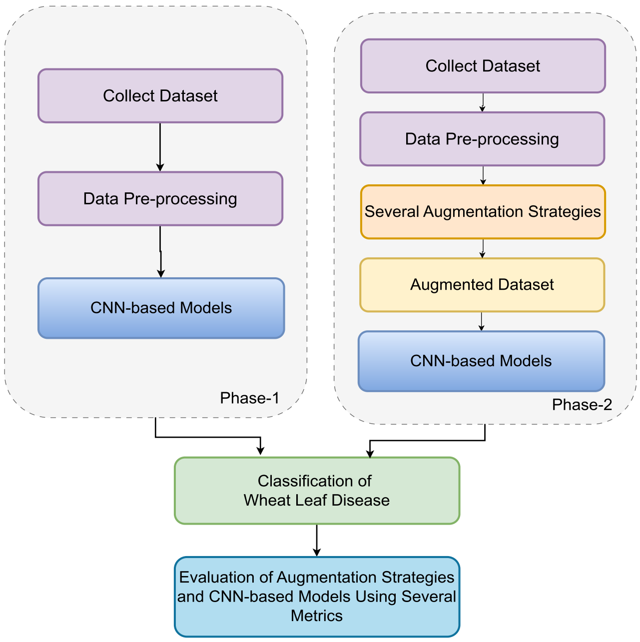

FIGURE 1. Abstract view of wheat leaf disease classification methodology.

CNN model for wheat disease detection, addressing the challenges posed by diverse disease types and complex field backgrounds. Integrating CBAM and ECA modules enhances disease identification and mitigates the impact of intricate backdrops. This proposed technique achieves an outstanding accuracy rate of 98.7% on the test dataset, surpassing both standard CNN models and lightweight models. Its minimal parameter size and low processing demands make it particularly suitable for mobile devices. Shrestha et al. ([14]) propose a CNN-based approach for plant disease identification, incorporating image processing techniques to analyze sample images of damaged and healthy plant leaves. The suggested method attains a test accuracy of 88.80%, providing insights into performance through various performance matrices. Lu et al. ([15]) present an in-field automatic wheat disease diagnosis system based on deep multiple instance learning, a weakly supervised deep learning framework. Leveraging alternative structures, the system achieves high recognition accuracies of 97.95% and 95.12% on the Wheat Disease Database 2017 (WDD2017) through 5-fold cross-validation.

Hossen et al. ([16]) introduce a two-dimensional CNN model designed for the detection and classification of wheat diseases. Demonstrating a high reliability of 98.84% accuracy on a dataset comprising 4800 images, the proposed model effectively distinguishes between infected and healthy wheat fields. To enhance pre-processing and disease detection accuracy, the study employs techniques such as segmentation, resizing, and feature extraction. Jin et al. ([17]) contribute a deep neural network classification system for distinguishing healthy and Fusarium head blight-infected wheat heads in hyperspectral images. Employing a hybrid neural network with convolutional and bidirectional recurrent layers, the model restructures pixel spectra data into a two-dimensional structure. Results indicate superior performance compared to existing models, with an F1 score of 0.75 and an accuracy of 0.743 on the testing dataset. Hussain et al. ([18]) proposes a convolutional neural network (CNN)-based system for identifying wheat crop diseases, achieving an accuracy of 84.54%. The dataset used for training includes representations of four types of wheat diseases, demonstrating the potential utility of the model for farmers in safeguarding their wheat harvests. Bao et al. ([19]) introduce SimpleNet, a portable CNN model for automatically detecting wheat ear disorders in field-captured images. SimpleNet incorporates a Convolutional Block Attention Module (CBAM) and a feature fusion module to enhance representation capabilities and prevent loss of detailed information. Experimental findings reveal an identification accuracy of 94.1%. Alharbi et al. ([20]) described a novel approach to disease classification in wheat crops that uses a few-shot learning mechanism. The study uses EfficientNet as a foundation and includes an attention technique to assist in efficient feature selection. The suggested network produces impressive outcomes, with an accuracy of 93.19% on a dataset of 18 wheat diseases manually gathered from the internet. Savary's et al. ([21]) use scenario analysis to investigate the effects of global changes on wheat health until 2050. The research proposes a scenario analysis methodology for three hypothetical agrosystems, taking into account elements such as climate change, pesticide use, and host plant resistance. The findings show an overall increase in risk probabilities and magnitudes across agrosystems, emphasizing the importance of scenario analysis in analyzing crop health development.

A. REVIEW OF GAN-BASED DATA AUGMENTATION STUDIES

In the research conducted by Kumar and Kukreja ([22]), a hybrid strategy for wheat yellow rust disease classification is presented. This approach combines a generative adversarial network (GAN), specifically the State of the Art GAN (STARGAN), for data augmentation, and a convolutional neural network (CNN) for classification. At the medium severity level, the proposed method achieves a classification accuracy of 95.6%. Usha Ruby et al. ([23]) used a deep learning-based system to classify wheat leaf diseases, with picture imputation performed by a collaborative generative adversarial network. The paper suggests modifying the ResNet50 architecture to include Conv, Batch Normaliz, and Activation Leaky Relu layers to increase feature extraction and discrimination. Extensive testing reveals that the proposed model outperforms ResNet50, InceptionV3, and DenseNet, with an excellent identification accuracy of 98.44%. Ramadan et al. ([24]) utilized CycleGAN for data augmentation in mango leaf disease classification, achieving 100% accuracy using the models ViT-B/16, ViT-B/32, and BiT-M-R50 × 1. In the realm of tomato disease detection, Abbas et al. ([25]) leverage transfer learning with synthetic images generated by a Conditional Generative Adversarial Network (C-GAN). The DenseNet121-based model achieves impressive accuracy levels of 99.51%, 98.65%, and 97.11% for classification tasks involving 5, 7, and 10 classes, respectively. Nazki et al. ([26]) address imbalances in small plant disease datasets using a data-level synthetic sampling approach based on Generative Adversarial Networks (GANs). The method enhances learning by resolving concerns with data distribution and class imbalance, resulting in more accurate classification results. Vasudevan and Karthick ([27]) present a hybrid method for plant disease detection using E-GAN and CapsNet. The three-stage approach includes leaf area extraction, dataset expansion using E-GAN, and disease diagnosis and stage identification using CapsNet, achieving an accuracy of 97.63% on real-time grape leaf images and the PlantVillage dataset.

Despite these advancements in the broader field of plant disease detection, there is a noticeable research gap specifically in the area of wheat disease detection. The presented paper aims to fill this gap by exploring the potential of augmentation methods such as SMOTE, SMOTETomek, and ADASYN in combination with GAN for more accurate wheat disease detection. The study emphasizes the need for further investigation to determine the most effective combination of augmentation techniques and classifiers, particularly for datasets with limited samples, such as septoria and stripe rust.

B. RESEARCH GAP

1) Despite the increasing popularity of Generative Adversarial Networks (GANs) across various computer vision applications, a substantial research void persists in the application of GANs, specifically CycleGAN, for the identification of wheat leaf diseases. Previous studies have predominantly focused on traditional deep learning techniques, overlooking the untapped potential advantages that GANs could offer in this specific domain.
2) The research gap encompasses a lack of comprehensive examinations regarding GAN-based data augmentation strategies aimed at tackling challenges posed by smaller datasets in the classification of wheat leaf diseases.
3) Another research gap involves the limited evaluation of the quality of GAN-generated images and their impact on model performance. To ensure the trustworthiness and generalizability of disease classification models, it is essential to understand the authenticity and realism of the synthetic images generated by GANs.

This research aims to assess various augmentation methods to identify an effective strategy for augmenting small datasets in wheat disease classification. The objective is to optimize CNN-based models, enhancing accuracy in the identification of wheat leaf diseases within limited data. The investigation will focus on CycleGAN and other augmentation approaches, to pinpoint the most effective augmentation technique and refine the CNN architecture to adeptly address the challenges of classification in smaller datasets.

III. METHODOLOGY

The methodology of the study unfolds in two distinct phases. In the initial phase (*Phase-1*), the study commenced by assembling a dataset of wheat leaf diseases, encompassing healthy leaves as well as those affected by septoria and stripe rust [28]. This compiled dataset underwent meticulous preprocessing, including essential steps such as normalization, resizing, and feature extraction, aimed at refining the image data for subsequent analysis. Following this preparatory stage, six advanced CNN-based classifiers were employed to train and validate the dataset. These classifiers underwent extensive fine-tuning to discern minute details and patterns characteristic of stripe rust and septoria, ultimately resulting in an accurate categorization of these wheat leaf diseases (refer to Figure 2).

Transitioning into *Phase-2*, this stage involved data augmentation utilizing four distinct techniques: CycleGAN, ADASYN, SMOTE, and SMOTETomek. The primary objective was to enhance dataset diversity and rectify inherent class imbalances. The augmentation process entailed generating synthetic images, effectively expanding the dataset while preserving its authenticity and relevance to wheat leaf diseases. The same six CNN-based classifiers employed in Phase 1 were applied to the augmented dataset. These classifiers underwent training and evaluation using the augmented dataset to assess their efficacy in classifying septoria and stripe rust. The study's dual-phase approach was strategically designed to enhance dataset quality and diversity, ultimately bolstering the robustness and accuracy of the disease categorization process.

Following the training of models with both augmented and non-augmented datasets, the results were compared to ascertain which techniques exhibited superior performance in predicting wheat-diseased leaf images. This comparison relied on various performance metrics, including accuracy, precision, recall, and F1 score. Algorithm 1 outlines the detailed implementation of the entire process, describing the step-by-step procedure for predicting wheat disease. This encompasses creating a dataset with images of diseased wheat leaves, segmenting the dataset into subsets for training, validation, and testing, utilizing different CNN models for training and forecasting, and evaluating various data augmentation techniques to determine their effectiveness in predicting wheat leaf disease.

A. WHEAT DISEASE CLASSIFICATION WITHOUT DATA AUGMENTATION: UTILIZING THE RAW DATASET

The transition from raw data to pre-processing and the subsequent utilization of machine learning models represent a crucial phase in the data processing pipeline, particularly in tasks like image classification. To derive meaningful insights from raw data, this study employs a diverse

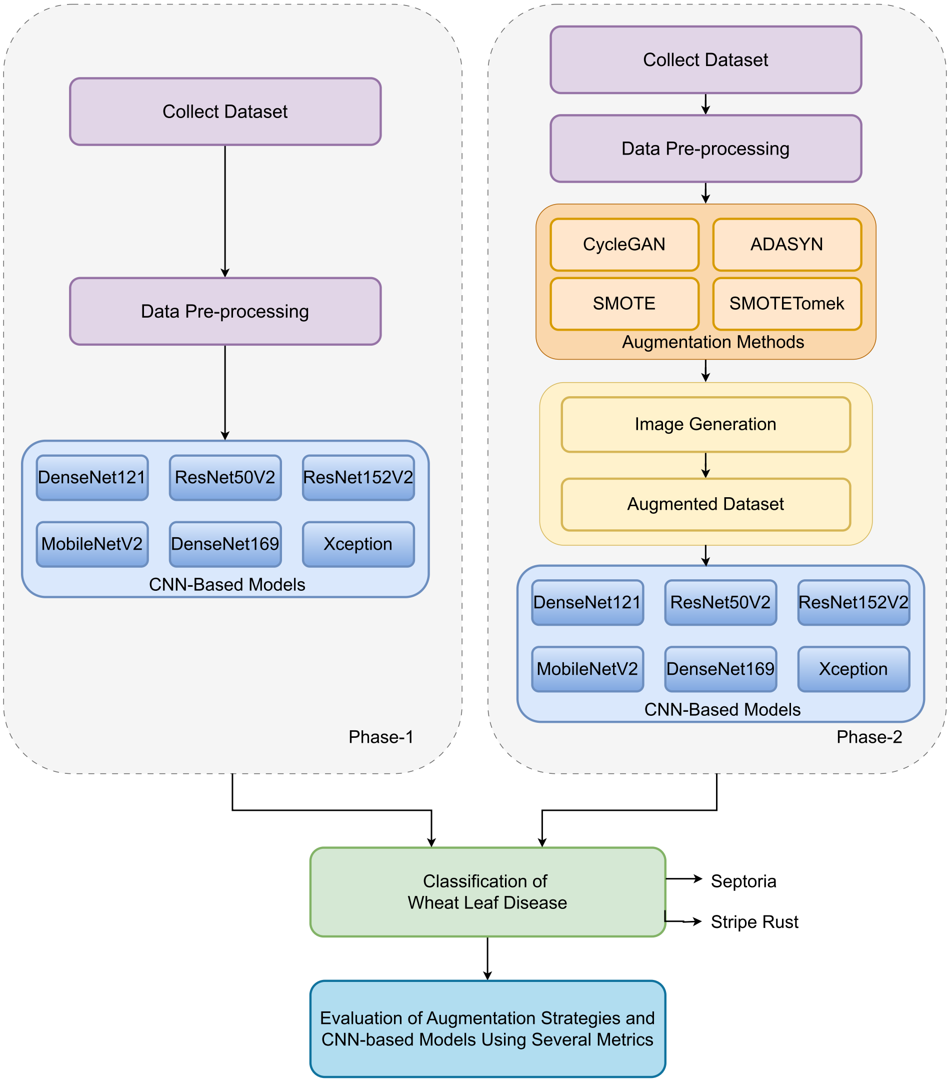

FIGURE 2. **Wheat leaf disease classification methodology: A holistic overview.**

set of machine learning models, including DenseNet121, ResNet50V2, DenseNet169, MobileNetV2, ResNet52V2, and Xception.

The dataset detailing wheat leaf diseases is presented in Table 1, offering an overview of sample distribution across various categories. The first category, labeled "Healthy," comprises wheat leaves devoid of visible diseases, representing the typical disease-free state and serving as a baseline or reference class. This dataset incorporates 102 samples of healthy wheat leaves.

The second category, "Stripe Rust," encompasses instances where wheat leaves exhibit indications of stripe rust, a common fungal disease affecting cereals. This category, crucial for studying pathogenic effects on wheat, possesses a larger sample size of 208 instances. The third category, "Septoria," includes wheat leaves infected with Septoria, another prevalent fungal infection, comprising 97 samples. Figure 3 illustrates a sample from the raw dataset.

The initial phase involves pre-processing the raw data, a procedure that often incorporates tasks such as scaling, normalization, and augmentation. Resizing ensures uniform dimensions of input images, and normalization standardizes pixel values to facilitate model convergence and effective learning.

The pre-processed data is subsequently fed into machine learning models for training and evaluation. Various architectures, each with distinct features, are employed for this purpose. DenseNet121, renowned for its dense connectivity patterns, ResNet50V2, and ResNet52V2, leveraging residual learning structures, DenseNet169 for enhanced feature extraction, MobileNetV2 for efficiency in mobile applications, and Xception, utilizing depthwise separable convolutions, represent examples of these architectures, each offering varying levels of complexity. The detailed implementation is outlined in Algorithm 2.

Algorithm 1 Overall Process for Wheat Disease Classification

```
Prepare the dataset of wheat sick leaf pictures, denoted by D.
Split D into training, validation, and test sets: D_train, D_val, and D_test.
Define the CNN models for training: C = {DenseNet121, ResNet50V2, DenseNet169, MobileNetV2, ResNet52V2, Xception}.
for each CNN model C_i in C do
    Train C_i using D_train.
    Use C_i to forecast wheat-diseased leaf images in D_test.
end for
Define the set of data augmentation methods: A = {CycleGAN, ADASYN, SMOTE, SMOTETomek}.
for each augmentation method A_i in A do
    Apply A_i to the basic dataset D to generate synthetic samples.
    Combine augmented data with D_train to create a new training set: D_train_augmented.
    for each CNN model C_i in C do
        Retrain C_i using D_train_augmented.
        Fine-tune C_i using D_val.
        Test C_i using D_test and evaluate its performance.
    end for
end for
Compare the performance of non-augmented and different augmentation methods based on evaluation results.
```

TABLE 1. Dataset of wheat leaf disease.

Leaf Type	Number of Samples
Healthy	102
Stripe Rust	208
Septoria	97

Stripe Rust Septoria Healthy

FIGURE 3. Sample images of initial dataset.

B. WHEAT LEAF CLASSIFICATION WITH DATA AUGMENTATION

In the second phase, the input dataset undergoes augmentation using four distinct methods: CycleGAN, ADASYN, SMOTE, and SMOTETomek. These augmentation techniques are employed to generate synthetic data samples, effectively augmenting the dataset's size. The expanded dataset is subsequently employed to retrain the same CNN models utilized in *Phase-1*.

Algorithm 2 Training CNN Models for Wheat Disease Prediction

```
Prepare the dataset of wheat unhealthy leaf images, denoted by D_unhealthy.
Preprocess D_unhealthy for training.
Split D_unhealthy into training, validation, and test sets: D_train, D_val, and D_test.
Define the CNN models for training: C = {DenseNet121, ResNet50V2, DenseNet169, MobileNetV2, ResNet52V2, Xception}.
for each CNN model C_i in C do
    Train C_i using D_train with epoch=50, optimizer=SGD, batch size=64, learning rate=0.0001.
    Use C_i to predict wheat-diseased leaf images in D_test.
    Store C_i loss, accuracy, precision, recall, and f1-score.
end for
```

1) DATA AUGMENTATION OF WHEAT LEAF DISEASE USING CycleGAN

CycleGAN stands as a Generative Adversarial Network (GAN) designed for image-to-image translation. Unlike conventional GANs that generate new images from random noise, CycleGAN is specifically trained to learn a mapping between two distinct image domains. Its unique feature lies in utilizing cycle consistency, ensuring the preservation of the original image throughout the translation process.

Cycle consistency is achieved through the incorporation of two generators, G_{AB} and G_{BA}, along with two discriminators, D_A and D_B. The generators take images from domain A and domain B as input, aiming to translate them into domain B and domain A, respectively. Simultaneously, the discriminators work to distinguish between the translated images and real images within their respective domains (refer to Figure 5).

The cycle consistency loss (CCL) quantifies the disparity between the original image and the image obtained by translating it from one domain to the other and then back. Mathematically, this loss is represented in Equation 1.

$$CCL = ||G_{BA}(G_{AB}(A)) - A|| + ||G_{AB}(G_{BA}(B)) - B|| \quad (1)$$

where $G_{BA}(G_{AB}(A))$ represents the image translated from domain A to domain B and back to domain A, and $||.||$ denotes the L1 or L2 norm, which measures the difference between the original and translated images. G_{AB} and G_{BA} are the generator functions of the CycleGAN model. G_{AB} takes an image from domain A as input and generates a corresponding image in domain B, while G_{BA} takes an image from domain B and generates its corresponding image in domain A. These generator functions are trained to minimize the cycle consistency loss (CCL) by ensuring that the images translated

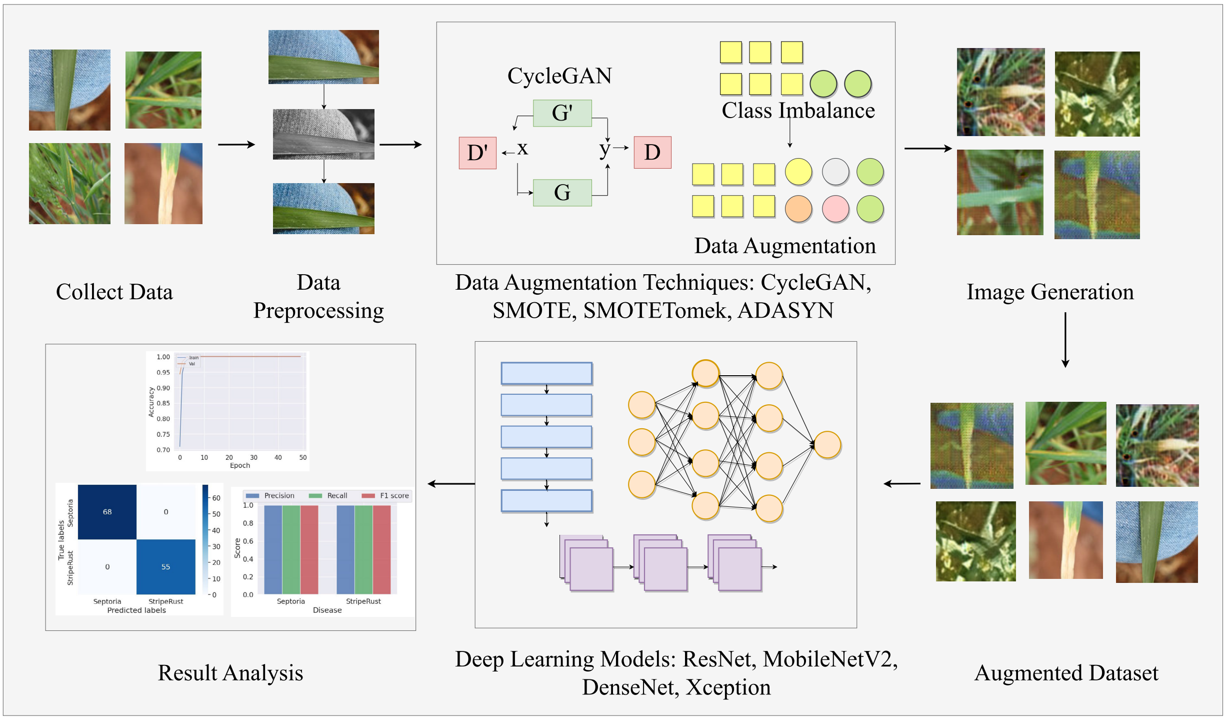

FIGURE 4. Visual demonstration of wheat leaf disease classification using various augmentation methods.

back and forth between the two domains are close to the original images.

CycleGAN possesses the capability to learn a mapping between two domains even in the absence of paired data, achieved by incorporating the cycle consistency loss into the comprehensive GAN loss function. This characteristic renders CycleGAN particularly advantageous for tasks necessitating paired data, such as picture style transfer, object transfiguration, and artistic image production. The complete implementation process of wheat leaf disease classification using CycleGAN is detailed in Algorithm 3.

2) DATA AUGMENTATION OF WHEAT LEAF DISEASE USING ADASYN

ADASYN (Adaptive Synthetic Sampling) is a widely recognized data augmentation strategy designed to address class imbalance in machine learning datasets, particularly effective in scenarios where one class significantly outnumbers the others. ADASYN's main objective is to generate synthetic samples for the minority class, with a focus on adaptively emphasizing instances that pose greater difficulty in correct identification.

The ADASYN approach initiates by determining the quantity of synthetic samples to be generated for each minority class instance. This calculation is influenced by the class distribution, resulting in more samples being created for instances that are in closer proximity to the decision boundary of the classifier.

Algorithm 3 Wheat Leaf Disease Classification With CycleGAN

Prepare the dataset of wheat sick leaf pictures, denoted by D.
Split D into training, validation, and test sets: D_train, D_val, and D_test.
Define the CycleGAN architecture.
Initialize the generator and discriminator networks of CycleGAN: G and D.
Train G and D using D_train with the CycleGAN loss function.
Use the trained G to generate synthetic wheat-diseased leaf images: $G(D\_train)$.
Combine synthetic images with D_train to create a new training set: $D\_train\_cyclegan$.
Define the CNN models for training: C = {DenseNet121, ResNet50V2, DenseNet169, MobileNetV2, ResNet52V2, Xception}.
for each CNN model C_i in C **do**
    Train C_i using $D\_train\_cyclegan$.
    Fine-tune C_i using D_val.
    Test C_i using D_test and evaluate its performance.
end for

The number of synthetic samples for each instance is calculated in Equation 2.

$$N_{\text{gen}} = \text{int}(ratio \times N_{\min}) \tag{2}$$

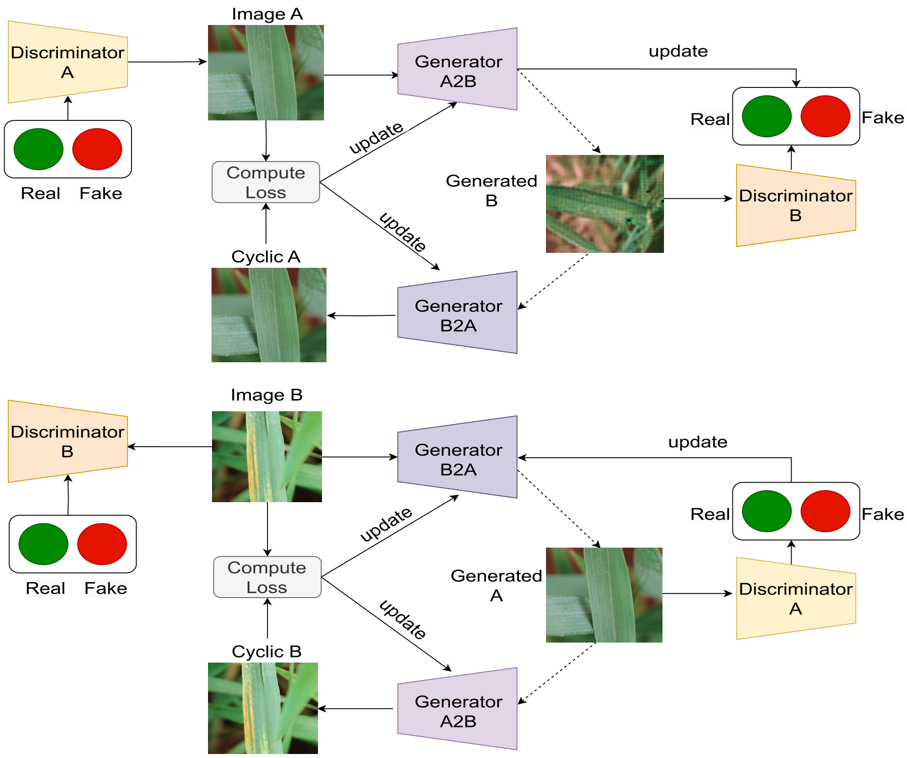

FIGURE 5. Architecture of CycleGAN for wheat leaf.

The ratio (ratio) in Equation 2 determines the proportion of synthetic samples to be generated for each instance. It is calculated as the product of the specified ratio and the minimum number of instances ($N_{\min}$) in the dataset. The resulting value is then rounded to the nearest integer using the int function to determine the number of synthetic samples (N_{gen}) to be generated for each instance.

Following the determination of the required number of synthetic samples, ADASYN proceeds to select N_{gen} nearest neighbors for each instance in the minority class. Synthetic samples are then generated by interpolating the feature values of the instance with those of its nearest neighbors. The resultant dataset is a fusion of the original data and the synthetically generated samples.

3) DATA AUGMENTATION OF WHEAT LEAF DISEASE USING SMOTE

SMOTE (Synthetic Minority Over-sampling Technique) stands out as a prominent technique in machine learning for augmenting imbalanced datasets. In numerous real-world scenarios, the minority class often contains significantly fewer instances than the majority class, leading to biased model predictions. SMOTE addresses this issue by balancing the dataset, and generating synthetic samples for the minority class.

The SMOTE augmentation process involves randomly selecting a sample from the minority class along with its k-nearest neighbors. Synthetic samples are then generated by interpolating the feature values of the chosen sample with those of its neighbors. The number of synthetic samples created is typically determined by a user-defined oversampling ratio.

Mathematically, for a minority class sample x_i and its k-nearest neighbors x_j, the synthetic sample x_{new} is generated using Equation 3.

$$x_{new} = x_i + \lambda \cdot (x_j - x_i) \tag{3}$$

Here, λ represents a randomly chosen value within the range of 0 to 1. This process is iteratively applied to multiple minority class samples, culminating in a balanced dataset that enhances the effectiveness of training machine learning models.

4) DATA AUGMENTATION OF WHEAT LEAF DISEASE USING SMOTETomek

SMOTETomek is an amalgamation of machine learning and unbalanced classification techniques designed to address imbalanced datasets. In tackling this issue, it integrates the Synthetic Minority Over-sampling Technique (SMOTE) with the Tomek connections approach.

In imbalanced datasets, the majority class significantly outnumbers the minority class, leading to biased model predictions. SMOTE addresses this imbalance by generating synthetic samples for the minority class. This involves randomly selecting a sample from the minority class and its k-nearest neighbors, followed by creating additional samples along the connecting line segments. However, SMOTE may introduce noisy or irrelevant synthetic samples, potentially impacting the model's performance. This is where Tomek linkages play a role. Tomek connections identify nearest neighbors of two samples—one from the majority class and one from the minority class. These connections are treated as ''noise'' and are removed to enhance the separation between the two classes.

The SMOTETomek method combines both SMOTE and Tomek linking strategies to construct a balanced and clearer dataset. It begins by oversampling the minority class, followed by using Tomek linkages to eliminate noisy and borderline samples from both classes. This approach enhances class separation, reducing the risk of overfitting.

Mathematically, the SMOTE-Tomek augmentation can be represented as follows:

1) Let X be the input dataset with n samples and d features, and y be the corresponding labels. Let X_{minority} be the subset of X that contains the minority class samples, and X_{majority} be the subset of X that contains the majority class samples.
2) Apply Tomek links to identify pairs of samples (x_i, x_j) where x_i belongs to X_{minority} and x_j belongs to X_{majority}, and their Euclidean distance is the minimum between all pairs of samples in X_{minority} and X_{majority}.
3) Remove the samples x_i from X_{minority} and x_j from X_{majority} that form Tomek links, resulting in new subsets $X_{\text{minority\_new}}$ and $X_{\text{majority\_new}}$, using Equations (4–5)

$$X_{\text{minority\_new}} = X_{\text{minority}} \setminus \{x_i\} \tag{4}$$

$$X_{\text{majority\_new}} = X_{\text{majority}} \setminus \{x_j\} \tag{5}$$

4) Apply SMOTE to $X_{\text{minority\_new}}$ to generate synthetic samples for the minority class, increasing its size to match the majority class.

The augmentation technique of SMOTE-Tomek thus generates a balanced dataset suitable for training machine learning models, enhancing their performance on imbalanced datasets and effectively addressing the challenge of class imbalance.

5) COMPUTATIONAL COMPLEXITY ANALYSIS

We used a GPU provided by Google Colab to conduct research. This made it possible to train and evaluate CNN-Based models more efficiently, as well as to use data augmentation techniques. The following analysis explores the complexities of training CNN-Based models, data augmentation methods like CycleGAN, ADASYN, SMOTE, and SMOTETomek, as well as retraining and fine-tuning processes.

a. *Training CNN-Based Models:* For each CNN-Based model C_i in the set C, the computational complexity primarily depends on the architecture of the model and the size of the training dataset D_{train}. Let N be the number of training samples, M be the number of training iterations, and $O(f)$ denote the complexity of a forward pass through the network. The overall complexity of training all CNN-Based models in C is $O(|C| \times N \times M \times O(f))$.

b. *Data Augmentation:* The complexity of data augmentation methods depends on their specific implementations. Training for CycleGAN entails simultaneously optimizing two discriminator networks and two generator networks. Given the number of training iterations for stripe rust and septoria, denoted as M_{stripe} and M_{septoria} respectively, the overall complexity is $O(M_{\text{stripe}} + M_{\text{septoria}})$. Similarly, the complexities of other augmentation methods like ADASYN, SMOTE, and SMOTE-Tomek depend on their respective algorithms and parameters.

c. *Training and Fine-Tuning CNN-Based Models:* After data augmentation, each CNN model C_i is retrained using the augmented training set $D\_\text{train\_augmented}$. The computational complexity is similar to the initial training step. Fine-tuning involves additional passes through the validation dataset D_val, contributing to the overall complexity.

d. *Overall Complexity:*

· The overall computational complexity of the algorithm is dominated by the training and testing steps of the CNN-Based models, as well as the data augmentation process. Let N_aug be the size of the augmented training set after applying augmentation methods.

· The total complexity can be approximated as $O(|C| \times N \times M \times O(f) + |A| \times N\_\text{aug} \times M\_\text{aug} \times O(g))$, where M_aug represents the maximum number of iterations for data augmentation method, $|A|$ represents the total number of data augmentation methods used in the algorithm, and $O(g)$ denotes the complexity of applying the augmentation method.

IV. RESULTS ANALYSIS

Based on the experimental analysis, Table 2 provides an in-depth examination of the effectiveness of various CNN models in the prediction of wheat leaf diseases, utilizing a range of data augmentation strategies. Each row in the table corresponds to a distinct data augmentation approach, and the columns provide comprehensive information on the CNN model employed, accuracy percentages, and the corresponding loss values.

In the context of the CycleGAN augmentation dataset, the DenseNet169 and MobileNetV2 models stand out prominently, achieving remarkable accuracy rates of 99.19% and 100%, respectively. This highlights their robust performance in handling the specific characteristics introduced by the CycleGAN augmentation. Additionally, the ResNet50V2 and Xception models demonstrate comparable accuracy levels, further emphasizing the effectiveness of these models in the given context.

TABLE 2. The result obtained by various augmentation strategies and CNN-based models.

Augmentation Methods	CNN-Based Model	Accuracy(%)	Loss
CycleGAN	DenseNet121	**98.37**	0.0318
	ResNet50V2	97.56	0.0849
	DenseNet169	**99.19**	0.0563
	Xception	**97.56**	0.0967
	MobileNetV2	**100**	0.0025
	ResNet152V2	**99.19**	0.0145
ADASYN	DenseNet121	96.47	0.0579
	ResNet50V2	**97.65**	0.1286
	DenseNet169	98.82	0.0155
	Xception	94.12	0.2132
	MobileNetV2	**100**	0.0033
	ResNet152V2	97.65	0.1245
SMOTE	DenseNet121	72.13	0.5048
	ResNet50V2	80.33	0.3598
	DenseNet169	78.69	0.4525
	Xception	67.21	0.6091
	MobileNetV2	62.3	0.6664
	ResNet152V2	85.25	0.3903
SMOTETomek	DenseNet121	84.52	0.4074
	ResNet50V2	91.67	0.2008
	DenseNet169	83.33	0.4351
	Xception	64.29	0.6533
	MobileNetV2	78.57	0.6471
	ResNet152V2	84.52	0.3772
No Augmentation	DenseNet121	96.72	0.1164
	ResNet50V2	96.72	0.3332
	DenseNet169	98.36	0.0431
	Xception	95.08	0.2178
	MobileNetV2	98.36	0.0771
	ResNet152V2	95.08	0.1150

Shifting focus to the ADASYN augmentation strategy, MobileNetV2 once again distinguishes itself by achieving a perfect accuracy score of 100%, reinforcing its efficacy in addressing the challenges posed by imbalanced datasets. Furthermore, both DenseNet169 and ResNet50V2 showcase high accuracy, exceeding 98%, indicating their proficiency in handling the augmented dataset generated by the ADASYN technique. These results collectively underscore the diverse strengths of CNN models and their adaptability to different data augmentation strategies in the realm of wheat leaf disease prediction.

In the case of SMOTE augmentation, the models display varying accuracy levels, with ResNet152V2 leading the pack at 85.25%. Notably, MobileNetV2 exhibits a decrease in accuracy to 62.3% with this augmentation technique, indicating a certain sensitivity to this particular method.

Moving on to SMOTETomek augmentation, the models generally exhibit improved performance compared to SMOTE alone. However, Xception experiences a reduction in accuracy, highlighting the nuanced impact of augmentation approaches on different models. Even without augmentation, DenseNet169 and MobileNetV2 maintain strong performance, both achieving accuracy levels exceeding 98%. This underscores the importance of selecting an appropriate augmentation approach, as models respond differently to various techniques. The chart illustrates the significance of this choice in influencing the behavior of models under diverse augmentation strategies.

Figures 6 a) and b) depict the loss of generators A and B, respectively, caused by the Septoria disease. The x-axis shows the epoch, while the y-axis represents the loss. The loss appears to be decreasing over time for both generators. Graphs c) and d) demonstrate the loss of discriminators A and B, respectively. The x and y axes are interpreted the same way as in graphs a) and b). The loss tends to reduce over time for both discriminators. Figure 7 shows the same trend with Stripe Rust.

In summary, CycleGAN augmentation emerges as highly effective in the classification of wheat leaf images, showcasing notable performance enhancements. While ADASYN also demonstrates merit in this task, it does not quite reach the exceptional results achieved by CycleGAN augmentation, firmly establishing CycleGAN as the most effective method for wheat leaf disease classification (refer to Figure 8). Conversely, other augmentation techniques, namely SMOTE and SMOTETomek, prove less effective in forecasting wheat leaf diseases.

A. SYNTHETIC IMAGE GENERATION USING DIFFERENT AUGMENTATION METHODS

Table 3 provides an overview of dataset augmentation for diseases such as Septoria and Stripe Rust using various

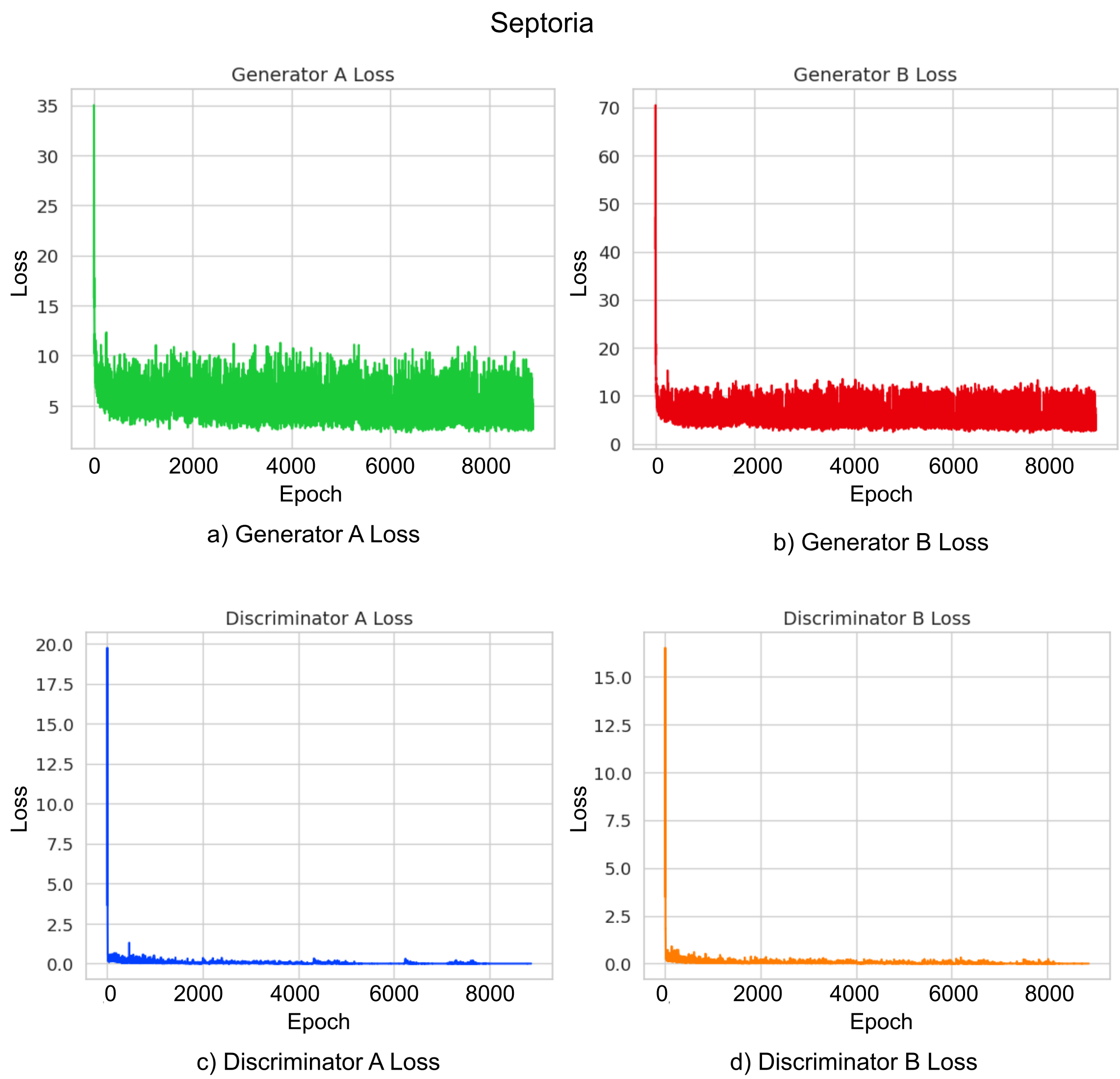

FIGURE 6. CycleGAN losses of generator and discriminator for Septoria.

augmentation techniques. The initial number of instances remains consistent across all methods, serving as a baseline for each disease type. However, CycleGAN augmentation notably amplifies the overall dataset size, leading to an increase in the number of samples for both disease types—Stripe Rust grows from 208 to 310 samples, and Septoria from 97 to 301 samples.

In contrast, the ADASYN technique focuses on achieving a balanced class distribution by significantly increasing the sample sizes of the minority class. This brings Septoria's count to 215, while Stripe Rust's samples remain the same. Similarly, both SMOTE and SMOTETomek techniques augment the total number of samples to 208 for the minority class, aligning it with the majority class and maintaining dataset balance.

These augmentation techniques not only aim to expand dataset sizes but also address class imbalances. This dual objective may contribute to enhancing the robustness and performance of machine learning models developed using these augmented datasets.

The visually represented synthetic samples obtained through various augmentation strategies are depicted in Figure 9. CycleGAN, in particular, showcases a distinctive pattern and enriches the dataset for both Septoria and Stripe Rust diseases by generating samples for both classes.

In contrast, ADASYN, SMOTE, and SMOTETomek exhibit a clear focus on augmenting the minority class, Septoria leaf disease, while maintaining the same number of samples for the majority class, Stripe Rust. These three approaches aim to rectify class imbalances by supplementing Septoria samples, addressing concerns about dataset imbalance without altering the existing count of Stripe Rust samples. Notably, they achieve this by significantly boosting the number of examples for the minority class.

By specifically addressing the needs of the minority class, these augmentation techniques effectively mitigate class

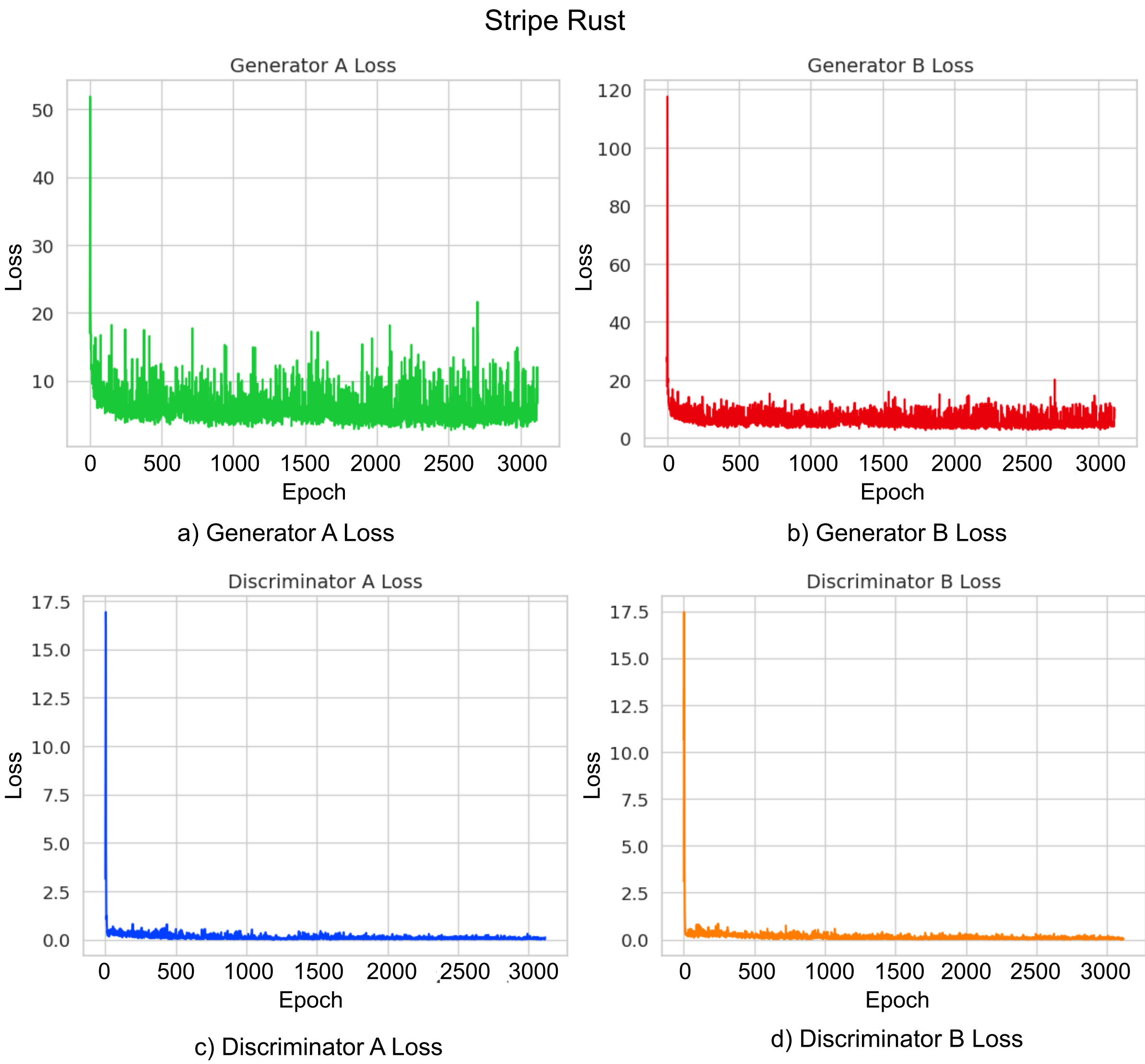

FIGURE 7. CycleGAN losses of generator and discriminator for stripe rust.

imbalance issues. This targeted approach has the potential to result in the development of a more robust and balanced dataset, thereby enhancing the performance and training of machine learning models.

B. CLASSIFIER PERFORMANCE ANALYSIS USING CycleGAN AUGMENTED DATASET

The application of CycleGAN augmentation proved highly effective in generating a diverse set of synthetic images, significantly enriching the wheat leaf disease dataset for classification purposes. The incorporation of this augmented dataset resulted in noticeable improvements in both diversity and balance. Examining the classifiers' performance through confusion matrices revealed distinct trends. Notably, the MobileNetV2 model demonstrated accuracy in correctly categorizing each class of wheat leaf disease. Conversely, the Xception model accurately identified every image of Stripe Rust but misclassified five septoria leaf images. Other models, however, only had one or two misclassifications per wheat leaf disease class. Furthermore, the models successfully identified every Stripe Rust image, underscoring their robustness when employed with the CycleGAN-augmented dataset (refer to Figure 10). In addition, precision, recall, and F1-scores, employed as evaluation metrics, provided further insight into the accurate identification of Septoria and Stripe Rust diseases (refer to Figure 11). These measures proved valuable in assessing the classifiers' performance, revealing their relative strengths and weaknesses in disease classification. The accuracy graph (Figure 12) showcased stabilized accuracy after 10 epochs for classifiers such as ResNet50V2, MobileNetV2, DenseNet169, Xception, and ResNet152V2, indicating consistent convergence in predicting the target classes. On the other hand, DenseNet121 exhibited varying accuracy, suggesting potential challenges in achieving

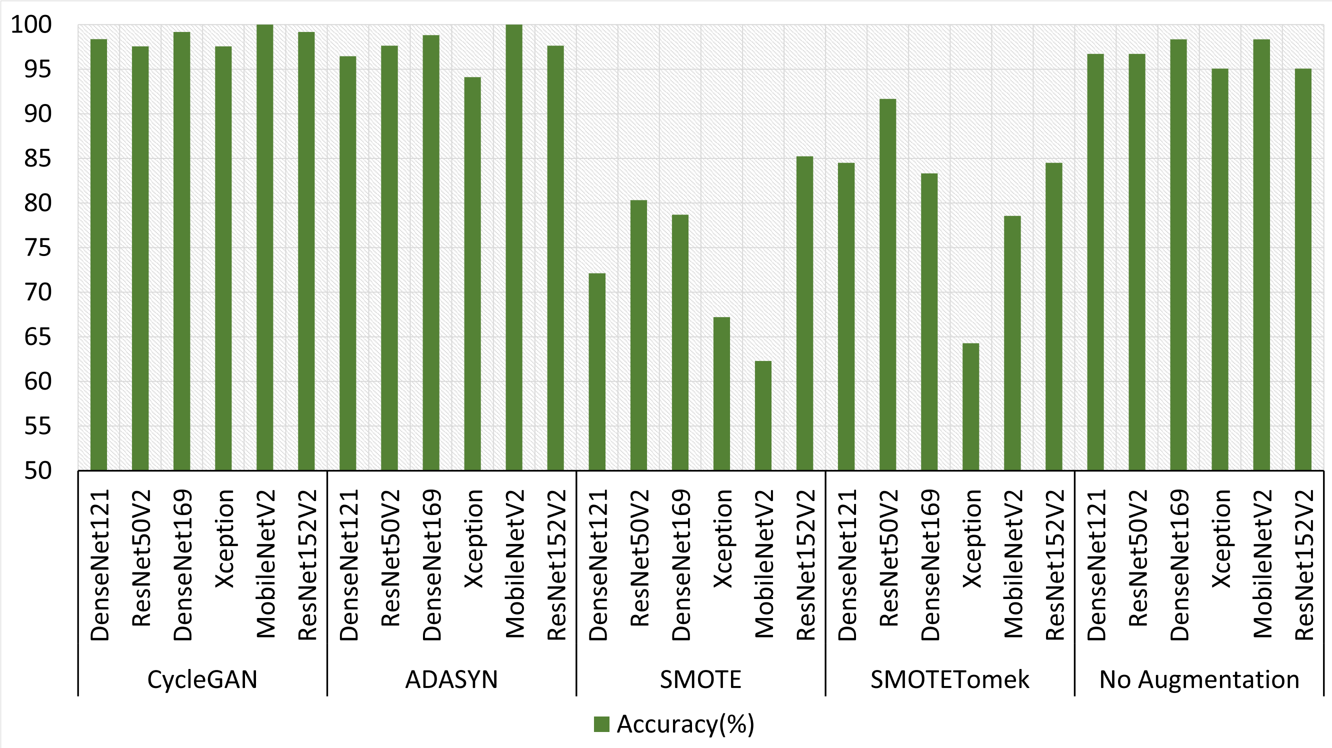

FIGURE 8. **Comparitive analysis of different data augmentation methods and no augmentation.**

convergence comparable to other models. A similar pattern was observed in the loss graph (Figure 13), with ResNet50V2, MobileNetV2, DenseNet169, Xception, and ResNet152V2 showing minimal variation in loss measures after 10 epochs, signifying consistent learning and convergence. In contrast, DenseNet121 displayed irregular loss values.

C. CLASSIFIER PERFORMANCE ANALYSIS USING ADASYN AUGMENTED DATASET

The dataset augmented with ADASYN demonstrated outstanding performance, notably achieving 100% accuracy with the MobileNetV2 model. In comparison with other augmented or non-augmented datasets, the ResNet50V2 model also exhibited impressive performance. However, some other models did not perform as well and did not reach the same level of accuracy achieved by the CycleGAN-augmented dataset.

Furthermore, the confusion matrices of various classifiers utilizing the ADASYN-augmented datasets depicted accurate classifications by the MobileNetV2 model, while some images were misclassified by other models (refer to Figure 14). Additionally, the evaluation metrics—precision, recall, and F1-score—offered insights into the effectiveness of the ADASYN-augmented datasets (see Figure 15).

The accuracy of ResNet50V2, ResNet152V2, and MobileNetV2 stabilized after 5 epochs and continued to perform consistently thereafter. In contrast, the accuracy scores of DenseNet121 and DenseNet169 stabilized after 20 epochs. Unlike the other models, Xception displayed notable oscillations in accuracy throughout the training process, indicating ongoing variability and potentially facing challenges with learning convergence (refer to Figure 16). In the loss graph (Figure 17), after 10 epochs, the majority of the models exhibited stabilized loss patterns, signifying learning convergence. However, Xception stood out due to the constant variations in its loss curve.

Even with the limited sample set available for the classification of wheat leaf disease, our investigation showed that CNN-based models did not exhibit overfitting. This finding underscores the robustness and generalization capacity of our approach, despite data limitations. The steady performance observed indicates that augmentation techniques, particularly CycleGAN and ADASYN were essential in improving model performance. These augmentation strategies substantially reduced the danger of overfitting and aided in the learning of meaningful representations of wheat leaf disease patterns by varying the training data through the creation of synthetic samples and adaptive sampling procedures.

D. COMPARATIVE ANALYSIS WITH THE STATE-OF-THE-ART METHOD

Table 4 provides a comprehensive overview of the effectiveness of various classification models employed in wheat disease identification, encompassing both our study's models and those detailed in other research publications. Each entry in the table specifies the dataset used, the classification

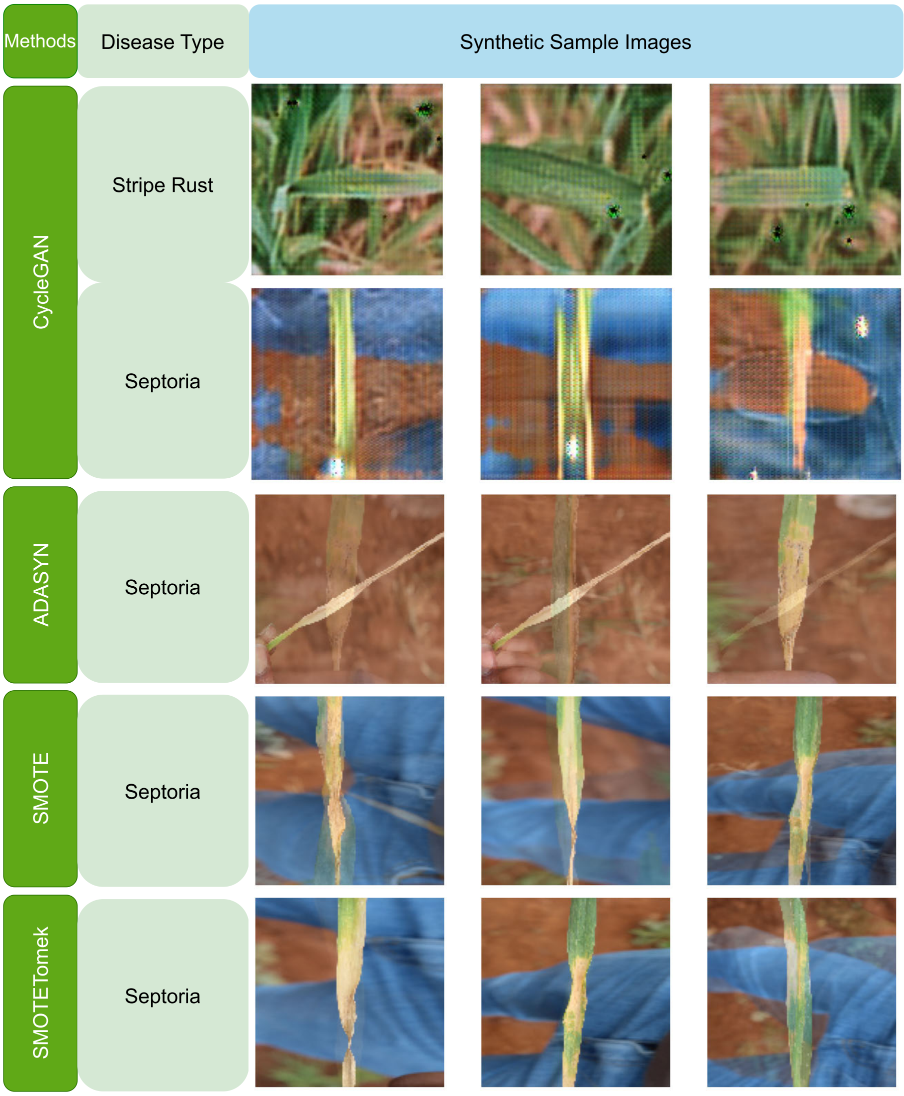

FIGURE 9. Synthetic sample images from different augmentation methods.

TABLE 3. Augmentation of dataset using different augmentation strategies.

Augmentation Methods	Disease Type	No. of Initial Samples	Total No. of Samples After Augmentation
CycleGAN	Septoria	97	301
	Stripe Rust	208	310
ADASYN	Septoria	97	215
	Stripe Rust	208	208
SMOTE	Septoria	97	208
	Stripe Rust	208	208
SMOTETomek	Septoria	97	208
	Stripe Rust	208	208

algorithm applied, the publication year, and the corresponding performance metrics expressed in percentages.

Comparing our work with others reveals diverse approaches and their associated performance outcomes. For instance, Jiang et al.'s 2022 study utilized Inception-v3 on the PlantVillage-37,721 and Field-based Wheat Diseases Images (FWDI) datasets, achieving an accuracy of 92.5%. Similarly, Kumar et al. [22] employed N-CNN-based Transfer Learning on the Powdery Mildew Wheat Disease Dataset, obtaining a commendable accuracy of 94.5%.

Our study introduces novel methodologies utilizing ResNet50V2, DenseNet169, and MobileNetV2 models in conjunction with ADASYN and CycleGAN techniques. Leveraging the ADASYN-enhanced dataset, our ResNet50V2 model surpassed prior studies, achieving an impressive accuracy of 97.65%. Furthermore, employing the CycleGAN-enhanced dataset with DenseNet169 and MobileNetV2 models yielded superior results, reaching accuracies of 99.19% and 100%, respectively. The MobileNetV2

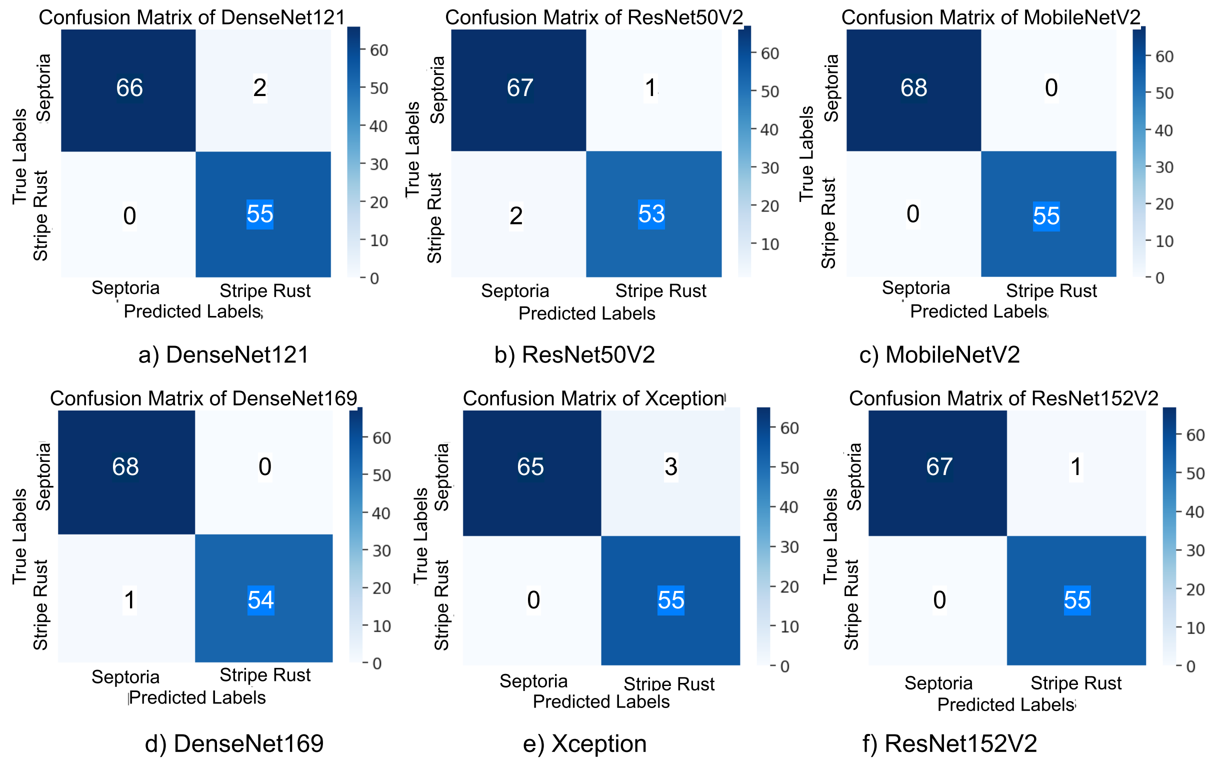

FIGURE 10. **Confusion matrices of different classifiers on CycleGAN-augmented dataset.**

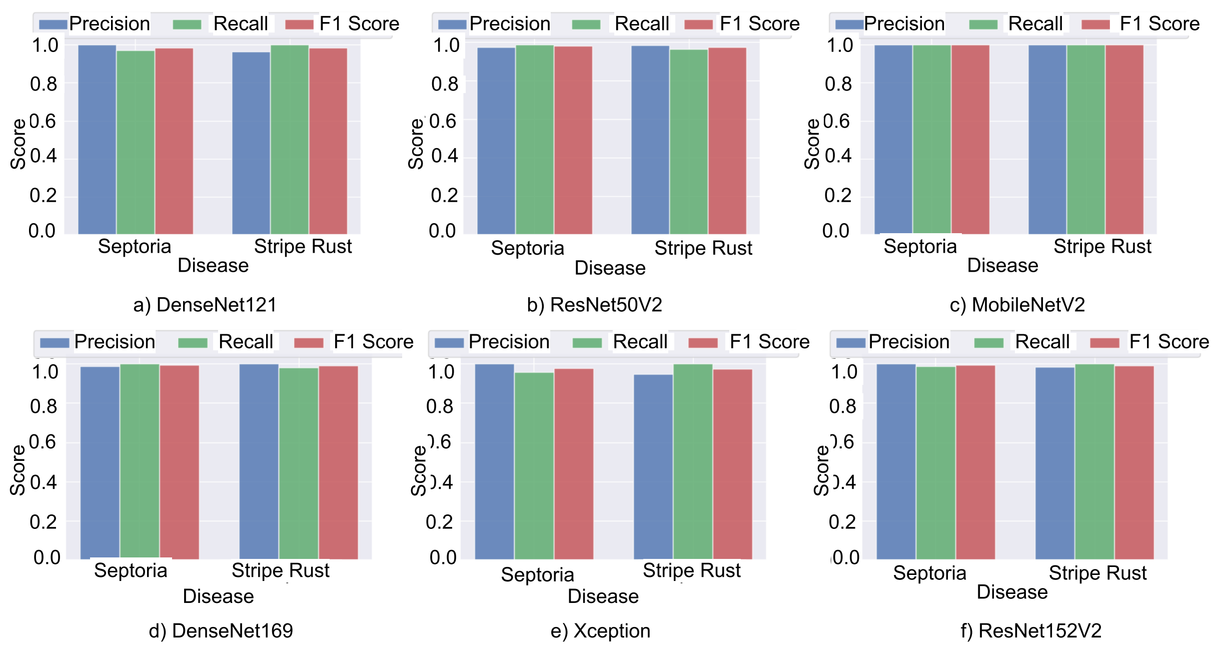

FIGURE 11. **Precision, Recall, and F1-score of different classifiers on CycleGAN-augmented dataset.**

model with the ADASYN-enhanced dataset also demonstrated perfect accuracy of 100%. These findings underscore the efficacy of our proposed strategies in enhancing wheat disease classification accuracy compared to existing methods.

The table illustrates that our approaches significantly enhance performance metrics, particularly when coupled with CycleGAN augmentation, showcasing its utility in advancing wheat disease classification models. These results emphasize how innovative augmentation methods can

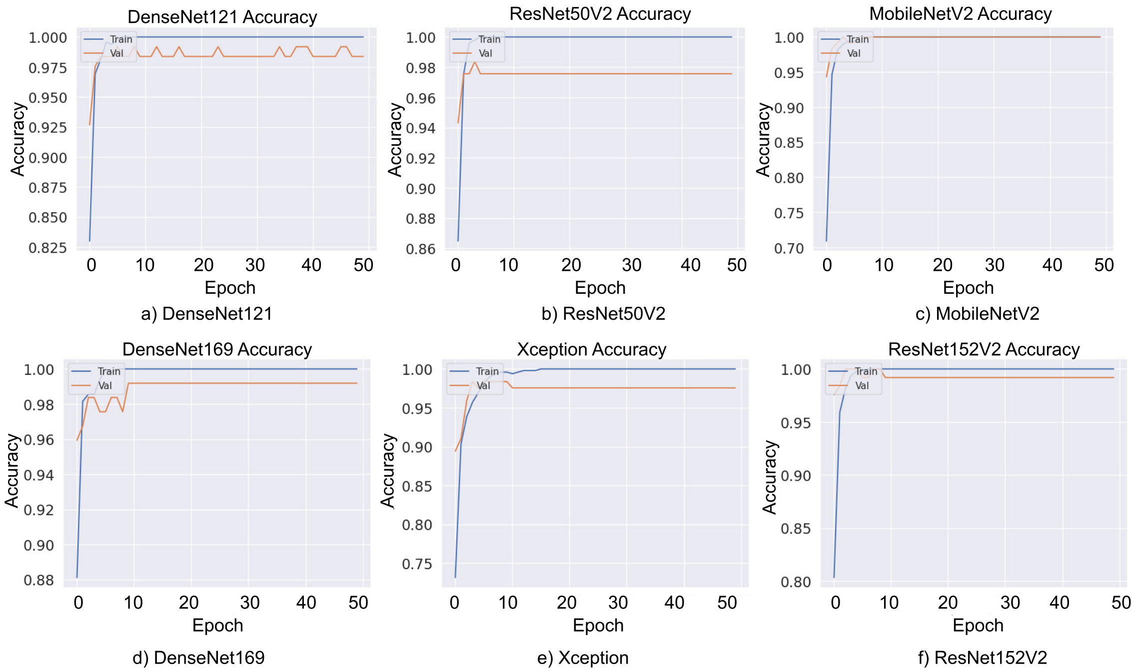

FIGURE 12. Accuracy of different classifiers on CycleGAN-augmented dataset.

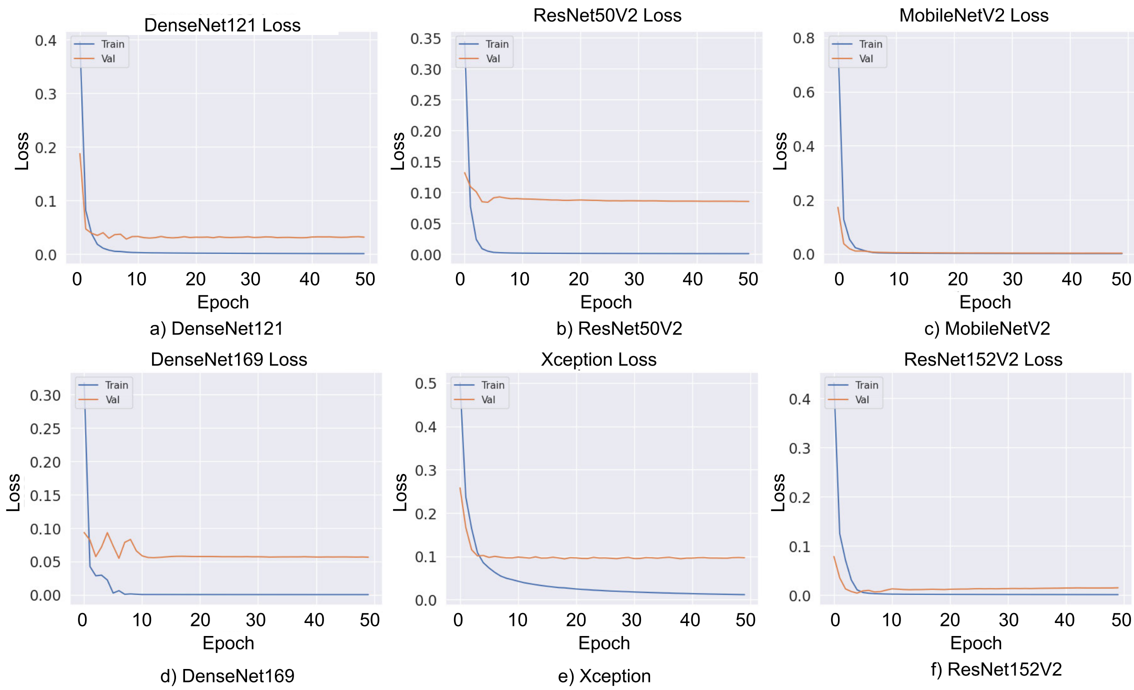

FIGURE 13. Loss of different classifiers on CycleGAN-augmented dataset.

enhance the precision and reliability of models, paving the way for more robust solutions in the realm of agricultural disease detection.

V. NOVELTY OF THE RESEARCH

The study introduces a pioneering approach to diagnosing wheat leaf diseases through the application of Generative Adversarial Network (GAN). This novel method explores a hitherto uncharted area in existing literature—employing GAN to generate synthetic images of wheat leaf diseases. The investigation addresses a critical challenge faced in the field: the scarcity of datasets pertaining to wheat disease identification.

Confronted with a limited number of images for each disease class, the research acknowledges the inherent challenge posed by data scarcity. Specifically, with only 208 images

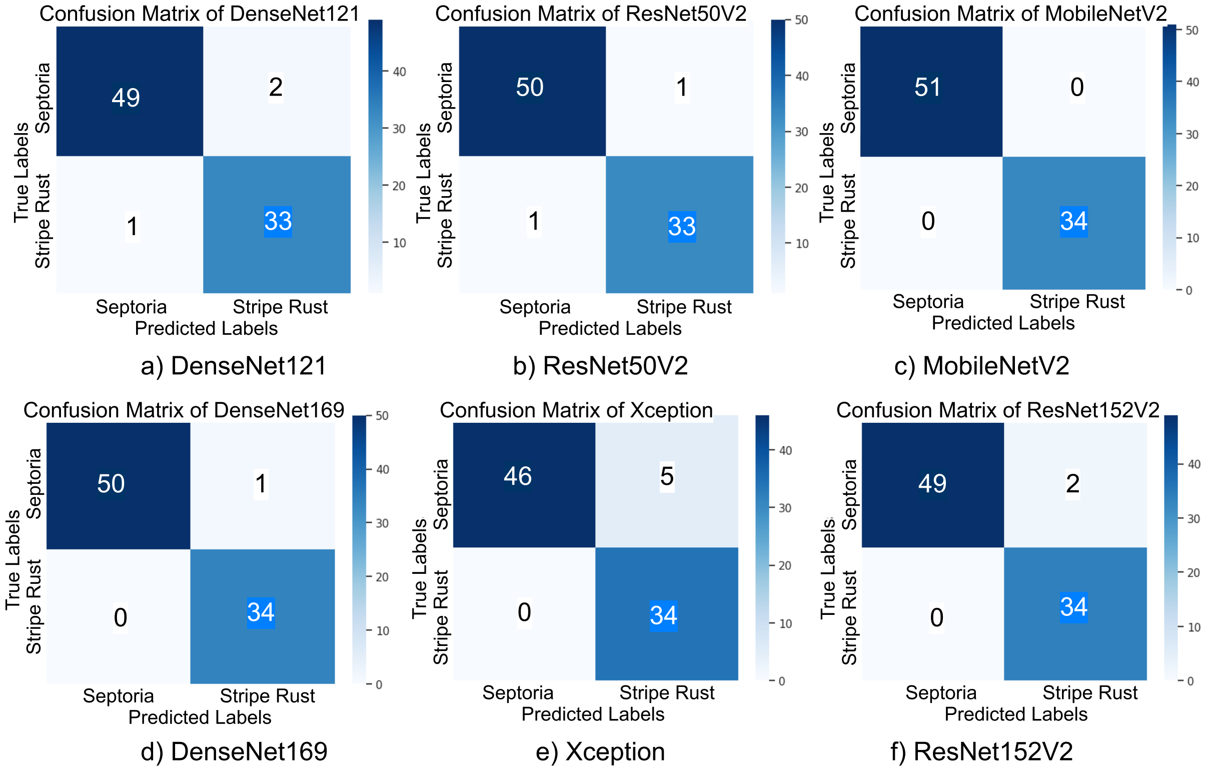

FIGURE 14. Confusion matrices of different classifiers on ADASYN-augmented dataset.

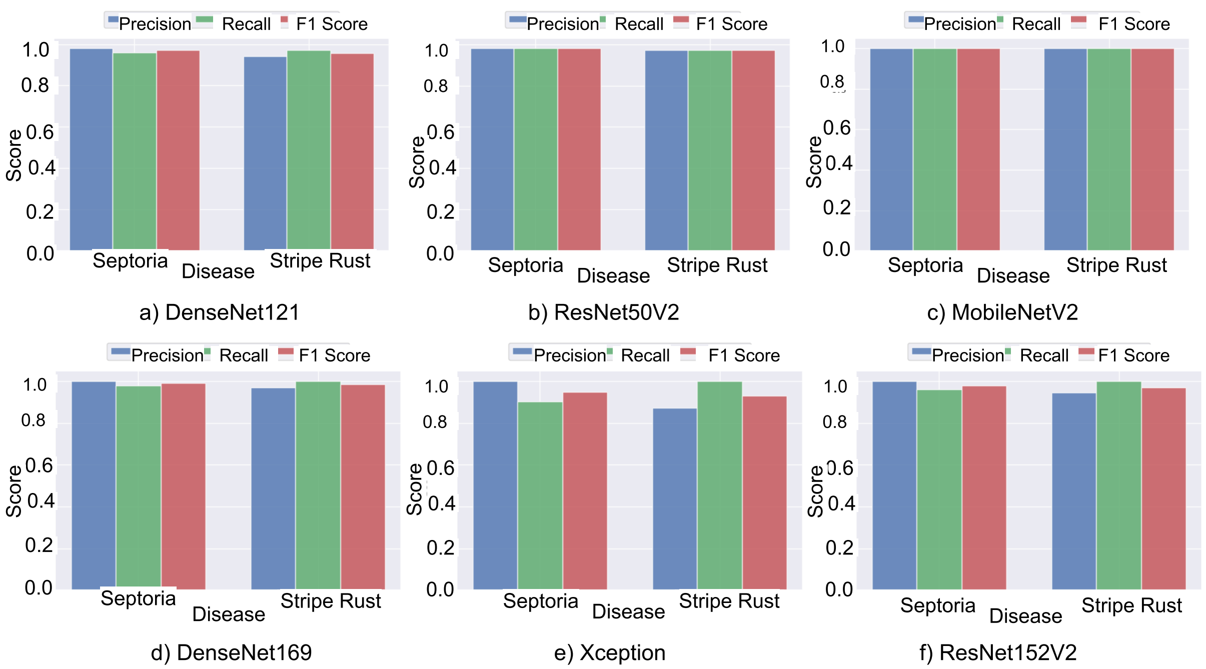

FIGURE 15. Precision, Recall, and F1-score of different classifiers on ADASYN-augmented dataset.

for "stripe rust" and 97 images for "septoria," the study recognizes and tackles the difficulty of training models with such constrained datasets. To overcome this hurdle, the research thoroughly explores various augmentation strategies. It evaluates the effectiveness of GAN techniques, including CycleGAN, and other augmentation methods

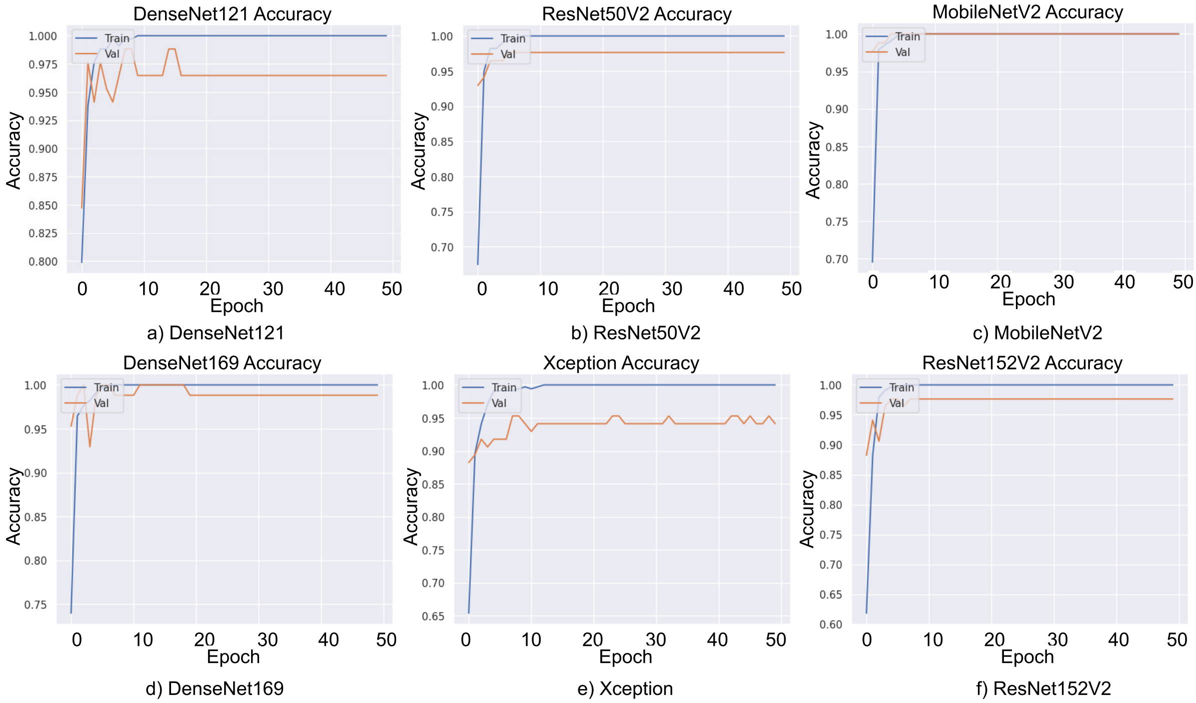

FIGURE 16. **Accuracy of different classifiers on ADASYN-augmented dataset.**

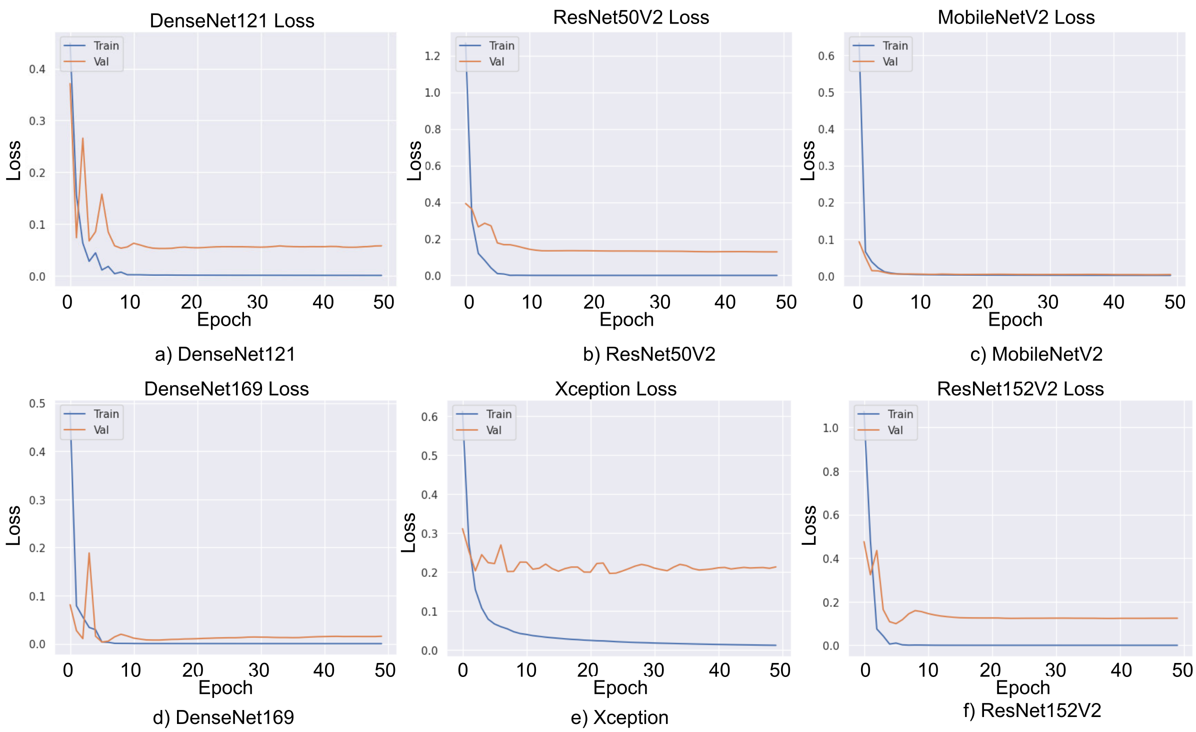

FIGURE 17. **Loss of different classifiers on ADASYN-augmented dataset.**

specifically targeting minority classes, such as SMOTE, SMOTETomek, and ADASYN. The study aims not only to increase the quantity of training data but also to enhance its quality. Through comprehensive assessments, the research elucidates the utility and efficacy of samples generated by CycleGAN in bolstering model performance when dealing with limited datasets.

In essence, this study breaks new ground in the realm of wheat disease detection by introducing innovative methodologies and harnessing GAN to augment small datasets,

TABLE 4. Performance of different classification models for wheat disease detection.

Paper	Year	Classification Method	Dataset	Performance (%)
Jiang et al., [11]	2022	Inception-v3	PlantVillage-37,721 Field-based Wheat Diseases Images (FWDI) - 2643	92.5
Kumar et al., [12]	2019	N-CNN Based Transfer Learning	Powdery Mildew Wheat Disease Dataset	94.5
Fang et al., [13]	2022	Proposed IRCE Model	Wheat Disease Detection Dataset	98.7
Lu et al., [15]	2017	Proposed DMIL-WDDS Model	Wheat Disease Database 2017 (WDD2017)	97.95
Hossen et al., [16]	2022	Proposed CNN Model	Kaggle and Github, Total-48,00 images	98.84
Kumar et al., [22]	2023	Combined CNN with STARGAN	Wheat Disease Dataset	95.6
Jin et al., [17]	2018	DCNN	Three types of Diseases: background pixels, healthy pixels, and the diseased pixels, Total-581,716 images	84.6
Hussain et al., [18]	2018	CNN(AlexNet)	4 different classes, Total-8,828 images	84.54
Bao et al., [19]	2021	Proposed SimpleNet model	Total-568 wheat ear images, wheat glume blotch images-183 images, wheat scab images-280 images, healthy wheat ear images- 105	94.10
Our Paper	-	ResNet50V2 model with ADASYN augmented dataset	Total-423, Septoria-215, Stripe Rust-208	97.65
Our Paper	-	DenseNet169 model with Cycle-GAN augmented dataset	Total-611, Septoria-301, Stripe Rust-310	99.19
Our Paper	-	ResNet152V2 model with Cycle-GAN augmented dataset	Total-611, Septoria-301, Stripe Rust-310	99.19
Our Paper	-	MobileNetV2 model with ADASYN augmented dataset	Total-423, Septoria-215, Stripe Rust-208	100
Our Paper	-	MobileNetV2 model with Cycle-GAN augmented dataset	Total-611, Septoria-301, Stripe Rust-310	100

thereby elevating the accuracy of deep-learning models. The central focus of this work lies in exploring the potential of GAN and other augmentation approaches in generating synthetic images of wheat leaf diseases and examining their impact on model performance—an area that has been notably underexplored in the existing body of research.

VI. DISCUSSION

This study sheds light on the efficacy of diverse augmentation strategies in the classification of wheat leaf diseases, addressing the crucial need for accurate identification in global agriculture. The research delves into key aspects, including preprocessing methodologies, the significance of augmentation strategies, novelty in approach, and the practical implications of the findings. The methods employed for preprocessing and feature extraction play a pivotal role in ensuring the reliability of input data for deep learning models. By grounding these choices in an extensive literature review, the study establishes a robust foundation for effective model training, particularly in overcoming challenges such as limited datasets, class imbalances, and computational demands, thereby ensuring the precision of disease identification.

The constraints imposed by limited datasets are substantially mitigated through the application of GAN-based data augmentation techniques, such as CycleGAN, along with additional augmentation methods like SMOTE, SMOTE-Tomek, and ADASYN. CycleGAN, uniquely, enhances both majority and minority classes by introducing a diverse array of images depicting both healthy and unhealthy wheat leaves to the dataset. Conversely, to address imbalanced datasets, SMOTE, SMOTETomek, and ADASYN focus exclusively on augmenting the minority class. While Cycle-GAN contributes to enhancing both the overall quality and quantity of data, the other methods predominantly tackle class imbalances. The remarkable accuracy of the MobileNetV2 model on datasets augmented by CycleGAN and ADASYN, reaching 100%, underscores the revolutionary potential of both methods in refining the accuracy of wheat leaf disease classification. However, subtle performance variations, such as the improved performance of ResNet50V2 with ADASYN-augmented data, emphasize the need for tailored model optimization. This highlights that different model architectures may require specific adjustments to fully harness the augmentation potential offered by CycleGAN.

In order to generate realistic synthetic examples without requiring paired data, CycleGAN preserves important visual properties, which makes it an ideal option for image translation. Furthermore, the domain adaption capacity allows synthetic data to be seamlessly integrated into real-world circumstances. Conversely, ADASYN is adept in adaptive oversampling and uses minority class density distribution to create synthetic samples that effectively address class imbalance. ADASYN improves model performance, especially with imbalanced datasets, by maintaining the original data distribution and lowering the chance of overfitting.

The choice between CycleGAN and ADASYN is contingent upon the type of data, task specifications, and processing capacity available. While CycleGAN excels at image translation and domain adaptation tasks using adversarial and cycle-consistency loss functions for realistic image transformation, ADASYN effectively addresses class imbalance by creating synthetic instances through interpolation between minority class examples and their nearest neighbors. We need to take into account variables including task complexity, computing limitations, and dataset characteristics when choosing the best augmentation technique for a certain application domain.

In summary, the study's approach not only addresses challenges stemming from limited datasets and class disparities but also showcases the transformative potential of CycleGAN in alleviating data shortages. CycleGAN emerges as a valuable tool for precision agriculture and disease management, exhibiting resilience and versatility in resource-constrained machine learning applications within the agricultural domain. Further exploration of CycleGAN's capabilities and its integration with diverse models holds promise for revolutionizing wheat disease categorization, contributing significantly to advancements in agricultural technology.

VII. LIMITATION AND FUTURE WORK

The limitations of the study encompass several crucial areas warranting further exploration. While the chosen augmentation strategies—CycleGAN, ADASYN, SMOTE, and SMOTETomek—are well-established methodologies, the exclusive focus on these techniques might constrain our understanding of their effectiveness in diverse settings. A more comprehensive analysis involving a broader spectrum of augmentation techniques, including novel or hybrid approaches, could offer a nuanced assessment of their merits and drawbacks across different datasets and class types.

Moreover, the study's concentration on specific diseases or classifications may limit the transferability of findings to other domains or datasets with unique characteristics, thus constraining the generalizability of conclusions to real-world scenarios. Additionally, the lack of a thorough investigation into any biases or constraints introduced by augmented data is noteworthy. Exploring how augmentation techniques impact interpretability, risks of overfitting, or model biases could yield valuable insights into the reliability and robustness of augmented datasets.

In terms of future prospects, a critical area requiring further exploration is the interpretability and explainability of augmented datasets and the resultant models. Ensuring transparency in the decision-making processes of machine learning models trained on augmented data becomes pivotal, especially in high-stakes applications. Subsequent investigations may focus on developing approaches that prioritize maintaining the semantic significance and domain relevance of augmented instances alongside increasing dataset sizes. By placing a strong emphasis on interpretability, models aim to build trust and facilitate stakeholders' understanding of the rationale behind their results.

Furthermore, integrating edge computing into these advancements could make a significant difference in achieving interpretability and explainability in real time. Implementing interpretability mechanisms directly on edge devices enables farmers or field experts to gain on-the-ground insights into model predictions and instills confidence in the decision-making process of AI-driven systems, without relying on intricate backend infrastructures. This integration aligns with the goal of enhancing model performance and ensuring practical deployment and comprehension of AI technologies in agriculture, thereby improving grassroots agricultural management and decision-making.

VIII. CONCLUSION

This research thorough analysis explores the efficacy of various augmentation techniques and a non-augmented dataset in classifying severely diseased wheat leaves. CycleGAN augmentation emerges as the most effective method, surpassing other augmentation techniques and the raw dataset. It notably outperforms ADASYN, showcasing remarkable results in classifying even a severely limited dataset of wheat leaves.

The analysis also indicates that the ResNet50V2 model performs better with ADASYN-augmented data, suggesting the necessity for further investigation into model optimization to fully harness the potential of CycleGAN augmentation across diverse model architectures.

Despite the challenges posed by severely limited datasets for septoria (97 images) and stripe rust (208 images), CycleGAN demonstrates its efficacy in generating synthetic images and achieving high accuracy in classifiers. This outcome underscores the revolutionary potential of CycleGAN in addressing data scarcity, showcasing its resilience and versatility in handling even the most challenging scenarios in resource-constrained machine learning agricultural applications.

In conclusion, this study provides compelling evidence for the effectiveness of CycleGAN augmentation in classifying wheat leaf disease. Its capacity to handle limited data and achieve high accuracy positions it as a valuable tool for precision agriculture and disease management, unlocking new possibilities for ensuring healthy crops and sustainable food production. As research progresses, further exploration of CycleGAN's capabilities and integration with diverse models holds immense promise for revolutionizing wheat disease classification and advancing the field of agricultural technology.

CONFLICT OF INTEREST STATEMENT

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ACKNOWLEDGMENT

This work was supported in part by Multimedia University, Malaysia, Grant no MMUI/230023.

REFERENCES

[1] *World Food and Agriculture—Statistical Pocketbook 2021*, Food Agricult. Org. (FAO), Rome, Italy, 2021.

[2] S. Savary, A. Ficke, J.-N. Aubertot, and C. Hollier, "Crop losses due to diseases and their implications for global food production losses and food security," *Food Secur.*, vol. 4, no. 4, pp. 519–537, Dec. 2012.

[3] M. S. Islam, S. Sultana, F. A. Farid, M. N. Islam, M. Rashid, B. S. Bari, N. Hashim, and M. N. Husen, "Multimodal hybrid deep learning approach to detect tomato leaf disease using attention based dilated convolution feature extractor with logistic regression classification," *Sensors*, vol. 22, no. 16, p. 6079, Aug. 2022.

[4] R. Bhuiyan, J. Abdullah, N. Hashim, F. Al Farid, W. N. M. Isa, J. Uddin, and N. Abdullah, "Deep dilated convolutional neural network for crowd density image classification with dataset augmentation for Hajj pilgrimage," *Sensors*, vol. 22, no. 14, p. 5102, Jul. 2022.

[5] I. Khandokar, M. Hasan, F. Ernawan, S. Islam, and M. N. Kabir, "Handwritten character recognition using convolutional neural network," *J. Phys., Conf. Ser.*, vol. 1918, Jun. 2021, Art. no. 042152.

[6] M. R. Golzarian and R. A. Frick, "Advances in machine vision applications for plant disease detection," *Annu. Rev. Phytopathol.*, vol. 58, no. 1, pp. 231–253, Aug. 2020, doi: 10.1146/annurev-phyto-010820-012841.

[7] S. T. Y. Ramadan, T. Sakib, M. M. Ul Haque, N. Sharmin, and M. M. Rahman, "Generative adversarial network-based augmented rice leaf disease detection using deep learning," in *Proc. 25th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2022, pp. 976–981.

[8] S. T. Y. Ramadan, T. Sakib, M. Ahsan Rahat, and S. Mosharrof, "CycleGAN-based data augmentation with CNN and vision transformers (ViT) models for improved maize leaf disease classification," in *Proc. IEEE 64th Int. Sci. Conf. Inf. Technol. Manage. Sci. Riga Tech. Univ. (ITMS)*, Oct. 2023, pp. 1–6.

[9] A. Dixit and S. Nema, "Wheat leaf disease detection using machine learning method—A review," *Int. J. Comput. Sci. Mob. Comput.*, vol. 7, no. 5, pp. 124–129, 2018.

[10] S. T. Y. Ramadan, T. Sakib, M. M. Ul Haque, N. Sharmin, and M. M. Rahman, "Wheat leaf disease synthetic image generation from limited dataset using GAN," in *Human-Centric Smart Computing* (Smart Innovation, Systems and Technologies), vol. 376, S. Bhattacharyya, J. S. Banerjee, and M. Kööppen, Eds. Singapore: Springer, 2024, doi: 10.1007/978-981-99-7711-6_40.

[11] J. Jiang, H. Liu, C. Zhao, C. He, J. Ma, T. Cheng, Y. Zhu, W. Cao, and X. Yao, "Evaluation of diverse convolutional neural networks and training strategies for wheat leaf disease identification with field-acquired photographs," *Remote Sens.*, vol. 14, no. 14, p. 3446, Jul. 2022.

[12] D. Kumar and V. Kukreja, "N-CNN based transfer learning method for classification of powdery mildew wheat disease," in *Proc. Int. Conf. Emerg. Smart Comput. Informat. (ESCI)*, Mar. 2021, pp. 707–710.

[13] X. Fang, T. Zhen, and Z. Li, "Lightweight multiscale CNN model for wheat disease detection," *Appl. Sci.*, vol. 13, no. 9, p. 5801, May 2023.

[14] G. Shrestha, Deepsikha, M. Das, and N. Dey, "Plant disease detection using CNN," in *Proc. IEEE Appl. Signal Process. Conf. (ASPCON)*, Oct. 2020, pp. 109–113.

[15] J. Lu, J. Hu, G. Zhao, F. Mei, and C. Zhang, "An in-field automatic wheat disease diagnosis system," *Comput. Electron. Agricult.*, vol. 142, pp. 369–379, Nov. 2017.

[16] M. H. Hossen, M. Mohibullah, C. S. Muzammel, T. Ahmed, S. Acharjee, and M. B. Panna, "Wheat diseases detection and classification using convolutional neural network (CNN)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 11, 2022.

[17] X. Jin, L. Jie, S. Wang, H. Qi, and S. Li, "Classifying wheat hyperspectral pixels of healthy heads and fusarium head blight disease using a deep neural network in the wild field," *Remote Sens.*, vol. 10, no. 3, p. 395, Mar. 2018.

[18] A. Hussain, M. Ahmad, I. A. Mughal, and H. Ali, "Automatic disease detection in wheat crop using convolution neural network," in *Proc. 4th Int. Conf. Next Gener. Comput.*, vol. 34, 2018, pp. 22–25.

[19] W. Bao, X. Yang, D. Liang, G. Hu, and X. Yang, "Lightweight convolutional neural network model for field wheat ear disease identification," *Comput. Electron. Agricult.*, vol. 189, Oct. 2021, Art. no. 106367.

[20] A. Alharbi, M. U. G. Khan, and B. Tayyaba, "Wheat disease classification using continual learning," *IEEE Access*, vol. 11, pp. 90016–90026, 2023.

[21] S. Savary, A. Djurle, J. Yuen, A. Ficke, V. Rossi, P. D. Esker, J. M. C. Fernandes, E. M. Del Ponte, J. Kumar, L. V. Madden, P. Paul, N. McRoberts, P. K. Singh, L. Huber, C. P. de Vallavielle, S. Saint-Jean, and L. Willocquet, "A white paper on global wheat health based on scenario development and analysis," *Phytopathology*, vol. 107, no. 10, pp. 1109–1122, Oct. 2017.

[22] D. Kumar and V. Kukreja, "Combined CNN with STARGAN for wheat yellow rust disease classification," *Int. J. Comput. Digit. Syst.*, vol. 13, no. 1, pp. 1081–1095, Apr. 2023, doi: 10.12785/ijcds/130187.

[23] A. U. Ruby, J. G. C. Chandran, B. N. Chaithanya, T. J. S. Jain, and R. Patil, "Wheat leaf disease classification using modified ResNet50 convolutional neural network model," *Multimedia Tools Appl.*, pp. 1–19, Jan. 2024.

[24] S. T. Y. Ramadan, T. Sakib, M. A. Rahat, S. Mosharrof, F. I. Rakin, and R. Jahangir, "Enhancing mango leaf disease classification: ViT, BiT, and CNN-based models evaluated on CycleGAN-augmented data," in *Proc. 26th Int. Conf. Comput. Inf. Technol. (ICCIT)*, vol. 1, no. 1, 2023, pp. 1–6, doi: 10.1109/ICCIT60459.2023.10441374.

[25] 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.

[26] H. Nazki, J. Lee, S. Yoon, and D. S. Park, "Image-to-image translation with GAN for synthetic data augmentation in plant disease datasets," *Korean Inst. Smart Media*, vol. 8, no. 2, pp. 46–57, Jun. 2019.

[27] N. Vasudevan and T. Karthick, "A hybrid approach for plant disease detection using E-GAN and CapsNet," *Comput. Syst. Sci. Eng.*, vol. 46, no. 1, pp. 337–356, 2023.

[28] H. Getachew, "Wheat leaf dataset," Mendeley Data, Elsevier, Ethiopia, Tech. Rep. wgd66f8n6h.1, 2021.

SYED TAHA YEASIN RAMADAN received the Bachelor of Science degree from the Department of Computer Science and Engineering, Military Institute of Science and Technology (MIST), Dhaka, in 2023. His research interests include artificial intelligence, deep learning, machine learning, natural language processing (NLP), computer vision, bioinformatics, human brain, and human–computer interaction. He has made significant contributions to his field through published research articles. Concurrently, he actively participates in diverse research projects, showcasing exceptional programming proficiency in artificial intelligence, machine learning, and deep learning.

TANJIM SAKIB received the Bachelor of Science degree from the Department of Computer Science and Engineering, Military Institute of Science and Technology (MIST), Dhaka, in 2023. His research interests include sentiment analysis, deep learning, machine learning, natural language processing (NLP), bioinformatics, human brain, and human–computer interaction. He has contributed to his field by publishing research articles. Alongside his professional engagements, he is actively involved in diverse research projects, showcasing strong programming expertise in machine learning and deep learning.

FAHMID AL FARID (Member, IEEE) received the B.S. degree in computer science and engineering from the University of Chittagong, Bangladesh, in 2010, the M.S. degree from the Faculty of Computer Science and Electrical Engineering, University of Ulsan (UOU), South Korea, in 2015, and the Doctor of Philosophy (Ph.D.) degree in information technology from Multimedia University, Cyberjaya, Malaysia. From 2013 to 2014, he was a Research Assistant with the Embedded System Laboratory, UOU. In 2015, he was a Research Assistant with the Ubiquitous Computing Technology Research Institute (UTRI), Sungkyunkwan University, South Korea. He is currently a Postdoctoral Scientist with the Faculty of Engineering, Multimedia University. He received Korean BK21 PLUS Scholarship supported by Korean Government for his M.S. degree, from 2012 to 2014. He also received an ICT Fellowship from Bangladesh Government, in 2014. His current research interests include artificial intelligence, algorithm design, computer vision, human–computer interaction, image and video analysis, power generation, and green technology.

MD. SHOFIQUL ISLAM received the B.Sc. degree in computer science and engineering from Islamic University, Kushtia, Bangladesh, in 2015, and the M.Sc. (research-based) degree from the Faculty of Computing, University Malaysia Pahang, Malaysia. After the consummation of his B.Sc. degree, he joined as a Lecturer with the Department of Computer Science, First Capital University of Bangladesh (FCUB), Chuadanga, Bangladesh, in 2015. In 2016, he joined as a Senior Lecturer of computer science and engineering with the Atish Dipankar University of Science of Technology (ADUST), Dhaka, Bangladesh. He received the Graduate Research Scheme Scholarship from the University Malaysia Pahang for his M.Sc. studies. He was a Senior Lecturer with the World University of Bangladesh (WUB), Uttara, Dhaka. He is currently an Assistant Professor with the Department of CSE, Military Institute of Science and Technology (MIST), Dhaka. His research interests include sentiment analysis, deep learning, machine learning, bioinformatics, human brain, and human–computer interaction. He has published more than 13 research high-impact research articles in his research field. In addition to his professional activities, he is working on various research projects with excellent programming knowledge in machine learning and deep learning.

JUNAIDI BIN ABDULLAH (Member, IEEE) received the Bachelor of Engineering (B.Eng.) degree (Hons.) from the University of Bristol, U.K., and the Ph.D. degree in computer science in the area of computer vision and augmented reality from the University of Southampton, U.K., in 2005. He is currently an Associate Professor with the Faculty of Computing and Informatics, Multimedia University (MMU), Malaysia. He is also the Dean of FCI, MMU. He is also the Chairperson of the Assistive Technology Special Interest Group. He has published more than 60 internationally multi-disciplinary refereed conference papers, journal articles, and books. His research interests include augmented reality, image and video processing, and computer vision.

MD ROMAN BHUIYAN received the master's degree in software engineering from the Faculty of Computer Science and Engineering, FTMS College Malaysia, in collaboration with Leeds Beckett University, U.K., in 2017, and the Doctor of Philosophy (Ph.D.) degree in information technology from Multimedia University, Cyberjaya, Malaysia, in 2022. Currently, he is a Postdoctoral Fellow with Fraunhofer IGD, Rostock, Germany. He is also a Lecturer of IT with Berlin School of Business and Innovation (BSBI). His current research interests include artificial intelligence, algorithm design, computer vision, deep learning, machine learning, and image and video analysis.

SARINA MANSOR received the B.Eng. degree (Hons.) in electronics and electrical engineering from the University of College London, in 1998, the M.Eng.Sc. degree from Multimedia University (MMU), Malaysia, in 2002, and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 2009. She is currently a Senior Lecturer with the Faculty of Engineering, MMU. She is also the Program Coordinator for B.Eng. studies. Her research interests include signal and image analysis, medical imaging, computer vision, machine learning, and the Internet of Things.

HEZERUL BIN ABDUL KARIM (Senior Member, IEEE) received the B.Eng. degree in electronics with communications from the University of Wales Swansea, U.K., in 1998, the M.Eng. degree in science from Multimedia University, Malaysia, in 2003, and the Ph.D. degree from the Center for Communication Systems Research (CCSR), University of Surrey, U.K., in 2008. He has been an Associate Professor with Multimedia University, since November 2015. He is currently the Deputy Dean of Student Affair and Alumni with the Faculty of Engineering, Multimedia University. He has been teaching multimedia and computing engineering subjects. His research interests include telemetry, 2D/3D image/video coding and transmission, error resilience and multiple description video coding, and deep learning in image and video. He is currently supervising and co-supervising several postgraduate students. He is the Vice Chair of the IEEE Signal Processing Society, Malaysia Section.