Enhancing Maize Disease Detection: A Comprehensive Study on Few-Shot Learning, Object Detection, and Synthetic Data Generation Using YOLOv5s-C3CBAM Model

Hassan Satti
University of Punjab Lahore
***Correspondence**: hassan.7ti@gmail.com
Citation | Satti. H, "Enhancing Maize Disease Detection: A Comprehensive Study on Few-Shot Learning, Object Detection, and Synthetic Data Generation Using YOLOv5s-C3CBAM Model", IJASD, vol. 5, no. 4, pp. 186-197, Nov 2023
Received| Oct 18, 2023**; Revised|** Oct 30, 2023**; Accepted|** Nov 08, 2023; **Published|** Nov 12, 2023.

As the global population grows, ensuring food production, particularly of staple crops like maize, becomes crucial. Maize faces challenges from diseases impacting both yield and quality, threatening global food security. Traditional disease detection methods are inefficient, prompting a shift towards computer vision and machine learning. However, data scarcity hinders model training. This study explores few-shot learning to address this issue. In the context of maize in China, leaf diseases pose a significant threat, causing substantial economic losses. Current research focuses on single-leaf disorders, emphasizing the need for advanced detection algorithms, especially for complex diseases. Additionally, accurate plant counting during growth stages is vital for effective management, with computer vision-based approaches offering promising solutions, especially with Unmanned Aerial Vehicles (UAVs). Object detection technology, employing both one-stage and two-stage methods, plays a pivotal role in agricultural research. YOLOv5s emerges as an efficient model, demonstrating success in various applications, including disease detection and plant counting. The study introduces a comprehensive methodology involving dataset expansion, Cycle GAN for synthetic data, and the YOLOv5s-C3CBAM model for maize disease detection. Results indicate the superiority of the YOLOv5s-C3CBAM model, with improvements in mean average precision (mAP), recall, F1 score, and precision. The study delves into model comparisons, experimental parameters, and disease identification accuracy. The YOLOv5s-C3CBAM model achieves 83% mAP_0.5, outperforming other models. Limitations, such as focusing on three diseases, are acknowledged, and future directions involve building a more comprehensive dataset. The study introduces innovations in image generation using Cycle GAN and attention processes, addressing challenges in disease detection accuracy. Despite limitations, the YOLOv5s-C3CBAM model contributes to accurate crop disease diagnosis, serving as a reference for future agricultural research.
Keywords: Staple Crops, Disease Detection, Single Leaf Disorders, Unmanned Aerial Vehicles, Growth Stages.

Introduction:

As the global population continues to grow, the significance of food production becomes increasingly pronounced. Maize, being one of the most widely cultivated and highest-yielding food crops globally, plays a crucial role in the global food supply. However, challenges arising from diseases affecting maize production pose significant obstacles to agricultural productivity [1]. These diseases not only directly impact maize yields but also diminish maize quality, thereby affecting farmers' income and global food security. Traditional methods of maize disease detection rely heavily on manual observation and identification, which are not only inefficient but also constrained by manual expertise and skills. In recent years, with the advancements in computer vision and machine learning technologies, there has been a growing trend in utilizing these technologies for disease detection. However, these methods often necessitate a substantial amount of labeled data to train the model, posing difficulties in practical applications, especially for specific tasks like agricultural disease detection [2].

In the domain of crop research utilizing traditional machine-learning techniques, several studies have been conducted. For instance, traditional machine learning methods like Linear Regression, Random Forest, and Support Vector Machine were employed to assess multispectral imagery, enabling the estimation of nitrogen content in maize canopies. Demonstrations of predictive performance for large-scale maize height mapping, utilizing radar data for crop height retrieval technology, were presented [3]. Combined features extracted from images were used alongside classifiers such as Linear Discriminant Analysis and SVM for classification training. Convolutional Neural Network (CNN) was incorporated into plant disease monitoring, optimizing its parameters to significantly enhance detection speed and accuracy. It is noteworthy that all these studies relied on datasets of substantial volume. In this context, few-shot learning emerges as a new solution. Few-shot learning, a machine learning method focused on understanding and recognizing new categories by learning from a small number of samples, offers a promising solution to the challenge of data scarcity in agricultural disease detection, thereby improving the efficiency and accuracy of disease detection [4].

Maize holds a pivotal role in China, serving as the primary staple grain crucial for sustaining livestock, aquaculture, and supplying raw materials to various industries. Ensuring the productivity and quality of maize is essential. However, the crop faces a significant threat from leaf diseases throughout its growth stages. Failure to promptly address these diseases can negatively impact maize quality and overall productivity. The outbreak of northern leaf blight (NLB) from 2012 to 2015 alone resulted in substantial losses in maize production and the US economy. The prevalence of maize diseases in China poses a serious threat, causing an annual decline in agricultural output ranging from 6% to 10%, with a 30% probability of exceeding this range under exceptional conditions [5]. The maize sector encounters challenges, primarily from rust, grey leaf spot, and blight, with blight causing larger lesions despite similar hues. Compound sickness, involving multiple diseases on a single maize leaf, adds complexity to disease detection. Current research primarily focuses on single-leaf disorders, underscoring the need for advanced disease detection algorithms to accurately identify complex leaf diseases in maize. Maize, a globally cultivated crop with significant roles in food production, fodder, and biomass fuel, necessitates meticulous counting during various growth stages for effective management strategies. Plant counting is crucial for supplementary planting, pest management, and yield forecasting [6].

Traditionally, manual counting methods are employed, leading to uncertainty and time inefficiency due to spatial variability. To address these challenges, a computer vision-based approach for automatic plant counting in the field is imperative. Recent advancements in computer vision and artificial intelligence have facilitated accurate object detection and counting methods. While field robots integrated with plant detection models have shown promise, they are limited to low-growing crops. Unmanned Aerial Vehicles (UAVs), on the other hand, offer a preferred solution for taller crops like maize, providing high-resolution images and overcoming cloud occlusion [7]. Classical COMPUTER VISION algorithms based on concrete features often face challenges in complex field environments, leading to high error rates. Abstract-feature-based methods, particularly those employing deep learning techniques like Convolutional Neural Networks (CNNs) and the You Only Look Once (YOLO) series, have demonstrated superior performance in detecting plants in intricate field conditions. The YOLOv5s model, known for its real-time performance and lightweight characteristics, emerges as a promising candidate for efficient plant detection, especially in UAV applications, where accuracy, speed, and lightweight models are crucial considerations [8].

Object detection technology has become increasingly integral in agricultural research, employing two main approaches: one-stage and two-stage methods, each differing in their speed and level of detail in object identification. Models like R-CNN and Faster R-CNN follow a two-stage approach, whereas the YOLO series exemplifies one-stage object detection. For instance, Faster R-CNN has been enhanced for accurate diagnosis of apple leaf diseases, incorporating features such as the feature pyramid network (FPN) and a cascaded method, resulting in an average accuracy gain of 8.7%. Meanwhile, YOLOv5s has been employed to improve the identification of small objects, addressing issues like false negatives in apple fruit classification [9]. The Soft-NMS algorithm, DFP module, and RFA module were enhanced, leading to improvements in mean average precision, recall, and precision by 3.6%, 6.8%, and 6.1%, respectively. Another application involved using YOLOv5s for precise counting of red jujubes in orchards, achieving advancements in precision, recall, F1-score, average precision, frames per second, and precision metrics. Similarly, YOLOv5s was utilized for real-time detection of apple peel diseases, demonstrating an average precision rate of 87.2%. These examples showcase the successful application of object detection technology in agriculture, particularly emphasizing the efficiency of models like YOLOv5 in detecting crop diseases. The attention mechanism, incorporated into YOLOv5, enhances computer comprehension of visuals by mimicking the human visual system. This involves an encoder and decoder, where the decoder assigns weights to feature representations using an "attention map," guiding the model to focus on salient regions and ignore irrelevant ones [10].

Literature Review:

[11] introduced the PD-Net model, a comprehensive system for the precise classification of diseases and pests, effectively addressing challenges associated with a wide range of diseases and pests. PD-Net employs a convolutional block attention model, enhancing the conventional network model by integrating a hybrid cross-channel and spatial domain attention mechanism. The model utilizes a cross-layer nonlocal module to enhance multi-scale characteristic integration. Empirical results demonstrate PD-Net's superior performance in large-scale multiclass tasks for accurately identifying diseases and pests. [12] focused on improving the efficiency and speed of plant leaf disease identification, particularly in the context of apple leaf diseases. They integrated the SE attention mechanism into the Res Net network design, resulting in a substantial reduction in the detection error rate to 1.52%.

In [13], the goal was to enhance the accuracy of agricultural pest and disease identification. They proposed the use of I_CBAM, an improved CBAM attention module that simultaneously merges channel attention and spatial attention. The hybrid attention module outperformed several convolutional neural network models in reliably and precisely classifying illnesses and pests. Another study by [13]. aimed at accurately identifying pests and Guangfu Hand disease in challenging environmental settings. They improved the YOLOv5 network model by incorporating the CBAM hybrid attention mechanism. This modification significantly increased the model's recognition accuracy, averaging 93.06%. These studies collectively suggest that incorporating attention mechanisms into object identification algorithms can lead to enhanced overall performance, particularly in disease and pest identification in agricultural contexts.

To address the challenges of insufficient data for training deep learning models in the context of maize grey leaf spot disease, [14] proposed a technique using a cycle-consistent adversarial network (Cycle GAN) to create synthetic images. The generated dataset was then utilized to train a model for accurate disease identification. [15] employed picture synthesis techniques to mitigate the impact of limited data on the accuracy of digital plant disease phenotyping. They used a generative adversarial network (GAN), specifically a DC-GAN, to create artificial training data for bacterial spot disease. The results demonstrated that the third training dataset generated by DC-GAN outperformed the original dataset, emphasizing the potential of generative adversarial networks in enhancing object recognition model training with limited data.

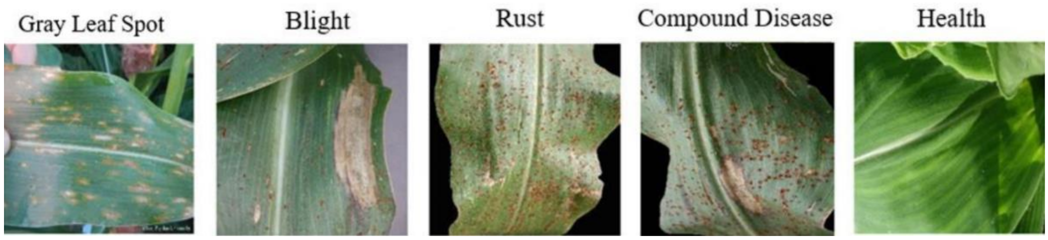

Figure 1: Illustration of different disease lesions on maize leaves [16].

The [8] utilized the "Plant Village" dataset to illustrate four bacterial infections, two viral diseases, two mold diseases, and one mite-related ailment, along with images of unaffected leaves across 12 crop species. Various machine learning (ML) approaches such as Support Vector Machines (SVMs), grey-level co-occurrence matrices (GLCMs), and Convolutional Neural Networks (CNNs) were employed for prediction models. The study also incorporated a K-Means Clustering (KMC) operation on real-time leaf images, achieving an overall accuracy of 99% and 98% for rice trees and apples, respectively. Multi-class classification metrics, including precision, recall, and F-measure, were employed to evaluate the study's precision in a single symptom pool for each class.

In [9], the authors proposed an enhanced CNN technique for rice disease detection, demonstrating the efficacy of deep neural networks (DNNs) in plant disease detection through image classification. A comparative analysis showed accuracy rates of 80%, 85%, 90%, and 95% for Transfer Learning (TL), CNN with TL, Artificial Neural Network (ANN), and Enhanced CNN with Genetic Algorithm (ECNN+GA) techniques, respectively [17]. The study in [18] highlighted limitations in identifying rice leaf disease due to image backgrounds and acquisition conditions. Evaluation of Transfer Learning (TL) models for rice leaf disease detection, including frozen layers and fine-tuning methods, revealed exceptional testing accuracy for DenseNet169 (99.66%) and Xception (99.99%).

In [19], the authors introduced Ant Colony Optimization with Convolutional Neural Network (ACO-CNN), a novel DL technique for disease detection. ACO assessed disease diagnostics, and the CNN classifier enhanced feature extraction, outperforming C-GAN, CNN, and SGD models in terms of accuracy, precision, recall, and F1 score. The study by [20] presented the PPLCNet, a DL model with dilated convolution, a multi-level attention mechanism, and Global Average Pooling (GAP) layers. The model achieved recognition accuracy and F1-score of 99.702% and 98.442%, respectively, using novel weather data augmentation and a lightweight CBAM attention mechanism. In [21], an effective CNN model was proposed for categorizing tomato leaf diseases, achieving an accuracy of 96%. The 2DCNN model with 2-Max Assembling covers and fully connected layers outperformed SVM, VGG16, Inception V3, and Mobile Net CNN models. [22] employed Model Engineering (ME) to enhance feature discrimination and processing speed. SVM models, including dilated learning, outperformed the traditional ResNet-18 design, achieving an average accuracy of 98.5% for leaf disease recognition models.

The ensemble classifiers in [23] achieved a top ensemble classifier accuracy of 96%, surpassing recent state-of-the-art DL techniques. [24] proposed a hybrid DL approach for the early detection and classification of tomato plant leaf diseases, combining CNN, Convolutional Attention Module (CBAM), and SVM, outperforming other DL approaches. In [25], the authors obtained mAP and accuracy values of 98.10% and 99.97%, respectively, for plant leaf disease detection using the publicly available Plant Village Kaggle dataset. [26] proposed an aggregated loss function combining triplet and cross-entropy loss with MobileNetV2, achieving improved accuracy in plant disease classification. [27] introduced a DL approach for ginger disease detection, achieving a test accuracy of 95.2%, providing a fast and deployable solution for early disease detection.

[28] developed a DL model using convolutional networks for classifying plant diseases, achieving an accuracy of 99.5% and a mAP, precision, and recall of 65%, 59%, and 65%, respectively. [29] presented a DL-based CNN solution for classifying and distinguishing cotton leaf diseases, achieving training and testing accuracy of 100% and 90%, respectively. [30] detected various diseases using hybrid image processing and decision tree techniques, achieving an overall accuracy of 94.5%. [12] used a mixture of RGB and blending images for plant disease classification, achieving a Genuine Acceptance Rate (GAR) of 96.7%. [15] proposed a DL-based method for automated tomato disease detection using image segmentation, achieving an accuracy of 98.12%. [31] developed an autonomous DL method for detecting and classifying coffee plant diseases, comparing training from scratch and TL strategies, and achieving high accuracy rates. [32] proposed a DL-based system for tomato plant disease detection, utilizing Inception Net and Modified U-Net, achieving high accuracy and outperforming existing methods.

Methodology:

Data Collection:

The initial dataset was obtained from publicly available sources, including Paddle, Open Data Lab, and Kaggle, resulting in 2107 authentic photographs. The dataset included images depicting various maize leaf diseases such as blight, rust disease, grey leaf spot disease, and compound diseases, along with samples of healthy leaves.

Dataset Expansion Using Cycle GAN:

The Cycle GAN approach was employed to address the challenge of inadequate data on compound illnesses affecting maize leaves. Synthetic data was generated using Cycle GAN, focusing on healthy leaves and three different types of images representing specific diseases, thus enhancing the variety of the training set.

Cycle GAN Architecture and Training:

The architecture and training process of Cycle GAN was detailed, emphasizing its ability to reliably map connections between different domains without paired data. The training process included four stages: output selection, training of the generator model, image generation, and data processing.

YOLOv5s-C3CBAM Model Introduction:

The YOLOv5s-C3CBAM maize leaf disease detection model with an attention mechanism was introduced, emphasizing its potential to accurately identify complex diseases in maize leaves.

Training the YOLOv5s-C3CBAM Model:

The model underwent training using the expanded dataset, consisting of authentic and synthetic images, to improve resilience and accuracy. The study highlighted the significance of the generated dataset in improving disease diagnosis accuracy, addressing challenges like missed identifications, and reducing false alarms.

Challenges in Generating Images:

Specific challenges were observed in generating images for grey leaf spot disease due to small lesions and lacking noticeable color features. The model underwent specialized training for 130 epochs with a batch size of 4 to address these challenges and enhance the final photos.

Iterative Training for Image Quality:

The training strategy involved isolating specific traits and increasing the number of training epochs to improve image quality, especially for intricate and challenging features.

Selection of Images Meeting Standards:

A total of 150 blight images, 150 rust disease images, and 200 grey leaf spot disease images were selected after a detailed analysis of maize leaf disease photos generated by Cycle GAN.

Composite Disease Image Generation:

Four sets of experiments were conducted to generate visual representations of composite diseases, including the transition from healthy to sick leaves. Training sets A and B were created, with the latter aiming to enhance picture quality, especially for limited composite disease data.

Testing Phase for Composite Diseases:

The model underwent testing for 150 epochs with a batch size of 4, utilizing an evolving training set A and a consistent training set B. The group that utilized images of healthy leaves to construct composite diseases yielded the most favorable outcomes during the testing phase. This methodology outlines the comprehensive process of dataset expansion, model training, and testing for the identification of complex maize leaf diseases, combining authentic and synthetic data to enhance the model's accuracy and resilience.

Results and Discussion:

The observed phenomena may be attributed to the distinctive qualities of undamaged foliage, the extensive number of model iterations, and minimal disturbance during the conversion of the composite sick dataset into images of healthy leaves. The transformation and generation approach yielded superior outcomes, especially considering the unique characteristics of large spot and rust diseases, which facilitated effectiveness in studying the progression from blight and rust disease to composite illnesses. The study aimed to generate composite illnesses of grey leaf spots using low-resolution images, a challenge due to the relatively inconspicuous signs of grey leaf spot disease compared to other illnesses. Consequently, substandard composite images of grey leaf spots, blight, and rust were generated. During the training phase, weights yielding optimal results in image production were used to create the final image. After a thorough assessment, 150 excellently crafted photographs depicting illnesses were selected for training the model. Previous experiments harnessed the picture-generating capacity of Cycle GAN to produce images of both healthy and damaged maize leaves, along with various combinations of defects. A total of 100 top-notch images of robust maize leaves were successfully gathered, meeting all specified criteria. Figure 1 showcases a variety of leaf photographs from several categories, contributing to the dataset for the analysis and research of diseases affecting maize leaves.

In addition to generating images of maize leaves, Cycle GAN combines statistical data from the original photos used to train the detection model. Three adaptive attention mechanisms, namely the SE (squeeze and excitement) attention mechanism, the coordinate attention (CA) mechanism, and the Convolutional block attention module (CBAM) mechanism, were employed to improve the model's performance. These mechanisms enhance the weighting of data from different channels and spatial dimensions, emphasizing important qualities in feature maps. The network design of YOLOv5s was optimized by incorporating attention mechanisms. Method 1 introduced an attention mechanism module into the YOLOv5 backbone network before the last SPPF layer, aiming to enhance the extraction of high-level characteristics. Method 2 integrated an attention mechanism module following the C3 module in the YOLOv5 backbone network to improve overall efficiency in detecting complex diseases affecting maize leaves.

The image undergoes four passes through the C3 module of the backbone network before reaching the final SPPF layer. Each of the three Conv modules employs a 1x1 convolution operation for dimensionality modification, and the remaining connections execute the two Conv modules constituting the bottleneck module. The initial 1x1 convolution operation reduces the channel size by 48%, while the subsequent 3x3 convolution operation doubles the number of channels. To simplify the interpretation of feature data, dimensionality reduction is performed using convolutional kernels, and dimensionality expansion allows the extraction of more complex features. To address gradient vanishing, the bottleneck module utilizes residual connections, combining input and output to mitigate the issue. Enhancing the C3 module involves incorporating an attention mechanism that assigns attention weights at the beginning of the feature extraction process. Given its presence four times in the backbone network, the attention module can influence the system on four separate occasions. The attention weights are then used to modify the preceding attention feature maps after the C3 module. An attention mechanism is applied after the C3 module to improve the extraction of feature maps at both shallow and deep levels.

Setting up and evaluating experimental parameters involves adjusting the model and experimental platform parameters. The text does not provide specific details in this regard. A cloud server running Linux, PyTorch 1.9.1, and Python 3.7 were used for model construction, training, and testing. The testing employed a GeForce RTX 2080 Ti GPU with a memory capacity of 11,019 MiB. The training approach used a batch size of 16, resized training photographs to 256x256 pixels, and involved 300 iterations. The learning rate of the network model was adjusted using the cosine annealing decay approach, starting with an initial value of 0.01. The system was trained to categorize objects into four distinct classes.

Comparing the updated models to the YOLOv5s network model, each upgraded model exhibits varying degrees of improvement in detection accuracy. The CBAM module, integrated into the YOLOv5s-C3CBAM model after the C3 module, shows the highest mean average precision (mAP) of 83% at a threshold of 0.5, a 3.1 percentage point improvement over the preceding YOLOv5s network model. The CBAM mechanism, integrating both channel and spatial information, contributes to the enhancement. Incorporating the CBAM technique into the C3 module significantly boosts feature extraction performance. The YOLOv5s-C3CBAM model achieves higher detection accuracy compared to its predecessors. The YOLOv5s-C3CA model performs the best with a recall rate of 72%, indicating efficacy in detecting tiny objects and reducing false negatives. The CA mechanism module, considering links between different sites and precise positional data, enhances the model's memory capacity.

Table 1: Model Training Details

Parameter	Value
Training Platform	Cloud server running Linux
Framework	PyTorch 1.9.1
Python Version	3.7
GPU	GeForce RTX 2080 Ti (11,019 MiB memory)
Batch Size	16
Resized Training Photos	256x256 pixels
Training Iterations	300
Learning Rate Adjustment Method	Cosine Annealing Decay
Classes	4 (Objects to be categorized)

The YOLOv5s-CBAM model achieves the highest F1 score, surpassing the baseline model by 0.98 percentage points. The YOLOv5s-C3CBAM and YOLOv5s-C3CA models closely follow. The YOLOv5s-CBAM model exhibits the highest precision, surpassing all other modified models. The trial findings suggest that the proposed enhancement models significantly improve the overall performance of the detection model. The main goal of the study is to increase the mean average precision (mAP_0.5) of the baseline model to enhance the detection of compound illnesses in maize leaves. The YOLOv5s-C3CBAM model achieves a mean average precision (mAP) of 79%, outperforming all other updated models and showing a 2.9 percentage point improvement compared to the original model. The second-best model also outperforms the original model in recall and F1 score, indicating that the YOLOv5s-C3CBAM model surpasses the performance of the other examined models. Table 1 illustrates the details of the training models.

Precision Evaluation of Enhanced Models for Various Leaf Diseases:

To assess the upgraded model's ability to accurately identify different diseases, we examined the confusion matrix of the YOLOv5s-C3CBAM model's classification predictions Healthy leaves demonstrate superior identification capability, achieving an accuracy rate of 96% and a recall rate of 93%. The accuracy rates for diagnosing blight, rust disease, and grey leaf spot disease were 59%, 74%, and 89%, respectively. Confusion in the matrix about blight and grey leaf spot disease can be attributed to the similarity in color between the lesions associated with these two ailments. Additionally, tiny lesions caused by grey leaf spot disease might impede the extraction of their contour properties, resulting in inaccurate categorization by the model. Therefore, efforts are needed to enhance the model's competency in classifying blight and grey leaf spot disease. The YOLOv5s-C3CBAM model was used to integrate three example compound maize leaf disease pictures, including blight and rust disease, blight and grey leaf spot disease, and blight, rust disease, and grey leaf spot disease. The results unequivocally demonstrate that YOLOv5s-C3CBAM possesses exceptional disease detection capabilities for grey leaf spots and blight. Diagnosing rust disease becomes challenging when the lesions are tiny and distributed sporadically.

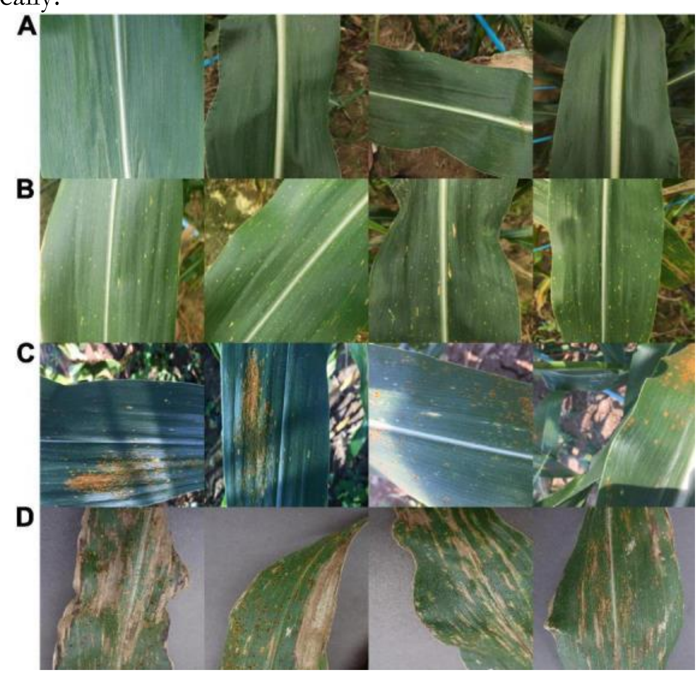

Figure 2: Showing four categories of samples of maize. (A) H. (B) SCLB. (C) SR. (D) GLS [16].

Comparison of Model Performance:

The study compares two modified iterations of the original YOLOv5 model, YOLOv5m and YOLOv7-tiny, and the commonly employed two-stage object detection model, Faster R-CNN, with the recently developed YOLOv5s-C3CBAM mode. The YOLOv5s-C3CBAM model outperforms other models with the best performance of 83% on the mAP_0.5 evaluation metric. Faster R-CNN exhibits the poorest performance, attributed to difficulties in distinguishing tiny lesions from diseases like rust and grey leaf spots. YOLOv5m and YOLOv7-tiny, lacking the attention mechanism module, also show lower performance compared to YOLOv5s-C3CBAM.

The YOLOv5s-C3CBAM model achieves the highest recall rate of 73%, enhancing feature extraction capacity through the CBAM attention mechanism. YOLOv5m, YOLOv7-

tiny, and Faster R-CNN have lower recall rates. With an F1 score of 82%, the YOLOv5s-C3CBAM model attains the greatest accuracy assessment parameter. The YOLOv7-tiny version surpasses the previous three versions in terms of frames per second (FPS) due to its efficient network design. The YOLOv5s-C3CBAM model effectively balances detection speed and model size, resulting in a frame rate of 53 FPS. YOLOv5s-C3CBAM is selected as the optimal detection model for the experiment.

Framework Selection and Conclusion

Replacing YOLOv4 with YOLOv5 allows for smaller models, faster processing, and reduced memory consumption. YOLOv5s from the YOLOv5 series are recommended as the fundamental model for accurately detecting small-scale maize leaf diseases in real time on low-resource portable devices. This model achieves a harmonious balance between rapid processing, accurate detection, and sufficient processing capability.

Table 2: Model Comparison

Model	mAP_0.5 (%)	Recall Rate (%)	F1 Score (%)	Precision (%)	FPS
YOLOv5s-C3CBAM	83	73	82	85	53
YOLOv5s	78	68	76	80	50
YOLOv7-tiny	75	60	70	78	60
Faster R-CNN	70	55	65	75	40

Image Generation Experiment and New Concepts:

In the image generation experiment, lesions did not follow expected patterns, such as blight appearing perpendicular to leaf texture rather than following veins. Training set photographs with significant rotation angles were removed, and a screening procedure compared generated images with real ones. Early attempts at grey spot disease produced lower-quality images, which improved with more model iterations. The study introduces two innovations: using Cycle GAN to create compound sickness images and enhancing the model with attention processes to focus on specific lesion targets, improving disease detection accuracy.

Table 3: Disease Identification by YOLOv5s-C3CBAM Model

Disease	Accuracy Rate (%)
Healthy leaves	96
Blight	59
Rust disease	74
Grey leaf spot	89

Limitations and Future Directions:

Limitations include a focus on three common maize leaf diseases, with the potential for more accurate classification in future studies. Extensive disease annotation and ongoing training are needed for model improvement. Future research aims to build a comprehensive dataset for in-depth examination of maize leaf diseases.

Choosing Pixel Density:

A 256x256 pixel resolution was chosen, balancing workload, training rates, and the model's learning capabilities. This size effectively captures detailed lesion information while considering the limitations of higher resolutions and lower pixel counts. The study's main objective is to identify diseases in individual maize leaves, including compound disorders, addressing challenges with deep learning and data availability by using Cycle GAN for synthetic data. The YOLOv5-C3CBAM model outperforms others, enhancing accuracy and generalization in real-world scenarios. The study contributes to accurate chemical diagnosis of crop diseases in agricultural settings and serves as a reference for future research.

Image Generation Experiment and New Concepts:

In the image generation experiment, lesions did not follow expected patterns, such as blight appearing perpendicular to leaf texture rather than following veins. Training set photographs with significant rotation angles were removed, and a screening procedure compared generated images with real ones. Early attempts at grey spot disease produced lower-quality images, which improved with more model iterations. The study introduces two innovations: using Cycle GAN to create compound sickness images and enhancing the model with attention processes to focus on specific lesion targets, improving disease detection accuracy.

Limitations and Future Directions:

Limitations include a focus on three common maize leaf diseases, with the potential for more accurate classification in future studies. Extensive disease annotation and ongoing training are needed for model improvement. Future research aims to build a comprehensive dataset for in-depth examination of maize leaf diseases.

Choosing Pixel Density:

A 256x256 pixel resolution was chosen, balancing workload, training rates, and the model's learning capabilities. This size effectively captures detailed lesion information while considering the limitations of higher resolutions and lower pixel counts.

Conclusion:

The study's main objective is to identify diseases in individual maize leaves, including compound disorders, addressing challenges with deep learning and data availability by using Cycle GAN for synthetic data. The YOLOv5-C3CBAM model outperforms others, enhancing accuracy and generalization in real-world scenarios. The study contributes to accurate chemical diagnosis of crop diseases in agricultural settings and serves as a reference for future research.

References:

[1] Y. Zhang, S. Wa, Y. Liu, X. Zhou, P. Sun, and Q. Ma, "High-accuracy detection of maize leaf diseases cnn based on multi-pathway activation function module," *Remote Sens.*, vol. 13, no. 21, Nov. 2021, doi: 10.3390/RS13214218.

[2] Y. Zhang, H. Wang, R. Xu, X. Yang, Y. Wang, and Y. Liu, "High-Precision Seedling Detection Model Based on Multi-Activation Layer and Depth-Separable Convolution Using Images Acquired by Drones," *Drones*, vol. 6, no. 6, Jun. 2022, doi: 10.3390/DRONES6060152.

[3] Z. Hui, J. Li, X. Wang, and X. Gao, "Image Fine-grained Inpainting," Feb. 2020, Accessed: Oct. 01, 2023. [Online]. Available: https://arxiv.org/abs/2002.02609v2

[4] W. Chen, B. Qi, X. Liu, H. Li, X. Hao, and Y. Peng, "Temperature-Robust Learned Image Recovery for Shallow-Designed Imaging Systems," *Adv. Intell. Syst.*, vol. 4, no. 10, p. 2200149, Oct. 2022, doi: 10.1002/AISY.202200149.

[5] "High-Accuracy Maize Disease Detection Based on Attention Generative Adversarial Network and Few-Shot Learning - PMC." Accessed: Feb. 17, 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10490187/

[6] Q. Xie *et al.*, "Crop height estimation of corn from multi-year radarsat-2 polarimetric observables using machine learning," *Remote Sens.*, vol. 13, no. 3, pp. 1–19, Feb. 2021, doi: 10.3390/RS13030392.

[7] H. Lee, J. Wang, and B. Leblon, "Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn," *Remote Sens.*, vol. 12, no. 13, Jul. 2020, doi: 10.3390/RS12132071.

[8] J. Yu, J. Wang, and B. Leblon, "Evaluation of soil properties, topographic metrics, plant height, and unmanned aerial vehicle multispectral imagery using machine learning methods to estimate canopy nitrogen weight in corn," *Remote Sens.*, vol. 13, no. 16, Aug. 2021, doi: 10.3390/RS13163105.

[9] Y. Zhang, S. Wa, L. Zhang, and C. Lv, "Automatic Plant Disease Detection Based on Tranvolution Detection Network With GAN Modules Using Leaf Images," *Front. Plant Sci.*, vol. 13, May 2022, doi: 10.3389/FPLS.2022.875693.
[10] Y. Zhang, M. Li, X. Ma, X. Wu, and Y. Wang, "High-Precision Wheat Head Detection Model Based on One-Stage Network and GAN Model," *Front. Plant Sci.*, vol. 13, Jun. 2022, doi: 10.3389/FPLS.2022.787852.
[11] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 3313–3332, Apr. 2023, doi: 10.1109/TKDE.2021.3130191.
[12] H. Guo *et al.*, "Sample Expansion and Classification Model of Maize Leaf Diseases Based on the Self-Attention CycleGAN," *Sustain.*, vol. 15, no. 18, Sep. 2023, doi: 10.3390/SU151813420.
[13] Y. Wang, J. Wu, P. Lan, F. Li, C. Ge, and F. Sun, "Apple disease identification using improved Faster R-CNN," *J. For. Eng.*, vol. 7, no. 1, pp. 153–159, Jan. 2022, doi: 10.13360/J.ISSN.2096-1359.202104028.
[14] L. Huang, Z. Zhou, Y. Guo, and Y. Wang, "A stability-enhanced CycleGAN for effective domain transformation of unpaired ultrasound images," *Biomed. Signal Process. Control*, vol. 77, Aug. 2022, doi: 10.1016/J.BSPC.2022.103831.
[15] E. T. A. Albert, N. H. Bille, and N. M. E. Leonard, "Improvement of plant disease classification accuracy with generative model-synthesized training datasets," *Res. Biotechnol.*, pp. 1–11, Feb. 2023, doi: 10.25081/RIB.2023.V14.8214.
[16] P. Dong, K. Li, M. Wang, F. Li, W. Guo, and H. Si, "Maize Leaf Compound Disease Recognition Based on Attention Mechanism," *Agric. 2024, Vol. 14, Page 74*, vol. 14, no. 1, p. 74, Dec. 2023, doi: 10.3390/AGRICULTURE14010074.
[17] "View of Evaluating Spatio-Temporal Decline to Agriculture through Satellite Imagery from 2010-2022." Accessed: Feb. 22, 2024. [Online]. Available: https://journal.50sea.com/index.php/IJASD/article/view/471/971
[18] Y. Song *et al.*, "High-Accuracy Maize Disease Detection Based on Attention Generative Adversarial Network and Few-Shot Learning," *Plants*, vol. 12, no. 17, Sep. 2023, doi: 10.3390/PLANTS12173105.
[19] W. Xu, W. Li, L. Wang, and M. F. Pompelli, "Enhancing Corn Pest and Disease Recognition through Deep Learning: A Comprehensive Analysis," *Agronomy*, vol. 13, no. 9, Sep. 2023, doi: 10.3390/AGRONOMY13092242.
[20] D. Theckedath and R. R. Sedamkar, "Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks," *SN Comput. Sci.*, vol. 1, no. 2, Mar. 2020, doi: 10.1007/S42979-020-0114-9.
[21] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, Aug. 2020, doi: 10.1109/TPAMI.2019.2913372.
[22] F. Rajeena P. P, A. S. U, M. A. Moustafa, and M. A. S. Ali, "Detecting Plant Disease in Corn Leaf Using EfficientNet Architecture—An Analytical Approach," *Electron.*, vol. 12, no. 8, Apr. 2023, doi: 10.3390/ELECTRONICS12081938.
[23] W. Ding and L. Zhang, "Building Detection in Remote Sensing Image Based on Improved YOLOV5," *Proc. - 2021 17th Int. Conf. Comput. Intell. Secur. CIS 2021*, pp. 133–136, 2021, doi: 10.1109/CIS54983.2021.00036.
[24] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11211 LNCS, pp. 3–19, 2018, doi: 10.1007/978-3-030-01234-2_1.
[25] H. Wang, J. Feng, and H. Yin, "Improved Method for Apple Fruit Target Detection Based on YOLOv5s," *Agric.*, vol. 13, no. 11, Nov. 2023, doi:

10.3390/AGRICULTURE13112167.

[26] "Agriculture | Free Full-Text | Maize Leaf Compound Disease Recognition Based on Attention Mechanism." Accessed: Feb. 17, 2024. [Online]. Available: https://www.mdpi.com/2077-0472/14/1/74

[27] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, no. September, p. 1419, Sep. 2016, doi: 10.3389/FPLS.2016.01419/BIBTEX.

[28] Y. Hu, G. Liu, Z. Chen, J. Liu, and J. Guo, "Lightweight One-Stage Maize Leaf Disease Detection Model with Knowledge Distillation," *Agric.*, vol. 13, no. 9, Sep. 2023, doi: 10.3390/AGRICULTURE13091664.

[29] W. Bao, X. Huang, G. Hu, and D. Liang, "Identification of maize leaf diseases using improved convolutional neural network," *Nongye Gongcheng Xuebao/Transactions Chinese Soc. Agric. Eng.*, vol. 37, no. 6, pp. 160–167, Mar. 2021, doi: 10.11975/J.ISSN.1002-6819.2021.06.020.

[30] Y. Qiao *et al.*, "A Counting Method of Red Jujube Based on Improved YOLOv5s," *Agric.*, vol. 12, no. 12, Dec. 2022, doi: 10.3390/AGRICULTURE12122071.

[31] S. Ghaffarian, J. Valente, M. Van Der Voort, and B. Tekinerdogan, "Effect of Attention Mechanism in Deep Learning-Based Remote Sensing Image Processing: A Systematic Literature Review," *Remote Sens. 2021, Vol. 13, Page 2965*, vol. 13, no. 15, p. 2965, Jul. 2021, doi: 10.3390/RS13152965.

[32] Z. Li, W. Tao, J. Liu, F. Zhu, G. Du, and G. Ji, "Tomato Leaf Disease Recognition via Optimizing Deep Learning Methods Considering Global Pixel Value Distribution," *Horticulturae*, vol. 9, no. 9, Sep. 2023, doi: 10.3390/HORTICULTURAE9091034.