iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81180

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Self-Supervised Learning for Agricultural Image Classification: Detecting Plant Diseases and Pests in Cotton

Narendra Narsingh Chauhan[1], Dr. Sheetal Rathi[2]
[1]*Department of Computer Engineering, Thakur College of Engineering & Technology, India, Mumbai*
[2]*Department of Computer Engineering, Thakur College Of Engineering & Technology, Kandivali, Mumbai, India*

Abstract: The early detection of plant diseases and pest infestations is essential for improving agricultural productivity and ensuring crop health. Cotton, as a major commercial crop, is highly susceptible to a wide range of diseases and pests that can significantly affect yield and quality if not identified at an early stage. Conventional deep learning approaches for plant disease classification rely heavily on large labeled datasets, which are often difficult and costly to obtain in real-world agricultural environments. In this paper, a self-supervised learning-based framework is proposed for agricultural image classification, focusing on the detection of plant diseases and pests in cotton. The proposed approach begins with dataset curation and cleaning using an automated image inspection technique to remove noisy, blurry, and duplicate samples, thereby improving data quality. Data augmentation techniques are then applied to enhance model generalization. A contrastive learning strategy inspired by SimCLR is employed to learn robust feature representations from unlabeled images using a ResNet18 encoder. The extracted features are further refined using Principal Component Analysis (PCA) for dimensionality reduction and K-Means clustering to improve feature separability. Finally, a pretrained VGG16 model is fine-tuned using labeled data for classification.

The integration of self-supervised learning with feature refinement and clustering enhances representation quality while reducing dependency on labeled datasets. The proposed framework is scalable, efficient, and suitable for real-world agricultural applications, including automated crop monitoring and intelligent farming systems.

Keywords: Self-Supervised Learning, Cotton Disease Detection, Pest Detection, Contrastive Learning, Principal Component Analysis (PCA), K-Means Clustering, Deep Learning, Agricultural Image Classification, VGG16.

I. INTRODUCTION

Agriculture plays a fundamental role in sustaining global food security and supporting economic development, particularly in developing countries. Among various crops, cotton is one of the most important commercial crops due to its extensive use in the textile industry and its significant contribution to national economies. However, cotton plants are highly vulnerable to a wide range of diseases and pest infestations, such as Alternaria leaf spot, Fusarium wilt, whiteflies, and thrips, which can severely impact crop yield and quality if not detected at an early stage [21]–[24]. According to recent agricultural reports, plant diseases and pests are responsible for substantial losses in cotton production worldwide, highlighting the need for efficient monitoring and early detection systems. Traditionally, the detection of plant diseases has relied on manual inspection by agricultural experts. While this method can be accurate, it is time-consuming, labor-intensive, and not scalable for large agricultural fields. Moreover, the effectiveness of manual inspection depends heavily on the expertise of the observer, making it prone to human error and inconsistency. With the rapid advancement of artificial intelligence, particularly deep learning, automated image-based disease detection systems have emerged as a promising solution. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in extracting hierarchical features from images and have been widely used for plant disease classification tasks [1], [2].

Several studies have applied deep learning techniques for plant disease detection. For instance, Mohanty et al. [1] demonstrated the effectiveness of deep CNNs for large-scale plant disease classification, while Ferentinos [2] explored multiple deep learning architectures for crop disease diagnosis. Similarly, Sladojevic et al. [4] proposed a deep neural network-based system for leaf image classification, achieving promising results. Comparative studies have also shown that fine-tuned deep learning models such as VGG16 and ResNet outperform traditional machine learning approaches in agricultural image analysis [5], [15]. Despite these advancements, most of these methods rely heavily on large labeled datasets, which are often difficult and expensive to obtain in real-world agricultural environments.

One of the major challenges in agricultural image analysis is the scarcity of labeled data. Annotating plant disease images requires domain expertise and is time-intensive. Furthermore, real-world datasets often contain noisy, blurry, and redundant images, which can degrade model performance [6], [17]. Data augmentation techniques are commonly used to address dataset limitations; however, they do not fully eliminate the dependency on labeled data [25]. These challenges necessitate the development of alternative learning approaches that can effectively utilize unlabeled data.

Self-supervised learning (SSL) has recently emerged as a powerful paradigm for learning meaningful feature representations from unlabeled data. Unlike supervised learning, SSL leverages pretext tasks to learn intrinsic patterns within the data without requiring manual annotations. Contrastive learning-based approaches such as SimCLR [8] and MoCo [9] have shown significant success in visual representation learning by maximizing the similarity between augmented views of the same image while minimizing similarity with other images. Similarly, methods such as BYOL [10] and SimSiam [26] have further advanced the field by eliminating the need for negative samples. These approaches enable the extraction of robust and generalized features, making them particularly suitable for agricultural datasets where labeled data is limited but unlabeled data is abundant [27].

In addition to feature learning, dimensionality reduction and clustering techniques play an important role in improving data representation and interpretability. Principal Component Analysis (PCA) is widely used to reduce feature dimensionality while preserving important variance in the data [3]. Clustering algorithms such as K-Means help in grouping similar data points and identifying underlying patterns in feature space [4]. These techniques can enhance feature separability and improve downstream classification performance when combined with deep learning models.

Motivated by these challenges and advancements, this paper proposes a hybrid framework for detecting plant diseases and pests in cotton using self-supervised learning. The proposed approach begins with dataset curation and cleaning to remove low-quality images, followed by data augmentation to improve generalization. A SimCLR-based contrastive learning model is then employed to learn feature representations from unlabeled images. The extracted features are further refined using PCA for dimensionality reduction and K-Means clustering for analyzing feature distribution and improving separability. Finally, a pretrained VGG16 model is fine-tuned using labeled data for classification.

II. RELATED THEORY

Numerous biotic stressors can directly impact leaf structure, physiological functions, and total crop output in cotton plants. Understanding the biological and visual traits of the main cotton leaf disorders is crucial because this work focuses on image-based disease categorization.

A. Bacterial Blight

Xanthomonas citri pv. malvacearum is the causative agent of bacterial blight in cotton. It mostly affects bolls, stems, and leaves. Early on, the leaf surface develops tiny, wet sores. These lesions get larger and turn dark brown or black as the infection worsens, frequently forming angular patterns confined by leaf veins.Bacterial blight causes uneven dark patches, edge necrosis, and, in extreme situations, leaf ripping. The leaf's visual structure is considerably altered by these texture and color differences, which enable image-based analysis to recognize it. The disease can cause defoliation and a significant decrease in yield if left untreated.

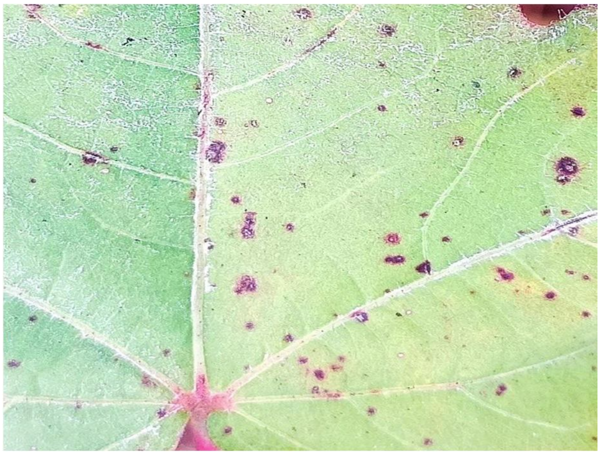

Fig. 1. Bacterial Blight

B. Cotton Leaf Curl Virus (CLCuV)

The Cotton Leaf Curl Virus is a serious viral disease that mostly spreads through whiteflies. The infection causes leaves to curl up or down, veins to thicken, and plants to grow more slowly.From the point of view of image analysis, the leaf curl virus changes the structure of the leaf instead of just changing its color.

The edges of the leaves become twisted, and the lamina may grow in strange ways. Yellowing may also happen in later stages. These changes in shape make geometric patterns that are easy to tell apart and can be captured by deep feature extraction methods.

Fig. 2. Curl Virus

C. Healthy Cotton Leaf

A healthy cotton leaf should be a uniform green color with clear veins and smooth edges. The surface texture stays the same, and there are no visible lesions, color changes, or changes in shape. From a computational point of view, healthy leaves give us baseline structural and color patterns that we can use to tell the difference between diseased samples. Changes from this normal shape often mean that an infection or pest damage is starting to happen.

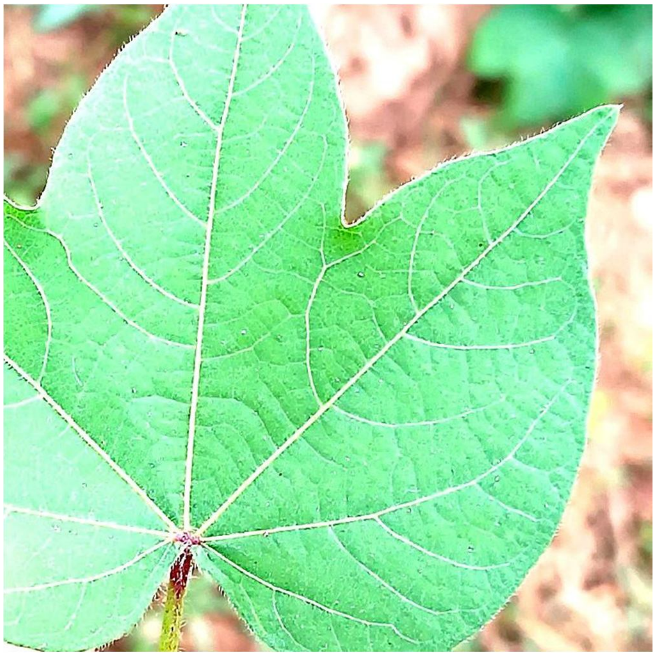

Fig. 3. Healthy Leaf

III. LITERATURE SURVEY

The application of deep learning in agriculture, particularly for plant disease detection, has gained significant attention in recent years. Early studies demonstrated the effectiveness of Convolutional Neural Networks (CNNs) in image-based plant disease classification. Mohanty et al. [1] utilized deep CNN architectures for large-scale plant disease detection and reported high classification performance across multiple crop species. Similarly, Ferentinos [2] evaluated various deep learning models and highlighted their capability in accurately diagnosing plant diseases under controlled conditions. Sladojevic et al. [4] further proposed a deep neural network-based system for leaf image classification, demonstrating the potential of automated disease recognition systems.

Subsequent research focused on improving model performance and generalization. Kamilaris and Prenafeta-Boldú [3] provided a comprehensive survey of deep learning applications in agriculture, emphasizing the advantages of CNN-based approaches over traditional machine learning techniques. Too et al. [5] conducted a comparative study on fine-tuning deep learning models such as VGG16 and ResNet, concluding that transfer learning significantly improves classification accuracy, especially when labeled data is limited. Additional studies have also explored the challenges associated with plant disease recognition, including variations in lighting, background noise, and image quality [6].

Despite the success of supervised deep learning models, their dependence on large labeled datasets remains a major limitation. In agricultural domains, labeling data requires expert knowledge and is often time-consuming. To address this issue, researchers have explored data augmentation techniques to artificially increase dataset size and diversity. Shorten and Khoshgoftaar [25] provided a detailed survey on image augmentation methods, demonstrating their effectiveness in improving model generalization. However, augmentation alone cannot fully overcome the limitations of insufficient labeled data.

To reduce dependency on labeled datasets, self-supervised learning (SSL) has emerged as a promising alternative. SSL methods enable models to learn meaningful feature representations from unlabeled data by leveraging pretext tasks. Chen et al. [8] introduced SimCLR, a contrastive learning framework that learns visual representations by maximizing agreement between augmented views of the same image. Similarly, He et al. [9] proposed Momentum Contrast (MoCo), which utilizes a dynamic dictionary for contrastive learning. Grill et al. [10] introduced BYOL, which eliminates the need for negative samples, further advancing the SSL paradigm. Chen and He [26] proposed SimSiam, a simplified Siamese network for representation learning. These methods have demonstrated state-of-the-art performance in various computer vision tasks.

In addition to SSL, representation learning techniques have been extensively studied to improve feature extraction. Bengio et al. [12] provided a foundational overview of representation learning, highlighting its importance in capturing underlying data patterns. Advanced architectures such as Vision Transformers [11] and EfficientNet [13] have also been proposed to improve model scalability and performance. However, CNN-based architectures such as VGG16 [14] and ResNet [15] remain widely used due to their effectiveness and simplicity in image classification tasks.

Dimensionality reduction and clustering techniques are also important for analyzing high-dimensional feature spaces. Principal Component Analysis (PCA) is commonly used to reduce feature dimensionality while preserving important variance [3]. Clustering algorithms such as K-Means help group similar data points and identify patterns in feature space. These methods are particularly useful when combined with deep learning features to improve separability and interpretability. In the agricultural domain, several studies have explored the use of machine learning and image processing techniques for plant disease detection. Singh et al. [7] proposed a machine learning-based approach for leaf disease detection, while Saleem et al. [18] compared different CNN architectures for plant disease classification. Additionally, domain-specific studies have highlighted the impact of pests such as whiteflies and thrips on cotton crops [21], [22], as well as diseases like Alternaria leaf spot and Fusarium wilt [23], [24]. Although significant progress has been made, existing methods still face challenges related to data quality, limited labeled data, and feature representation. Most studies focus either on supervised learning or self-supervised learning independently, with limited integration of feature refinement techniques such as PCA and clustering. To address these gaps, the present work proposes a hybrid framework that combines dataset cleaning, self-supervised learning, dimensionality reduction, clustering, and transfer learning. This integrated approach aims to improve feature representation, reduce dependency on labeled data, and enhance classification performance for cotton disease and pest detection.

IV. CONCEPTUAL FRAMEWORK

The proposed conceptual framework offers a thorough and organized method for finding pests and diseases in cotton plants. It uses a mix of self-supervised learning, feature engineering, and deep learning-based classification. The framework is meant to solve some of the biggest problems in analyzing agricultural images, such as noisy datasets, not enough labeled data, and too many features.

The system works in a series of steps, with each step helping to improve data quality, get useful representations, and make classification work better. The combination of self-supervised learning with dimensionality reduction and clustering is the main new idea.

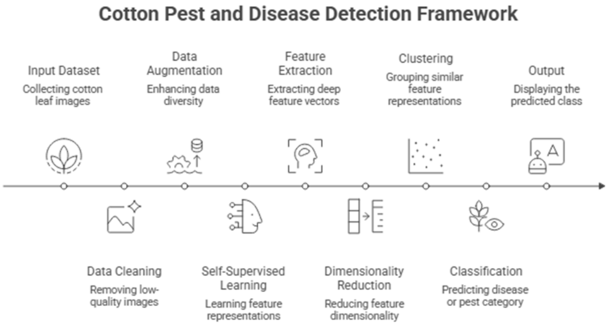

Fig. 4. Conceptual Framework

A. *Input Dataset*

The framework begins with a dataset consisting of cotton leaf images representing various disease conditions and pest infestations. These images are collected under different environmental conditions, making the dataset diverse but also prone to noise and inconsistencies.

B. *Data Cleaning*

Real-world agricultural datasets often contain low-quality images that negatively impact model performance. Therefore, an automated data cleaning process is applied to improve dataset quality.

The following issues are identified and removed:

- Blurry images
- Overexposed or underexposed images
- Low-information samples
- Duplicate and near-duplicate images

This step ensures that only high-quality and informative images are used for training, leading to better feature learning.

C. *Data Augmentation*

To improve generalization and prevent overfitting, data augmentation techniques are applied. These include:

- Random rotation
- Horizontal flipping
- Image resizing

Augmentation simulates real-world variations such as lighting conditions, leaf orientation, and background noise, making the model more robust.

D. *Self-Supervised Learning (SimCLR)*

A self-supervised learning approach is used to learn feature representations from unlabeled data. The framework employs a contrastive learning method inspired by SimCLR.

Two augmented views of the same image are generated and passed through a shared encoder (ResNet18). The model is trained to:

- Maximize similarity between representations of the same image
- Minimize similarity with other images

This enables the model to learn meaningful visual features such as texture patterns, color variations, and disease characteristics without requiring labeled data.

E. Feature Extraction

Once the self-supervised model is trained, the encoder is used to extract deep feature vectors from images. These features represent high-level semantic information and serve as input for further processing.

METRIC	VALUE (%)
ACCUARCY	94.185
RECALL	94.279
PRECISION	94.158
F1 SCORE	94.217

F. Dimensionality Reduction (PCA)

The extracted features are high-dimensional and may contain redundant information. Principal Component Analysis (PCA) is applied to:

- Reduce feature dimensionality
- Remove noise and redundancy
- Improve computational efficiency

This step ensures that only the most significant features are retained.

G. Clustering (K-Means)

K-Means clustering is applied to group similar feature representations. This helps in:

- Understanding feature distribution
- Improving class separability
- Identifying hidden patterns in data

The optimal number of clusters is determined using the elbow method.

H. Classification (VGG16)

In the final stage, a pretrained VGG16 model is fine-tuned for classification. The model uses labeled data to:

- Learn disease-specific patterns
- Perform final classification

Transfer learning improves performance by leveraging pretrained weights.

I. Output

The system outputs the predicted class of the cotton leaf image, identifying the disease or pest category. This enables automated and accurate diagnosis.

V. PROPOSED SYSTEM

This study proposes a self-supervised learning-based framework for the classification of cotton leaf diseases and pests to overcome the limitations of purely supervised approaches. Self-supervised learning has become a useful way to learn representations without needing to add labels by hand [7], [8].The proposed system consists of two major stages: representation learning using unlabeled images and supervised fine-tuning using a smaller labeled dataset.

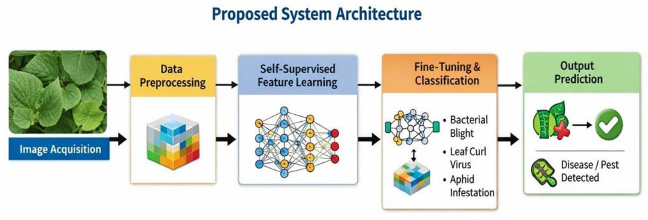

Fig. 5. Proposed System

A. Self-Supervised Learning

The first step is to use unlabeled images of cotton leaves to train a deep neural network backbone. The model learns through a contrastive learning mechanism that is similar to the ones used in SimCLR and MoCo [7], [9] instead of using class labels.Two augmented views are made for each image by changing things like cropping, rotating, and color jittering. The model learns to make embeddings of the same image as similar as possible while making embeddings of different images as different as possible. This objective of contrast helps the network learn about intrinsic structural features like:

- Texture variations caused by fungal or bacterial infections
- Leaf curling patterns
- Spot distribution and lesion boundaries
- Color discoloration caused by pest damage

Because this stage does not rely on manual labeling, the system can utilize large volumes of raw field images.

B. Feature Embedding Extraction and Analysis

After training the model on its own, feature embeddings are taken from the backbone network. A high-dimensional feature vector represents each image of a cotton leaf.To evaluate the quality of the learned representations, dimensionality reduction techniques such as t-distributed Stochastic Neighbor Embedding (t-SNE) are applied for visualization [10].

If the learned features are significant, images associated with analogous disease categories inherently aggregate in the embedding space.Also, clustering algorithms like K-Means are used to look at how features group together. The clustering performance shows how well the model has learned disease-specific patterns without using labels.

C. Supervised Fine-Tuning

In the second stage, a smaller labeled dataset is used to fine-tune the pretrained backbone. An entirely connected classification layer is included to link the learned embeddings to disease and pest groups. The model doesn't need as many labeled samples to work well because it has already learned general visual features during the self-supervised stage. This makes annotation less important than starting from scratch with full supervision. During the classification stage, cross-entropy loss is used, and metrics like accuracy, precision, recall, and F1-score are used to measure performance.

VI. METHODOLOGY

Data preparation, self-supervised representation learning, embedding analysis, and supervised fine-tuning are the four primary phases of the suggested methodology.To guarantee consistent input proportions, cotton leaf photos gathered from field settings are first shrunk and normalized. To increase robustness against real-field variations, data augmentation techniques like rotation, flipping, and random cropping are used [1]. These actions lessen overfitting and improve generalization.A contrastive learning approach is used in the self-supervised stage to learn visual representations without the use of labels [2], [3]. A shared backbone network is used to create two augmented views for every image. The model is trained to minimize similarity between different images and maximize similarity between embeddings of the same image. This makes it possible for the network to identify structural patterns like lesions. This makes it possible for the network to record structural patterns including leaf deformation brought on by pests and diseases, lesion texture, and color distortion. Feature embeddings are taken from the backbone network following pretraining. Dimensionality reduction using t-SNE is used for visualization in order to assess representation quality [4].
Additionally, grouping behavior in the embedding space is examined using K-Means clustering.Lastly, a tagged cotton leaf dataset is used to refine the pretrained model. Cross-entropy loss is employed for optimization, and a classification layer is added. Accuracy, precision, recall, and F1-score are used to assess the model's performance.

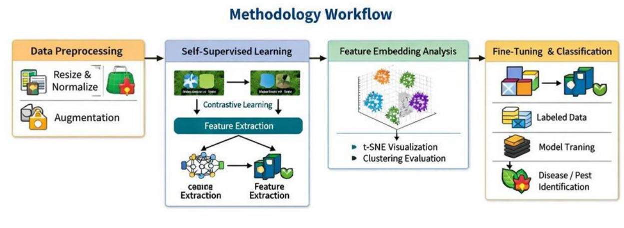

Fig. 6. Methodology Workflow

IMPLEMENTATION & RESULTS.

A. Dataset Description and Preparation

The dataset used in this study consists of approximately **6600 cotton leaf images**, representing multiple disease and pest categories. To improve the robustness and generalization of the model, an augmented dataset was generated using various image transformation techniques.

Data augmentation techniques applied include:

- Rotation
- Horizontal flipping
- Resizing
- Random transformations

These techniques increase the diversity of the dataset and help the model learn invariant features under different environmental conditions such as lighting, orientation, and background variations.

After preprocessing and cleaning, the dataset is divided as follows:

- Training set: 80% (~5280 images)
- Testing set: 20% (~1320 images)

This ensures a balanced evaluation of the model on unseen data.

B. Training Configuration

The model was implemented using the PyTorch framework. The training process consists of two major stages:

1. Self-Supervised Learning Stage
    - Encoder: ResNet18
    - Objective: Learn feature representations using contrastive learning
    - Loss Function: Contrastive Loss
2. Supervised Fine-Tuning Stage
    - Model: VGG16 (pretrained)
    - Loss Function: Cross-Entropy Loss
    - Optimizer: Adam
    - Learning Rate: 0.0001
    - Batch Size: 32
    - Epochs: 10

C. Loss Function (Cross-Entropy)

The classification model is trained using Cross-Entropy Loss:

$$\mathbf{Loss = - \Sigma\, yi\, log(\hat{y}i)}$$

where:

- yi = true label
- yi^ = predicted probability
- C = number of classes

D. Performance Evaluation Metrics

The model performance is evaluated using the following metrics:

Accuracy - Out of all the predictions, accuracy indicates how many of them the model was correct. When one class is more dominant than the other, it might be deceptive even though it provides an idea of overall performance. For instance, a model may have high accuracy but miss crucial information about other classes even if it accurately predicts the majority class most of the time.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

Precision - The precision of the model's positive predictions is its main focus. It indicates the proportion of "positive" forecasts that came true. When detecting fraud or spam emails, for example, it is crucial to reduce false positives. The formula of precision is:

$$\text{Precision} = TP / (TP + FP)$$

Recall - The model's recall indicates how well it predicts positives. It shows the proportion of true positives detected out of all the actual positive instances. When missing positive cases has serious repercussions, like in medicine, high recall is crucial.

$$\text{Recall} = TP / (TP + FN)$$

F1 Score - F1-score combines precision and recall into a single metric to balance their trade-off. It gives a more accurate picture of a model's overall performance, especially when dealing with unbalanced datasets. It is helpful when both false positives and false negatives are important though it assumes precision and recall are equally important but in some situations one might matter more than the other.

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

E. Experimental Results

The proposed model demonstrates strong classification performance on the augmented cotton dataset.

F. Confusion Matrix Analysis

The confusion matrix provides a detailed evaluation of the classification performance across different classes.

- True Positives (TP): Correct predictions
- False Positives (FP): Incorrect predictions
- False Negatives (FN): Missed detections
- True Negatives (TN): Correct rejections

The confusion matrix shows strong diagonal values, indicating that the majority of samples are correctly classified. The low off-diagonal values suggest minimal confusion between different disease and pest classes.

G. Discussion

The experimental results demonstrate that the proposed hybrid framework is highly effective for cotton disease and pest detection. The key findings are:

- The use of augmented dataset (~6600 images) significantly improves model generalization
- Self-supervised learning enhances feature extraction without requiring extensive labeled data
- PCA reduces feature redundancy and improves efficiency
- K-Means clustering improves feature separability
- Transfer learning with VGG16 results in high classification performance

Overall, the integration of these techniques leads to a robust and scalable model suitable for real-world agricultural applications.

VII. CONCLUSION

This research proposes a hybrid system that combines deep learning-based categorization, feature refinement, and self-supervised learning to detect plant diseases and pests in cotton. This study's main goal was to address important issues in agricultural image analysis, such as large feature dimensionality, noisy datasets, and a lack of labeled data.

In order to increase data quality and diversity, the suggested method uses a thorough pipeline that starts with dataset cleaning and augmentation. To extract useful feature representations from unlabeled data, a contrastive learning-based self-supervised learning model is used. Principal Component Analysis (PCA) is used to reduce dimensionality in these features, while K-Means clustering is used to increase feature separability. Ultimately, a pretrained VGG16 model is optimized for precise cotton disease categorization.

The experimental findings show that the suggested framework achieves high classification performance, demonstrating its usefulness in practical agricultural applications. While feature refinement approaches improve the quality of learnt representations, self-supervised learning greatly lowers reliance on huge labeled datasets. Furthermore, the application of transfer learning enhances accuracy and speeds up convergence. All things considered, the suggested approach offers a scalable and effective way to identify plant diseases automatically. Early diagnosis and prompt action are made possible by its effective application in smart farming systems, precision agriculture, and real-time crop monitoring. This study demonstrates how self-supervised learning and conventional deep learning methods can be combined to enhance agricultural picture processing.

VIII. ACKNOWLEDGEMENTS

The Department of Computer Engineering provided the infrastructure and technical assistance required to conduct this research, for which the authors are deeply grateful. We also thank our project supervisor for his advice and insightful recommendations, which were crucial to finishing this study

.We would like to express our gratitude to the field workers and agricultural specialists who helped gather and validate the cotton leaf picture data. Understanding disease traits and pest symptoms was made possible by their subject expertise

.Lastly, we would like to express our gratitude to our author, institution, and guide for supporting research endeavors and offering an academic setting that enabled the effective completion of this work.

REFERENCES

[1] M. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, pp. 1–10, 2016.
[2] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, vol. 145, pp. 311–318, 2018.
[3] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, vol. 147, pp. 70–90, 2018.
[4] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," Computational Intelligence and Neuroscience, vol. 2016, pp. 1–11, 2016.
[5] J. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.
[6] Y. Liu, J. Wang, C. Long, and H. Li, "Challenges and solutions for plant disease recognition using deep learning," IEEE Access, vol. 8, pp. 186487–186503, 2020.
[7] R. Singh, S. Srivastava, and R. Singh, "Detection of plant leaf diseases using image processing and machine learning," International Journal of Computer Applications, vol. 975, pp. 8887–8890, 2016.
[8] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in Proc. ICML, 2020, pp. 1597–1607.
[9] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in Proc. CVPR, 2020, pp. 9729–9738.
[10] J. Grill et al., "Bootstrap your own latent: A new approach to self-supervised learning," in Proc. NeurIPS, 2020, pp. 21271–21284.
[11] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in Proc. ICLR, 2021.
[12] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798–1828, 2013.
[13] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. ICML, 2019, pp. 6105–6114.
[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, 2015.
[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. CVPR, 2016, pp. 770–778.
[16] R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. ICCV, 2017, pp. 618–626.
[17] S. Minaee et al., "Image classification using deep learning: A survey," IEEE Access, vol. 9, pp. 13537–13565, 2021.
[18] A. Saleem, J. Potgieter, and K. Arif, "Plant disease classification: A comparative evaluation of CNN architectures," IEEE Access, vol. 8, pp. 113641–113653, 2020.
[19] P. Blackman and V. Eastop, Aphids on the World's Crops, 2nd ed., Wiley, 2020.
[20] S. Agrios, Plant Pathology, 5th ed., Elsevier Academic Press, 2015.
[21] H. Oliveira, R. Duque-Gamboa, and M. D. A. Ribeiro, "Whitefly pests and their impact on cotton," Crop Protection, vol. 64, pp. 1–9, 2014.
[22] J. Morse and M. Hoddle, "Invasion biology of thrips," Annual Review of Entomology, vol. 51, pp. 67–89, 2006.
[23] R. Verma, S. K. Singh, and R. K. Sharma, "Alternaria leaf spot disease in cotton," Plant Disease, vol. 102, no. 4, pp. 703–710, 2018.
[24] D. Jones et al., "Fusarium wilt disease management in cotton," Plant Pathology Journal, vol. 33, no. 4, pp. 321–330, 2017.
[25] C. Shorten and T. Khoshgoftaar, "A survey on image data augmentation," Journal of Big Data, vol. 6, no. 60, pp. 1–48, 2019.
[26] X. Chen and K. He, "Exploring simple Siamese representation learning," in Proc. CVPR, 2021, pp. 15750–15758.
[27] Y. Liu et al., "Self-supervised learning for agricultural image analysis," Plant Methods, vol. 17, no. 96, pp. 1–15, 2021.
[28] A. Vaswani et al., "Attention is all you need," in Proc. NeurIPS, 2017, pp. 5998–6008.
[29] P. Mohan, R. Thirumalai, and S. Prakash, "Artificial intelligence in precision agriculture," IEEE Internet of Things Journal, vol. 9, no. 6, pp. 4512–4525, 2022.
[30] FAO, "Cotton production and productivity report," Food and Agriculture Organization of the United Nations, Rome, Italy, 2022.