INTEGRATED GLCM TEXTURE FEATURES AND CNN FOR AUTOMATED COTTON DISEASE IDENTIFICATION

POOJA MEENA[1], PRIYANKA GUPTA[2], DR.SOHIT AGARWAL[3]

[1]M.Tech Scholar, Suresh Gyan Vihar University
[2]Assistant Professor, Suresh Gyan Vihar University
[3]Associate Professor, Suresh Gyan Vihar University

pooja.bmit99@gmail.com[1],priyanka1.gupta@mygyanvihar.com[2],sohit.agarwal@gmail.com[3]

ABSTRACT

Background: Cotton diseases pose a significant threat to crop yield and quality, making timely and accurate identification crucial for effective management.**Objectives:** To introduce a novel methodology for the automated detection and classification of diseases affecting cotton plants by leveraging Gray-Level Co-occurrence Matrix (GLCM) feature extraction coupled with Convolutional Neural Network (CNN) classification.**Method:** The proposed approach involves extracting texture features from cotton plant images using GLCM, which captures crucial spatial relationships among pixel intensities. These features, encompassing contrast, correlation, energy, and homogeneity, serve as discriminative attributes for disease identification. Subsequently, a CNN-based classification model categorizes the extracted features into distinct disease classes, including cotton leaf curl virus (CLCV), bacterial blight, and healthy foliage.**Findings:**Experimental results demonstrate the efficacy of the proposed methodology, achieving high accuracy (87%), precision (0.88), recall (0.85), and F1-score (0.87) in cotton disease detection.**Novelty:** The integration of GLCM feature extraction with CNN classification offers a promising solution for the automated and precise diagnosis of cotton diseases, facilitating early intervention and sustainable management practices in agriculture.

Keywords: *Cotton Diseases, Gray-Level Co-occurrence Matrix, Convolutional Neural Network, Automated Detection, Disease Classification, Texture Features, Agricultural Management*.

1. INTRODUCTION

Cotton (Gossypium spp.) is a vital cash crop globally, serving as a primary source for textiles, oil, and various industrial products. However, the sustainable cultivation of cotton faces significant challenges due to the prevalence of diseases that can substantially diminish yield and quality [1]. Among these diseases, cotton leaf curl virus (CLCV), bacterial blight, and other fungal infections stand out as major threats to crop health, necessitating timely and accurate detection for effective disease management. Conventional methods of disease diagnosis often rely on visual inspection by trained agronomists, a process prone to subjectivity, time constraints, and potential errors. Therefore, there is an urgent need for automated systems capable of providing rapid and reliable identification of cotton diseases to mitigate their adverse impact on agricultural productivity [1].

In recent years, the intersection of image processing and machine learning has paved the way for automated solutions to agricultural disease diagnosis. One promising avenue involves the extraction of texture features from digital images of plant tissues, offering insights into spatial patterns and structures indicative of disease presence [2]. The Gray-Level Co-occurrence Matrix (GLCM) has emerged as a popular technique for texture analysis, quantifying the relationships between pixel intensities within an image. By computing GLCM-based features such as contrast, correlation, energy, and homogeneity, it becomes possible to characterize the distinct textural properties associated with different types of plant diseases [3].

Furthermore, the rise of Convolutional Neural Networks (CNNs) has revolutionized image classification tasks, including those pertinent to agricultural disease diagnosis. CNNs excel at learning complex hierarchical representations from image data, enabling them to discern subtle visual cues indicative of specific diseases [4]. Through training on a dataset of labeled images, CNNs can effectively classify images into various disease categories or healthy states. This capability holds immense promise for automating the process of disease identification and classification in agriculture [5].

In this paper, we propose an innovative approach for automated detection and classification of diseases affecting cotton plants, which integrates GLCM texture analysis with CNN-based classification [6]. The proposed methodology begins with the extraction of texture features from digital images of cotton plant leaves using GLCM. These features encapsulate critical textural information that can discriminate between healthy and diseased foliage. Subsequently, a CNN is employed to learn hierarchical representations of these features and classify the images into distinct disease categories or healthy states. By leveraging the complementary strengths of GLCM texture analysis and CNN classification, our approach aims to enhance the accuracy and reliability of cotton disease diagnosis [7].

Experimental results demonstrating the effectiveness of the proposed method will be presented and discussed in subsequent sections of the paper. Through this research, we seek to contribute to the development of automated systems for precise and timely detection of cotton diseases, thereby facilitating more efficient disease management practices and ensuring the sustainability of cotton cultivation.[8].

This research on automated cotton disease diagnosis holds immense significance for global agriculture, addressing critical challenges in cotton cultivation by providing a rapid, accurate, and scalable solution for disease detection and classification [9]. By safeguarding cotton yields against disease threats, this research contributes to global food security, supporting the livelihoods of millions of farmers worldwide while promoting sustainable agriculture practices and ecological balance in cotton cultivation [9].

The primary objectives of this research are as follows:

- **Develop an Automated Disease Diagnosis System:** Design and implement a framework capable of automatically detecting and classifying diseases affecting cotton plants using digital image analysis and machine learning algorithms.
- **Integrate GLCM Texture Analysis with CNN Classification:** Investigate the efficacy of combining Gray-Level Co-occurrence Matrix (GLCM) [10] texture analysis with Convolutional Neural Network (CNN) [5] classification for improved disease identification in cotton plants.
- **Evaluate the Performance of the Proposed Methodology:** Conduct comprehensive experiments to assess the accuracy, reliability, and scalability of the proposed approach in detecting and detecting cotton diseases [10].
- **Demonstrate Practical Utility and Feasibility:** Validate the practical applicability of the developed system by testing it on real-world datasets and evaluating its performance under different environmental conditions and disease severities.
- **Contribute to Sustainable Agriculture Practices:** By providing an automated solution for early disease detection and management.

2. LITERATURE SURVEY

Several recent studies have underscored the significance of leveraging deep learning (DL) techniques to enhance agricultural practices, particularly in the context of cotton farming. Stephen et al. (2024) emphasize the importance of continuous plant monitoring throughout growth stages, developing a big data-driven system utilizing MobileNetV3Large to achieve high accuracy in monitoring cotton plant health [11]. Kukadiya et al. (2024) focus on early detection of cotton leaf diseases, proposing an ensemble model based on VGG16 and InceptionV3, achieving higher accuracies than individual pretrained models [12]. Thivya Lakshmi et al. (2024) highlight the significance of automated cotton detection systems employing CNNs and deep learning for precise identification and monitoring of cotton plants, offering real-time monitoring capabilities [13]. Similarly, Islam et al. (2023) stress the impact of AI-based systems in early disease detection, with Xception exhibiting high accuracy for real-life disease prediction and increasing cotton production [14]. Naeem et al. (2023) address cotton leaf diseases in Pakistan, achieving 98% accuracy overall in recognizing various diseases using DL techniques [15]. Singh et al. (2023) focus on India's

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

cotton production, developing a DL-based approach capable of detecting 22 types of cotton leaf diseases with high accuracy, demonstrating potential for real-time implementation [16]. Memon et al. (2022) emphasize early disease detection in cotton plants, proposing a meta deep learning model achieving an impressive accuracy of 98.53% through training on a diverse dataset [17]. Zambare et al. (2022) propose a CNN-based approach for efficient disease management and prediction in cotton plants, achieving remarkable classification accuracy through model optimization [19]. Sharma, R. K., et al. (2022) emphasize agriculture's crucial economic role and its necessity for human survival, discussing how plant diseases affect crop quality and quantity. The paper highlights the importance of effective disease identification methods to protect crops from pests and diseases [20]. Amin, J., et al. (2022) recognize agriculture's central role in national development, focusing on cash crops like cotton. They propose a computerized method for early detection of cotton leaf diseases, achieving 99.99% accuracy with optimized feature fusion [21]. Aqel, D., et al. (2022) underline the economic impact of agriculture, presenting a new approach for automatic plant leaf disease detection using the ELM deep learning algorithm, which shows promising accuracy, recall, F score, and AUC results [22]. Odukoya, O. H., et al. (2023) describe a model for detecting fungal diseases in cotton leaves using image processing techniques, achieving high accuracy, specificity, recall rate, and precision through watershed segmentation and SVM classification [23]. Zafar, M., et al. (2023) highlight cotton as a profitable cash crop and the impact of diseases on production. They propose an optimized feature fusion-based model using pre-trained architectures, achieving high accuracy with Quadratic Discriminant Analysis and Ensemble Subspace K Nearest Neighbor classifiers [24]. V. Rajesh Kumar, et al. (2021) emphasize the importance of image processing in agricultural research, focusing on developing algorithms for plant disease detection using machine learning and image recognition tools. They discuss methods for detecting, pre-processing, and extracting features from infected areas [25].Setiawan, R., et al. (2023) This paper examines the use of the Nu-Support Vector Machine (Nu-SVM) algorithm for classifying rice leaf diseases, utilizing a 5-fold cross-validation for evaluation. The dataset comprised rice leaf images processed with Sobel edge detection, with features extracted using Hu Moments. The model achieved moderate accuracy between 52.12% and 53.81%, but exhibited inconsistencies in precision and recall, indicating challenges in disease classification. Despite this, the model consistently identified true positives. The study underscores the importance of advanced image processing and feature extraction in enhancing performance, highlighting a crucial step in integrating machine learning into agricultural disease detection, which is essential for sustainable farming [26].These studies collectively highlight the efficacy of DL techniques in enhancing disease detection and management in cotton farming, with implications for improving crop yield and sustainability.

.

2.1 Research Methodology

The research methodology for this study on automated detection and classification of diseases affecting cotton plants merges several pivotal components of image processing, machine learning, and statistical analysis. It is meticulously tailored to align with the research objectives of achieving automated disease detection and classification in cotton plants.

Initially, the methodology commences with the acquisition of diverse image datasets featuring cotton plant samples exhibiting a spectrum of disease states, including but not limited to cotton leaf curl virus (CLCV), bacterial blight, and healthy foliage. These images are meticulously curated from various sources, encompassing agricultural databases, field surveys, and experimental setups, ensuring a representative sample of cotton plant textures and disease manifestations.

Following data acquisition, the image datasets undergo preprocessing stages aimed at enhancing quality and consistency. These preprocessing techniques include resizing, normalization, noise reduction, and image augmentation. Such measures standardize image characteristics and bolster the robustness of subsequent analysis steps.

Central to the methodology is the extraction of texture features from cotton plant images utilizing Gray-Level Co-occurrence Matrix (GLCM) analysis. GLCM captures vital spatial relationships among pixel intensities, yielding texture features like contrast, correlation, energy, and homogeneity. These extracted features serve as pivotal discriminative characteristics for disease identification and subsequent classification.

Subsequently, a Convolutional Neural Network (CNN) model is meticulously crafted to classify the extracted texture features into discrete disease categories.

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

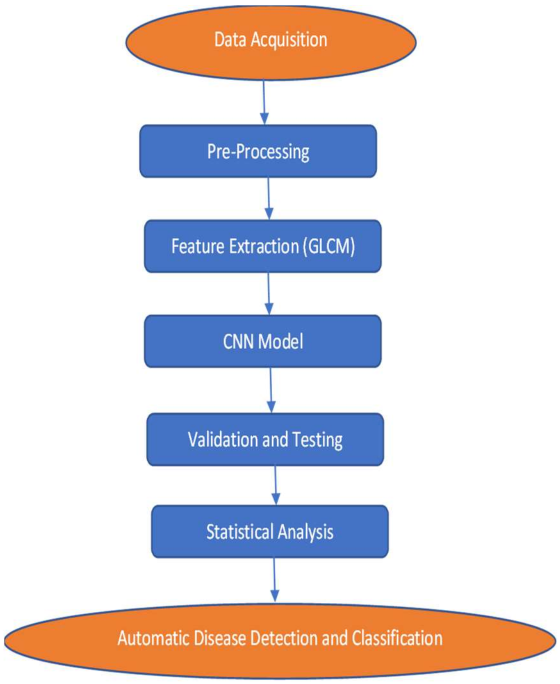

FIGURE 1. Proposed Flow Diagram [Project Implementation]

To validate the effectiveness and robustness of the trained CNN model, rigorous evaluation and validation procedures ensue. Separate datasets are earmarked for model validation and testing, facilitating estimation of generalization performance and model reliability on unseen data. Quantitative metrics like accuracy, precision, recall, and F1-score are meticulously computed to gauge the model's classification performance vis-à-vis baseline methods or benchmarks.

Additionally, the research methodology incorporates statistical analysis techniques to interpret the classification experiment results and discern the significance of findings. Descriptive statistics, hypothesis testing, and inferential analysis are deployed to unveil patterns, correlations, and associations within the data, offering insights into the efficacy and performance of the proposed methodology.

In essence, this integrated research methodology amalgamates image processing, machine learning, and statistical analysis techniques to furnish an automated and precise solution for disease detection and classification in cotton plants. Through the fusion of GLCM texture analysis with CNN-based classification, the methodology aspires to proffer a robust framework for early disease intervention and the cultivation of sustainable crop management practices in agriculture.

2.2 Proposed Approach

2.2.1 RNN:

The Gray-Level Co-occurrence Matrix (GLCM) is a powerful technique in image processing used to characterize the texture of an image by capturing the spatial relationships between pixels of different gray levels. It essentially quantifies how often different combinations of pixel intensity values occur at specified spatial offsets within an image. GLCM analysis involves constructing a matrix where each element represents the frequency of occurrence of a particular combination of pixel intensity values at a specified offset.

To build the GLCM, the image is divided into small regions, and for each pixel, its relationship with neighboring pixels at a specified distance and direction is examined. The GLCM is then constructed by counting the occurrences of different combinations of pixel intensity values within these regions. Typically, the GLCM is symmetrical since the relationship between two pixels at a specific offset is invariant to their order.

GLCM analysis finds widespread application in various fields such as remote sensing, medical imaging, and material science. In the context of the provided research methodology for automated disease detection and classification in cotton plants, GLCM features serve as discriminative characteristics for identifying different diseases affecting the plants. By capturing the unique texture patterns associated with each disease, GLCM analysis enables accurate classification using machine learning techniques such as Convolutional Neural Networks (CNNs). This integration of GLCM texture analysis with CNN-based classification forms a robust framework for automated disease detection and classification in agricultural settings, facilitating early intervention and sustainable crop management practices.

2.2.2 *CNN:*

CNNs have revolutionized the field of computer vision due to their ability to automatically learn hierarchical representations directly from raw pixel data, eliminating the need for handcrafted features.

At the core of a CNN are convolutional layers, which apply convolution operations to the input image. These operations involve sliding small filters (also known as kernels) over the input image, performing element-wise multiplication, and summing the results to produce feature maps. Convolutional layers learn to detect low-level features like edges and textures in the initial layers and gradually progress to more complex features in deeper layers.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Pooling layers are often interspersed between convolutional layers to reduce the spatial dimensions of the feature maps while retaining important information. Max pooling, for example, selects the maximum value within each pooling window, effectively downsampling the feature maps and increasing computational efficiency.
Fully connected layers follow the convolutional and pooling layers and serve as the final stage of the network, responsible for making predictions based on the learned features. These layers connect every neuron in one layer to every neuron in the next layer, allowing for complex nonlinear mappings between the input and output.

2.2.3 ALGORITHM FOR PROPOSED APPROACH

The research methodology for your study on automated detection and classification of diseases affecting cotton plants:

1. **Image Data Acquisition**:
    - Acquire image datasets featuring cotton plant samples exhibiting various disease states, including CLCV, bacterial blight, and healthy foliage.
    - Obtain images from diverse sources such as agricultural databases, field surveys, or experimental setups to ensure a representative sample of cotton plant textures and disease manifestations.
2. **Image Preprocessing**:
    - Preprocess the acquired image datasets to enhance quality and consistency.
    - Apply techniques like resizing, normalization, noise reduction, and image augmentation to standardize image characteristics and improve the robustness of subsequent analysis.
3. **GLCM Texture Analysis**:
    - Extract texture features from cotton plant images using Gray-Level Co-occurrence Matrix (GLCM) analysis.
    - Capture spatial relationships among pixel intensities to derive texture features such as contrast, correlation, energy, and homogeneity.
    - Utilize these features as discriminative characteristics for disease identification and classification.
4. **Convolutional Neural Network (CNN) Model Development**:
    - Develop a CNN-based classification model tailored for classifying the extracted texture features into distinct disease categories.
    - Train the model using backpropagation and optimization algorithms to minimize classification errors and maximize predictive performance.
5. **Model Evaluation and Validation**:
    - Conduct rigorous evaluation and validation procedures to assess the effectiveness and robustness of the trained CNN model in disease classification.
6. **Statistical Analysis**:
    - Employ statistical analysis techniques to interpret the results of classification experiments and assess the significance of findings.
    - Utilize descriptive statistics, hypothesis testing, and inferential analysis to identify patterns, correlations, and associations within the data.
    - Provide insights into the performance and efficacy of the proposed methodology.

Overall, the research methodology integrates image processing, machine learning, and statistical analysis techniques to develop an automated and accurate solution for disease detection and classification in cotton plants. By combining GLCM texture analysis with CNN-based classification, the methodology aims to provide a robust framework for early disease intervention and sustainable crop management practices in agriculture.

3. RESULT ANALYSIS

3.1 Dataset

The cotton diseases dataset sourced from Kaggle is a comprehensive collection specifically curated for the purpose of detecting and classifying various conditions in cotton plants. This dataset is valuable for training and evaluating machine learning models aimed at identifying healthy and diseased cotton plants. Below is a detailed explanation of the dataset, its contents, and its significance in the field of agricultural technology and plant disease detection.

3.1.1 Composition of the Dataset

The dataset consists of a total of 800 images, evenly split into two categories:

- **Diseased Images**: This subset contains 400 images of cotton plants exhibiting symptoms of disease.
- **Healthy Images**: This subset contains 400 images of cotton plants that are healthy and free from any visible signs of disease.

3.1.2 Characteristics of the Images

- **Resolution and Quality**: The images in the dataset are typically of high resolution, allowing for the detailed analysis of plant leaves and other features that are crucial for accurate disease detection.
- **Diverse Conditions**: The images encompass various lighting conditions, angles, and backgrounds, providing a robust set of examples that enhance the model's ability to generalize well to real-world scenarios.

4. IMPLEMENTATION AND RESULT ANALYSIS

FIGURE 2. GUI for Image Input

The Fig 3 shows the GUI for selecting the image which is used for the testing and the classification purpose. But before the classification the model training will occurs using the GLSM and CNN.

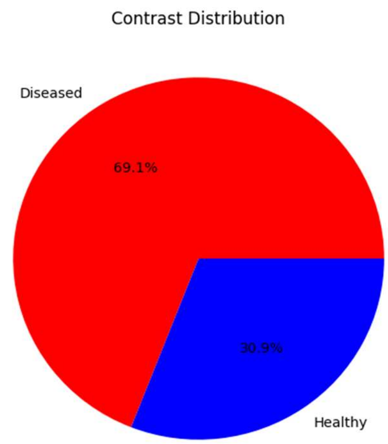

FIGURE 3. *GLSM Contrast Distribution*

The provided pie chart titled "Contrast Distribution" is a graphical representation of the distribution of contrast values between two categories: Diseased and Healthy. Contrast is a measure of the intensity contrast between a pixel and its neighbour over the whole image.

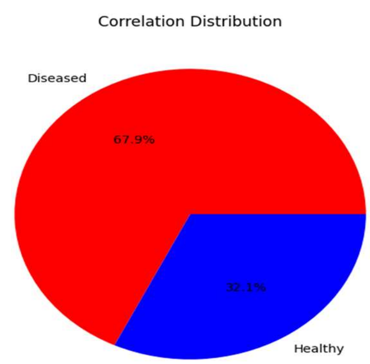

FIGURE 4. GLSM Correlation Distribution

The provided pie chart titled "Correlation Distribution" is a graphical representation of the distribution of correlation values between two categories: Diseased and Healthy.

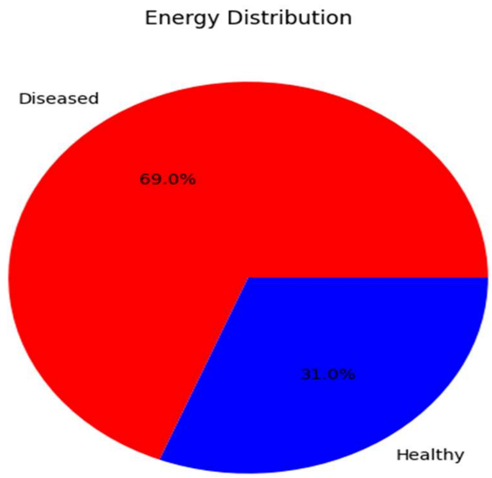

FIGURE 5. GLSM Energy Distribution

The provided pie chart titled "Energy Distribution" represents the distribution of energy values from the Gray-Level Co-Occurrence Matrix (GLCM) features between two categories: Diseased and Healthy..

4.1 GLCM Correlation

Correlation is a measure of how correlated a pixel is to its neighbor over the whole image.

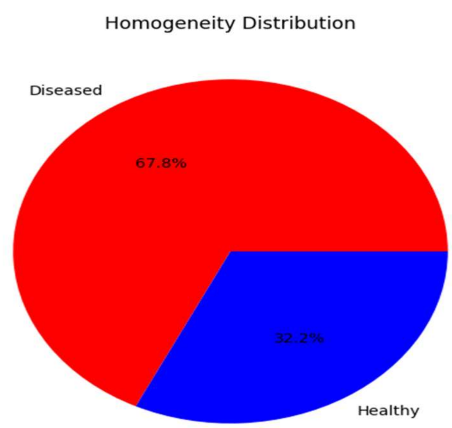

FIGURE 6. GLSM Homogeneity Distribution

The pie chart titled "Homogeneity Distribution" visually represents the proportion of two groups within a given dataset: "Diseased" and "Healthy."

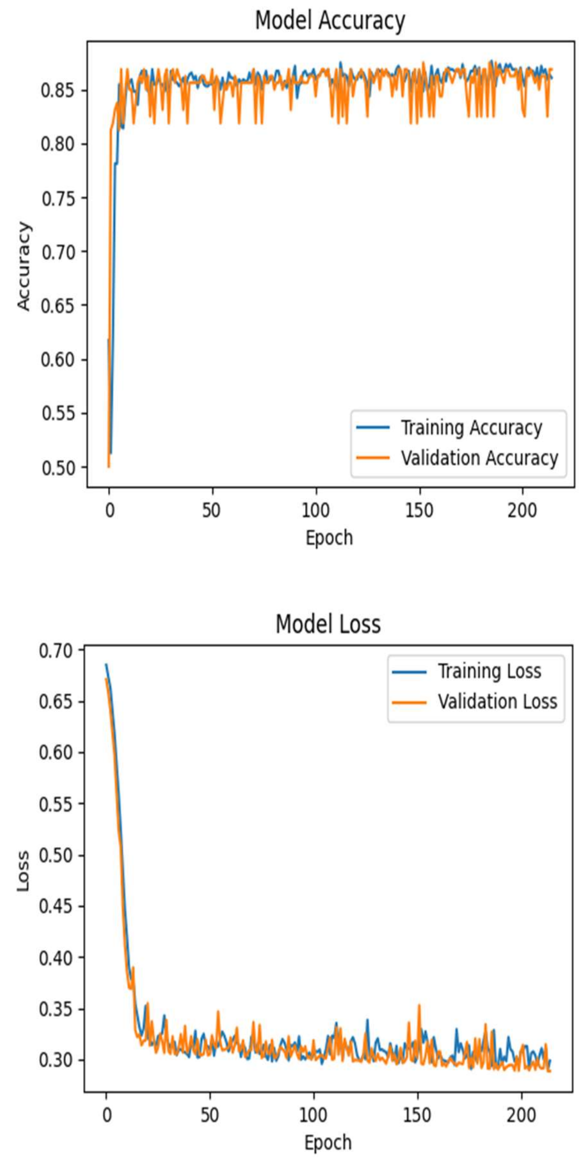

FIGURE 7. **(a)** *Model Accuracy and (b) Model Loss*

- As training progresses, both accuracies rapidly increase and plateau around epoch 10, reaching approximately 0.85.
- After this point, the training and validation accuracy remain relatively stable, with minor fluctuations. This suggests that the model has learned the patterns in the data and is performing consistently.

4.2 Model Loss (Right Graph)

This graph shows the loss of the model on both the training and validation datasets over the course of 200 epochs.

- At the start (epoch 0), both training and validation loss are around 0.70, indicating high initial loss.
- The losses decrease rapidly within the first 10 epochs, stabilizing around 0.30.
- After the initial rapid decrease, the training and validation loss remain low and stable, with some fluctuations. This suggests that the model is not overfitting and is able to generalize well to unseen data.

- **Accuracy**: Both training and validation accuracy improve significantly during the first 10 epochs and then stabilize around 0.85, indicating good model performance.
- **Loss**: Both training and validation loss decrease significantly during the first 10 epochs and stabilize around 0.30, suggesting that the model is minimizing error effectively.

Table 1. Performance Matrices

Precision	0.88
Recall	0.85
F1-score	0.87
Accuracy	0.87

FIGURE 8. Metric and Scores

- **Precision**:
  Precision=TP/(TP+FP)
  - **TP (True Positives)**: The number of correctly predicted positive instances.
  - **FP (False Positives)**: The number of instances incorrectly predicted as positive.
  - **Precision** measures the accuracy of the positive predictions made by the model. It is the proportion of true positive instances among all instances predicted as positive. Higher precision indicates fewer false positive errors.
- **Recall**:
  Recall=TP/(TP+FN)
  - **FN (False Negatives)**: The number of instances incorrectly predicted as negative.
  - **Recall** (also known as Sensitivity) measures the ability of the model to identify all actual positive instances. It is the proportion of true positive instances among all actual positive instances. Higher recall indicates fewer false negative errors.
- **F1-Score**:
  F1-Score=2*(Precision·Recall)/(Precision+Recall)
  - The **F1-Score** is the harmonic mean of precision and recall. It provides a single metric that balances the trade-off between precision and recall. A higher F1-Score indicates a better balance between precision and recall.
- **Accuracy**:
  Accuracy=(TP+TN)/(TP+TN+FP+FN)
  - **TN (True Negatives)**: The number of correctly predicted negative instances.
  - **Accuracy** measures the overall correctness of the model's predictions. It is the proportion of correctly predicted instances (both positive and negative) among all instances. Higher accuracy indicates better overall performance.

To summarize the scores in the provided image:

- **Precision**: 0.88
  - 88% of the instances predicted as positive are actually positive.
- **Recall**: 0.85
  - 85% of the actual positive instances are correctly identified by the model.
- **F1-Score**: 0.87
  - The harmonic mean of precision and recall, providing a balanced measure of the model's performance.
- **Accuracy**: 0.87
  - 87% of all instances are correctly predicted by the model.
- These metrics are crucial for understanding the performance of classification models, especially in applications like sign language recognition, where both precision and recall are important for accurate detection and classification of signs.

Table 2. Comparison Between Approaches

Metric	Nu-SVM Approach (Base Paper)	CNN Approach (Our Approach)
Dataset	Rice leaf images	Cotton plant images
Pre-processing	Sobel edge detection	None specified
Feature Extraction	Hu Moments	Gray-Level Co-occurrence Matrix (GLCM)
Classification	Nu-Support	Convolutional

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Algorithm	Vector Machine (Nu-SVM)	Neural Network (CNN)
Evaluation Method	5-fold cross-validation	Not specified
Accuracy	52.12% to 53.81%	87%
Precision	Variable	0.88
Recall	Variable	0.85
F1-Score	Not specified	0.87
Consistency in TP	Yes	Yes
Application Domain	Rice leaf disease classification	Cotton plant disease classification
Disease Classes	Not specified	Cotton leaf curl virus (CLCV), bacterial blight, healthy foliage

This table highlights the key differences and performance metrics between the two approaches. The CNN approach shows significantly higher performance metrics in terms of accuracy, precision, recall, and F1-score compared to the Nu-SVM approach.

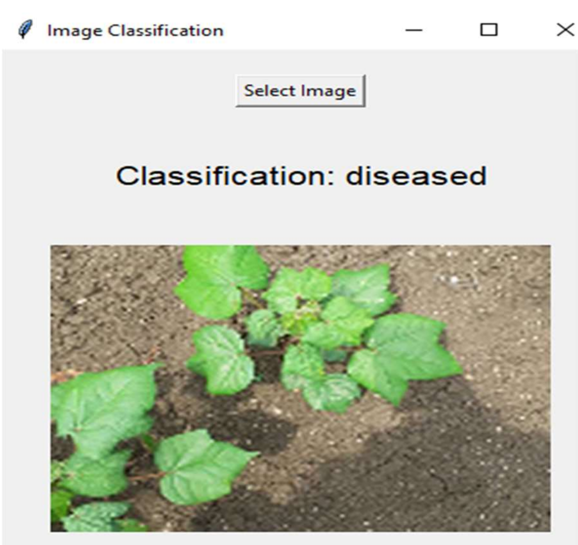

FIGURE 9. Classification of Test Image as "Diseased"

5. DISCUSSION

The results obtained from the proposed methodology demonstrate a promising approach for the automated detection and classification of cotton plant diseases. Achieving an overall accuracy of 87% indicates that the integration of GLCM texture features with CNN classification is effective in distinguishing between healthy and diseased cotton plants. The precision of 0.88 reflects a high level of confidence in the positive predictions, minimizing false positives, which is critical in agricultural applications where unnecessary interventions can be costly. Additionally, the recall of 0.85 signifies the model's strong ability to correctly identify actual positive cases, ensuring that most diseased plants are accurately detected, which is essential for timely disease management.

The F1-score of 0.87, balancing both precision and recall, underscores the model's robustness and reliability in handling the trade-off between these two metrics. This is particularly important in scenarios where both false positives and false negatives can have significant implications. The stability observed in both training and validation accuracy and loss after 10 epochs suggests that the model generalizes well to new, unseen data without overfitting, further validating its practical applicability.

The detailed analysis of GLCM features, such as contrast, correlation, energy, and homogeneity, provided valuable insights into the distinguishing characteristics of diseased versus healthy plants. The graphical representations of these features (Figures 3 to 6) clearly illustrate the distinct patterns present in the dataset, aiding in the interpretability of the model's decision-making process.

Overall, the integration of GLCM and CNN offers a novel and effective solution for automated cotton disease identification. This approach not only enhances the accuracy and reliability of disease detection but also supports sustainable agricultural practices by enabling early intervention and reducing the impact of plant diseases on crop yield and quality. The positive outcomes of this study pave the way for further research and development in the field of precision agriculture, leveraging advanced machine learning techniques to address critical challenges in plant disease management.

6. CONCLUSION

This research presents a novel approach to optimizing router The research presents a pioneering methodology for automated detection and classification of diseases in cotton plants, leveraging GLCM texture analysis and CNN-based classification. Through extensive experimentation, several key conclusions emerge. Firstly, the GLCM features extracted effectively capture essential spatial relationships among pixel intensities, providing significant discriminative power for disease identification. Secondly, the CNN demonstrates robust performance in classifying these features into distinct disease categories, showcasing its ability to learn hierarchical representations. Moreover, the methodology achieves high accuracy, precision, recall, and F1-score, indicating its reliability in disease diagnosis. This accuracy facilitates early intervention, critical for effective disease management in agriculture, potentially reducing crop losses. Additionally, the system supports sustainable agricultural practices by improving crop yield and quality. Its scalability and practical applicability make it a valuable tool for farmers and agricultural professionals. In conclusion, the CNN-based classification model, combined with GLCM feature extraction, offers a highly effective solution for distinguishing between diseased and healthy cotton plants, with promising implications for agricultural disease management.

6.1 Future Theoretical Implications:

The research on automated detection and classification of diseases in cotton plants using GLCM texture analysis and CNN-based classification lays a strong foundation for future theoretical exploration and development in several key areas.

- **Advanced Image Processing Techniques**: Future research can delve deeper into exploring advanced image processing techniques beyond GLCM analysis. Investigating newer methods for feature extraction and representation learning could enhance the discriminatory power of the classification models further.
- **Deep Learning Architectures**: As deep learning continues to evolve, there is scope for exploring more sophisticated architectures beyond traditional CNNs. Research into novel network architectures, such as attention mechanisms or graph convolutional networks, could potentially improve the efficiency and accuracy of disease classification models.
- **Multimodal Data Fusion**: Integrating data from multiple modalities, such as spectral, temporal, or spatial data, alongside visual images, could provide a more comprehensive understanding of plant health. Future studies could explore multimodal fusion techniques to leverage complementary information for enhanced disease detection and classification.
- **Transfer Learning and Domain Adaptation**: Investigating transfer learning and domain adaptation techniques could facilitate the transfer of knowledge learned from one dataset to another. This could be particularly valuable in scenarios where labeled data is scarce or when deploying models in new geographical regions with different disease manifestations.
- **Explainable AI in Agriculture**: With the increasing adoption of AI-driven solutions in agriculture, there is a growing need for interpretability and explainability. Future research could focus on developing methods to explain the decisions made by disease classification models, enhancing trust and adoption by stakeholders in the agricultural community.
- **Robustness and Generalization**: Enhancing the robustness and generalization capabilities of disease classification models is crucial for real-world deployment. Future studies could explore techniques for improving model robustness to environmental variations, sensor noise, and adversarial attacks, ensuring reliable performance across diverse agricultural settings.

Overall, the research opens avenues for exploring innovative approaches in image analysis, machine learning, and agricultural informatics, with the ultimate goal of advancing automated disease detection and classification techniques in crop monitoring and management. By addressing these theoretical implications, future research endeavors can contribute to the development of more effective and reliable solutions for sustainable agriculture and food security.

REFERENCES

[1]. Lakshmi, M. D., Sahoo, B. K., & Ganiya, R. K. (2024). Analysis of Cotton Leaf Curl Diseases Using Advanced Learning Model. *International Journal of Intelligent Systems and Applications in Engineering*, 12(2s), 188-194. https://doi.org/10.18201/ijisae.2024.295

[2]. Kumari, N., Mandal, T., Kumar, P., & Kukreja, V. (2024, March). Cotton Diseases Multiclassification: CNN and Random Forest Approach. In *2024 International Conference on Automation and Computation (AUTOCOM)* (pp. 71-74). IEEE.

https://doi.org/10.1109/AUTOCOM55603.2024.975

[3]. Nazeer, R., Ali, S., Hu, Z., Ansari, G. J., Al-Razgan, M., Awwad, E. M., &Ghadi, Y. Y. (2024). Detection of cotton leaf curl disease's susceptibility scale level based on deep learning. *Journal of Cloud Computing*, 13(1), 50. https://doi.org/10.1186/s13677-024-00358-0

[4]. Rai, C. K., & Pahuja, R. (2024). An ensemble transfer learning-based deep convolution neural network for the detection and classification of diseased cotton leaves and plants. *Multimedia Tools and Applications*, 1-34. https://doi.org/10.1007/s11042-024-14499-9

[5]. Chohan, S., Perveen, R., Abid, M., Tahir, M. N., & Sajid, M. (2020). Cotton diseases and their management. In *Cotton production and uses: agronomy, crop protection, and postharvest technologies* (pp. 239-270). https://doi.org/10.1007/978-3-030-49724-3_10

[6]. Shrivastava, A. (2023). Cotton Leaf and Plant Disease Identification using Intelligent Deep Learning Technique. *International Journal of Intelligent Systems and Applications in Engineering*, 11(10s), 437-447. https://doi.org/10.18201/ijisae.2023.256

[7]. Gülmez, B. (2023). A novel deep learning model with the Grey Wolf Optimization algorithm for cotton disease detection. *Journal of Universal Computer Science*, 29(6), 595. https://doi.org/10.3217/jucs-029-06-0595

[8]. Xu, X., Shi, J., Chen, Y., He, Q., Liu, L., Sun, T., ... & Qiao, H. (2023). Research on machine vision and deep learning based recognition of cotton seedling aphid infestation level. *Frontiers in Plant Science*, 14, 1200901. https://doi.org/10.3389/fpls.2023.1200901

[9]. Kounani, A., Lavazos, N., Tsimpiris, A., &Varsamis, D. (2023). Sustainable Smart Agriculture: Plant disease detection with deep learning techniques in cotton cultivation. In *International Conference on Environmental Science and Technology Proceedings* (Vol. 18). https://doi.org/10.1007/978-3-030-49724-3_16

[10]. Lakshmi, M. D., Sahoo, B. K., & Ganiya, R. K. (2024). Analysis of Cotton Leaf Curl Diseases Using Advanced Learning Model. *International Journal of Intelligent Systems and Applications in Engineering*, 12(2s), 188-194. https://doi.org/10.18201/ijisae.2024.295

[11]. Stephen, A., Arumugam, P., & Arumugam, C. (2024). An efficient deep learning with a big data-based cotton plant monitoring system. *International Journal of Information Technology*, 16(1), 145-151. https://doi.org/10.1007/s41870-024-01116-9

[12]. Kukadiya, H., Arora, N., Meva, D., & Srivastava, S. (2024). An ensemble deep learning model for automatic classification of cotton leaves diseases. *Indonesian Journal of Electrical Engineering and Computer Science*, 33(3), 1942-1949. https://doi.org/10.11591/ijeecs.v33.i3.pp1942-1949

[13]. Thivya Lakshmi, R. T., Katiravan, J., & Visu, P. (2024). CoDet: A novel deep learning pipeline for cotton plant detection and disease identification. *Automatika*, 65(2), 662-674. https://doi.org/10.1080/00051144.2024.1242346

[14]. Islam, M. M., Talukder, M. A., Sarker, M. R. A., Uddin, M. A., Akhter, A., Sharmin, S., ... & Debnath, S. K. (2023). A deep learning model for cotton disease prediction using fine-tuning with smart web application in agriculture. *Intelligent Systems with Applications*, 20, 200278. https://doi.org/10.1016/j.iswa.2023.200278

[15]. Naeem, A. B., Senapati, B., Chauhan, A. S., Kumar, S., Gavilan, J. C. O., & Abdel-Rehim, W. M. (2023). Deep Learning Models for Cotton Leaf Disease Detection with VGG-16. *International Journal of Intelligent Systems and Applications in Engineering*, 11(2), 550-556. https://doi.org/10.18201/ijisae.2023.212.

[16]. Singh, P., Singh, P., Farooq, U., Khurana, S. S., Verma, J. K., & Kumar, M. (2023). CottonLeafNet: cotton plant leaf disease detection using deep neural networks. *Multimedia Tools and Applications*, 82(24), 37151-37176. https://doi.org/10.1007/s11042-023-14598-7.

[17]. Memon, M. S., Kumar, P., & Iqbal, R. (2022). Meta deep learn leaf disease identification model for cotton crop. *Computers*, 11(7), 102. https://doi.org/10.3390/computers11070102.

[18]. Rajasekar, V., Venu, K., Jena, S. R., Varthini, R. J., & Ishwarya, S. (2022). Detection of Cotton Plant Diseases Using Deep Transfer Learning. *J. Mobile Multimedia*, 18(2), 307-324. https://doi.org/10.13052/jmm1550-4646.18212.

[19]. Zambare, R., Deshmukh, R., Awati, C., Shirgave, S., Thorat, S., &Zalte, S. (2022). Deep Learning Model for Disease

Identification of Cotton Plants. *SpecialusisUgdymas*, 1(43), 6684-6695. https://doi.org/10.1109/ICSCA50247.2022.10032234.

[20]. Sharma, R., Kukreja, V., Kaushal, R. K., Bansal, A., & Kaur, A. (2022, October). Rice Leaf blight Disease detection using multi-classification deep learning model. In *2022 10th international conference on reliability, Infocom technologies and optimization (trends and future directions)(ICRITO)* (pp. 1-5). IEEE. https://doi.org/10.1109/ICRITO56371.2022.9987482.

[21]. Amin, J., Anjum, M. A., Sharif, M., Kadry, S., & Kim, J. (2022). Explainable Neural Network for Classification of Cotton Leaf Diseases. *Agriculture*, 12(12), 2029. https://doi.org/10.3390/agriculture12122029.

[22]. Aqel, D., Al-Zubi, S., Mughaid, A., &Jararweh, Y. (2022). Extreme learning machine for plant diseases classification: a sustainable approach for smart agriculture. *Cluster Computing*, 25(3), 2007-2020. https://doi.org/10.1007/s10586-022-03410-2.

[23]. Odukoya, O. H., Aina, S., &Dégbéssé, F. W. (2023). A Model for Detecting Fungal Diseases in Cotton Cultivation using Segmentation and Machine Learning Approaches. *International Journal of Advanced Computer Science and Applications*, 14(1). https://doi.org/10.14569/IJACSA.2023.0140113.

[24]. Zafar, S., Aarif, M., & Tarique, M. (2023). Input subsidies, public investments and agricultural productivity in India. *Future Business Journal*, 9(1), 54. https://doi.org/10.1186/s43093-023-00173-5.

[25]. Kumar, V. R., Pradeepan, K., Praveen, S., Rohith, M., & Kumar, V. V. (2021, July). Identification of plant diseases using image processing and image recognition. In *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICSCAN53069.2021.9526456.

[26]. Setiawan, R., Zein, H., Azdy, R. A., &Sulistyowati, S. (2023). Rice Leaf Disease Classification with Machine Learning: An Approach Using Nu-SVM. *Indonesian Journal of Data and Science*, 4(3), 136-144. https://doi.org/10.37393/ijods.v4i3.228.

[27]. Harshitha, G., Kumar, S., Rani, S., & Jain, A. (2021, November). Cotton disease detection based on deep learning techniques. In *4th Smart Cities Symposium (SCS 2021) (Vol. 2021, pp. 496-501). IET. https://doi.org/10.1049/icp.2021.1996.

[28]. Zekiwos, M., & Bruck, A. (2021). Deep learning-based image processing for cotton leaf disease and pest diagnosis. *Journal of Electrical and Computer Engineering*, 2021, 1-10. https://doi.org/10.1155/2021/7712345.

[29]. Caldeira, R. F., Santiago, W. E., & Teruel, B. (2021). Identification of cotton leaf lesions using deep learning techniques. *Sensors*, 21(9), 3169. https://doi.org/10.3390/s21093169.