JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Improvised Tulsi Leaf Disease Detection using CNN

Garima Paliwal [1], Riddhi Raj Rathore[2], Siddhi Raj Rathore[3] , Dr.Mayank Patel[4]

[123]UG Scholar, Computer Science and Engineering, Geetanjali Institute of Technical Studies, Dabok, Udaipur, Rajasthan

[4] Professor, Computer Science and Engineering Geetanjali Institute of Technical Studies, Dabok, Udaipur, Rajasthan

[1]garimapaliwal262002@gmail.com, [2]riddhirajrathorewrk26@gmail.com, [3] rathoresiddhiraj26@gmail.com, [4]mayank999_udaipur@yahoo.com

Abstract :-This study has been undertaken to investigate An automatic method for classifying the diseases that affect the Tulsi (Holy Basil) leaf. It aims to apply the features of the CNN (Convolutional Neural Network) to classify the Tulsi leaf into different classes such as healthy leaf or infected leaf. The study investigates how CNN improves prediction accuracy and refers its importance with respect to disease management and potential yield of Tulsi plant.

***IndexTerms* - Disease detection, CNN, image detection, Machine learning, Tulsi leaf**

I. INTRODUCTION

Tulsi is an important medicinal herb which has been traditionally used in several systems of medicine and has become more popular today. Diseases if not detected early, quality and quantity of yield of Tulsi plants can be greatly compromised. Manual approaches for the diagnosis of diseases have clear limitations in that they are time-consuming largely because they rely on the human eye interpretive skill to make decisions. This therefore creates a need for the development of fast and efficient ways of diagnosis. It is believed that having a tulsi plant in the courtyard of house brings happiness & prosperity in house. The most common type of tulsi found in Indian households is Rama tulsi, also known as green leaf tulsi or holy basil which is slightly sweet in taste It's also used to make tea and other remedies for colds, sore throats, headaches, and sinusitis.

II PROBLEM STATEMENT

Human error is a liability with the manual methods of inspections and concentrated effort may not be able to address diseases at the early stages. Machine learning based technologies especially, CNNs can provide automatic diagnosis of diseases in a timely manner and in most cases with high accuracy.

III LITERATURE REVIEW

CNN Based Plant Disease Detection Methods:

- Lu et al. (2021): Provided a brief explanation on the principles of deep learning, outlined the issues of plant disease classification, and further emphasized the presence of datasets that are large and of good quality.
- ugrul et al. (2022): Reported the importance of CNNs in the early stages of disease detection and further stressed the significance of quality datasets.
- Pacal et al. (2024): In their study that involved reviewing 160 articles, they acknowledged CNNs as useful tools in the detection of plant disease and also noted the drawbacks of small dataset availability.
- Sheetal S. Patil et al-This study directly worked on Tulsi leaf disease detection using CNN.It reported 75% accuracy, so it’s likely a practical implementation or experiment. It’s relevance is very strong source for Tulsi leaf CNN-based detection with measurable results.
- IEEE Xplore Conference Paper-A conference paper from IEEE, which is credible, focus on Tulsi leaf disease detection using CNNs. It is highly relevant and reliable reference.

Fig.1:Healthy tulsi plant

Source: Google

As shown in Fig 1, the Tulsi leaf appears healthy, characterized by its vibrant green color, smooth surface, and uniform shape. The absence of spots, yellowing, or deformities indicates that the plant is in good condition, making it an ideal example of a non-infected leaf for comparative analysis.

IV OBJECTIVES

To design a CNN based univariate model for identification and classification-

- Formulate a CNN-based model that can accurately detect and classify the diseases affecting Tulsi leaves.
- Train and validate the model until a prediction accuracy that is satisfactory is achieved.

CONTRIBUTIONS:

Collection process and organizing of the dataset comprising images of Tulsi leaves labeled as healthy or sick.
Suggestion of a CNN model that is oriented towards disease recognition.
Model evaluation, prediction analysis of the model's accuracy

V TYPES OF TULSI LEAF DISEASE

1. **Powdery mildew disease**

Powdery mildew is also a fungal infection which is caused by parasites. Powdery mildew fungi belong to order Erysiphales are most common obligate parasites . On the host surface, these fungi produce enormous conidia, which leave a white powdery covering and dusty look.

Fig.2 Powdery mildew diseased leaves
Source: Google

As illustrated in Fig 2, the infected Tulsi leaf shows visible signs of Powdery Mildew, including whitish or gray powder-like patches on the surface. These fungal spots typically reduce the plant's ability to perform photosynthesis, leading to reduced vitality and yield.

2. **Leaf spot disease**

Basil downy mildew is a disease that affects basil leaves, branches, and stems. Symptoms incude

- Basil downy mildew is a disease that affects basil leaves, branches, and stems. Symptoms include:
- Yellowing: Leaves turn yellow, often in sections restricted by veins
- Fuzzy growth: A grayish-purple fuzzy growth appears on the lower surface of leaves in humid conditions
- Brown leaves: Infected leaves turn brown and scorched as the disease progresses
- Black spots: Irregular black spots appear on infected leaves as they age.

Fig.3 Downy mildew diseases leaves
Source: Google

Fig 3 highlights a Tulsi leaf infected with Downy Mildew, a disease caused by water mold pathogens. The leaf exhibits irregular yellow spots on the top surface and a downy, mold-like layer underneath. These symptoms disrupt normal leaf function and are commonly observed in humid conditions.

VI. DATASET COLLECTION AND PREPROCESSING

Data collection:-
Source- online images,online repositories Classes of Diseases: Healthy, powdery mildew, and leaf spot.

Table 1 : dataset collection

Disease type	Sample count
Healthy leaves	200
Powdery mildew disease	150
Leaf spot or downy mildew disease	150

Image Preprocessing: - Image Resizing : Resize images to 128x128 pixels. - Normalization : Scale pixel values to the [0, 1] range. -Data Augmentation :Apply techniques such as rotation, flipping and brightness adjustments to increase dataset. Diversity

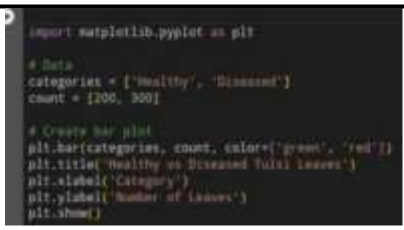

Fig 4 : image of code used for making a graph of healthy vs diseased tulsi leaves

Fig 4 illustrates the Python code used to generate a graphical comparison between healthy and diseased Tulsi leaves. The graph visually represents the distribution or count of classified leaf types, aiding in the interpretation of the model's effectiveness. By plotting this comparison, the analysis provides clearer insights into how well the CNN model distinguishes between the two categories, enhancing the overall evaluation of classification accuracy.

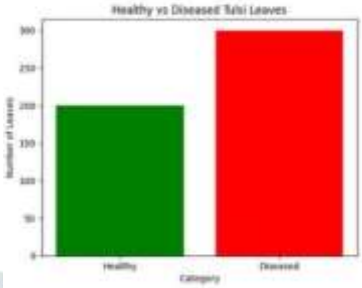

Fig 5 : healthy vs diseased tulsi leaves

Fig 5 displays a bar graph comparing the number of healthy and diseased Tulsi leaves. The noticeably taller bar for diseased leaves indicates that the dataset contains a higher count of infected samples compared to healthy ones. This imbalance can play a critical role in model training and evaluation, as a larger number of diseased examples may lead to better recognition of infections, but also raises the need for careful handling to avoid biased predictions.

VII. Methodology

Using traditional machine learning approaches, achieving the classification problem takes numerous sequential phases, including pre-processing and feature extraction . For picture categorization, the project's technique employs neural networks unsupervised machine learning techniques such as neural network (NN) A CNN is a form of machine learning method that works with input that has a grid layout, such digital photos. Pixels are placed in grids in the image structure, and each pixel has a unique property. CNN draws its name from "convolution matrices," a mathematical linear process between matricesIn the earliest layers of the CNN algorithm, basic characteristics such as lines and curves are detected, while more complex features (patterns and faces) are detected in advance levels.A CNN typically has three layers:

- Convolutional layer
- Pooling layer
- Fully connected layer

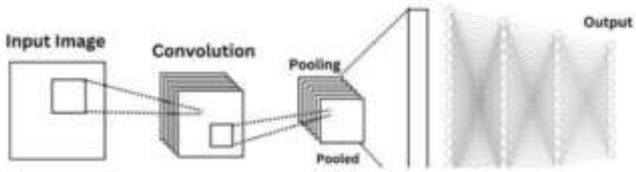

Fig 6 : CNN model diagram

Figure 6 presents the architecture of the Convolutional Neural Network (CNN) used for Tulsi leaf disease detection. The model consists of multiple layers, including convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. This structured flow allows the model to learn and identify complex patterns in leaf images, enabling it to distinguish between healthy and various diseased conditions effectively.

The model of the system takes a database of leaf images. Then the data is cleaned and classified into two parts: Training and Testing Dataset The training dataset is now pre-processed for further implementations. Pre-processing of images include:

- Setting the width and height of the images.
- Setting the color orientation.
- Setting the input size of the images

After the pre-processing the image will go to the CNN layers and it will first get classified into categories: Infected and non-infected. Finally, the images categorized as infected will be differentiate further for prediction of different types of Model Training: - Loss Function: The categorical cross-entropy. - Optimizer: Adam optimizer with a learning rate of 0.001. - Hyperparameters: A batch size of 32 with 50 epochs by making use of the early stopping technique to avoid overfitting. Image segmentation in CNN: The goal of the image segmentation is to segment the plant leaf digital image into various parts and partitioning is called segmentation. It's a tool for detecting image features Picture segmentation recognizes a variety of features in a picture, including Line, Point, and Boundary of specific areas . It assist with the recognition of distinguishing aspects in a photograph. Feature Extraction: We do feature extraction after segmentation. It's a method for determining unique elements that represent a picture by taking into account the color and texture of the image. To test the model's accuracy, the testing dataset will be introduced into the trained CNN model.

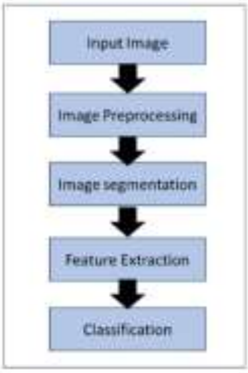

Fig 7: Flowchart for leaf disease recognition

As shown in Figure 7, the flowchart outlines the complete process of Tulsi leaf disease recognition, starting from image input and preprocessing to classification using the CNN model and final disease prediction. This structured approach ensures accurate detection by following a systematic pipeline.

VIII. Technological Outcomes And Result

Evaluation Metrics -Accuracy: Overall proportion of correctly classified images.
- Precision, Recall, and F1 Score: Classify performance for each class.
- Confusion Matrix: Display performance of each class in classification.

Results and Analysis:
Model Performance vs. Traditional Methods:
Compare the performance of CNN with traditional methods like KNN, CNN model does all succeeds to obtain more features with classification accuracy regarding complicated images also.
Discussion- Effectiveness of CNNs
The CNN method proved to be quite effective in detecting Tulsi leaf diseases as it automatically learned distinguishing features from images.
Data Augmentation: It notably enhanced the robustness and generalization abilities of the model.

Table 2 : performance matrix

Type	Precision	Recall	F1-Score
Healthy	0.75	1.00	0.86
powdery mildew	1.00	0.33	0.50
Leaf spot	0.80	1.00	0.89

Accuracy = 0.80

Table 2 shows that the model performs well for Healthy and Leaf Spot categories, as seen by high recall and F1-scores. However, the performance for Powdery Mildew is poor, primarily due to low recall, meaning the model struggles to identify all true cases of this condition.

Fig 8 : Image of code used for confusion matrix

Figure 8 displays the code used to generate the confusion matrix, which evaluates the performance of the CNN model in classifying Tulsi leaf images. The confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives, offering deeper insights into the model's accuracy and potential misclassifications.

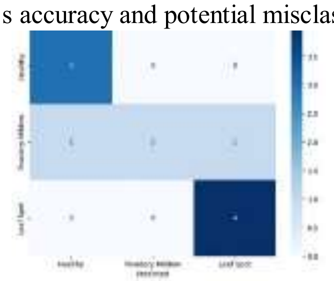

Fig 9: confusion matrix

As shown in Fig.9 The confusion matrix shows that the model performs well for the Healthy and Leaf Spot classes, but struggles slightly with the Powdery Mildew class. One powdery mildew leaf was misclassified as Healthy and other as Leaf Spot.

```
import matplotlib.pyplot as plt

# Sample data
epochs = range(1, 11)
accuracy = [70, 75, 78, 80, 82, 84, 85, 86, 87, 88]

plt.plot(epochs, accuracy, marker='o')
plt.title('Model Accuracy vs. Epochs')
plt.xlabel('Epochs')
plt.ylabel('Accuracy (%)')
plt.show()
```

Fig 10 : image of code used for making model accuracy vs Epochs graph

Fig10 illustrates the code implemented to plot the model's accuracy over training epochs. This graph helps visualize how the CNN model improves (or fluctuates) in accuracy as it learns from the dataset over time.

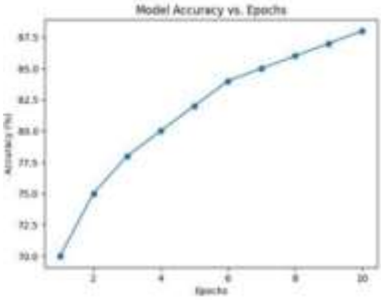

Fig 11 : model accuracy vs Epochs

As shown in Fig.11 this graph visualizes how increasing the number of epochs enhances model performance, though the rate of improvement slows as accuracy approaches its upper limit.

IX. Conclusion

In so doing, the study confirms the potential of CNNs for automation in the detection of Tulsi leaf disease, so high accuracy is achieved and then practical applications are presented for agriculture purposes. Future research should focus on how to expand the dataset and explore other architectures for CNN, further developing real-time mobile applications that can be applied in the field. This will reduce the time and efforts required for the process of manual inspection and will find early diagnosis and treatment of the plant. It is expected that due to technological advancements and increased availability of data, the accuracy and efficiency of the CNN models are also going to improve, making it a valuable tool in agriculture. CNNs offer a revolutionary approach toward the detection of diseases in plants, precision, scalability, and cost- effectiveness. There is a great potential for this technology in the management of plant health of Tulsi and other crops by overcoming the present issues and advanced integration strategies.

X. Future Scope

- Real-Time Detection: Develop mobile applications that would execute a trained CNN model to instantaneously identify the disease in the field. such an application might employ the camera of a mobile device to capture and instantly classify the type of disease in leaf.
- IoT Integration :IoT integration within the detection system for continuous monitoring and agricultural field-based early warning systems.
- Future research can be highly contributed towards the development and advancements in capabilities and applications of CNN-based disease-detection systems, leading to sustainable agricultural practices and improved crop management.
- It is a small dataset size and easy overfitting ; higher resolution images can be used to make performance even better. so will be using big dataset size for additional features (lighting and other variations) that may change the model performance.

XI.Acknowledgment

I would like to express my sincere gratitude to my project guide/mentor Dr. Mayank Patel(Hod CSE) for their continuous support, guidance, and valuable suggestions throughout this research work on Tulsi leaf disease detection using Convolutional Neural Networks.

I also extend my thanks to Geetanjali Institute of Technical Studies for providing the resources and environment necessary for completing this work. I would like to acknowledge the authors and researchers whose works provided a strong foundation and reference for my study.

XII. References

[1] Tugrul, C., Ozgen, A. P., & Sevim, A. (2022). Using of "Early Disease Detection in plants Convolutional Neural Network". Journal Agricultural Informatics,13(2), 56-63
[2] Pacal, I., Aydin, G., & Hussain, A. (2024). "Comprehensive Review of CNN Applications in Plant Disease Detection". Artificial Intelligence Review, 57(4), 993-1022.
[3] Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. Disease Detection: A Review". Computers and Electronics in Agriculture, 164, 104946.
[4] Patil, M. S., & Kumar, R. (2023). "Automated Disease Detection in Tulsi Leaves Using Deep International Conference on Artificial Intelligence and Machine Learning, 54, 23-27.
[5] Shen, W., Zhou, M., Yang, B., & Tong, Y. (2018). "Multilayer Convolutional Neural Networks
[6] Sheetal S. Patil et al "Tulsi Leaf Disease Detection using CNN" detecting diseases in Tulsi leaves, highlighting the model's 75% accuracy rate

[7] IEEE Xplore Conference Paper: "Tulsi Leaf Disease Detection using CNN" published in the 2022 lEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation

[8] Kanchan Ashok Taksale and Dr.Sachin Bhoite "A Review of Machine Learning and Deep Learning Approaches for Tulsi Classification in Health Monitoring and Species Identification" This review explores various ML and DL techniques, including CNNs, for identifying and categorizing diseases in Tulsi leaves

[9] A.Maru, A. K. Sharma and M. Patel, "Hybrid Machine Learning Classification Technique for Improve Accuracy of Heart Disease," *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, 2021, pp. 1107-1110, doi: 10.1109/ICICT50816.2021.9358616.

[10] Tiwari, K., Patel, M. (2020). Facial Expression Recognition Using Random Forest Classifier. In: Mathur, G., Sharma, H., Bundele, M., Dey, N., Paprzycki, M. (eds) International Conference on Artificial Intelligence: Advances and Applications 2019. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-15-1059-5_15

[11] Patel M (2018) Data Structure and Algorithm With C. Educreation Publishing

[12] Taunk, D., Patel, M. (2021). Hybrid Restricted Boltzmann Algorithm for Audio Genre Classification. In: Sheth, A., Sinhal, A., Shrivastava, A., Pandey, A.K. (eds) Intelligent Systems. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-2248-9_11

[13] Sardogan et al "A convolutional neural network-driven computer vision system toward plant leaf recognition and leaf disease detection" This article covers the application of CNNs for plant leaf recognition and disease detection.