**CSE / EEE / ETE 499A (Section 4)**

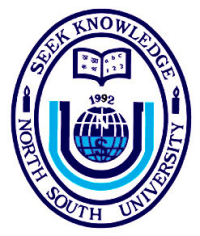
**Project Proposal (CO1)**

**Project Title:** Plant Diseases Detection Using Image Processing

**Submitted To**

**Dr. Shazzad Hosain (SZZ)**

**Date: 7/10/2024**



**Semester: Summer 24**

**Group No: G-3**

**Group Members**

|  |  |
| --- | --- |
| **ID** | **Name** |
| 2011892042 | Akib Hasan |
| 2012464042 | Abroor Zahin Niloy |
| 2011566042 | Tania Akter Lima |
| 2012959042 | Naima Homaira Khan |

**Introduction**

Agriculture is increasingly leveraging artificial intelligence to address major challenges, particularly in crop health monitoring and disease management. Traditional methods of identifying plant diseases—often through manual inspection—are slow, inconsistent, and impractical for large-scale agricultural practices. These limitations have led to the development of automated, AI-powered solutions that use machine learning (ML) and deep learning (DL) to provide rapid and precise disease detection. Currently, deep learning models, especially Convolutional Neural Networks (CNNs), are among the most promising tools, offering high accuracy in image classification tasks critical to detecting various plant diseases.

This capstone project seeks to address the need for an efficient, accurate, and scalable disease detection system that outperforms conventional methods and even traditional ML models. By developing a framework that incorporates CNNs, VGG-ICNN, and image processing techniques like K-means clustering and GLCM-based feature extraction, the project aims to enhance disease classification accuracy and adaptability. Further, by integrating UAV-based imaging and real-time monitoring, this project extends disease detection capabilities to larger areas, supporting early intervention on a broad scale.

Through a combination of image processing, neural networks, and transfer learning, this project aspires to deliver a practical and high-performing solution that meets the evolving demands of precision agriculture. Solutions like John Deere’s AI-driven monitoring systems and Prospera Technologies’ UAV-enabled crop surveillance demonstrate the real-world applicability and impact of such technologies, underscoring the potential of our project to support sustainable agriculture and contribute to global food security.

**Literature Review**

[1] The research explores machine learning techniques for plant disease detection, utilizing Random Forests, CNNs, and feature descriptors like Hu moments, Haralick texture, and Color Histograms. Random Forest outperformed other models, achieving 70.14% accuracy. Key studies include Sannakki and Rajpurohit’s 97.30% accuracy with neural networks and Rothe and Kshirsagar’s 85.52% through pattern recognition. The methodology highlights image preprocessing, feature extraction via HoG, and classification with Random Forest. Results suggest increasing dataset size and incorporating advanced techniques like SIFT and SURF can further improve accuracy.

[2] The study reviews machine learning (ML) and deep learning (DL) techniques for detecting plant diseases, highlighting the superior accuracy of DL methods like VGG-ICNN, which achieved 99.16% accuracy, compared to traditional ML methods such as SVM, which reached 98.97%. Technological advancements, including Precision Agriculture and UAVs, are revolutionizing disease detection, enabling early intervention. Challenges such as climate change and land scarcity emphasize the need for integrated pest control and precision farming. This research underscores the potential of DL for accurate, efficient plant disease detection and calls for further advancements to improve agricultural sustainability.

[3]This study explores machine learning (ML) and deep learning (DL) approaches for plant disease detection, highlighting limitations in traditional visual inspection methods. Models like SVM, decision trees, CNNs, and the Xception architecture are analyzed, with a focus on DL's ability to improve accuracy and automate classification. Challenges such as data imbalance and model interpretability are discussed, alongside the potential of frameworks like TensorFlow, PyTorch, and Scikit-learn. The paper emphasizes the importance of transfer learning and ensemble methods to improve ML accuracy, advocating for further research to enhance model scalability, robustness, and adaptability to real-world conditions.

[4]This paper explores advanced AI-driven approaches for plant disease detection, leveraging machine vision and deep learning for early, accurate diagnosis by analyzing leaf image features. It examines methods like thresholding, region growing, and clustering for image segmentation, and emphasizes the importance of color, texture, and shape in feature extraction. Despite significant improvements, challenges remain due to limited high-quality, labeled datasets representing diverse real-world conditions. The study suggests prioritizing AI for early detection and monitoring disease progression, while also highlighting the need for scalable models that manage high computational demands and detect multiple plant diseases simultaneously.

[5]This study presents an efficient plant disease detection and classification system utilizing statistical machine learning, image processing, and the random forest classifier to achieve a high average accuracy of 93% with a 0.93 F1 score. By detecting 20 diseases across 5 common plants, the proposed system is computationally less demanding and faster than deep learning approaches. Techniques like K-fold cross-validation, confusion matrix analysis, and feature extraction (including contrast, dissimilarity, and homogeneity) enhance accuracy and robustness. The balanced accuracy and F1 scores indicate effective handling of false positives and negatives, making it a promising solution for real-world agricultural applications.

[6]This paper emphasizes the significance of early plant disease detection for agriculture, addressing the limitations of traditional manual inspection methods. It reviews advancements in AI, particularly machine learning (ML) and deep learning (DL) techniques, which have improved accuracy in disease identification using leaf images. The study highlights the effectiveness of Convolutional Neural Networks (CNNs) for image processing, discusses various ML and DL models like Naive Bayes, K-Nearest Neighbors, and Artificial Neural Networks, and proposes innovative real-time detection methods. This comprehensive analysis aids future researchers in identifying suitable models for enhancing multi-class and multi-label plant disease classification.

[7]This study examines the application of various object detection algorithms and machine learning techniques for effective plant disease detection. It highlights the SmartPlantCare system, integrating algorithms such as K-Nearest Neighbors, Naïve Bayes, and Random Forest, while emphasizing feature extraction methods like color, texture, and shape analysis. The research evaluates model performance using multiple datasets, including the Apple-Cucumber-Tomato-Grape and Plant Detector datasets, and employs segmentation methods such as thresholding and watershed algorithms. The findings stress the importance of feature extraction in enhancing model accuracy and reliability, thus improving agricultural disease management practices.

[8]This research paper investigates plant disease detection and classification through various object detection algorithms and machine learning techniques. It employs algorithms such as SSD, DSSD, and R-SSD, alongside segmentation methods like K-means clustering, Chan-Vese segmentation, and Grab-cut segmentation tailored for different crops. Feature extraction techniques include the Color Co-occurrence method, Local Binary Patterns, and wavelet-based extraction. The study utilizes a neural network with backpropagation for classification and evaluates performance using SVM, ANN, KNN, and others. The findings reveal detection accuracies of 90.723% for fruit crops, 87.825% for vegetables, and emphasizes the efficacy of SVM as a robust classifier.

[9]This research paper outlines a methodology for detecting and classifying leaf diseases using image processing, K-means clustering, and Convolutional Neural Networks (CNNs). The process involves image segmentation through K-means clustering, where similar pixels are grouped to form distinct clusters. Feature extraction utilizes the Gray-Level Co-occurrence Matrix (GLCM) to derive statistical features such as mean, variance, and contrast. A CNN with 10 hidden layers, trained for 2000 epochs, classifies the diseases. The methodology demonstrates a mean accuracy of 89.8% for pomegranate leaves and 91% for potato leaves, highlighting its effectiveness as an automated diagnostic tool in agriculture.

[10]This study explores advanced machine learning and image processing techniques for detecting apple crop diseases like Alternia leaf spot, brown spot, and rust. Utilizing image segmentation methods (e.g., k-means clustering, Otsu’s method) and feature extraction techniques (e.g., GLCM), the paper compares classifiers, finding Naive Bayes slightly more accurate than Random Forest (99.03% vs. 97.71% accuracy). Challenges in adapting models to real-world conditions are discussed, highlighting the importance of dataset diversity. This research underscores effective approaches for high-accuracy apple disease detection and classification, providing insights for agricultural applications.

**Business viability of the project**

The business viability of our plant disease detection project is grounded in its innovative use of CNN, VGG-ICNN, and UNet algorithms for accurate and early disease classification in crops, a critical need for agricultural sustainability. By leveraging publicly available datasets, our system can target diverse market segments, including commercial agriculture firms, research institutions, and individual farmers seeking precision solutions. The project's primary competitors provide general detection systems; however, our model's focus on high accuracy in disease classification—enhanced through targeted image processing techniques like segmentation and feature extraction—gives it a unique advantage in reliability and usability for agriculture.

Success will be measured by maintaining classification accuracy rates above 98% and achieving robust, multi-class disease identification adaptable across crops, which are must-have outcomes. A feature that will be nice to have is additional data augmentation for future adaptability. Notable risks involve data limitations, potential accuracy issues with newer disease variants, and concerns around model adaptability to different environmental factors. Mitigating these risks while focusing on accuracy and model flexibility will establish the project as a valuable asset in agricultural technology, delivering clear benefits to the agricultural sector through cost-effective and impactful disease management.

**Project analysis**

1. **Major components of the project**

* Dataset Collection
* **Image Preprocessing**
* **Model Creation**
* Training the Model
* Model Evaluation

1. **Target for 499A and 499B**

**Target for 499A:**

In the CSE 499A, we will focus on several key components that are crucial for the successful development of our plant disease detection model. Our goals include:

* **Dataset Collection:** We will gather a well-structured dataset that includes images of healthy and diseased plants.
* **Image Preprocessing:** Once the dataset is collected, we will process the images to improve their quality and make them suitable for analysis.
* **Model Creation:** In this phase, we will design and build a convolutional neural network (CNN) architecture for detecting plant diseases.
* **Model Training:** After creating the model, we will train it using our prepared dataset. This involves feeding the model our images and adjusting its parameters to minimize prediction errors.
* **Performance Evaluation:** Finally, we will evaluate the model's performance using various metrics such as accuracy, precision, and recall. This will help us understand how well the model is performing and identify areas for improvement.

**Target for 499B:**

In CSE 499B, we will shift our focus to further enhancing our model and deploying it effectively. Our goals will include:

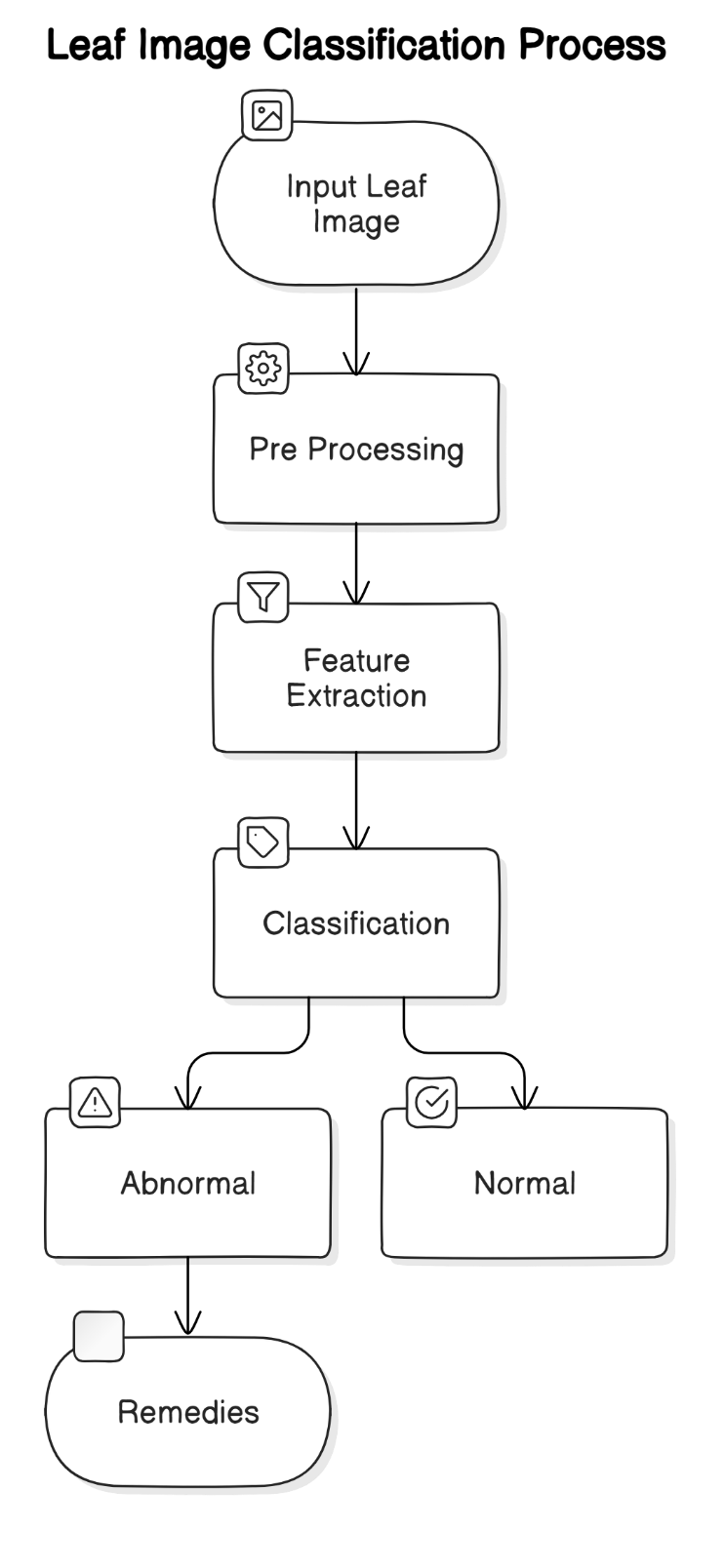
* **Optimizing Model:** We will explore techniques to improve the model's accuracy and efficiency, such as fine-tuning hyper parameters, using advanced training techniques, or experimenting with different architectures.
* **Web App Deployment:** Finally, we will develop a web application that allows users to easily access the model for plant disease detection. This will involve integrating the model into a user-friendly interface and ensuring it functions seamlessly for end users.

1. **Task analysis, activity diagram, critical path analysis**

**Task analysis**

* **Dataset Collection:** We will collect an existing dataset of plant images from publicly available sources like Kaggle or specialized agricultural repositories. This dataset will include both healthy and diseased plant leaves, labeled with the type of disease.
* **Preprocessing:** Preprocessing will involve resizing, normalization, and augmentation (flipping, rotation, etc.) to ensure a balanced dataset for training.
* **Convolutional Neural Network (CNN) Design:** The core of the project will be building a CNN model. CNNs are ideal for image-based tasks like this due to their ability to automatically detect important features like edges, textures, and patterns, making them suitable for plant disease identification.
* **Training the Model:** The CNN will be trained on the preprocessed images. This includes splitting the dataset into training, validation, and test sets to avoid overfitting. We will also fine-tune the model by adjusting parameters like learning rate, optimizer choice, and layers.
* **Performance Evaluation:** Model performance will be measured using metrics like accuracy, precision, recall, and F1-score on the test dataset. Confusion matrices and graphs will be used to visualize the performance across different plant disease categories.
* **Result Analysis**: We will analyze the performance of the model by comparing its predictions with actual results. Any discrepancies will be used to fine-tune the model.
* **Deployment:** The trained model can be deployed on a simple web application using frameworks like Flask, allowing users to upload leaf images for disease classification.

**Activity Diagram**



**Figure:** Activity Diagram

**Critical Path Analysis**

**Dataset Collection**

This task involves identifying and sourcing a relevant dataset for plant disease detection. It is the starting point for the project and has no prior dependencies. Without this, the subsequent tasks cannot begin.

**Image Preprocessing**

After the dataset is acquired, images need to be preprocessed to make them suitable for training the CNN model. This includes operations like resizing, normalization, and augmentation. This task depends on the completion of dataset collection.

**CNN Model Design**

Designing the Convolutional Neural Network architecture based on the input image dimensions and project requirements is crucial. The model must be built after image preprocessing is done, as the input dimensions of the dataset will influence the design. Hence, this task depends on the completion of both dataset collection and image preprocessing.

**Model Training**

The designed model is trained on the preprocessed dataset. Training the model requires significant computational resources and time. It can only begin after the model design is finalized and the images are preprocessed.

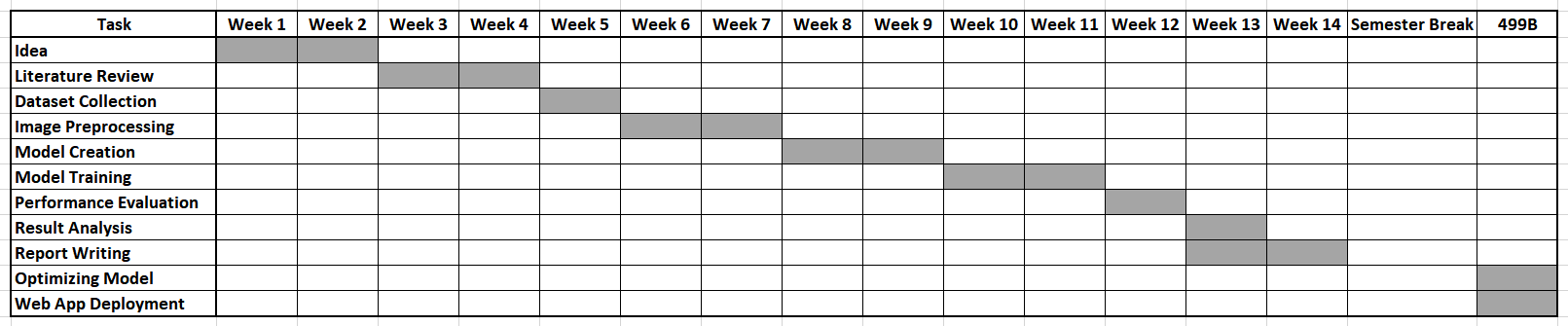
**Model Evaluation**

After training, the model is evaluated using a test dataset to measure its performance. This phase is crucial for understanding the accuracy, loss, and other metrics. It depends on the completion of the model training phase.

**Result Analysis and Fine-tuning**

Based on the evaluation results, the model’s performance will be analyzed. If necessary, the model will be fine-tuned to improve its accuracy. This task depends on the evaluation phase.

**d. Timeline**



# References

|  |  |
| --- | --- |
| [1] | M. R. H. N. M. P. R. P. B. N. S. N. Shima Ramesh, "Plant Disease Detection Using Machine Learning," *International Conference on Design Innovations for 3Cs Compute Communicate Control,* 2018. |
| [2] | P. G. R. S. Jadon, "Plant Disease Detection Using Machine Learning Models," *7th international joint conference on computing sciences,* 2023. |
| [3] | D. R. S. Nikhil S. Band, "Machine Learning Models for Plant Disease Detection and Classification," *International Journal ofINTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING,* 2024. |
| [4] | A. B. K. R. K. A. V. Y. A. Seyed Mohamad Javidan, "Feature engineering to identify plant diseases using image processing and artificial intelligence: A comprehensive review," *Smart Agricultural Technology,* 2024. |
| [5] | A. K. T. K. S. K. A. J. M. W. Pranesh Kulkarni, "Plant Disease Detection Using Image Processing and Machine Learning". |
| [6] | W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data,* 2024. |
| [7] | R. K. A. K. J. Bharti, "Smartplantcare: Plant Disease Detection Using Machine Learning Algorithms," *Journal of Engineering Design and Computational Science,* 2024. |
| [8] | D. G. M. H. K. J. M. K. Mr.V Suresh, "Plant Disease Detection using Image Processing," *International Journal of Engineering Research & Technology (IJERT),* 2020. |
| [9] | B. .. Y. V. Rajaneni Deepika sai chowdary, "PLANT DISEASE DETECTION USING IMAGE PROCESSING AND MACHINE LEARNING ALGORITHM," *Journal of Xidian University,* 2022. |
| [10] | P. C. V. A. F. J. G. P. Anupama Mishra, "Plant Disease Detection using Image," *International Conference on Cyber Security, Privacy and Networking (ICSPN 2022),* 2023. |