**Cassava Disease Detection and Classification – A Comparative study of Fuzzy Logic and Traditional Machine Learning.**

Michael Acquah

*ID: 84978417*

*University* *of Michigan-Dearborn*

*Dearborn, United States*

*macquah@umich.edu*

Aksheya Kannan Subramanian

*ID: 25502870*

*University of Michigan-Dearborn* *Dearborn,* United *States*

*aksheya@umich.edu*

***Abstract*—** **Plant diseases pose a significant threat to global food security, impacting crop yields and economic stability. Early detection and intervention are crucial in mitigating these losses. Current AI-driven plant disease detection heavily relies on Convolutional Neural Networks (CNNs). However, the need for large datasets and high computational power restricts their use in environments with limited resources. This research explores the potential of fuzzy logic as an alternative approach for cassava disease detection and classification, aiming to address the inherent uncertainty and imprecision in plant disease expression. Our study develops a fuzzy image processing framework, integrating fuzzy logic-based algorithms to accommodate variable disease manifestations. Through comparative analysis with conventional machine learning techniques, we assess the efficacy of fuzzy logic and hybrid model in terms of accuracy, precision, and computational efficiency. The project culminates in the creation of a prototype application designed to provide farmers with real-time disease dete ction capabilities. This work validates the use of fuzzy logic in agriculture and provides a practical tool for improving cassava disease management**

***Keywords—*** ***Plant Disease Detection, Fuzzy logic, Cassava, Disease, Convolutional Neural Networks (CNN), Real Time Detection***

1. Introduction

Food security serves as a foundational element for both human survival and economic stability, therefore, responsible agricultural resource management is a top priority. Several significant challenges impede effective crop production, plant diseases stand out as a great threat, capable of causing devastating losses to both farmers and the larger economy. Cassava, a staple crop for millions of people in tropical regions, is particularly vulnerable to a range of diseases that can severely impact yield and quality. Addressing these challenges requires innovative solutions that enable early detection and intervention to mitigate potential losses.

In recent years, artificial intelligence (AI) and machine learning (ML) have emerged as powerful tools in the realm of agricultural analysis and assessment. Convolutional Neural Networks (CNNs) have been at the forefront of this technological revolution, offering promising results in image-based disease detection. However, despite their potential, the practical application of CNNs in agriculture is often hampered by the scarcity of large, annotated datasets and sufficient computing power.

This research seeks to explore fuzzy logic as an alternative to traditional machine learning techniques for cassava disease detection and classification. Fuzzy logic, with its ability to handle imprecision and uncertainty, offers a promising approach to dealing with the complex and variable symptoms presented by plant diseases. Fuzzy image processing is employed in this study to create a disease detection solution that is both adaptable and accessible, even in environments with limited resources.

The primary objectives of this research include the development of a fuzzy image processing framework tailored to cassava diseases, the integration of fuzzy logic-based algorithms to enhance data interpretation, and a comparative analysis of these techniques against conventional image processing methods. Ultimately, the project aspires to deliver a prototype application capable of assisting farmers with real-time, reliable disease detection, thereby contributing to improved management practices and enhanced food security.

1. Literature review

The paper [1] tackles the critical challenge of early plant disease identification, essential for improving agricultural productivity and ensuring food security. Leveraging the EfficientNetV2S model, a pre-trained convolutional neural network (CNN), the study fine-tunes the model to handle noisy datasets, simulating real-world conditions such as low-resolution images, cluttered backgrounds, and varying brightness. Using the Plant Diseases Dataset, which contains 38 classes across various crops, the model achieves a remarkable validation accuracy of 95.01%. This demonstrates its effectiveness in addressing complex scenarios and its potential for practical deployment in agricultural settings.

A key innovation of the research is its deliberate introduction of noise into the dataset, enhancing the model’s robustness and ability to generalize under imperfect conditions. The EfficientNetV2S model’s lightweight architecture makes it suitable for field deployment, where computational resources may be limited. This study highlights the transformative potential of deep learning in agriculture, offering a reliable solution to combat crop loss and improve precision farming. By bridging the gap between theoretical research and real-world applications, the findings lay a strong foundation for future advancements in multi-crop disease detection and real-time monitoring systems.

This research compares supervised maximum likelihood (ML) and fuzzy logic image classification methods using a SPOT image [2]. Fuzzy logic, implemented in MATLAB, utilized a Sugeno-type inference system with Gaussian membership functions derived from the ML classification's statistics. The fuzzy approach achieved 89% accuracy in classifying land cover (water, urban, crops, vegetation), demonstrating its potential despite sensitivity to cloud cover and parameterization. While computationally potentially comparable to ML including pre-processing, fuzzy logic offers advantages in handling data uncertainty and providing an intuitive, rule-based classification framework.

In this research [3], Sari et all uses a Fuzzy Naïve Bayes Classifier (FNBC), incorporating fuzzy logic and the Saaty scale, to diagnose papaya diseases. Fuzzy logic handles ambiguous symptom descriptions, while the Naïve Bayes Classifier proficiently classifies diseases based on symptom weights. The system achieves 88% (FNBC) and 90% (forward chaining) accuracy against expert knowledge. Strengths include efficient processing and handling of uncertainty; weaknesses include limited scope (41 symptoms, 13 diseases), reliance on user input accuracy, and a need for broader dataset validation and algorithm exploration.

Mohd et al proposes a neuro-fuzzy system for early paddy disease detection in Malaysia, focusing on Bacterial Leaf Blight (BLB), Leaf Blast Disease (LBD), and Bacterial Sheath Blight (BSB). It uses a Canon 550D camera to capture images, MATLAB for image processing (including feature extraction via Canny edge detection and LAB color space analysis), and a Back-Propagation Neural Network (BPNN) combined with fuzzy logic for classification. While aiming for automated early detection, the achieved accuracy of 74.21% is modest. The methodology lacks detail (BPNN architecture, fuzzy rules), and the dataset's characteristics are not fully described, hindering reproducibility and assessment of generalizability. The reliance on visual features also limits the system's overall diagnostic capability[4].

This research uses computer vision and fuzzy logic to automate leaf disease detection and grading[5]. Leaf images are pre-processed, and features are extracted using GLCM and other methods. K-means clustering segments the leaf, diseased area, and background. An ANN classifies leaf type and disease, while a fuzzy logic system grades disease severity based on percentage infection (diseased area/total leaf area). While offering automation and quantitative assessment via a user-friendly GUI, the system's accuracy is limited by a small dataset and potentially suboptimal algorithms. Also, further research is needed to improve accuracy and generalizability.

This research paper analyzes existing methods for automated papaya plant disease detection. Most studies use image processing and machine learning, including CNNs (like ResNet50), SVMs, and Naive Bayes classifiers. While these offer faster, more accurate diagnosis than manual methods, a critical limitation is the scarcity of large, publicly available datasets representing diverse disease types and plant parts. This data shortage hinders model development and reproducibility. Future research needs to focus on creating and sharing comprehensive datasets to improve the reliability and applicability of automated papaya disease detection systems[6].

The next significant contribution comes from Awotunde et al. who presents an Interval Type-2 Fuzzy Logic System (IT2FLS) for plant disease diagnosis, achieving 97% accuracy on a Kaggle dataset [7]. IT2FLS's strength lies in handling the inherent uncertainty in plant disease identification, outperforming Type-1 Fuzzy Logic. However, the study lacks detail on input variables and dataset specifics, limiting assessment of generalizability. While implementation complexity is low, the inherent computational cost of T2FL remains a potential concern. Further research is needed to address these limitations and compare performance against other machine learning methods.

This research proposes an automated plant disease detection system using digital image processing and fuzzy classification [8]. Leaf images are acquired, pre-processed (noise reduction, Otsu's thresholding, edge detection), segmented (K-means clustering), and features (color, shape, texture) are extracted. Fuzzy classification, aided by a Random Forest (RF) classifier (though the integration isn't fully detailed), compares the image to a database, achieving 93% accuracy. Strengths include high accuracy and automation; weaknesses include a small dataset (20 samples), insufficient detail on the fuzzy logic implementation, and limited generalizability. Further research is needed to address these limitations.

This paper presents a novel ANFIS Fuzzy Convolutional Neural Network (CNN) model for detecting pepper bell leaf disease. The model integrates Local Binary Patterns (LBP) for improved efficiency and accuracy. Achieving more than 99% accuracy with LBP and demonstrating robustness through cross-validation, this method outperforms several existing ML/DL models. However, its reliance on a specific dataset (PlantVillage) limits generalizability, and further research is needed to assess its performance on other plant species and diseases, along with improving model interpretability [9].

This paper presents a plant leaf disease detection system using image processing and fuzzy logic. Leaf images undergo pre-processing, Otsu's thresholding segmentation, and feature extraction (GLCM and statistical measures). A fuzzy logic classifier achieves 88% accuracy on a dataset of 57 images representing three diseases. However, limitations include a small dataset and lack of detailed comparative analysis, requiring further validation and broader testing [10].

This study, Sibiya et al. presents an automated system for assessing maize common rust disease severity using image analysis, fuzzy logic, and deep learning [11]. Otsu thresholding segments diseased leaf areas, feeding into fuzzy rules that classify severity into four stages. A VGG-16 CNN then classifies these stages, achieving 95.63% validation and 89% testing accuracy. This surpasses human-based methods in accuracy and objectivity but requires a large, high-quality training dataset and computationally intensive processing. The grayscale conversion may lose valuable color information, and Otsu's method's limitations need consideration. Further testing on diverse datasets is crucial to validate its generalizability to other plant diseases.

This paper proposes a plant leaf disease detection system using a hybrid approach: K-means clustering segments diseased areas, fuzzy logic estimates infection likelihood (iron deficiency or fungal), and an SVM classifier determines disease severity. While claiming improved accuracy, the evaluation is limited by a small dataset (500 images) and lacks detailed performance metrics. The reliance on tuned fuzzy logic parameters and potentially suboptimal feature selection raises concerns about generalizability. Further testing on diverse datasets is needed to validate its effectiveness and compare it to other methods [12].

This research proposes IoT\_FBFN [13], an automated plant leaf disease detection system using IoT devices, SIFT feature extraction, a Fuzzy Based Function Network (FBFN) for classification, and the Firefy Algorithm (FFA) for optimization. It offers real-time monitoring and improved accuracy compared to manual methods, but its current scope is limited to \*Pauropsyllatuberculate\* galls on \*Alstonia Scholaris\*. Further research is needed to broaden its applicability and address potential limitations related to data dependency, computational resources, and network reliability.

1. Methodology

**3.1 Dataset**

The Cassava leaf disease dataset was obtained from Kaggle, a leading platform for research datasets. The dataset contains two folders namely the test folder and train folder with each containing five different classes of the cassava leaf. These classes are Cassava Bacterial Blight (CBB), Cassava Brown Streak Disease (CBSD), Cassava Green Mottle (CGM), Cassava Mosaic Disease (CMD) and Healthy. Overall, the dataset comprises of 21.4k plant leaf images.

**3.2 Preprocessing**

Data preprocessing is a very essential step in preparing the data for model training. The process begins with loading the Cassava Disease Plant dataset. The images are then resized to a uniform resolution, as CNN architectures require consistent input dimensions for optimal performance. Noise reduction techniques are also applied to reduce the impact of irrelevant background details. Also, pixel values are normalized in a range [0,1] to standardize the image data, enabling faster convergence during training. This preprocessing pipeline is very important in our work because it ensures that the model is working with high quality data, leading to better prediction accuracy.

A close-up of a plant

Description automatically generated

A close-up of plants

Description automatically generated

**Fig 3.2.1: Five classes Cassava Leaf Images**

3.2 Describe CNN model

The Convolutional Neural Network (CNN) model without fuzzy logic was developed using a transfer learning approach, leveraging the MobileNetV2 architecture pre-trained on the ImageNet dataset. The model's design involved fine-tuning the MobileNetV2 base for the task of cassava disease classification by adding a custom classification head. This included a GlobalAveragePooling2D layer to reduce the spatial dimensions of the extracted features, followed by a dense layer with 1024 neurons and ReLU activation for feature learning, and a final dense layer with softmax activation to output probabilities for each disease class. During training, the base layers of MobileNetV2 were frozen to retain the pre-trained feature extraction capabilities, while the newly added layers were optimized using the Adam optimizer with a learning rate of 0.0001, and categorical cross-entropy as the loss function.

The dataset was augmented for training using the ImageDataGenerator class to mitigate overfitting and improve generalization. Augmentation techniques included rescaling pixel values to [0, 1], random rotations, width and height shifts, shearing, zooming, horizontal flips, and filling missing pixels, ensuring that the model was exposed to a wide variety of image transformations. The training data was further divided into training and validation subsets, with 15% reserved for validation. Additionally, the test dataset was prepared using rescaling only, without augmentation, to ensure consistent evaluation.

The model training utilized early stopping to prevent overfitting by monitoring validation loss and restoring the best model weights after five epochs of no improvement. Model checkpoints were also employed to save the best-performing model during training. This carefully designed CNN model demonstrated the ability to accurately classify cassava diseases, serving as a strong baseline for comparative analysis against fuzzy logic-based approaches.

3.4 Describe fuzzy model

The proposed fuzzy hybrid model integrates fuzzy logic-based feature extraction with a deep learning architecture to enhance the classification of cassava plant diseases. The model begins by preprocessing image data using TensorFlow's ImageDataGenerator, scaling pixel values to the range [0, 1] and splitting the training dataset into training and validation sets. Fuzzy membership functions are employed to extract two primary features from the images: average color intensity and texture roughness. These features are fuzzified using Gaussian and trapezoidal membership functions, representing low, medium, and high color intensities, as well as smooth, moderate, and rough textures.

The hybrid model architecture combines a convolutional neural network (CNN) and a fuzzy neural network (FNN). MobileNetV2 serves as the CNN backbone, extracting high-level spatial features from the images. Its pre-trained layers are frozen to leverage transfer learning, while additional dense layers are added for fine-tuning. Simultaneously, the FNN processes the fuzzified features through dense layers designed to capture non-linear relationships. The outputs of the CNN and FNN are concatenated, forming a unified representation of both spatial and fuzzy features. A final dense layer with a softmax activation function predicts the class probabilities for five potential disease categories.

The hybrid model is trained using a custom data generator, which combines image data and fuzzified features as input, optimizing the categorical cross-entropy loss function with the Adam optimizer. This integration of fuzzy logic and deep learning exploits both human-like interpretability of fuzzy systems and the automated feature learning capability of CNNs, resulting in a robust framework for disease classification.

1. Implementation

The implementation of the proposed system involves two distinct approaches: a traditional CNN-based model without fuzzy logic, and a hybrid model integrating fuzzy logic into the detection pipeline. This section provides an overview of the coding and algorithmic workflows for both systems, along with the steps undertaken during their development. Screenshots of the code and algorithms are added here to provide a visual representation of the implementation process. Additionally, MATLAB was utilized to develop and simulate the fuzzy inference system, and relevant images from MATLAB will be included to complement this section. The non-fuzzy CNN model utilized the MobileNetV2 architecture pre-trained on ImageNet as a feature extractor. A custom classification head was added to adapt the model for the cassava disease dataset. The implementation was carried out in Python using TensorFlow and Keras.

The dataset was preprocessed using ImageDataGenerator for augmentation, which included techniques like rescaling, rotation, zooming, and flipping. Screenshots of the dataset preprocessing code will be included to demonstrate the augmentation pipeline. The transfer learning workflow involved freezing the MobileNetV2 base and training only the custom layers. The model was compiled with the Adam optimizer and categorical cross-entropy loss function, and callbacks like early stopping and model checkpointing were employed to improve training efficiency.

The fuzzy logic-based system was designed to complement CNN by handling uncertainties and imprecise data during image analysis.

Fuzzy image processing techniques were applied to enhance the input images for the CNN or to refine the classification outputs. Screenshots of the MATLAB code for fuzzy image enhancement and segmentation will be included. This step involved defining fuzzy membership functions for features like intensity, texture, and color to identify diseased regions more effectively.

The fuzzy rules in the system were designed to classify symptoms of cassava diseases based on linguistic variables.

A graph of color intensity

Description automatically generated

**Fig 4.1: Color Intensity Membership Function**

**A graph of roughness and roughness

Description automatically generated**

**Fig 4.2: Texture Roughness Membership Function**

MATLAB was extensively used to visualize the fuzzy logic system. Images showing:  
- Membership functions for inputs like leaf color intensity, and texture irregularity.  
- Surface plots representing the decision-making process of the fuzzy inference system.

By combining the strengths of traditional CNNs and fuzzy logic, this implementation demonstrates two distinct approaches to cassava disease detection. The non-fuzzy CNN provides a robust baseline, while the fuzzy logic system addresses challenges like imprecision and uncertainty in real-world agricultural datasets. The screenshots and MATLAB visualizations included in this section illustrate the practical application of these methodologies, showcasing their potential for real-time disease detection in cassava farming.

A colorful graph on a grid

Description automatically generated

**Fig 4.3: Fuzzy Inference Surface**

1. Testing and results

This section presents the evaluation of both the CNN model without fuzzy logic and the fuzzy logic-integrated system for cassava disease detection. The performance of each system was assessed using testing datasets, and confusion matrices were generated to visualize classification accuracy and errors. A comparative analysis of both systems is provided, highlighting their strengths and weaknesses. Additionally, the fuzzy logic system is compared to related work discussed in the literature review, emphasizing its contributions and improvements over previous methods.

The CNN model without fuzzy logic was tested on the dataset containing images of cassava leaves with different diseases. The model's performance was evaluated based on metrics such as accuracy, precision, recall, and F1-score. A confusion matrix was generated to summarize the classification results.

Screenshots of the confusion matrix for the CNN model will be included. The matrix illustrates the true positive, false positive, and false negative rates for each class (e.g., healthy, Cassava Mosaic Disease, Cassava Brown Streak Disease).

The non-fuzzy CNN achieved an overall accuracy of approximately 67%, with strong performance in identifying healthy plants but occasional misclassifications between diseases with similar visual symptoms.

A close-up of a plant

Description automatically generated

**Fig 5.1: Traditional System Predictions**

The fuzzy logic-enhanced system was evaluated using the same testing dataset. The fuzzy logic component was applied during preprocessing to improve the quality of input features or post-classification to refine the predictions. The system's performance metrics were calculated, and a confusion matrix was generated for comparison.

Screenshots of the confusion matrix for the fuzzy logic system will be included. The results show improved classification accuracy, particularly in cases where symptoms were ambiguous or overlapping between disease classes.

The hybrid system achieved an overall accuracy of approximately 92%, outperforming the non-fuzzy CNN model by a noticeable margin. The fuzzy logic system demonstrated higher precision and recall in detecting diseases with subtle or variable symptoms, highlighting its ability to handle uncertainty effectively.

The fuzzy logic-integrated system outperformed the non-fuzzy CNN in terms of overall accuracy and robustness.

A close-up of a plant

Description automatically generated

**Fig 5.2: Fuzzy System Predictions**

The fuzzy logic system excelled in managing uncertainty and imprecision, particularly in borderline cases where disease symptoms were not distinctly visible. Its rule-based approach provided interpretability, allowing a clearer understanding of how classifications were made.

The non-fuzzy CNN struggled with misclassifications between diseases with similar symptoms, such as Cassava Mosaic Disease and Cassava Brown Streak Disease, due to its rigid feature extraction process.

The fuzzy logic system's performance was benchmarked against similar methods discussed in the literature review. Previous studies have explored both machine learning and fuzzy logic-based approaches for plant disease detection.

Traditional CNN-based systems reported high accuracy but encountered challenges in handling noisy datasets and required extensive computational resources. The fuzzy logic-enhanced system in this study demonstrated improved resilience to such challenges by integrating uncertainty-handling mechanisms

1. Conclusion

This work explored the use of fuzzy logic as a novel approach for Cassava Disease detection and classification. Fuzzy image processing successfully addressed the uncertainty and variability in the manifestation of plant disease symptoms, thereby highlighting the power of fuzzy logic in this domain. Through a comparative analysis with conventional machine learning methods, it can be said that the fuzzy logic approach proved to be a more resource-efficient and competitive alternative, especially when data and computing power are limited.

While the results are encouraging, this study also opens more possibilities for future research. Future research could focus on expanding the dataset to include different environmental conditions, integrating hybrid approaches that combine fuzzy logic with other AI methods and optimizing the fuzzy inference system which will increase the systems accuracy and robustness.

Overall, this research shows the significant potential of fuzzy logic to transform agricultural disease management, promoting sustainable and durable farming practices.

1. References

[1] Hamed, Bahaa & Hussein, Mahmoud & Mousa, Afaf. (2023). Plant Disease Detection Using Deep Learning. International Journal of Intelligent Systems and Applications. 15. 38-50. 10.5815/ijisa.2023.06.04.

[2] Nedeljkovic, I. (2004). Image classification based on fuzzy logic. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 34.

[3] W. E. Sari, Y. E. Kurniawati, and P. I. Santosa, “Papaya Disease Detection Using Fuzzy Naïve Bayes Classifier,” in 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2020, Institute of

Electrical and Electronics Engineers Inc., Dec. 2020, pp. 42–47. doi: 10.1109/ISRITI51436.2020.9315497.

[4] I. J.Rudas, Recent advances in mathematical and computational methods : Proceedings of the 17th International Conference on Mathematical and Computational Methods in Science and Engineering (MACMESE ’15), Kuala Lumpur, Malaysia, April 23-25, 2015. WSEAS Press, 2015.

[5] 2nd International Conference on Signal Processing and Integrated Networks (SPIN) 2015: 19-20 February, 2015, Amity School of Engineering and Technology, Noida, India. IEEE, 2015.

[6] H. Kaur, D. Prashar, and V. Kumar, “Disease Identification in Papaya Plant and their Dataset,” in Proceedings of 5th International Conference on Contemporary Computing and Informatics, IC3I 2022, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1220–1224. doi: 10.1109/IC3I56241.2022.10072453.

[7] J. B. Awotunde et al., “Plant Disease Diagnosis and Detection using Type-2 Fuzzy Logic System,” in 2023 International Conference on Science, Engineering and Business for Sustainable Development Goals, SEB-SDG 2023, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/SEB-SDG57117.2023.10124608.

[8] N. A. Kishore, “Plant Disease Detection Using Fuzzy Classification,” 2021. [Online]. Available: http://annalsofrscb.ro

[9] T. H. Kim et al., “ANFIS Fuzzy convolutional neural network model for leaf disease detection,” Front Plant Sci, vol. 15, 2024, doi: 10.3389/fpls.2024.1465960.

[10] V. Mahajan and N. R. Dhumale, “Engineering and Technology (A High Impact Factor,” International Journal of Innovative Research in Science, vol. 7, 2018, doi: 10.15680/IJIRSET.2018.0706067.

[11] M. Sibiya and M. Sumbwanyambe, “Automatic fuzzy logic-based maize common rust disease severity predictions with thresholding and deep learning,” Pathogens, vol. 10, no. 2, pp. 1–17, Feb. 2021, doi: 10.3390/pathogens10020131.

[12] S. S. Chouhan, U. P. Singh, and S. Jain, “Automated Plant Leaf Disease Detection and Classification Using Fuzzy Based Function Network,” Wirel Pers Commun, vol. 121, no. 3, pp. 1757–1779, Dec. 2021, doi: 10.1007/s11277-021-08734-3.

[13] M. Sowmya and B. Subramani, “Disease detection on plant leaf using K-means segmentation with fuzzy logic SVM algorithm,” Int J Health Sci (Qassim), pp. 4749–4756, Apr. 2022, doi: 10.53730/ijhs.v6ns2.612.