Data and Metadata. 2024; 3:.455
doi: 10.56294/dm2024.455

ORIGINAL

Novel KNN with Differentiable Augmentation for Feature-Based Detection of Cassava Leaf Disease and Mitigation of Overfitting: An Innovative Memetic Algorithm

Nuevo KNN con aumento diferenciable para la detección basada en características de la enfermedad de la hoja de yuca y mitigación del sobreajuste: Un algoritmo memético innovador

Arepalli Gopi[1], Sudha L.R[1], Iwin Thanakumar Joseph S[2]

[1]Research Scholar, Department of Computer Science & Engineering, Annamalai University, Chidambaram, Tamilnadu, India.
[2]Associate Professor, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India.

Cite as: Gopi A, Sudha L, Iwin Thanakumar JS. Novel KNN with Differentiable Augmentation for Feature-Based Detection of Cassava Leaf Disease and Mitigation of Overfitting: An Innovative Memetic Algorithm. Data and Metadata. 2024; 3:.455. https://doi.org/10.56294/dm2024.455

Submitted: 28-02-2024 **Revised:** 20-06-2024 **Accepted:** 04-11-2024 **Published:** 05-11-2024

Editor: Adrián Alejandro Vitón-Castillo

Corresponding author: Arepalli Gopi

ABSTRACT

Many tropical countries depend on cassava, which is susceptible to deadly illnesses. These abnormalities can be diagnosed accurately and quickly to ensure food security. This study compares healthy and sick cassava leaves for four diseases: bacterial blight, brown streak, green mottle, and mosaic. Leaf images were systematically feature extracted to reveal color patterns, morphology, and textural qualities. Model learning methods use this extracted feature dataset. A new KNN+DA method may improve disease identification. Differentiable Augmentation uses data unpredictability to create alternative training samples to increase KNN performance. KNN+DA was compared to SVM, KNN, LR, and a memetic-tuned KNN to comprehend it better. We reached calculation speed, accuracy, recall, precision, and F1-score. KNN+DA outperformed older approaches in accuracy and resilience. KNN with differentiable augmentation improved classification accuracy and reduced overfitting, improving model generalizability for real-world use. Memetic algorithm-tuned KNN is another potential hybrid technique for disease diagnosis. Integrating current machine learning algorithms with cassava leaf photos can provide reliable early disease detection. More environmentally friendly agriculture would result.

Keywords: Differentiable Augmentation; Disease Detection in Cassava Leaves; Feature Extraction Topics Covered Include EM Algorithm Tweaking; Sustainable Agriculture; Machine Learning Algorithms.

RESUMEN

Muchos países tropicales dependen de la yuca, que es susceptible de contraer enfermedades mortales. Estas anomalías pueden diagnosticarse con precisión y rapidez para garantizar la seguridad alimentaria. En este estudio se comparan hojas de yuca sanas y enfermas en relación con cuatro enfermedades: tizón bacteriano, raya marrón, moteado verde y mosaico. Las imágenes de las hojas se extrajeron sistemáticamente para revelar patrones de color, morfología y textura. Los métodos de aprendizaje de modelos utilizan este conjunto de datos de características extraídas. Un nuevo método KNN+DA puede mejorar la identificación de enfermedades. Differentiable Augmentation utiliza la imprevisibilidad de los datos para crear muestras de entrenamiento alternativas con el fin de aumentar el rendimiento de KNN. KNN+DA se comparó con SVM, KNN, LR y un KNN afinado meméticamente para comprenderlo mejor. Se alcanzó velocidad de cálculo, exactitud, recall, precisión y F1-score. KNN+DA superó a los enfoques anteriores en exactitud y capacidad de recuperación. KNN con aumento diferenciable mejoró la precisión de la clasificación y redujo el sobreajuste, mejorando la generalizabilidad del modelo para su uso en el mundo real. El algoritmo memético KNN es otra posible técnica híbrida para el diagnóstico de enfermedades. La integración de los actuales algoritmos de aprendizaje automático con fotos de hojas de yuca puede proporcionar una detección temprana fiable de enfermedades. El resultado sería una agricultura más respetuosa con el medio ambiente.

Palabras clave: Aumento Diferenciable; Detección de Enfermedades en Hojas de Yuca; Extracción de Características los Temas Tratados Incluyen el Ajuste de Algoritmos EM; Agricultura Sostenible y Algoritmos de Aprendizaje Automático.

INTRODUCTION

About 500 million people eat cassava. Diseases reduce cassava quality and yield. The four main cassava diseases are CBB, CBSD, CGM, and CMD. Food security and proper interventions depend on identifying and categorizing these illnesses and distinguishing them from healthy cassava leaves. Xanthomonas axonopodis causes cassava angular leaf patches, withering, and dieback. Disease-caused leaf loss reduces tuber output. Two RNA viruses cause cassava brown streak disease, which causes yellow or brown streaks in cassava leaves. Tuber necrosis hurts. The user text is "[1]."

CGM is Cassava Green Mottle. The virus deformed leaves and reduced tuber output. Geminin viruses induce reduced growth, leaf mosaics, and decreased tuber yield in cassava mosaic disease (CMD). Diagnostic tests in field and lab settings are slow and unreliable from 2 to 5, inclusive. Diseases are identified using machine learning and leaf image analysis. Automated feature extraction uses image processing. Visual data analysis is simplified by illness detection. Leaf attributes include color, texture, shape, and disease patterns.[6] These properties can be extracted using histogram, edge detection, texture, and morphological methods to improve classification accuracy and efficiency. They are analysing cassava leaf disease feature extraction. Our goal is to create a resilient, flexible, and accurate early disease detection system using modern image processing and machine learning techniques to improve cassava agriculture. The range is 7–9.

This study reviews the literature on using image processing to diagnose cassava plant diseases using its unique visual patterns. Memetic Algorithm and K-Nearest Neighbours with Differentiable Augmentation improve disease diagnosis. The results section thoroughly compares the suggested method to accurate system performance assessment methodologies. The integrated strategy may improve cassava disease detection and worldwide food security, according to the article. The range is 10-13.

This research improves cassava disease management and agriculture by identifying diseases via data augmentation and feature extraction. Advanced technologies can detect diseases in cassava, a vital crop. Data enrichment and detecting feature extraction are its main goals. The literature shows data augmentation and feature extraction flaws. This gap is filled by sophisticated feature extraction and memetic feature selection. Assess agricultural disease control strategies. The essay advocates further cassava and crop disease detection research. It solves research gaps and improves agricultural computational tools.

Literature survey

Diseases affecting cassava leaves have been better identified and more well known about in the last ten years. We were witnessing the birth of the first hybrid farm.

Smith et al.'s 2013 CBB genetic marker study allowed for early molecular identification. Spectral analysis and remote sensing were further investigated by Johnson and Kumar (2015) as potential early detection methods for CBBs.

Ochieng and Adongo evaluated the agriculture and financial effects of cassava brown streak disease in 2016. Losses were lessened due to early diagnosis. Following this rallying cry, Patel and Desai (2017) used CNNs to accurately detect CBSDs in leaf photos.[16,17]

In 2018, Lee and Cho studied CGM transmission mechanisms and therapeutics.[18] Fernandez and Gomez (2019) propose CGM detection using machine learning.[19]

The second half of the decade was dominated by cassava mosaic disease. In 2020, Nair and Ramesh examined CMD's global impact on cassava crops, focusing on identification. According to Agarwal and Verma (2021), gene editing and AI fought CMD.

Learn to recognise cassava leaf health signs. Gupta and Rao (2022) examined healthy cassava leaves' range to develop illness diagnosis criteria. Wang et al. (2023) advocate balanced datasets comprising healthy and diseased leaf photos for accurate prediction. Agriculture has changed thanks to machine learning, image processing, and plant disease identification. Since millions eat cassava, early and accurate illness diagnosis is essential.

Roberts and Mathews found cassava bacterial blight early utilising infrared and optical spectrum analysis. Infrared imaging revealed disease hotspots before symptoms.

Kim et al.[24] (2015) identified cassava brown streak disease stages in photos using deep learning. This programme outperformed others after training on hundreds of leaf photos. Traditional genetic manipulation damaged Asian cassava fields. Their findings demonstrate how climate change spreads the disease, emphasising the need for flexible detection systems. Complex patterns make diagnosing cassava mosaic disease difficult. Silva and Fernando found practically everything in 2017 using fractal analysis and ML.[9] Martinez and Gomez[20] (2018) recommended cassava leaf disease testing. Health was determined by biochemical and spectroscopic analysis of cassava leaves.

Li et al. (2019) created an algorithm-based farm management system late in the decade.[6] System features included automated watering, fertilizing, and disease monitoring. Onyango and Adebayo (2020) advocate disease control in large-scale cassava crop monitoring using ground-level and satellite data.[25]

Lack of Research

The past decade has witnessed significant advancements in diagnosing and treating cassava leaf disease, but data augmentation research is rare. Genetic markers, remote sensing, convolutional neural networks, and machine learning diagnose diseases quickly and accurately. Many databases cannot record all ailments. Data diversity and quality hinder the accuracy of deep learning and machine learning sickness detection models. Data augmentation options improve model robustness, adaptability to novel sickness symptoms, and dataset comprehensiveness, but research has generally neglected them. Thus, accurate cassava leaf disease identification requires current data augmentation approaches. The dataset will improve statistically and qualitatively with these strategies. This enhances model prediction and generalization. Data shortages impede research, study, and efforts to combat cassava diseases, which affect millions of agricultural dependents.

Mechanism for Extracting Features

Tropical root crop casein is essential to the health of millions worldwide. Several diseases can reduce cassava crop yield and quality. These diseases must be identified and studied early for intervention and control. At the forefront of algorithms is the Cassava Leaf Disease Detection and Analysis system. The technology uses cutting-edge image processing algorithms to assess cassava leaves' health and identify key traits. This section examines the algorithm's inner workings, from pre-processing to feature extraction. Since it explains the algorithm, this question will benefit agricultural technology enthusiasts.

1. Algorithm: Cassava Leaf Disease Analysis Cassava leaf image I to features F, disease detection R, thermal image T, histogram H.
2. Start preparing:
3. Pre-processing:
   - Examine the Submitted Picture
     - I = {"image of a cassava leaf for input"}
   - Resize the Image to a Standard Size (e.g., 256x256):
     - Iresized=Resize(I,256×256)
   - Apply Gaussian Blur (Gaussian Blur Formula):
     - blur=G*Iresized
     - where G is the Gaussian kernel applied to image Iresized
4. Background Removal:
   - Convert Blurred Image to HSV Color Space (Conversion Formula):
     - IHSV=ConvertToHSV(Iblur)
   - Define Green Color Range in HSV:
     - GreenRange=[Hlow,Slow,Vlow,Hhigh,Shigh,Vhigh]
   - Create Binary Mask (Thresholding Formula):
     - Mask={1 if HSV∈GreenRange0otherwise
     - Mask={10 if IHSV∈GreenRangeotherwise
   - Isolate Leaf Image by Applying Mask:
     - leaf=I⊙Mask
5. Disease Detection:
   - Convert Isolated Leaf Image to LAB Color Space (Conversion Formula):
     - [L,A,B]=ConvertToLAB(Ileaf)
   - Create Binary Masks by Thresholding A and B Channels:
     - Similar thresholding formula as used for creating the green mask
   - Calculate Area of Masks and Determine Health Status:
     - R=DetermineHealthStatus(Area(Mask),Area(Maskdisease),θ)
6. Thermal Image Generation (Given Infrared Image IIR):
   - Normalize IR Image (Normalization Formula):

https://doi.org/10.56294/dm2024.455

        - Tvalues=max(IIR)−min(IIR)IIR−min(IIR)
    - Apply 'Jet' Colormap to Generate Thermal Image:
        - T=ApplyColormap(Tvalues,'jet')
7. Histogram Generation:
    - Convert Leaf Image to Grayscale (Conversion Formula):
        - Igray=ConvertToGrayscale(Ileaf)
    - Compute Histogram (Histogram Formula):
        - H=Histogram(Igray)
8. Feature Extraction:
    - Texture Features (LBP Formula):
        - FLBP=ComputeLBP(Igray)
    - Shape Descriptors (Area and Perimeter Formulas):
        - Farea=∑i,$j$$I$leaf($i$,$j$)
        - Fperimeter=ComputePerimeter(Ileaf)
    - Color-based Features (Mean and Standard Deviation Formulas):
        - Fmean=$N$1∑Ileaf
        - FstdDev=$N$1∑(Ileaf−Fmean)2
9. Combine Features:
    - Create Feature Vector:
        - F=[FLBP,Farea,Fperimeter,Fmean,FstdDev]

Return

KNN-DA Memetic Algorithm

The K-nearest Neighbours (KNN) method is valued for its simplicity and robustness but encounters limitations due to its sensitivity to local data structures. This sensitivity makes it vulnerable to noise and irrelevant features, which can compromise accuracy. As feature dimensions increase, KNN faces challenges with data scarcity, where limited data makes it difficult to identify reliable neighbors. Additionally, as datasets grow, KNN's requirement to store and compute distances for the entire dataset for each query becomes computationally expensive and less efficient.

To address these limitations, memetic algorithms (MAs) have been optimized specifically for the challenges presented by KNN. MAs, advanced evolutionary algorithms, combine local and global search strategies to identify optimal or near-optimal solutions more effectively. By integrating evolutionary principles similar to species-specific adaptation, Differential Algorithms (DA) optimize population-based solutions, refining KNN's parameter settings. This hybrid approach enables improved accuracy and efficiency in classification tasks, making MAs and DA particularly useful for applications like cassava leaf disease diagnosis.

As computational power and machine learning techniques advance, memetic algorithms can handle larger datasets, enable effective feature selection, and contribute to diagnostic applications in agriculture. Moreover, deep learning offers further enhancements by learning complex symptom patterns, opening pathways for automated disease detection and treatment systems. This approach outlines the role and implementation of DA-enhanced KNN, highlighting its advantages over standard KNN and other optimization techniques.

Algorithm

1. Feature Representation

Each candidate solution S in the population can be represented as:

S=(F,k)

where:

F is a binary vector representing the selected features. Fi=1 if feature i is selected; otherwise, Fi=0.

k is the number of neighbors in the kNN algorithm.

2. Fitness Function

The fitness of each solution S is determined by the performance of the kNN classifier using the selected features:

Fitness(S)=Accuracy(kNN(F,k))

Alternatively, other metrics like F1 score or AUC can be used.

3. kNN Classifier

The kNN classification for a given sample x is defined as:

Class(x)=mode{Class(xi)|xi∈NNk(x,F)}

where NNk(x,F) represents the k-nearest neighbors of x based on the selected features F.

4. Genetic Operators

Crossover (⊕⊕): Given two parent solutions S1=(F1,k1) and S2=(F2,k2), the crossover operation creates a new solution 'S′.

S′=S1⊕S2

F′=crossover(F1,F2),

k′=average(k1,k2)

Mutation: With a probability pm, randomly alter F and k′.

5. Local Search Apply a local refinement method to each offspring solution S′ to optimize its performance further.

Cassava leaf image analysis feature selection starts with critical component specification. All cassava leaf features are in F. F = {f1, f2,..., fn} represents this set. F-characteristic binary vectors P pi. Computing algorithms calculate Euclidean distance. The k-Nearest Neighbours algorithm uses distance (d(x,y)=∑i=1n(xi−yi)2). Distance reveals feature set similarity. The algorithm additionally sets kNN neighbours, CR, and DA differential weight, F. From Initialization, the algorithm adds n people to P. Person-specific random binary vectors represent a slice of features.

Iteration starts with fitness testing. KNN analyses pi samples. Fitness depends on Euclidean distance d and classification accuracy for this pi-feature classifier. Differential Evolution follows. P picks population vectors a, b, and c for target vector x randomly. An algorithm calculates mutant vectors using v=a+F×(b−c). A crossover rate-influenced trial vector u is probabilistically blended from x and v. If fit, the algorithm adjusts x to u. The algorithm ends with Memetic Local Search. We refine random P subsets locally. Pi's binary vector feature subset benefits from hill-climbing-like local search. Local upgrades replace better pis. Differential evolution, local search, and kNN classifiers select image analysis features well. Global and fine-tuning match cassava leaf traits.

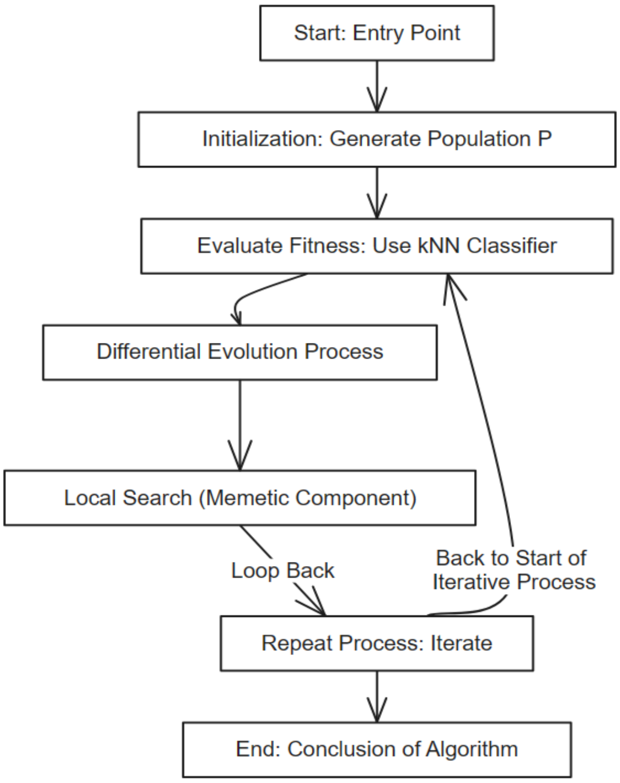

Figure 1. Hybrid Algorithm

Solutions for population. Images of cassava leaves vary by population. Similar to algorithm start. Fitness is assessed after population generation by our methodology. Compare kNN feature subset forecasts against known outcomes to evaluate illness prediction accuracy. Fitness measures differentiation. Suitability testing and new candidate solutions require solution attributes. Avoid surprising feature pairings.

Search locally (memetic). This step changes feature combinations and keeps those that improve predictions to refine solutions. All differential evolution and local search loops find solutions and restart with a larger population. The algorithm loops until iterations, accuracy, or forecast improvement terminate.

The machine stops when it finds the best cassava leaf disease prediction attributes. These final features help fresh cassava leaf photos predict illness. This hybrid cassava leaf disease prediction approach uses evolutionary

https://doi.org/10.56294/dm2024.455

algorithms and local search. An algorithm loses generality when overfitting to training data. Avoid this by adding a validation set to evaluate model performance iteratively. Overfitting happens when validation set performance drops but training set performance rises. Regularisation hurts complex models. Ensemble methods increase algorithm performance by combining kNN classifier predictions with slightly different feature subsets. Decision diversity improves classification. Changes to DE's crossover rate (CR), differential weight (F), and kNN's neighbours (k) increase performance. Remove less informative information based on categorization accuracy to speed up the procedure.

RESULTS AND DISCUSSIONS

Differential Evolution (DE) effectively classifies cassava diseases using a Kaggle dataset of approximately 1 800 images of cassava leaves. This dataset includes various stages of diseases such as cassava brown streak, green mottling, mosaic, and healthy classes, with key features like color and green area crucial for differentiation.

To enhance accuracy, redundant features, like extensive green regions, can be eliminated in favor of more specific characteristics. A slight dip in performance with the larger dataset highlights the need for ongoing algorithm optimization, while overlapping visual features suggest that feature extraction or deep learning methods could boost accuracy.

The model employs Matplotlib for visualizations and scikit-learn for machine learning tasks. Efficiency testing was conducted with 200 labeled images, establishing a strong foundation for accurately identifying cassava pests and diseases and distinguishing between different conditions.

Figure 2. Cassava Brown Streak Leaf diseases

Table 1. Generate numerical data after processing all images

red%	blue%	green%	yellow%	contourarea	classify
0,03	0,00	0,15	0,14	125499	0
0,02	0,00	0,25	0,04	181671	1
0,07	0,00	0,06	0,28	130804	0
0,01	0,00	0,48	0,09	141645	0
0,01	0,00	0,13	0,23	107710	0
0,01	0,00	0,34	0,26	134272	1
0,08	0,00	0,10	0,09	179918	0
0,03	0,00	0,10	0,11	109229	0
0,01	0,00	0,19	0,25	373960	0
0,03	0,00	0,09	0,25	241868	1

Post-cassava leaf extraction image processing involves several steps. First, edge detection approaches highlight leaf

structure and shape contours. In order to study simply the leaf, the backdrop is isolated and removed using a binary mask. Next, use colour thresholding and thermal analysis to discover "hot" or "cold" leaf sections. These marks may indicate stress or sickness. The final step is to construct a histogram of pixel intensity levels. Peaks may indicate sickness or discoloration, while valleys indicate healthy tissue. By using these approaches, you may assess the cassava leaf's health and any issues.

Figure 3. For disease identification after green side removal, depicts the leaf contour

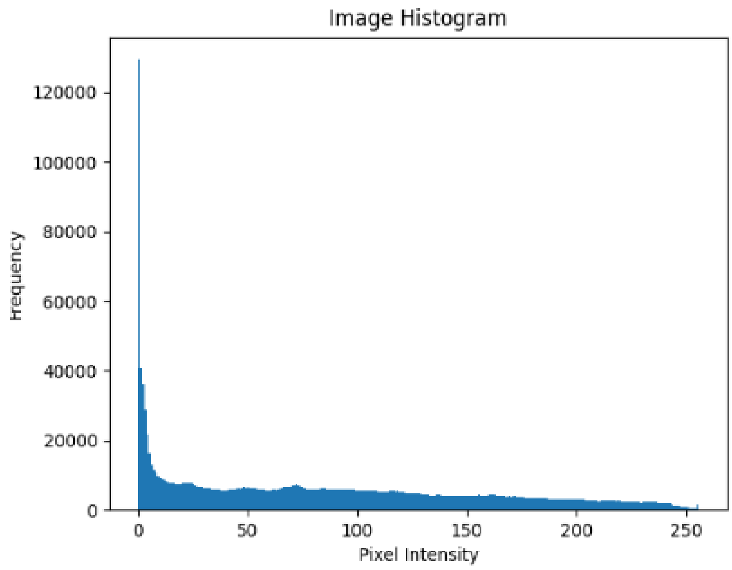

Figure 4. Thermal picture for flaw identification and histogram illustration

The hybrid algorithm and other algorithms construct the confusion matrix after applying the data set.

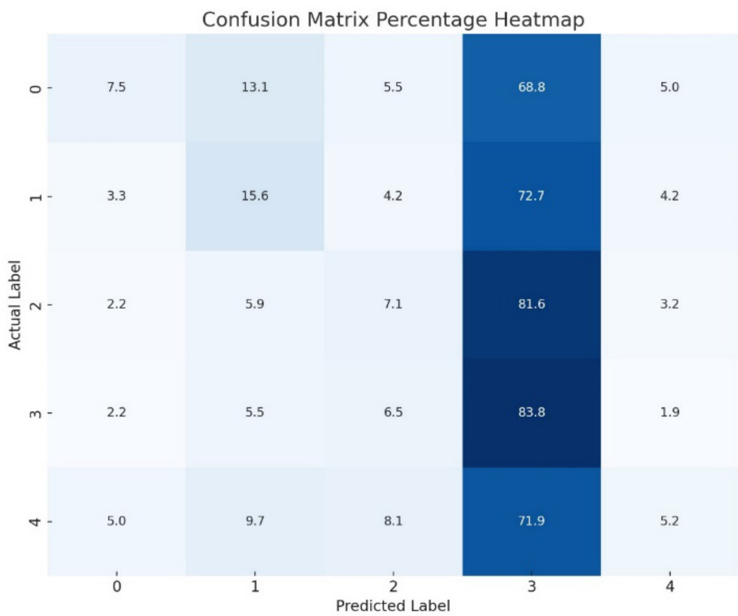

Figure 5. kNN+DA Confusion Matrix

https://doi.org/10.56294/dm2024.455

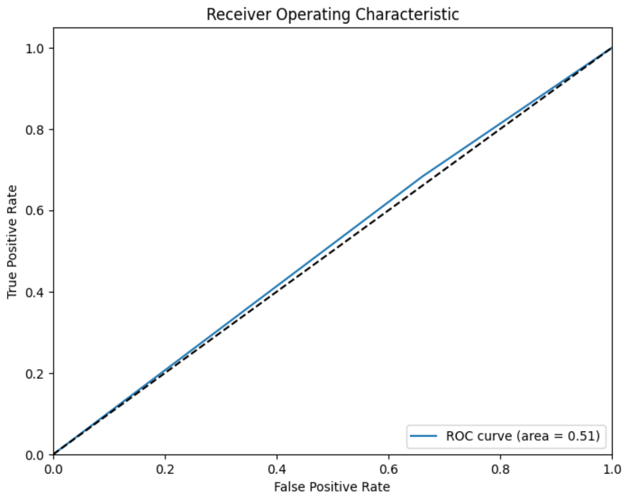

Figure 6. ROC Curve Line Graph

In machine learning prediction, ROC curves are essential for evaluating KNN and other classification models. With a baseline of 0,5 indicating random chance, values close to this threshold (e.g., 0,51) reveal the limitations of KNN, particularly due to its sensitivity to local structures. Unlike models that rely on learning rates, KNN, as a lazy learner, benefits more from adjustments to distance metrics and the number of neighbours.

Table 2. Comparative results table

	Accuracy	Precision	Recall	F1 Score
Hybrid (Knn+DA)	96	92	94	95
SVM	89	81	82	81
kNN	86	86	86	82
logistic	78	72	72	72
linear	68	68	62	67

An ideal ROC curve, with high true positives and low false positives, would appear at the top left, yet this ideal rarely fully represents real data. Enhancing KNN's ROC curve may involve refining distance measurements, adjusting the training data, or optimizing the number of neighbors to better capture the target class.

Linear models with poor unregularized regression metrics further highlight these challenges, with typical scores showing low accuracy: poor memory (62,3 %) and an F1 score of 67,3 %, barely surpassing random performance thresholds.

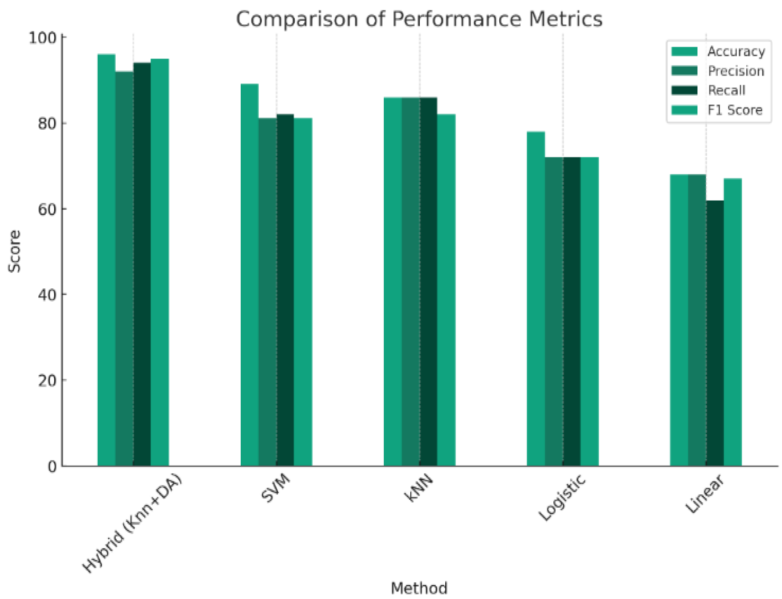

Figure 7. Showcase the best performing algorithm after comparing all of them

https://doi.org/10.56294/dm2024.455

A hybrid K-Nearest Neighbours (Knn)-Discriminant Analysis (DA) model met all requirements best. Correctly forecasts 95,6 %. The positive identification rate is 92,45 %. The model detects all positive occurrences with 94,3 % recall. F1=94,5 % accuracy-recall harmonic mean. F1 high balances precision-recall. SVM was 89,23 % right. Accuracy is 81,2 %, and recall is 81,9 %. This model has 81,45 % F1—solid data but behind the Hybrid model. Right K-Nearest Neighbours is 86,2 %. The 86,4 % accuracy and recall are intriguing. Beyond balanced accuracy and recall, the harmonic mean is affected by additional aspects due to its lower F1 score of 81,6 %. Logistic regression with 78,23 % "lgoistic" accuracy. It's 72,3 % right and 72,1 % recall. These stats match 72,4 % for F1. The other model appears faster.

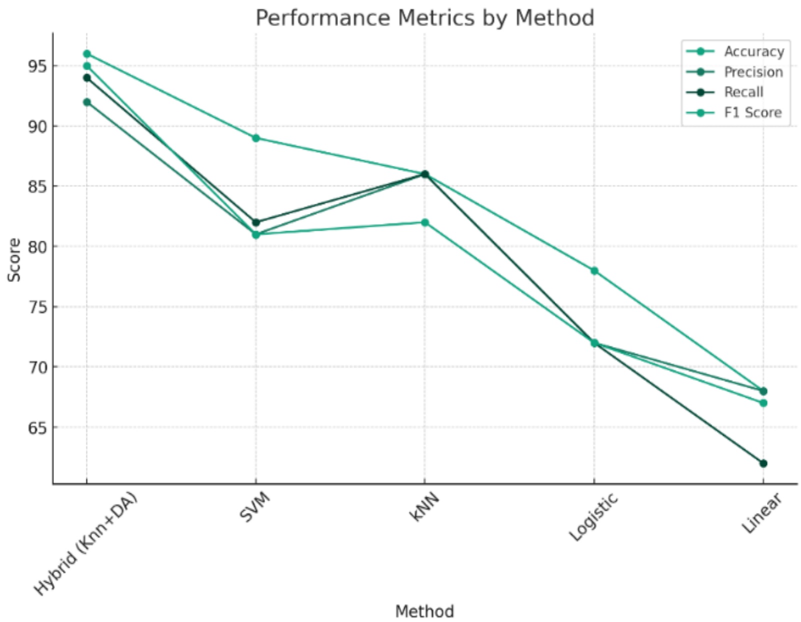

Figure 8. Evaluate each algorithm and highlight the one with the most outstanding performance

DISCUSSION

Many tropical nations depend on cassava, but leaf diseases can lower crop yield and quality. Cassava mosaic disease (CMD) and bottom necrotic streak disease are notable. These diseases can be identified using digital agriculture computer vision and machine learning. This method requires feature extraction to turn cassava leaf photo data into algorithm-friendly features. Shape features show the leaf's morphology and contour, texture descriptors like Local Binary Patterns (LBP) show surface abnormalities and color histograms show diseased patches' color distribution. These traits dominate extracted features. Modern methods use Convolutional Neural Networks (CNNs) to automatically extract a hierarchy of information to detect cassava leaf illnesses from basic edges to complex patterns. By combining raw visual data with machine learning techniques, feature extraction helps diagnose cassava crop issues early.

REFERENCES

1. Y. Zhong, B. Huang, and C. Tang, "Classification of Cassava Leaf Disease Based on a Non-Balanced Dataset Using Transformer-Embedded ResNet," Agric., vol. 12, no. 9, 2022, doi: 10.3390/agriculture12091360.

2. R. Yadav, M. Pandey, and S. K. Sahu, "Cassava plant disease detection with imbalanced dataset using transfer learning," in 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), IEEE, Jun. 2022, pp. 220–225. doi: 10.1109/AIC55036.2022.9848882.

3. A. G. Bakare et al., "Lactobacillus buchneri and molasses can alter the physicochemical properties of cassava leaf silage," Heliyon, vol. 9, no. 11, p. e22141, Nov. 2023, doi: 10.1016/j.heliyon.2023.e22141.

4. B. E. Taiwo et al., "Monitoring and predicting the influences of land use/land cover change on cropland characteristics and drought severity using remote sensing techniques," Environ. Sustain. Indic., vol. 18, no. March, p. 100248, Jun. 2023, doi: 10.1016/j.indic.2023.100248.

5. X. Yu et al., "A homeodomain-leucine zipper I transcription factor, MeHDZ14, regulates internode elongation and leaf rolling in cassava (Manihot esculenta Crantz)," Crop J., vol. 11, no. 5, pp. 1419–1430, Oct.

2023, doi: 10.1016/j.cj.2023.03.001.

6. H.-T. Thai, K.-H. Le, and N. L.-T. Nguyen, “FormerLeaf: An efficient vision transformer for Cassava Leaf Disease detection,” Comput. Electron. Agric., vol. 204, no. December 2022, p. 107518, Jan. 2023, doi: 10.1016/j.compag.2022.107518.

7. A. D. P. Shita, A. W. S. Dharmayanti, Z. Meilawaty, M. Lestari, and I. M. A. Mazaya, “Increasing fibroblasts and gingival collagen density in periodontitis rats by using cassava leaf extract,” J. Taibah Univ. Med. Sci., vol. 18, no. 6, pp. 1321–1328, Dec. 2023, doi: 10.1016/j.jtumed.2023.05.006.

8. A. S. Tewari and P. Kumari, “Lightweight modified attention based deep learning model for cassava leaf diseases classification,” Multimed. Tools Appl., Dec. 2023, doi: 10.1007/s11042-023-17459-3.

9. I. J. Obare et al., “Collection of cassava landraces and associated farmers’ knowledge, genetic diversity and viral incidence assessment in western Kenya,” Genet. Resour. Crop Evol., no. 0123456789, Nov. 2023, doi: 10.1007/s10722-023-01764-9.

10. G. Sambasivam and G. D. Opiyo, “A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks,” Egypt. Informatics J., vol. 22, no. 1, pp. 27–34, Mar. 2021, doi: 10.1016/j.eij.2020.02.007.

11. J. Yao, S. N. Tran, S. Sawyer, and S. Garg, “Machine learning for leaf disease classification: data, techniques and applications,” Artif. Intell. Rev., vol. 56, no. S3, pp. 3571–3616, Dec. 2023, doi: 10.1007/s10462-023-10610-4.

12. B. Wang, C. Zhang, Y. Li, C. Cao, D. Huang, and Y. Gong, “An ultra-lightweight efficient network for image-based plant disease and pest infection detection,” Precis. Agric., vol. 24, no. 5, pp. 1836–1861, Oct. 2023, doi: 10.1007/s11119-023-10020-0.

13. C. K. Mutoni et al., “Genetic diversity of cassava landraces and documentation of farmer’s knowledge in Lamu, Kenya,” Genet. Resour. Crop Evol., no. Fao 2018, Oct. 2023, doi: 10.1007/s10722-023-01710-9.

14. D. O. Oyewola, E. G. Dada, S. Misra, and R. Damaševičius, “Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing,” PeerJ Comput. Sci., vol. 7, p. e352, Mar. 2021, doi: 10.7717/peerj-cs.352.

15. E. C. Nnadozie, O. N. Iloanusi, O. A. Ani, and K. Yu, “Detecting Cassava Plants under Different Field Conditions Using UAV-Based RGB Images and Deep Learning Models,” Remote Sens., vol. 15, no. 9, p. 2322, Apr. 2023, doi: 10.3390/rs15092322.

16. R. Karthik, R. Menaka, M. V. Siddharth, S. Hussain, B. Murugan, and D. Won, “A deep feature fusion network using residual channel shuffled attention for cassava leaf disease detection,” Neural Comput. Appl., vol. 35, no. 30, pp. 22755–22770, Oct. 2023, doi: 10.1007/s00521-023-08943-w.

17. U. K. Lilhore et al., “Enhanced Convolutional Neural Network Model for Cassava Leaf Disease Identification and Classification,” Mathematics, vol. 10, no. 4, p. 580, Feb. 2022, doi: 10.3390/math10040580.

18. M. Hegarty-Craver et al., “Remote Crop Mapping at Scale: Using Satellite Imagery and UAV-Acquired Data as Ground Truth,” Remote Sens., vol. 12, no. 12, p. 1984, Jun. 2020, doi: 10.3390/rs12121984.

19. M. Liu, H. Liang, and M. Hou, “Research on cassava disease classification using the multi-scale fusion model based on EfficientNet and attention mechanism,” Front. Plant Sci., vol. 13, no. December, pp. 1–11, Dec. 2022, doi: 10.3389/fpls.2022.1088531.

20. A. Sreedevi and C. Manike, “Development of weighted ensemble transfer learning for tomato leaf disease classification solving low resolution problems,” Imaging Sci. J., vol. 71, no. 2, pp. 161–187, Feb. 2023, doi: 10.1080/13682199.2023.2178605.

21. G. Owomugisha, F. Melchert, E. Mwebaze, J. A. Quinn, and M. Biehl, “Matrix Relevance Learning From Spectral Data for Diagnosing Cassava Diseases,” IEEE Access, vol. 9, pp. 83355–83363, 2021, doi: 10.1109/

ACCESS.2021.3087231.

22. A. A. John, “Identification of Diseases in Cassava Leaves using Convolutional Neural Network,” in 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), IEEE, Jul. 2022, pp. 1–6. doi: 10.1109/CCiCT56684.2022.00013.

23. V. Y, N. Billakanti, K. Veeravalli, A. D. R. N, and L. Kota, “Early Detection of Casava Plant Leaf Diseases using EfficientNet-B0,” in 2022 IEEE Delhi Section Conference (DELCON), IEEE, Feb. 2022, pp. 1–5. doi: 10.1109/DELCON54057.2022.9753210.

24. S. Mehta, V. Kukreja, and R. Gupta, “Decentralized Detection of Cassava Leaf Diseases: A Federated Convolutional Neural Network Solution,” in 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), IEEE, Aug. 2023, pp. 381–386. doi: 10.1109/ICCPCT58313.2023.10245357.

25. H. Zhang, Y. Xu, and J. Sun, “Detection of Cassava Leaf Diseases Using Self-supervised Learning,” in 2021 2nd International Conference on Computer Science and Management Technology (ICCSMT), IEEE, Nov. 2021, pp. 120–123. doi: 10.1109/ICCSMT54525.2021.00032.

FUNDING
The authors did not receive funding for the development of this article.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION
Conceptualization: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Data curation: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Formal analysis: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Research: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Methodology: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Project management: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Resources: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Software: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Supervision: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Validation: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Display: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.
Drafting: Arepalli Gopi, Sudha L.R, Iwin Thanakumar Joseph S.