2026 Volume: 02 Issue: 01 www.ijrpa.com ISSN 2456-9995 Review Article

International Journal Research Publication Analysis

Page: 01-08

EFFICIENT PREPROCESSING TECHNIQUE FOR ACCURATE AND ROBUST FEATURE EXTRACTION IN MACHINE LEARNING BASED TOMATO LEAF DISEASE DETECTION

Hemalatha B*[1], Jayakumar B L[2], Dr Smitha Kurian,[3] Dr Krishna Kumar P R,[4]

[1]M Tech Student, Department of CSE & Technology, SEA College of Engineering & Technology.

[2, 4]Faculty, Department of CSE, SEA College of Engineering & Technology.

[3]Professor & HoD Department of CSE, HKBK College of Engineering & Technology, Bangalore.

Article Received: 07 December 2025
Article Revised: 27 December 2025
Published on: 15 January 2026

***Corresponding Author: Hemalatha B**
M Tech Student, Department of CSE & Technology, SEA College of Engineering & Technology.

DOI: https://doi-doi.org/101555/ijrpa.1642

ABSTRACT

Crop yield and efficiency in farming are significantly dependent on the early identification and prediction of plant leaf diseases. In Many Times machine learning algorithms have surfaced as potent instruments for mechanizing this procedure, offering farmers a precise and effective way to recognize and handle leaf illnesses. With a focus on early disease prediction before the formation of observable symptoms. In the field of agriculture, sustaining high yields and guaranteeing food security depends on the early diagnosis of diseases in crops like tomato plants.which have demonstrated promise in automating disease diagnosis procedures.

KEYWORDS: *Plant leaf diseases, GPU computing, CUDA-ResNet50 Classifier.*

INTRODUCTION

Achieving sustainability and reducing environmental impact while satisfying the world's food demand presents the agriculture sector with previously unheard-of issues. Thise are one of the important crops grown in The India. They are a main ingredient in many cuisines and a major driver of the agricultural economy. there is a danger to agricultural productivity, quality, and economic stability due to the widespread spread of diseases, especially those that harm tomato leaves. The agricultural industry has experienced a notable increase Conventional disease detection techniques [1] mostly rely on manual examination, which is time-consuming, labor- intensive, and frequently prone to mistakes. An increasing number of people are interested in automating illness diagnosis procedures by utilizing cutting-edge technology like computer vision and parallel computing to get beyond these constraints

PROBLEM DEFINITION

Millions of people throughout the world rely on tomato farming as their main source of food and income, making it an essential component of global agriculture

Tomato leaf disease detection is a challenging problem that impedes agricultural output and efficient disease control. Among these difficulties some of them are:

1. Limitations of Manual Inspection
2. Diversity and Complexity of Diseases
3. Volume and Variability of Data:
4. Resource Constraints

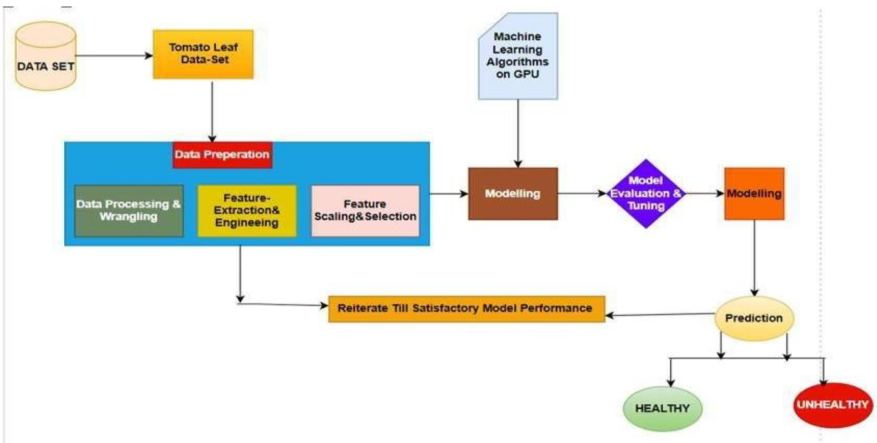


Figure 1: Proposed Model Architecture

LITERATURE SURVEY

Title	Year	Methodology	Key Findings
Tomato Leaf Disease Detection	2018	Deep learning, CNNs, GPU computing	Real-time disease detection was
Using Deep Learning and GPUs			achieved with high accuracy. GPU acceleration significantly improves inference speed.
Parallel SVM-	2018	Image processing,	Achieved 95% accuracy in disease detection using parallel

based Tomato Disease Detection Using Image Processing		Support Vector Machine (SVM)	SVM training
GPU- accelerate d Feature Extraction for Tomato Disease Identification	2018	Feature extraction, K-nearest neighbors (KNN)	Improved classification accuracy by 10% using GPU- accelerated KNN
Accelerating Tomato Plant Disease Detection Using CNNs on GPU	2019	CNNs, GPU computing, optimization techniques	Significant speed up The detection was achieved compared to CPU-based implementations. Various CNN architectures were explored for improved inference speed while maintaining accuracy.
"Distributed Machine Learning for Tomato Leaf Disease Detection"	2019	Genetic algorithm, Feature selection	dentified optimal feature subset for disease detection efficiently
Efficient Tomato Leaf Disease	2020	Deep learning, lightweight CNNs, GPU computing	Lightweight CNN architecture optimized

METHODOLOGY

A most concern for farmers today is plant diseases. Often, they are not sure which insecticide or pesticide to apply to a particular infected plant due to a lack of knowledge about its disease

1. Damping Off
2. Septoria leaf spot
3. Bacterial stem and fruit canker
4. Early blight
5. Bacterial leaf spot
6. Bacterial wilt
7. Leaf curl
8. Mosaic
9. IPM for Tomato
10. Tomato spotted wilt disease

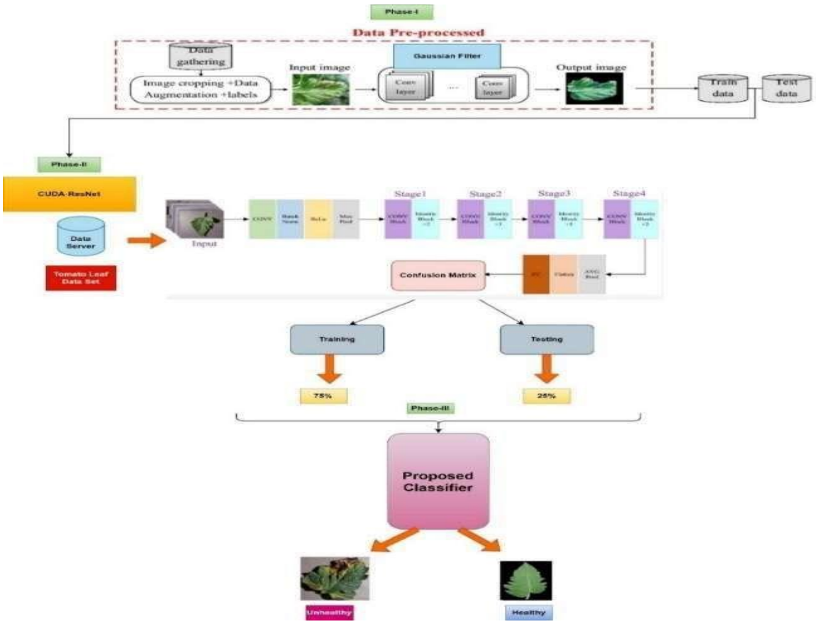


Figure 2: Overall Architecture of The Proposed Classifier

Data Set:

Plant Village Data set UCI Repository

The Open Agriculture Foundation

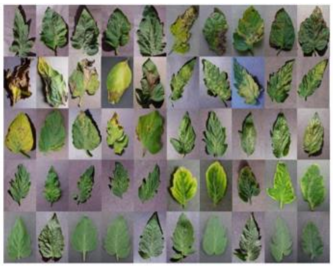


Figure 3: Tomato Leaf Dataset.

Feature Extraction**:**

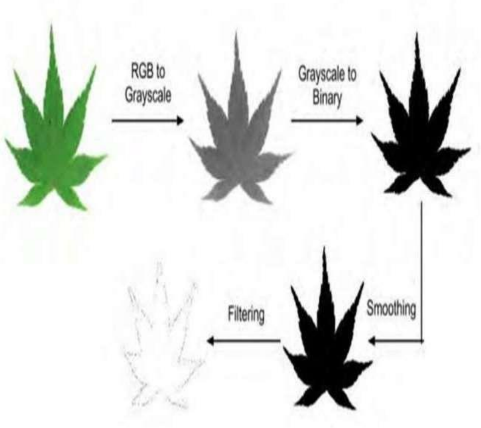


Figure 4: After Processing Image of Tomato Leaf Figure 5: Image Preprocessing of Tomato Leaf

Why CUDA?

A parallel computing platform and a programming model, CUDA (Compute Unified Device Architecture), were developed by NVIDIA for general-purpose computing on GPUs

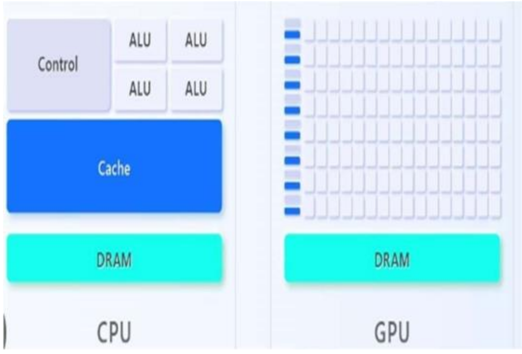


Figure 6: CPU vs GPU

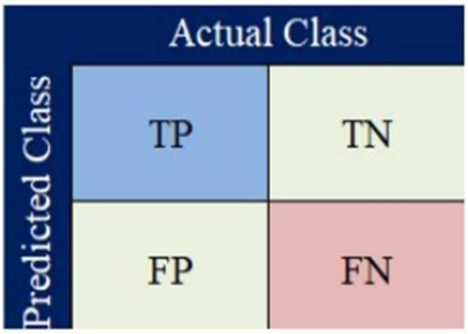


Figure 7: Confusion Matrix for a 2-Class Problem.

Confusion Matrix:

Accuracy : Accuracy=(TP+TN+FP+FN)/(TP+TN) (1)

Precision:

Precision= (TP)/ (TP+FP) (2)

Recall:

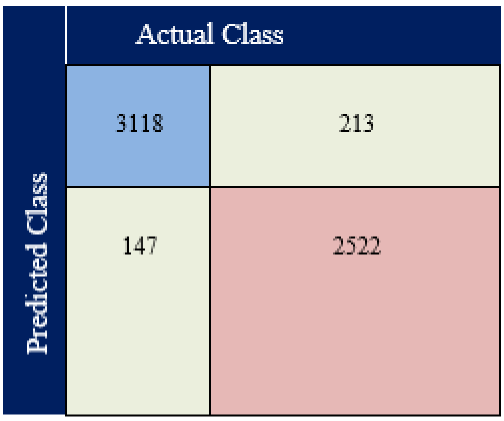
Recall=(TP) / (TP + FN) (3)

F1-Score:

F1-Score=2× (Precision × Recall) / (Precision+ Recall))) (4)

Proposed Classifier

Table 4: CONFUSION MATRIX GENERATED FOR THE PROPOSED CLASSIFIER



The Records using CUDA-ResNet50 Classify

CUDA-ResNet50 GPU Time	No of Records sec/12k	No of Records sec/32k	No of Records sec/52k	No of Records sec/72k
Classification Time	0.662	1.224	1.865	2.445
CPU-Time	0.714	1.321	1.887	2.674
GPU-Time	0.552	0.984	1.223	1.786
Acceleration-Ratio	1.259	1.108	1.168	1.165

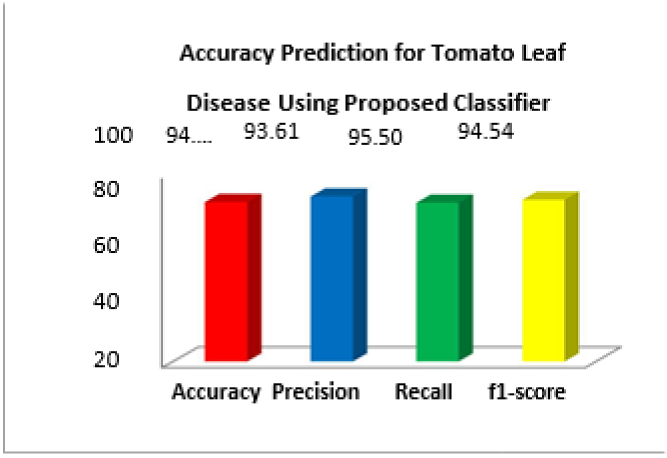


Figure 8: Performance Metrics of Proposed Classifier.

Support Vector Machine

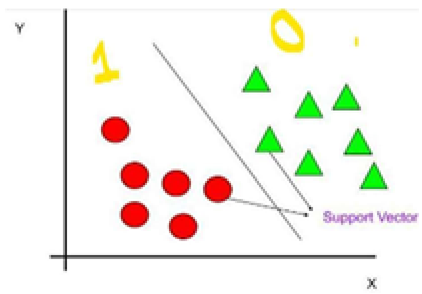


Figure 9: Class Labels Division Using SVM

	Actual Class	
Predicted Class	2888	421
	325	2366

Table 2: Confusion Matrix Generated for the SVM Classifier

Decision Tree

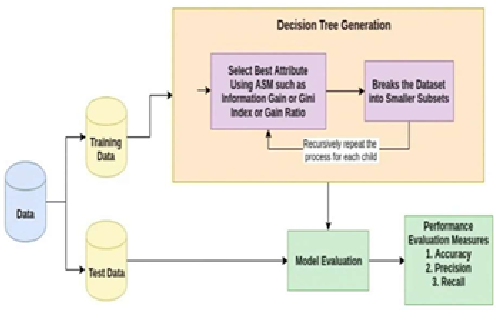


Figure 10: Process of Decision Tree Classifier.

Label	Precision	Recall	F1-Score	Support
0(Healthy)	95.48	88.72	91.56	3120
1(Unhealthy)	94.98	87.56	90.44	2880
Accuracy			90.00	6000
MacroAvg	95.88	88.27	91.56	6000
WeightedAvg	94.10	87.21	90.12	6000

Table3:Validation Table Generated for the DT Classifier

CONCLUSION & WORK

This paper aims to develop a tomato leaf detection algorithm that uses a pre-trained Rest Net model accelerated with CUDA. To enhance performance and speed up the detection process, we utilized GPU-accelerated computation for feature extraction and feature extraction from deep learning.

There are some challenges to be solved in the furthure i.e.,

1. Class Imbalance Handling
2. Transfer Learning and Domain Adaptation.
3. Large-Scale Dataset Challenges

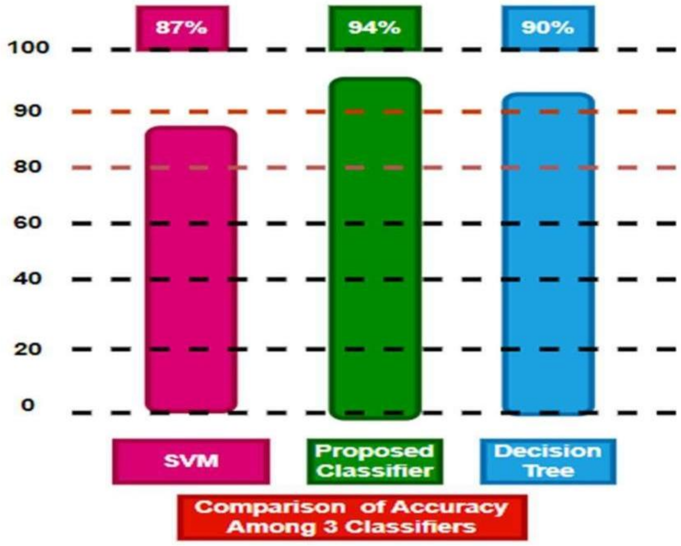


Figure 11: Accuracy Comparison of 3 Classifiers.

REFERENCES

1. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Frontiers in Plant Science, 7, 1419. doi: 10.3389/fpls.2016.01419.
2. Ferentinos, K. P. (2018). doi: 10.1016/j.compag.2018.01.009.
3. Sladojevics, S., Arsenovica, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). 2016, 1-11. doi: 10.1155/2016/3289801.
4. Raza, S. E. A., Cheema, H. M., & Shafique, M. (2017). Deep learning-based tomato plant diseases identification: A comparative study. In 2017 International Conference on Frontiers of Information Technology (FIT) (pp. 163-168).
5. IEEE. doi: 10.1109/FIT.2017.00035
6. Barbeda, J. G. A. (2019). Effect of data set size and Different on the effectiveness and transfer learning for plant disease classification. Computers in Agriculture, 161, 272-281. doi: 10.1016/j.compag.2019.02.024
7. Agarwal, G., & Dadhich, A. (2020). In 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 316-320). IEEE.
   Doi: 0.1109/SPIN48961.2020.9071148
8. Li, Jinyu, et al. Robotics and Cybernetics (CRC), pp. 192-195.

9. IEEE, 2020. [DOI: 10.1109/CRC48926.2020.9071910].
10. Mhamdi, Mohamed Amine, et al. "Machine Learning Approaches for Tomato Plant Diseases Detection and Classification." In 2020 IEEE 5th International Conference on Advanced Robotics and Mechatronics (ICARM), pp. 301-306. IEEE, 2020. [DOI:10.1109/ICARM49283.2020.9161609]