Indian Journal of Computer Science and Technology
Volume 5, Issue1 (January-April 2026), PP: 421-425.
www.indjcst.com
ISSN No: 2583-5300

A Deep Learning Framework for Papaya Crop Health Identification

Nisha Sahu[1], Dr. Gargishankar Verma[2]
[1]*M. Tech, Department of Computer Science Engineering, Krishna's Vikash Institute of Technology, Raipur, Chhattisgarh, India.*
[2] *Professor, Department of CSE, Krishna's Vikash Institute of Technology, Raipur, Chhattisgarh, India.*

To Cite this Article: *Nisha Sahu[1], Dr. Gargishankar Verma[2], "A Deep Learning Framework for Papaya Crop Health Identification ", Indian Journal of Computer Science and Technology, Volume 05, Issue 01 (January-April 2026), PP: 421-425.*

Abstract: *Worldwide, papaya disease is a serious hazard to farmers and causes large losses every year. Understanding how urgent it is to mitigate these losses, scientists have been concentrating more on creating systems to identify papaya diseases. However, farmers frequently don't know how to spot diseases, so they don't identify them until the papayas are already impacted, which results in lost crops and financial losses. As a result, many farmers are reluctant to carry on growing papaya. We carried a study using deep learning technologies for papaya disease detection and classification in order to address this problem.*

Key Words: *Convolutional Neural Networks (CNN), Deep Learning, Papaya Disease, SVM, ResNet101, Disease Identification.*

I. INTRODUCTION

The papaya (Carica papaya L.), a tropical fruit that originated in southern Mexico and Central America, is now grown in many tropical regions of the world. It is abundant in papain, a digestive enzyme, and vitamins A, C, and E. The cultivation of papaya is important to the economics of many poor nations, and because of its nutritious benefits and delectable flavor, it is also becoming more and more popular in industrialized nations. However, there are several obstacles to papaya production, and a number of illnesses have a major effect on fruit quality and productivity. Some of the most prevalent and damaging diseases impacting papaya crops include Papaya Ring spot Virus (PRSV), Papaya Black Spot (PBS), and fungal diseases like Anthracnose and Phytophthora.

These illnesses have the potential to result in catastrophic losses, including decreased fruit yield, subpar fruit, and even total crop failure. For successful disease management techniques, papaya illnesses must be identified early and accurately. Traditional disease detection techniques depend on farmers' visual observation. Nevertheless, this strategy has drawbacks. First of all, it is based on the farmer's experience and skill and is therefore subjective. Farmers with less expertise may find it difficult to detect infections in their early stages, which could result in a delay in intervention and higher production losses. Second, especially on large farms, visual assessment takes a lot of time. Lastly, some illnesses can have subtle signs that are hard to see with the unaided eye.

These drawbacks emphasise the need for more impartial, automated, and trustworthy methods of identifying papaya diseases. This thesis investigates how deep learning, a branch of machine learning, might be used to do this. Deep learning algorithms have the potential to completely transform the area of plant disease diagnosis by providing farmers with precise, effective, and user-friendly solutions. Furthermore, the necessity for efficient disease management techniques in papaya production is made even more urgent by the growing concerns about food security and sustainable agriculture. Deep learning algorithms have demonstrated potential in a number of agricultural applications and have the potential to greatly improve papaya crop disease identification and control.

II.RELATED WORK

A: Inspiration

This idea was motivated by the urgent need for a more precise and quick way to identify papaya diseases. Although they offer a baseline for illness detection, traditional approaches are constrained by human error, subjectivity, and time limits.

The following are some of the particular motivations behind this project: The Papaya's Economic Impact infections: Due to decreased productivity and fruit quality, papaya infections result in large financial losses for producers. To minimize these losses and execute effective disease management methods, early and accurate disease detection is essential.

- **Limited Disease Diagnosis Expertise:** Farmers, especially smallholder farmers in underdeveloped nations, might not have the skills needed to correctly identify papaya illnesses, especially in their early stages. An approachable and dependable substitute can be provided by a deep learning-based tool. Why Time Restrictions on Large plantations: To cover large areas, large papaya plantations need effective disease detection techniques. Time and money can be saved by automating the disease identification process using deep learning models.

- **Potential for Early Disease Detection:** By using small visual clues that human examination could overlook, deep learning models may be able to identify diseases in their early stages. For prompt measures to stop the spread of illnesses, early detection is essential.
- **Developments in Deep Learning Technology:** New methods for creating reliable and effective disease detection models have been made possible by recent developments in deep learning architectures and image categorization techniques.

B: Papaya Disease Recognition:

Traditionally, farmers have relied on their own eyes and experience to spot papaya diseases, a method that is as subjective as it is time-consuming. For those new to farming, telling a healthy plant from a sick one can be especially tricky, particularly when early symptoms are faint and easy to miss. On sprawling farms, the challenge only grows, and some diseases barely leave a trace for the naked eye to catch. These hurdles often lead to late interventions, overlooked infections, and significant crop losses. Enter deep learning: a game-changing approach that harnesses the power of artificial intelligence to identify papaya diseases from plant images. Unlike human inspection, deep learning models offer fast, objective, and reliable results. Trained on vast collections of labeled images, these algorithms learn to spot the subtle patterns that signal disease, even in its earliest stages. Their speed and accuracy make them an invaluable tool for large-scale papaya farms, transforming disease detection from a guessing game into a science.

C: Machine Learning Techniques:

In machine learning, supervised learning algorithms are essential, especially for applications like predicting papaya diseases. When supervised learning is used, a labeled dataset, which consists of data points (papaya photos) associated with a label (healthy or diseased), is crucial. The machine learning model analyses these labeled examples to identify trends and links between the prevalence of particular diseases and attributes taken from the photographs. In essence, the knowledge required to distinguish between healthy and diseased papaya is encoded in these learned patterns and connections. The trained algorithm then uses these discovered connections to forecast the illness state of fresh, unobserved papaya photos. By improving disease detection efficiency, this procedure helps with prompt interventions and reduces yield gains. Farmers may protect crop production and livelihoods by taking proactive steps to efficiently manage papaya diseases by utilizing supervised learning. By using cutting-edge technology like supervised learning, the agricultural industry can adjust to changing problems and support the resilience of global food production.

D: Convolutional Neural Networks:

A particular kind of deep learning architecture called Convolutional Neural Networks (CNNs) is made to examine visual data, such as pictures. They work over several layers, each of which recognizes distinct patterns like edges, textures, and forms. These patterns are integrated as the data passes through deeper layers to identify more intricate structures, such as faces or objects. Because of their exceptional capacity to recognize patterns, CNNs are frequently employed for tasks like object identification and image classification.

Crucial elements of CNNs consist of:

1. Convolutional Layers: These layers scan the input image using filters, sometimes known as kernels, to produce feature maps that emphasize significant visual elements.
2. Pooling Layers: These layers use a technique known as max-pooling to reduce the size of the feature maps. This lessens the processing strain while aiding in the retention of important information.
3. Fully Connected Layers: Following the extraction of pertinent characteristics, these layers carry out the last stage of classification, classifying the image according to the learned features.

E. SVM:

Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification tasks. It works by finding an optimal boundary called a hyper plane that separates data into different classes. The data points closest to the boundary are known as support vectors, which help define the decision boundary. SVM can handle both linear and nonlinear data using kernel functions such as the Radial Basis Function (RBF) kernel. In CNN-based disease detection systems, SVM is used to classify deep features extracted from CNN, resulting in improved accuracy and better performance.

F. ResNet101:

The "Residual" Concept Traditional AI tries to learn the entire image at every layer. ResNet only tries to learn the difference (the "residual") between the input and the output.

Three Key Features

- Skip Connections: It has "highways" that allow information to jump over layers. This prevents the "Vanishing Gradient" problem, where the model forgets what it's learning because it's too deep.
- 101 Layers: It is divided into 4 stages of deep processing, allowing it to see very tiny details (like the texture of a fungus) that shallow models miss.
- Bottleneck Design: It uses 1×1 convolutions to "squeeze" the data before processing it, making a 101-layer network faster and more efficient than older, shallower models. Papaya diseases like Ring Spot and Anthracnose can look very similar. ResNet-101 is deep enough to spot the microscopic differences in pattern and color that a human eye might overlook.

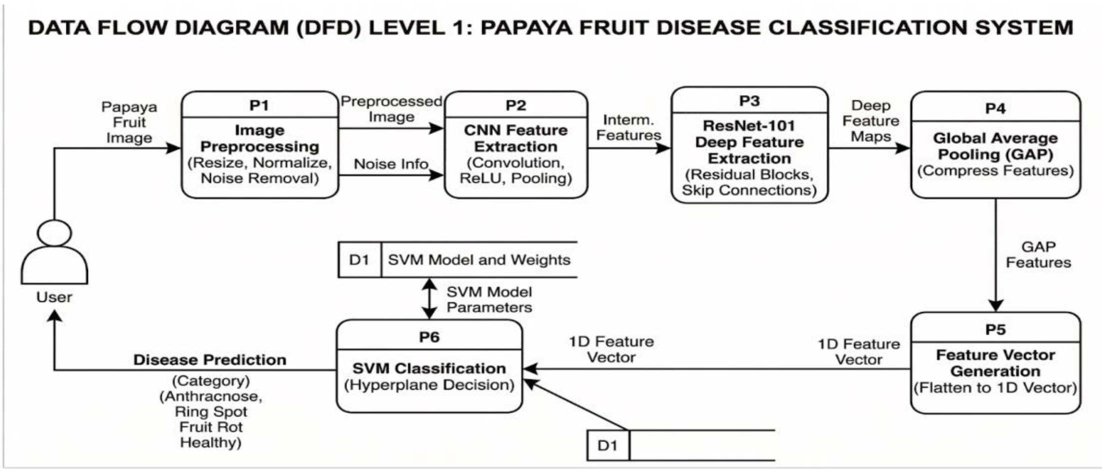
III.LITERATURE REVIEW

Convolutional Neural Networks (CNNs), in particular, are deep learning techniques that have become effective tools for automated disease detection in a variety of crops. CNNs are ideal for tasks like image classification, which is crucial for identifying diseases, because they can learn intricate patterns from picture data. CNN models may be trained to discriminate between healthy and diseased samples as well as between different papaya diseases by examining big datasets of photos of papaya leaves and fruit with disease symptoms.

The performance parameters show that ResNet101 is the best by achieving the maximum 97.50% of accuracy in classifying diseased and healthy papaya fruit varieties [1]. The study reported that the combination of Random Forest and Fuzzy C-Means Clustering achieved the highest accuracy of 96.2% for papaya disease detection [2]. The research utilized a custom implementation of YOLOv9c and achieved an accuracy of 89% for papaya disease classification with improved disease localization [3]. The proposed model identified Ring Spot disease with 77% confidence compared to Anthracnose and Healthy classes [4].The Vision Transformer model achieved a peak accuracy of 91%, outperforming traditional CNN models such as AlexNet, VGG16, and ResNet-50[5].Convolutional Neural Network (CNN) model achieves a high accuracy of 91%, specifically designed to provide village farmers with a rapid and reliable tool for early disease diagnosis to ensure optimal crop yield [6]. The Yolo-Papaya detector, utilizing Convolutional Block Attention Modules (CBAM), achieves a significant 86.2% mean Average Precision (mAP) across a massive dataset of 23,158 samples to support industrial automation and small-scale rural farming [7]. A high-accuracy (98%) CNN-based digital tool that provides farmers with affordable, instant, and reliable papaya disease diagnosis in the field [8]. The combination of Fuzzy C-Means segmentation and Random Forest proved most effective, reaching 97.14% accuracy in identifying fruit health. [9]. An online expert system using K-means and SVM achieved over 90% accuracy, offering a robust automated diagnostic tool for remote agricultural areas[10]. The review identifies SVM as the most popular classifier, noting that Linear SVM paired with ORB features yields the highest overall accuracy (99%)[11 Hybrid models like EfficientNetV2 and ViTs achieved 99.34% accuracy for leaf diseases and nearly 99% for fruit maturity, suitable for mobile monitoring. [12]. Proved Fuzzy segmentation with Random Forest (97.14% accuracy) is the superior method for detecting Anthracnose disease.[13]. A framework using texture and structural features to identify infected fruit regions more effectively than standard DL. [14]. The study developed a reliable deep learning model using the Keras API and CNN architecture that achieved 98% accuracy in recognizing papaya diseases to help farmers reduce cultivation losses[15]. The study achieved an impressive overall accuracy of 98.11% using the proposed three-level CNN topology [16]. YOLOv8 outperformed other versions with a 0.95 mAP, proving highly effective for detecting fruit maturity stages in complex natural environments [17]. A **CNN-based model** was integrated with a pesticide sprayer system, achieving an overall accuracy of **93.1% [18].**

IV. PROPOSED METHODOLOGY

This hybrid Papaya Fruit Disease Detection System integrates CNN, ResNet-101, and SVM for high-accuracy classification. In this pipeline, the CNN captures initial visual patterns, while ResNet-101 utilizes deep residual blocks to extract complex feature representations. Finally, an SVM classifier categorizes the results into Anthracnose, Ring Spot, or Fruit Rot, providing a robust solution for automated smart agriculture.



1. Data Source and External Entity

- **User (External Entity):** The primary actor who interacts with the system. The user provides the Papaya Fruit Image as input and receives the Disease Prediction as the final output.

2. Process Breakdown (P1 – P6)

P1: Image Preprocessing

The raw image undergoes initial refinement to ensure the deep learning model can read it effectively.

- **Resizing:** Adjusts the image to a fixed resolution (e.g., 224×224 pixels) required by the ResNet-101 architecture.
- **Normalization:** Scales pixel values (usually to a range of $[0, 1]$ or $[-1, 1]$) to speed up mathematical convergence.
- **Noise Removal:** Filters out visual artifacts (graininess) that might interfere with feature detection.

P2: CNN Feature Extraction

The system uses standard Convolutional Neural Network (CNN) layers to identify low-level features.

- **Convolution:** Applies filters to detect edges, textures, and simple patterns.
- **ReLU (Rectified Linear Unit):** Adds non-linearity to the model, allowing it to learn complex patterns.
- **Pooling:** Reduces the spatial dimensions (width and height) of the data, keeping only the most important information and reducing computational load.

P3: ResNet-101 Deep Feature Extraction

This is the core "brain" of the system. ResNet-101 is a deep residual network 101 layers deep.

- **Residual Blocks:** These use skip connections (shortcuts) that allow the gradient to flow through very deep layers without vanishing. This enables the model to learn high-level, sophisticated features of specific diseases (like the unique spotting of Anthracnose).

P4: Global Average Pooling (GAP)

Instead of using traditional fully connected layers which are heavy on parameters, GAP takes the average of each feature map. This generates a compact representation of the features, making the system more robust to spatial translations of the fruit in the image.

P5: Feature Vector Generation

The processed data is "flattened." It converts the multi-dimensional feature maps into a 1D Feature Vector. Think of this as a numerical "fingerprint" that uniquely represents the health status of the papaya.

P6: SVM Classification

While ResNet extracts the features, the Support Vector Machine (SVM) makes the final decision.

- Hyper plane Decision: The SVM plots the feature vector in a high-dimensional space and uses a mathematical boundary (hyper plane) to separate the data into categories.

3. Data Stores (D1)

- **D1:** SVM Model and Weights: This represents the database or stored file where the pre-trained weights and mathematical parameters are kept. Process P6 retrieves these parameters to compare the current fruit's features against known patterns of disease.

4. Proposed Output

The final data flow returns a Disease Prediction to the user. The system classifies the papaya into one of four distinct states:

1. Anthracnose: Small, water-soaked spots.
2. Ring Spot: Circular patterns or rings.
3. Fruit Rot: Soft, decaying areas.
4. Healthy: No visible symptoms.

V.CONCLUSION AND FUTURE SCOPE

This research successfully addresses the challenge of Papaya Disease Recognition by implementing a robust framework utilizing various machine learning techniques. By integrating feature extraction and data augmentation, the proposed approach enhances existing algorithms to accurately identify specific papaya illnesses. While this thesis establishes a strong methodological foundation, current results were limited by the availability of a perfectly balanced dataset. Moving forward, the future scope of this work focuses on acquiring a more comprehensive and high-quality dataset to further refine model performance. By training on superior data, we aim to significantly improve classification accuracy and eliminate under fitting, ultimately providing a more reliable tool for real-time agricultural disease management.

REFERENCES

1. A. K. Ratha, N. K. Barpanda, P. K. Sethy, S. K. Behera, and A. Nanthaamornphong, "Optimizing papaya disease classification: A hybrid approach using deep features and PCA-enhanced machine learning," Journal of Intelligent Systems, vol. 34, 2025, Art. no. 20240523. https://doi.org/10.1515/jisys-2024-0523
2. H. Nishad, A. Sharma, G. K. P, M. Prajapati, and M. D. Hamza, "Proposed Image-Based Papaya Health Classification Using Machine Learning," International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), vol. 11, no. 3, pp. 498-504, May-June 2025. Available: https://doi.org/10.32628/CSEIT25113110
3. S. S. Shetty, S. Shetty, S. B. Shetty, Y. N. Pai, and S. T. K, "Papaya Disease Classification Using Machine Learning," International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), vol. 13, no. 3, pp. 520-525, March 2024. Available: https://doi.org/10.17148/IJARCCE.2024.13384
4. V. Chacko, "Image Classification for Papaya Disease Detection Using Deep Learning," International Journal of Innovative Research in Technology (IJIRT), vol. 11, no. 4, pp. 611-615, Sept. 2024. Available: https://ijirt.org/article?manuscript=167880
5. M. Parmar and S. Degadwala, "Deep Learning for Accurate Papaya Disease Identification Using Vision Transformers," International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), vol. 10, no. 2, pp. 420-426, Mar.-Apr.

2024. Available: https://doi.org/10.32628/CSEIT2410235
6. M. Parmar and S. Degadwala, "A Comprehensive Review on Deep Learning for Accurate Papaya Disease Identification," International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), vol. 9, no. 10, pp. 276-282, Sept.-Oct. 2023. Available: https://doi.org/10.32628/CSEIT2361047
7. J. L. de Moraes, J. de Oliveira Neto, C. Badue, T. Oliveira-Santos, and A. F. de Souza, "Yolo-Papaya: A Papaya Fruit Disease Detector and Classifier Using CNNs and Convolutional Block Attention Modules," Electronics (Switzerland), vol. 12, no. 10, 2023.
8. M. S. Hossen et al., "Deep Learning Based Classification of Papaya Disease Recognition," in Proc. 3rd Int. Conf. Intell. Sustain. Syst. (ICISS), 2020, pp. 1-6.
9. N. Bais and N. Rawat, "Image-Based Papaya Fruit Health Classification Using Machine Learning," Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET), vol. 13, no. 5, pp. 3874-3882, May 2025.
10. M. T. Habib et al., "Machine vision based papaya disease recognition," J. King Saud Univ. - Comput. Inf. Sci., vol. 32, no. 3, pp. 300-309, 2020.
11. L. Goel and J. Nagpal, "A Systematic Review of Recent Machine Learning Techniques for Plant Disease Identification and Classification," IETE Tech. Rev., vol. 40, no. 3, pp. 423-439, 2023.
12. M. V. N. Godapitiya and T. K. N. Mandira, "A Deep Learning Approach for Papaya Disease, Pest and Maturity Identification via Mobile Imaging," in Proc. 7th Int. Conf. Adv. Comput. (ICAC), 2025, pp. 1-6.
13. P. Kumaran et al., "Anthracnose Disease Detection in Papaya Fruit Using Machine Learning Techniques," in Proc. 5th Int. Conf. Innov. Trends Inf. Technol. (ICITIIT), 2024, pp. 1-6.
14. K. Harika and M. N. Nachappa, "A Study on Deep Learning Methods to Identify the Infected Regions from Papaya Fruit Images," in Proc. Int. Conf. Ambient Intell. Knowl. Inform. Ind. Electron. (AIKIIE), 2023, pp. 1-6.
15. M. S. Hossen et al., "Deep Learning Based Classification of Papaya Disease Recognition," in Proc. 3rd Int. Conf. Intell. Sustain. Syst. (ICISS), 2020, pp. 1353-1360.
16. R. R. Kumar et al., "Enhancing Precision in Papaya Crop Health Management: A CNN-Based Approach for Early Disease Detection and Classification," in Proc. IEEE Int. Conf. Inf. Technol. Electron. Intell. Commun. Syst. (ICITEICS), 2024, pp. 1-6.
17. A. Prabhu et al., "YOLO Papaya: Towards Identification of Maturity Stages of Papaya on Trees in Natural Environment," in Proc. 6th Int. Conf. Emerg. Technol. (INCET), 2025, pp. 1-6.
18. "Deep Learning for Papaya Leaf Disease Detection and Pesticide Spraying," Int. J. Creat. Res. Thoughts (IJCRT), vol. 13, no. 5, May 2025