JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Deep Learning Approaches for Tomato Leaf Disease Detection Implementation

[1]Mrs.Vanishri Sataraddi, [2]Sanjana Kumari, [2]Shreeya P, [2]Shwetha D
[1]Assistant Professor, Department of Information Science and Engineering, RNS Institute of Technology, Bangalore, India
[2]B.E Students, Department of Information Science and Engineering, RNS Institute of Technology, Bangalore, India

***Abstract*—The goal of this research is to develop an automated system that properly classifies nine different tomato leaf diseases using convolutional neural networks (CNNs). The system tackles many diseases, including Bacterial Spot, Late Blight, and Tomato Mosaic Virus, by using a large dataset that was acquired from Kaggle. The main objective is to meet the pressing need for quick and accurate diagnosis in tomato crops by facilitating early disease detection through the application of deep learning techniques. The current process entails training a CNN model and incorporating it into a user-friendly graphical interface with the goal of giving farmers and agriculture enthusiasts a useful and accessible tool for quickly identifying illnesses.**

Keywords—Convolutional Neural Networks (CNNs), ResNet, DenseNet.

I. INTRODUCTION

Human Activity Recognition (HAR) has become increasingly significant across a wide array of domains including human- machine interaction, healthcare, surveillance, and autonomous driving. There is a growing demand for precise activity detection and pose estimation, catalyzed by the need for efficient time management, contactless interactions amid pandemics, and improved engagement in rehabilitation programs. This survey delves into the aims of training Convolutional Neural Network (CNN) models for automated human activity detection, assessing outcomes, and integrating abnormality detection for real-time surveillance. The focus of the human action recognition system lies in accurately identifying the type of behavior exhibited within a sequence of frames.

II. LITERATURE REVIEW

[1] The article titled "Deep Learning-Based Recognition of Tomato Diseases" represents a significant leap forward in agricultural technology, as it delves deeply into the utilization of Convolutional Neural Networks (CNNs) for the precise detection and classification of tomato diseases. This pioneering research not only underscores the growing intersection of cutting-edge technology and agriculture but also highlights the pivotal role that deep learning can play in addressing the myriad challenges faced by modern farmers. By meticulously examining methodologies and datasets, the study not only unveils the intricate workings of deep learning applications but also sheds light on the multifaceted relationship between technology and agriculture. Through a thorough exploration of common leaf diseases afflicting tomato crops, the research lays a robust foundation for understanding the nuanced complexities of disease identification and management in tomato cultivation. Furthermore, by harnessing the formidable capabilities of deep learning, this study not only holds the promise of significantly enhancing disease diagnosis accuracy but also offers invaluable insights into potential mitigation strategies to safeguard crop yields. As the agricultural landscape continues to evolve in response to technological advancements, this research serves as a beacon of innovation, propelling the industry towards more sustainable and resilient farming practices, ultimately ensuring food security for future generations.

[2] The article titled "A Survey of Deep Learning for Plant Disease Detection," this expansive investigation delves into a comprehensive exploration of diverse deep learning methodologies utilized in the intricate task of detecting plant diseases, with a particular emphasis on those affecting the crucial tomato crop. Embarking on this scholarly journey, the survey meticulously unravels the intricate nuances of cutting- edge approaches, datasets utilized, and the myriad obstacles encountered along the path to more effective disease detection and management strategies. With a keen eye for detail and a commitment to exhaustively covering every facet of the subject matter, this survey transcends mere documentation, emerging as a veritable compendium of knowledge and insights that are indispensable for professionals and researchers within the agricultural domain. By delving deep into the complexities of deep learning techniques, it not only provides a roadmap for navigating the rapidly evolving landscape of agricultural technology but also offers profound insights into the transformative potential inherent in these methodologies. Through its meticulous analysis and comprehensive coverage, the survey equips practitioners with the tools and resources essential for addressing the pressing challenges posed by plant diseases in an increasingly interconnected and dynamic world. As the agricultural industry stands on the cusp of a new era characterized by rapid technological advancement and unprecedented global challenges, this survey serves as a beacon of innovation, guiding the way towards a future where technology and agriculture converge to ensure food security, sustainability, and resilience for generations to come.

[3] The article titled "Tomato Diseases and Pests: Detection and Classification Using Deep Learning," this comprehensive research endeavor delves into the critical task of identifying and categorizing diseases and pests that afflict tomato plants, employing cutting-edge machine learning techniques. As plant diseases and pests continue to pose significant threats to agricultural productivity, the need for accurate and efficient detection methods becomes increasingly pressing. Leveraging the advanced capabilities of deep learning, the study explores a wide array of methodologies aimed at detecting and classifying the myriad diseases and pests that impact tomato crops. Central to the research is the investigation of various deep learning architectures, including but not limited to Faster Region-based Convolutional Neural Network (Faster R-CNN), Region-based Fully Convolutional Network (R-FCN), and Single Shot Multibox Detector (SSD). By combining these architectures with deep feature extractors such as VGG net and Residual Network (ResNet), the study seeks to identify the most effective approach for detecting and classifying tomato diseases and pests. Additionally, the research proposes innovative techniques for local and global class annotation and data augmentation, aimed at enhancing the accuracy andreliability of the detection system.reliability of the detection system.

The performance of the proposed methodology is rigorously evaluated using a meticulously curated dataset specifically tailored for tomato diseases and pests. This dataset encompasses a wide range of scenarios and variations, including different types of diseases and pests, as well as various infection statuses and locations within the plant. Through extensive experimentation and analysis, the research demonstrates the efficacy of the deep learning- based approach in accurately detecting and classifying nine different types of diseases and pests affecting tomato crops. Moreover, the study highlights the potential implications of its findings for agricultural practice, emphasizing the significance of early detection and classification in mitigating agricultural losses and ensuring the health and productivity of tomato crops. By providing a comprehensive overview of state-of-the-art techniques and methodologies, the research aims to pave the way for the development of more robust and effective solutions for combating plant diseases and pests in agricultural settings. As the agricultural industry continues to evolve in response to emerging challenges and technological advancements, the insights gleaned from this study are poised to play a crucial role in shaping the future of crop management and sustainability.

[4] The article titled "CNN-powered Leaf Disease Diagnosis and Treatment Suggestion System" delves into the intricate challenges posed by plant diseases and pests in agriculture. Despite the availability of resources like Kisan call centers, effective disease management remains elusive due to limited accessibility and communication barriers. Farmers often struggle to convey disease concerns accurately over the phone, emphasizing the need for more advanced solutions. Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), offer promising avenues for addressing these challenges. By analyzing real-time images captured by camera devices, CNNs enable swift and accurate disease detection, revolutionizing traditional diagnostic methods. The paper aims to identify the most effective deep learning architecture among Faster R-CNN, R-FCN, and SSD, evaluating their performance in plant disease detection. Furthermore, the article explores the integration of deep learning architectures with advanced feature extractors such as VGG net and ResNet to enhance accuracy and robustness. This comprehensive approach seeks to optimize disease detection systems, ensuring reliable performance across diverse agricultural settings. Additionally, innovative techniques for data annotation, preprocessing, and augmentation are discussed, aiming to improve model performance and generalizability.

In conclusion, the article highlights the transformative potential of CNN-powered leaf disease diagnosis and treatment suggestion systems in revolutionizing agriculture. By leveraging the capabilities of deep learning, coupled with advanced image analysis techniques, farmers can now combat plant diseases and pests with unprecedented accuracy and efficiency. This paradigm shift not only safeguards agricultural productivity but also ensures food security and sustains livelihoods in a rapidly evolving world.

[5] The article titled "Self-Supervised Collaborative Multi- Network for Fine-Grained Visual Categorization of Tomato Diseases" introduces a novel approach to the artificial recognition of tomato diseases. This recognition process is often laborious, time-consuming, and subjective, particularly due to the challenges in discerning small discriminative features between different diseases within tomato images. To address these challenges, the proposed model, named LFC-Net, consists of three interconnected networks: a Location network, a Feedback network, and a Classification network. Notably, the model incorporates a self-supervision mechanism, eliminating the need for manual annotation such as bounding boxes or parts, by effectively detecting informative regions within tomato images. Operating on a novel training paradigm that considers the consistency between image category and informativeness, the Location network identifies informative regions within the tomato image, refining its iterations under the guidance of the Feedback network. Subsequently, the Classification network utilizes the informative regions proposed by the Location network, in conjunction with the entire tomato image, for accurate classification. This collaborative multi- network approach allows for concurrent progress, demonstrating promising advancements in the field of fine- grained visual categorization of tomato diseases.

[6] The article titled "Less Is More: More Compact and Quicker A Comprehensive Neural Framework for Tomato Leaf Disease Grouping" delves into the critical realm of plant disease detection, particularly focusing on the challenges associated with identifying diseases in tomato leaves. This endeavor is essential for safeguarding global food security and ensuring the economic viability of stakeholders within the agricultural sector. While deep learning-based image classification has introduced numerous solutions, their applicability to low-end devices remains a significant obstacle. In response to this challenge, the study proposes a novel lightweight transfer learning-based approach specifically tailored for detecting tomato leaf diseases. Through the integration of effective preprocessing techniques, including illumination correction, the system aims to significantly enhance disease classification accuracy, thereby providing more reliable results. Leveraging a sophisticated combined model comprising a pretrained MobileNetV2 architecture and a robust classifier network, the framework facilitates the extraction of comprehensive features essential for precise disease prediction. To further augment performance and address inherent challenges like data leakage and class imbalance, traditional augmentation methods are replaced with dynamic runtime augmentation strategies. By dynamically augmenting the dataset during training, the system learns to adapt more effectively to diverse scenarios, resulting in improved generalization and robustness. Extensive evaluation conducted on a comprehensive dataset of tomato leaf images sourced from the PlantVillage dataset validates the efficacy and reliability of the proposed framework. The results showcase an impressive level of accuracy, exceeding 99.30%, while maintaining a compact model size of 960MB and 487M floating-point operations. This remarkable performance, combined with its computational efficiency, positions the proposed framework as a promising solution for real-world deployment in low-end devices, thereby advancing the field of plant disease detection and contributing to the sustainability of agriculture on a global scale.

[7] The Article Titled as: Fine-Grained Classification of Plant Leaf Diseases Employing Convolutional Neural Networks. In the realm of plant pathology, the fine-grained classification of leaf diseases presents a multifaceted challenge that requires nuanced solutions. While the variations within disease classes can be substantial, distinguishing between different disease categories often relies on subtle differences, emphasizing the need for methodologies capable of capturing intricate local area features. Moreover, the practical implementation of such solutions is hindered by the computational demands imposed by complex neural network architectures, which often surpass the computational resources available on low-cost terminals. Consequently, there is a pressing need for innovative approaches that reconcile these challenges and pave the way for more accessible and efficient disease classification systems.

In response to these challenges, this study proposes a novel approach to fine-grained disease categorization that harnesses the power of convolutional neural networks (CNNs). By leveraging the inherent capabilities of CNNs in image recognition tasks, the proposed methodology aims to achieve robust and accurate classification of plant leaf diseases at a granular level. Central to this approach is the utilization of attention mechanisms within the CNN architecture, which enables the model to focus on relevant local features while disregarding irrelevant background information. This attention-based strategy not only enhances the model's ability to discriminate between subtle disease patterns but also mitigates the computational burden associated with processing large-scale image datasets.

[8] The article titled "Plant Disease Detection using Internet of Things (IoT)" introduces the innovative application of Internet of Things (IoT) technology in agriculture. It explores the role of IoT in monitoring and controlling agricultural diseases and pest infestations, encompassing systems for disease and pest monitoring, data collection through sensor nodes, and data processing and analysis. The proposed IoT-based disease and pest control system consists of three levels and three subsystems, offering a novel approach to accessing agricultural information for farmers. The paper focuses on the development of an automated system to identify plant diseases, recognizing the significant impact of diseases on plant growth, yield, and agricultural product quality. By leveraging sensors such as temperature, humidity, and color to assess plant leaf health conditions, the automated disease detection system aims to differentiate between healthy and infected plants. Through analysis of temperature, humidity, and color parameters, the system can detect the presence of plant diseases, providing farmers with timely information for effective disease management and crop protection.

[9] The article titled "Detection of Tomato Leaf Diseases for Agro-Based Industries Using Novel PCA Deep Net" underscores the significant strides made by Deep Learning and Computer Vision in revolutionizing agriculture, particularly in the realm of disease detection. Effective classification and detection of healthy and diseased crops are crucial for optimizing production rates and ensuring high-quality yields. In this context, the study introduces a novel approach to detecting tomato leaf diseases by leveraging Deep Neural Networks, aimed at bolstering agro-based industries. The proposed framework combines classical Machine Learning techniques, such as Principal Component Analysis (PCA), with a customized Deep Neural Network dubbed PCA DeepNet. Additionally, the hybridized framework incorporates Generative Adversarial Networks (GANs) to obtain a diverse range of datasets. Disease detection is facilitated through the use of the Faster Region-Based Convolutional Neural Network (F-RCNN). Impressively, the study reports a classification accuracy of 99.60% and an average precision of 98.55%, with a promising Intersection over Union (IOU) score of 0.95 in detection. These results demonstrate the superiority of the presented approach over existing state-of-the-art methods.

[10] The article titled "Crop Disease Detection using Deep Learning" sheds light on the pressing issue of crop diseases exacerbated by drastic climate changes and weakened crop immunity. This surge in crop diseases not only results in widespread crop destruction but also diminishes agricultural productivity, ultimately leading to financial losses for farmers. With the ever-growing variety of diseases and the limited knowledge available to farmers, accurately identifying and treating these diseases has become a significant challenge. The unique texture and visual similarities of crop leaves play a crucial role in disease identification, making computer vision combined with deep learning an ideal solution. To address this challenge, the paper proposes a deep learning-based model trained on a publicly available dataset comprising images of both healthy and diseased crop leaves. This model effectively classifies leaf images into diseased categories based on distinctive patterns of defects, offering a promising approach to combatting crop diseases and enhancing agricultural sustainability.

[11] The article titled "SqueezeNet Model-based Identification of Tomato Plant Diseases by Leaf Image Analysis" introduces a groundbreaking approach utilizing the SqueezeNet design for highly accurate diagnosis of tomato plant diseases through leaf image analysis. This innovative methodology underscores the significance of smartphone connectivity in enabling early disease detection, a pivotal aspect of modern agriculture. By harnessing the capabilities of the SqueezeNet model, the study aims to provide a robust and efficient solution for identifying various ailments affecting tomato plants, thereby empowering farmers with timely insights to mitigate crop losses and optimize yield. Through advanced leaf image analysis techniques, this approach not only enhances disease diagnosis accuracy but also underscores the transformative potential of smartphone- based technologies in revolutionizing agricultural practices.

[12] The article titled "On Using Artificial Intelligence and the Internet of Things for Crop Disease Detection: A Contemporary Survey" delves into the cutting-edge realm of automated agricultural disease diagnosis, specifically focusing on the context of Morocco. This comprehensive study traverses the landscape of advancements in machine learning, deep learning, and Internet of Things (IoT) technologies, elucidating their role in revolutionizing crop disease detection methodologies. By conducting a thorough comparative analysis of existing approaches, the study not only highlights current trends and challenges but also proposes valuable insights and recommendations for further research avenues. Through an exploration of diverse techniques and their applicability in the agricultural domain, this survey aims to provide a holistic understanding of the evolving landscape of crop disease detection, ultimately paving the way for more robust and efficient solutions tailored to the unique challenges faced by farmers in Morocco and beyond.

[13] The article titled "Advanced Real-Time Detection of Tomato Plant Issues Using Deep Learning" delves into a groundbreaking exploration of leveraging deep learning methodologies for real-time identification of issues plaguing tomato plants. By harnessing on-site photos, this research showcases how diseases and pests afflicting tomato plants can be swiftly and accurately identified through the application of sophisticated deep learning architectures. This innovative approach not only ensures the prompt detection of a wide array of problems but also underscores its potential to revolutionize agricultural practices. By streamlining the process of issue recognition, farmers can benefit from enhanced efficiency and productivity, thereby realizing significant economic advantages. Through a meticulous examination of on-site data and deep learning structures, this study contributes to the advancement of techniques aimed at addressing the myriad challenges faced by tomato growers, offering promising prospects for sustainable and profitable agricultural operations.

[14] The article titled "Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network" introduces a groundbreaking advancement in the realm of deep learning models tailored for accurate identification of tomato leaf diseases. This research presents the innovative RRDN (Restructured Residual Dense Network), a sophisticated deep learning architecture meticulously engineered to achieve a remarkable accuracy rate of 95% in disease identification. By strategically restructuring the

conventional residual dense network, the RRDN model optimizes parameters and enhances information flow, thereby surpassing the performance of existing approaches in crop leaf identification. Notably, despite its superior accuracy, the RRDN model exhibits remarkable efficiency by demanding fewer computational resources, thus offering a cost-effective and scalable solution for crop disease management. This pioneering research not only demonstrates the potential of deep learning techniques in addressing agricultural challenges but also underscores the importance of innovation in optimizing model architectures for enhanced performance and resource efficiency in real- world applications.

[15]The article titled "Detection of Tomato Leaf Diseases for Agro-Based Industries Using Novel PCA Deep Net" presents a cutting-edge approach that integrates Principal Component Analysis (PCA) with a customized deep learning architecture, aptly named PCA Deep Net. This innovative hybrid framework marks a significant advancement in the field of plant pathology, particularly in the detection of tomato leaf diseases. By harnessing the complementary strengths of PCA and deep learning, the proposed methodology achieves unparalleled accuracy in identifying various diseases affecting tomato leaves. The incorporation of PCA enriches the feature representation, enabling the deep learning model to extract and leverage essential information effectively. Furthermore, the tailored deep learning architecture is adept at capturing intricate patterns and nuances present in the leaf images, thereby enhancing the overall disease identification process. Through rigorous experimentation and evaluation, this study demonstrates the superiority of the PCA Deep Net approach over existing methods, showcasing remarkable classification accuracy and robustness. The integration of PCA not only enhances the interpretability of the deep learning model but also contributes to its efficiency and scalability, making it a promising solution for disease detection in agro-based industries. This research not only contributes to advancing the state-of-the-art in plant disease diagnostics but also holds immense potential for practical applications in agricultural settings, where accurate and timely disease detection is critical for ensuring crop health and productivity.

III. ANALYSIS

A. *Problem Statement:*

The agricultural sector faces significant challenges in effectively managing and mitigating plant diseases, particularly in crops like tomatoes. Leaf diseases, if not detected early, can lead to substantial yield losses and economic impacts. Traditional methods of disease identification are often labor-intensive and time- consuming, hindering timely intervention. This project aims to address the pressing problem of efficient and rapid tomato leaf disease detection using deep learning techniques. The lack of accessible tools for farmers to identify and classify diseases in real-time contributes to delayed responses and increased crop susceptibility. Additionally, the absence of user-friendly interfaces exacerbates the adoption of advanced technologies in the agricultural community. The project seeks to develop a solution that leverages Convolutional Neural Networks (CNN), ResNet, and DenseNet for accurate disease classification. Integrating these models into a web-based application using Flask, with HTML/CSS interfaces, aims to provide a user-friendly platform for farmers and stakeholders. The project's primary challenge lies in enhancing early disease detection, minimizing crop losses, and facilitating informed decision-making in agriculture by bridging the gap between advanced deep learning models and practical, accessible tools for the farming community.

B. *Objectives*

- Model Development and Training: The primary objective is to develop and train deep learning models, specifically Convolutional Neural Network (CNN), ResNet, and DenseNet, for the accurate classification of tomato leaf diseases. This involves constructing robust models that can effectively differentiate between healthy and diseased leaves based on diverse input images.
- Web Application Integration: Integrate the trained deep learning models into a web application using Flask. The aim is to create a user-friendly platform with HTML/CSS interfaces, allowing users, especially farmers, to easily upload tomato leaf images for disease detection. The web application should facilitate real-time inference, providing prompt feedback on the detected disease type.
- Optimization for Real-world Conditions: Optimize the deep learning models to perform effectively under real- world agricultural conditions. This involves addressing the challenge of detecting diseases in low-quality images, simulating scenarios where farmers may have limited access to high-quality imaging equipment. The objective is to enhance the practical applicability of the models in agricultural settings and promote early disease detection for improved crop management.

C. *Aims of the Project:*

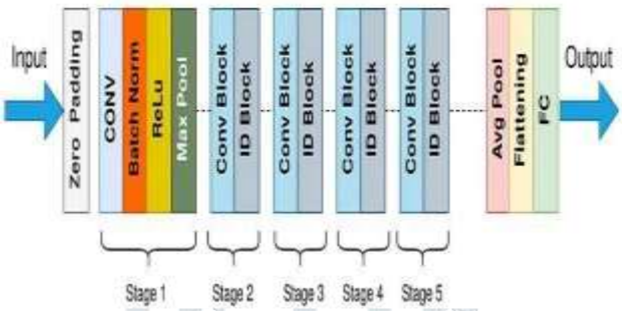
- Develop an automated system for efficient and accurate diagnosis of diseases in tomato crops.
- Utilize deep learning methods, specifically Convolutional Neural Networks (CNNs), for disease recognition and classification.
- Source a comprehensive dataset covering nine distinct classes of tomato leaf diseases from Kaggle.
- Address the challenge of timely identification and management of tomato leaf diseases.
- Create a user-friendly graphical interface integrated with a trained CNN model.
- Contribute to healthier yields in tomato agriculture through early detection and intervention.

Empower individuals involved in tomato cultivation to swiftly and accurately manage diseases for improved crop health and productivity.

IV. METHODOLOGY

A. Data Collection Model:

The basis for accurate disease detection in tomato leaf, the data collection model is a crucial component of the suggested technique. This step involves obtaining a large dataset that includes pictures of various tomato leaf diseases, such as Tomato Mosaic Virus, Target Spot, Bacterial Spot, Tomato Yellow Leaf Curl Virus, Late Blight, Leaf Mold, Early Blight, Spider Mites (Two-spotted Spider Mite), and Septoria Leaf Spot, in addition to pictures of leaves that are in good health. It necessitates careful labeling and making sure the dataset accurately represents real-world situations, which is essential to guaranteeing the necessary diversity needed for reliable model training



C. Model Training:

Figure 3: ResNet

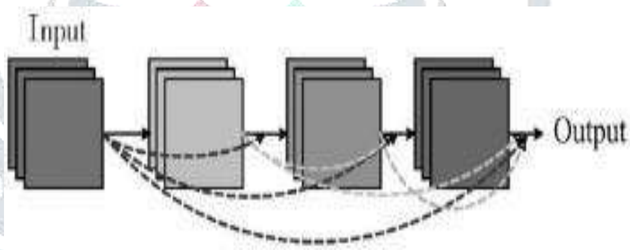


Figure 4: Dense Net

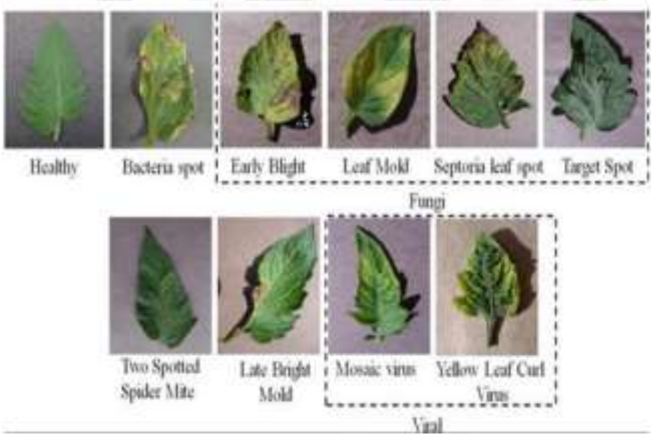


Figure 1: Images of leaves in the Dataset

A. Data Pre-Processing:

Resize and standardize the images to a common resolution to maintain consistency.
Implement data augmentation techniques to increase the diversity of the dataset, which can include random rotations, flips, and brightness adjustments.
Normalize pixel values to a common range (e.g., 0-1) to aid in model convergence.

B. Model Selection:

Choose pre-trained architectures, such as CNN, ResNet and DenseNet, which have demonstrated success in image classification tasks.

Initialize the selected model with pre-trained weights on a large dataset (e.g., ImageNet) to leverage transfer learning.

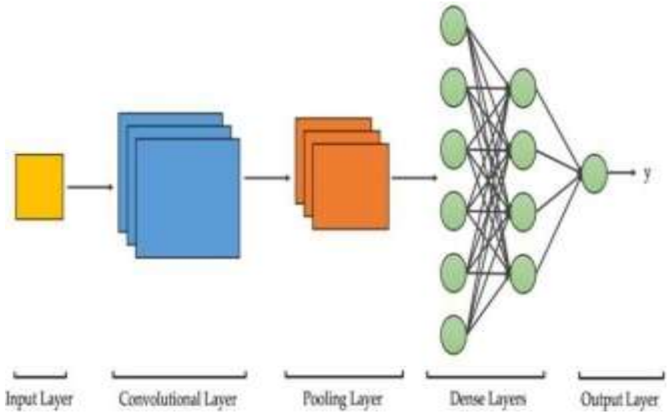


Figure 2: CNN Architecture

Fine-tune the selected pre-trained model on the tomato leaf disease dataset. Utilize an appropriate loss function, such as categorical cross-entropy, and select an optimizer for training, like Adam or SGD. Implement early stopping to prevent overfitting by monitoring performance on the validation set.

D. Model Evaluation:

Assess the trained model's performance using multiple evaluation metrics, including accuracy, precision, recall, and F1 score. Examine confusion matrices and ROC curves for a more in-depth understanding of classification results. Fine-tune hyperparameters as necessary to optimize model performance

E. Model Selection and Integration:

Based on the evaluation results, select the best- performing model for disease detection.

Integrate the chosen model into a Flask-based graphical user interface (GUI) to create a user-friendly application for end-users.

V. SYSTEM DESIGN

Systems design implies a systematic approach to the design of a system. It may take a bottom-up or top-down approach, but either way, the process is systematic wherein it takes into account all related variables of the system that needs to be created—from the architecture, to the required hardware and software, right down to the data and how it travels and transforms throughout its travel through the system. Systems design then overlaps with systems analysis, systems engineering, and systems architecture.

A. System Architecture

The proposed system architecture as shown in Figure 5 consists of several steps that facilitate efficient user interaction with the interface. The system follows a supervised learning approach for image classification. It starts with data pre-processing, where raw data is cleaned and prepared. Exploratory data analysis (EDA) helps understand the data's characteristics and relationships between variables. Then, multiple models, potentially including CNNs (Convolutional Neural Networks), ResNets, and DenseNets, are trained on the pre-processed data. These models learn to identify patterns in images and associate them with specific classes (labels). Next, the models are evaluated using a separate validation set to assess their performance. Metrics like accuracy, precision, recall, or F1-score are used for comparison. data is chosen for deployment.

This selection process ensures that the system effectively classifies new images based on the knowledge it has acquired from the training data. The selected model is integrated with the GUI. The image is given as input to the model after analysis the output is given back to the user through the GUI.

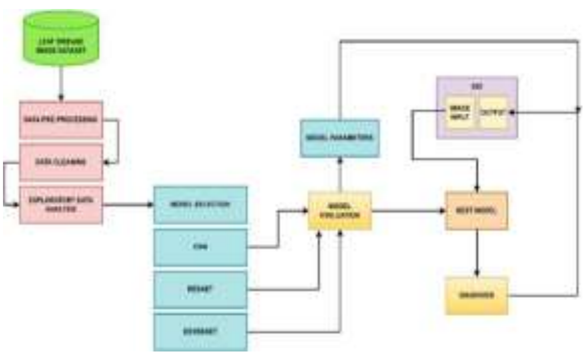


Figure 5: System architecture

VI. IMPLEMENTATION

A. Overview of System Implementation

- Data Collection: Data collection is a crucial aspect of the system. It involves gathering images or sensor data from tomato plants. This data serves as input for disease detection algorithms.
- Image Processing and Disease Detection: Once the data is collected, image processing techniques are applied to extract relevant features from the images. Machine learning or computer vision algorithms can then be used to detect diseases based on these features. These algorithms may need to be trained on a dataset of labeled images to accurately identify diseases.
- User Interface: A user interface is developed to allow users to interact with the system. This could be a web application, mobile app, or desktop software. The interface allows users to upload images of tomato leaves, view the results of disease detection, and access historical data.

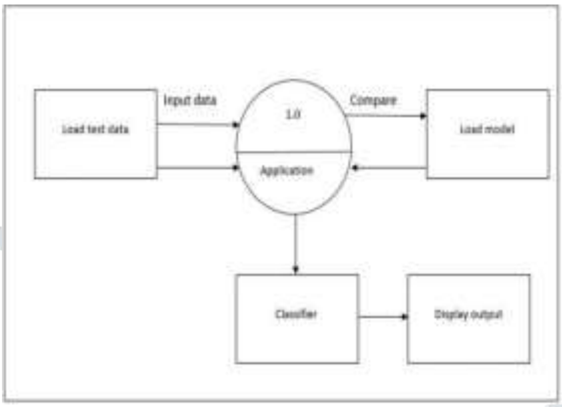
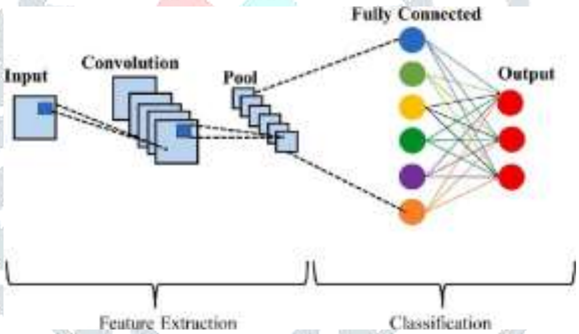


Figure 6: Dataflow Diagram

Algorithms:

1. Convolutional Neural Networks (CNNs):



Convolutional Neural Network (CNN) is a deep learning architecture specifically designed for processing and analyzing visual data, such as images and videos. CNNs are composed of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers apply filters to input images, extracting features at different spatial hierarchies through convolutions. Pooling layers reduce the dimensionality of feature maps, capturing the most relevant information while preserving spatial relationships. Fully connected layers integrate the extracted features and perform classification or regression tasks. CNNs excel at tasks like image classification, object detection, and semantic segmentation, thanks to their ability to automatically learn hierarchical representations of visual data. Their hierarchical feature extraction process makes them highly effective in capturing patterns and structures present in images, enabling robust and accurate analysis across a wide range of applications in computer vision.

2. ResNet:

ResNet, short for Residual Network, is a deep convolutional neural network architecture that has revolutionized the field of computer vision. Introduced by Kaiming He et al. in 2015, ResNet addresses the challenge of training very deep neural networks by introducing skip connections, or "residual connections," that enable the network to learn residual mappings instead of directly learning the underlying mapping functions.

This architectural innovation effectively alleviates the vanishing gradient problem and enables the training of extremely deep networks with hundreds or even thousands of layers. ResNet architectures, including variants like ResNet- 50, ResNet-101, and ResNet-152, have achieved state-of-the- art performance on various image classification, object detection, and segmentation tasks, earning widespread adoption in both research and practical applications. With its remarkable depth, robustness, and scalability, ResNet continues to be a cornerstone in the development of advanced deep learning models for visual recognition tasks.

3. DenseNet:

DenseNet, short for Dense Convolutional Network, is a groundbreaking neural network architecture that introduces dense connectivity patterns among layers. Proposed by Gao Huang et al. in 2017, DenseNet revolutionizes the traditional convolutional neural network (CNN) design by establishing direct connections between all layers within a dense block. In contrast to traditional architectures where each layer is connected only to its subsequent layers, DenseNet connects each layer to every other layer in a feed-forward fashion.

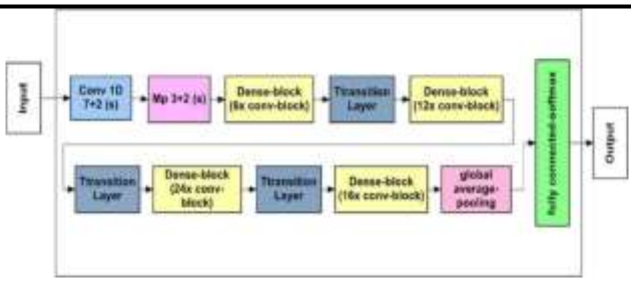


Figure 7: DenseNet

This dense connectivity facilitates feature reuse, encourages feature propagation, and alleviates the vanishing gradient problem, leading to improved gradient flow and better information flow throughout the network. DenseNet architectures, such as DenseNet-121, DenseNet-169, and DenseNet-201, have demonstrated superior performance on various image classification tasks, achieving state-of-the-art accuracy with significantly fewer parameters compared to other CNN architectures. With its unique dense connectivity structure and impressive efficiency, DenseNet has become a cornerstone in deep learning research and has found widespread applications in image analysis, medical imaging, and computer vision tasks.

VII RESULTS AND DISCUSSIONS

The tomato leaf disease detection system represents a significant advancement in agriculture, providing farmers with a powerful tool for monitoring and managing crop health. By integrating advanced machine learning models like CNNs and DenseNet, the system enables rapid and accurate disease detection in tomato plants. This empowers farmers to make informed decisions, mitigate disease impact, and optimize crop yields. Overall, this system demonstrates the transformative potential of technology in addressing agricultural challenges, contributing to food security, environmental stewardship, and economic prosperity.

A. User Interface

At the start of their interaction, users encounter a login interface, where they're prompted to enter their username and password for authentication. This login page serves as the main gateway for users to enter the system, where they confirm their credentials to gain access. It prominently displays fields for users to input their username and password. Upon verification of the provided credentials, users gain entry; otherwise, an error message indicating an invalid username or password is shown. Once successfully logged in, users are welcomed with a user-friendly menu featuring four buttons, each corresponding to a project objective: Home, Model Parameters, and Detector. Within the Detector menu, users can upload images from the dataset and utilize the model to predict diseases in tomato plant leaves.

Figure 8: Login page

Figure 9: Main menu

B. Exploratory Data Analysis

1. The Exploratory Data Analysis (EDA) reveals insights into a dataset comprising 10,000 entries for tomato leaf detection: File Path: Each entry includes a file path denoting the location of associated images. Image Height and Width: The dataset consists of images with a consistent height and width, represented as integers. Class: The "class" column indicates the category or class of each entry, pertaining to different types of tomato leaf diseases .Data Types: Two columns are of integer type (image_ height and image_ width), while the other two are of object type (file_ path and class).

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 4 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   file_path     10000 non-null  object
 1   image_height  10000 non-null  int64
 2   image_width   10000 non-null  int64
 3   class         10000 non-null  object
dtypes: int64(2), object(2)
memory usage: 312.6+ KB
```

Figure 10: Basic information of training data

2. The statistical description of the training data for tomato leaf detection reveals key insights into the dataset's characteristics. Both the image height and width encompass 10,000 entries, reflecting a complete dataset. With mean dimensions of 256 pixels for both height and width, the data showcases uniformity in image size. Notably, a standard deviation of 0 indicates no variability in image dimensions, further affirming consistency. The minimum and maximum values, along with percentiles (25th, 50th, and 75th), all registering at 256 pixels, reinforce the uniformity observed. This comprehensive statistical summary provides crucial information about the dataset's size and uniformity, pivotal for comprehending the data's nature in tomato leaf detection.

	image_height	image_width
count	10000.0	10000.0
mean	256.0	256.0
std	0.0	0.0
min	256.0	256.0
25%	256.0	256.0
50%	256.0	256.0
75%	256.0	256.0
max	256.0	256.0

Figure 11: statistical description of training data

3. The count plot depicts the distribution of classes in the dataset through a bar graph. Each class is represented by a bar, with the count values ranging from 100 to 1000, as indicated on the y-axis with intervals of 200 (0, 200, 400, 600, 800, 1000). The classes include Bacterial Spot, Early Blight, Healthy Tomato Leaves, Leaf Mold, Septoria Leaf Spot, Target Spot, Spider Mites Two-Spotted Spider Mite, Tomato Mosaic Virus, and Tomato Yellow Leaf Curl Virus. This visualization offers a clear overview of the frequency distribution of different tomato leaf disease classes, facilitating easy comparison of their occurrence within the dataset.

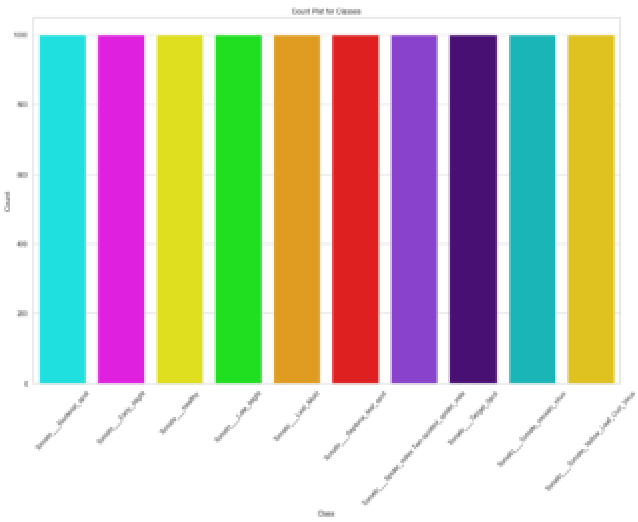


Figure 12: Count plot of classes

4. The pie chart illustrates the distribution of various classes of tomato leaf diseases in the dataset. Each segment represents a specific disease category, with the following distribution: Tomato Mosaic Virus, Tomato Yellow Leaf Curl Virus, Target Spot, Bacterial Spot, Spider Mites Two- Spotted Spider Mite, Early Blight, Septoria Leaf Spot, Healthy Leaves, Leaf Mold, and Late Blight, each accounting for 10% of the dataset. This distribution showcases the diversity of diseases present in tomato plants, with an equal representation of each class. Such visual representation aids in understanding the relative prevalence of different diseases, providing valuable insights for disease management and mitigation strategies in tomato cultivation.

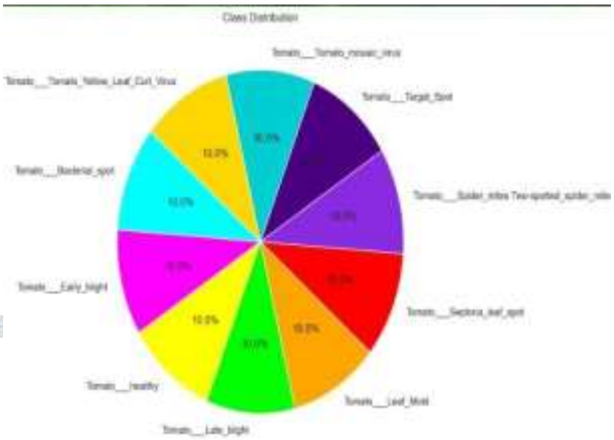


Figure 13: Pie plot of class distribution

C. Model Parameters

1. **Basic CNN:** Our basic Convolutional Neural Network achieves an accuracy of 88% on the test dataset.

CNN Classification Report:

	precision	recall	f1-score	support
Tomato___Bacterial_spot	0.93	0.90	0.91	100
Tomato___Early_blight	0.79	0.86	0.82	100
Tomato___Late_blight	0.90	0.89	0.89	100
Tomato___Leaf_Mold	0.98	0.85	0.91	100
Tomato___Septoria_leaf_spot	0.76	0.90	0.82	100
Tomato___Spider_mites Two-spotted_spider_mite	0.93	0.85	0.89	84
Tomato___Target_Spot	0.85	0.82	0.83	100
Tomato___Tomato_Yellow_Leaf_Curl_Virus	0.98	0.89	0.93	100
Tomato___Tomato_mosaic_virus	0.92	0.86	0.89	100
Tomato___healthy	0.84	0.98	0.91	100
accuracy			0.88	984
macro avg	0.89	0.88	0.88	984
weighted avg	0.89	0.88	0.88	984

Figure 14: Classification report of **CNN** model

- Confusion matrix for basic CNN model

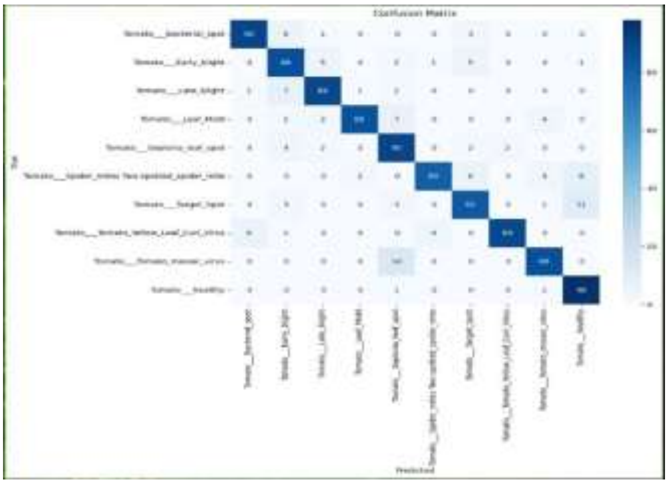


Figure 15: Confusion matrix for basic CNN model

DenseNet: Utilizing DenseNet architecture, our model achieves an improved accuracy of 91% on the test dataset compared to the basic CNN.

DENSENET Classification Report:

	precision	recall	f1-score	support
Tomato Bacterial spot	0.87	0.97	0.92	100
Tomato Early blight	0.94	0.90	0.92	100
Tomato Late blight	0.75	1.00	0.85	100
Tomato Leaf Mold	0.88	0.94	0.91	100
Tomato Septoria leaf spot	0.85	0.87	0.86	100
Tomato Spider mites Two-spotted spider mite	0.92	1.00	0.96	84
Tomato Target Spot	1.00	0.79	0.88	100
Tomato Tomato Yellow Leaf Curl Virus	0.99	0.99	0.99	100
Tomato Tomato mosaic virus	1.00	0.88	0.94	100
Tomato healthy	1.00	0.76	0.86	100
accuracy			0.91	984
macro avg	0.92	0.91	0.91	984
weighted avg	0.92	0.91	0.91	984

Figure 16: Classification report of denseNet model

- Confusion matrix for basic Dense net model

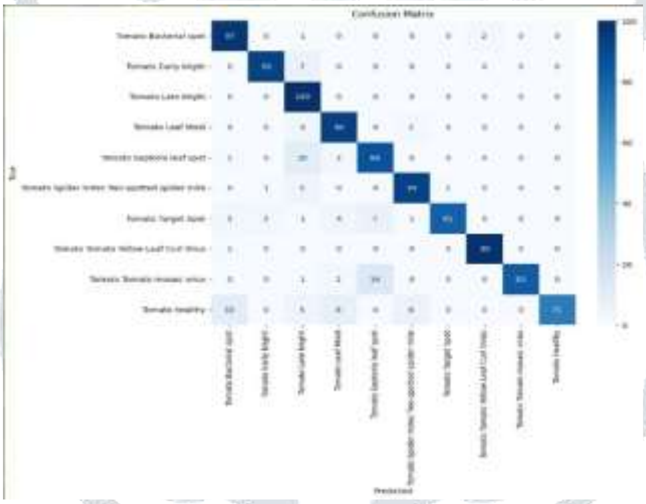


Figure 17: Confusion matrix for basic Dense net model

2. ResNet: Employing Residual Neural Network architecture, our model achieves the highest accuracy of 95% on the test dataset, outperforming both the basic CNN and DenseNet architectures.

RESNET Classification Report:

	precision	recall	f1-score	support
Tomato___Bacterial_spot	0.99	0.94	0.96	100
Tomato___Early_blight	0.92	1.00	0.96	100
Tomato___Late_blight	0.95	0.97	0.96	100
Tomato___Leaf_Mold	0.98	0.99	0.99	100
Tomato___Septoria_leaf_spot	0.85	0.87	0.86	100
Tomato___Spider_mites Two-spotted_spider_mite	0.96	0.89	0.93	84
Tomato___Target_Spot	0.93	0.93	0.93	100
Tomato___Tomato_Yellow_Leaf_Curl_Virus	0.97	0.99	0.98	100
Tomato___Tomato_mosaic_virus	1.00	0.93	0.96	100
Tomato___healthy	0.96	0.98	0.97	100
accuracy			0.95	984
macro avg	0.95	0.95	0.95	984
weighted avg	0.95	0.95	0.95	984

Figure 18: Classification report of ResNet model

ResNet: After examining the confusion matrix of ResNet, it's evident that it is diagonally dominant, suggesting excellent model performance. The predictions are precise with minimal False Positives and False Negatives, indicating reliable performance across the board.

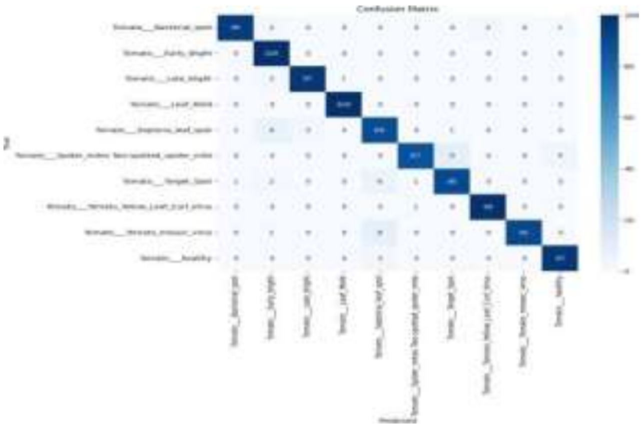


Figure 19: Confusion matrix of Resnet model

VIII APPLICATION

By harnessing the power of advanced technologies like deep learning, particularly through machine learning and computer vision, tomato leaf disease detection achieves unprecedented levels of accuracy and efficiency. Deep learning algorithms, such as Convolutional Neural Networks (CNNs), DenseNets and ResNet, enable the automated analysis of large volumes of image data, allowing for rapid and precise identification of various diseases affecting tomato plants. These algorithms can learn intricate patterns and features from images, distinguishing between healthy and diseased leaves with remarkable accuracy. By leveraging deep learning, the system not only enhances the speed and accuracy of disease detection but also streamlines the process, empowering farmers with timely insights for effective crop management and disease control.

IX CHALLENGES AND OPEN PROBLEMS

Detecting diseases in tomato leaves presents multifaceted challenges and open problems that require innovative solutions. The foremost challenge lies in ensuring the quality and resolution of captured images, as these factors profoundly influence the accuracy of disease identification. Efforts to enhance image quality and resolution through advancements in camera technology and image enhancement techniques are crucial for improving disease detection accuracy. Moreover, detecting subtle symptoms or early-stage indicators of diseases demands innovative approaches, such as leveraging advanced camera technologies like multispectral or thermal imaging to identify early signs of disease stress or physiological changes in tomato plants. Additionally, real-time analysis and processing of large volumes of image data pose significant computational challenges. Developing efficient real-time image processing algorithms, potentially integrated directly into camera devices, can streamline data analysis and reduce processing overhead. Ensuring scalability and adaptability across different environments, plant varieties, and disease types is vital, requiring the development of scalable algorithms capable of handling diverse datasets and environmental conditions while maintaining accuracy. Designing intuitive user interfaces further enhances accessibility and utility for farmers and agricultural stakeholders. Finally, building accurate and comprehensive datasets for training disease detection algorithms necessitates innovative labeling approaches, including semi-supervised or unsupervised learning methods and automated labeling techniques, to facilitate the creation of large-scale, labeled datasets for training and validation. Addressing these challenges promises to advance tomato leaf disease detection, ultimately benefiting crop health, productivity, and agricultural sustainability.

REFERENCES

[1] Kementerian Pertanian Direktorat Jenderal Hortikultura, "Statistik Produksi Hortikultuia Tahun 2014," Stat. Produksi Hortik. Tahun 2014, p. 286, 2014.

[2] N.Otsu, "A Threshold Selection Method from Gray- Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, No. 1, pp. 62-66, Jan. 1979.

[3] J. Franke, and G. Menz, "Multi-temporal wheat disease detection by multi-spectral remote sensing," Precision Agriculture, vol. 8, no. 3, pp. 161–172, June 2007.

[4] Y. H. Bhosale and K. Sridhar Patnaik, ‘‘IoT deployable lightweight deep learning application for COVID-19 detection with lung diseases using RaspberryPi,’’ in Proc. Int. Conf. IoT Blockchain Technol. (ICIBT), May 2022, pp. 1–6.

[5] Sankaran, Sankaran, Mishra, Ehsani, and Davis, June 2010, "A review of advanced techniques for detecting plant diseases," Computers and Electronics in Agriculture, vol. 72, no. 1, pp. 1–13

[6] D. W. Zhang and J. Wang, “Design on image features recognition system of cucumber downy mildew based on BP algorithm”, Journal of Shenyang Jianzhu University (Natural Science), May 2009, vol. 25, pp. 574–578.

[7] "Attention embedded residual CNN for disease detection in tomato leaves," by R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, Appl. Soft Comput., vol. 86, Jan. 2020, Art. no. 105933.

[8] In the proceedings of the 6th International Conference on Agro Geoinformatics, held in August 2017 in Fairfax, Virginia, USA (IEEE Press), pp. 1–5, H. Durmus, E. O. Gunes, and M. Kirci discussed "Disease detection on the leaves of the tomato plants by using deep learning." doi: 10.1109/AGRO-GEOINFORMATICS.2017.8047016.

[9] F. Kuusisto, J. W. Steill, Z. Kuang, J. A. Thomson, D. Page, and R. M. Stewart, ‘‘A simple text mining approach for ranking pairwise associations in biomedical applications,’’ in Proc. AMIA Jt Summits Transl. Sci., vol. 2017, 2017, pp. 166–174.

[10] . Khan and M. Narvekar, "A novel combination of superpixel-based approach and color balancing for the detection of tomato plant diseases in a complex natural environment," Early access published in J. King Saud Univ.-Comput. Inf. Sci. on September 15, 2020; doi: 10.1016/j.jksuci.2020.09.006.

[11] S. D. Khan and H. Ullah, ‘‘A survey of advances in vision-based vehicle re-identification,’’ Comput. Vis. Image Understand., vol. 182, pp. 50–63, May 2019.

[12] Y. Yu, K. Zhang, L. Yang, and D. Zhang, ‘‘Fruit detection for strawberry harvesting robot in non- structural environment based on mask-RCNN,’’ Comput. Electron. Agricult., vol. 163, Aug. 2019, Art. no. 104846, doi: 10. 1016/J.COMPAG.2019.06.001.

[13] M. Krishnan and G. Jabert, Using Image Processing for Pest Control in Agricultural Plantations. 2013; 6(4): p. 68–74; IOSR Journal of Electronics and Communication Engineering (IOSR-JECE).

[14] I. Gogul and V. S. Kumar, ‘‘Flower species recognition system using convolution neural networks and transfer learning,’’ in Proc. 4th Int. Conf. Signal Process., Commun. Netw. (ICSCN), Mar. 2017, pp. 1–6, doi: 10. 1109/ICSCN.2017.8085675.

[15] “Tomato crop disease classification using pre-trained deep learning algorithm," A. K. Rangarajan, R. Purushothaman, and A. Ramesh, Procedia Comput. Sci., vol. 133, pp. 1040–1047, Jan. 2018, doi: 10.1016/ J.PROCS.2018.07.070.