Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 10, Issue 3, pp: (33-43), Month: September - December 2023, Available at: www.noveltyjournals.com

A Cascaded Intelligent Recommender Model for Multi-Class Tomato Leaf Diseases Identification and Control

Nkechi F. Esomonu[1], Udoka F. Eze[2], A. M. John-Otumu[3], I. G. Ayogu[4]

[1,2,3]Department of Information Technology, Federal University of Technology Owerri

[4]Department of Computer Science, Federal University of Technology Owerri

DOI: https://doi.org/10.5281/zenodo.10053690

Published Date: 28-October-2023

Abstract: **In this paper, we present an integrated mobile system designed to empower tomato farmers with effective disease management strategies. Our approach harmonizes advanced machine learning techniques: Pre-trained Networks (MobileNetV2), Support Vector Machines (SVM), and Universal Sentence Encoder (USE), enabling precise disease identification and classification. This innovation introduces a pioneering recommender system aiding farmers in disease recognition through symptom analysis and preventive measures. In contrast to conventional methods, our system offers a holistic solution uniting technological innovation with pivotal information science components, thereby enhancing sustainability in tomato agriculture. We assess model performance in multi-class classification—encompassing ten classes (healthy and unhealthy)—utilizing an extensive dataset of 11,000 tomato images from Plant Village. Existing tomato leaf disease classification models are scrutinized, encompassing model types, datasets, recommender systems, and mobile integration. Predominantly, consulted research employs customized Convolutional Neural Networks (CNNs), demonstrating enhanced performance through hyper-parameter tuning. However, these models primarily focus on classification, lacking a recommender system that informs farmers of symptoms and preventive measures at the early disease stages. Our investigation reveals that the pretrained MobileNetV2 in conjunction with SVM yields exceptional performance, achieving 93.68% precision, 93.84% F-1 score, 96.81% accuracy, and 95.4% recall—surpassing alternative pre-trained models.**

Keywords: **AI, Deep Learning, CNN, Classification, Images, USE.**

1. INTRODUCTION

Agriculture is the mother of nature and a major resource needed for the growth of any economy. In agriculture, innovative technologies can help improve both quality and quantities of agricultural products in the country. For instance, some technologies can be used by farmers for weather forecasting, disease prediction, identification of phenology and crop harvesting. Due to extremely high infant mortality, the human population of the planet increased slowly until the year 1700. The first billion was reached in ca. 1800, followed by the second billion in 1928, the third billion in 1960. In 2017, the world's population reached its seventh billion. (Domingues *et al.*, 2022).

The rise in population growth over recent decades is attributed to improved medical care. According to predictions from the United Nations, the world's population is expected to reach 9.7 billion in 2050, and 10.9 billion in 2100 (Ahmad et al., 2020, Roser & Ortiz-Ospina, 2013).

The demand and consumption of agricultural products in recent years has appreciated due to the rapid population growth globally, which in turn has led to a large expansion of cultivation(Frona *et al*., 2019). It is of paramount importance to double output production of crop yield by 2050 in order to meetup with the high demand for food, bio-fuels, and animal products. To attain this goal, a 2.4% key improvement of crop yield each year must be considered, based on the poor estimated increase of 1.3% per year of present day.(Ray *et al*., 2013).

Statistics has shown that Agricultural sector is a very significant sector of Nigeria's economy based on it's gross domestic product (GDP) contribution of about 23% in the first half of the year (Izuaka, 2020). Small scale farming system is predominately the most practiced agricultural system in Nigeria within an average farm size of 0.2-3 hectares. This subsistence farming contributes to low scale food production of over 70% that sustains just about 75% food demand of Nigeria's population demand. (FAO, 2017). Based on this report, advancement and management of small-scale farming is critical to food security. As demand for food grows in Nigerian, urban populations are increasingly depending on small-scale farmers for their household supply of food. Food availability for urban and peri-urban populations are highly dependent on small-scale farmers.

Tomatoes is one among the very common vegetable crops grown on small scale and also one of the most extensively used vegetable among Nigerian crops in Nigeria and globally.(Adekunle et *al*., 2019). Tomato growing plays a critical role which has been linked to many health benefits, that generates revenue and creates employment for rural and urban populations. The tomato (*Lycopersicon esculentum Mill*), a member of the Solanaceae family and an edible fruit and vegetable, is a South American native. Probably from Lycopersicon esculentum Var, it evolved. Early Mexican Indians brought it under domestication and began cultivating it throughout Central America. Based on Spanish colonization of the Americans, tomato cultivation spread throughout the world (Modi *et al*., 2006). The tomato plant is an annual crop planted outdoors in temperate climates that normally develops to a height of 1-3 meters (3-10 feet).

Based on annual crop production, tomatoes can be grouped into two: namely, determinant and indeterminant with more than 10,000 species. It has hairy leaves that varies in sizes and colours; like red, pink and yellow. When the seeds are caught, the change form to be slightly curved, flattened and have become hairy as light brown seeds. The fruit is a tomato's most important component. Botanically speaking, the fruit is a berry with two or more curved surfaces that soften as the fruit ripens and the seeds are fully formed (Nell *et al*., 2006).

Mkhabela(2005), has it that a lot of novel tomato cultivars are frequently grown in green houses and in milder areas, like RomaVF, where they can be grown from winter through spring and summer. The maturation period ranges from 120 to 150 days. The quality of the fruit, adaptability, diseases, pests, plant growth habits, the market, and planting time all factor into the selection of the cultivar. Again, (Bhowmik *et al*., 2012), believes that tomato fruits can be used in a variety of ways. Fresh tomatoes are eaten raw, or they can be processed into purees, pastes, powders, ketchup (tomato sauce), soup, or whole fruits that are canned. Tomato extracts have been utilized in traditional medicine to treat a variety of illnesses, according to research. According to the production guidelines for tomato (2001), it is believed that tomato poultice has been used to treat edema in pregnant women in Japan, Greece, Peru, and Guatemala while hot water extract of dried fruits has been used to heal ulcers, hemorrhoids, and burns. Report has also shown that fresh tomato fruits can improve digestion and alleviate liver and renal problems. Additionally, the fruit is a good source of vitamin C and lycopene, an organic molecule that is a red carotenoid pigment present in tomato fruit.

As reported by Kheyrodin (2017), tomatoes have great nutritional value and are good a source of vitamins A and B which are essential for the human diet. In many ways, the crop has significant economic worth such as disease-resistant tomatoes or booster types of output.

There is a profitable income from the sales of tomato, according to Oladejo & Laudia, (2011), the tomato economy has not only increased the agricultural sector's part of the national economy but also represents one of Nigeria's most promising sectors with a competitive advantage in the liberalized market. It is a vegetable agricultural product that has an all year-round demand for most homes, globally.

However, tomato production posed with numerous threats related to pest and diseases. There has been 80-100% depreciation in tomatoes production due to disease outbreak (Baudoin *et al*., 2017; FAO, 2019). Pest and diseases have negatively affected production success of small-scale farming. Pest and diseases are prominent factors that has contributed to substantial losses based on high mortality in crops and reduced productivity. For example, in accordance with Premium times Newspapers reported that in April 2020, Tuta absoluta, one of the most devastating pests affecting tomato crops, reappeared in Kaduna and Katsina states. It was said to have caused about 80 per cent loss of tomato production in the country, astronomically increasing the market price of the essential vegetable nationwide (Izuaka, 2020). As a result, critical measures taken are aimed at reducing the negative impact, in order to improve the productivity by small-scale farmers as well.

In addition, early detection of plant diseases is a crucial management method in prevention of large-scale damage to crops. This proactive action of early treatment will prevent spread of the diseases or advancing of the diseases to untreatable levels(Buja *et al*., 2021). Most small-scale farmers depend on experience and human observation in identification of crop infestation. While effective, these methods are most of the time inaccurate or characterized by misinformation of the actual problem.

Advancement in technology has provided an opportunity for adoption of better crop diseases management practices. Among these methods, is the integration of machine and deep learning models with IoT devices which are deployable in real-time. Early classification and detection of the various classes of diseases affecting tomato can be identified and can help farmers to have favourable yields thereby boosting sales.

This research delves into examining the impact of utilizing pre-trained CNN models as feature extractors to enhance the identification of tomato diseases and pests. The investigation employs the renowned pre-trained CNN model, MobileNetV2, for feature extraction. The extracted deep features are subsequently utilized to train the SVM classifier. Additionally, the Universal Sentence Encoder serves as a recommender system. The outcomes reveal that the fusion of these two models results in a significant achievement in accurately detecting tomato diseases and pests, achieving an impressive overall accuracy of 96.9%. The study compares its experimental findings with both internal analyses and pertinent studies in the existing literature.

The subsequent sections of the paper are organized as follows: Section 2 provides an overview of related works, Section 3 elucidates the methodology. Sections 4 and 5 encompass the Results/Discussion and Conclusion, respectively. Lastly, Section 6 offers insights into future directions for research.

2. RELATED WORKS

This section is based on personal observations and assessments of numerous research scholars' works in the area of artificial intelligence for the classification and prediction of tomato leaf diseases.

A CNN-based model was created by Khatoon *et al*. (2020) with the goal of forecasting 24 different diseases in tomato leaves and fruits. The Python programming language and the Keras library were implemented. The studies were carried out on a GPU-based workstation with 16 GB of RAM, an NVIDIA GTX 1080ti graphics card with a memory capacity of 8 GB, and CUDA with 1920 cores. Three pre-trained network models, including DenseNet, RestNet, and VGG16, were also employed for better comparison. 23,716 images from adjacent farms and publicly available datasets were used in total. 90% of the dataset was used for training and 20% for testing in order to prevent overfitting. Only the DenseNet model has a documented accuracy rate of 95.31%, according to the findings. It is advised that the training data sets be retrieved from various sources in light of ongoing research. Again, no management approach was given.

Altuntaş and Kocamaz (2021) built a deep convolutional neural network for the purpose of forecasting tomato leaf diseases with the aid of 18835 datasets from the plant village on Kaggle.com, According to the study, 10% of the dataset was used for testing, and 40% of it was used to train the model. The experiment was carried out using a PC with the following specifications: an i5-8250U CPU, 8 GB of RAM, a 2 GB GPU, and a 256 SSD HDD. According to the results, ResNet50 with 50 convolution layers had 96.99% accuracy, followed by Alexnet with 92.18% accuracy, and GoogleNet with 89.31% accuracy. Important stratagies were adopted that will optimize efficiency of the model, the authors suggested the fusion of several methodologies or methods. But additional observations showed that just 50% of the dataset was utilized, and overfitting could have an impact on the model's overall performance. Once more, no documented procedure for managing diseases and pests was found.

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 10, Issue 3, pp: (33-43), Month: September - December 2023, Available at: www.noveltyjournals.com

Chowdhury *et al.* (2021) created a model for diagnosing tomato leaf diseases using convolutional neural network based on deep learning technique. 18,162 datasets from the Plant Village dataset on Kaggle.com was downloaded and utilized. Based on deep learning's need for huge datasets, which was achievable by application of dada augmentation procedure that increases data sets. ResNet, MobileNet, DenseNet201, and InceptionV3 were the other four pre-trained models used in the trials. To improve the model's performance, some of the parameters were tuned using 70% of the datasets for training, 10% for testing, and 10% for evaluation. The PC with property of 64 GB of RAM, and a 16 GB NVIDIA GeForce ran under the Python 3.7 and the PyTorch library. Results showed that the proposed model was 98.05% accurate. Additional observations also showed that the proposed model included a feedback system that offers the proper understanding, therapies, disease prevention, and control approaches in order to increase agricultural yields, but no preventive measures were noted.

A deep learning-CNN model was developed by (Gunarathna & Rathmayaa, 2020) using the Keras framework to recognize 10 tomato illnesses from digital photographs of the leaves. The experiment was carried out using an ACER Aspire3 laptop with an Intel Core i5-8265U 1.66GHz processor, 8GB of DDR4 memory, and an NVIDIA GeForce MX230 graphics card. Along with the suggested CNN, three pre-trained models—including VGG16, MobileNet, and Inceptionv3—were used. The 22,930 datasets that were collected through the Internet were not accompanied by a source citation. The accuracy of the proposed CNN model was 90%, that of the VGG16 model was 89%, that of MobileNet was 91%, and that of Inceptionv3 was 87%, according to the results. Once more, crop management was not included in the suggested paradigm.

Using 18835 datasets from the plant village on Kaggle.com, Altuntaş & Kocamaz (2021) constructed a deep convolutional neural network for the goal of forecasting tomato leaf diseases. The study found that 40% of the dataset was used to train the model, while 10% was used for testing. A computer with the following components was used for the experiment: an i5-8250U CPU, 8 GB of RAM, a 2 GB GPU, and a 256 SSD HDD. The accuracy of ResNet50 with 50 convolution layers was 96.99%, followed by Alexnet with 92.18% accuracy and GoogleNet with 89.31% accuracy, according to the data. Important measures were taken to improve the model's effectiveness, and the authors recommended combining several approaches. However, additional observations revealed that just 50% of the dataset was used, suggesting that overfitting can affect the model's overall performance. Upon comprehensive scrutiny and evaluation of various scholarly works, the study conducted by (Altuntaş & Kocamaz, 2021) titled "Deep Feature Extraction for Detection of Tomato Plant Diseases and Pests based on Leaf Images" reveals several shortcomings. The examination reveals that merely 40% of the dataset was employed for training purposes, with a mere 10% allocated for testing. A conspicuous observation during the analysis is the omission of any mention of the remaining 50% of the dataset, a gap that may lead to overfitting, thereby casting doubt on the reliability of the outcome's accuracy.

Furthermore, the data presented in their table 1.2 highlights a multi-class classification quandary compounded by an imbalanced dataset. Given this scenario, the most suitable approach for an accurate model performance evaluation would be the utilization of micro average metrics, which regrettably were not employed. Moreover, the absence of any indication of control measures following the detection of tomato diseases poses a notable limitation in their study.

Liu and Wang (2020) used Ubuntu 16.04, Python, OpenCV, CUDA, and Mainboard to create a CNN-based model for forecasting tomato leaf diseases. WS X299 SAGE by Asus 1. Intel Core i7-9800X processor, 16 GB of DDR4 memory, GeForce GTX 1080 Ti graphics card, and 256 GB solid state drive is the configuration of the PC used in conducting the experiment. SSD, Faster R-CNN, and You Only Look Once (YOLO) version 3 were the three pre-trained models used in the proposed model. A total of 1,952 tomato leaves were collected from the field as the primary datasets, but the images from the leaves were of poor quality. This can be because of the capturing tools employed. Additional research revealed that while the Yolo v3 model is capable of learning from data to anticipate bounding boxes, it is less successful at generalizing objects with novel or odd aspect ratios or configurations. Additionally, it suffers with clusters of little things. Results showed that YOLO had a detection time of 20.39 ms and an accuracy of 92.39%.

After extensively analyzing classification models for tomato leaf diseases, considering factors like model type, dataset size, techniques, programming language, integration with recommender systems, and mobile compatibility, it's evident that many reviewed research papers employed personalized Convolutional Neural Network (CNN) models to extract features and perform classification. These studies demonstrated superior performance with their custom CNN models, albeit relying heavily on intricate hyperparameter tuning and improving the accuracy of the model, a laborious process. Moreover, unexplored latent features persist within the input data.

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 10, Issue 3, pp: (33-43), Month: September - December 2023, Available at: www.noveltyjournals.com

Presently, all existing tomato leaf disease prediction models understudied conclude at the classification stage, lacking a recommender system that would guide farmers with symptom insights and preventive measures in early disease stages. Furthermore, integration into a mobile platform, essential for farmers and horticulturists, is notably absent in most tomato leaf disease models. Model accuracy and recall/precision tradeoffs might be commendable, yet subpar performance in model integration may stem from various factors linked to model selection.

Table 2.1: Summary of related works on tomato leaf diseases using AI approach.

S/N	Author(s)/year	Image domain	Technique(s)/ Experimental environment used	Pre-trained Models	Size/source of dataset	Preprocessing of dataset	Performance Metric	Comments for Gap Analysis
1	(Zaki *et al.*, 2020)	Tomato leaf	Deep learning, Python	MobileNet V2	4,671 images from PlantVillage dataset	Training = 80%, testing=20%,	Accuracy = 90%	No pest & disease management
2	(ALTUNTAŞ & KOCAMAZ, 2021)	Tomato Leaf	Deep learning -CNN MATLAB 2019b,	3 Pre-trained CNN models; Alexnet, GoogleNet ResNet50	18,835 from Plant Village (Kaggle.com)	Data segmentation Training=40% Testing=10%	Accuracy Alexnet =92.18 GoogleNet =89.31 ResNet50 = 96.99%.	No procedure on how to manage pest and diseases
3	(Hong *et al.*, 2020)	Tomato Leaf	Deep learning-CNN	ShuffleNet, Xception, Resnet50, MobileNet, and DenseNet121 Xception	41263 images from plant village (Kaggle.com)	Data augmentation Training=80% Testing=20%	Accuracy DenseNet121_ Xception= 91.10%	No pest & disease management
4	(Liu & Wang, 2020)	Tomato Leaf	Deep learning –CNN Ubuntu 16.04, python, OpenCV, CUDA,	3 pre-trained models; SSD, Faster R-CNN and YOLO v3	1,952 from Field collection Images collected are of low quality	NA	Accuracy YOLO V3 = 92.39% detection time =20.39 ms	No management strategy
59	(Prajwala et al., 2020)	Tomato leaf	Deep Learning-CNN Keras, a neural network API written in Python	LeNet	18,160 images from Plant Village (Kaggle.com)	Training=70% Testing=30%	Accuracy = 94-95%	No pest & disease management
6	(Gonzalez-Huitron *et al.*, 2021)	Tomato Leaf	Deep learning & Machine learning IDE Spyder, OpenCV	MobileNetV2 Xception MobileNetV3	18,215 images from plant village (Kaggle.com)	Data augmentation Training=70% Testing=30%	Accuracy Xception = 94.5%	No pest & soil management
7	(Hiremath & Gowda, n.d.)	Tomato leaf	Deep learning-CNN	VGG16, InceptionV3 and MobileNet.	5,000 images from Plant village dataset (Kaggle.com)	Data augmentation Image segmentation	Accuracy CNN = 91.2%	No management strategy given
8	(Agarwal *et al.*, 2020)	Tomato Leaf	Deep learning-CNN	VGG16 MobileNet, and InceptionV3	17,500 images from plant village (Kaggle.com)	Training=70% Testing=10% Validation=20%	Accuracy = 91.2%	No pest & disease management

9	(Damicone & Brandenberger, 2021)	Tomato Leaf	Deep Learning	DNN CNN RNN FFNN BPNN MAT- LAB®	1,500 images from field collection	Training=1300 Testing =200	Overall Accuracy: DNN=86.18%	No management measure
10	(Mkonyi *et al.*, 2020)		Deep learning-CNN	VGG16, VGG19, ResNet50,	2,145 from internet	Data augmentation, Training=85% Testing=15%	Accuracy VGG16=91.9%	No control measure

3. METHODOLOGY

3.1.a Describing Components: The integrated system comprises four key components that synergistically enhance tomato disease management:

i. **MobileNetV2 (Pre-trained CNN Model):** MobileNetV2 is a well-established pre-trained Convolutional Neural Network (CNN) model that excels in feature extraction from images. Its architecture is optimized for mobile devices, ensuring efficient computation while maintaining robust performance.

ii. **Support Vector Machines (SVM):** SVM, a powerful supervised learning algorithm, utilizes the extracted deep features from MobileNetV2 to classify tomato leaf images into disease categories. It creates an optimal hyperplane to separate different classes, enabling accurate disease identification.

iii. **Universal Sentence Encoder (USE):** The Universal Sentence Encoder transforms textual symptom descriptions into high-dimensional vector representations. It captures semantic meaning and context, enabling effective comparison and recommendation of disease-related information.

iv. **Recommender System:** The recommender system integrates the outputs from SVM and USE. It analyzes the disease classification results and symptom descriptions to provide tailored recommendations to farmers. By matching symptoms with potential diseases and suggesting preventive measures, the system empowers informed decision-making.

b. Explaining Component Operations: MobileNetV2 employs depth-wise separable convolutions and inverted residuals to efficiently capture hierarchical features in images. It learns to extract distinctive patterns from tomato leaf images, which are subsequently utilized by the SVM for classification. SVM, on the other hand, computes decision boundaries using the extracted features to distinguish between various disease classes. USE utilizes a transformer architecture to encode sentences into fixed-length vectors. It captures the semantic essence of symptom descriptions, making them amenable to comparison. The recommender system combines SVM's disease predictions and USE's encoded symptom vectors. It quantifies the similarity between symptom vectors and known disease vectors, generating relevant recommendations for farmers.

3.2 Data Collection and Preprocessing

This section encompasses the methodology for gathering data employed in this research endeavor. The researcher adopts a dual approach, drawing upon both primary and secondary data sources to comprehensively grasp the study's context. For the evaluation of model efficacy and system performance using real-time imagery, primary data collection involves procuring live photographs depicting both healthy and diseased tomato specimens. Supplementary data stems from the Kaggle platform, a hub that furnishes datasets for the study of AI and intelligent systems. Kaustubh (2019) curated the dataset, subsequently making it available on Kaggle. With a size of 188 MB, the dataset comprises 11,000 files organized into training and validation set folders. Table 1 provides an inclusive list of disease classifications that impact tomato leaves, including a reference to healthy leaves. Preprocessing involves data augmentation techniques such as rotation, scaling, and flipping to enhance model robustness. Images are normalized to ensure consistent feature extraction across the dataset. Textual symptom descriptions are encoded into USE-compatible formats, preserving their semantic meaning for effective recommendation.

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 10, Issue 3, pp: (33-43), Month: September - December 2023, Available at: www.noveltyjournals.com

Table 3.1

S/N	Disease Name	Family Name	Symptoms
1	Mosaic Virus	Tobamoviridae	Tiles, wrinkle, reduction and curvature of leaflets, and irregular ripening of fruits.
2	Tomato Blind Spot	Corynesporascaceae	Pinpoint-sized, water-soaked spots on the upper leaf surface, Initial symptoms on stems and petioles are lesions that are slightly sunken, brown flecks.
3	Bacteria Spot	Xanthomonadaceae	Small, round, water-soaked spots that gradually turn dark-brown or black and are surrounded by yellow halo. The size of lesions is fairly variable, but rarely develops to more than 1/10 inch in diameter.
4	Tomato yellow leaf curl virus	Geminiviridae	Small leaves that become yellow between the veins. The leaves also curl upwards and towards the middle of the leaf.
5	Tomato late blight	Peronosporaceae	Tomato leaves are irregularly shaped, water-soaked lesions, often with a lighter halo or ring around them.
6	Leaf Mold	Mycosphaerellaceae	The oldest leaves are infected first. Pale greenish-yellow spots, usually less than 1/4 inch, with no definite margins, form on the upper sides of leaves.
7	Early Blight	Pleosporaceae	Small dark spots form on older foliage near the ground. Leaf spots are round, brown and can grow up to 1/2 inch in diameter
8	Tomato red spider mite	Tetranychida	Tomato red spider mite feeding causes whitening or yellowing of leaves, which then dry out and eventually fall off.
9	Septoria leaf spot	Mycosphaerellaceae	Appears on the lower leaves after the first fruit sets. Spots are circular, about 1/16 to 1/4 inch in diameter with dark brown margins and tan to gray centers with small black fruiting structures
10	Tomato Healthy	Solanaceae	Healthy

3.2.1 System Component Architecture

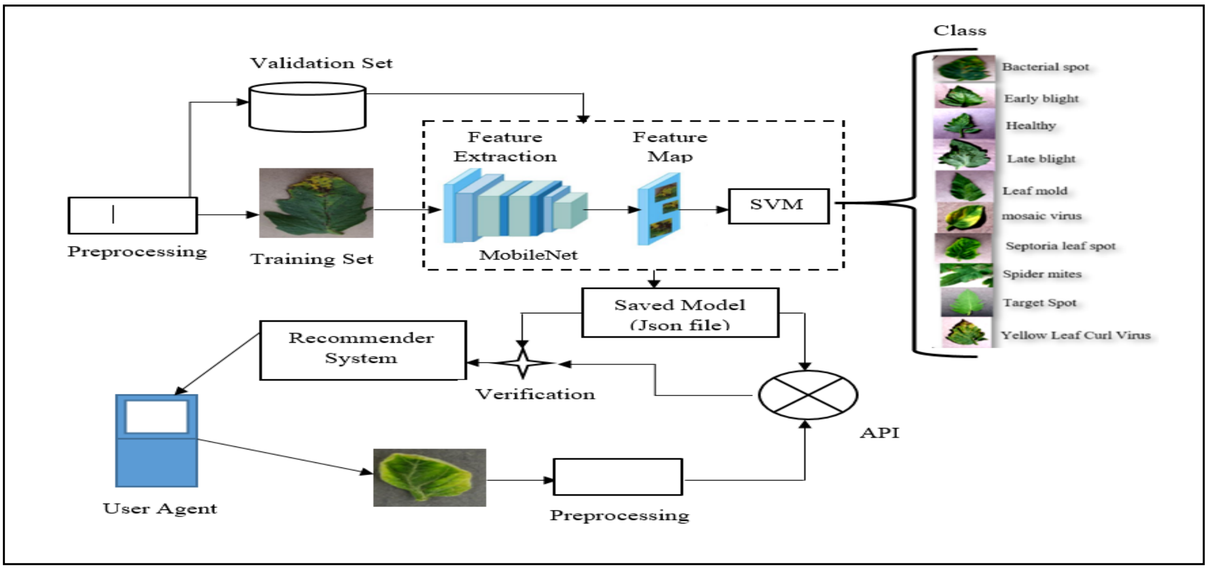

Figure 1: components of model architecture

Novelty Journals
ISSN 2394-7314
International Journal of Novel Research in Computer Science and Software Engineering
Vol. 10, Issue 3, pp: (33-43), Month: September - December 2023, Available at: www.noveltyjournals.com

3.2.2 Data Flow Diagram (DFD)

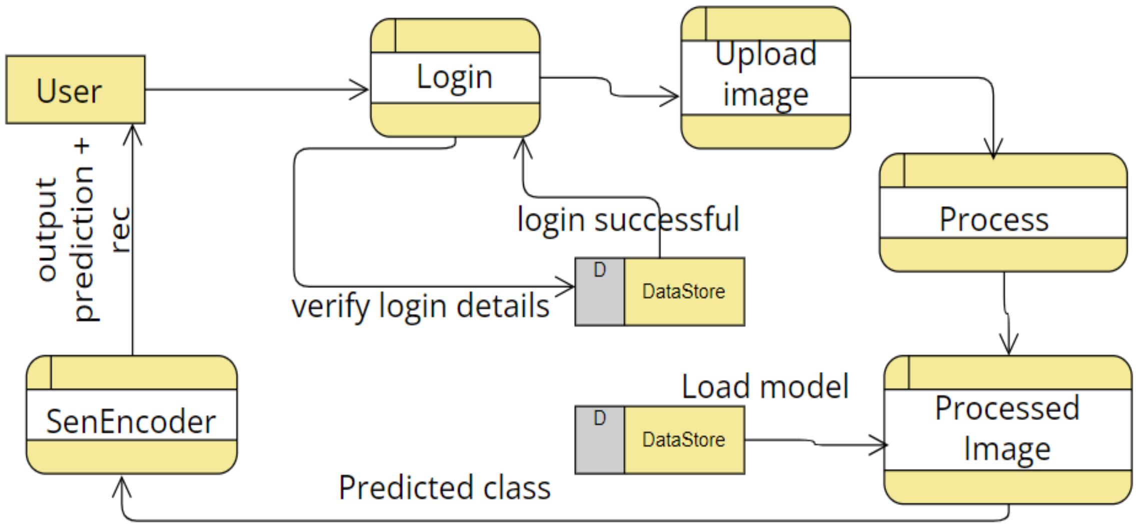

Figure 2: DFD of the proposed system

The diagram in Figure 2 of the proposed system's Data Flow Diagram illustrates the progression of information from the user, an external entity, through various stages to the final output, which then returns to the user. When utilizing the software, individuals must initially collect different tomato leaves. Subsequently, they can either create an account or log in using existing credentials. The system includes a verification stage to ascertain if the entered password was previously recorded during enrollment. Upon receiving a response from the database via the server, the user can upload an image and perform predictions. The procedural steps for these processes remain undisclosed to end users. Since the intended users are farmers or botanists aiming to identify tomato leaf diseases and receive advice on disease control through the proposed model, all stages of the process, including up to the SenEncoder phase, are maintained as confidential components.

4. RESULTS

For our experimental setup, we conducted a rigorous process to train and evaluate the performance of our integrated system for tomato disease classification. We gathered a comprehensive dataset comprising 11,000 tomato leaf images, categorized into healthy and diseased classes. This dataset was divided into training and validation sets. Prior to training, images were preprocessed by resizing them to a uniform dimension of 224x224 pixels and normalizing pixel values. We employed the MobileNetV2 architecture as a feature extractor, followed by SVM for classification. The Universal Sentence Encoder was used to process symptom descriptions. The system was trained using the Adam optimizer and categorical cross-entropy loss function. Performance assessment was conducted using a hold-out validation set. The integrated system's precision, F-1 score, accuracy, and recall were the primary metrics employed. The performance metrics of our integrated system on the tomato disease classification task are as follows: Precision: 93.68%, F-1 Score: 93.84%, Accuracy: 96.81%, Recall: 95.4%. These metrics collectively illustrate the high accuracy and reliability of our system in identifying and classifying tomato diseases. To ascertain the significance of our findings, we compared our integrated system's performance with other pre-trained models reported in related studies. Our system consistently outperformed alternative models, exhibiting notably higher precision, F-1 score, accuracy, and recall rates. The superior performance of our integrated system emphasizes its potential as a robust tool for real-world implementation, aiding farmers in early disease identification and effective agricultural management practices.

5. DISCUSSIONS

The mobile performance of our integrated system is reflected in the achieved accuracy of 96.9% in tomato disease and pest detection. This high accuracy rate underscores the effectiveness of the MobileNetV2-SVM fusion in precisely identifying a wide range of diseases. These results hold significant implications for tomato disease management, as our approach equips farmers with a reliable tool to swiftly identify and address potential threats to their crops. Timely intervention based on accurate disease identification can lead to improved crop yield and reduced economic losses. One of the key strengths of our integrated system is its ability to provide holistic disease management support, combining image-based disease identification with textual symptom analysis and preventive recommendations. This comprehensive approach enhances decision-making for farmers. During the development process, we encountered a challenge related to the diversity of disease symptoms and their visual representation in images. Some diseases exhibited subtle variations, making them challenging to distinguish accurately. Additionally, unexpected variations in lighting conditions and image quality posed challenges to consistent feature extraction. Despite these challenges, the hybrid model demonstrated remarkable adaptability and robustness, underscoring its potential to handle variations inherent in real-world scenarios. The initial problem centered around enhancing tomato disease management through an integrated mobile system. Our findings directly address this problem by presenting a comprehensive solution that amalgamates advanced technology with practical utility. The accurate disease classification and proactive recommendations provided by the system address the primary concern of timely disease detection and management. Furthermore, the integration of SVM and USE adds a novel dimension by enabling symptom-based recommendations. Thus, our findings not only validate the effectiveness of our approach but also showcase its potential to revolutionize tomato agriculture through informed decision-making and sustainable practices.

6. CONCLUSION

The fundamental challenge at the heart of this research was the effective management of tomato diseases for sustainable agriculture. Addressing this, our approach involved the development of an integrated mobile system that combined advanced machine learning techniques, including MobileNetV2 for image feature extraction, SVM for classification, USE for symptom analysis, and a recommender system for informed decision-making. Our investigation yielded compelling results. The integrated system demonstrated remarkable accuracy, achieving an impressive 96.9% success rate in detecting and classifying tomato diseases. This outcome signifies a substantial leap forward in early disease identification, enabling timely interventions. The significance lies in equipping farmers with a powerful tool to protect their tomato crops, enhancing yield and reducing the environmental impact of disease control methods. The potential impact of our integrated system on sustainable tomato agriculture is substantial. By providing farmers with timely disease identification, tailored recommendations, and preventive measures, we empower them to make informed decisions. This translates to reduced crop losses, optimized resource utilization, and ultimately contributes to a more ecologically balanced and economically viable tomato cultivation process. As we look ahead, there are promising avenues for further exploration. First, optimizing the recommender system's algorithms could enhance the precision and relevance of recommendations, enabling even more effective disease management. Moreover, extending our integrated approach to other crops holds the potential to revolutionize agriculture on a broader scale. Investigating the system's adaptability to different agricultural contexts and exploring its integration with emerging technologies like IoT sensors and blockchain could open new dimensions for research and practical implementation.

REFERENCES

[1] Adekunle, C. P. 2 (2019) Differentials in willingness to pay for biofortified tomato fruits in Abeokuta metropolis, Ogun state, Nigeria." Nigeria Agricultural Journal 50, no.: 1-9.

[2] Ahmad, I., Hamid, M., Yousaf, S., Shah, S. T., & Ahmad, M. O. (2020). Optimizing pretrained convolutional neural networks for tomato leaf disease detection *Complexity*, *2020*. https://doi.org/10.1155/2020/8812019

[3] ALTUNTA, Y., & KOCAMAZ, F. (2021). Deep Feature Extraction for the Detection of Tomato Plant Diseases and Pests Based on Leaf Images *Celal Bayar Üniversitesi Fen Bilimleri Dergisi*, *17*(2), 145–152. https://doi.org/10.18466/cbayarfbe.812375

[4] Baudoin, W., Nersisyan, A., Shamilov, A., Hodder, A., & Gutierrez, D. (2017). Good Agricultural Practices for greenhouse vegetable production in the Southeast European countries Good Agricultural Practices for greenhouse vegetable production in the Southeast European countries http://www.fao.org/3/a-i6787e.pdf

[5] Bhowmik, D., Kumar, K. P. S., Paswan, S., & Srivastava, S. (2012). Tomato: A Natural Medicine and Its Health Benefits Phytojournal, 1(1), 33–43.

[6] Buja, I., Sabella, E., Monteduro, A. G., Serena, M., Bellis, L. De, Luvisi, A., & Maruccio, G. (2021). Advances in Plant Disease Detection and Monitoring: From Traditional Assays to In-Field Diagnostics

[7] Chowdhury, M. E. H., Rahman, T., Khandakar, A., Ayari, M. A., Khan, A. U., Khan, M. S., Al-Emadi, N., Reaz, M. B. I., Islam, M. T., & Ali, S. H. M. (2021). Automatic and Reliable Leaf Disease Detection Using Deep Learning Techniques AgriEngineering, 3(2), 294–312. https://doi.org/10.3390/agriengineering3020020

[8] Damicone, J. P., & Brandenberger, L. (2021). Common Diseases of Tomatoes Figure 1, 1–6. http://pods.dasnr.okstate.edu/docushare/dsweb/Get/Document-1323/EPP-7625web.pdf

[9] Domingues, T., Brando, T., & Ferreira, J. C. (2022) Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey Agriculture, 12(9), 1350. https://doi.org/10.3390/agriculture12091350 Environmental Sciences, "Cite." 28.311. https://doi.org/10.14569/IJACSA.2022.0130138

[10] FAO. (2017). The future of food and agriculture: trends and challenges In The Future of Food and Agriculture: Trends and Challenges (Vol. 4, Issue 4), www.fao.org/publications%0A http://www.fao.org/3/a-i6583e.pdf%0A http://siteresources.worldbank.org/INTARD/825826-1111044795683/20424536/Ag_ed_Africa.pdf %0A www.fao.org/cfs%0A http://www.jstor.org/stable/4356839%0A https://ediss.uni-goettingen.de/bitstream/han

[11] FAO. (2019). Integrated Pest Management in Tomatoes in Eritrea In Sustainable Management of Arthropod Pests of Tomato (Vol. 2, Issue 1) https://doi.org/10.1016/B978-0-12-802441-6.00012-7%0A http://dx.doi.org/10.1016/B978-0-12-394314-9.00002-6%0A https://doi.org/10.1016/B978-0-12-802441-6.00004-8%0A https://www.weforum.%0A https://doi.org/10.1016/B978-0-12-802441-6.00002-4%0A www.elsevier.com/l

[12] Frona, D., Janos, S., & Harangi-Rakos, M. (2019). The challenge of feeding the poor MDPI Journal for Sustainability, 2017, 3–4.

[13] Gunarathna, M. M., Rathnayaka, R. M. K. T., & Kandegama, W. M. W. (2020). Identification of an Efficient Deep Leaning Architecture for Tomato Disease Classification Using Leaf Images Journal of Food and Agriculture, 13(1), 33 https://doi.org/10.4038/jfa.v13i1.5230

[14] Hiremath, V. B., & Gowda, S. (n.d.) DISEASE PREDICTION OF TOMATO LEAF USING CNN AND DEEP LEARNING TECHNIQUES in the International Research Journal of Modernization in Engineering, Technology, and Science. www.irjmets.com

[15] Hong, H., Lin, J., & Huang, F. (2020). Tomato Disease Detection and Classification by Deep Learning *Proceedings of the 2020 International Conference on Big Data, Artificial Intelligence, and Internet of Things Engineering, ICBAIE 2020*, 25–29 https://doi.org/10.1109/ICBAIE49996.2020.00012

[16] Islam, M. P., Hatou, K., Aihara, T., Seno, S., Kirino, S., & Okamoto, S. (2022). Performance prediction of tomato leaf disease by a series of parallel convolutional neural networks Smart Agricultural Technology, 2 (March), 100054. https://doi.org/10.1016/j.atech.2022.100054

[17] Izuaka, M. (2020a). Traders report a tomato shortage as a pest reappears in Kano. In premiumtimesng. https://www.premiumtimesng.com/agriculture/agric-news/394536-traders-report-tomato-shortage-as-pest-reappears-in-kano.html. Federal Ministry of Agriculture and Rural Development (FMARD,2022)?tztc=1

[18] Khatoon, S., Hasan, M. M., Asif, A., Alshmari, M., & Yap, Y. K. (2021). Image-based automatic diagnostic system for tomato plants using deep learning *Computers, Materials, and Continua*, *67*(1), 595–612. https://doi.org/10.32604/cmc.2021.014580

[19] Kheyrodin, H., & Kheyrodin, S. (2017). Importance of the Tomato as such as medical plant. International Journal of Advanced Research in Biological Sciences (IJARBS), 4(4), 106-115

[20] Kumar, A., & Vani, M. (2019). Image-Based Tomato Leaf Disease Detection *2019 10th International Conference on Computing, Communication, and Networking Technologies, ICCCNT 2019*, 1–6 https://doi.org/10.1109/ICCCNT45670.2019.8944692

[21] Liu, J., & Wang, X. (2020). Tomato Diseases and Pest Detection Based on an Improved Yolo V3 Convolutional Neural Network Frontiers in Plant Science, 11 (June), 1–12. https://doi.org/10.3389/fpls.2020.00898

[22] Lu, J., Ehsani, R., Shi, Y., de Castro, A. I., & Wang, S. (2018). Detection of multi-tomato leaf diseases (late blight, target, and bacterial spots) in different stages by using a spectral-based sensor Scientific Reports, 8(1) https://doi.org/10.1038/s41598-018-21191-6

[23] Mkhabela, T. (2005). Technical efficiency in a vegetable-based mixedcropping sector in Tugela Ferry, Msinga District, Kwazulu-Natal, Agrekon, 44(2), 187–204. https://doi.org/10.1080/03031853.2005.9523709

[24] Mkonyi, L., Rubanga, D., Richard, M., Zekeya, N., Sawahiko, S., Maiseli, B., & Machuve, D. (2020). Early identification of Tuta absoluta in tomato plants using deep learning Scientific African, 10. https://doi.org/10.1016/j.sciaf.2020.e00590

[25] Modi, M., Modi, A., & Hendriks, S. (2006). POTENTIAL ROLE FOR WILD VEGETABLES IN HOUSEHOLD FOOD SECURITY. African Journal of Food, Agriculture, Nutrition, and Development, 6(1), 1–13.

[26] Nell, W., Wessels, B., Mokoka, J., & Machedi, S. (2006). A creative multidisciplinary approach towards the development of food gardening. Development Southern Africa, 17(5), 807–819. https://doi.org/10.1080/03768350020013642

[27] Nithish, E. K., Kaushik, M., Prakash, P., Ajay, R., & Veni, S. (2020). Tomato leaf disease detection using a convolutional neural network with data augmentation Proceedings of the 5th International Conference on Communication and Electronics Systems, ICCES 2020, Icces, 1125–1132. https://doi.org/10.1109/ICCES48766.2020.09138030

[28] Oladejo, J. A., & Laudia, O. (2011). Technical efficiency of tomato production in Oyo State. June.

[29] Prajwala TM, Alla Pranathi, Kandiraju Sai Ashritha, Nagaratna B. Chittaragi, and S. G. K. (2018). Tomato Leaf Disease Detection and Classification Using Convolutional Neural Networks *Proceedings of the 2020 Innovations in Intelligent Systems and Applications Conference, ASYU 2020*, 2–4. https://doi.org/10.1109/ASYU50717.2020.9259832.

[30] Ramakrishna, R. (2020). A Machine Learning-Based Approach to the Detection and Classification of Tomato Plant Leaf Diseases http://trap.ncirl.ie/4275/

[31] Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS ONE, 8(6) https://doi.org/10.1371/journal.pone.0066428

[32] Roser, M., & Ortiz-Ospina, E. (2013). World Population Growth for Semester 2 (H) Geography Students Paper-CC3 (TH) Topic-7. 2, 1–21

[33] Trivedi, N. K., Gautam, V., Anand, A., Aljahdali, H. M., Villar, S. G., Anand, D., Goyal, N., & Kadry, S. (2021). Early detection and classification of tomato leaf disease using a high-performance deep neural network Sensors, 21(23). https://doi.org/10.3390/s21237987

[34] Vijay, N. (2021). Detection of Plants. Using Deep Learning Techniques AgriEngineering, 3(2), 294–312. https://doi.org/10.3390/agriengineering3020020

[35] Zaki, S. Z. M., Zulkifley, M. A., Mohd Stofa, M., Kamari, N. A. M., & Mohamed, N. A. (2020). Classification of tomato leaf diseases using MobileNet v2 *IAES International Journal of Artificial Intelligence*, *9*(2), 290–296, https://doi.org/10.11591/ijai.v9.i2.pp290-296