



OPEN ACCESS

EDITED BY

Jo Burgess,
Isle Group Ltd, United Kingdom

REVIEWED BY

Ferdinand Oswald,
The University of Auckland, New Zealand

*CORRESPONDENCE

An Thanh Le
✉ let25@myunitec.ac.nz

RECEIVED 07 February 2024

ACCEPTED 29 April 2024

PUBLISHED 20 May 2024

CITATION

Le AT, Shakiba M and Ardekani I (2024)
Tomato disease detection with lightweight
recurrent and convolutional deep learning
models for sustainable and smart agriculture.
Front. Sustain. 5:1383182.
doi: 10.3389/frsus.2024.1383182

COPYRIGHT

© 2024 Le, Shakiba and Ardekani. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Tomato disease detection with lightweight recurrent and convolutional deep learning models for sustainable and smart agriculture

An Thanh Le^{1*}, Masoud Shakiba¹ and Iman Ardekani^{1,2}

¹School of Computing, Electrical and Applied Technology, Unitec – Te Pukenga, Auckland, New Zealand, ²School of Arts and Sciences, The University of Notre Dame Australia, Fremantle, WA, Australia

KEYWORDS

smart agriculture, plant disease detection, deep learning, Convolutional Neural Network, Recurrent Neural Network, Liquid Time-Constant Networks, internet of things, sustainable agriculture

1 Introduction

The detection of plant diseases is a critical concern in agriculture, as it directly impacts crop health, yields, and food security (Fang and Ramasamy, 2015). Traditionally, this task has relied heavily on the observations of farmers and agricultural experts, which is fraught with many shortcomings, including human error and the inability to identify latent or early-stage infections. In response to these limitations, the scientific community has developed multiple innovative solutions. Among these, image classification techniques have gained widespread adoption due to their cost-efficiency (Chhillar et al., 2020) and the ability to enable real-time monitoring, allowing farmers to promptly detect diseases and take timely action (Chen et al., 2020). Additionally, these techniques are highly scalable, adaptable, non-invasive, and non-destructive and can be applied to different crops and disease scenarios (Ramcharan et al., 2017).

However, despite these advantages, image classification methods come with their unique set of challenges. One of the most prominent issues is finding the delicate balance between computational cost and accuracy (Barbedo, 2016). Researchers have delved into various strategies to tackle this challenge, including model pruning (Jiang et al., 2022), transfer learning (Shaha and Pawar, 2018), and hybrid models (Tuncer, 2021). The proposed research seeks to make a valuable contribution to this field by creating a new combination of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), known as hybrid CNN-RNN models for detecting tomato plant diseases. This model effectively addresses the computational cost challenge while upholding high accuracy.

The successful implementation of the model marks a significant milestone, paving the way for real-world applications of disease detection in the agriculture industry. The benefits extend beyond reduced crop loss and increased crop quality, leading to higher farmer income and strengthened food security. It also contributes to global economic improvement, reduced pesticide use that positively impacts the environment, and enhanced relationships among farmers, researchers, and corporations.

2 Objectives

The primary goal of this research is to develop and optimize lightweight CNN-RNN models for effectively detecting tomato plant diseases using images. This objective stems from the necessity to create efficient and accessible solutions for disease detection in agricultural settings, particularly where resources are limited. To reach this objective, we will concentrate on several key aspects. Firstly, this research delves into various integrations of CNNs and RNNs to capture both spatial and sequential information within plant images. Additionally, we will evaluate the effectiveness of incorporating Liquid Time-Constant Networks (LTC) (Hasani et al., 2021) alongside CNN models for the image classification task, aiming to understand how LTC enhances our models' ability to capture temporal dependencies within plant images.

Additionally, we will explore the feasibility of deploying our optimized models on Raspberry Pi IoT devices. These devices have become widely used in various image-processing applications worldwide due to their affordability, speed, and efficiency (Kondaveeti et al., 2022). In the agricultural industry, Raspberry Pi devices have found particular prevalence. For example, studies highlighted in Mhaski et al. (2015) and Mustaffa and Khairul (2017) demonstrate the use of Raspberry Pi for real-time image processing to evaluate fruit maturity based on color and size, utilizing CNN models. In another domain, research (Wardana et al., 2021) addresses the challenge of accurate air quality monitoring on resource-constrained edge devices by designing a novel hybrid CNN-RNN deep learning model for hourly PM2.5 pollutant prediction, which is also implemented on Raspberry Pi. Our evaluation of Raspberry Pi will prioritize assessing inference times and energy consumption to ensure that the model remains both functional and efficient within these constraints, facilitating its practical application in real-world scenarios.

By achieving these targets, we aim to make a valuable contribution to the field of plant disease detection, providing accurate and lightweight solutions accessible to a broader agricultural community. Moreover, targeting to solve smart agriculture problems also aligns with and contributes to various Sustainable Development Goals (SDGs).

3 Dataset

The data collection process for this research involves using the Tomato Leaves dataset, which is sourced from Kaggle (Motwani and Khan, 2022). This dataset comprises over 20,000 images categorized into 10 different diseases and a healthy class as shown in Figure 1. The images are collected from two distinct environments: controlled lab settings and real-world, in-the-wild scenes. This diverse dataset provides a comprehensive representation of tomato plant conditions, making it suitable for training and evaluating plant disease detection models.

4 Methodology

The novelty of our approach relies on a combination of transfer learning, and specific neural network components such as CNN,

RNN, and LTC models. Deep Neural Networks are at the core of our research, mimicking the complexity of the human brain to excel in learning intricate patterns from raw data. Particularly suited for image classification tasks like plant disease detection, these networks play a pivotal role in extracting hierarchical features from tomato leaf images.

4.1 Transfer learning

To enhance the efficiency of our disease detection models, we plan to implement a technique known as transfer learning. This method aims to reduce computational costs and training time by leveraging knowledge gained from solving one problem and applying it to a related problem (Torrey and Shavlik, 2010). Unlike traditional machine learning approaches, which train models from scratch on specific datasets, transfer learning reuses pre-trained models and adapts them to new tasks.

Transfer learning involves transferring knowledge from a source domain, where labeled data is abundant, to a target domain, where labeled data may be limited. By doing so, it reduces the need for large amounts of labeled data for training new models, making it more practical and cost-effective in real-world scenarios.

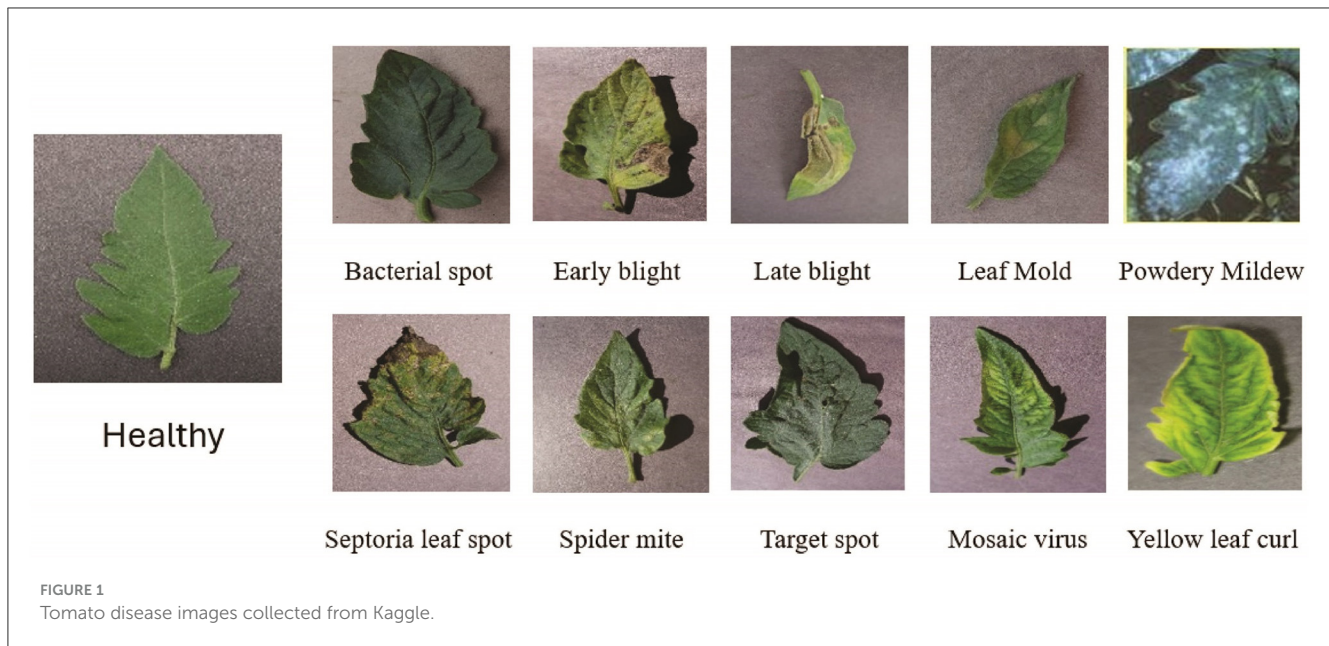
There are two primary types of transfer learning. The first is feature extraction, where the learned representations (features) from a pre-trained model are extracted and used as input to a new model. This involves removing the output layer of the pre-trained model and adding a new output layer tailored to the new task, allowing the model to be fine-tuned on the target dataset.

The second type, which we will apply in this research, is fine-tuning. Instead of freezing the parameters of the pre-trained model, the entire model is further trained on the target dataset with a small learning rate. This enables the model to adapt its learned representations to better suit the different characteristics of the new task while still retaining the knowledge gained from the source domain.

4.2 Convolutional Neural Network

Among neural network types, CNNs are ideal for image tasks that can capture spatial features through convolutional layers (Albawi et al., 2017). CNN is a powerful class of deep learning models inspired by the human visual system. CNNs excel at extracting meaningful features from images, making them highly effective tools in computer vision applications. Therefore in this research, we will use CNNs to analyze tomato leaf images and extract visual cues linked to different diseases.

CNNs typically consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. Convolutional layers perform the feature extraction process by applying convolution operations. This filter captures local patterns such as edges, textures, and shapes. By repeatedly applying convolutional operations across the entire image, CNNs can learn hierarchical representations of features, starting from simple patterns in the lower layers to more complex ones in the higher layers. Pooling layers reduce the spatial dimensions of the feature



maps, making the network more computationally efficient while preserving important information. Finally, fully connected layers combine the extracted features to make predictions.

One of the key advantages of CNNs is their ability to automatically learn hierarchical representations of features directly from raw pixel data. This end-to-end learning approach eliminates the need for manual feature engineering, allowing CNNs to adapt to a wide range of visual recognition tasks.

However, not all information can be deduced from static spatial features alone, so we will implement RNNs for modeling sequential data and capturing temporal dependencies to further improve the accuracy of the model.

4.3 Recurrent Neural Network

RNN is a class of artificial neural networks specially designed to process sequential data, such as time series, text, and speech (Salehinejad et al., 2017). Unlike traditional feedforward neural networks (like CNNs), RNNs have connections that form directed cycles, allowing them to maintain internal memory and capture temporal dependencies within the input data.

At each time step, an RNN takes an input vector and combines it with the previous hidden state to produce an output and update the hidden state. This process is repeated iteratively for each element in the sequence. The hidden state acts as a memory unit that retains information from previous time steps, allowing the network to incorporate context and make predictions based on the entire sequence.

One of the key advantages of RNNs is their ability to handle input sequences of varying lengths, making them well-suited for tasks such as natural language processing, speech recognition, and time series prediction. Additionally, RNNs can learn long-term dependencies in sequential data, thanks to their recurrent connections and memory cells.

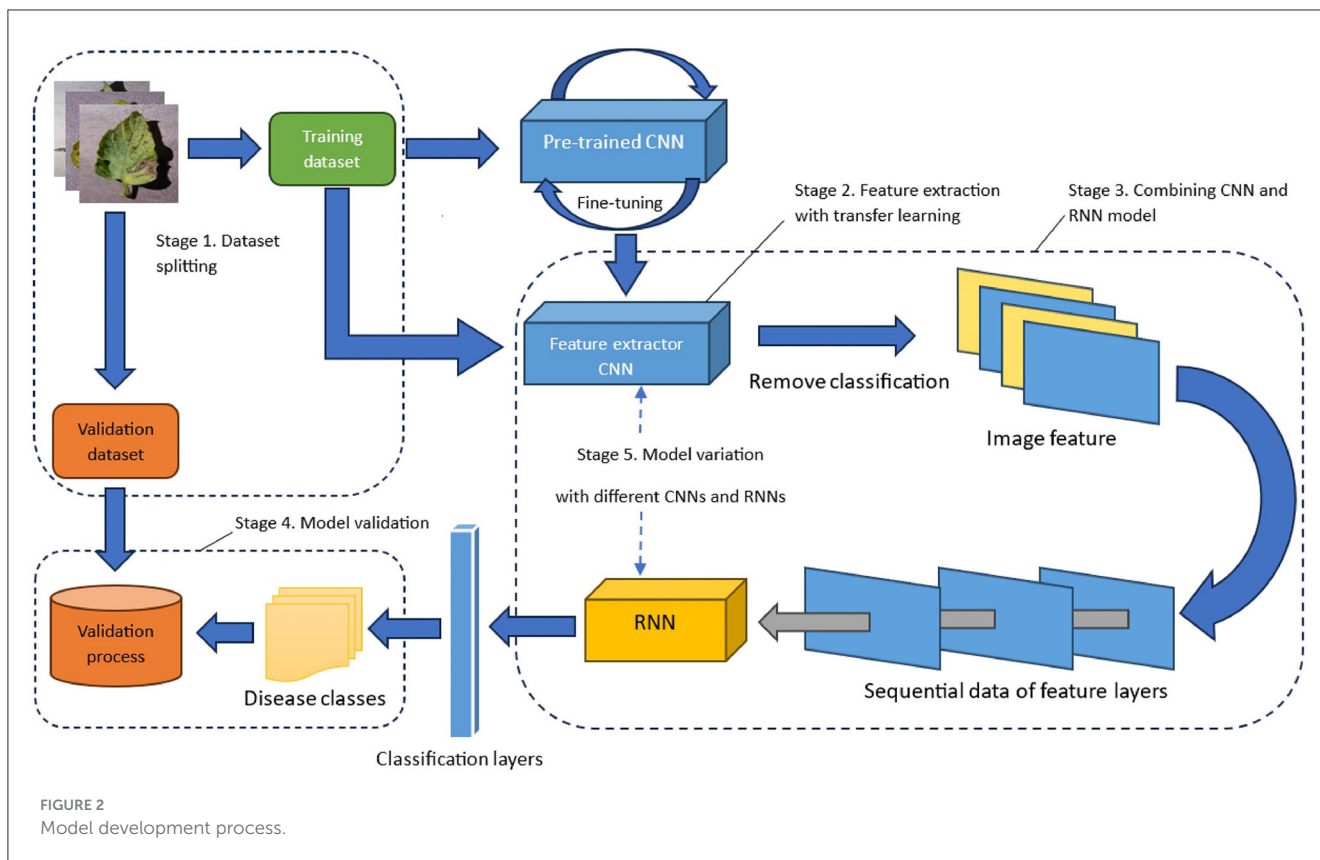
This research has utilized RNNs as we realize that images also contain sequential information. For example, for handwriting recognition tasks (Dutta et al., 2018), after CNNs have extracted the features of the image—each feature stands as a separate character—instead of building another convolutional layer to detect the whole string, we can pass those “character” features as sequential data into RNNs, and it can easily predict the output.

By using similar mechanisms, we can detect special patterns in different disease classes in our tomato leaf images with a lower computational burden on the CNNs structure, resulting in a more efficient model.

It is also worth noting that previous studies on CNN-RNN models have predominantly focused on the combination of CNNs with the Long Short-Term Memory (LSTM) network. However, these researchers commonly face challenges in dealing with image variability and noise in real-world data. The LSTM model struggles to handle these issues effectively, resulting in reduced accuracy in practical scenarios. In response to this challenge, we propose to combine CNNs with the new LTC model, drawing inspiration from liquid-state machines. By dynamically adjusting their hidden state over time in response to input data, LTC models autonomously identify infected regions in plant images. This innovative approach enhances robustness and generalization across various disease symptoms in different real-life scenarios, overcoming the limitations associated with traditional hybrid CNN-RNN approaches.

In this research, a systematic approach in Figure 2 is followed to develop and evaluate models for accurate disease detection while considering computational efficiency. The process involves five key stages, each contributing to the success of the research.

- i. Dataset splitting: after data collection, the dataset is divided into two subsets: training and validation data. This partition, typically with an 80% training and 20% validation ratio, is vital for model evaluation and to prevent overfitting.



- ii. Feature extraction with transfer learning: transfer learning is a fundamental aspect of this project. We will implement multiple pre-trained CNN models as feature extractors and validate their effect on the hybrid model.
- iii. Combining CNN and LTC model: the next phase involves combining the fine-tuned CNN model with the LTC model. Initially, the classification layer of the CNN model is removed, leaving it prepared for feature extraction rather than class prediction. The output features from the CNN model are concatenated into a sequential format before inputting into the LTC model, allowing it to capture sequential relationships between these features.
- iv. Model validation: metrics such as accuracy, precision, recall, F1-score, and confusion matrix are calculated to assess the model's effectiveness in disease detection. Hyperparameter tuning will be conducted to optimize the model's performance, exploring variations in parameters like the number of RNN units, learning rates, and batch sizes.
- v. Model variations: in the project's final phase, various model combinations are explored to assess accuracy, training time, and computational cost. Emphasis is placed on developing lightweight architectures suitable for resource-constrained environments like IoT devices.

At the end of the research, we plan to deploy the model that can achieve the best balance between accuracy and computational cost on an IoT device like Raspberry Pi and validate the accuracy with actual tomato plant images captured from our lab.

5 Discussion

Our research has culminated in the development of an exceptionally efficient hybrid CNN-RNN model. This model significantly reduces training time and computational costs, making it highly suitable for resource-constrained environments, particularly IoT and edge devices. Despite its simplified design, it maintains an impressive level of accuracy in disease detection when compared to more complex models. Furthermore, the innovative implementation of this new LTC model substantially enhances the model's ability to handle noisy and outlier data.

Moreover, this research offers significant potential to contribute to various Sustainable Development Goals (SDGs), addressing challenges not only in New Zealand Aotearoa, where the research is conducted, but also in other regions around the world. By enabling early and precise plant disease detection, this research helps to minimize the damage caused by diseases, and farmers can preserve a larger portion of their harvest, thereby safeguarding their livelihoods. With fewer crop losses, farmers can generate higher yields, leading to increased incomes and economic stability. This reduction in crop losses directly addresses the goal of "No Poverty" by improving the financial wellbeing of farmers and their communities.

Furthermore, by implementing more accurate disease detection models, farmers can maintain a more consistent and reliable food supply. Early intervention helps prevent large-scale crop failures, ensuring that agricultural produce remains available for consumption. This consistency in food availability contributes to

the goal of “Zero Hunger” by ensuring that people have access to an adequate and nutritious diet throughout the year.

Additionally, this model also contributes to the Sustainable Development Goal (SDG) of “Good Health and Wellbeing” by preventing contaminated crops from entering the food supply chain. Detecting diseases early ensures that only safe produce reaches consumers, reducing the risk of foodborne illnesses and related healthcare costs. This model safeguards consumer health, minimizes health risks, and indirectly lowers healthcare expenditures, promoting overall wellbeing and safety.

Economic growth is another area significantly influenced by this research. When crops remain healthy, they yield higher quantities of produce, which farmers can then sell in the market. This increased productivity not only boosts farmers’ incomes but also creates demand for additional labor within the agricultural sector. More workers are needed for various tasks such as planting, harvesting, and maintenance of healthy crops. Additionally, the demand for skilled technicians and researchers may also increase to develop and implement advanced disease prevention measures. As the agricultural sector expands to accommodate these needs, it generates more job opportunities, thereby contributing to the SDG of “Economic Growth.”

In terms of “Industry, Innovation, and Infrastructure,” the development of this model signifies innovation within the agriculture sector. It introduces advanced technology and data-driven solutions that benefit not only farmers but also the broader agricultural industry. Implementing this technology might require the establishment of suitable infrastructure for efficient disease monitoring and control.

By accurately identifying diseases early, farmers can apply treatments only where needed, minimizing pesticide use and promoting environmental sustainability, which promotes “Responsible Consumption.” This also aligns with the SDG of “Climate Action” by reducing the environmental footprint of agriculture and its impact on the environment.

Finally, the successful implementation of this research may necessitate strategic partnerships between researchers, governmental bodies, agricultural organizations, and technology companies. Collaborative efforts are essential to deploy and scale the model effectively, contributing to the SDG of forming “Partnerships” to achieve sustainable goals.

6 Conclusion

Successfully implementing the model in resource-constrained settings is a significant achievement, paving the way for practical disease detection in agriculture. By improving crop health and

reducing costs and energy consumption, this outcome holds the potential to make a substantial impact not only in enhancing food security but also in other sustainable development aspects. However, it’s crucial to acknowledge the limitations and areas for further exploration.

The model may require customization for specific agricultural contexts and broader crop types. Ongoing investigations into its performance under various environmental conditions are warranted. In summary, this research contributes to more efficient disease detection in agriculture, with the need for ongoing refinement and adaptation in diverse settings. It represents a step toward accessible, efficient, and cost-effective disease detection, benefiting the agricultural industry and beyond.

Author contributions

AL: Conceptualization, Formal analysis, Investigation, Methodology, Software, Testing, Validation, Writing—original draft, Writing—review & editing. MS: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Writing—original draft, Supervision, Resources, Writing—review & editing. IA: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Writing—original draft, Supervision, Resources, Writing—review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). “Understanding of a convolutional neural network,” in *2017 International conference on engineering and technology (ICET)* (Antalya: IEEE), 1–6. doi: 10.1109/ICEngTechnol.2017.8308186
- Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosyst. Eng.* 144, 52–60. doi: 10.1016/j.biosystemseng.2016.01.017

- Chen, J., Lian, Y., and Li, Y. (2020). Real-time grain impurity sensing for rice combine harvesters using image processing and decision-tree algorithm. *Comput. Electron. Agric.* 175:105591. doi: 10.1016/j.compag.2020.105591

- Chhillar, A., Thakur, S., and Rana, A. (2020). “Survey of plant disease detection using image classification techniques,” in *2020 8th International Conference on*

- Reliability, *Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (Noida: IEEE), 1339–1344. doi: 10.1109/ICRITO48877.2020.9197933
- Dutta, K., Krishnan, P., Mathew, M., and Jawahar, C. V. (2018). “Improving CNN-RNN hybrid networks for handwriting recognition,” in *2018 16th international conference on frontiers in handwriting recognition (ICFHR)* (Niagara Falls, NY: IEEE), 80–85. doi: 10.1109/ICFHR.2018.2018.00023
- Fang, Y., and Ramasamy, R. P. (2015). Current and prospective methods for plant disease detection. *Biosensors* 5, 537–561. doi: 10.3390/bios5030537
- Hasani, R., Lechner, M., Amini, A., Rus, D., and Grosu, R. (2021). Liquid time-constant networks. *Proc. AAAI Conf. Artif. Intell.* 35, 7657–7666. doi: 10.1609/aaai.v35i9.16936
- Jiang, Y., Wang, S., Valls, V., Ko, B. J., Lee, W. H., Leung, K. K., et al. (2022). Model pruning enables efficient federated learning on edge devices. *IEEE Trans. Neural. Netw. Learn. Syst.* 34, 10374–10386. doi: 10.1109/TNNLS.2022.3166101
- Kondaveeti, H. K., Bandi, D., Mathe, S. E., Vappangi, S., and Subramanian, M. (2022). “A review of image processing applications based on Raspberry-Pi,” in *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Vol. 1* (Coimbatore: IEEE), 22–28. doi: 10.1109/ICACCS54159.2022.9784958
- Mhaski, R. R., Chopade, P. B., and Dale, M. P. (2015). “Determination of ripeness and grading of tomato using image analysis on Raspberry Pi,” in *2015 Communication, Control and Intelligent Systems (CCIS)* (Mathura: IEEE), 214–220. doi: 10.1109/CCIntelS.2015.7437911
- Motwani, A., and Khan, Q. (2022). *Data from: “Tomato Leaves Dataset”, Kaggle*. Available online at: <https://www.kaggle.com/datasets/ashishmotwani/tomato> (accessed April 15, 2024).
- Mustaffa, I. B., and Khairul, S. F. B. M. (2017). “Identification of fruit size and maturity through fruit images using OpenCV-Python and Raspberry Pi,” in *2017 International Conference on Robotics, Automation and Sciences (ICORAS)* (Melaka: IEEE), 1–3. doi: 10.1109/ICORAS.2017.8308068
- Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., and Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Front. Plant Sci.* 8:1852. doi: 10.3389/fpls.2017.01852
- Salehinejad, H., Sankar, S., Barfett, J., Colak, E., and Valaee, S. (2017). Recent advances in recurrent neural networks. *arXiv [preprint]*. arXiv:1801.01078. doi: 10.48550/arXiv.1801.01078
- Shaha, M., and Pawar, M. (2018). “Transfer learning for image classification,” in *2018 second international conference on electronics, communication and aerospace technology (ICECA)* (Coimbatore: IEEE), 656–660. doi: 10.1109/ICECA.2018.8474802
- Torrey, L., and Shavlik, J. (2010). “Transfer learning,” in *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, eds. E. S. Olivas, and M. Martinez Sober (Hershey, PA: IGI global), 242–264. doi: 10.4018/978-1-60566-766-9.ch011
- Tuncer, A. (2021). Cost-optimized hybrid convolutional neural networks for detection of plant leaf diseases. *J. Ambient Intell. Humaniz. Comput.* 12, 8625–8636. doi: 10.1007/s12652-021-03289-4
- Wardana, I. N. K., Gardner, J. W., and Fahmy, S. A. (2021). Optimizing deep learning at the edge for accurate hourly air quality prediction. *Sensors* 21:1064. doi: 10.3390/s21041064