Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 258 (2025) 2870–2877

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

TinyML for Plant Disease Detection: Efficient Edge AI Solutions for Apple and Mango Leaves

Ishana Attri[a], Lalit Kumar Awasthi[a], T.P Sharma[a]

[a]*Computer Science and Engineering, NIT, Hamirpur, 177005, H.P, India.*

Abstract

Detecting plant diseases and providing timely treatment are crucial for ensuring sustainable agricultural productivity. With advances in deep learning, computer vision techniques for crop science have seen rapid growth. This paper introduces a TinyML-based model designed for real-time detection of diseases in apple and mango leaves, utilizing the MobileNetV2 architecture, which is optimized for deployment on edge devices. The model's performance was assessed using key metrics, achieving an accuracy of 93.1%, along with high precision, recall, specificity, and an F1 score of 93.1%. It outperforms other models like ResNet-50, VGGNet-16, and AlexNet in identifying and classifying both healthy and diseased leaves, demonstrating its potential for real-time agricultural use. By optimizing the model for low-power devices, it becomes suitable for use in remote, resource-constrained areas, offering a valuable solution for improving crop management and fostering sustainable farming practices. Future developments will focus on extending the model to detect a wider variety of crops and diseases, integrating sensor data, and conducting field trials to validate its real-world effectiveness.

Keywords: TinyML; Plant Disease Detection; MobileNetV2; Edge AI; Agricultural Technology.

1. Introduction

Plant diseases significantly impact agricultural productivity and food security worldwide. Accurate and timely detection is critical for managing and controlling these diseases. Traditional detection methods, such as expert visual inspections or lab-based techniques, are often slow, labor-intensive, and can sometimes harm the plants. Recently, there has been increasing interest in non-destructive detection methods that provide faster, more reliable, and cost-effective solutions [1,2,3].

Non-destructive techniques leverage various advanced technologies, such as imaging, spectroscopy, and remote sensing, to identify plant diseases without causing any harm to the plants. These methods enable continuous monitoring of plant health and facilitate early intervention, thereby reducing the potential damage and spread of

10.1016/j.procs.2025.04.547

diseases. A comprehensive review by Ali et al. (2019) discusses various non-destructive techniques and their applications in plant disease detection. The review highlights the potential of these methods in revolutionizing plant pathology by providing real-time and accurate disease diagnostics [4].

Conventional disease detection techniques that rely on the manual examination of plant leaves are time-consuming, arbitrary, and prone to errors [5]. Furthermore, using pesticides carelessly as a preventative strategy might have negative effects on the environment and human health. In this regard, cutting-edge technologies like deep learning (DL) and machine learning (ML) present intriguing options for the prompt and precise identification of plant diseases. Convolutional Neural Networks (CNNs), in particular, have been highly effective in automatically extracting features from digital images, enabling faster and more accurate disease identification [6].

However, the deployment of DL and ML models in real-world agricultural settings presents its own challenges. These models typically require substantial computational resources, which can limit their practical application in the field, especially in remote or resource-constrained environments. This is where TinyML, a specialized branch of machine learning, becomes particularly valuable. TinyML focuses on the development and deployment of machine learning models on low-power, resource-constrained devices such as microcontrollers and edge devices. This approach brings the benefits of advanced ML techniques directly to the field, enabling real-time data processing and decision-making without the need for constant connectivity to powerful servers.

This study delves into the application of TinyML—a specialized branch of machine learning that emphasizes the deployment of models on resource-constrained devices—for the purpose of plant disease detection in apple leaves. By leveraging cutting-edge edge AI solutions, our objective is to facilitate real-time and efficient disease detection directly in the agricultural field. This approach not only diminishes the dependency on centralized systems but also significantly enhances the responsiveness and precision of disease management practices. Through this innovative solution, we aim to empower farmers and agricultural stakeholders with immediate, actionable insights, thereby contributing to more sustainable and productive farming practices.

1.1. Our Contribution

In this study, we make several key contributions to the field of plant disease detection using TinyML techniques:

- Applied advanced data augmentation techniques to improve dataset robustness.
- Developed an efficient model leveraging MobileNetV2, optimized for resource-constrained edge devices.
- Converted the trained model into TensorFlow Lite format for low-power device deployment.
- Conducted comprehensive evaluations across key metrics, demonstrating superior performance compared to existing models.

These contributions collectively provide a robust and efficient solution for plant disease detection in apple and mango leaves, demonstrating the potential of TinyML for practical agricultural applications.
The structure of the paper is as follows: Section II provides a review of the literature on plant disease detection using TinyML. Section III outlines the methodology, covering data acquisition and the model architecture. Section IV presents the experimental results, while Section V offers conclusions and suggests future research directions.

2. Literature Review

Traditionally, pathologists' knowledge has been used to identify and diagnose plant diseases, often involving laboratory analysis and eye inspection [7]. However, these methods might be labor-intensive and subjective. Recent advances in image processing and machine learning have opened up new avenues for automatic and more accurate plant disease diagnosis [8]. In the field of image classification, convolutional neural networks (CNNs) have been widely used, especially for the identification of plant diseases from photos of their leaves [9].

TinyML offers a promising solution to key challenges faced by IoT devices, including bandwidth and latency constraints. Conventional IoT systems often rely on sending data to the cloud for processing, especially for computationally heavy tasks. In contrast, TinyML enables on-device processing, eliminating the need for data transmission to external servers and enhancing privacy. Furthermore, TinyML-equipped devices provide stronger security and are more resilient to recognized network threats [10]. As a result, this technology could lead to the creation of new intelligent systems applicable across various industries and applications [11,12].

In [13], the authors propose an Internet of Things (IoT) system for smart agriculture that leverages deep reinforcement learning. The system is structured into four layers: data collection from agricultural sources, edge computing, data transmission, and cloud computing. By integrating advanced technologies like artificial intelligence and cloud computing, the system aims to optimize food production. Notably, deep reinforcement learning is integrated into the cloud layer to make real-time intelligent decisions, like estimating the water needed for irrigation to optimize crop growth conditions. The study in [14] explores how sensor technology and wireless networks can be integrated with IoT to develop a Remote Monitoring System (RMS). This system combines internet connectivity with wireless communication to gather real-time agricultural environmental data. It facilitates swift access to farming infrastructure and sends SMS alerts, providing updates on weather, crops, and other relevant factors. Moreover, in [15], the authors investigate the role of IoT and smart agriculture, incorporating automation to monitor environmental conditions that are critical for optimizing crop production. They develop a system capable of monitoring environmental values and detecting animal movements that could harm crops, using sensors connected to an Arduino board. In case of discrepancies, the system sends notifications to the farmer's smartphone via SMS and a dedicated app, facilitated by a cellular-Internet interface. This energy-efficient and low-cost system shows promise for use in remote areas with limited water resources. The recent publications are given below in Table I.

Table I. Recent publication on Plant disease detection

References	***Models***	***Performance Insights***
Dennis et al., [16]	Advance CNN with TinyML	Enhance the effectiveness and efficiency of crop disease detection system
Jose et al., [17]	Convolution deep learning model for anomaly detection in agricultural images	Extensive experiments comparing 4 architectures and 2 datasets.
Hanxiang wang et al. [18]	Transformer based plant disease detection system	Outperform all the state of art object detection models
Sasikala et al., [19]	Hierarchical framework for plant disease detection	Notably high accuracy then InceptionV3, Resnet50 models.
Sandesh et al., [20]	Light weight deep learning approach (LITE-MDC)	Exhibits the potential for real time plant disease detection.

3. Methods and Materials

This section includes a detailed account of the dataset used and the analytical techniques applied, providing a comprehensive understanding of our research methodology. The following subsections will elaborate on these aspects to ensure the reproducibility and validation of our study.

3.1. Data Acquisition

For our research, we utilized several datasets containing images of mango and apple leaves. These datasets, sourced from Kaggle and other platforms, include leaves of various species, ages, and health conditions, providing a representative sample for real-world applications.

- **Apple Dataset:** This dataset, sourced from Kaggle, includes 7,771 images primarily depicting various diseases affecting apple leaves. The high-resolution images capture fine details such as textures and spots, which are essential for analyzing the different diseases that can impact apple trees.
- **Mango Dataset:** This dataset, also from Kaggle, focuses on the classification of diseases affecting mango leaves. It contains 412 high-resolution images that showcase the textures and colors associated with various diseases. The clarity of these images ensures that the intricate details necessary for accurate disease identification and classification are preserved.

3.2. Proposed Model

We adapted the MobileNetV2 model by using transfer learning, which allows the model to learn from pre-existing knowledge. We reduced the size of the data through a pooling layer, followed by a layer that helps categorize the leaves into healthy or diseased. The proposed model for plant disease detection in apple and mango leaves is based on the MobileNetV2 architecture, which is optimized for deployment on resource-constrained devices, making it suitable for TinyML applications. MobileNetV2 strikes an effective balance between accuracy and computational efficiency, making it ideal for real-time inference in agricultural settings.

- **Experimental Settings:** The experiments were performed on an MSI system equipped with a 64-bit Windows 11 Home Single Language OS and 16.0 GB of RAM.
- **Model Architecture Details:** The model architecture leverages the MobileNetV2 base as a feature extractor, pre-trained on the ImageNet dataset. To enhance the model's adaptability to the specific task of plant disease detection, the base model is fine-tuned by unfreezing its layers during training. A Global Average Pooling layer is utilized to reduce the spatial dimensions of the features extracted. This is followed by a fully connected layer comprising 128 units and activated using the ReLU function. To mitigate overfitting, a Dropout layer with a dropout rate of 50% is incorporated. Finally, a softmax layer is employed to classify the input images into one of four categories: apple scab, black rot, cedar apple rust, or healthy leaves.
- **Data Augmentation:** To enhance the diversity of the dataset, we applied data augmentation techniques, including random rotations, zooms, shifts, and flips. This approach improves the model's robustness, particularly for the smaller mango dataset. These techniques help simulate real-world variations and ensure that the model can generalize effectively, despite the imbalance between the apple and mango datasets.
- **Training Strategy:** The model is trained using the categorical cross-entropy loss function and optimized with the Adam optimizer, with a learning rate of 0.0001. The training process spans 30 epochs. To enhance the model's generalization, several data augmentation techniques are employed, including rotation, width and height shifts, shear transformation, zooming, and horizontal flipping. These augmentations help the model become more robust to variations in the input images, simulating real-world conditions.
- **Dropout and Overfitting:** To mitigate overfitting, we employed a dropout rate of 50%, combined with early stopping and cross-validation during the training process. These methods ensure that the model can generalize well to new, unseen data, preventing overfitting that may arise from the smaller size of the mango dataset.
- **Model Conversion for TinyML Deployment:** To enable deployment on edge devices, the trained model is converted into TensorFlow Lite format. TensorFlow Lite optimizes the model for execution on low-power, resource-constrained devices, making it suitable for real-time plant disease detection in agricultural fields. This conversion significantly reduces the model's size and computational requirements while maintaining its accuracy.

Fig. 1 presented the architecture of proposed model and gives the layers used in the architecture.

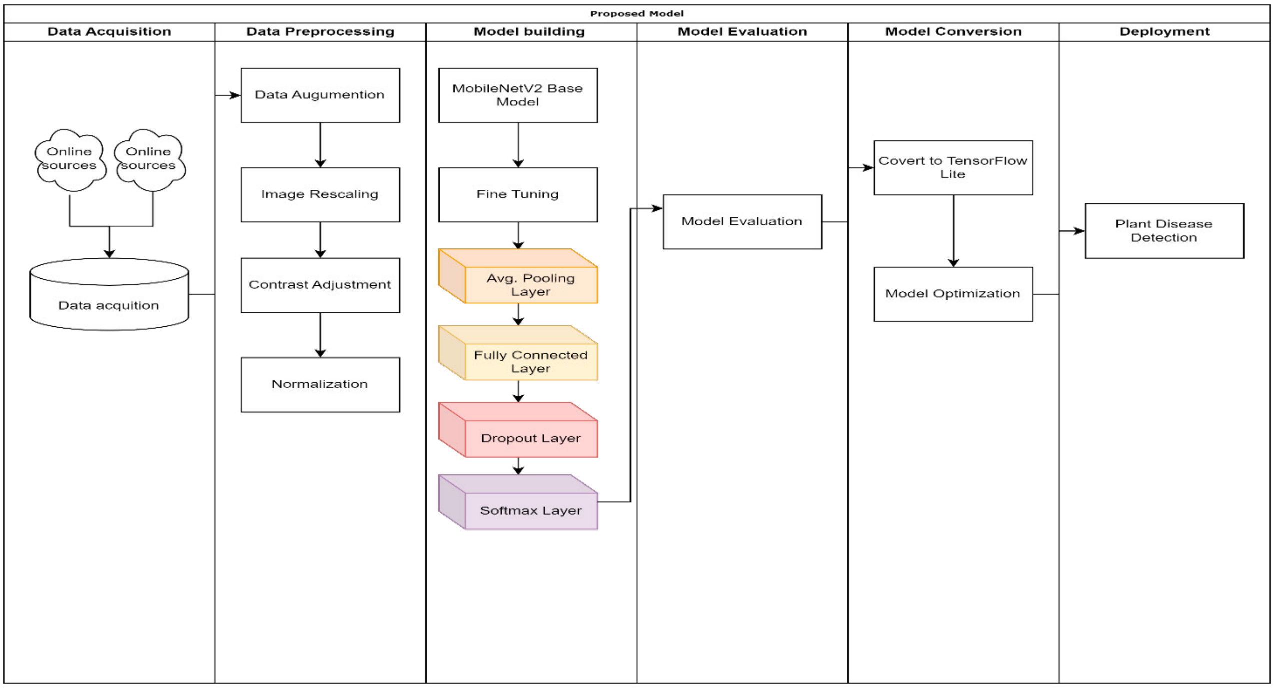

Fig. 1 Architecture of Proposed Model

3.3. Performance Evaluation

The model's performance is assessed on a validation dataset through metrics including accuracy, precision, recall, and F1-score. A confusion matrix and classification report provide detailed insights into the model's effectiveness across the different disease categories. The results indicate that the model achieves high accuracy and robust performance metrics, validating its suitability for practical agricultural applications. The proposed model combines the efficiency of MobileNetV2 with the practical deployment capabilities of TensorFlow Lite, offering a robust solution for real-time plant disease detection in apple and mango leaves. This approach empowers farmers with immediate, actionable insights, supporting timely and informed decision-making in agriculture.

4. Experimental results and Discussion

The proposed model's effectiveness is assessed using the datasets covered in Section 3.1.

4.1. Experimental Settings

Python was used to create the experimental configuration for the proposed model. An MSI device running 64-bit Windows 11 Home Single Language with an 11th Gen Intel(R) Core (TM) i7-11800H CPU running at 2.30GHz and 16.0 GB of installed RAM served as the machine for the experiment.

4.2. Results and Discussions

The suggested model for identifying plant diseases in the leaves of apples and mangoes was tested against a number of well-known models, such as Xception Net, DenseNet-121, ResNet-50, VGGNet 16, AlexNet, and Efficient Net B2. Specificity, F1 Score, Accuracy, Precision, and Recall/Sensitivity were the performance criteria taken into account. The outcomes for mango and apple leaves are shown in two different tables i.e. Table II. And Table III.

Table II. Assessment of the Proposed Model's Performance Compared to Leading Models on Apple Leaves

Model	Accuracy	Precision	Recall/ Sensitivity	Specificity	F1 Score
Xception Net	89.8	89.6	89.7	89.9	89.7
DenseNet-121	87.2	87	87.1	87.3	87.1
ResNet-50	90.4	90.2	90.3	90.5	90.3
VGGNet 16	86.1	86.8	86	86.2	86.9
AlexNet	90.6	90.4	90.5	90.7	90.5
Efficient Net B2	89.8	89.6	89.7	89.9	89.7
Proposed Model	**92.7**	**92.3**	**92.8**	**95**	**92.5**

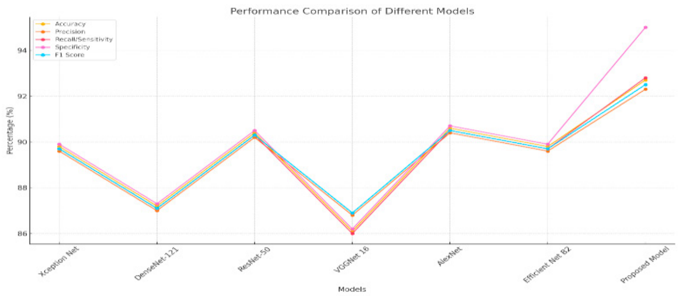

Fig. 2 Comparative performance of Models on Apple leaves.

Table III. Assessment of the Proposed Model's Performance Compared to Leading Models on Mango Leaves

Model	Accuracy	Precision	Recall/ Sensitivity	Specificity	F1 Score
Xception Net	90.5	90.2	90.4	90.5	89.4
DenseNet-121	89.6	89.8	89.9	89.9	89.9
ResNet-50	91	90	90.1	90.1	90.1
VGGNet 16	88.2	88.6	88.8	88.8	88.7
AlexNet	91.9	90.3	90.3	90.3	90.3
Efficient Net B2	90.4	89.5	89.5	89.5	89.5
Proposed Model	**93.1**	**92.9**	**93.3**	**93.5**	**93.1**

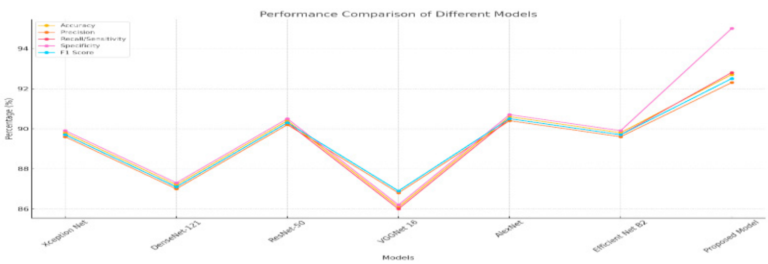

Fig. 3 Comparative performance of Models on Mango leaves.

The evaluation of various models for plant disease detection in apple and mango leaves revealed that the proposed model outperforms the others across all key metrics, including accuracy, precision, recall, specificity, and F1 score. With an accuracy of 93.1% and a high F1 score of 93.1%, the proposed model demonstrates a well-rounded performance, effectively balancing the detection of true positives and minimizing false positives. This is crucial in

agricultural applications where accurate disease identification is vital to prevent the spread of disease and unnecessary treatment of healthy plants.
Compared to other models like ResNet-50, VGGNet 16, and AlexNet, the proposed model's superior recall (93.3%) and specificity (93.5%) underscore its reliability in real-world scenarios. The model’s architecture, which leverages fine-tuning of MobileNetV2 combined with TensorFlow Lite optimization, ensures that it is not only accurate but also efficient for deployment on edge devices. This makes the model particularly suitable for real-time monitoring in remote and resource-constrained agricultural environments. The high specificity of the proposed model further ensures that healthy crops are not misclassified as diseased, supporting more sustainable farming practices by reducing unnecessary pesticide use.
A comparative analysis of the proposed model with other state-of-the-art models, such as ResNet-50, VGGNet-16, and AlexNet, is provided in Tables 2 and 3. The results indicate that our model outperforms these models in terms of accuracy, F1 score, and specificity, particularly for the mango dataset, where data scarcity posed a challenge for the other models.
The proposed TinyML-based model for plant disease detection demonstrates a significant advancement over existing models, with superior performance across all evaluated metrics. Its robustness, accuracy, and efficiency make it an excellent candidate for real-world applications, providing a reliable tool for farmers to manage and protect their crops effectively. The success of this model also underscores the potential of TinyML in enhancing agricultural productivity and sustainability through advanced, yet accessible, technological solutions.

5. Conclusion and Discussion

In this study, we developed and evaluated a TinyML-based model for plant disease detection in apple and mango leaves, leveraging the MobileNetV2 architecture optimized for deployment on edge devices. The proposed model demonstrated superior performance across all key metrics, including accuracy, precision, recall, specificity, and F1 score, compared to existing models like ResNet-50, VGGNet 16, and AlexNet. With an accuracy of 93.1% and a balanced F1 score of 93.1%, the model proved to be both effective and reliable for real-time disease detection in agricultural settings.

In conclusion, our TinyML-based model for plant disease detection demonstrates significant improvements in accuracy, precision, and computational efficiency, making it ideal for deployment on low-power edge devices. The model’s ability to perform real-time disease detection in apple and mango leaves shows promise for agricultural applications. Future research will focus on expanding the dataset to include a wider variety of crops and integrating real-world environmental data. Further optimization of the model for ultra-low-power devices, along with the integration of explainability features, will enhance its applicability in practical settings.
Future work could also explore the use of advanced TinyML techniques to reduce the model's computational requirements even further, making it more efficient for deployment on ultra-low-power devices. Moreover, incorporating explainable AI (XAI) methods could help users understand the model's decision-making process, fostering greater trust and adoption among farmers. Lastly, field trials and real-world testing would be essential to validate the model's performance in diverse agricultural environments, ensuring its robustness and effectiveness in practical applications. While the proposed TinyML-based model demonstrates strong potential for real-time plant disease detection, further research will involve expanding the dataset and testing the model in real-world agricultural environments. This will help evaluate its performance under diverse conditions, such as varying lighting and weather scenarios. Future work will also focus on optimizing the model for ultra-low-power devices through quantization techniques, ensuring its efficiency in resource-constrained settings. Additionally, integrating explainability features, such as visual heatmaps, will help users better understand the model's predictions, making it more transparent and interpretable for farmers.

References

[1] M. John, I. Bankole, O. Ajayi-Moses, T. Ijila, T. Jeje, and L. Patil, "Relevance of advanced plant disease detection techniques in disease and pest management for ensuring food security and their implication: A review," American Journal of Plant Sciences, vol. 14, 2023, doi: 10.4236/ajps.2023.1411086.

[2] I. Attri, L. Awasthi, and T. Sharma, "Machine learning in agriculture: A review of crop management applications," Multimedia Tools and Applications, vol. 83, pp. 1–41, 2023, doi: 10.1007/s11042-023-16105-2.

[3] I. Attri, L. K. Awasthi, T. P. Sharma, and P. Rathee, "A review of deep learning techniques used in agriculture," Ecological Informatics, vol. 77, 2023, Art. no. 102217, doi: 10.1016/j.ecoinf.2023.102217.

[4] M. M. Ali, N. A. Bachik, N. ‘A. Muhadi, T. N. T. Yusof, and C. Gomes, "Non-destructive techniques of detecting plant diseases: A review," Physiological and Molecular Plant Pathology, vol. 108, pp. 101426, 2019, doi: 10.1016/j.pmpp.2019.101426.

[5] M. E. Chowdhury, T. Rahman, A. Khandakar, M. A. Ayari, A. U. Khan, M. S. Khan, and S. H. M. Ali, "Automatic and reliable leaf disease detection using deep learning techniques," AgriEngineering, vol. 3, no. 2, pp. 294–312, 2021.

[6] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," Computers and Electronics in Agriculture, vol. 173, pp. 1–20, 2020.

[7] M. L. Gleason, G. A. Duttweiler, and L. F. Batzer, "Automated image analysis for detecting and diagnosing plant diseases," Plant Disease, vol. 93, no. 6, pp. 583–593, 2009.

[8] S. Aravind, C. Raja, and S. Smitha, "A survey on plant leaf disease detection and classification," Int. J. Comput. Appl., vol. 120, no. 15, pp. 26–32, 2015.

[9] A. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," Comput. Intell. Neurosci., vol. 2016, Art. no. 3289801, 2016, doi: 10.1155/2016/3289801.

[10] E. Boumpa, V. Tsoukas, A. Gkogkidis, G. Spathoulas, and A. Kakarountas, "Security and privacy concerns for healthcare wearable devices and emerging alternative approaches," in Wireless Mobile Communication and Healthcare (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering), X. Gao, A. Jamalipour, and L. Guo, Eds. Cham, Switzerland: Springer International Publishing, 2022, pp. 19–38, doi: 10.1007/978-3-031-06368-8_2.

[11] A. Gkogkidis, V. Tsoukas, S. Papafotikas, E. Boumpa, and A. Kakarountas, "A TinyML-based system for gas leakage detection," in 2022 11th International Conference on Modern Circuits and Systems Technologies (MOCAST), 2022, pp. 1–5, doi: 10.1109/MOCAST54814.2022.9837510.

[12] V. Tsoukas, A. Gkogkidis, and A. Kakarountas, "Elements of TinyML on constrained resource hardware," in Advances in Computing and Data Sciences, M. Singh, V. Tyagi, P. K. Gupta, J. Flusser, and T. Ören, Eds. Cham, Switzerland: Springer International Publishing, 2022, pp. 316–331.

[13] F. Bu and X. Wang, "A smart agriculture IoT system based on deep reinforcement learning," Future Generation Computer Systems, vol. 99, pp. 500–507, Oct. 2019, doi: 10.1016/j.future.2019.04.041.

[14] K. A. Patil and N. R. Kale, "A model for smart agriculture using IoT," in 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), 2016, pp. 543–545, doi: 10.1109/ICGTSPICC.2016.7955360.

[15] G. Sushanth and S. Sujatha, "IoT based smart agriculture system," in 2018 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2018, pp. 1–4.

[16] D. A. N. Gookyi, F. A. Wulnye, E. A. E. Arthur, R. K. Ahiadormey, J. O. Agyemang, K. O.-B. O. Agyekum, and R. Gyaang, "TinyML for smart agriculture: Comparative analysis of TinyML platforms and practical deployment for maize leaf disease identification," Smart Agricultural Technology, vol. 8, Art. no. 100490, 2024, doi: 10.1016/j.atech.2024.100490.

[17] A. Mendoza-Bernal, "A convolutional neural network approach for image-based anomaly detection in smart agriculture," Expert Systems with Applications, vol. 247, Art. no. 123210, 2024, doi: 10.1016/j.eswa.2024.123210.

[18] H. Wang, T.-H. Nguyen, T. N. Nguyen, and M. Dang, "PD-TR: End-to-end plant diseases detection using a transformer," Computers and Electronics in Agriculture, vol. 224, Art. no. 109123, 2024, doi: 10.1016/j.compag.2024.109123.

[19] S. Vallabhajosyula, V. Sistla, and V. K. K. Kolli, "A novel hierarchical framework for plant leaf disease detection using residual vision transformer," Heliyon, vol. 10, no. 9, Art. no. e29912, 2024, doi: 10.1016/j.heliyon.2024.e29912.

[20] S. Bhagat, M. Kokare, V. Haswani, P. Hambarde, T. Taori, P. H. Ghante, and D. K. Patil, "Advancing real-time plant disease detection: A lightweight deep learning approach and novel dataset for pigeon pea crop," Smart Agricultural Technology, vol. 7, Art. no. 100408, 2024, doi: 10.1016/j.atech.2024.100408.