AgriLoRa: A Digital Twin Framework for Smart Agriculture

Pelin Angin[1*], Mohammad Hossein Anisi[2], Furkan Göksel[1], Ceren Gürsoy[1], and Asaf Büyükgülcü[1]
[1]Middle East Technical University, Ankara, 06800 Turkey
pangin@ceng.metu.edu.tr, {furkan.goksel, ceren.gursoy, asaf.buyukgulcu}@metu.edu.tr
[2]University of Essex, Colchester, Essex, CO4 3SQ UK
m.anisi@essex.ac.uk

Received: October 10, 2020; Accepted: December 14, 2020; Published: December 31, 2020

Abstract

Throughout history, farmers and agricultural engineers have focused on the issue of increasing the yield of crops using different farming methods. In today's digitalized world, these techniques have been combined with IoT technology and machine learning algorithms, which have given rise to smart agriculture systems. However, farmers who live in developing countries hesitate to use such systems because of their hardware and maintenance costs. To address this issue, this paper proposes a low-cost farmland digital twin framework called AgriLoRa for smart agriculture. AgriLoRa consists of a wireless sensor network established in the farmland and cloud servers that run computer vision algorithms to detect plant diseases, weed clusters and plant nutrient deficiencies. In order to assess the feasibility of accurate plant disease detection, we have performed initial experiments with agricultural vision datasets using two different algorithms, the MobileNet and UNet models, and achieved successful results. AgriLoRa is promising to achieve a low-cost, high-precision smart agriculture solution to address the growing high-yield production needs of farmers worldwide.

Keywords: smart agriculture, digital twins, wireless sensor networks

1 Introduction

Smart agriculture producing high yields with optimal use of resources has become a must in the whole world due to the increasing population and scarcity and cost of resources such as water, fuel and fertilizers. According to recent studies, the fast-growing world population, which is expected to reach 9.7 billion by 2050, will create the need to increase crop production by 25-70% to meet the global food demands [15]. Use of advanced technologies in agriculture is especially needed in countries where agriculture is a major contributor to the economy and the population is increasing at a steady rate, resulting in the disturbance of the balance between food supplies and demand.

Community welfare and resilience largely depends on agriculture. The world economy has been devastated by COVID-19, which restricted food imports/exports and led to an unexpected shortage of human labor. Occurrences of infectious disease may continue to increase, according to prominent studies, through changing climate [10] and the resulting loss of biodiversity [18]. This has shown the importance of automation and high productivity in agriculture to build resilient communities.

The rise of Internet of Things (IoT) in the past decade has allowed physical devices of various types to be accessible via the Internet, making it possible to create automated monitoring and decision-making systems to reduce resource consumption and production costs in agriculture. While many countries have

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 11(4):77-96, Dec. 2020
DOI:*10.22667/JOWUA.2020.12.31.077*

*Corresponding author: Department of Computer Engineering, Middle East Technical University, Universiteler Mahallesi Dumlupinar Bulvari No 1, Cankaya, Ankara, 06800, Turkey, Tel: +90-312-210-5532.

started using smart agriculture applications relying on IoT and artificial intelligence, most developing countries have not been able to adopt these systems at large due to barriers such as high cost and high power consumption, which may not offer a fair trade-off with the yield increase to be achieved. Considering these issues, in this work we propose a high precision, low-cost IoT-based smart agriculture framework. The proposed framework aims to create farmland *digital twins*, which allow for watching the status of farmlands in near real-time through their digital replicas and taking appropriate actions recommended by intelligent processing of the gathered data. The overarching goal of the framework is to create an extensible system architecture for sustainable precision agriculture, which provides early detection of crop diseases and nutrient deficiencies and improves irrigation and fertilization strategies to achieve the highest yield and lowest production costs in a highly automated manner. The specific objectives that will contribute to the achievement of this goal are as follows:

- Creation of an extensible digital twins framework for IoT-based precision sustainable agriculture, which achieves near real-time cloud-based fusion and processing of multi-sensor data obtained using wireless sensor networks (WSNs) and multispectral imaging to detect crop diseases and nutrient deficiencies.
- Development of a low-overhead WSN communication model, utilized in monitoring field conditions, which does not require cellular connectivity in the field for long ranges.
- Creation of an optimal placement and activation strategy WSN nodes in a field, which is adaptable to different contexts for reliable and efficient operation, achieving a long battery life for the sensor nodes.

The proposed framework is built upon two main components. For the data gathering and communication component, an optimized LoRa based Wireless Sensor Network (WSN) is established. LoRa (Long Range) is one of the many Low Power Wide Area Network (LPWAN) technologies. It offers low data rate communication over long distances with very low cost and low power consumption. Because of these features, it is suitable for use in agricultural areas. End nodes of the WSN gather environmental data from the farmland and send these data to cloud servers through a gateway using multi-hop routing. The agricultural intelligence component relies on utilizing machine learning algorithms for tasks like disease detection and nutrient deficiency detection on plant leaves, as well as the correlation of the results obtained with the data gathered using the WSN to provide a complete view of the farm field.

The rest of the paper is structured as follows: Section 2 provides an overview of related work in smart agriculture. Section 3 describes our proposed digital twins-based smart agriculture framework. Section 4 provides experimental evaluation of the plant disease detection algorithms on public agricultural vision datasets as well as the performance of the simulated WSN. Section 5 concludes the paper with future work directions.

2 Related Work

The rise of the IoT paradigm in the past decade has created immense opportunities for precision agriculture worldwide through providing means for remote monitoring of fields as well as automation of important processes such as irrigation and planting. Significant research efforts have been put into the design and development of WSN-based systems, weather warning systems, as well as advanced image capture and processing systems, and many companies providing precision agriculture services to farmers were established.

Kang et al. [16] proposed a system that gathers data from a greenhouse and allows farmers to check this information remotely via the Web. This system consists of wireless sensor nodes utilizing the Zigbee protocol for communication, and an actuator node database. Each sensor node measures temperature, humidity, leaf temperature, and leaf wetness of the greenhouse, and then transfers these data to the sink node. The received sensor data are stored in a database and the sink node transfers the behavior signal to the actuator node.

In [7], important soil parameters such as moisture, nutrient levels, and pH values are measured using a smart sensing system. The measured data from the smart farm sensing system are sent to the smart irrigator via a GSM module. Based on these measured values, the required quantity of green manure, compost, and water is applied on the crops using the smart irrigator.

In [8] a low-cost, modular and Long-Range Wide-Area Network (LoRaWAN)-based IoT platform called LoRaFarM was proposed with the purpose of improving the management of generic farms in a highly customizable way. The authors of the paper experimented with the compatibility between different wireless technologies and LoRa technology.

With the developments in IoT, smart irrigation systems have become popular. In [27] irrigation nodes mainly composed of the LoRa communication modules, solenoid valves and hydroelectric generators were proposed. In the proposed system, the irrigation node sends data to the cloud through LoRa gateways. The system can be controlled remotely by mobile applications.

Image processing is used in many areas of agriculture such as disease detection from plant leaves, detection of arid regions on the land with aerial imagery, as well as separation of weeds and plants in the field. [25] provides a comprehensive survey of image processing techniques for smart agriculture, including weed detection, fruit grading and disease detection among other applications. Disease detection on plants can be achieved with various machine learning algorithms. In [24], convolutional neural networks (CNN) were used for the classification of diseased portions of cotton plant images. Mainly two CNN architectures, AlexNet and GoogleNet were used, achieving successful results.

For image transmission with LoRa technology, [6] proposed a new reliable delivery protocol named Multi-Packet LoRa (MPLR). In order to achieve reliable transmission, they proposed an acknowledgement mechanism for the delivery of image packets. Moreover, instead of sending data packets one by one, an acknowledgment is expected after the number of packets sent reaches the predetermined window size.

While recent work on precision agriculture has mostly focused on the use of IoT and imaging processing technologies, an end-to-end framework that employs multi-sensor data fusion from the farmland WSNs and precisely matches the data with the digital representation of the field is still lacking.

3 Proposed Approach

In this paper, we propose an extensible, high-precision IoT-based farm field digital twin framework, which enables near real-time monitoring of the field conditions, intelligent cloud-based processing of the rich image and sensor data gathered from the field and recommendation of field treatment to address the detected problems at optimal cost. The proposed framework differs from existing IoT-based precision agriculture systems in that it aims to build a context-adaptive, shareable and secure digital twins-based model under computation and communication cost constraints, which is significant for adoption by farmers having limited resources and countries in need of cost-efficient farming technologies.

At a high level, a digital twin can be defined as a virtual replica of objects or systems from the physical world [12]. A digital twin continuously receives data from its physical counterpart to provide an up-to-date virtual model and the virtual model can also provide feedback to the physical world through the same communication channel. In the case of the farmland digital twin framework, the objects modeled consist of the plants in the farm field, the soil sensor nodes and the gateways responsible for communicating the data gathered from the field to the cloud computing platform for processing. The digital twin will provide an accurate, up-to-date representation of the status of the farm field both visually with drone-captured imagery, and in terms of the soil parameters including pH, salinity, nitrogen, phosphorus and potassium levels, temperature and humidity by periodically gathering data from the sensor nodes.

A significant advantage of the digital twin framework is providing remote access to an up-to-date view of the field over the Web, greatly facilitating the job of the farmers by obviating the need for manual monitoring processes, especially in large farm fields. A high-level architecture of the proposed digital twin framework is provided in Figure 1.

The proposed digital twin model works as follows: To overcome high costs of the deployment of high-capacity cellular connectivity capable sensors in the field to monitor soil parameters, WSNs consisting of commercial-off-the-shelf sensors with low power requirements based on the LoraWAN technology are created. The WSN Lora nodes stream data to a cloud platform through a common LoraWAN gateway device, capable of covering a large area that may contain multiple fields in some cases. The gathered data are securely stored in the cloud, correlated with drone-based imagery for the same location through GPS-based location mapping. The multi-sensor data are periodically processed using computer vision and deep learning algorithms with high predictive power to detect crop diseases and nutrient deficiencies.

The system should be capable of utilizing data from multiple fields to provide more accurate estimates for the field state, however the privacy of the data for each field should be strictly preserved by implementing data isolation techniques for separate users in the system on the client side. Continuous collection of data from the field will form a growing dataset on which supervised deep learning models are trained to predict the current levels of nutrients, as well as make estimates about how the induction of varying levels of fertilizers will affect the situation based on historical data available, including the aforementioned soil parameter values. The results are compiled and sent to the client interface of the digital twin platform periodically, where farmers are able to see the status of the field, results of nutrient deficiency and plant disease detection algorithms and recommendations for fertilizer application.

In the subsections below, we describe two main components of the system: The LoraWAN-based WSN and image processing module for plant disease/nutrient deficiency detection.

3.1 LoRaWAN-based Wireless Sensor Network

In the proposed smart agriculture framework, wireless sensor nodes in the farm field gather important soil parameters periodically and send these data packets to a gateway. The role of the gateway is collecting incoming packets from sensor nodes and forwarding these packets to the cloud servers. Machine learning algorithms in the cloud servers process these data packets to detect anomalies or any unfavorable conditions that could affect crop yield.

As discussed before, one of the primary considerations of the framework is cost minimization. Specifically, the framework needs to consist of low-cost sensor nodes that consume minimal energy for their functioning. For these two purposes, we propose the use of LoRa [21] modules for communication in the proposed WSN. Since LoRa modules are low-cost and they can be used for constructing a Low-Power Wide-Area Network, they are the current best option for the construction of a low-cost, high-precision smart agriculture system.

Technically, LoRa is a radio modulation scheme — a way of manipulating a radio wave to encode information using a chirped, multi-symbol format [19]. LoRa technologies operate in licensed free bands in different regions. For instance, the 863 to 870 MHz frequency band is reserved for many European countries, while it is the 902 to 928 MHz band in the United States[1]. LoRa technology only covers the physical layer of the OSI model.

For the data-link and network layers of the OSI model, LoRa Alliance [3] offers the LoRaWAN protocol [4]. LoRaWAN is a media access control (MAC) protocol that allows IoT devices to communicate with Internet-connected applications over long range wireless connections. The LoRa protocol stack is

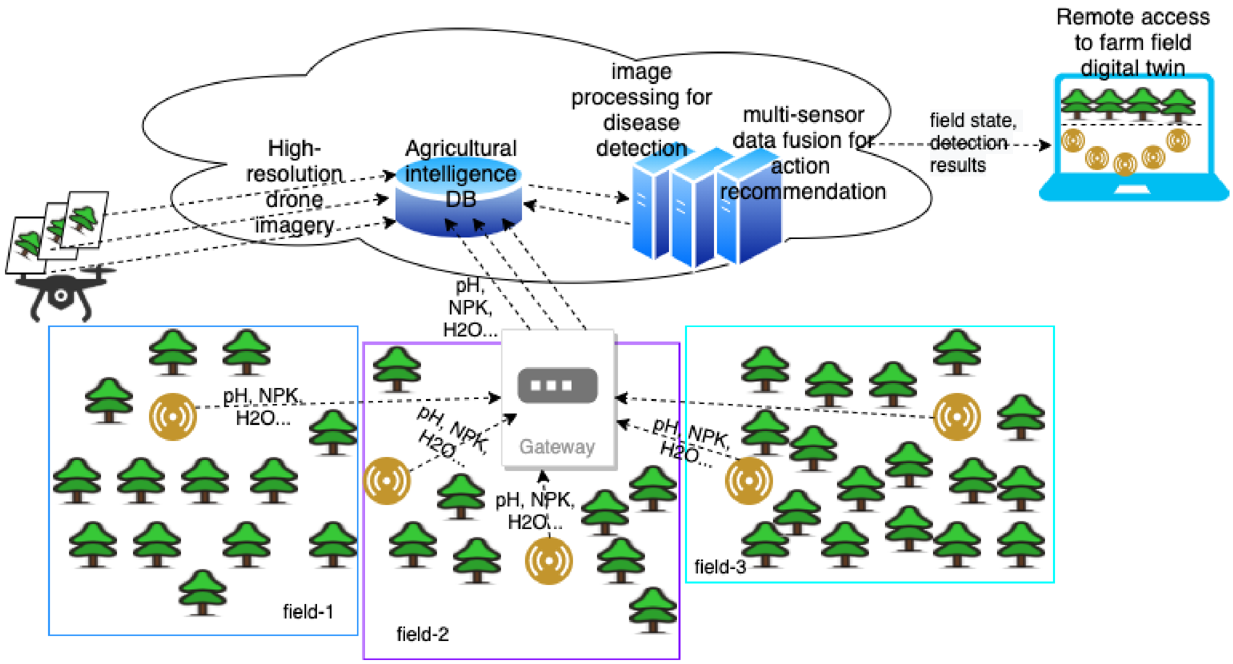


Figure 1: Farm field digital twin framework

shown in Figure 2.

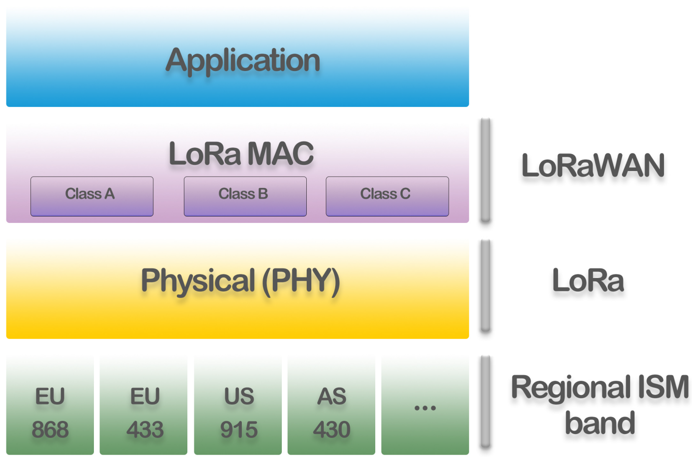


Figure 2: LoRa protocol stack

The LoRaWAN specification defines three types of devices: Class A, Class B, and Class C. The difference between these classes is their packet receiving period. Class A devices can receive packets in two short downlink windows. Class B devices can open extra receiving windows at scheduled times. Class C devices can receive packets continuously. In other words, they are listening for packets all the time. Among these, the most energy-efficient one is Class A devices and in the proposed WSN, sensor nodes are these types of devices.

The LoRaWAN protocol proposes a star topology for the network model. According to this model, end devices in the network cannot communicate with each other. The data packets that are generated in sensor nodes are sent to the gateway directly. Since a LoRa gateway's range is up to 2.6 km in rural areas [5], a whole farmland can be monitored using one gateway in many cases and even the furthest node can send its measurements without using multi-hop transmission.

In LoRaWAN networks, all gateways need to be connected to the Internet. Therefore, if a gateway is not connected to the Internet, even if it gets a packet from any of the sensor nodes, it cannot forward them to the network server. This is a serious problem for rural areas that have no Internet infrastructure. In order to overcome this issue, we propose a modified version of the AODV routing protocol between gateways. With this protocol, gateways that are not connected to the Internet send their packets by "hopping" over other gateways to one that has the Internet connection.

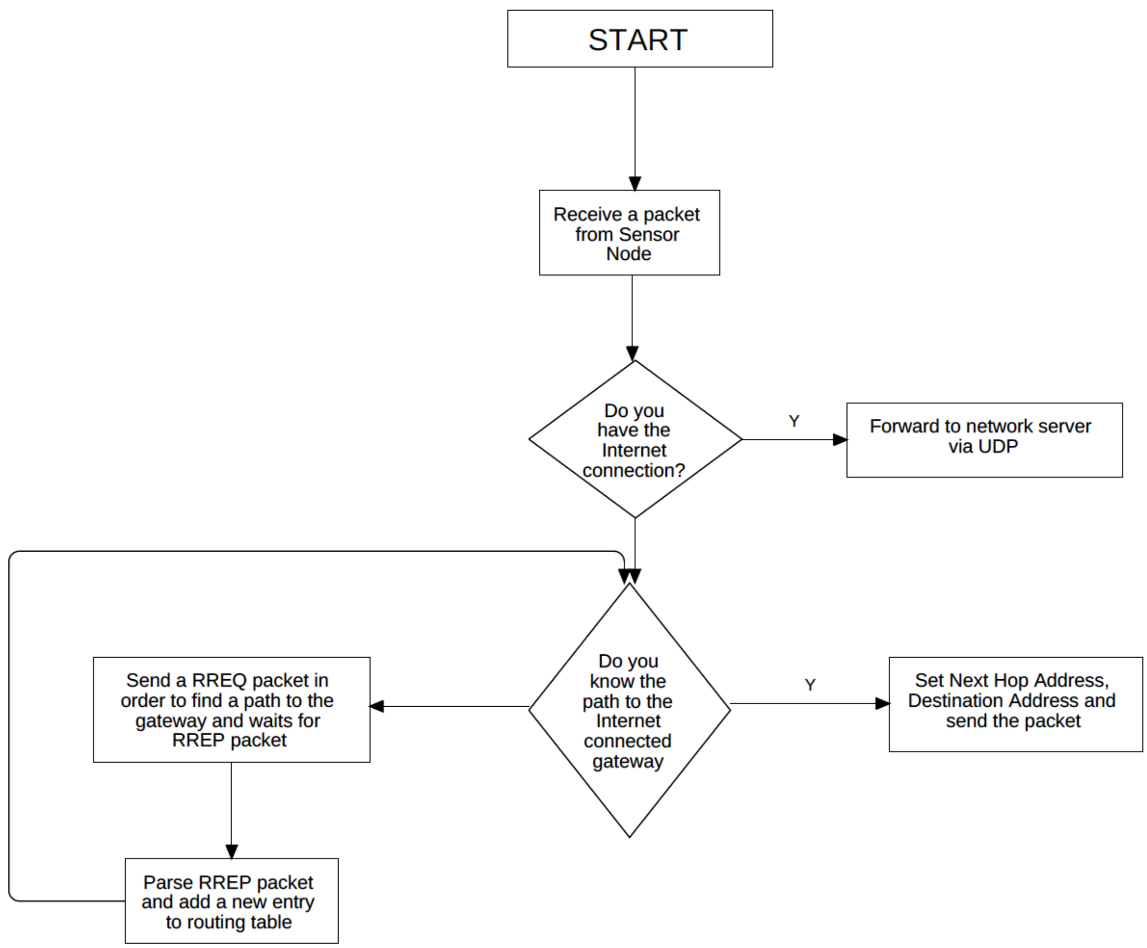


Figure 3: LoRaWAN Gateway Communication Flowchart

In Figure 3, the proposed communication protocol between LoRa gateways is summarized. Whenever a gateway receives a packet from sensor nodes using the LoRaWAN protocol, it checks whether it is connected to the Internet or not. If it is connected, it simply forwards the coming packet via UDP, otherwise, it checks for the route to the predefined Internet-connected node. If it knows the route, it simply forwards the packet to the next node. If the route is unknown, then an RREQ packet is prepared for route finding.

As mentioned above, nodes that do not have Internet connectivity need to have an entry in their routing table for a node that has Internet connectivity. The determination of the route takes place via RREQ and RREP packages.

Whenever a node receives a RREQ Packet, it first checks whether the destination node is itself or not. If it is, it parses the coming RREQ packet to add a new entry into its routing table. Then it prepares the RREP packet as the response to this packet. If it is an intermediate node rather than a destination node, it broadcasts the packet if this packet is received the first time. If the packet has not reached this node for the first time, it is simply discarded to avoid a broadcast storm. Each time an intermediate node or the destination node receives a packet, it parses the destination address, previous node address, and source address from the packet and adds a new entry to its routing table. This process is summarized in the flowchart in Figure 4.

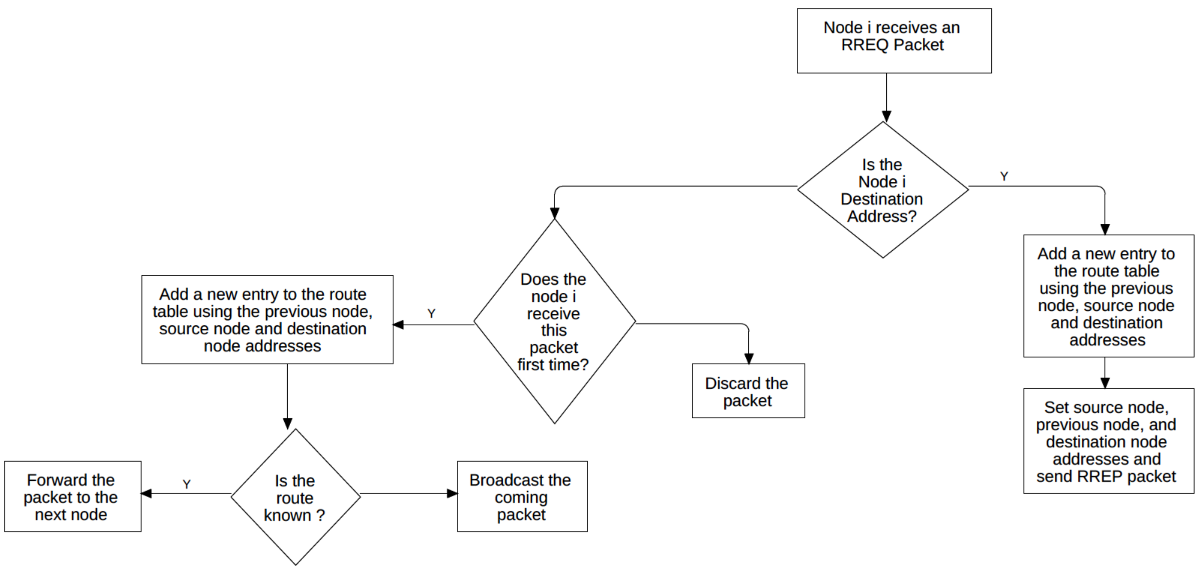


Figure 4: Flowchart of Receiving a RREQ Packet

RREQ packets determine the route to the destination node, and the destination node adds this path to its routing table, but the source node is not aware of this path yet. Therefore, RREP packets are sent so that the source node is also informed of this route.

RREP packets are used to notify the source node of the specified route. It is sent by the destination node, which is the Internet-connected node in our case, to the source node. Since the route has already been determined, RREP packets are sent directly to the specific intermediate node instead of broadcasting. With the help of RREQ and RREP packets, intermediate nodes can be aware of their neighbor nodes and record the path to any node in their routing table. In the case of multiple paths to the destination node, the path with the minimum hop count is chosen. Figure 5 summarizes this process.

3.2 Disease Detection

Although the architecture of AgriLoRa has been designed to allow a variety of agricultural intelligence tasks, in this paper we focus on two of those tasks: disease and weed detection from plant imagery. AgriLoRa utilizes deep learning algorithms running on cloud servers for detection of diseases in plants from drone-based imagery. Accurate disease detection is achieved by building an evolving agricultural intelligence database with images of healthy and diseased plants, which comprise the training set of the deep learning algorithms to be used for real-time detection of plant diseases. Below we provide details

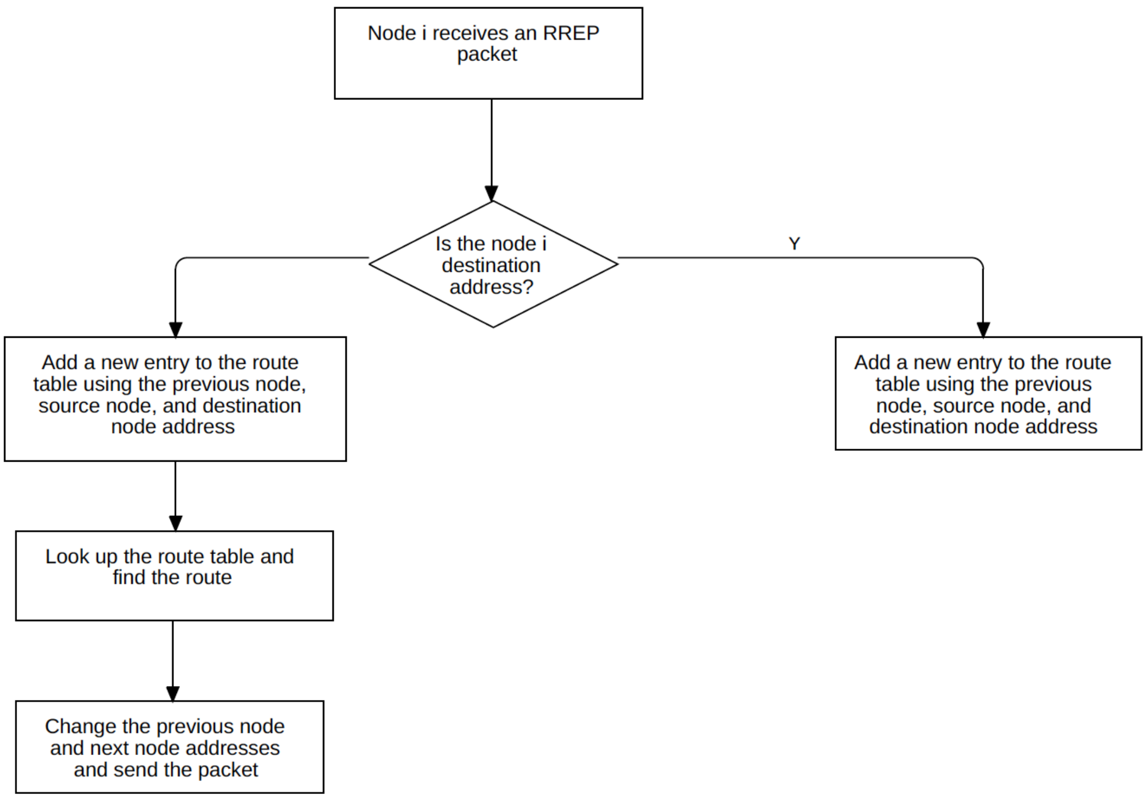


Figure 5: Flowchart of Receiving a RREP Packet

of the algorithms utilized for plant disease detection and weed detection in farm fields.

3.2.1 Disease Detection on Plant Leaves

When a large farm area with many plants is considered, noticing whether a plant is healthy or not, or deciding the type of its disease can take a long time. However, this is an issue that should be dealt with as early as possible to take precautions before the damages become irreversible. Research in computer vision for agriculture has resulted in image datasets that could be utilized for building accurate disease detection models. The PlantVillage dataset [14] is one such resource containing 54303 healthy and diseased leaf images divided into 38 categories by species and diseases as shown in Table 1.

Due to their longstanding success in image classification tasks, we propose the use of convolutional neural networks (CNN) for detecting plant diseases in the framework. CNN is a class of deep neural networks that has achieved remarkable success in many machine learning tasks, especially in the computer vision domain. A neural network is a learning architecture consisting of layers with filters and functions, each of which resembles neurons in the human brain. Layers in a CNN take inputs in three dimensions, which are width, height, and depth. This architecture is convenient for image inputs.

In the CNN architecture, there are three main layers, which are Convolutional Layer, Pooling Layer, and Fully-Connected Layer. Convolutional Layer and Pooling Layer provide feature extraction from the input to analyze it and the Fully Connected Layer transforms them into a single vector, learns the relation between features and their labels, then creates an output. The number of layers and parameters in each function varies in different models. Figure 6 shows the basic structure of a CNN.

MobileNet [13] is a CNN architecture, where convolutional layers are in are in the form of depthwise

	Diseased	Healthy
Apple (3172)	Gymnosporangium juniperivirginianae (276) Venturia insequalis (630) Botryospaeria obtuse (621)	(1645)
Blueberry (1502)		(1502)
Cherry (1906)	Podosphaera spp (1052)	(854)
Corn (3852)	Cercospora zeaemaydis (513) Puccinia sorghi (1192) Exserohilum turcicum (985)	(1162)
Grape (4063)	Guignardia bidwellii (1180) Phaeomoniella spp. (1384) Pseudocercospora vitis (1076)	(423)
Orange (5507)	Candidatus Liber ibacter (5507)	
Peach (2657)	Xanthomonas campestris (2291)	(360)
Bell Pepper (2475)	Xanthomonas campestris (997)	(1478)
Potato (2152)	Alternaria solani (1000) Phytophthora Infestans (1000)	(152)
Raspberry (371)		(371)
Soybean (5090)		(5090)
Squash (1835)	Erysiphe cichoracearum/ Sphaerotheca fuliginea (1835)	
Strawberry (1565)	Diplocarpon earlianum (1109)	(456)
Tomato (18162)	Alternaria solani (1000) Septoria lycopersici (1771) Corynespora cassiicola (1404) Fulvia fulva (952) Xanthomonas campestris pv. Vesicatoria (2127) Phytophthora Infestans (1910) Tomato Yello Leaf Curl Virus (5357) Tomato Mosaic Virus (373) Tetranychus urticae (1676)	(1592)

Table 1: Contents of the PlantVillage Dataset

separable convolutions, which consist of depthwise convolutions and pointwise convolutions as shown in Figure 7. Depthwise convolutions apply a single filter to each input channel and a pointwise convolution applies a 1×1 convolution to the output of depthwise convolution to create its linear combination. The difference of depthwise separable convolution from standard convolution is that depthwise separable convolution does filtering and combining in separate layers.

After depthwise convolution and poolwise convolution, batchnorm, and ReLu activation functions and after Fully Connected Layer, softmax function are used as shown in Figure 8. With convolutional, pooling, and fully connected layers, the MobileNet architecture consists of 28 layers. As mentioned in [13], MobileNet provides between 8 to 9 times less computation than standard convolution by separating convolutions into depthwise and pointwise.

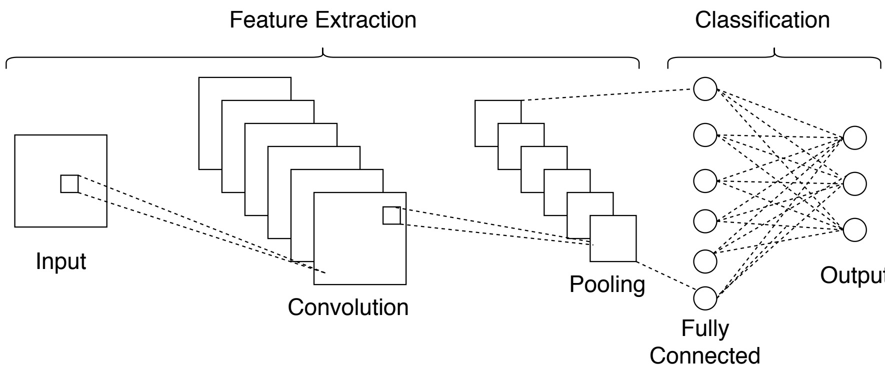
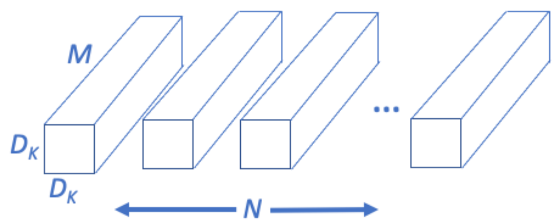
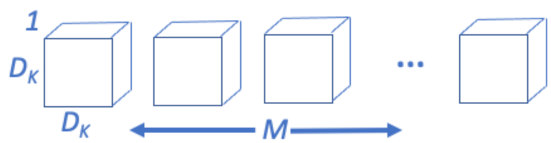


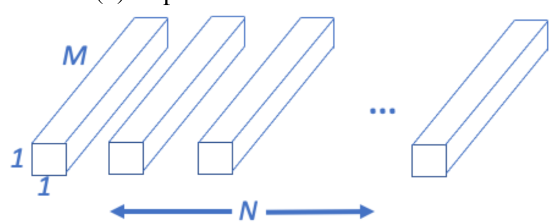
Figure 6: Convolutional neural network architecture



(a) Standard convolution filters



(b) Depthwise convolutional filters



(c) 1 x 1 convolutional filters called pointwise convolution in the context of depthwise separable convolution

Figure 7: Standard convolution layers vs depthwise convolutional layers[13]

3.2.2 Image Downsampling

Image downsampling is an image scaling method that compresses images. In this technique, redundant bits in images are removed and their sizes are reduced. Since high-resolution images have high storage space requirements and the cost of transferring them is expensive in terms of bandwidth and energy consumption, downsampling will be an important component of agricultural image processing systems with the growing data with widespread use. When utilizing downsampling, it is important not to lose essential information that affect the accuracy of the disease detection models. Three popular image compression methods include nearest neighbor interpolation, bilinear interpolation, and bicubic interpolation. While nearest neighbor interpolation simply changes the value of a pixel with the value of its nearest neighbor

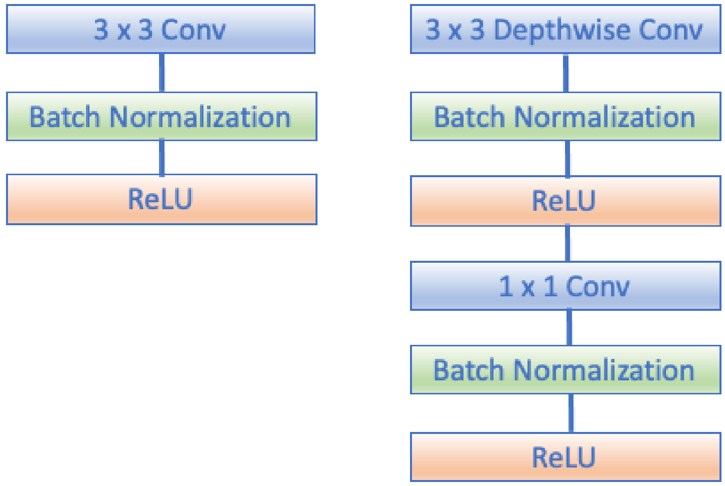


Figure 8: Left: Standard convolution layers with batchnorm and ReLu. Right: Depthwise separable convolutional layers followed by batchnorm and ReLu [13]

pixel, the bilinear interpolation method does this by taking the weighted average of the values of its four closest neighbor pixels on the horizontal and vertical line, and finally, the bicubic interpolation method does this with its sixteen neighbor pixels.

Images in the PlantVillage dataset are in JPG format and their sizes are 256x256. In order to evaluate the success of each compression method in this work, we have utilized these three compression methods and ratios of 0.5, 0.25, and 0.125 to downsample the images to the sizes of 128x128, 64x64, and 32x32, respectively. The effectiveness of each method has been evaluated using their Mean Squared Error (MSE) and Peak Signal to Noise Rate (PSNR) values that are given by:

$$MSE = \frac{1}{uv}\sum_{i=1}^{u}\sum_{j=1}^{v}(m_{ij}-n_{ij})^2$$
$$PSNR = 10log_{10}\frac{255^2}{MSE}$$

where u, v are the numbers of rows and columns respectively and m_{ij}, n_{ij} denote the original and reconstructed signals respectively, where i = 1 : u and j = 1 : v.

4 Evaluation

In order to assess the feasibility of utilizing the proposed framework for low-cost precision agriculture, we have evaluated the performances of the simulated LoRa-based WSN using a LoRa simulator and the accuracy of plant disease detection algorithms using a public dataset. The below subsections provide the evaluation results.

4.1 Proposed WSN Performance Results

For evaluating the implementation and observing network performance metrics, we used the OmNet++ tool [2], and for modeling LoRa nodes and gateways we utilized the FLoRa framework[23]. The device that we simulated is Semtech SX1272/73[9] with a voltage supply of 3.3 V. Figure 9 shows a screenshot of the Flora simulation framework.

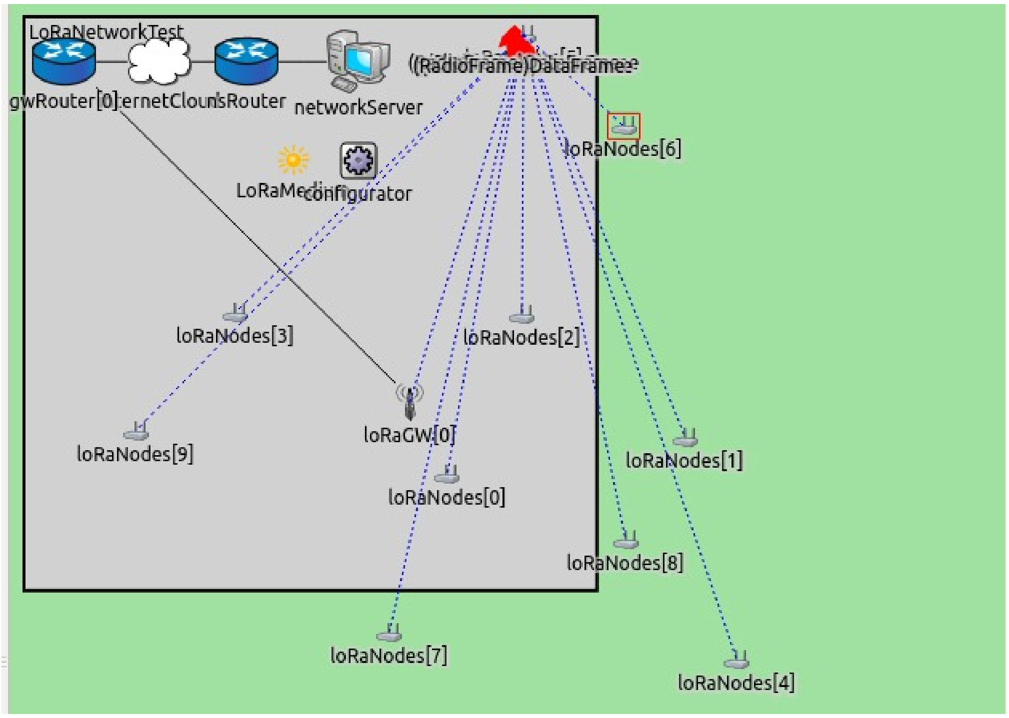


Figure 9: Flora simulation framework screenshot

What we aim to observe in these performance experiments is the effect of different device parameters on both cover range and energy consumption. We tested three important parameters in the device, and they are Spreading Factor, Bandwidth, and Transmission Power.

Spreading Factor: This metric simply equals the number of chips per data symbol. The value of the Spreading Factor can be between 7 and 12 in LoRa devices.

Bandwidth: This metric represents the width of frequencies in the transmission band. In LoRa communication, it should be either 125 kHz, 250 kHz, or 500 kHz.

Transmission Power: This is a term related to the distance that a signal can reach. The value of transmission power can be between 2 dBm and 20 dBm in LoRa devices.

4.1.1 Cover Range Experiments

In LoRa-based communication, there are two terms about the cover range. The first one is RSSI (Received Signal Strength Indicator) and the other one is sensitivity. Simply put, if the RSSI is greater than or equal to the sensitivity, the packet can be captured by the LoRa module. From the below graphs, we can see the effects of the Spreading Factor and Transmission Power on the cover range.

We observe that LoRa provides a long cover range to its users. Also, since there are no specific optimal values according to the results, we should adjust the parameters considering the size of the farmland where the system will be placed.

4.1.2 Energy Consumption Experiments

In order to observe the energy consumption of a LoRa device, we simulated the device that sends the packets every four hours. The daily energy consumption of this device is shown in Figure 12 below.

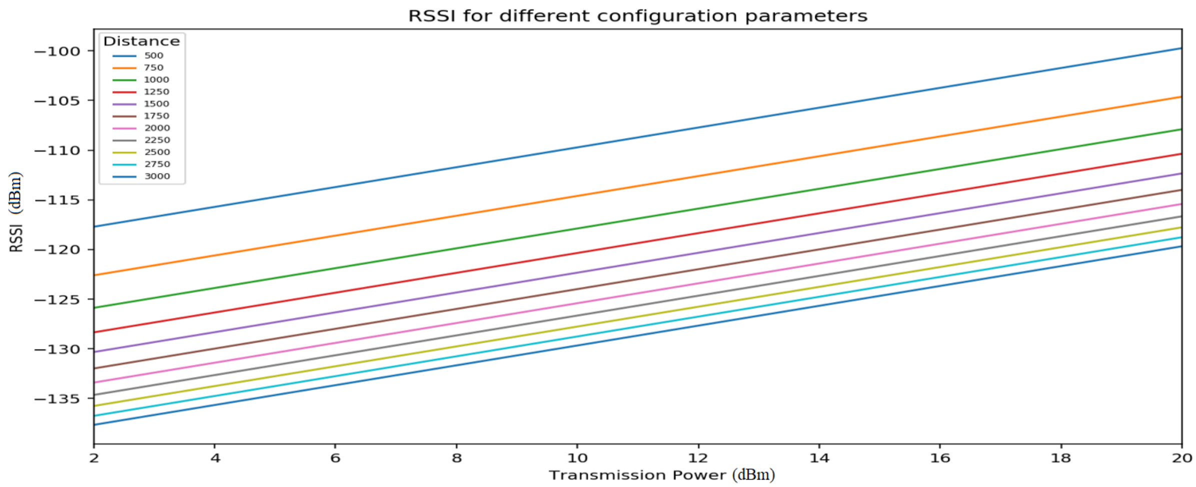


Figure 10: RSSI values

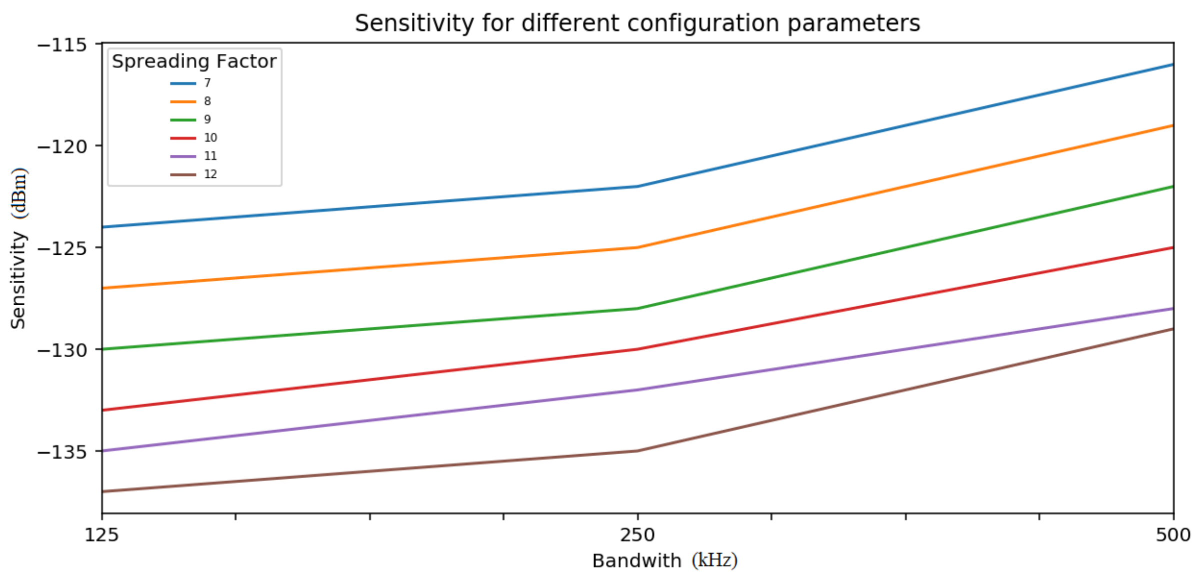


Figure 11: Sensitivity values

We can observe that when transmission power is increased, energy consumption also increases, but not in a linear fashion. For the optimal energy consumption, the transmission power parameter should be chosen considering the farmland size and features (e.g. density of plants and trees etc.) too.

4.2 Plant Disease Detection Experiments

For optimizing a deep learning model there are two different strategies: Training the model from scratch and transfer learning. While in training from scratch, a model learns from only the given dataset, in transfer learning it uses another, especially very large, dataset for feature extraction and then applies its knowledge to the given dataset, making connections between the pre-learned dataset and the given dataset. Since the transfer learning technique decreases training time, we used it by processing our data with MobileNet's pre-trained version where its weights were originally obtained by training on the

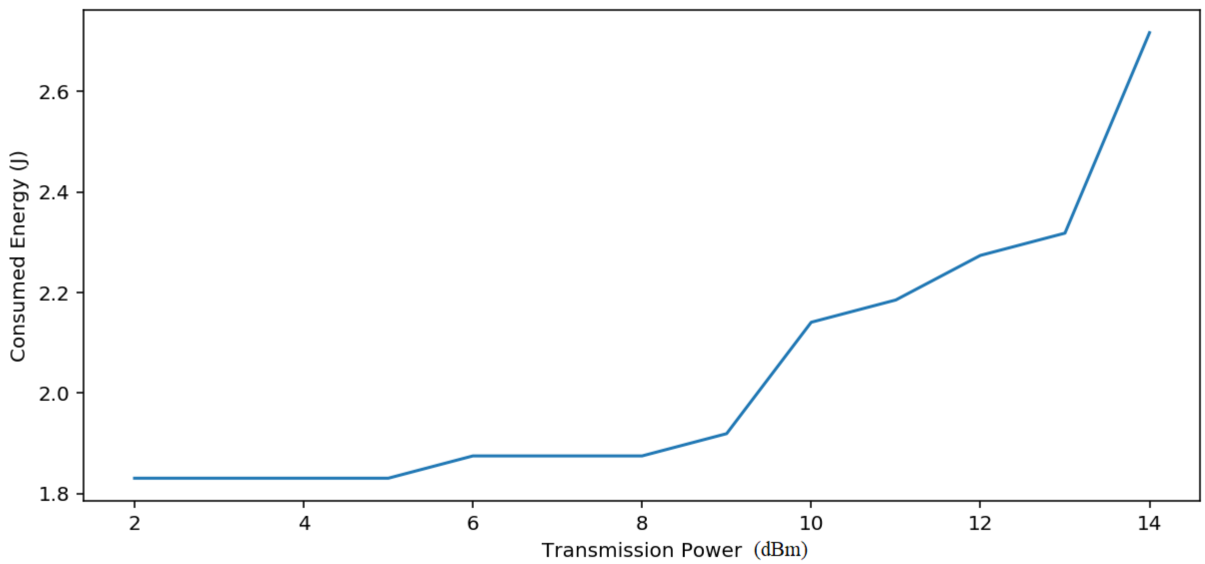


Figure 12: Daily energy consumption with respect to transmission power

ILSVRC-2012-CLS dataset [20], which belongs to ImageNet [11] and contains more than 1 million images.

The evaluation dataset was separated into training and testing data with 80% and 20% ratios, respectively. Therefore, the machine learning model was trained with 43428 images and tested with 10875 images, each in RGB format. Before training the model, data augmentation was performed, which is a technique to increase the size of the dataset by making changes in the dataset itself. We rotated, flipped horizontally, shifted width and height, zoomed, sheared, and rescaled the dataset, and finally, we trained the model with 100 training steps.

After training the MobileNet model with the PlantVillage dataset in 80% training and 20% testing ratios, we reached 0.9262 training accuracy, 0.9567 validation accuracy, 0.2339 training loss, and 0.1446 validation loss as shown in Figure 13.

4.2.1 Image Downsampling

Images in the PlantVillage dataset were downsampled in MATLAB R2020a. An example image belonging to the Apple-Venturia insequalis class is shown in Figure 14 and its downsampled versions are shown in Figure 15, Figure 16 and Figure 17.

In order to compare these compression methods, we can view their MSE and PSNR values shown in Table 2 and 3. We can see that images downsampled with bicubic interpolation have the highest PSNR and the lowest MSE values. The bicubic interpolation method provides images that are more similar to their original versions.

	Bicubic Interpolation	Bilinear Interpolation	Nearest Neighbor Interpolation
0.5	167.308	195.551	344.595
0.25	240.134	267.691	491.311
0.125	326.922	367.736	667.510

Table 2: MSE Values of Downsampling Methods

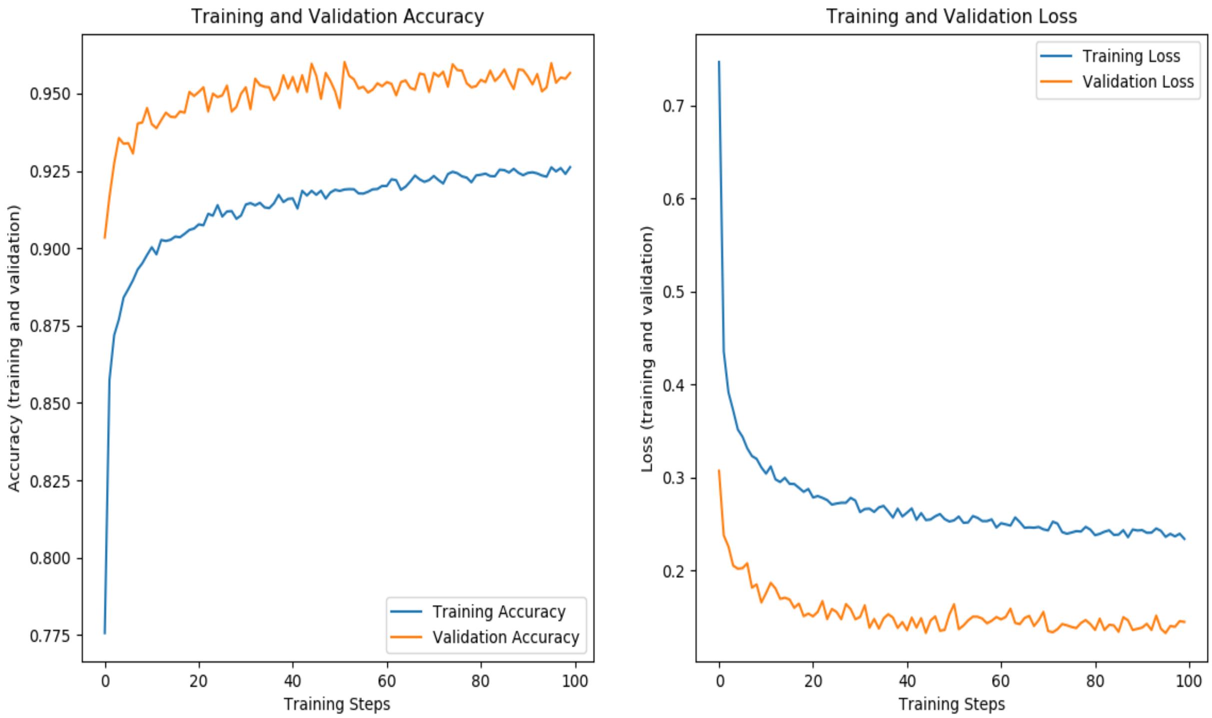


Figure 13: Changes in Training and Validation Accuracy and Training and Validation Loss in 100 Training Steps

Figure 14: Original sized image

Figure 15: Image downsampled with Bicubic Interpolation and ratios of 0.5, 0.25, 0.125

Moreover, when we tested the MobileNet method trained with the original-sized image dataset, we

Figure 16: Image downsampled with Bilinear Interpolation and ratios of 0.5, 0.25, 0.125

Figure 17: Image downsampled with Nearest Neighbor Interpolation and ratios of 0.5, 0.25, 0.125

	Bicubic Interpolation	Bilinear Interpolation	Nearest Neighbor Interpolation
0.5	28.277	27.256	24.787
0.25	26.086	25.408	22.891
0.125	24.303	23.656	21.163

Table 3: PSNR Values of Downsampling Methods

achieved the results shown in Table 4. We can observe that when we downsample the images with the Nearest Neighbor Interpolation method, the MobileNet model gives more accurate results than those of the other compression models.

	Bicubic Interpolation	Bilinear Interpolation	Nearest Neighbor Interpolation
0.5	68.27%	61.16%	73.19%
0.25	17.05%	13.94%	17.73%
0.125	5.12%	3.63%	4.57%

Table 4: Accuracy Values for Compressed Images

5 Conclusion

In this work, we proposed a digital twins-based smart agriculture framework utilizing LoRaWAN for sensor networks in the farm fields and intelligent processing of aerial imagery for plant disease and nutrient deficiency detection. LoRaWAN provides a robust solution for communication in rural fields without Internet access. We simulated the proposed wireless sensor network communication model with different configurations to observe the cover range and energy consumption. We performed experiments for detecting plant diseases from leaf images using the MobileNet CNN model and achieved 0.95 accuracy. We also tested different compression techniques and observed that while images compressed with bicubic interpolation are more similar to original ones, images compressed with nearest neighbor interpolation provide more accurate results with MobileNet.

The proposed digital twins-based approach will lay the ground for exciting research in IoT-based sustainable smart agriculture, by design and development of a framework that digitally represents a farm field to enable constant monitoring of soil parameters, automated detection of crop diseases and prescription of fertilizer treatment plans to achieve high yields at low cost. In future work, we plan to deploy the proposed system in a real farm field after extensive simulations for finding the best WSN configurations, taking into consideration the security of LoRaWAN [26] and the cloud [22]. The lifetime of the WSN can also be further optimized by using machine learning techniques as proposed by Kang et al. [17]. The end product will have important contributions for the agricultural economy of countries worldwide, especially those currently lagging in the use of smart technologies, causing abandonment of farms due to the difficulties faced. It will not only help farmers achieve increased crop yield at reduced production costs, but will also provide improved nutritional value of crops, which is important both for export potential and the overall health of the citizens that consume the produce.

6 Acknowledgments

This research is supported by the Royal Society, UK under grant number NMG\R1\191128.

References

[1] Lorawan frequency plans and regulations by country. `https://www.thethingsnetwork.org/docs/lorawan/frequencies-by-country.html` [Online; accessed December 20, 2020].

[2] Omnet++ discrete event simulator, April 2020. `"https://omnetpp.org/` [Online; accessed December 20, 2020].

[3] L. Alliance. About lora alliance. `https://lora-alliance.org/about-lora-alliance` [Online; accessed December 20, 2020].

[4] L. Alliance. What is the lorawan® specification? `https://lora-alliance.org/about-lorawan` [Online; accessed December 20, 2020].

[5] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso. Do lora low-power wide-area networks scale? In *Proc. of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM'16), Malta*, page 59–67. ACM, November 2016.

[6] T. Chen, D. Eager, and D. Makaroff. Efficient image transmission using lora technology in agricultural monitoring iot systems. In *Proc. of the 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom'19) and IEEE Cyber, Physical and Social Computing (CPSCom'19) and IEEE Smart Data (SmartData'19), Atlanta, Georgia, USA*, pages 937–944. IEEE, July 2019.

[7] Chetan Dwarkani M, Ganesh Ram R, Jagannathan S, and R. Priyatharshini. Smart farming system using sensors for agricultural task automation. In *Proc. of the 2015 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development (TIAR'15), Chennai, India*, pages 49–53. IEEE, July 2015.

[8] G. Codeluppi, A. Cilfone, L. Davoli, and G. Ferrari. Lorafarm: A lorawan-based smart farming modular iot architecture. *Sensors*, 20(7):2028, March 2020.

[9] S. Corporation. Sx1272/73 datasheet, 2017. `http://www.mouser.com/datasheet/2/761/sx1272-1277619.pdf` [Online; accessed December 20, 2020].

[10] D. Curseu, M. Popa, D. Sirbu, and I. Stoian. Potential impact of climate change on pandemic influenza risk. In I. Dincer, A. Hepbasli, A. Midilli, and T. H. Karakoc, editors, *Global Warming: Engineering Solutions*, pages 643–657. Springer US, October 2010.

[11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and F.-F. Li. Imagenet: A large-scale hierarchical image database. In *Proc. of the 2009 International Conference on Computer Vision and Pattern Recognition (ICPR'09), Miami, Florida, USA*, pages 248–255. IEEE, June 2009.

[12] M. Eckhart and A. Ekelhart. Digital twins for cyber-physical systems security: State of the art and outlook. In S. Biffl, M. Eckhart, A. Lüder, and E. Weippl, editors, *Security and Quality in Cyber-Physical Systems Engineering*, pages 383–412. Springer-Cham, November 2019.

[13] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861, April 2017. `https://arxiv.org/abs/1704.04861` [Online; accessed December 20, 2020].

[14] D. P. Hughes and M. Salathé. An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing. arXiv:1511.08060, November 2015. `https://arxiv.org/abs/1511.08060` [Online; accessed December 20, 2020].

[15] M. Hunter, R. Smith, M. Schipanski, L. Atwood, and D. Mortensen. Agriculture in 2050: Recalibrating targets for sustainable intensification. *BioScience*, 67(4):386–391, January 2017.

[16] B. Kang, D. Park, K. Cho, C. Shin, S. Cho, and J. Park. A study on the greenhouse auto control system based on wireless sensor network. In *Proc. of the 2008 International Conference on Security Technology (SecTech'08), Hainan Island, China*, pages 41–44. IEEE, December 2008.

[17] J. Kang, J. Kim, and M. M. Sohn. Supervised learning-based lifetime extension of wireless sensor network nodes. *Journal of Internet Services and Information Security*, 9(4):59–67, November 2019.

[18] F. Keesing et al. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, 468(7324):647–652, December 2010.

[19] B. Ray. What is lora? a technical breakdown, June 2018. `https://www.link-labs.com/blog/what-is-lora` [Online; accessed December 20, 2020].

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, April 2015.

[21] SemTech. What is lora? `https://www.semtech.com/lora` [Online; accessed December 20, 2020].

[22] C. Shih, J. Chou, and K. Lin. Wukong: Secure run-time environment and data-driven iot applications for smart cities and smart buildings. *Journal of Internet Services and Information Security*, 8(2):1–17, May 2018.

[23] M. Slabicki, G. Premsankar, and M. Di Francesco. Adaptive configuration of lora networks for dense iot deployments. In *Proc. of the 2018 IEEE/IFIP Network Operations and Management Symposium (NOMS'18), Taipei, Taiwan*, pages 1–9. IEEE, April 2018.

[24] P. Udawant. Diseased portion cassification & recognition of cotton plants using convolution neural networks. *International Journal of Engineering and Advanced Technology*, 8(6):3492–3496, August 2019.

[25] A. Vibhute and S. K. Bodhe. Applications of image processing in agriculture: A survey. *International Journal of Computer Applications*, 52(2):34–40, August 2012.

[26] I. You, S. Kwon, G. Choudhary, V. Sharma, and J. Seo. An enhanced lorawan security protocol for privacy preservation in iot with a case study on a smart factory-enabled parking system. *Sensors*, 18(6):1888, June 2018.

[27] W. Zhao, S. Lin, J. Han, R. Xu, and L. Hou. Design and implementation of smart irrigation system based on lora. In *Proc. of the 2017 IEEE Globecom Workshops (GLOCOMW'17), Singapore*, pages 1–6. IEEE, December 2017.

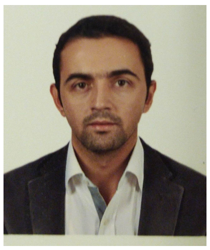
Author Biography

Pelin Angin is an Assistant Professor of Computer Engineering at Middle East Technical University. She completed her B.S. in Computer Engineering at Bilkent University in 2007 and her Ph.D. in Computer Science at Purdue University, USA in 2013. Between 2014-2016, she worked as a Visiting Assistant Professor and Postdoctoral Researcher at Purdue University. Her research interests lie in the fields of cloud computing and IoT security, distributed systems, 5G networks and blockchain. She is among the founding members of the Systems Security Research Laboratory and an affiliate of the Wireless Systems, Networks and Cybersecurity Laboratory at METU. She serves on the editorial boards of multiple journals on IoT, wireless networks and mobile computing.

Mohammad Hossein Anisi is currently an Assistant Professor with the School of Computer Science and Electronic Engineering, University of Essex, and head of the Internet of Everything Laboratory. Prior to that, he worked as a Senior Research Associate with the University of East Anglia, U.K., and a Senior Lecturer with the University of Malaya, Malaysia, where he received Excellent Service Award for his achievements. As a computer scientist, he has designed and developed novel architectures and routing protocols for the Internet of Things (IoT) enabling technologies including wireless sensor and actuator networks, vehicular networks, heterogeneous networks, body area networks, and his research results have directly contributed to the technology industry. He has a strong collaboration with industry and working with several companies in the U.K., with the focus on monitoring and automation systems based on IoT concept capable of reliable and seamless generation, transmission, processing, and demonstration of data. He has published more than 80 articles in high-quality journals and several conference papers and won two medals for his innovations from PECIPTA 2015 and IIDEX 2016 expositions. His research has focused specifically on real world application domains such as energy management, transportation, healthcare and other potential life domains. He has received several international and national funding awards for his fundamental and practical research as PI and Co-I. He is a Senior Member of the Institute of Research Engineers and Doctors (the IRED), a Member of ACM, IEEE Council on RFID, the IEEE Sensors Council, the IEEE Systems Council and International Association of Engineers (IAENG). He is a Fellow of Higher Education Academy. He is also a Technical Committee Member of Finnish-Russian University Cooperation in Telecommunications (FRUCT). He has been also serving as an executive/technical committee member of several conferences. He is an Associate Editor of a number of journals including IEEE ACCESS, the Ad Hoc & Sensor Wireless Networks, the IET Wireless Sensor Systems, International Journal of Distributed Sensor Networks. the KSII Transactions on Internet and Information Systems journals and Journal of Sensor and Actuator Networks. He has been a guest editor of special issues of the journals and Lead organizer of special sessions and workshops at the IEEE conferences such as the IEEE CAMAD, the IEEE PIMRC, and the IEEE VTC.

Furkan Göksel is an undergraduate computer engineering student at Middle East Technical University (METU), Turkey. His main research interests are in the fields of exploitation, operating systems, reverse engineering, and application security.

Ceren Gürsoy is a fourth year computer engineering student at Middle East Technical University. Her research interests focus mainly on artificial intelligence and machine learning.



Asaf Büyükgülcü is a B.S. student at the Department of Computer Engineering at Middle East Technical University (METU). His research interests include mobile computing and computer networks.