2024 23rd ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN) | 979-8-3503-6201-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IPSN61024.2024.00026

# Demo Abstract: Underground Potato Root Tuber Sensing via a Wireless Network

Tao Wang
Harbin Institute of Technology
Shenzhen, Guangdong, China
hitszwangtao@gmail.com

Yang Zhao
Harbin Institute of Technology
Shenzhen, Guangdong, China
yang.zhao@hit.edu.cn

Jie Liu
Harbin Institute of Technology
Shenzhen, Guangdong, China
jieliu@hit.edu.cn

Yujie Zhuang
Harbin Institute of Technology
Shenzhen, Guangdong, China
zhuangyujiecs@gmail.com

## ABSTRACT

We propose to demonstrate a novel underground potato root tuber sensing framework using deep learning algorithms. We build a data acquisition system for capturing the ground truth of the tuber shape and location underground as well as the received signal strength (RSS) measurements from a wireless network. Then we design a two-stage neural network to reconstruct the cross-section images of potato tubers. Our initial experimental results show that the reconstructed images can be used to predict the size and location of various potato tubers buried underground with high accuracy. We will demonstrate the real-time performance of our prototype.

## KEYWORDS

wireless network, deep learning, underground wireless sensing

## 1 INTRODUCTION

Accurate and non-destructive crop biomass sensing is critical for crop growth monitoring, crop phenotyping, and other smart farming applications. While various sensors and methods have been proposed for crop above-ground biomass sensing, below-ground biomass sensing remains largely understudied. With the advancement of wireless sensing technology, radio frequency (RF) sensors provide a promising approach to capture the shape, size and location of underground objects due to their "see-through" soil capability. Recently, [2] proposes to use a ground penetrating radar (GPR) to detect and reconstruct images of below-ground roots. It uses a deep neural network (DNN) to locate root branches in each 2D GPR sensing slice, and then reconstruct the root structure from multiple scenes. However, the sensing range of an GPR sensor is constrained by the field of view of its antennas, and the applicability is hindered by its high cost. In this demo, we propose an alternative approach, a novel underground root tuber sensing framework, using RF measurements from a low-cost wireless network. Our framework can not only reconstruct the maximum cross-section of below-ground potato tubers but also recognize tuber sizes.

## 2 FRAMEWORK

As shown in 1, in this demo, we propose to demonstrate a below-ground potato tuber sensing framework using a wireless network and DNN algorithms. Specifically, we first develop a data acquisition system with a ZigBee wireless network. Then, we design a two-stage neural network to reconstruct their maximum cross-sections and recognize the size category of potato tuber using RF measurements from the wireless network. Finally, we project our estimated cross-section images onto the soil via a projector tied to the PVC bracket. A video of this demo is available online[1].

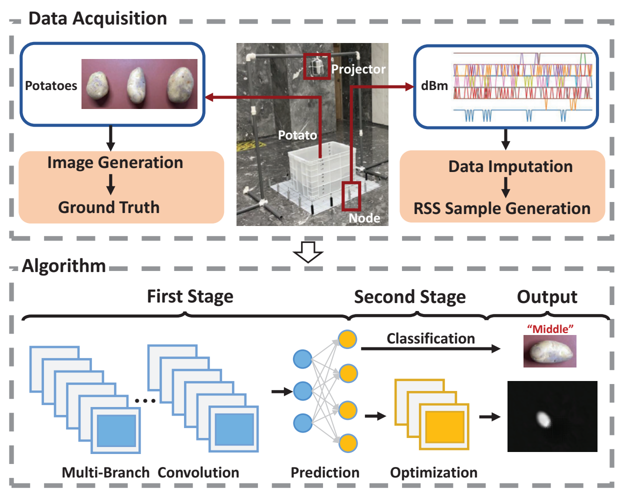


**Figure 1: An overview of our framework for potato tuber below-ground sensing. RSS measurements from a network of wireless nodes operating on 16 ZigBee channels are fed to a two-stage neural network to reconstruct the cross-section images of tubers, which will be projected to the surface of the soil.**

### 2.1 Data Acquisition System

*Testbed:* In our demo, we build a ZigBee mesh network using 16 TI CC2531 nodes. The wireless nodes operate on the 2.4$GHz$ ISM band and can transmit on one of 16 frequency channels, each separated by 5$MHz$. The wireless nodes utilize a multi-channel time multiple access (TDMA) communication protocol, as described in [1].

In our experimental setup, a sink node receives all packets transmitted by the nodes. The sink node is connected to a laptop where

[1]Online demo video: https://youtu.be/gqGjgKJ9BHE

979-8-3503-6201-5/24/$31.00 ©2024 IEEE
DOI 10.1109/IPSN61024.2024.00026

Authorized licensed use limited to: University Town Library of Shenzhen. Downloaded on September 04,2025 at 01:42:45 UTC from IEEE Xplore. Restrictions apply.

the RSS measurements are collected and stored for subsequent processing.

*Data imputation:* During the data collection process, three possible loss cases may occur. In the first scenario, a node in the receiving mode encounters an unsuccessful reception of packets transmitted by other nodes. The second scenario occurs when the other nodes in the network do not receive packets transmitted by a particular node. In the third scenario, all packets within a specific channel may be lost. To tackle the challenges mentioned above, we utilize contextual information for imputing the missing data. Specifically, when a node is unable to receive packets transmitted by other nodes, we employ values from the same frequency channel and transmitted node at different time instances for imputation. Then, when packets broadcast by a specific node are not received by other nodes, we impute the missing data by searching for values transmitted from the same node in the same frequency channel at a later time. Additionally, in the case where packets within a specific frequency channel are lost, we employ values associated with the same channel at a later time to perform imputation. Ultimately, the sample dimension utilized for training and testing is set to $16 \times 16 \times 16$, covering data from 16 frequency channels and 16 transceiver nodes.

*Ground truth generation:* To generate ground truth data for supervised deep learning, we initially measured the size and position of each potato tuber within the RF sensor coordinate system. Then, we utilize an image to represent the monitored area, where each pixel corresponds to a $1cm \times 1cm$ region within the monitored area. Based on the pre-measured size and position of the potato tuber, we label the pixel values corresponding to the potato tuber as 1, while assigning a value of 0 to all other pixel values in the monitored area. Finally, we incorporate the Gaussian smoothing method to refine the ground truth with a kernel size of $5 \times 5$. Furthermore, for potato tuber size recognition, we categorize each tuber into one of three classes: “large”, “middle” or “small”.

## 2.2 Algorithm

In this subsection, we present our data-driven image reconstruction algorithm, which includes two stages: (1) a multi-branch CNN-based network maps RSS measurements to reconstruct images of below-ground tubers and predict tuber size classes. (2) an UNet-based [3] network aims to eliminate noise and enhance the contrast ratio of the estimated results from the first stage.

*Multi-branch CNN-based network:* In this demo, we utilize a multi-branch CNN-based network, with each employing a pyramid architecture. In particular, each branch is composed of multiple sequential convolutional blocks, each of which contains a convolutional layer for feature extraction, a BatchNorm layer for managing covariate shift and facilitating model convergence, and a ReLU activation function for introducing nonlinearity to the neurons. Through using distinct convolutional sequences in the three branches, each characterizes RSS data in a specific dimension. Subsequently, the integration of features from the three convolutional branches is accomplished through addition, followed by the application of a dropout layer to mitigate potential overfitting issues. The output from the dropout layer undergoes a non-linear mapping process involving two fully connected layers and a ReLU activation function, enabling the prediction of the cross-section image. Simultaneously, it is fed into another non-linear mapping submodule, consisting of two fully connected layers and a ReLU activation function, facilitating the prediction of the potato tuber’s size class label.

*UNet-based network:* In this demo, we employ a lightweight UNet architecture with three convolutional blocks in both the encoder and decoder, leveraging convolutional and deconvolutional operations as well as skip connections to generate high-resolution results. In the encoder, each block utilizes a convolutional layer and the second and third blocks incorporate a max pooling layer to downsize the feature map. In the decoder, both the first and second blocks consist of a deconvolutional layer followed by a convolutional layer. The last convolutional block of the decoder has a single convolutional layer employing a $1 \times 1$ kernel size, specifically designed to reduce the channel dimensions of the decoder’s output. The ultimate output of the decoder is a monochromatic image, and all intermediate outputs are activated by a ReLU function. In this demo, we conduct skip connections between the results obtained from the second and third blocks in the encoder and those acquired from the second and first blocks in the decoder, respectively.

The mean square error (MSE) loss function is used in both stages to optimize the construction of cross-section images, while the first stage also employs the cross entropy loss function to enhance accuracy in identifying the size label for different potato tubers. Note that, the training dataset for the first stage and the second stage are different, which contributes to enhancing the universality of the neural network in the second stage.

# 3 DEMONSTRATION DESCRIPTION

In the demo, we plan to deploy a wireless network with 16 wireless nodes around a $60cm$ by $60cm$ area, where a container filled with soil will be positioned within the sensing area. Demo audience can pick one of the three potato tubers and bury it in the soil. Then, we will run our data acquisition system to capture RSS measurements and run our data imputation, DNN algorithms to reconstruct cross-section images and predict the size category of the potato tuber. Considering that environmental changes induce domain gaps, we will also employ a few-shot domain adaptation algorithm to fine-tune our algorithm, facilitating rapid adaptation to new environments. Without strong wireless interference, we expect to show audience the potato tuber shape and location prediction results immediately after all algorithms finishing onsite. Note that before placing the potato tuber into the container, a calibration process is conducted without any tuber present. As an interesting test, we can perform calibration for varying durations to observe the impact of calibration measurements on the imaging and prediction accuracy of the prototype.

# REFERENCES


[1] Ossi Kaltiokallio, Maurizio Bocca, and Neal Patwari. 2012. Enhancing the accuracy of radio tomographic imaging using channel diversity. In *2012 IEEE 9th international conference on mobile ad-hoc and sensor systems (MASS 2012)*. IEEE, 254–262.

[2] Yawen Lu and Guoyu Lu. 2022. 3D Modeling Beneath Ground: Plant Root Detection and Reconstruction Based on Ground-Penetrating Radar. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. IEEE, 68–77.

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 234–241.




Authorized licensed use limited to: University Town Library of Shenzhen. Downloaded on September 04,2025 at 01:42:45 UTC from IEEE Xplore. Restrictions apply.