2024 International Conference on Intelligent Computing and Robotics (ICICR) | 979-8-3503-7227-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICR61203.2024.00013

FPDTN: Fast Plant Digital Twin Model for IoT Automated Greenhouse Based on NeRF

Fazheng Xu[1]
Computer Science and Technology
Beijing Institute of Technology, Zhuhai
Zhuhai, China
stu2000507@cgt.bitzh.edu.cn

Juchi He[1]
Computer Science and Technology
Beijing Institute of Technology, Zhuhai
Zhuhai, China
Stellaris299792458@outlook.com

Haifei Huang*
Computer Science and Technology
Beijing Institute of Technology, Zhuhai
Zhuhai, China
haifei.huang@cgt.bitzh.edu.cn

***Abstract*—This study aims to develop an innovative and cost-effective digital twin technology called Fast Plant Digital Twin (FPDTN), leveraging Neural Radiance Fields (NeRF) for efficient 3D modeling of plant morphology in IoT-enabled greenhouses. FPDTN addresses scalability and financial constraints inherent in traditional digital twins, offering a novel solution. Through the creation of a unique dataset and methodology, FPDTN demonstrates superior accuracy in simulating plant morphology compared to existing models. The findings highlight FPDTN's significant advancements in digital twin technology tailored for botanical applications, emphasizing the critical necessity for scalable, cost-effective solutions adaptable across diverse domains. This paper concludes by emphasizing FPDTN's substantial contributions to the field and advocating ongoing exploration and innovation in digital twin technologies.**

Keywords—digital twin, NeRF, plant greenhouse, IoT.

I. Introduction

Digital twin technologies serve as foundational solutions within the metaverse, encompassing diverse interpretations of digital twins [1]. A prevalent perspective defines a Digital Twin as a representation mirroring the physical and functional attributes of complex systems, such as construction projects. Essentially, a Digital Twin consists of three primary components: the physical entity, its virtual counterpart, and the linkage enabling their interaction. This linkage constitutes a flow of data, information, and knowledge, facilitated by advanced technologies including computer vision, the Internet of Things (IoT), high-speed networking, and sophisticated analytical methods. Data essential for this interaction are gathered from sensors embedded in the physical object to construct its virtual representation [2-4]. Originating from its initial application in aerospace production and product lifecycle management by NASA during the Apollo 13 moon mission [5], the concept of digital twins has extended to various sectors, including the greenhouse industry [6], and smart shop-floor management [7]. Over the years, numerous researchers have delved into and employed digital twin technologies, envisioning computational digital twins as reliable surrogates for tangible objects. These visions draw upon modeling platforms such as Unreal Engine (UE) and Unity [8], integrating past experiences concerning management practices, trends, and physical characteristics. They also harness data collected either in real-time or at predetermined intervals to convincingly simulate real-world objects. Consequently, digital twins facilitate enhanced observation and analysis.

While numerous projects endeavor to construct large-scale digital twins as a technical foundation, the inherent variability and intricate complexity of the natural world, coupled with the need for rapid model establishment, present significant challenges. Current 3D modeling heavily relies on manual input and complex formulaic modeling [9]. Despite the remarkable applications enabled by manual 3D engine modeling and laser scanning, ensuring model credibility, swift creation, and the reusability of digital twin models remains a formidable challenge. Success demands an approach accessible to diverse individuals, including car modifiers, crafters, or home remodelers, offering a quick and straightforward means to produce understandable yet reliable 3D digital twin models. Addressing this gap, our paper delves into a specific academic application: plant digital twins.

For botanists or agriculturalists, documenting various plant conditions is pivotal, with the plant's physical and environmental characteristics being particularly crucial. Many plants cannot be easily relocated to labs, and creating physical models or specimens on-site proves arduous and inefficient. Thus, the primary method has been photography and written records, which may not offer an intuitive representation of the plant and its surroundings. Digital twin technology presents a potential solution, but as mentioned, there exists a gap in applying digital twin technology to tackle this issue.

This paper introduces FPDTN, a digital twin leveraging NeRF models for believable plant 3D digital twin modeling that is convenient and quick to use. FPDTN necessitates workers to capture specific photographs of plants, and through well-designed data processing methods, it can generate credible digital twin models of plants.

The development of FPDTN entails a digital twin architecture that streamlines plant imagery capture, data processing, AI-driven 3D modeling, and integrates IoT for real-time monitoring. This architecture comprises three main components: a multi-camera method for capturing plant images from different angles, an automated process for organizing captured data into npz datasets for model training, and a NeRF-based model for creating 3D digital twins of plants. Additionally, it features an IoT-enabled greenhouse system for collecting and analyzing plant data in real-time, enhancing the digital twin's analytical capabilities. This architecture supports a broad range of applications, from academic research to entertainment, enabling rapid digital

[1]These authors contributed equally to this work.

979-8-3503-7227-4/24/$31.00 ©2024 IEEE
DOI 10.1109/ICICR61203.2024.00013

twin modeling via smartphones for outdoor and lab use, facilitating plant monitoring, identification, and online engagement. FPDTN evaluations encompass comparing 3D models to real plants, assessing AI-rendered image similarity to actual photos, analyzing dataset attribute effects on model quality, and testing IoT data reliability.

In summary, this paper offers the following contributions:

- FPDTN introduces a low-barrier, quick, and convenient method for creating digital twins of plants.
- A novel approach for the collection and processing of plant digital twin datasets.
- A maintenance and operation method for plant digital twins based on real-time IoT monitoring.

II. Related Work

A. NeRF

NeRF, short for Neural Radiance Fields, developed by Ben Mildenhall in 2020, is a technique for creating new views of complex scenes through a deep network from limited input views. It marks a major progress in view synthesis and 3D scene reconstruction by optimizing a continuous volumetric function with sparse inputs to produce high-quality scene visualizations. NeRF maps 5D coordinatesn location and viewing direction to determine volume density and radiance, synthesizing views with volume rendering techniques [10].

FPDTN leverages NeRF for digital twin modeling in agriculture and botany, aiming to revolutionize the creation and use of plant digital twins with rapid, accurate 3D plant reconstructions in automated greenhouses. Integrating IoT technology, FPDTN enhances precision monitoring and maintenance, utilizing NeRF's computational efficiency and scalability. This synergy is designed to improve plant growth monitoring, optimize greenhouse operations, and support plant research, species classification, and education, ultimately fostering more sustainable and productive practices.

B. Botany Greenhouse, Digital Twin and IoT

Digital twins (DTs) have become pivotal in Industry 4.0, providing detailed virtual replicas of physical entities. Their integration into botany is emerging, facing challenges like depicting dynamic plant growth [11]. However, DTs hold promise for enhancing agricultural yields, refining management practices, and educational engagement. Extending DTs to urban ecosystems, notably for tree species modeling in cities, highlights their transformative potential for urban planning and environmental sustainability [12]. Accurately capturing tree characteristics and interactions aids in designing greener cities, conserving biodiversity, and improving urban living conditions, addressing issues like urban heat and carbon sequestration.

Advancements in real-time monitoring technologies have been applied to agriculture, enhancing the study of plant physiological processes. A notable example by Shimomoto et al. [13] involves a system for monitoring photosynthesis and transpiration in tomato plants, showing precision agriculture progress. This real-time system facilitates understanding plant responses to environmental stresses and improves greenhouse management. By providing insights into plant efficiency and resource use, these technologies contribute to developing resilient agricultural systems, highlighting the integration potential of real-time monitoring with digital twin models for better agricultural and urban green space management.

III. Preliminaries

In the realm of computer image processing, NeRF represents a departure from traditional modeling techniques such as meshes, point clouds, or voxels. Instead, NeRF adopts a method where model information is implicitly stored within a neural network. This network is trained using images captured from known camera perspectives, and then, by inputting parameters from different camera viewpoints, it predicts images from previously unseen perspectives. Consider a set of camera poses in 3D space, defined by the tuple (x,y,z,θ,ϕ), along with a corresponding set of sampled image outputs, defined by (R,G,B,σ). These elements denote specific meanings as outlined in the accompanying Table I.

TABLE I. Variable Definitions for NeRF Modeling

Notation	Definition
(x, y, z)	Spatial location coordinates in 3D space, representing the position of a point in the scene.
θ	Azimuth angle of the viewing direction, indicating the horizontal angle in spherical coordinates.
ϕ	Elevation angle of the viewing direction, indicating the vertical angle in spherical coordinates.
(R, G, B)	The RGB color values emitted from a point in space, as observed from a specific viewing direction.
σ	The volume density at a point in space, acting as a differential opacity controlling how much radiance is accumulated by a ray passing through.

In the context of NeRF for synthesizing views, (θ, ϕ) represent the spherical coordinates that describe the direction of rays emitted from the camera position to the pixels on the imaging plane. These directions are analogous to the Yaw and Pitch in the Yaw-Pitch-Roll system, excluding Roll. Here, θ denotes the azimuthal angle, which is the angle between the projection of the point on the $x - y$ plane and the positive direction of the x-axis, typically within the range of $[0, 2\pi]$. ϕ represents the polar angle, or the angle between the point and the positive half of the z-axis, with a range of $[-\pi, \pi]$. It's important to note that the actual input to the network is a three-dimensional vector indicating the ray's direction (x, y, z), rather than the physical camera position. Therefore, each pixel's corresponding ray and camera location form a single piece of training data. This approach allows the rendering information of the 3D model to be implicitly stored within the deep learning model. The output includes the color values (R, G, B) and opacity σ for a set of sample points along the ray. Integration over these sample points along the ray direction is performed, where the first peak encountered has the most significant impact on coloring the pixel. The term "opacity" here can also be referred to as "density", which is independent of the observation angle (Fig. 1).

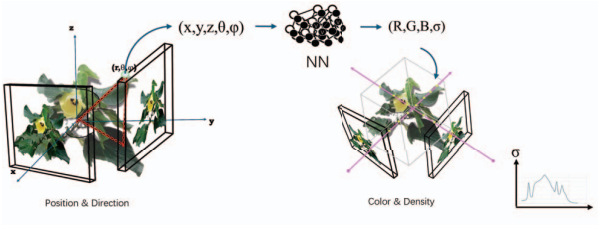


Fig. 1. Variables for NeRF

To accurately capture high-frequency details within images, position encoding is applied to the spatial coordinates (x, y, z). This encoding significantly enhances the detail quality of images without being affected by the smoothing of surrounding pixels. Position encoding integrates information by concatenating (parallelly aligning in the input dimension of the neural network) each basic encoding component, which consists of a pair of sine and cosine functions. For a total of L terms, the position encoding can be expressed as Eq. (1):

$$E(p) = (sin(2^0\pi p), cos(2^0\pi p)), \\ sin(2^{L-1}\pi p, cos(2^{L-1}\pi p)) \quad (1)$$

Where p represents x or y or z. The correlated embedded position input dimension in NN is 2L. The Fourier Feature Mapping can be thought of as the principle to support this method.

In the FPDTN model, NeRF's positional encoding is enhanced with a transformer-based encoder, critical for capturing spatial nuances in digital twin representations of IoT automated greenhouses. This encoder translates spatial coordinates (x, y, z) into a high-dimensional space, leveraging a transformer's self-attention mechanisms to emphasize spatial relationships. The process begins with a linear projection of coordinates into a higher dimensional space, represented by:

$$\boldsymbol{P}proj = \boldsymbol{W}p \cdot (x, y, z)^T + \boldsymbol{b}p \quad (2)$$

where $\boldsymbol{W}_p$ and $\boldsymbol{b}_p$ are the linear projection's weights and biases. The transformer encoder then processes $\boldsymbol{P}_{\{proj\}}$ through multiple layers, each comprising multi-head self-attention (MHA) and position-wise feed-forward networks (FFN) [14], to produce the encoded position vector $\boldsymbol{E}_{\{transformer\}}$:

$$\boldsymbol{E}_{\{transformer\}} = TransformerEncoder(\boldsymbol{P}_{\{proj\}}) \quad (3)$$

For MHA, input vectors are transformed into queries $(\boldsymbol{Q})$, keys $(\boldsymbol{K})$, and values $(\boldsymbol{V})$ using learnable weights, allowing the encoder to focus on relevant spatial positions through the attention mechanism:

$$\mathbf{Q}, \mathbf{K}, \mathbf{V} = \mathbf{P}proj^{(l-1)}\mathbf{W}^Q, \mathbf{P}proj^{(l-1)}\mathbf{W}^K, \mathbf{P}_{\text{proj}}^{(l-1)}\mathbf{W}^V \quad (4)$$

Attention and subsequent processing lead to:

$$\mathbf{E}_{\text{transformer}} = \mathbf{P}\text{proj}^{(N)} \quad (5)$$

where $\mathbf{P}_{\text{proj}}^{(N)}$ is the position vector after N layers, combining MHA and FFN, capturing complex spatial relationships essential for high-fidelity plant growth imagery in digital twins. This transformer-based approach provides a nuanced understanding of spatial relationships, vital for precise monitoring and control in automated greenhouses.

To optimize NeRF for faster rendering times, which can range from under a second to about 30 seconds, the hierarchical sampling technique is crucial, dividing the sampling process into "coarse" and "fine" stages to efficiently focus computational resources on areas of the scene that contribute most to the final image. This approach significantly reduces the number of required network evaluations.

Specifically, a two-network system is used: a coarse network to determine areas of interest and a fine network for detailed rendering in those areas. Mathematically, the optimization process can be outlined as follows:

Coarse Network Sampling: Firstly sample a set of N_c locations using stratified sampling and evaluate the coarse network to obtain a rough estimate of the scene.

Fine Network Sampling: Based on the output of the coarse network, a more focused sampling of N_f locations is performed, guided by the distribution of significant features identified in the first step.

Combined Rendering: The final image is rendered using samples from both the coarse and fine networks, optimizing the rendering process, and reducing computation time without sacrificing image quality.

The equation for rendering a color $C(r)$ of a camera ray r using the coarse-to-fine approach can be formalized as Eq. (6):

$$C(r) = \sum_{i=1}^{N_c+N_f} T_i(1 - \exp(-\sigma_i\delta_i))c_i \quad (6)$$

where T_i is the accumulated transmittance to sample i, σ_i is the density at sample i, δ_i is the distance between samples i and $i+1$, and c_i is the color at sample i. This method allows NeRF to render complex scenes with high fidelity at practical computation times.

IV. FPDTN Architecture

A. Data Description

To train and test the NeRF model, the pioneering NeRF dataset by Ben Mildenhall is selected. as our foundational dataset [10]. This dataset comprises path-traced images of eight objects, including ferns, flowers, fortresses, horns, leaves, and orchids, all showcasing items made of genuine and complex non-Lambertian materials. Among these, images from six objects were drawn from viewpoints sampled over the upper hemisphere, while the remaining two were from viewpoints sampled across the entire sphere. Each image, rendered at a resolution of 800 × 800 pixels, is accompanied by detailed camera parameters for every 3D viewpoint. These parameters include the camera's field of view along the X and Y axes ($camerangle_x$ and $camerangle_y$), determining the width and height of the

camera's view; the focal lengths along the X and Y axes (fl_x and fl_y), linked to the camera angle and controlling the scene's perspective effect; radial and tangential distortion coefficients (k_1, k_2, p_1, p_2) for correcting lens- induced image distortions; the coordinates of the optical center (cx, cy), indicating the shift between the image center and the camera sensor's center; and the image dimensions (w, h). Additionally, each image frame includes a transformation matrix ($transform\_matrix$) that details the specific location and orientation of the camera in 3D space, along with an image sharpness score ($sharpness$).

It is worth emphasizing that we focus on the development of NeRF models in the field of botany, addressing the inherent high dynamics and complexity of plant growth and morphology, which are often overlooked in traditional 3D modeling of static objects. The variation of plant morphology is not only controlled by gene expression, but also by environmental factors such as light, temperature and humidity. Together, these elements contribute to significant individual differences in plant structure over time. Therefore, the author designed and made a comprehensive plant data acquisition and observation system based on Arduino platform. The system is designed to accurately capture the growth process of the plant and the conditions of its surrounding environment. The observation system adheres to the data acquisition standard of the NeRF model, and four 1280*720 high-definition wide-angle camera arrays are arranged in the plant data acquisition and observation framework to capture the growth state of plants from all angles. The cameras are precisely calibrated to ensure that the images captured from the four viewpoints are correctly aligned and reconstructed. The system integrates a variety of sensors, including temperature (Celsius), humidity (%), photosynthetically active radiation intensity μmol/m²/s), carbon dioxide concentration (ppm, parts per million), to monitor and record various environmental parameters in real time. To ensure the diversity and complexity of the data, the authors simulated different growth environments of the plants, including controlled indoor environments and natural outdoor conditions. This allows us to collect data under a variety of light, temperature, and humidity conditions, thus enhancing the generalization ability of the dataset. The data contains the whole growth process from seed germination to plant maturity, covering a variety of plant species to ensure the wide applicability of model training. To precisely capture this biological complexity, 3D modeling techniques must adapt to the continuous changes in plant growth while capturing subtle morphological details, posing greater demands on data collection methods and modeling technologies. Especially in simulating plant growth environments and physiological responses, it is crucial to accurately control and monitor environmental conditions to better understand how they influence plant morphological development.

B. Digital Twin Generation Model Architecture

FPDTN is engineered to create high-fidelity, three-dimensional representations of plants within an IoT-enabled automated greenhouse by leveraging NeRF. One core of this system lies a Transformer-based encoder, which processes spatial coordinates (x, y, z) and viewing angles (θ, ϕ) as inputs. These are then embedded using positional encoding, enhancing the model's ability to capture fine-grained spatial variances. The encoder outputs a feature representation that is fed into a series of fully connected layers to infer the volume density and color for each point in space, ultimately rendering the scene with implicit lighting and view-dependent effects.

The NeRF model's architecture, consists of a series of dense layers (Fig. 2). With an input dimensionality expanded through positional encoding, the input data undergoes a transformation via a Transformer encoder to capture the complex interactions between different spatial locations. Following this, a sequence of linear transformations takes place. Notably, every fourth layer reintegrates the initial encoded features with the current layer's outputs, ensuring that the model maintains an awareness of the original spatial context throughout deeper layers. This design allows for a deeper network, with D representing the depth and W the width of the network, which can be flexibly adjusted based on the complexity of the plant models and computational resources.

The rendering pipeline is a crucial component of the FPDTN, translating the learned implicit field into color and density outputs at sampled points along rays cast through the scene. This process begins with ray generation, using the get rays function that casts rays from the camera position through each pixel on the image plane, directed towards the scene. The rays are sampled at uniform intervals between the near and far bounds to generate a set of points in 3D space. The model then predicts the color and density at each point. The render rays function computes the final pixel color by accumulating contributions along the ray according to the volume rendering equation, simulating how light interacts with the plant material.

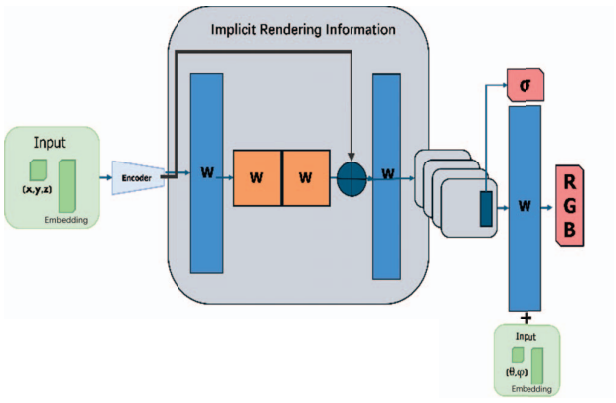


Fig. 2. NN Architecture for FPDTN

In training the FPDTN model, stochastic gradient descent is performed to minimize the discrepancy between the generated images and ground truth, using mean squared error as the loss function. Optimization is facilitated by the Adam optimizer, a choice driven by its adaptive learning rate capabilities that help in converging to high-quality solutions efficiently. During the iterative process, random subsets of the greenhouse scene are sampled, and the corresponding rays are rendered to simulate various lighting conditions and viewing angles. This approach enables the model to generalize well across different plant species and environmental setups within the greenhouse, effectively

creating accurate digital twins for precision agriculture applications.

C. Plant Greenhouse and IoT Structure

1) Overview of System Architecture: In this study, the author has successfully designed and implemented a plant data collection and observation system based on IoT technology greenhouse. At the heart of this system lies the Arduino UNO R4, which, in conjunction with a variety of sensors and control modules, facilitates the meticulous management of the plant growth environment by precisely monitoring and automatically adjusting environmental parameters such as temperature, humidity, and light levels. Furthermore, the aesthetic design of the entire system was conceived through 3D modeling techniques. The original design of the plant greenhouse was created using Sharp3D software, followed by fabrication and reshaping with acrylic materials, endowing the system with a unique appearance and practical structure. The greenhouse's enclosure comprises five parts, including a lid that can be opened and closed, a control compartment, a base, and a control panel. These components are both aesthetically pleasing and functional, ensuring overall system stability and ease of operation.

To enhance the data collection capabilities of our system further, the observation system is equipped with a quartet of 1280*720 high-definition ultra-wide-angle cameras, based on the data collection standards of the NeRF model [15], to capture the growth state of plants from all angles comprehensively. These cameras are precisely calibrated to ensure that images of plants captured from four different angles can be accurately aligned and reconstructed, thereby providing high- quality training and testing data for our NeRF model.

The system's architecture encompasses both hardware and software subsystems. The hardware component primarily consists of a DTH11 temperature sensor, humidity sensor, Risym12V soil moisture sensor, and SGP30 gas quality sensor, and utilizes an ESP32 wireless networking module along with a DX-BT05 Bluetooth module for cloud data transmission and storage. The software component, on the other hand, leverages the Blinker platform and various library functions to achieve data collection, processing, and remote monitoring capabilities. Through this system, the author has achieved the anticipated stable, accurate, and intelligent monitoring of specific data during the plant growth process.

2) Hardware subsystem design: Following a comprehensive overview of the system architecture, the author delves deeper into the design details of the hardware subsystem of the Plant Data Collection and Observation System—greenhouse. The architecture of this system is designed to accurately monitor and regulate the environmental conditions necessary for plant growth through an integrated network of sensors and highly automated control modules. The hardware composition of the system is divided into two main parts: sensors and controllers. Below is the wiring diagram of the system (Fig. 3).

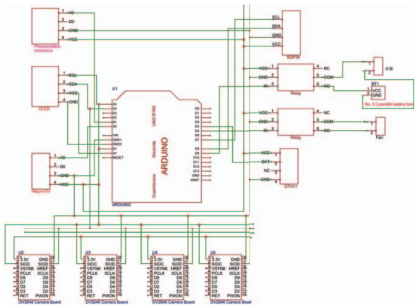


Fig. 3. System bus design diagram

Design of the Sensor Component:

- Temperature and Humidity Sensor (DHT11): Uses capacitive sensing elements to measure humidity and temperature in the air. The sensor determines the air humidity by measuring the capacitance change, the capacitance value increases with the increase of air humidity, and the temperature is measured by the thermistor, whose resistance value decreases with the increase of temperature. Its humidity measurement range is 20-90%RH (relative humidity), temperature measurement range is 0-50°C, humidity and temperature measurement accuracy can reach ±5%RH and ±2°C, respectively. The data is transmitted to the microcontroller via a single digital signal.
- Gas Sensor (SGP30): The sensor element using metal oxide (MOX) technology changes its resistance value by heating the MOX material, which is proportional to the detected gas concentration. It can detect a variety of volatile organic compounds (VOCs) and carbon dioxide levels in the air and has high sensitivity to the gas and fast response time. This is essential for assessing plant photosynthesis and respiration conditions.
- Photosensitive sensor: By measuring the ambient light intensity through a photosensitive resistor or photodiode, it can adjust its resistance value or generate current according to the change of ambient light intensity to ensure that the plant receives the right amount of light. The light intensity measurements can be used to adjust the brightness of the LED lamp array to simulate different insolation conditions and optimize plant photosynthesis.
- Soil Moisture sensor: It estimates the moisture content by measuring the electrical conductivity of the soil, usually using the principle of resistance or capacitance measurement. Resistive moisture sensors measure the resistance of the soil through two electrodes, and the resistance value decreases as the moisture of the soil increases. Capacitive moisture sensors measure the effect of the soil on the capacitance, the higher the moisture in the soil, the greater the capacitance value, and these data are used to automatically control the irrigation system to ensure that the soil maintains the appropriate moisture level.

Design of the Control Component:

- HD cameras: An array of four 1280*720 HD wide-angle cameras is used to capture continuous video streams or timed photos, providing users with real-time monitoring of plant growth. This data can be used to analyze the health status, growth rate and growth pattern of plants.
- Ventilation fan: Based on the computer silent cooling fan transformation, it can provide 6.2CFM air volume at the speed of 6000+-10%RPM. It can be automatically adjusted or manually turned on according to the temperature and humidity data of DHT11 sensor to control the environmental conditions by improving the air circulation in the greenhouse.
- 3v DC motor water pump: The small DC motor water pump is connected to a water tank and can automatically control irrigation based on data from the soil moisture sensor or manually turn on according to user Settings. When the soil moisture sensor detects that the soil moisture is below a preset value, the water pump automatically starts and delivers water from the tank to the plant roots until the soil moisture reaches a desired level.
- LED light array provides auxiliary lighting, automatically or manually turned on to adjust light intensity and spectral composition based on light sensor data and plant growth stage, simulating natural light conditions.
- 0.96-inch OLED Display module: This small display module can display environmental parameters (such as temperature, humidity, light intensity) and plant growth status in real time, providing users with intuitive information feedback.
- Air humidification module: A special microporous atomizer is driven by an atomization drive circuit board to provide air humidity for the system environment, and the environmental humidity is adjusted automatically or manually according to the humidity data of DHT11 sensor. When the environmental humidity is lower than the set value, the atomization module automatically starts, and the water is atomized into small water droplets to increase the air humidity until it reaches the ideal humidity level.
- DX-BT05 Bluetooth module: The DX-BT05 module supports Bluetooth low energy technology, which can wirelessly transmit data to smart devices or cloud platforms for remote monitoring and management. The Bluetooth module allows users to remotely view environmental parameters and plant growth status or even remotely adjust system Settings by collecting data from sensors and control systems and then sending them to pre-paired devices via Bluetooth signals.

3) Exterior design: The design of the system's enclosure is entirely original, utilizing Sharp3D software for 3D modeling, followed by printing and reshaping using acrylic material to form the main structure of a plant greenhouse. The system's enclosure comprises five main components: a hinged trans- parent cover, a control chamber, a base, and a control panel, ensuring both practicality and aesthetic appeal. As shown in Fig. 4.

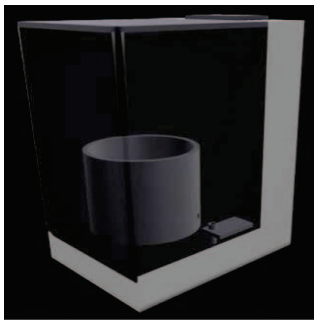


Fig. 4. Exterior Structure Modeling Design Drawing

Hinged Cover Design The design philosophy for the hinged cover incorporates several aspects, As shown in Fig. 5(a):

- Hinge Structure: To enable the lid to open and close freely, the designer installed a pair of appropriately sized hinges between the lid and the enclosure, securing the hinges and the top cover with hot glue. This design not only facilitates the opening and closing of the plant greenhouse but also better protects the plants inside.
- Transparent Material: Considering that plants require ample sunlight, the lid is made of transparent acrylic. Optionally, UV-resistant materials can be chosen to prevent issues like fading or aging due to intense ultraviolet rays, ensuring the lid's performance and appearance.
- Handle Design: A handle is positioned in the center of the lid, connected to the top cover with hot glue for easy lifting and handling. The design of the handle adheres to ergonomic principles, offering comfort and a stable grip. Servo Design: Two servos can be installed on the inner wall of the lid, providing the power to maintain the lid's open position, and enabling its opening and closing capabilities.

The design considerations for the control chamber's exterior structure include, as shown in Fig. 5(b):

- Slide Cover Design: A sliding cover is placed on top of the control chamber, fixed with suitable-sized rails for easy opening and closing, ensuring the safety of the internal equipment.
- Shell Material Choice: The shell of the control chamber is made of acrylic, offering lightness, better waterproofing, and allowing users to view the internal equipment's status. Integration of a Small Fan: A small fan is installed on the inner wall of the control chamber to ensure air circulation, maintaining a stable internal environment.

Base Design: The base's exterior structural modeling includes several considerations, as shown in Fig. 5(c) :

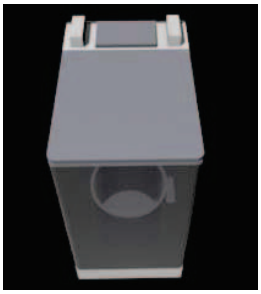
- Detachable Design: The base is designed to be detachable for easy cleaning and maintenance. A detachable lid connects the upper and lower parts of

the base, facilitating water addition for easy connection and separation.

- Silicone Hose Holes: Appropriately sized holes for silicone hoses are created on the base's top surface, allowing for easy insertion, and securing of water pipes while preventing displacement and leakage.
- Built-in Water Pump and Tank: The base contains ample space to accommodate a water pump and tank, with the pump responsible for transporting water from the tank to the plant soil via hoses.
- Durable, Waterproof Material: The base is made of sturdy and durable acrylic to support the weight of the water tank and equipment while providing waterproofing to ensure the greenhouse's temperature and humidity stability.

Control Panel Design: The design considerations for the control panel include, as shown in Fig. 5(d):

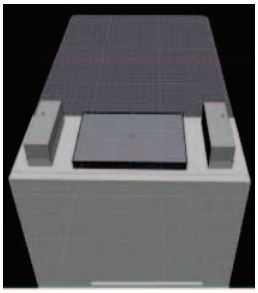
- Design Specifications: Features a 2.42-inch OLED screen and two joystick buttons.
- Button Layout: The two joystick buttons are ergonomically designed and symmetrically positioned on either side of the screen.
- Screen Selection: A 2.42-inch OLED screen is used to display more information content, offering lower power consumption and better viewing angles.



(a) Switchable cover design

(b) Control warehouse modeling design diagram

(c) Bottom warehouse modeling design



(d) Control panel modeling design diagram

Fig. 5. Exterior Designs

V. Experiments

A. 3D Model Scale and Proportion Evaluation

To evaluate the scale and proportion accuracy of our FPDTN model, the author tracked the model's development over iterations, beginning with its initialization. Fig. 6 illustrates the iterative refinement process of the 3D model. Initial stages

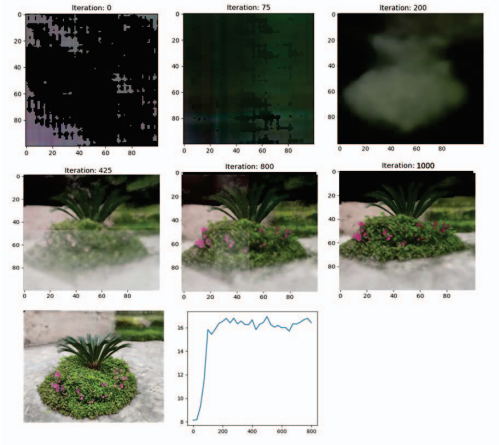


Fig. 6. DT Result of FPDTN

(Iterations 0, 75, 200) show a fragmented and indistinct plant structure, reflecting the early neural network's uncertainty in representing the spatial distribution of the plant's features. As training progresses (Iterations 425, 800), the model begins to coalesce into a coherent form, demonstrating the network's increasing proficiency in understanding the plant's geometry. By Iteration 1000, the model has attained a level of detail and precision that closely matches the real-world plant, with ac- curate representation of scale and proportions, showcasing the effectiveness of our FPDTN model in capturing the intricate details necessary for a reliable digital twin.

B. AI-Rendered vs. Real-World Image Similarity

To assess the fidelity of our AI-rendered images to real-world photographs, the author adopted a two-pronged approach. First, the author utilized a similarity metric that quantitatively measured the likeness between the NeRF-generated images and their real-world counterparts. Secondly, the author conducted a visual assessment, and compared the nuanced characteristics such as texture, color fidelity, and light interplay. This combined evaluation aimed to ensure that the digital twins not only met objective criteria but also passed the subjective scrutiny of human perception, solidifying the FPDTN's utility in practical applications where visual accuracy is paramount.

C. Dataset Attributes Impact Analysis

The author performed an exhaustive analysis to understand the influence of dataset characteristics on the NeRF model's performance. By altering attributes such as image resolution, the number of training images, and varied lighting conditions, the author gauged the model's sensitivity to data quality and quantity. This analysis provided insights into the minimum dataset requirements for achieving a high-fidelity plant digital twin, guiding users in data collection processes for effective FPDTN deployment.

D. IoT Data Reliability Test

The author collected a comprehensive dataset by Green-house, simulating the indoor growth conditions of

Guangzhou evergreen, recording parameters such as temperature, air humidity, soil moisture, CO_2 content, O_2 content, and hourly light intensity. A total of 745 data were collected for 30 consecutive days from 2023/5/13 to 2023/6/13. As shown in Fig. 7.

	DateTime	Temperature (°C)	Air Humidity (%)	Soil Humidity (%)	CO2 Content (ppm)	O2 Content (%)	Light Intensity (μmol/m²/s)
1	2023-05-13 00:00:00	19.49816047538945	86.98723719823705	64.75088261748655	518.6443140214768	29.7188306640741	4.643152810576971
2	2023-05-13 01:00:00	21.80285722563966	71.51673664119635	67.4650358325868	578.4604996279304	29.61748670249778	0.3090423034461316
3	2023-05-13 02:00:00	20.927975762456	76.30658583341966	64.54539254526303	437.0660359057428	29.85243550502169	0.4618026405568926
4	2023-05-13 03:00:00	20.39463393678814	87.19416332893641	61.46391847536887	415.7938324381657	29.90856071007792	0.8139549460311468
5	2023-05-13 04:00:00	18.62407456176974	78.72713987741976	72.06897186762755	447.9020339819076	29.77321515803892	2.976064170634374
6	2023-05-13 05:00:00	18.62397808134481	63.50694122125092	73.36425597110032	558.9156557400381	29.98351763728072	0.7523631550299874
7	2023-05-13 06:00:00	22.2323344486728	88.19496370840426	72.38980692019229	406.9340533429666	29.525834360828	195.2114572176100
8	2023-05-13 07:00:00	25.46470458309974	78.83124159214253	69.26880876011444	516.5613670166979	29.75239779915226	117.0071790887780
9	2023-05-13 08:00:00	24.40446004697283	70.04716843971258	67.5957156035846	599.0875032382	29.85922697366127	125.2276940525249
10	2023-05-13 09:00:00	24.83229031118418	64.17816217990162	77.26667299143651	571.1392187675078	29.93132023557459	86.99879237081097
11	2023-05-13 10:00:00	22.08233797718320	83.82075567810888	70.38163570272437	504.2891378087179	29.58962780595389	120.6440378652100
12	2023-05-13 11:00:00	25.87963940864798	78.60218267785541	69.58363755259462	412.7281818573285	29.90000174088184	149.2870448955517
13	2023-05-13 12:00:00	25.32977056320168	76.00383275928965	60.51284131612871	566.274702877142	29.77635353785927	162.6765705431064
14	2023-05-13 13:00:00	22.84935644271310	86.81677749152873	66.82495655252313	519.7957016354421	29.69827684099498	163.0802124064869
15	2023-05-13 14:00:00	22.72729986882840	83.65791633673592	67.60391237571714	422.9865994119133	29.56585751428954	136.6836690411324
16	2023-05-13 15:00:00	22.73618039413734	64.55024639198254	67.97645561793802	418.7714569646712	29.93264787945445	185.6539662490786
17	2023-05-13 16:00:00	23.21696897183815	69.35166203386645	71.60344738454664	581.9253638873184	29.57863660408986	172.6111770959164
18	2023-05-13 17:00:00	24.09902572652895	67.45467419443398	70.67205093453757	533.8400525876157	29.65489392960455	163.5572172047938
19	2023-05-13 18:00:00	23.72778007456846	82.31838877771803	72.15810185586889	565.8573598633511	29.64502765981217	58.34082825206909
20	2023-05-13 19:00:00	19.16491656079216	61.00597304207338	75.29766523084257	575.7957800564534	29.93570717095428	0.034389811098246

Fig. 7. The IoT Collected Data

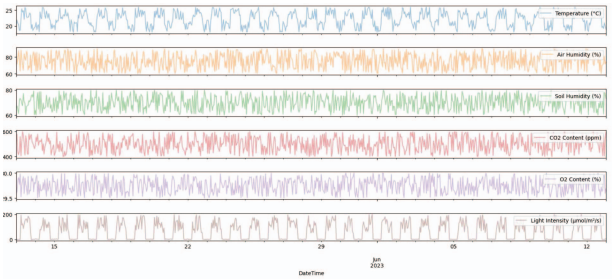


Fig. 8. The IoT Data Visualization

Reliability Testing Framework: To assess the reliability of IoT data, the author implemented a two-step approach: Temporal consistency check: the author compared hourly records over consecutive days to identify any sudden deviations beyond the expected natural variation. This test is designed to detect potential sensor failures or data transmission errors. Enhanced reliability with control group sensors to ensure the accuracy and reliability of IoT data, the author implemented a rigorous approach that includes multiple control group sensors and host sensor arrays. The method aims to mitigate any potential bias or error associated with the performance of individual sensors, thus providing a more robust dataset for analysis. A visualization is shown in Fig. 8. Specifically in this study, each environmental parameter (temperature, air humidity, soil moisture, CO_2 content, O_2 content, and light intensity) is monitored not only by a primary set of sensors, but also by two additional control sensors located in the same indoor environment. These control sensors were selected based on the proven accuracy and reliability in previous studies as a benchmark to validate the data collected by the main sensor. Results: Temporal consistency checks revealed a high degree of stability in the dataset, with indoor conditions fluctuating slightly within the expected range. These results show that the simulated IoT system can reliably capture daily and hourly changes in environmental conditions without significant anomalies. The integration of control group sensors significantly enhances the reliability of the IoT data. Inter-device reliability checks show a high degree of agreement between sensors with error bounds that are acceptable for scientific research. It is important to note that the few instances of significant deviation prompted an immediate recalibration of the affected sensors to ensure that the final dataset reflected accurate environmental conditions.

VI. Conclusion

This study successfully introduces the Fast Plant Digital Twin (FPDTN), an innovative approach that harnesses Neural Radiance Fields (NeRF) technology and IoT capabilities to produce digital twins of plants with unparalleled accuracy and efficiency. The findings illustrate FPDTN's effectiveness in addressing the scalability and financial limitations of traditional digital twins, underscoring its potential to enhance automated greenhouse operations by enhancing the monitoring and maintenance of plant health.

Despite these advancements, the author acknowledges the limitations of the current iteration of FPDTN. Specifically, further development is needed to accommodate a wider range of plant species and the intricacies of their complete life-cycle modeling. Additionally, optimizing the system's performance across various lighting conditions and enhancing real-time interactivity, particularly in mobile applications, are essential areas for improvement. Addressing these challenges through continuous iteration and incorporating user feedback will be crucial to ensure that FPDTN is robust, versatile, and capable of meeting diverse user needs.

Looking ahead, this work will focus on refining the user interface (UI) to enhance the usability of the application, particularly in capturing the required plant images for model generation via mobile devices. The author also aims to explore the integration of Artificial General Intelligence (AGI) with large general models, such as Large Language Models (LLMs), alongside NeRF. This exploration is expected to further enhance the convenience and intelligence of digital twin generation, targeting a broader range of plant groups and making FPDTN more accessible to users across various domains, including botany, horticulture, and education.

In conclusion, the development of FPDTN represents a significant advancement in the application of digital twin technology within the precision agriculture domain. As the author continues to refine and expand FPDTN's capabilities, it holds the promise of becoming an invaluable asset for agricultural professionals, botanists, gardening enthusiasts, and beyond. The ongoing evolution of FPDTN is poised to make substantial contributions to the fields of agriculture and botany, facilitating interactive science communication, assisting in the classification and cultivation of plant species, and ultimately advancing the digital twin technology landscape.

References

[1] F. Tao, H. Zhang, A. Liu, et al., “Digital twin in industry: State-of-the-art,” *IEEE Transactions on Industrial Informatics*, vol. 15, pp. 2405–2415, 2019.

[2] H. H. Hosamo, H. Nielsen, A. Alnmr, et al., “A review of the digital twin technology for fault detection in buildings,” in *Frontiers in Built Environment*, vol. 8, pp .1013196, 2022.

[3] F. Jiang, L. Ma, T. Broyd, et al., “Digital twin and its implementations in the civil engineering sector,” *Automation in Construction*, vol. 130, p. 103838, 2021.

[4] R. Sacks, I. K. Brilakis, E. Pikas, et al., “Construction with digital twin information systems,” *Data-Centric Engineering*, vol. 1, pp, e14, 2020.

[5] P. Augustine, “The industry use cases for the digital twin idea.” Advances in Computers, vol. 117, no. 1, pp. 79–105, 2020.

[6] D. A. Howard, Z. G. Ma, C. T. Veje, et al., "Greenhouse industry 4.0 – digital twin technology for commercial greenhouses," *Energy Informatics*, vol. 4, pp.1-3, 2021.

[7] F. Tao and M. Zhang, "Digital twin shop-floor: A new shop-floor paradigm towards smart manufacturing," *IEEE Access*, vol. 5, pp. 20418–20427, 2017.

[8] W. Yang, X. Bao, Y. ting Zheng, et al., "A digital twin framework for large comprehensive ports and a case study of qingdao port," *The International Journal of Advanced Manufacturing Technology*, pp.1-8 2022.

[9] A. Fuller, Z. Fan, and C. Day, "Digital twin: Enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020.

[10] B. Mildenhall, P. P. Srinivasan, M. Tancik, et al., "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communication of ACM*, vol. 65(1), pp.99-106, 2020.

[11] C. Mitsanis, W. Hurst, and B. Tekinerdogan, "A 3d functional plant modelling framework for agricultural digital twins," *Computers and Electronics in Agriculture*, vol. 218, pp. 108733, 2024.

[12] L. Gobeawan, D. J. Wise, S. T. Wong, et al., "Tree species modelling for digital twin cities," *Transactions on Computational Science XXXVIII*, pp. 17–35, 2021.

[13] K. Shimomoto, K. Takayama, N. Takahashi, et al., "Real-time monitoring of photosynthesis and transpiration of a fully-grown tomato plant in greenhouse," *Environment Control in Biology*, vol. 58(3), pp.65-70, 2020.

[14] A. Vaswani, N. M. Shazeer, N. Parmar, J. Uszkoreit, et al., "Attention is all you need," in *Neural Information Processing Systems*, vol. 30, 2017.

[15] J. T. Barron, B. Mildenhall, D. Verbin, et al.,"Mip-nerf 360: Unbounded anti-aliased neural radiance fields," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5460–5469, 2021.