2023 20th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON) | 979-8-3503-0052-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/SECON58729.2023.10287522

CAS: Crop Aerial Sensing Simulation in Smart Farming

Jingyu Liu
International Research Institute for AI
Harbin Institute of Technology, Shenzhen
Shenzhen, China

Yang Zhao
International Research Institute for AI
Harbin Institute of Technology, Shenzhen
Shenzhen, China
yang.zhao@hit.edu.cn

Xinrui Xiao
International Research Institute for AI
Harbin Institute of Technology, Shenzhen
Shenzhen, China

Ran Meng
College of Resources and Environment
Huazhong Agricultural University
Wuhan, China

Jie Liu
International Research Institute for AI
Harbin Institute of Technology, Shenzhen
Shenzhen, China

***Abstract*—Unmanned aerial vehicles (UAV) with onboard sensors become a cost-effective way of crop remote sensing in large-scale farms. However, current vision-based crop aerial sensing methods suffer from occlusion issue and require large amount of annotation data. In this paper, we target one particular crop species, corn, and propose to use 3D modeling and simulation to help resolve the issues. We first develop a corn-field 3D model and a crop aerial sensing (CAS) simulation framework. Then we use the CAS framework to generate synthetic data to train various deep learning models for corn leaf segmentation. In addition, we change the 3D model parameters in CAS, e.g., distances between individual corn plants, to derive leaf area index (LAI) correction coefficients for various corn plant and row spacings. Our experimental results from real-world UAV images show that our leaf segmentation model using synthetic data from the CAS framework outperforms state-of-the-art segmentation models by 1.4-3.3%. Our simulation results show that the plant and row spacings of a corn-field have significant effects on correcting the UAV image-based LAI, which can be underestimated by a factor of 2.6, due to the overlap and occlusion issues.**

***Index Terms*—Remote Sensing, Smart Farming, Unmanned Aerial Vehicle, Image Segmentation**

I. Introduction

As the world population increases, more food needs to be produced with limited expansion of cultivable land. Smart farming, also known as Digital Agriculture, is the Agriculture 4.0 vision, in which farmers will use the minimum water, fertilizers and pesticides to increase crop yield with the help of digital technologies. The digital technologies being used in smart farming include artificial intelligence (AI), internet of things, robotics, etc. Among them, aerial sensing, *i.e.*, crop monitoring with unmanned aerial vehicle (UAV) sensors, becomes a cost-efficient way of data acquisition in large-scale crop fields. For example, crop leaf area index (LAI) can be estimated from RGB images taken by cameras onboard UAVs [1]. Defoliation can be detected by deep learning models trained from tens of thousands of UAV images collected from soybean fields [2]. A multi-agent reinforcement learning algorithm is recently proposed to schedule and manage UAV swarms to conduct crop monitoring missions [3].

In the aforementioned applications, UAV serves as a low-cost, fast and maneuverable sensor that monitors large-scale crop in an automatic and non-invasive way. However, these applications still have the following challenges. First, the performance of the AI algorithms heavily rely on manually-annotated training dataset, which can be time-consuming to obtain, especially for semantic segmentation tasks [4]. Second, it is difficult to "see-through" multi-layer and overlapping crops using optical cameras onboard UAV, due to the occlusion issue [1]. In this work, we look into the digital twin technology and develop computer simulation to help resolve the above issues in smart farm aerial sensing.

Specifically, for the data annotation issue, previous studies have used synthetic dataset and data augmentation methods to generate large amount of training data for semantic segmentation and other tasks [4]–[7]. For example, data augmentation preserving the photo-realistic appearance of plant leaves has been developed to achieve state-of-the-art results on leaf segmentation of rosette plants in the Leaf Segmentation Challenge hosted by Computer Vision Problems in Plant Phenotyping [5]. Manually-annotated field images and synthetically-generated images that mimic the distribution of real images have been used in segmentation of weeds and corn crop [6]. However, none of the studies above have targeted the aerial imagery, where challenges at altitude and range cannot be ignored [8]. The recent study in [6] set up two cameras at two different height levels, 2m and 1.6m, to collect corn images, but these height levels are not even close to the heights when UAVs collect field crop images. In this paper, we focus on aerial sensing for the corn (Zea mays) crop species. We create a corn-field 3D model (example shown in Figure 1), and develop a crop aerial sensing (CAS) simulation framework. We use CAS and data augmentation method to generate synthetic data that are specifically targeted on aerial images collected from high altitude.

Fig. 1. Corn-field 3D model (annotations of corn leaves are highlighted in blue in the bottom-right subgraph).

Second, for the occlusion issue in crop sensing, a recent study [9] uses the stereo vision technique and develops a depth discontinuity segmentation algorithm to separate the overlapped leaves based on a depth discontinuity criteria. This approach can significantly improve leaf segmentation performance. However, the working range of the depth cameras and the stereo vision technique is limited compared to RGB cameras, and thus has limited potentials in large-scale aerial sensing. While other sensors, such as synthetic aperture radar (SAR) can be used to overcome the line-of-sight issue of the optical sensors [10], in this work, we follow an alternative approach, without investigating new sensors and hardware. We propose to use computer simulation to find a calibration approach, i.e., a correction method for crop aerial sensing. Taking the LAI estimation for example, we can use computer simulation to find a mapping between the LAI ground-truth and the LAI derived from UAV images in various crop planting conditions, *e.g.*, various row and plant spacings. Then, an LAI correction coefficient can be used in a real-world crop aerial sensing mission. We use CAS to generate such a set of correction coefficients for various crop spacings in this paper.

The contributions of this paper are summarized as follows:

- We build a corn-field 3D model, and a crop aerial sensing (CAS) simulation framework, which simulates the UAV data acquisition of crop images at various conditions.
- We develop two use cases for the CAS framework: 1) generation of synthetic data for leaf semantic segmentation, 2) calculation of LAI correction coefficient.
- Our evaluation shows that the synthetic data from the CAS framework improves the accuracy of the state-of-the-art segmentation models by 1.4-3.3%. Our simulation shows that the LAI derived from UAV images can be underestimated by a factor of 2.6, due to the leaf overlap and occlusion issues. We plan to make the CAS framework and models publicly available upon paper acceptance.

II. Methods

A. Corn-field 3D Model Creation

Blender is an open-source 3D simulation software, which enables us to develop our own CAS framework. Previous work in [11] builds an open-source grid-based render farm in Blender, showing its ability in crop simulation.

To build a Blender-based simulation framework on crop sensing, we first use Blender to control the texture of the corn objects, since we need to change the original corn 3D model texture for leaf annotation. Texture can be seen as a regular and repeatable image, which is the starting point of color rendering and also the part that people can control most in the process. As a renderer, shader allows us to adjust various parameters, such as the basic tone, reflectivity, roughness of the model. For complex models such as the corn model, texture is made into multiple types of images, which are used as inputs for several parameters in shader respectively, so that they can be rendered correctly. Material is the result of this rendering process, that is, shader controls the blending of multiple textures to form materials with specific visual effects. Materials are independent and reusable, which can be assigned to any object to form objects with rich color details that people see in daily life.

After the object texture step, we generate multiple corn plants and arrange them to mimic real situation. This can be achieved by duplicating plant models and moving them precisely. We first duplicate the corn model into multiple plants. Then, for these duplicated plants, we can control their position on X-axis and Y-axis in the Blender environment. After these steps, we build a corn-field 3D model, as shown in Figure 1.

B. Crop Aerial Sensing Simulation

1) Analysis of Real-world Corn Data Collection: In order to produce high-fidelity synthetic image data, *i.e.*, make the synthetic data as close as possible to the data collected by experts in the real environment, we conducted a systematic analysis of the scene, where UAVs collect RGB images in an experimental cornfield. We obtained a large number of cornfield images from agricultural experts, which are RGB images captured by UAV flying over a cornfield. Those RGB images are orthorectified during the post processing. Then we quantitatively analyzed the pixels-per-meter(ppm) index of the orthophoto cornfield images and summarize the behavior of the UAV and corresponding camera when conducting data collection.

Based on the provided data, it is evident that in our data collection scenario, the pilot will maneuver the unmanned aerial vehicle to a height ranging from 30 to 50 meters above the ground. Then the image sensor will be adjusted to be parallel to the ground, enabling a top-down capture of cornfield images. Moreover, We know that the spacing of maize plants ranges from 0.3 m to 0.6 m according to the image scale. All these information helps us to build a realistic cornfield and conduct a realistic synthetic data collection in the following work.

2) Collection Configurations: Drawing upon the simulation methodology elucidated in Section II-A and accounting for the density of the cornfield, we constructed a realistic cornfield environment model along with a corresponding labeled model within our simulation environment, Blender. The labeled cornfield model is built based on the realistic cornfield model, but we replaced the texture of objects in the realistic model with meaningful flat colors. We use three RGB flat colors to distinguish different elements: corn leaf (255,255,255), sky (0,255,0), and ground and others (0,0,0).

Following the establishment of the environment models, we positioned our sensors within them and rendered images at identical locations, utilizing the same camera settings. This process facilitated the generation of pixel-to-pixel annotated images depicting the cornfield. To ensure an ample and diverse synthetic dataset for training the leaf segmentation model, we devised four different configurations for collecting synthetic cornfield dataset from various perspectives.

1) Config 1: Simulate the UAV collecting cornfield orthophoto image data in cornfield. We simulated the camera 30m above the cornfield and flew to the collection points shown in Figure 2 for synthetic data collection. The distance between each collection point is set to be 2 meters, resulting in a total of 25 collection points. The sensor simulation is configured in the orthophoto mode, with a collection range spanning 4 meters, and an image resolution of 480x320 pixels.
2) Config 2: Simulate the UAV collecting multi-angle synthetic data. In this configuration, the camera focuses on a specific corn plant, and the UAV carrying the camera will rotate to collect non-orthophoto synthetic data. The UAV rotates two circles around the center plant in the cornfield, collecting a total of 100 synthetic images.
3) Config 3: Enrich the local information of synthetic data by collecting partial feature images of corn leaves of a single corn plant. The camera spirals up around the single plant as the axis to collect local features of corn plants.
4) Config 4: Enrich the overall features of a single corn plant by collecting images of the plant from different angles. The camera spirals up around a single plant as the axis at a relatively far distance to collect overall features of corn plants.

C. Leaf Segmentation

Corn leaf image segmentation is an important technique for smart farming applications, such as pest detection [12], unmanned aerial vehicle spraying [13], and precise fertilizer calculation [14]. In this section, we discuss the segmentation baseline model and the segmentation models, on which we use synthetic dataset for training.

1) Segmentation Model: Due to the synthetic data contains corn leaf images of different dimensions and accuracy, we select the U-Net++ [15] model as the base model for corn leaf image segmentation. U-Net++ is an improved convolutional neural network (CNN) structure that enhances the generalization ability of image segmentation networks by extending and optimizing the original U-Net [16] encoder-decoder structure.

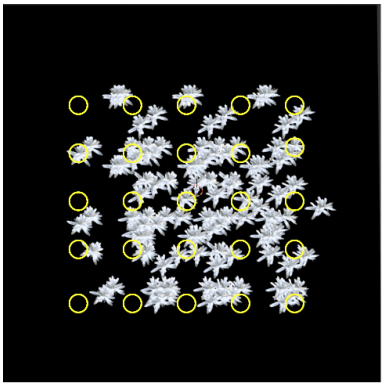

Fig. 2. Way-points (circles) of UAV for collecting aerial images in the CAS framework.

The leaf image segmentation task at hand is a binary classification task, in which the pixels in the image are classified as either corn leaves or non-corn leaves. For this problem, we have chosen to use the Intersection over Union (IoU) index to evaluate the effectiveness of the leaf segmentation model. This index represents the degree of overlap between the predicted segmentation result and the actual segmentation result, and is calculated by dividing the intersection area of the predicted and actual segmentation results by their union area. The value of the IoU index ranges from 0 to 1, with a higher value indicating better segmentation results.

2) Pre-process and Training: We implemented the U-Net++ model using PyTorch. Then we using the four different annotated datasets generated above and real-world human-labeled dataset to train the leaf segmentation model, following the standard pre-process scheme in semantic segmentation, such as cropping, color normalization and image augmentation. In pre-process stage, all input images are formalized to (480, 420) RGB image, and the mask are formalized to (480, 320) gray-scale images. Then we performed color normalization with the means and variation of ImageNet [17]. In the end, we apply image augmentation to improve data diversity. Then we tested 6 different segmentation models on this mission:

1) **U-Net++ Manual**: a segmentation model based on a U-Net++ model pre-trained on ImageNet, then trained on over one hundred manually labeled (before image data augmentation) real-world top-down cornfield images captured by UAV.
2) **U-Net++ Synthetic1**: a segmentation model based on a U-Net++ model pre-trained on ImageNet, and trained on Synthetic dataset from Config 3 and Config 4.
3) **U-Net++ Synthetic2**: a segmentation model based on a U-Net++ model pre-trained on ImageNet, and trained on three synthetic dataset, Config 2, 3, and 4.
4) **U-Net++ Synthetic3**:a segmentation model based on a U-Net++ model pre-trained on ImageNet, and trained on

all synthetic dataset we generated in the four configurations.

5) **U-Net++ Synthetic + Manual**: a segmentation model based on model U-Net++ Synthetic3, then trained on over one hundred manually labeled real-world top-down cornfield images.
6) **SAM**: the SOTA general segmentation model Segment Anything Model (SAM), which is a large-scale image segmentation model from Meta AI Research [18].

After pre-process stage, we converted the RGB raw images to (3, 480, 320) tensors, and converted the input gray-scale mask to (1, 480, 320) images. The size of output tensor is (1, 480, 320), which classify the corresponding pixels to either leaf or not-leaf. We used the pre-trained Resnext 50x24d model on ImageNet dataset as the encoder of our U-Net++ model. We used Adam optimizer and Dice Loss to train our model. Then we trained our model on the four annotated dataset we generated above in four different rounds with learning rate at 0.00008 (0.00001 after 15 epochs). Each round has 15 to 20 training epochs.

D. LAI Correction

As mentioned before, the LAI derived from UAV images are typically underestimated due to the leaf overlap and occlusion issues. To make a correction to LAI, we define an LAI correction coefficient φ as the ratio of the LAI ground-truth LAI_{True} to the LAI derived from UAV images LAI_{UAV}:

$$\varphi = \frac{LAI_{True}}{LAI_{UAV}}. \tag{1}$$

We explain how we calculate LAI_{True} and LAI_{UAV} next.

1) LAI Ground-truth: First, for calculating the LAI ground-truth, we use a built-in plugin Measureit in Blender, which can output the total area of the planes selected by users. To make the plane selection process easier, we can combine some small planes to minimize the number of planes we need to manually select. Some of the combined leaf can be seen from the highlighted blue area in Figure 1. After several rounds of combination, a single leaf will get few planes left. In this way, we can use the plug-in to read and calculate the true leaf area in total, and this leaf area is one part of the LAI calculation. For the land area part, we can take the circumscribed quadrilateral of 16 corn plants to calculate its area, approximately near to a square with side length of 4.8 meters. In this way, we can simply get the theoretical LAI ground-truth value from Blender.

2) LAI Derived from UAV Images: The next step is to obtain the LAI from UAV images. As mentioned in Section II-B, we use the camera plugin in Blender to take images at various waypoint locations for our simulated corn-field with high fidelity. Since too many color and texture levels of the 3D model bring difficulty in leaf segmentation, we first make the 3D model monochromatic before the UAV image acquisition step.

After that, we use semantic segmentation models, e.g., the SAM segmentation model, to performance leaf segmentation.

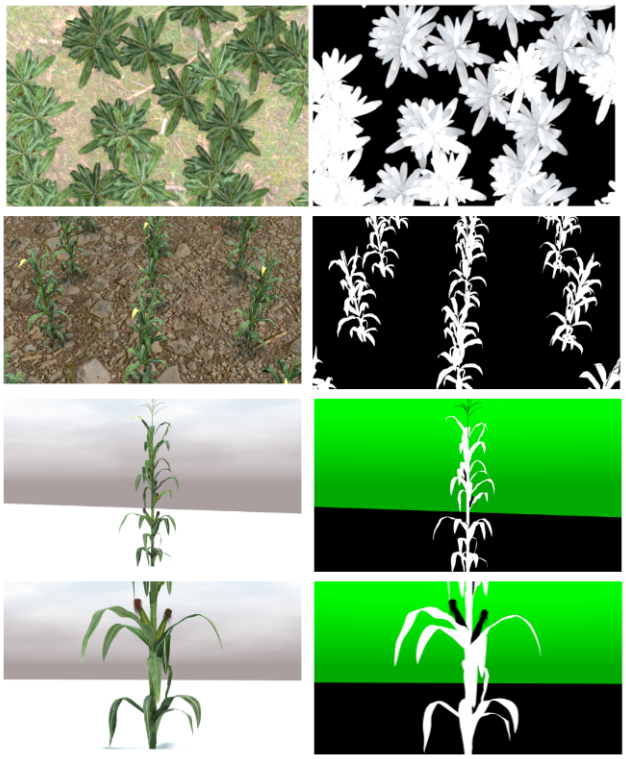

Fig. 3. Four pairs of raw (left) and annotated (right) synthetic images of cornfield.

Fig. 4. Top-down cornfield image(left) captured by UAV and the corresponding mask(right) labeled by human.

Finally, we can count the number of pixels that leaves occupy in the image.

III. Experiments and Results

A. Leaf Segmentation Results

1) Synthetic Dataset: Figure 3 shows a couple of realistic and flat-color images of the corn leaves, as they are captured in different aspect and resolution. Each pair is corresponded to a collection configuration in Section II-B2. The four pairs of images from top to bottom are the examples of synthetic dataset captured from Config 1 to Config 4. In this way, we generate a batch of annotated synthetic data of corn leaf, an exact correspondence at the pixel level has been achieved.

2) Test Results on Real-world Data: To demonstrate the efficacy of the labeled synthetic dataset of cornfield, We compare the segmentation result among the 6 different models

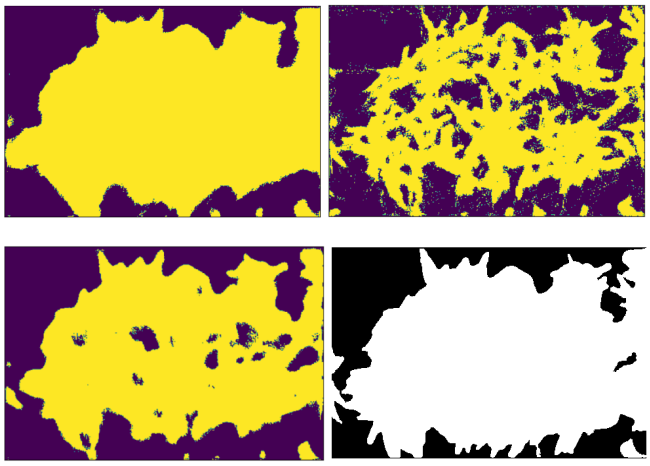

Fig. 5. Segmentation results for Figure 4 of U-Net++ Manual(top-left), U-Net++ Synthetic3(top-right), U-Net++ Synthetic+Manual(down-left) and SAM(down-right).

TABLE I
IoU FROM THREE SEGMENTATION MODELS.

Model	Average IoU	Max IoU	Min IoU
U-Net++ Manual	0.6535	0.8903	0.4531
U-Net++ Synthetic-1	0.4669	0.8256	0.1531
U-Net++ Synthetic-2	0.4964	0.8579	0.1863
U-Net++ Synthetic-3	0.5887	0.8360	0.2460
SAM	0.6343	0.9115	0.4033
U-Net++ Synthetic+Manual	0.6675	0.8511	0.4706

using 16 human-labeled top-down cornfield images collected by UAV (Figure 4).

Table I shows the IoU indices for corn leaf segmentation, and Figure 5 demonstrates the prediction result of the test image showed on Figure 4. Significantly, considering the lower-bound IoU index achieved by humans, which stands at 0.64, models *U-Net++ Manual, U-Net++ Synthetic+Manual*, and *SAM* exhibit human-level segmentation results, scoring 0.6535, 0.6675, and 0.6343, respectively. Our best model(U-Net++ Synthetic+Manual) exhibits a notable 3.3% increase in the IoU index compared to SAM. Furthermore, it outperforms a segmentation model trained exclusively on a human-labeled dataset by 1.4%. It is not surprising that model *U-Net++ Synthetic+Manual* performs the best, given its training on both synthetic dataset and real-world dataset, whereas *U-Net++ Synthetic3* solely relies on a synthetic dataset and model *U-Net++ Manual* solely relied on real-world dataset. As a versatile segmentation model, *SAM* achieves a commendable performance in corn leaf segmentation, only slightly trailing behind *U-Net++ Manual*.

The segmentation performance of the U-Net++ Synthetic3 model, which is trained exclusively on synthetic dataset, indicates the potential to develop a segmentation model based solely on synthetic data, thereby alleviating the laborious task of manual labeling. Note that our U-Net++ model has a compact size of 196 MB, whereas the SAM (SAM-1b) model size is 2.7 GB. Our lightweight model is more suitable for deployment on resource-limited IoT devices.

TABLE II
LAI CORRECTION COEFFICIENT φ FOR VARIOUS ROW SPACING AND PLANT SPACING VALUES.

Case no.	Plant spacing(m)	Row Spacing(m)	φ
1	1.2	0.381	1.876
2	1.2	0.508	1.605
3	1.2	0.635	1.344
4	1.2	0.762	1.202
5	0.381	1.0	2.611
6	0.508	1.0	2.217
7	0.635	1.0	1.894
8	0.762	1.0	1.683

B. LAI Correction Results

1) Row and plant spacing settings for LAI Correction: In the CAS framework, we can adjust the row spacing and plant spacing for our corn-field model. Since previous studies show that row spacing affects crop yield [19], in this paper, we perform experimental analysis for the effect of row spacing on the LAI correction coefficient. We fix the plant spacing as 1.2m, and analyze the change of the LAI coefficient as the row spacing is set to be 0.381m, 0.508m, 0.635m, and 0.762m.

As mentioned before, to calculate the LAI correction coefficient φ, we need to figure out three values: the leaf area ground-truth, the number of pixel that leaves and the reference plane take, and the area that the reference plane actually takes. The leaf area ground-truth can be counted by using the measureit plug-in in Blender, and the value is 16.74 square meters. Let's take row spacing of 0.762m as an example. From the image taken by the UAV camera, we can calculate that the corn leaves occupy 572,726 pixels, and the reference plane takes 59,205 pixels. The last thing we need to know is the area of our reference plane, and this can be easily known from the measureit plug-in, read as 1.44 square meters. With these data and the equivalent relationship of two ratios, we can calculate the actual area that corn leaves take, which is 13.93 square meters. Finally, based on the assumption we make, we can calculate the LAI coefficient of this simulation to be 1.202.

As for row spacing of 0.762m and plant spacing of 1.2m, we can see the plant arrangement like the following picture shows. Using the SAM model, we can figure out that the pixel which leaves take is about 572,614. Based on this and our reference plane, we can calculate that the area that leaves take is 13.93 square meters. As we mentioned above, we have already counted the leaf area ground-truth; so the corresponding LAI coefficient can be calculated to be 0.832. With the same step, we can also figure out the LAI coefficient in row spacing of 0.635m, 0.508m and 0.381m (listed in Table II).

In the analysis mentioned above, we fix the plant spacing while changing the row spacing, in which way we calculate the corresponding LAI coefficient at certain row spacing. Now we will change the plant spacing when fixing the row spacing. We choose to fix the row spacing to be 1 meter. As for the corresponding plant spacing, we still make use of the four values as before, which is 0.762m, 0.635m, 0.508m and

0.381m. The result are shown in the case 5 to 8 of Table II. From Table II, we see that the LAI can be underestimated by up to a factor of 2.6, for the 1.2m plant spacing, due to the leaf overlap and occlusion issues.

IV. Related work

Due to the "data hungry" nature of deep learning methods, more and more synthetic data are generated and used in data-driven model training. 3D modeling software such as Blender, Unity are being used for the creation of realistic computer graphic models. In this paper, we purchase a 3D model of a single corn plant, and use Blender to create our own corn field model. We also tested other simulation platforms such as Gazebo [20]. We find Blender can provide more photo-realistic crop models than the Gazebo simulator. Thus, we build our simulation framework in Blender.

The idea of using computer simulation for sensing and learning purposes was adopted in various domains, such as self-driving, home-care robotics, etc. For example, the CARLA simulator was an open-source simulation platform for studying the performance of various autonomous driving methods and generating synthetic data for AI model training and validation [21]. RCareWorld, a human-centric simulation platform for physical and social robots was developed for simulating care-giving scenarios, and provides the capability to plan, control, and learn both human and robot control policies [22]. Our work focuses on the smart farming scenario, in which we use a 3D corn-field model and simulation to avoid tedious and manual work of image ground-truth labeling. The corn-field model also gives us the flexibility to investigate the effects of plant spacing on LAI coefficient correction.

V. Conclusions

In this paper, we develop a crop aerial sensing simulation (CAS) framework using open-source 3D modeling software. We use the CAS framework to generate synthetic data to train deep learning models for corn leaf segmentation. We also use CAS to derive relationships between LAI correction coefficient and corn plant and row spacings at a particular corn growth stage. Our experimental results show that the segmentation model trained by synthetic and manual annotations outperforms state-of-the-art segmentation models. Our simulation results show that LAI can be significantly underestimated due to the overlap and occlusion issues.

References

[1] L. Roth, H. Aasen, A. Walter, and F. Liebisch, "Extracting leaf area index using viewing geometry effects—a new perspective on high-resolution unmanned aerial system photography," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 141, pp. 161–175, 2018.

[2] Z. Zhang, S. Khanal, A. Raudenbush, K. Tilmon, and C. Stewart, "Assessing the efficacy of machine learning techniques to characterize soybean defoliation from unmanned aerial vehicles," *Computers and Electronics in Agriculture*, vol. 193, p. 106682, 2022.

[3] J. Boubin, C. Burley, P. Han, B. Li, B. Porter, and C. Stewart, "Marble: Multi-agent reinforcement learning at the edge for digital agriculture," in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*. IEEE, 2022, pp. 68–81.

[4] L. M. Tassis, J. E. T. de Souza, and R. A. Krohling, "A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images," *Computers and Electronics in Agriculture*, vol. 186, p. 106191, 2021.

[5] D. Kuznichov, A. Zvirin, Y. Honen, and R. Kimmel, "Data augmentation for leaf segmentation and counting tasks in rosette plants," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2019, pp. 0–0.

[6] A. Picon, M. G. San-Emeterio, A. Bereciartua-Perez, C. Klukas, T. Eggers, and R. Navarra-Mestre, "Deep learning-based segmentation of multiple species of weeds and corn crop using synthetic and real image datasets," *Computers and Electronics in Agriculture*, vol. 194, p. 106719, 2022.

[7] D. Ward, P. Moghadam, and N. Hudson, "Deep leaf segmentation using synthetic data," *arXiv preprint arXiv:1807.10931*, 2018.

[8] D. Cornett, J. Brogan, N. Barber, D. Aykac, S. Baird, N. Burchfield, C. Dukes, A. Duncan, R. Ferrell, J. Goddard *et al.*, "Expanding accurate person recognition to new altitudes and ranges: The briar dataset," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 593–602.

[9] Z. M. Amean, T. Low, and N. Hancock, "Automatic leaf segmentation and overlapping leaf separation using stereo vision," *Array*, vol. 12, p. 100099, 2021.

[10] S. Fang and S. Nirjon, "Superrf: Enhanced 3d rf representation using stationary low-cost mmwave radar," in *International Conference on Embedded Wireless Systems and Networks (EWSN)...*, vol. 2020. NIH Public Access, 2020, p. 120.

[11] M. Z. Patoli, M. Gkion, A. Al-Barakati, W. Zhang, P. Newbury, and M. White, "An open source grid based render farm for blender 3d," in *2009 IEEE/PES Power Systems Conference and Exposition*. IEEE, 2009, pp. 1–6.

[12] K. P. Panigrahi, H. Das, A. K. Sahoo, and S. C. Moharana, "Maize leaf disease detection and classification using machine learning algorithms," in *Progress in Computing, Analytics and Networking: Proceedings of ICCAN 2019*. Springer, 2020, pp. 659–669.

[13] A. Hafeez, M. A. Husain, S. Singh, A. Chauhan, M. T. Khan, N. Kumar, A. Chauhan, and S. Soni, "Implementation of drone technology for farm monitoring & pesticide spraying: A review," *Information processing in Agriculture*, 2022.

[14] K. Manninen, S. Koskela, A. Nuppunen, J. Sorvari, O. Nevalainen, and S. Siitonen, "The applicability of the renewable energy directive calculation to assess the sustainability of biogas production," *Energy Policy*, vol. 56, pp. 549–557, 2013.

[15] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation," *IEEE transactions on medical imaging*, vol. 39, no. 6, pp. 1856–1867, 2019.

[16] O. Ronneberger, P. Fischer, and T. Brox, "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, 2015, pp. 234–241.

[17] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[18] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo *et al.*, "Segment anything," *arXiv preprint arXiv:2304.02643*, 2023.

[19] G. Á. Maddonni and J. Martínez-Bercovich, "Row spacing, landscape position, and maize grain yield," *International Journal of Agronomy*, vol. 2014, 2014.

[20] J. Liu, X. Xiao, Y. Zhao, and J. Liu, "Containerized mobile sensing simulation framework for smart agriculture," in *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, 2022, pp. 766–767.

[21] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.

[22] R. Ye, W. Xu, H. Fu, R. K. Jenamani, V. Nguyen, C. Lu, K. Dimitropoulou, and T. Bhattacharjee, "Rcare world: A human-centric simulation world for caregiving robots," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 33–40.