

Towards Autonomous Crop Monitoring: Inserting Sensors in Cluttered Environments

Moonyoung Lee , Aaron Berger , Dominic Guri , Kevin Zhang , Lisa Coffey , George Kantor ,
and Oliver Kroemer 

Abstract—Monitoring crop nutrients can aid farmers in optimizing fertilizer use. Many existing robots rely on vision-based phenotyping, however, which can only indirectly estimate nutrient deficiencies once crops have undergone visible color changes. We present a contact-based phenotyping robot platform that can directly insert nitrate sensors into cornstalks to proactively monitor macronutrient levels in crops. This task is challenging because inserting such sensors requires sub-centimeter precision in an environment which contains high levels of clutter, lighting variation, and occlusion. To address these challenges, we develop a robust perception-action pipeline to grasp stalks, and create a custom robot gripper which mechanically aligns the sensor before inserting it into the stalk. Through experimental validation on 48 unique stalks in a cornfield in Iowa, we demonstrate our platform’s capability of detecting a stalk with 94% success, grasping a stalk with 90% success, and inserting a sensor with 60% success. In addition to developing an autonomous phenotyping research platform, we share key insights obtained from deployment in the field.

Index Terms—Agricultural automation, field robots, grippers and other end-effectors.

I. INTRODUCTION

WITH advancements in AI and robotics, the agricultural sector is well-positioned to adopt precision agriculture methods that can enhance crop production and minimize environmental footprint [1]. For example, one of the predominant issues faced by farmers is the overuse of fertilizers, which can be alleviated with increased sensing accuracy [2], [3]. Many existing research works focus on automating vision-based crop monitoring and phenotyping [4], [5], [6], [7], [8], [9], specifically for detecting cornstalks [10], [11]. However, visible

Manuscript received 5 November 2023; accepted 23 March 2024. Date of publication 8 April 2024; date of current version 22 April 2024. This letter was recommended for publication by Associate Editor Z. Wang and Editor H. Moon upon evaluation of the reviewers’ comments. This work was supported in part by the NSF Robust Intelligence under Grant 1956163 and in part by the NSF/USDA NIFA AIIRA AI Research Institute under Grant 2021-67021-35329. (Corresponding author: Moonyoung Lee.)

Moonyoung Lee, Dominic Guri, Kevin Zhang, George Kantor, and Oliver Kroemer are with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15206 USA (e-mail: moonyoul@cs.cmu.edu; dguri@cs.cmu.edu; klzl@cs.cmu.edu; kantor@cs.cmu.edu; okroemer@cs.cmu.edu).

Aaron Berger is with the Department of Computer Science, Harvard University, Cambridge, MA 02138 USA (e-mail: aaronberger@college.harvard.edu).

Lisa Coffey is with the Department of Agronomy, Iowa State University, Ames, IA 50011 USA (e-mail: lmcoffey@iastate.edu).

Our research platform is open-sourced, with additional information available at <https://kantor-lab.github.io/cornbot>.

This letter has supplementary downloadable material available at <https://doi.org/10.1109/LRA.2024.3386463>, provided by the authors.

Digital Object Identifier 10.1109/LRA.2024.3386463



Fig. 1. Robot inserting sensors into cornstalks to monitor plant nitrate concentration in Curtiss Farm, Iowa. Our custom gripper utilizes contacts in a clutter-rich environment to precisely align with stalks and insert sensors.

plant features are a lagging indicator of nutrient deficiencies. Early detection of these deficiencies enables farmers to respond sooner, enhancing harvest yield. Consequently, our research emphasizes contact-based phenotyping. This involves inserting nitrate sensors developed at Iowa State University [12], [13] into cornstalks to monitor nitrogen concentration, allowing agronomists to proactively address crop deficiencies.

While less abundant than vision-based phenotyping, there are several studies that measure plant traits using contact-based sensors. Examples include robots with custom leaf-grasping grippers which measure chlorophyll content with a SPAD meter [14] or spectral reflectance with a spectrometer [15]. Sensor penetration through leaves has also been demonstrated in [16], [17], which use a needle-like fluorometer to detect chlorophyll fluorescence emission. In addition to interacting with leaves, there have been efforts to grasp cornstalks to measure diameter [18]. Our work is most similar to contact-based phenotyping methods deployed on sorghum plants, where a force-gauge sensor is probed into a stalk to measure strength [19], [20]. However, unlike reusing a single probe for multiple measurements, our task involves deploying sensors in multiple plants, which assists

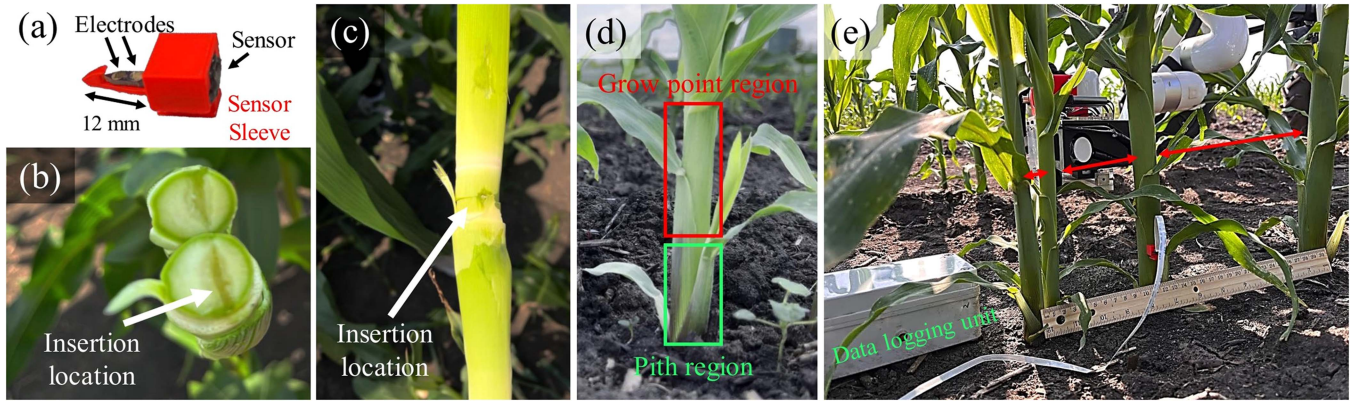


Fig. 2. Overview of the task: (a) Sensor to convert nitrate to voltage and the sensor sleeve for protection when penetrating into stalks; (b) Cross-section view of a stalk after a successful insertion; (c) View of the insertion location after a deployed sensor has been manually removed; (d) View of a stalk showing the pith region, a solid node near the root, and the grow-point region composed of layers of leaves; (e) Cluttered field environment with varying stalk spacing, occlusion from leaves, and uneven ground terrain. The deployed sensor is wired to the data logging unit to store data over the plant's life cycle.

a farmer in taking action to address deficiencies in a large area of a field.

Manually installing these nitrate sensors with corresponding data logging units across a farm is a laborious process that is difficult to scale [21], [22]. Thus, we present a robotic research platform capable of autonomously inserting nitrate sensors and deploying corresponding data logging units in the field. We believe our presented work can benefit the agricultural robotics community by identifying key challenges in deploying a robot for contact-based phenotyping. By sharing our design decisions and valuable insights obtained in the field, we hope to contribute to building a fully autonomous phenotyping research platform that can support farmers.

To summarize, the contributions of this paper are:

- An open-sourced platform capable of autonomously deploying nitrate sensors into stalks
- Evaluation of our robotic platform on 48 sensor insertion trials at Curtiss Farm, Iowa
- A dataset of 7600 cornstalks for segmentation tasks

II. CHALLENGES OF SENSOR INSERTION IN CORN

In contrast to robotics experiments conducted in a controlled indoor lab environment, agricultural robotics is challenging due to extreme variations in field conditions and limited accessibility to test a system due to the narrow time frame of seasonal crops. For the task of cornstalk sensor insertion, we specifically discuss the challenges that arise in near-ground precise manipulation tasks and the perception difficulties when operating in a cluttered, outdoor environment.

A. Contact Interaction Challenges

One of the main challenges is the sub-centimeter precision required to insert a small sensor (whose probe is 12x3x2 mm in dimension, shown in Fig. 2(a) for accurate data logging. The stalk diameter and probe length are of similar magnitude, so it is important to precisely insert sensors near the vertical center of the stalk. If the sensor is instead inserted near the edge, the insertion may not be deep enough to fully immerse the sensor's two electrode pads into the stalk, leading to faulty

sensor readings. Examples of successful insertions are shown in Fig. 2(b), (c), (e).

The variation of stalk shape and diameter observed in the field make this required precision difficult. For example, while some stalks have circular cross-sections, a large portion have elliptical cross-sections, making the direction of insertion important. For the purpose of monitoring crop nitrate concentration, it is ideal to insert sensors in corn plants in the V4-V8 stages of the plant's life cycle [13]. The stage of a corn plant is measured by the number of leaf collars: V4-V8 plants have four to eight leaf collars, a stalk diameter between 12–40 mm, and are typically grown in 14–28 days.

An additional requirement for the sensor to properly monitor nitrate concentration is to be placed in the pith region of the stalk, referred as the first node typically 2–10 cm above the ground (Fig. 2(d)). The pith region does not structurally change over the plant's growth cycle, allowing for consistent measurement from an embedded sensor. In the grow-point region, however, the stalk contains layers of leaves that continuously grow over the life cycle and thus force out the embedded sensor from within the stalk. This requirement to operate in close proximity to the soil poses a challenge: the terrain of a cornfield is highly uneven, and ditches and mounds may collide with the robot arm.

Another challenge is the high variation in corn plant spacing, as shown in Fig. 2(e). Even when corn seeds are planted with a predetermined density on a seeder machine, tillers (reproductive shoots growing upward from the same plant) and other growing conditions lead to narrower or wider gaps between stalks.

B. Perception Challenges

These described variations in stalk shape and spacing also make perception difficult, as heuristics like expected stalk spacing and maturity are inaccurate. Varying lighting conditions depending on time of day also pose a challenge: as the robot traverses through corn plant rows, the camera switches between facing east and west. In the morning and evening, these viewpoints introduce problems with sun flare or severe overcast shadows from the stalks and the robot itself. Another challenge arises from the frequent occlusion of stalks from nearby leaves, as shown in Fig. 1. Lastly, targeting the near-ground pith region is difficult as highly uneven terrain make vision-based

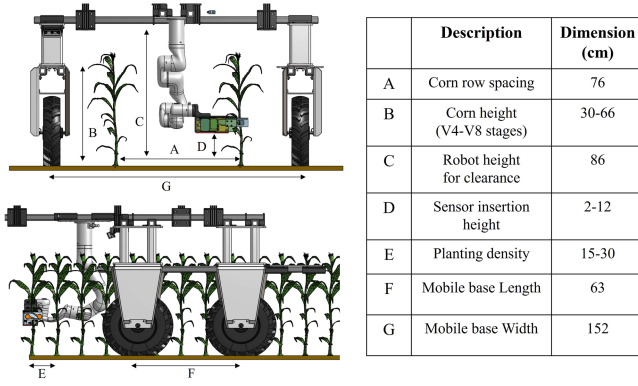


Fig. 3. Cornfield specification and key dimensions of the robot platform.

height predictions inaccurate—ground plane detection methods like RANSAC and image segmentation were tested, but did not accurately identify ground height around the corn plants.

III. ROBOT PLATFORM OVERVIEW

The robot platform consists of a six degree-of-freedom xArm robot attached to a four-wheel-drive mobile platform with a custom end-effector for the sensor insertion task. The mobile base is built upon the commercially available Amiga platform which allows for modular hardware adaptations to the required agriculture task as discussed in [23]. Our mobile base is modified by adjusting the bar extrusion for width and height according to the field specifications at Curtiss Farm, Iowa, as shown in Fig. 3. The width of the mobile base (between left and right wheel centers) is set at 1.5 m to accommodate the 0.75 m corn row spacing. The height of the bottom of the robot platform is set according to the expected corn plant height at the V4-V8 stages, such that the robot's platform hovers above the two straddled corn plant rows.

All software for the robot runs onboard on the mini-ITX motherboard containing an Intel i9 24-core 3 GHz CPU and an external RTX4070 GPU, enclosed in a weather-resistant box. The platform is powered by two Lithium-ion batteries providing a roughly 3 h run-time in the field.

IV. GRIPPER DESIGN

A. Design Motivation

As discussed in Section II, the main challenges for precise sensor insertion in the field are variation in stalk attributes (shape, diameter, spacing, and clutter) and the actual contact interaction with the stalk. While there are visual methods to reduce pose uncertainty such as multi-view pose estimation and visual servoing [24], these approaches are less applicable in cluttered environments due to occlusion and collision risks. Another approach to improve robustness for precision tasks is to model the physical contact interaction between the gripper and the deformable stalk. Previous works have shown plant interactions can be modeled analytically [25] or learned [26], but face difficulty in generalizing to the wide variations encountered in the real environment. Instead, to achieve the required precision, we design the gripper system to mechanically induce the alignment of the sensor and corn stalk using funneled edges and

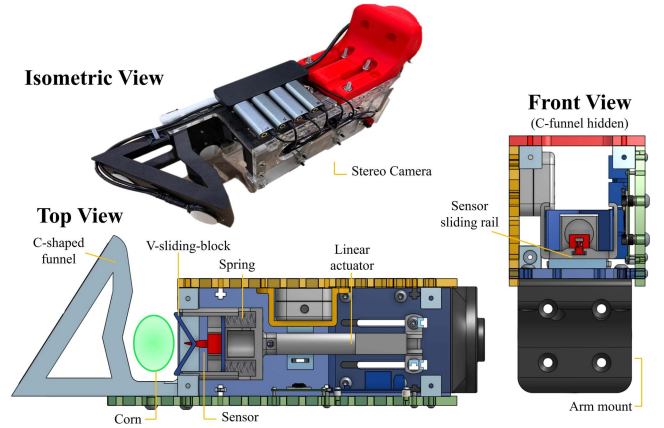


Fig. 4. Custom gripper for sensor insertion. The V-sliding block is actuated to press against the stalk, compressing the spring and exposing the sensor.

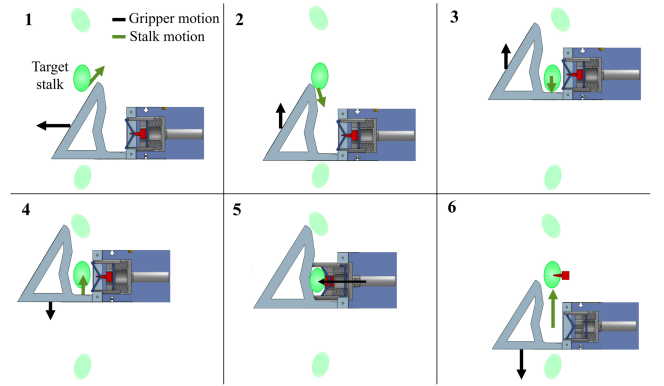


Fig. 5. Sequence of contacts with the stalk for sensor alignment.

spring-loaded compliance. For compactness, the gripper utilizes only one linear actuator to push the sensor into the stalk.

B. Gripper Mechanism

To fit between corn plants and insert sensors near the ground, the gripper is compact, measuring $254 \times 76 \times 76$ mm (L $\times$ W $\times$ H). The gripper contains an Intel D405 stereo camera for stalk detection, a 50 mm stroke linear actuator for linear sensor insertion, a spring-loaded V-sliding-block for compliance, and a rigid C-shaped funnel piece (Fig. 4) for centering the stalk for insertion. The linear actuator's force of 90 N and stroke length of 50 mm were empirically determined to best address stalk rigidity and diameter in a plant's V4-V8 stages.

The front end of the C-shaped funnel is designed as a wedge to traverse through corn plant leaves while approaching the targeted stalk. In cases where the gripper collides with the targeted stalk during the approach motion due to pose estimation or motion execution error, the wedge also guides the stalk until it enters the funnel. As the robot arm moves laterally to position the stalk inside the funnel, the funneled edge comes in contact with the compliant stalk and guides it to the region precisely in front of the V-sliding-block (Fig. 5). The linear actuator then pushes the V-sliding-block to the wall of the funnel where the stalk is located. Because of the stalk's compliance and roundness, as the V-sliding-block comes in contact with the stalk, the stalk further

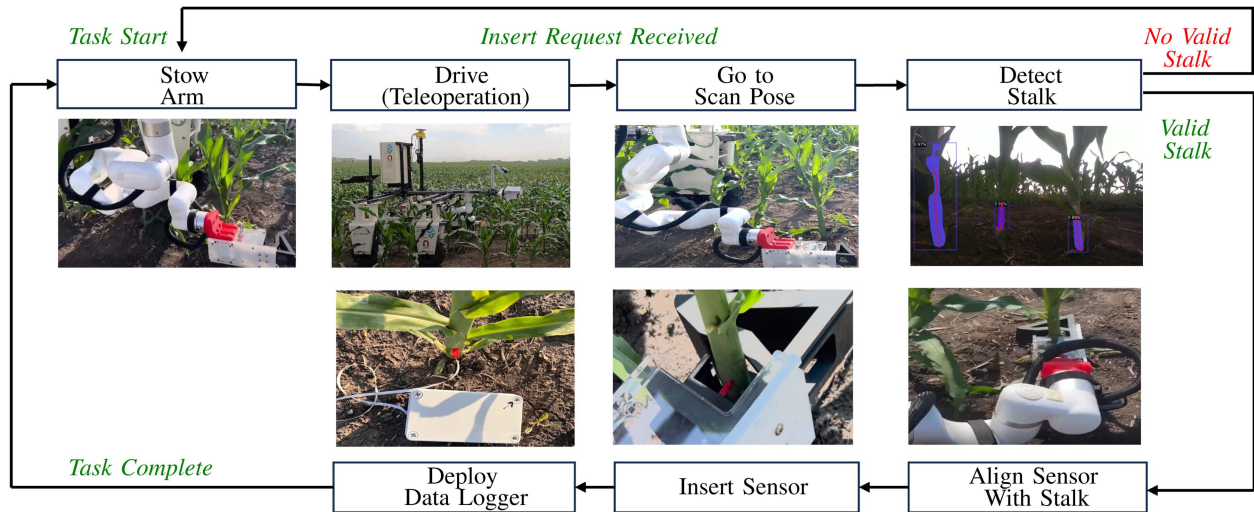


Fig. 6. Overview of the sensor insertion sequence. The user manually drives to a general location and initiates the sequence. The xArm scans for a valid stalk, executes an open-loop motion that utilizes mechanical alignment in the gripper, inserts the sensor, deploys the data logging unit, and awaits command in the next region.

aligns to the cusps of the V-sliding-block. Since the V-sliding-block is spring loaded, upon contact with the stalk, the block slides backward, exposing the rigid sensor to penetrate the stalk. This spring loading is critical, since the stalk is further aligned by the V-sliding-block before the sensor makes contact. Without the spring loaded compliance, the sensor would often initially prod the stalk off-center, at which point further alignment by the V-sliding-block would not re-align the sensor leading to failed insertions.

C. Sensor Deployment

To measure the nutrient content in cornstalks, we utilize the custom nitrate sensor developed by Ali et al. [13]. The sensor relies on the electrochemical reaction from the solution applied on the surface of the printed circuit board (PCB) to generate voltage across the two electrode pads. This solution on the surface of the PCB can wear and tear during the stalk penetration process. To prevent this sensor damage, the sensor is fitted inside a 3D printed sensor sleeve from PLA material, which protects the membrane, as shown in Fig. 2(a).

The sensor sleeve is positioned on a T-slot rail on the V-sliding-block, which restricts the sleeve motion laterally, as shown in Fig. 4. During the actuator extension process, the solid backend of the T-rail holds the sensor in place. During the actuator retraction process, the arrowhead grips to the stalk, providing sufficient force to pull the sensor out from the T-rail. The T-rail tolerance is balanced to provide sufficient friction to hold the sensor in place during motion while allowing the needed tolerance to easily slide in and out during insertion (and reloading a new sensor).

V. INSERTION MOTION

A. Arm Configuration

The xArm is mounted in an inverted configuration, aligned to the center of the mobile platform and towards the front, so that stalks on either side can be targeted as shown in Fig. 3, and the arm can be easily accessed for reloading nitrate sensors in

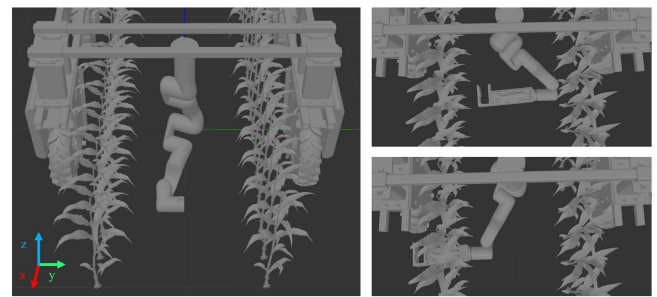


Fig. 7. Arm placement validation conducted through simulation by checking the motion range needed for insertion.

testing. In order to validate that the robot is able to kinematically reach targeted stalks in both rows, motion tests were performed with the robot platform in the Gazebo simulator by moving from the arm's initial position to various possible stalk X,Y,Z positions. As the stalks have no collision properties in simulation, the RRT* trajectory planner provided in MoveIt returns feasible trajectories that allow the arm to reach targeted stalks, as shown in Fig. 7.

B. Insertion Motion Sequence

The software for the insertion motion sequence is built on a ROS SMACH task-level state machine [27]. The finite state machine (FSM) enables modular development of each task. Fig. 6 shows the flow diagram of the insertion sequence, including the xArm motion primitives, stalk detection, sensor insertion, and data logger deployment.

The sequence starts with the robot arm in a stowed position, and the operator teleoperates the mobile base to a general region in the cornfield. Given an insert command, the robot arm moves out of the stowed position to a scan position in order to visually detect cornstalks. As later discussed in Section VI, a detect request to the detection pipeline returns either a reposition response—in which case the robot halts the motion sequence due

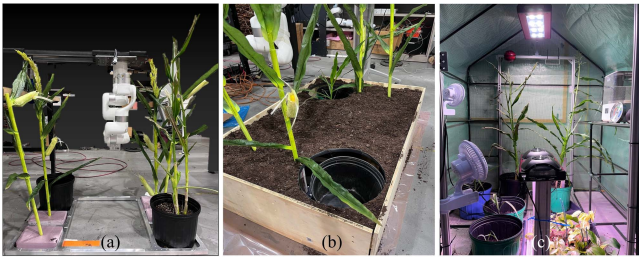


Fig. 8. Mock cornfield setup for indoor evaluation: (a) xArm configuration emulated for field condition; (b) Mock field with synthetic and real corn; (c) Greenhouse created to grow corn plant samples for insertion.

to lack of visible stalks—or a stalk insertion position. This insertion position never exceeds the xArm’s kinematic workspace. Given a valid stalk pose, the gripper approaches the stalk with a predetermined sequence of actions using a combination of joint and Cartesian space commands. In order to align the stalk with the sensor inside the funnel, the arm makes intentional contact with the stalk to guide it into position and locally align the sensor as discussed in Fig. 5. Once aligned, a serial communication to the Arduino inside the gripper triggers the two-channel relay to extend and retract the linear actuator, ramming the sensor into the stalk. As the actuator retracts, the arrow-hook in the sensor grips into the stalk causing the sensor to slide out of the T-slot rail as it retracts. Finally, to deploy a datalogging unit that stores sensor measurements over the plant growth cycle, a serial communication to a separate Arduino inside the deployment box opens one of the five available datalogging units to be released.

Although the robot is operating in a highly cluttered area, the motion sequence does not do 3D scene mapping to explicitly plan for collision-free trajectories. Instead, the motions are a heuristically determined sequence of open-loop trajectories to align the stalk inside the gripper based on the given insertion pose. Since most contacts occur with compliant leaves and plants, this implementation decision does not lead to critical collisions. By not handling external collision avoidance, the motion sequence is fast and lightweight. Furthermore, in the few instances of collision with the ground, the arm safely stops using current limits on all arm joints.

C. Validation on Mock Cornfield

Crucial gripper design decisions, such as the force required to penetrate a cornstalk and optimal dimensions for a funnel, are best determined from real corn plant samples. However, as there are no cornfields available before the summer, cornstalks were grown indoors using a greenhouse. The mock field setup in the lab (Fig. 8) enabled the integration of the arm motion sequence and detection pipeline discussed in Section VI prior to field testing. This included validating the in-hand calibration so that the gripper could grasp the targeted stalk given its detected position from the camera.

VI. VISUAL DETECTION

Sunlight induces strong interference with common active IR RGB-D cameras. We therefore utilize a stereo-based RGB-D camera on the robot gripper, which provides accurate 3D data at close ranges. The perception pipeline’s purpose is to reliably determine a sensor insertion point in 3D space, optimally in the

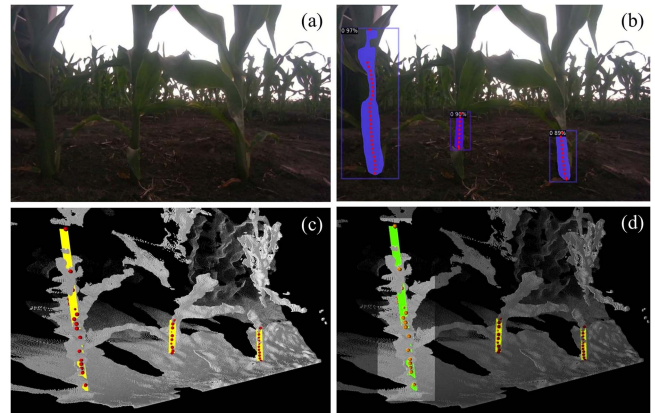


Fig. 9. Stalk detection pipeline: (a) Input image from the in-hand camera; (b) Mask segmentation and the extracted 2D feature points; (c) 3D feature points and fitted lines along the stalks; (d) Extraction of the best stalk for insertion based on scoring metrics from mask and stalk features.

pith region (the first node) of a stalk, which is generally the bottom 2–10 cm of the plant protruding from the ground, based on stalk maturity. The insertion point’s z -coordinate is therefore set as a tuned hyperparameter, since the insertion height varies across fields and vision-based ground plane height estimations proved to be inaccurate due to uneven terrain.

The input to the pipeline is a sequence of RGB-D frames and the output is a stalk 3D position relative to the robot. In each frame, 2D masks of stalks are segmented and the best stalk is determined from 3D stalk attributes extracted from the pointcloud. After all frames are processed, the best insertion position is determined by a consensus among frames.

A. Creating the Stalk Detection Dataset

Variations in lighting, time of day, plant maturity, and field conditions at test-time necessitate a diverse image dataset for training a deep segmentation model. Due to limited access to the cornfield with the robot, the dataset images are collected from videos captured on an iPhone 13 in various regions and sunlight conditions (at 10 AM and 5 PM) at the Curtiss Farm. During recording, the phone is positioned to emulate the viewpoint of the camera on the robot arm—approximately 15 cm above the ground.

First, one frame per second is taken from the video files while removing extraneous images, resulting in a collection of 2667 RGB images. These images are then manually labeled using Meta’s Segment Anything model [28], which efficiently assists the interactive annotation process. Masks are specifically applied to stalks in the foreground, with each mask extending from the bottom of the visible stalk up to the point of occlusion by the plant’s leaves. The labeling process resulted in 7681 stalk instance annotations. The dataset is then split randomly into 80% training, 10% validation, and 10% testing data. To our knowledge, this is the largest corn stalk dataset with segmentation annotation (3.5x the size of the dataset used in [10] and 7.4x that of [29]).

B. Stalk Detection

Given an RGB image as input, the segmentation model outputs 2D masks of all stalks in the foreground. The segmentation

method is based on the Mask-R-CNN architecture [30] with the ResNet-50 backbone and Feature Pyramid Networks for segmentation. A pretrained Mask R-CNN model is fine-tuned on our training data for 40,000 iterations, with a maximum learning rate of $2.5e-4$ and a batch size of 4. Training took 6.7 hours on an NVIDIA GeForce RTX 3060. Quantitative evaluation is discussed in Section VII-A.

C. Stalk Pose Estimation

From the 2D segmentation mask, corresponding depth values along the center points of the mask are used to project to 3D. However, since these masks may be occluded by nearby leaves, it is important to note that the bottom of the mask does not necessarily correspond to the ground. Thus, L number of points along the stalk is fitted to a single line. This is done by using RANSAC followed by Least Squares refinement along the stalk in 3D space. The line is then extended to the ground plane, under the assumption that the robot coordinate's ground plane is at zero height. Representing the stalk as a line allows us to deduce the insertion height from the ground as shown in Fig. 9(c). The fitted line, which robust to partial occlusion, is used to determine the insertion point at the specified z -height.

For N detected stalks $S[(x, y, z)_1, \dots, (x, y, z)_L]_{i=1, \dots, N}$, in order to select the best stalk S^* in which to insert a sensor, the detection pipeline uses a simple yet effective heuristic weighting function to spatially reason and reject candidates that pose difficulties for the robot arm. First, a binary function $R(x, y)$ called *inRange* rejects stalks that are outside the manipulator's workspace based on predetermined boundaries. Second, a binary function $G(y)$ called *Graspable* eliminates stalks that are positioned too close together, such that the gripper's width prevents it from reaching the targeted stalk. From the remaining valid candidates, S^* is selected to be one that has the highest score in the equation:

$$\text{score} = \arg \max(R(x, y) \times G(y) \times c^2 \times w \times \sqrt[3]{h} \times (1 - d)) \quad (1)$$

$$R(x, y) = \begin{cases} 1 & \text{if } x_{\min} \leq x \leq x_{\max} \\ & y_{\min} \leq y \leq y_{\max} \\ 0 & \text{else} \end{cases} \quad (2)$$

$$G(y) = \begin{cases} 1 & \text{if } \Delta y_{i,j} \leq \text{GripperWidth} \\ 0 & \text{else} \end{cases} \quad (3)$$

where c is the segmentation confidence score ($0 \leq c \leq 1$), w is the segmentation mask's width, h is the mask's height, and d is the mask's horizontal distance from the center of the image frame. Thus, preference is given to larger stalks with higher confidence which are closer to the center of the image. The powers on the c and h term are selected from empirical testing.

To evaluate how each function has an affect on scoring for S^* , we conduct an ablation study where individual terms are removed from (1). Evaluating on the 48 stalks in the recorded field data, we compare the success rate of choosing S^* after removal of each term. Ground truth for S^* is visually determined by a human, and success is when the ablated scoring function yields the same stalk as the ground truth. As summarized in Table I, we observe that *Graspable* function $G(y)$ has the most impact on the heuristic weighting score.

TABLE I
ABLATION BY REMOVING EACH TERM FOR BEST STALK SCORING

	None (ours)	inRange	Graspable	Mask Score	Width	Height	Distance
Success Rate $\uparrow$	0.91	0.86	0.62	0.91	0.91	0.91	0.91

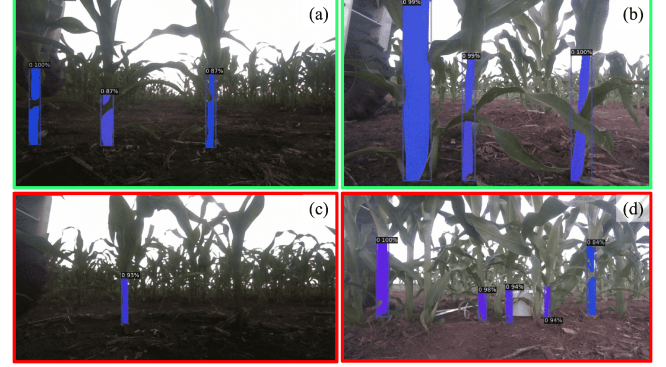


Fig. 10. Segmentation success and failure cases: (a) Successfully segmenting stalks around leaf occlusions; (b) Successfully segmenting stalks with highly varying width; (c) missing a stalk due to shadows in poor lighting conditions; (d) Missing multiple stalks due to occlusion.

TABLE II
EVALUATION OF INSTANCE SEGMENTATION MODELS

Architecture	Mask R-CNN	Mask R-CNN	Center Mask	Center Mask	Mask2Former	Mask2Former
Backbone	R50	R101	R50	VoVNet	R50	Swin-B
AP ₅₀ $\uparrow$	69.2	70.2	83.8	82.9	86.9	87.0

Finally, to reduce the effect of noise and faulty detection, this pipeline is repeated for five frames and the best stalk among all frames is determined by clustering the stalk lines and selecting the representative from the largest cluster.

VII. EXPERIMENTAL RESULTS

To evaluate our platform's ability to insert sensors into cornstalks, we deployed the robot at Curtiss Farm in Iowa during July, in an approximately 40x15 m region of the cornfield with V8 stalks (about 66 cm tall).

A. Stalk Detection Evaluation

Our 2D segmentation model shows accurate stalk segmentation resulting in 69% average precision (AP) at Intersection over Union (IoU) of 0.5 (and 49% at IoU 0.75) on the test dataset. Failure cases are mostly attributed to high amounts of clutter and shadows over the foreground stalks by the robot. Examples of results on challenging scenes are shown in Fig. 10. Occlusions by leaves (which are often mis-segmented by the model as part of the stalk) lead to lower average precision especially at high IoU. To investigate how these failure cases could be improved, after completing the field deployment, we evaluated recent segmentation models such as CenterMask [31] and Mask2Former [32] against our detection pipeline utilizing Mask R-CNN with a ResNet50 backbone. As shown in Table II, Mask2Former segmentation model showed a higher AP, indicating improved detection capabilities for future works.

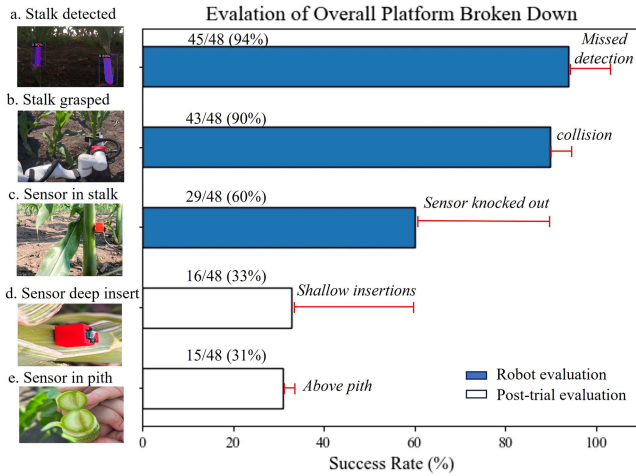


Fig. 11. Overall platform evaluation from 48 insertion trials. We measure 5 criteria in increasing complexity for insertion success.

B. Field Evaluation

We conduct insertion trials on 48 unique cornstalks to assess performance, which is 3x more evaluation samples than in a similar work for stalk grasping [20]. For each trial, the operator manually drove to a new region every couple meters where the robot would then select a target stalk to insert a sensor into. On average, each trial takes two minutes including the driving and sensor reloading time.

The result for each trial is categorized into five criteria for evaluation, as shown in Fig. 11.

The stalk detection pipeline achieves 94% success rate, returning 3 missed detections out of 48 trials when valid stalks are within the field of view. It is important to note that the cameras used during the training of the segmentation network (iPhone 13) and testing (Realsense D405) are different. Nonetheless, the pretrained Mask R-CNN model is robust in distinguishing the learned features of stalks.

The robot successfully grasps 43 stalks out of the 45 detected stalks. As the cornstalks are static targets, the open-loop insertion motion sequence proves sufficient, with a cumulative 90% grasp success rate. Overall, the robot autonomously deploys sensors into 29 of the 48 trial stalks, resulting in sensor insertion success rate of 60%.

The most frequent failure case is not penetrating the sensor deep enough into the stalk. When a sensor is improperly inserted, it may be knocked out of the stalk when the gripper retracts, as shown in insertion sequence step 6 in Fig. 5. Such shallow insertions occur either when the sensor insertion direction deviates too far from the surface normal at the insertion point, or when the insertion point is offset from the vertical center of the stalk. Contrary to our initial belief that cross-sections of cornstalks are mostly circular, stalk cross-sections are often elliptical in shape, and their growth orientation is stochastic in the field, making alignment difficult. For stalks with elliptical cross-sections, insertion positions along the minor axis achieve far greater success than those along the major axis, due to the less severe curvature.

We conduct post-trial evaluation after each insertion attempt by visually inspecting that both electrode pads of the sensor are fully covered inside the plant. We also cut open each stalk to

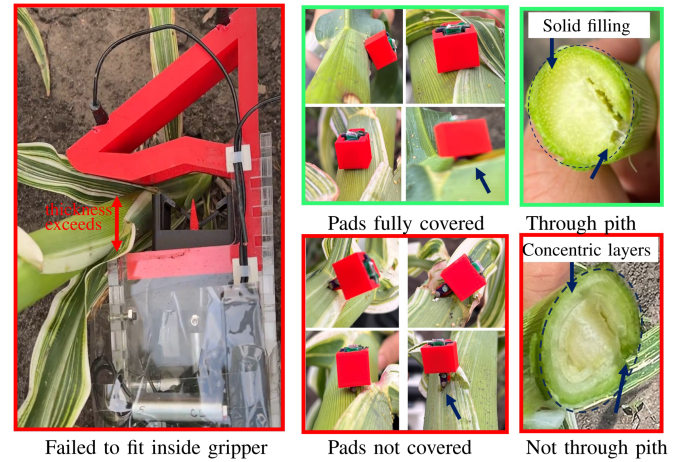


Fig. 12. Visual examples of successful (green box) and failure cases (red box) of sensor insertion.

determine whether the sensor is inserted into the pith region. Of the 29 successful insertions, inspection reveals that in 16 trials, both sensor pads are sufficiently covered. Given the size of the sensor and approximate dimension of the stalks, even a 2 mm error in insertion depth often leads to an incomplete covering of the pads, as shown in Fig. 12. All except one of these trials have successful insertion in the pith region, which is verified by cutting open the stalk and exposing the pith region composed of a solid filling, as opposed to the undesired growth-region which is composed of layers of concentric structures that grow into leaves as shown in Fig. 12.

VIII. DISCUSSION

Having built and deployed a contact-based phenotyping robot platform, we discuss the limitations and insights obtained from our experience. First insight is that stalks can be surprisingly rigid near the pith region we are interested in for sensor insertion. With the aim of improving sensor insertion precision, we designed the gripper with compliance to intentionally make contact with the stalk and slide into the center of the gripper for sensor-stalk alignment. While the spring-loaded V-sliding-block worked on young and compliant stalks, the sensor alignment was significantly worse on more mature and rigid stalks. To address this issue, for the gripper's future iteration, we suggest a mechanism which can fully envelope a stalk and conform to its shape rather than merely sliding against the stalk for alignment.

Second, the overall insertion evaluation shown in Fig. 11 indicate that the most frequent failure case is shallow sensor insertions. We can address this problem via improved mechanical design of the sensor or with a more robust perception algorithm. The addition of the sensor sleeve (red 3D printed material) increases the surface area, and therefore blunts the contact point. This can be improved with electromechanical solutions such as stabbing the stalk with another sharp material before inserting the raw sensor for deeper insertions. In addition, shallow insertions can be improved by also accounting for stalk orientation in the perception pipeline. On the field, we noticed stalks with highly elliptical cross-sections are common.

However, the current detection algorithm disregards this geometry and returns a single insertion position. Thus, improving the perception pipeline to estimate the pose of a 3D ellipse would allow the robot to align itself with the most favorable axis of the ellipse on which the insertion position lies.

Lastly, the xArm motion sequence may be improved by closing the perception-action loop and integrating sensor feedback to improve localization of the stalk. While the demonstrated open-loop motion sequence with the gripper design works well, it cannot account for high variation in stalk diameter and rigidity. Future work includes investigating additional sensing modalities, such as vibro-tactile feedback with an array of contact microphones [33], to reduce the uncertainty of the stalk position within the gripper.

IX. CONCLUSION

We demonstrate our robot platform's capability of inserting sensors into cornstalks, and share design decisions and improvements needed for future research. The task of sub-centimeter-precision sensor insertion in highly cluttered and varying cornstalks is inherently difficult. But understanding the real challenges in the field aids iteration on both hardware and software design to build a research platform that can autonomously insert sensors for phenotyping and crop management. We plan on further improving the research platform for more scalable deployment of sensors, as well as incorporating navigation capabilities to achieve a fully autonomous platform.

ACKNOWLEDGMENT

The authors would like to thank Vignesh Kumar, Professor Liang Dong, and Professor Patrick Schnable from Iowa State University for the help at Curtiss Farm.

REFERENCES

- [1] D. Xie, L. Chen, L. Liu, L. Chen, and H. Wang, "Actuators and sensors for application in agricultural robots," *Machines*, vol. 10, no. 10, 2022, Art. no. 913.
- [2] M. Grell et al., "Point-of-use sensors and machine learning enable low-cost determination of soil nitrogen," *Nature Food*, vol. 2, no. 12, pp. 981–989, Dec. 2021.
- [3] J. Dhakshayani and B. Surendiran, "M2F-Net: A deep learning-based multimodal classification with high-throughput phenotyping for identification of overabundance of fertilizers," *Agriculture*, vol. 13, 2023, Art. no. 1238.
- [4] C. Xie and C. Yang, "A review on plant high-throughput phenotyping traits using UAV-based sensors," *Comput. Electron. Agriculture*, vol. 178, 2020, Art. no. 105731.
- [5] R. Oberti et al., "Selective spraying of grapevines for disease control using a modular agricultural robot," *Biosyst. Eng.*, vol. 146, pp. 203–215, 2016.
- [6] S. Cubero, E. Marco-Noales, N. Aleixos, S. Barbé, and J. Blasco, "Rob-hortic: A field robot to detect pests and diseases in horticultural crops by proximal sensing," *Agriculture*, vol. 10, no. 7, 2020, Art. no. 276.
- [7] S. Kumar, W. Luo, G. Kantor, and K. Sycara, "Active learning with Gaussian processes for high throughput phenotyping," in *Proc. 18th Int. Conf. Auton. Agents Multiagent Syst.*, 2019, vol. 2017, pp. 6700–726152.
- [8] F. Y. Narvaez, G. Reina, M. Torres-Torriti, G. Kantor, and F. A. Cheein, "A survey of ranging and imaging techniques for precision agriculture phenotyping," *Trans. Mechatronics*, vol. 22, no. 6, pp. 2428–2439, Dec. 2017.
- [9] H. Freeman, E. Schneider, C. H. Kim, M. Lee, and G. Kantor, "3D reconstruction-based seed counting of sorghum panicles for agricultural inspection," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2023, pp. 9594–9600.
- [10] J. Yuan, J. Hong, J. Sattar, and V. Isler, "ROW-SLAM: Under-canopy cornfield semantic SLAM," in *Proc. Int. Conf. Robot. Automat.*, 2022, pp. 2244–2250.
- [11] A. N. Sivakumar et al., "Learned visual navigation for under-canopy agricultural robots," in *Proc. 17th Robot. Sci. Syst. RSS*, 2021.
- [12] M. A. Ali et al., "Continuous monitoring of soil nitrate using a miniature sensor with poly(3-octyl-thiophene) and molybdenum disulfide nanocomposite," *ACS Appl. Mater. Interfaces*, vol. 11, pp. 29195–29206, 2019.
- [13] H. Ibrahim, S. Yin, S. Moru, Y. Zhu, M. J. Castellano, and L. Dong, "In planta nitrate sensor using a photosensitive epoxy bioresin," *ACS Appl. Mater. Interfaces*, vol. 14, no. 22, pp. 25949–25961, 2022.
- [14] G. Alenyà, B. Dellen, S. Foix, and C. Torras, "Robotic leaf probing via segmentation of range data into surface patches," in *Proc. 2012 IROS Workshop Agricultural Robot. Enabling Safe, Efficient, Affordable Robots Food Prod.*, 2012.
- [15] A. Atefi, Y. Ge, S. Pitla, and J. Schnable, "In vivo human-like robotic phenotyping of leaf traits in maize and sorghum in greenhouse," *Comput. Electron. Agriculture*, vol. 163, 2019, Art. no. 104854.
- [16] Y. Bao, L. Tang, and D. Shah, "Robotic 3D plant perception and leaf probing with collision-free motion planning for automated indoor plant phenotyping," in *Proc. ASABE Annu. Int. Meeting. Amer. Soc. Agricultural Biol. Engineers*, 2017.
- [17] D. Shah, L. Tang, J. Gai, and R. Putta-Venkata, "Development of a mobile robotic phenotyping system for growth chamber-based studies of genotype x environment interactions," *IFAC-PapersOnLine*, vol. 49, no. 16, pp. 248–253, 2016.
- [18] A. Atefi, Y. Ge, S. Pitla, and J. Schnable, "Robotic detection and grasp of maize and sorghum: Stem measurement with contact," *Robotics*, vol. 9, no. 3, 2020, Art. no. 58.
- [19] T. Mueller-Sim, M. Jenkins, J. Abel, and G. Kantor, "The robotanist: A ground-based agricultural robot for high-throughput crop phenotyping," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2017, pp. 3634–3639.
- [20] T. Parhar, H. Bawaja, M. Jenkins, and G. Kantor, "A deep learning-based stalk grasping pipeline," in *Proc. Int. Conf. Robot. Automat.*, 2018, pp. 6161–6167.
- [21] R. Xu and C. Li, "A review of high-throughput field phenotyping systems: Focusing on ground robots," *Plant Phenomics*, vol. 2022, 2022, Art. no. 9760269.
- [22] A. Atefi, Y. Ge, S. Pitla, and J. Schnable, "Robotic technologies for high-throughput plant phenotyping: Contemporary reviews and future perspectives," *Front. Plant Sci.*, vol. 12, 2021, Art. no. 611940.
- [23] D. Guri, M. Lee, O. Kroemer, and G. Kantor, "Hefty: A modular reconfigurable robot for advancing robot manipulation in agriculture," 2024, *arXiv:2402.18710*.
- [24] R. Haugaard, J. Langaa, C. Sloth, and A. Buch, "Fast robust peg-in-hole insertion with continuous visual servoing," in *Proc. Conf. Robot Learn.*, 2021, pp. 1696–1705.
- [25] A. LaGrassa, M. Lee, and O. Kroemer, "Task-oriented active learning of model preconditions for inaccurate dynamics models," 2024, *arXiv:2401.04007*.
- [26] C. H. Kim, M. Lee, O. Kroemer, and G. Kantor, "Towards robotic tree manipulation: Leveraging graph representations," 2023, *arXiv:2311.07479*.
- [27] "Ros smach library," Accessed: Oct. 10, 2023. [Online]. Available: <https://wiki.ros.org/smach>
- [28] A. Kirillov et al., "Segment anything," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2023.
- [29] Z. Zhang, E. Kayacan, B. Thompson, and G. Chowdhary, "High precision control and deep learning-based corn stand counting algorithms for agricultural robot," *Auton. Robots*, vol. 44, pp. 1289–1302, 2020.
- [30] K. He et al., "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017.
- [31] Y. Lee and J. Park, "CenterMask: Real-time anchor-free instance segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 13903–13912.
- [32] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar, "Masked-attention mask transformer for universal image segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022.
- [33] M. Lee, M. Lee, K. Zhang, G. Kantor, and O. Kroemer, "Contact classification for agriculture manipulation in cluttered canopies," in *Proc. 2021 AAAI workshop AI Agriculture Food Syst.*, 2021.