2024 International Conference on Unmanned Aircraft Systems (ICUAS)
June 4-7, 2024. Chania, Crete, Greece

UAV–based Foliage Plant Species Classification for Semantic Characterization of Pre–Fire Landscapes

Prateek Arora*, Pau Alcolea Vila, Aiden Borghese, Stephen J. Carlson, David Feil-Seifer, Christos Papachristos

* *Consider for Best Student Paper Award*

***Abstract*— In this work we deal with the problem of establishing a system architecture to facilitate the real-time autonomous volumetric mapping alongside the semantic characterization of sagebrush ecosystem landscapes, in order to support the pre-fire modeling and analysis required to plan for wildfire prevention and/or suppression. The world, and more specifically the broader region of N. Nevada has been facing one of its most challenging periods over the course of the last decade, as far as uncontrolled wildfires are concerned. This has led to the development of research initiatives aimed at the ecosystem-specific modeling of the pre-, during-, and post-fire process effects in order to better understand, predict, and address these phenomena. However, to collect the required wide-field information that contains both centimeter-level volumetric mapping fidelity, as well as semantic details related to plant (sub)-species, which for the common case of sagebrush can only be identified based on close-up inspection of their foliage fine structure, satellite photography remains insufficient. To this end, we propose a perception and mapping architecture of an aerial robotic system that is capable of: a) LiDAR-based centimeter-level reconstruction, b) robust multi-modal sensor fusion Simultaneous Localization and Mapping (SLAM) leveraging LiDAR, IMU, Visual-Inertial Odometry, and Differential GPS in a global optimization mapping framework, as well as c) a gimbal-driven point-zoom camera for the efficient real-time collection of close-up imagery of foliage pertaining to specific target plants, in order to allow their real-time identification based on their leaf micro-structure, by leveraging Deep-Learned classification deployed on a Neural Processing Unit. We present the associated systems, the overall hardware and software architecture, as well as a series of field deployment studies validating the proposed aerial robotic capabilities.**

I. INTRODUCTION

Unmanned Aerial Systems (UASs) are increasingly being utilized across multiple application domains, including exploration of subterranean environments [1–6], infrastructure inspection operations [7–11], search and rescue [12–14], as well as futuristic concepts such as extraterrestrial discovery robots [15, 16] and perpetually-deployed self-sustainable systems [17–22]. The relevant advances in autonomy and multi-modal resilient perception [23–26] have been the backbone of their success. Across all these, aerial robots are broadly considered as remote sensing [12] platforms. More recently, the need has arisen for such airborne deployments to provide data that facilitate concurrent mapping and semantic characterization of the operating environment[27], in order to inform critical decision-making; one such prominent application domain is wildfire prevention and/or suppression, through high-fidelity landscape characterization to drive predictive fire process modeling.

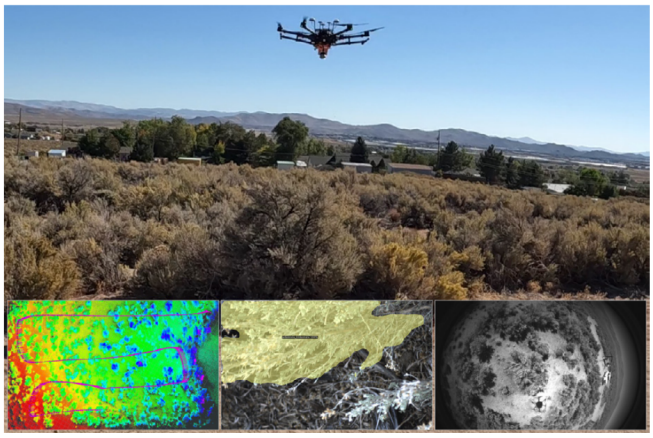

Fig. 1. Field demonstration of the Desert Ecology Nevada Drone (DENDrone) aerial system during data collection for desert landscape 3D mapping, and plant species characterization. Details: a) Path and derived LiDAR-inertial & D-RTK GPS-based 3D volumetric map; b) Foliage-based plant species classification; c) Indicative image of the challenging nature of appearance-based classification for desert environment shrubs.

A major requirement for predicting fire processes involves accurate characterization of the biomass or volume estimation as well as species characterization and identification of the foliage. Traditional estimation techniques are based on either ground measurements or satellite image remote sensing. Although ground-based methods are more accurate, they involve destructive methods that eliminate the vegetation of the landscape, are limited to a small scale, and are unable to provide a biomass estimate of a bounded region. On the other hand, satellite imaging-based estimation methods are non-destructive, but involve high-scale maps over large landscapes, hence suffering from low spatial resolution, as well as data retrieval challenges related to atmospheric conditions such as cloud coverage. In contrast, leveraging UASs for remote sensing provides the best of both approaches, essentially having the advantage of broad-scale landscape characterization with the flexibility of close-up examination for more precise observations.

This material is based upon work supported by the NSF Awards: *2148788: EPSCoR RII Track-1: Harnessing the Data Revolution for Fire Science*, and *IIS-2150394: Research Experiences for Undergraduates Site: Collaborative Human-Robot Interaction for Robots in the Field*. The presented content and ideas are solely those of the authors.

The authors are with the University of Nevada, Reno, 1664 N. Virginia, 89557, Reno, NV, USA prateeka@nevada.unr.edu

2024 International Conference on Unmanned Aircraft Systems (ICUAS) | 979-8-3503-5788-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICUAS60882.2024.10556861

In this work we propose a UAS-based system to facilitate the pre-fire characterization of desert-type landscapes encountered in the N. Nevada region; the Desert Ecology Nevada Drone (DENDrone). The DENDrone's architecture integrates multi-modal sensor fusion, relying on LiDAR / Inertial / GPS systems for consistent and accurate 3D mapping of the desert / shrub landscape, despite the robust perception challenges due to environment's "flat" structure. The accurate 3D mapping is a key requirement for fire fuel estimation contained within plants, as it allows an approximate but consistent assessment of the enclosed volume. More importantly however, knowledge of the specific plant species is required to increase fire fuel estimation accuracy. For plants endemic to this ecosystem, shrubs such as sagebrush (Artemisia tridentata), rabbitbrush, are ambiguous when observed at wide camera scales. We propose the use of a gimbaled mechanism with a high-zoom lens and camera that unlocks capturing fine detail features, such as the leaf microstructure. Such close-up foliage inspection is a common way for experts to distinguish such plant species. We also propose and experimentally validate the use of a properly trained Semantic Segmentation Deep Learning framework, to enable the automation and systematicity of this critical task.

The remainder of this paper is structured as follows: Section II presents the relevant prior work in the field. Section III describes our proposed approach for 3D reconstruction and mapping as well as semantic characterization of the landscape. Section IV evaluates the results of the semantic segmentation approach. Experimental results presenting the 3D reconstructed map, volume estimation, and semantic segmentation are presented in Section V, and our conclusions are drawn in Section VI.

II. RELATED WORK

Recent advancements in deep–learning approaches have led to its extensive adoption for various applications, especially in the field of agriculture robotics and remote sensing to characterize and identify tree species [28–32], vegetation [33–36], and weed/invasive species [37–42] in the presence of challenging lighting conditions and scale variation. These works pose the characterization problem either as classification and bounding box regression or as segmentation (semantic or instance) to identify the species in images acquired from RGB cameras. Several other works use images from acquisition devices such as infrared, hyperspectral, or multispectral [43–46] cameras to train the neural network allowing better and robust detections under varying lighting conditions. Given the extensive data requirement of deep–learning algorithms, these works either rely on existing datasets or make efforts to collect their own dataset. In contrast, many studies rely on satellite imagery as means for data collection and have successfully applied deep–learning based detection and segmentation techniques for vegetation identification [47–50]. However, satellite images suffer from poor spatial resolution required for the task, depend highly on cloud coverage, and have fixed revisit time which doesn't allow for real-time identification. In the past decade, realizing the potential of Unmanned Aerial Vehicles (UAVs) for remote sensing as an active and flexible tool capable of collecting high temporal and spatial resolution images, recent works have leveraged them for faster and more precise data collection as well as real-time online deployment for vegetation characterization [33, 51–56].

Apart from species identification, foliage volume or biomass estimation has been an area of interest in the fire evolution community. Traditionally approaches estimate biomass by destructive sampling methods and rely on manual calculation of dry weight which are time-consuming, laborious, and limited to small scales[57]. Over the past few years, non-destructive remote sensing techniques have emerged for biomass characterization for large-scale areas by using classical and deep–learning based methods on images[58–64], while a few works have successfully applied learning–based methods on other modalities such as lidar [65, 66].

Despite the enormous progress, sagebrush sub-species identification poses a unique challenge due to their similar structural characteristics from a distant view that can only be distinguished upon close-up examination of the leaf structures which are only a few centimeters long. Therefore, a system that is capable of flying over large areas and performing close-up visual inspection, as well as mapping of sagebrush, is required. In this work, we propose a UAV system that is capable of centimeter-level onboard 3D mapping of the scene that allows for sagebrush volume estimation as well as identification of sagebrush subspecies in RBG images captured by a high-optical-zoom-lens gimbal-mounted camera.

III. PROPOSED APPROACH

This section details the primary components of the proposed foliage plant species segmentation.

A. DENDrone UAV platform

In this work, we use the DJI Matrice M600 Pro hexacopter equipped with an A3 pro flight control system and DJI Differential Real-time kinematic (D-RTK) module as the primary localization sensor. The aircraft is modified with 3D printed structures to equip other sensor modalities, namely, Ouster OS0-120 LiDAR, VectorNav VN-100 IMU/AHRS, FLIR Blackfly S USB3 RGB camera as well as an onboard Intel NUC-I7 PC that allows us to execute onboard 3D Mapping and other state estimation pipelines. Additionally, a rigidly mounted Tarrort 3-axis gimbal holds the RGB camera equipped with a high-optical zoom lens that allows us to acquire a close-up view of the sagebrush leaf structure. The gimbal arms that connect to each motor have been replaced with slightly extended 3D-printed arms to accommodate the full range of motion of the camera. All the sensors, except the GPS sensor, gimbal-mounted camera, and the onboard companion PC are rigidly attached to the bottom of the UAV platform whereas the DJI D-RTK receiver module is mounted on the top along with the DJI GPS module. The primary localization sensor, i.e. D-RTK module offers horizontal positional accuracy of $1\ cm + 1\ ppm$ and vertical positional

accuracy of 2 $cm + 1$ ppm with orientation accuracy of $(0.2/R)°$, where R is the baseline distance (distance between the two Air System antennas) in meters, while the primary mapping sensor, i.e. Ouster OS0-128 LiDAR, offers 90° vertical FoV, 100 m maximum range and 128 Channels of resolution.

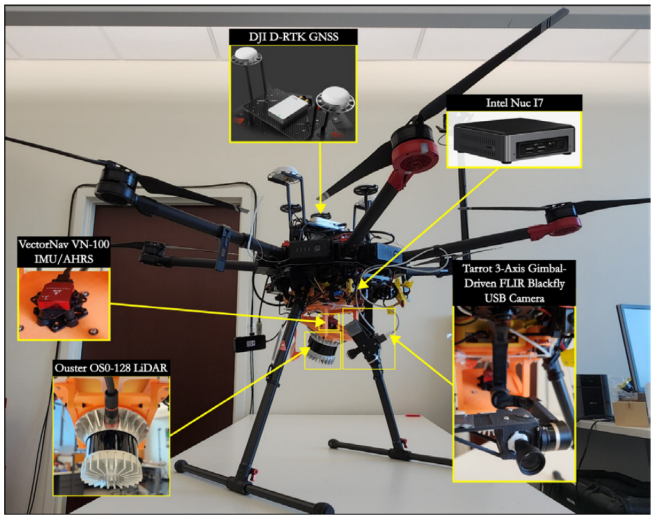

Fig. 2. The DENDrone platform: DJI Matrice M600 Pro accompanied with various sensor modalities such as LiDAR, IMU, D–RTK, GPS, and gimbal–driven point–zoom camera as well as i7 intel core companion computer. All the components are mounted to the UAV by leveraging custom-designed 3D printed parts.

B. 3D Reconstruction and Mapping

The first fundamental capability of the proposed system consists of multi–modal Simultaneous Localization and Mapping (SLAM) and volumetric reconstruction, Model Predictive Control (MPC), and state estimation pipeline based on established open–source works [67–70]. The SLAM pipeline adopts a graph–based approach incorporating multiple factors from various modalities, namely, LiDAR odometry factors, IMU preintegration factors, and D–RTK factors to create a consistent map of the environment. The landscape considered in this work presents a challenge for the reconstruction pipeline due to mostly flat nature of the terrain and therefore lack of prominent LiDAR features to constrain the optimization problem. Despite the challenge, the proposed pipeline is successfully able to construct a consistent map of the environment with centimeter-level accuracy. Given a reference local region based on the reconstructed map, the volume of the local bounded region containing the sagebrush shrub can be estimated. A Normal representation of the reference region containing the foliage can be used for ground plane estimation and thus segment the shrub corresponding to the sagebrush. Followingly, deploying surface generation technique over the locally segmented point cloud, a cad model can be derived which can be used to infer the corresponding volume.

C. Deep Learning Based Semantic Segmentation

The second capability of the proposed system is a pipeline that allows for semantic segmentation for sagebrush characterization. Given a query 3D location based on the 3D reconstructed map, the proposed approach aims to identify the sagebrush species. For close–up visual examination, the UAV positioned directly above the query 3D location with a downward–facing camera begins descending vertically to acquire a sequence of images $\mathcal{I}_W$ with the aim of obtaining the sharpest, in-focus image such that it contains the maximum discernible sagebrush leaf structure. The reason for obtaining a sharp image is two-fold: first, the shallow depth of field of the onboard zoom lens limits us to capture very few features of leaf structures if captured from a poorly selected viewpoint or if the camera is pointed towards a less dense area which is exacerbated by other factors such as insufficient light or sensor noise, and secondly, the downward gush from propellers causes additional swaying of the foliage leaves that further contributes to the motion blur. Therefore, acquiring a focused, sharp image is crucial for real-time computationally efficient implementation. This is achieved by denoising each image using Gaussian blur in the sequence of Images captured $\mathcal{I}_W$ and then retrieving the top five candidate images $\mathcal{I}_S$ with the maximum gradient, as follows: $argmax \ \sum_{i,j} |G_{i,j}|$, where $G_{i,j} = \left[\frac{\partial f_{i,j}}{\partial x}, \frac{\partial f_{i,j}}{\partial y}\right]$

The candidate images $\mathcal{I}_S$ are then used for forward inference to get images $\mathcal{I}_M$ with semantic segmentation mask corresponding to the sagebrush species. It is highlighted here, that we only consider semantic masks as positive candidates that have detection confidence over 85%. An alternate approach to the process of executing forward inference operation on a select few images is to instead continuously perform inference on a sub-sample of the captured image facilitated by hardware acceleration capabilities of neural compute stick. Deployment of the trained network on such hardware requires the model weights to be quantized, essentially converting the floating point weights to lower bandwidth integer weights, providing accelerated inference speed while trading off a marginal amount of accuracy.

For each surviving segmentation mask in the image, ray–casting operation is performed at the centroid of each mask to get corresponding 3D locations when the ray hits an occupied cell. These 3D locations can then be transformed in the world frame and marked for later processing for pre-, during-, and post-fire processing study. The entire detection pipeline continues the process for the next query point for further characterization of the remaining landscape.

It is noted that we use Mask R-CNN with ResNet+FPN backbone initialized with MSCOCO pre-trained weights provided by the detectron2 [71] framework that we fine-tune on our custom dataset. The backbone version used is ResNet50 which is lightweight and suitable for onboard deployment. The network adopts a two-stage procedure, with the first stage being the region proposal network that defines a region of interest (RoI) where the object may lie, while the second stage, predicts the class and box offset, a binary mask for each RoI in parallel.

We use a custom dataset to train the network that is collected with the UAV system described in the paper. The

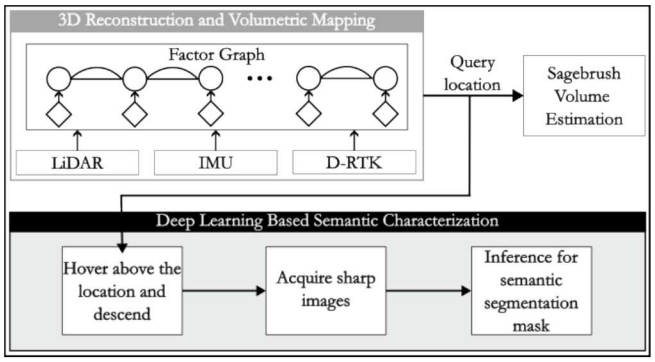

Fig. 3. The overarching architecture of the proposed approach, indicating how the individual capabilities combine together to achieve the envisioned task of sagebrush identification and characterization

dataset primarily contains images belonging to two species of sagebrush, namely, a) Artemisia tridentata and b) Rabbitbrush, with varying lighting conditions captured at different times of day in the Northern Nevada region, totaling over 200 images. The training, validation, and testing image split is 60% : 20% : 20% .The images have a spatial resolution of 1440x1080 pixels, are captured at an acquisition frame rate of 120 fps, and are hand-labeled [72] with semantic masks that correspond to each species where the leaves of the sagebrush are clearly visible and in focus, i.e. we mark the pixels that are in focus, have visible leaf structure, and are unaffected by huge motion blur. It is also noted that the boundary of the mask is arbitrary since the shallow depth of field gradually blurs one region of the sagebrush from the other and there is no one particular boundary that can be considered as a distinct outline for the segmentation mask. Figure 3 illustrated the overarching system architecture described in this section.

IV. SYSTEM PERFORMANCE

In this section, we discuss the performance of the trained Deep Learning network for generating semantic segmentation masks on our test datasets, essentially demonstrating the efficacy of the deployed network for our envisioned task.

A. Semantic Segmentation Qualitative Results

Figure 4 provides the inference results of unseen test image dataset with segmentation masks overlayed on the classified pixels of the input image. It can be observed that the network inherently learns to identify only the parts of the image that are focused and does not attempt to classify any blurry pixels, which severely reduces the chances of false positive detections. This is the expected behavior that we require for our envisioned task. Furthermore, the network is capable of differentiating between the two species even when the image sample contains the same color shades (yellow) as evident from the bottom two rows in Figure 4, essentially indicating that the network relies more on foliage structural features as compared to other features. Having the capability to discern different species based on structural features such as leaf shape and size, branching pattern, and shrub density is crucial for sagebrush identification.

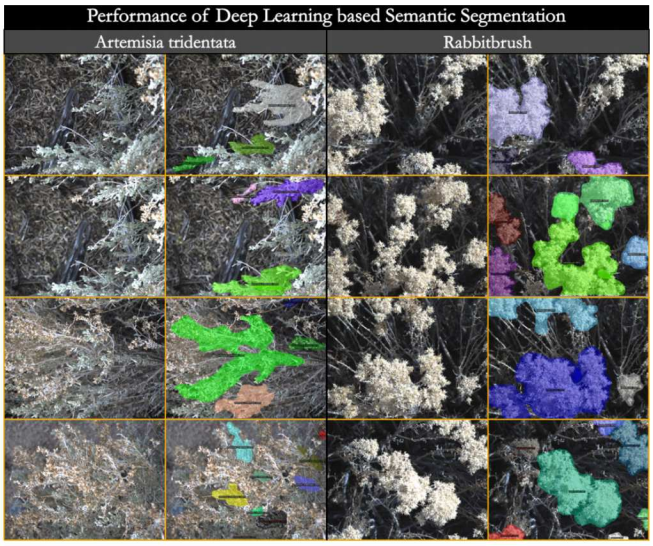

Fig. 4. The figure depicts the two species of Sagebrush and their corresponding segmentation mask. The left two columns contain images belonging to Artemisia tridentata while the two columns on the right contain images of Rabbitbrush. Out of the two columns in each category, the left image corresponds to the input to the network while the right one depicts the detected mask overlayed on the input image.

B. Semantic Segmentation Quantitative results

Table I provides the average precision (AP) score for various IoU thresholds and area thresholds. It can be observed that the mean AP is lower than AP 0.5 IoU threshold and as we increase the IoU threshold to 0.75 the AP drops even further. This is primarily because of the nature of the collected data and the absence of a single delineating boundary for ground truth segmentation mask as mentioned in the subsection III-C. Since a single clear mask boundary is hard to establish it is an expected behavior to observe a drop in the AP values for a higher threshold of IoU. A similar pattern can be observed for the AP over the average IoU when compared to different areas, i.e. for larger mask areas more sharp features are visible which leads to better model performance as compared to the smaller areas.

TABLE I
AVERAGE PRECISION SCORE OF THE SEMANTIC NETWORK TESTED ON OUR CUSTOM SAGEBRUSH DATASET

AP @[IoU=0.50:0.95 | area= all]	31.1
AP @[IoU=0.50 | area= all]	52.7
AP @[IoU=0.75 | area= all]	37.3
AP @[IoU=0.50:0.95 | area= small]	12.6
AP @[IoU=0.50:0.95 | area= medium]	35.7
AP @[IoU=0.50:0.95 | area= large]	51.9

V. EXPERIMENTAL STUDY

We present a comprehensive set of results that showcase the proposed DENDrone system's field capabilities related to the characterization of pre-fire landscapes in N. Nevada. More specifically, the collected dataset is from a desert / shrub ecosystem near Reno, Nevada, one of the candidate sites being considered for the deployment of the HDRFS project infrastructure, which aims toward the systematic

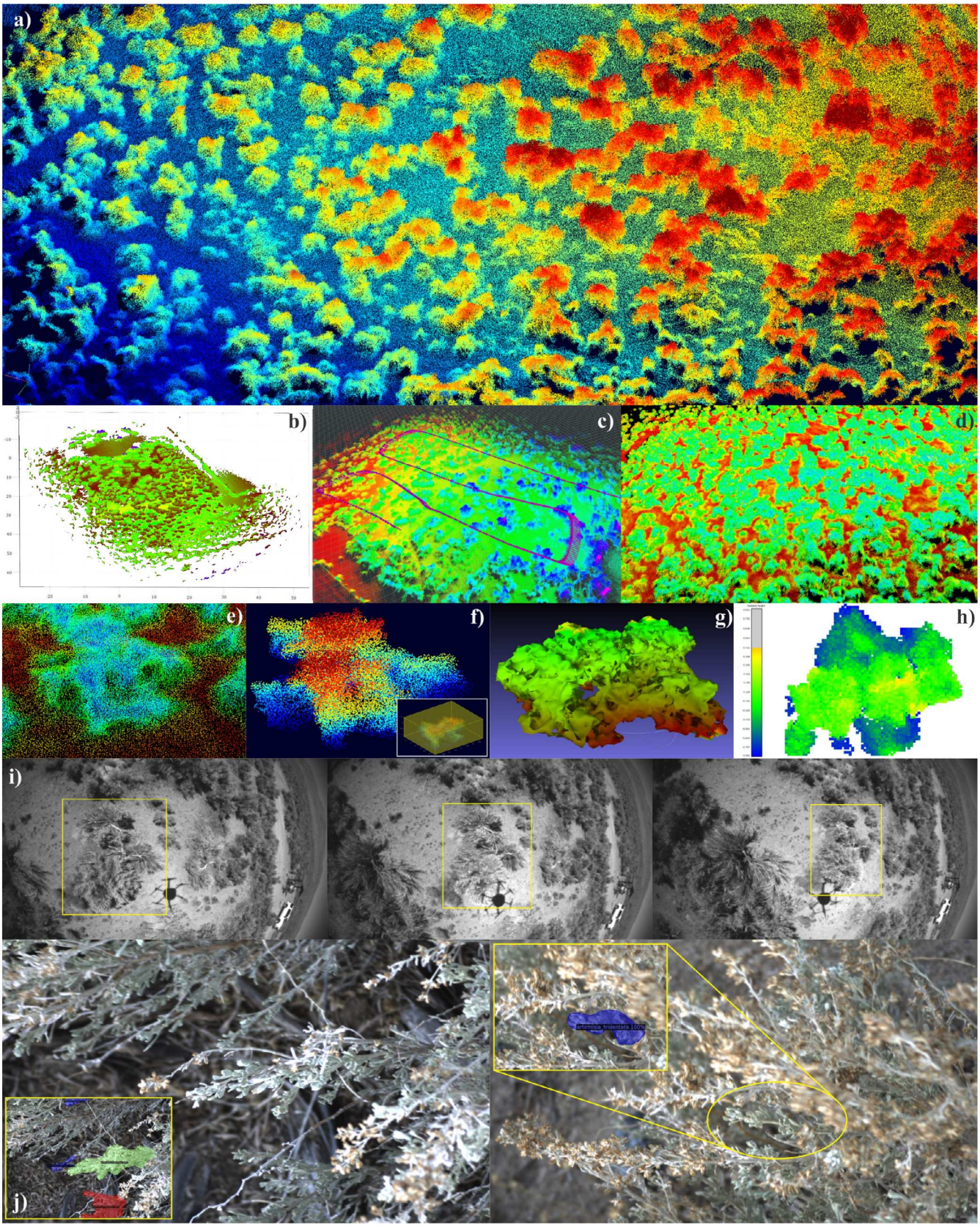

Fig. 5. Experimental Validation of the DENDrone's pre-fire landscape characterization capabilities. *a)*: The 3D-mapped desert environment comprising mostly sagebrush plants. *b)*: Size of experimental "plot". *c)*: Dataset collection path and online 3D-reconstructed result. *d)*: Normals-estimation colormap indicating the viability of ground/plant segmentation for desert-type landscapes. *e-h)*: Sample shrub volumetric segmentation and manifold-enclosing volume estimation. *i)*: Wide lens flyby indicating the lack of species-identifying features at this scale. *j)*: Zoomed-in image capture enabling foliage micro-structure-based classification with a Deep-Learned Semantic Segmentation framework.

year-round monitoring, characterization, and ecological modeling of an area to-scale with the presented experimental scenario. The mapped shrub ecosystem spans an area of approximately $3000m^2$ while the DENDrone followed a sweeping flight profile with varying altitudes of $3m - 5m$ above the ground. The corresponding results are illustrated in Figure 5.

Typically, multimodal sensor mapping methods rely on datasets that include ground truth information generated by GPS modules for benchmarking their algorithms. As one of the initial efforts in characterizing and mapping such challenging terrain, where existing LiDAR-inertial/GPS solutions fail, an existing dataset with ground truth trajectory is lacking. Leveraging the centimeter-level accuracy provided by the D-RTK (Differential Real-Time Kinematic) system for localization, the trajectory generated by this system can serve as a reliable ground truth reference for future benchmarking efforts.

A. Desert Terrain Ecology Consistent 3D Mapping

The first set of rows in Figure 5 demonstrates the 3D mapping consistency achieved by the system despite the challenging "flat" nature of the desert landscape.

In the top row, a) showcases the 3D pointcloud result of LiDAR / Inertial / D-RTK GPS mapping over the aforementioned area. As is evident by the elevation-based colormap, lack of vertically-high structures in the natural environment justifies our selection for a ground-pointed LiDAR configuration.

In the following row, b) presents the metric size of the 3D-mapped area of our experimental "plot", and c) illustrates the DENDrone's path that was followed to collect the associated data. It is mentioned that the indicated 3D mapping result of this picture is achieved onboard the system in flight; nevertheless, the collected raw data can be used for offline 3D mapping without suffering any real-time constraint requirements, in order to facilitate higher mapping consistency. Finally, d) shows a possible approach to facilitate the detection of Regions-Of-Interest that contain shrub vegetation. More specifically, this subfigure presents a normals-based colormap for the derived pointcloud, which is more consistent than an elevation map in separating the ground from the plant based on the estimated per-point normals values.

B. Shrub Volume Estimation for Pre-Fire Fuel Characterization

The subsequent row in Figure 5 illustrates the application of DENDrone's results to facilitate the critical process of fuel estimation for a pre-fire landscape.

More specifically, e) presents a close-up of the normals-colored pointcloud at an area which contains a shrub, illustrating how this information can support ground plane segmentation to isolate the vegetation. Subfigure f) shows the 3D segmented plant part corresponding to the previously mentioned shrub (points are height-color mapped in this instance), and followingly g) illustrates 3D mesh estimation of the isolated plant manifold. It is noted that actual fire fuel estimation is not a self-evident process: First, because mesh volume estimation assumes a solid manifold, which is not the case for plants whose actual volume is determined by structural details (branches, leaves) hidden underneath the foliage and impenetrable by LiDAR. Secondly, because the actual fuel capacity depends on the plant species, its characteristics, and the time-of-year. This is a goal of the HDRFS project's objectives which depends on the integration of data and methods between fire, ecology, and robotics science experts; nevertheless, it depends on the demonstrated DENDrone system's capacity for 3D reconstruction and can be supported by approximate volume estimation methods, such as the one shown in h).

C. Shrub Species Classification for Vegetation Characterization

The last set of rows in Figure 5 is focused on the foliage-based plant species classification capabilities of DENDrone. As previously highlighted, this is critical information to support the accurate pre-fire fuel estimation of desert-like landscapes.

Row i) showcases a flyby of the shrub which was discussed in the previous subsections, denoting it on the frame captured by a wide Field-of-View camera. As discussed, even if the segmentation of shrubs in such an ecosystem can be achieved with traditional UAV-borne sensing, the identification of (sub-)species between bush-like plants is exceptionally difficult based on their macroscopic appearance (even with color cameras).

Towards this goal, the DENDRone system supports the RGB-capture of high-optical-zoom images with the gimbal-mounted camera mechanism, as previously discussed. Row j) illustrates two instances of such an operation in this landscape, indicating the effectiveness of the employed approach. Not only is the leaf-micro structure and its details accurately captured, but as shown in the callout boxes, the proposed Semantic Segmentation framework achieves classification across widely differently-looking scenes: The first presented case (easy) contains multiple Artemisia tridentata (sagebrush) leaves, which are correctly identified by the trained Deep Learning network, but the second picture (hard) contains same-species flowers and branches mostly; nevertheless, the system correctly identifies the plant by even the very few leaves in the frame, without getting confused by its training for rabbitbrush-type flower structures.

Overall, the DENDrone system experimentally demonstrates its ability to achieve the goals of robust and consistent centimeter-level accurate 3D mapping of desert-like landscapes, as well as shrub (sub-)species classification by relying on appropriately tailored macro-lens airborne imagery and a Deep-Learned Semantic Segmentation framework.

VI. CONCLUSIONS

In this work, we presented a system for the systematic collection of data to support pre-fire fuel estimation of desert

ecosystems and landscapes. The facilitating system, DEN-Drone, is capable of centimeter-level accurate and consistent mapping of such landscapes through LiDAR / Inertial / D-RTK GPS fusion, but is additionally capable of macro-imagery collection of plants, in order to examine their foliage micro-structure. Through this, as well as a properly trained Deep-Learned Semantic Segmentation framework, it enables the plant species identification even between difficult-to-distinguish shrubs, which is a key capacity to unlock accurate fire fuel estimation enclosed within a plant's volume. The proposed system was demonstrated w.r.t. its operating performance in real-world field conditions, comprehensively showcasing the associated capabilities.

REFERENCES

[1] T. Dang, F. Mascarich, S. Khattak, H. Nguyen, N. Khedekar, C. Papachristos, and K. Alexis, "Field-hardened robotic autonomy for subterranean exploration," *Field and Service Robotics (FSR)*, 2019.

[2] S. Khattak, F. Mascarich, T. Dang, C. Papachristos, and K. Alexis, "Robust thermal-inertial localization for aerial robots: A case for direct methods," in *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2019, pp. 1061–1068.

[3] C. Papachristos, S. Khattak, and K. Alexis, "Uncertainty-aware receding horizon exploration and mapping using aerial robots," in *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, pp. 4568–4575.

[4] C. Papachristos, M. Kamel, M. Popović, S. Khattak, A. Bircher, H. Oleynikova, T. Dang, F. Mascarich, K. Alexis, and R. Siegwart, "Autonomous exploration and inspection path planning for aerial robots using the robot operating system," in *Robot Operating System (ROS)*. Springer, Cham, 2019, pp. 67–111.

[5] C. Papachristos, F. Mascarich, S. Khattak, T. Dang, and K. Alexis, "Localization uncertainty-aware autonomous exploration and mapping with aerial robots using receding horizon path-planning," *Autonomous Robots*, vol. 43, no. 8, pp. 2131–2161, 2019.

[6] C. Papachristos, S. Khattak, F. Mascarich, T. Dang, and K. Alexis, "Autonomous aerial robotic exploration of subterranean environments relying on morphology–aware path planning," in *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2019, pp. 299–305.

[7] A. Bircher, K. Alexis, M. Burri, P. Oettershagen, S. Omari, T. Mantel, and R. Siegwart, "Structural inspection path planning via iterative viewpoint resampling with application to aerial robotics," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 6423–6430.

[8] C. Papachristos, K. Alexis, L. R. G. Carrillo, and A. Tzes, "Distributed infrastructure inspection path planning for aerial robotics subject to time constraints," in *2016 international conference on unmanned aircraft systems (ICUAS)*. IEEE, 2016, pp. 406–412.

[9] F. Mascarich, T. Wilson, C. Papachristos, and K. Alexis, "Radiation source localization in gps-denied environments using aerial robots," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 6537–6544.

[10] C. Papachristos and K. Alexis, "Augmented reality-enhanced structural inspection using aerial robots," in *2016 IEEE international symposium on intelligent control (ISIC)*. IEEE, 2016, pp. 1–6.

[11] G. Paul, S. Webb, D. Liu, and G. Dissanayake, "Autonomous robot manipulator-based exploration and mapping system for bridge maintenance," *Robotics and Autonomous Systems*, vol. 59, no. 7-8, pp. 543–554, 2011.

[12] T. Tomic, K. Schmid, P. Lutz, A. Domel, M. Kassecker, E. Mair, I. L. Grixa, F. Ruess, M. Suppa, and D. Burschka, "Toward a fully autonomous uav: Research platform for indoor and outdoor urban search and rescue," *IEEE robotics & automation magazine*, vol. 19, no. 3, pp. 46–56, 2012.

[13] P. Arora and C. Papachristos, "Mobile manipulation–based deployment of micro aerial robot scouts through constricted aperture-like ingress points," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 6716–6723.

[14] P. Arora, S. J. Carlson, T. Karakurt, and C. Papachristos, "Deep-learned autonomous landing site discovery for a tiltrotor micro aerial vehicle," in *2022 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2022, pp. 255–262.

[15] W. Johnson, S. Withrow-Maser, L. Young, C. Malpica, W. Koning, K. WJF, M. Fehler, A. Tuano, A. Chan, A. Datta *et al.*, *Mars Science Helicopter Conceptual Design*. National Aeronautics and Space Administration, Ames Research Center, 2020.

[16] R. D. Lorenz, E. P. Turtle, J. W. Barnes, M. G. Trainer, D. S. Adams, K. E. Hibbard, C. Z. Sheldon, K. Zacny, P. N. Peplowski, D. J. Lawrence *et al.*, "Dragonfly: A rotorcraft lander concept for scientific exploration at titan," *Johns Hopkins APL Technical Digest*, vol. 34, no. 3, p. 14, 2018.

[17] S. J. Carlson, P. Arora, and C. Papachristos, "A multi-vtol modular aspect ratio reconfigurable aerial robot," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 8–15.

[18] S. J. Carlson and C. Papachristos, "Migratory behaviors, design principles, and experiments of a vtol uav for long-term autonomy," *ICRA 2021 Aerial Robotics Workshop on "Resilient and Long-Term Autonomy for Aerial Robotic Systems"*, 2021. [Online]. Available: https://www.aerial-robotics-workshop.com/uploads/5/8/4/4/58449511/icra2021-aerial_paper_8.pdf

[19] S. J. Carlson and C. Papachristos, "The MiniHawk-VTOL: Design, modeling, and experiments of a rapidly-prototyped tiltrotor uav," in *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2021, pp. 777–786.

[20] S. J. Carlson, B. Moore, T. Karakurt, P. Arora, T. Cooper, and C. Papachristos, "The gannet solar–vtol: An amphibious migratory uav for long–term autonomous missions," in *2023 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2023, pp. 419–424.

[21] S. J. Carlson, P. Arora, T. Karakurt, B. Moore, and C. Papachristos, "Towards multi-day field deployment autonomy: A long-term self-sustainable micro aerial vehicle robot," in *2023 International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, p. to appear.

[22] S. J. Carlson, T. Karakurt, P. Arora, and C. Papachristos, "Integrated solar power harvesting and hibernation for a recurrent-mission vtol micro aerial vehicle," in *2022 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2022, pp. 237–244.

[23] S. Khattak, C. Papachristos, and K. Alexis, "Keyframe-based direct thermal–inertial odometry," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 3563–3569.

[24] S. Khattak, C. Papachristos, and K. Alexis, "Keyframe-based thermal–inertial odometry," *Journal of Field Robotics*, vol. 37, no. 4, pp. 552–579, 2020.

[25] S. Khattak, C. Papachristos, and K. Alexis, "Visual-thermal landmarks and inertial fusion for navigation in degraded visual environments," in *2019 IEEE Aerospace Conference*. IEEE, 2019, pp. 1–9.

[26] S. Khattak, C. Papachristos, and K. Alexis, "Marker based thermal-inertial localization for aerial robots in obscurant filled environments," in *Advances in Visual Computing: 13th International Symposium, ISVC 2018, Las Vegas, NV, USA, November 19–21, 2018, Proceedings 13*. Springer International Publishing, 2018, pp. 565–575.

[27] M. Tranzatto, F. Mascarich, L. Bernreiter, C. Godinho, M. Camurri, S. M. K. Khattak, T. Dang, V. Reijgwart, J. Loeje, D. Wisth, S. Zimmermann, H. Nguyen, M. Fehr, L. Solanka, R. Buchanan, M. Bjelonic, N. Khedekar, M. Valceschini, F. Jenelten, M. Dharmadhikari, T. Homberger, P. De Petris, L. Wellhausen, M. Kulkarni, T. Miki, S. Hirsch, M. Montenegro, C. Papachristos, F. Tresoldi, J. Carius, G. Valsecchi, J. Lee, K. Meyer, X. Wu, J. Nieto, A. Smith, M. Hutter, R. Siegwart, M. Mueller, M. Fallon, and K. Alexis, "Cerberus: Autonomous legged and aerial robotic exploration in the tunnel and urban circuits of the darpa subterranean challenge," *Field Robotics*, pp. 274–324, arXiv.2201.07 067, 2021.

[28] X. Wang, Y. Wang, C. Zhou, L. Yin, and X. Feng, "Urban forest monitoring based on multiple features at the single tree scale by uav," *Urban Forestry & Urban Greening*, vol. 58, p. 126958, 2021.

[29] F. Schiefer, T. Kattenborn, A. Frick, J. Frey, P. Schall, B. Koch, and S. Schmidtlein, "Mapping forest tree species in high resolution uav-based rgb-imagery by means of convolutional neural networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 170, pp. 205–215, 2020.

[30] S. Egli and M. Höpke, "Cnn-based tree species classification using high resolution rgb image data from automated uav observations," *Remote Sensing*, vol. 12, no. 23, p. 3892, 2020.

[31] J. Zheng, H. Fu, W. Li, W. Wu, L. Yu, S. Yuan, W. Y. W. Tao, T. K. Pang, and K. D. Kanniah, "Growing status observation for oil palm trees using unmanned aerial vehicle (uav) images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 173, pp. 95–121, 2021.

[32] M. M. Moura, L. E. S. de Oliveira, C. R. Sanquetta, A. Bastos, M. Mohan, and A. P. D. Corte, "Towards amazon forest restoration: Automatic detection of species from uav imagery," *Remote Sensing*, vol. 13, no. 13, p. 2627, 2021.

[33] E. Raparelli and S. Bajocco, "A bibliometric analysis on the use of unmanned aerial vehicles in agricultural and forestry studies," *International Journal of Remote Sensing*, vol. 40, no. 24, pp. 9070–9083, 2019.

[34] P. Lottes, J. Behley, A. Milioto, and C. Stachniss, "Fully convolutional networks with sequential information for robust crop and weed detection in precision farming," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 2870–2877, 2018.

[35] A. K. Mortensen, M. Dyrmann, H. Karstoft, R. N. Jørgensen, R. Gislum *et al.*, "Semantic segmentation of mixed crops using deep convolutional neural network," in *CIGR-AgEng conference*, 2016, pp. 26–29.

[36] C. Potena, D. Nardi, and A. Pretto, "Fast and accurate crop and weed identification with summarized train sets for precision agriculture," in *Intelligent Autonomous Systems 14: Proceedings of the 14th International Conference IAS-14 14*. Springer, 2017, pp. 105–121.

[37] C.-L. Chang, B.-X. Xie, and S.-C. Chung, "Mechanical control with a deep learning method for precise weeding on a farm," *Agriculture*, vol. 11, no. 11, p. 1049, 2021.

[38] L. Zhang, R. Li, Z. Li, Y. Meng, J. Liang, L. Fu, X. Jin, and S. Li, "A quadratic traversal algorithm of shortest weeding path planning for agricultural mobile robots in cornfield," *Journal of Robotics*, vol. 2021, pp. 1–19, 2021.

[39] Z. Wu, Y. Chen, B. Zhao, X. Kang, and Y. Ding, "Review of weed detection methods based on computer vision," *Sensors*, vol. 21, no. 11, p. 3647, 2021.

[40] A. Salazar-Gomez, M. Darbyshire, J. Gao, E. I. Sklar, and S. Parsons, "Towards practical object detection for weed spraying in precision agriculture," *arXiv preprint arXiv:2109.11048*, 2021.

[41] A. Wang, Y. Xu, X. Wei, and B. Cui, "Semantic segmentation of crop and weed using an encoder-decoder network and image enhancement method under uncontrolled outdoor illumination," *Ieee Access*, vol. 8, pp. 81 724–81 734, 2020.

[42] G. Chaudhuri and N. B. Mishra, "Detection of aquatic invasive plants in wetlands of the upper mississippi river from uav imagery using transfer learning," *Remote Sensing*, vol. 15, no. 3, p. 734, 2023.

[43] I. Sa, Z. Chen, M. Popović, R. Khanna, F. Liebisch, J. Nieto, and R. Siegwart, "weednet: Dense semantic weed classification using multispectral images and mav for smart farming," *IEEE robotics and automation letters*, vol. 3, no. 1, pp. 588–595, 2017.

[44] C. McCool, I. Sa, F. Dayoub, C. Lehnert, T. Perez, and B. Upcroft, "Visual detection of occluded crop: For automated harvesting," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 2506–2512.

[45] A. Lucieer, Z. Malenovskỳ, T. Veness, and L. Wallace, "Hyperuas—imaging spectroscopy from a multirotor unmanned aircraft system," *Journal of Field Robotics*, vol. 31, no. 4, pp. 571–590, 2014.

[46] R. Khanna, I. Sa, J. Nieto, and R. Siegwart, "On field radiometric calibration for multispectral cameras," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 6503–6509.

[47] S. Watanabe, K. Sumi, and T. Ise, "Identifying the vegetation type in google earth images using a convolutional neural network: a case study for japanese bamboo forests," *BMC ecology*, vol. 20, pp. 1–14, 2020.

[48] S. L. Middleton, "Automating image segmentation for vegetation monitoring," *Nature Reviews Earth & Environment*, vol. 4, no. 12, pp. 807–807, 2023.

[49] H. Dong, Y. Gao, R. Chen, and L. Wei, "Mangroveseg: Deep-supervision-guided feature aggregation network for mangrove detection and segmentation in satellite images," *Forests*, vol. 15, no. 1, p. 127, 2024.

[50] E. Cassidy, "Sensing invasive species from space."

[51] M. Cabezas, S. Kentsch, L. Tomhave, J. Gross, M. L. L. Caceres, and Y. Diez, "Detection of invasive species in wetlands: practical dl with heavily imbalanced data," *Remote Sensing*, vol. 12, no. 20, p. 3431, 2020.

[52] S. Natesan, C. Armenakis, and U. Vepakomma, "Resnet-based tree species classification using uav images," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 475–481, 2019.

[53] F. Gambella, L. Sistu, D. Piccirilli, S. Corposanto, M. Caria, E. Arcangeletti, A. R. Proto, G. Chessa, and A. Pazzona, "Forest and uav: a bibliometric review," *Contemporary Engineering Sciences*, vol. 9, no. 28, pp. 1359–1370, 2016.

[54] S. Kentsch, M. L. Lopez Caceres, D. Serrano, F. Roure, and Y. Diez, "Computer vision and deep learning techniques for the analysis of drone-acquired forest images, a transfer learning study," *Remote Sensing*, vol. 12, no. 8, p. 1287, 2020.

[55] L. Tang and G. Shao, "Drone remote sensing for forestry research and practices," *Journal of Forestry Research*, vol. 26, pp. 791–797, 2015.

[56] G. Grenzdörffer, A. Engel, and B. Teichert, "The photogrammetric potential of low-cost uavs in forestry and agriculture," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 31, no. B3, pp. 1207–1214, 2008.

[57] J. Walter, J. Edwards, G. McDonald, and H. Kuchel, "Photogrammetry for the estimation of wheat biomass and harvest index," *Field Crops Research*, vol. 216, pp. 165–174, 2018.

[58] L. Kumar and O. Mutanga, "Remote sensing of above-ground biomass, remote sens., 9, 935," 2017.

[59] W. Zhu, Z. Sun, J. Peng, Y. Huang, J. Li, J. Zhang, B. Yang, and X. Liao, "Estimating maize above-ground biomass using 3d point clouds of multi-source unmanned aerial vehicle data at multi-spatial scales," *Remote Sensing*, vol. 11, no. 22, p. 2678, 2019.

[60] M. L. Blatchford, C. M. Mannaerts, Y. Zeng, H. Nouri, and P. Karimi, "Status of accuracy in remotely sensed and in-situ agricultural water productivity estimates: A review," *Remote sensing of environment*, vol. 234, p. 111413, 2019.

[61] X. Lian, H. Zhang, W. Xiao, Y. Lei, L. Ge, K. Qin, Y. He, Q. Dong, L. Li, Y. Han *et al.*, "Biomass calculations of individual trees based on unmanned aerial vehicle multispectral imagery and laser scanning combined with terrestrial laser scanning in complex stands," *Remote Sensing*, vol. 14, no. 19, p. 4715, 2022.

[62] B. Thapa, S. Lovell, and J. Wilson, "Remote sensing and machine learning applications for aboveground biomass estimation in agroforestry systems: a review," *Agroforestry Systems*, pp. 1–15, 2023.

[63] D. Yu, Y. Zha, Z. Sun, J. Li, X. Jin, W. Zhu, J. Bian, L. Ma, Y. Zeng, and Z. Su, "Deep convolutional neural networks for estimating maize above-ground biomass using multi-source uav images: A comparison with traditional machine learning algorithms," *Precision Agriculture*, vol. 24, no. 1, pp. 92–113, 2023.

[64] W. Castro, J. Marcato Junior, C. Polidoro, L. P. Osco, W. Gonçalves, L. Rodrigues, M. Santos, L. Jank, S. Barrios, C. Valle *et al.*, "Deep learning applied to phenotyping of biomass in forages with uav-based rgb imagery," *Sensors*, vol. 20, no. 17, p. 4802, 2020.

[65] S. Oehmcke, L. Li, J. C. Revenga, T. Nord-Larsen, K. Trepekli, F. Gieseke, and C. Igel, "Deep learning based 3d point cloud regression for estimating forest biomass," in *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*, 2022, pp. 1–4.

[66] L. Pan, L. Liu, A. G. Condon, G. M. Estavillo, R. A. Coe, G. Bull, E. A. Stone, L. Petersson, and V. Rolland, "Biomass prediction with 3d point clouds from lidar," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 1330–1340.

[67] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4758–4765.

[68] C. Brommer, R. Jung, J. Steinbrener, and S. Weiss, "MaRS : A Modular and Robust Sensor-Fusion Framework," 2020.

[69] M. Kamel, T. Stastny, K. Alexis, and R. Siegwart, "Model predictive control for trajectory tracking of unmanned aerial vehicles using robot operating system," in *Robot Operating System (ROS) The Complete Reference, Volume 2*, A. Koubaa, Ed. Springer.

[70] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, and J. Nieto, "Voxblox: Incremental 3d euclidean signed distance fields for on-board mav planning," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 1366–1373.

[71] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron2," https://github.com/facebookresearch/detectron2, 2019.

[72] K. Wada, "Labelme: Image Polygonal Annotation with Python." [Online]. Available: https://github.com/wkentaro/labelme