ORIGINAL RESEARCH **Open access**

Detection-Uncertainty-Guided UAV Path Planning for Agricultural Field Search Operations

Chinedu Okafor[1*], Amina Bello[1], Ibrahim Musa[2], Grace Adeyemi[3]

Abstract

This work presents an adaptive path planner for locating objects in agricultural fields using unmanned aerial vehicles. The approach relies on a high-altitude reconnaissance route, scheduling further low-altitude checks only when the detection network's certainty falls short. An offline simulation framework incorporating real-world imagery was used to test the planner. To demonstrate its capabilities, a YOLOv8 detection network was trained to spot artificial plants placed within grassy areas. Different measures of detection certainty were compared, path-planning parameters were fine-tuned, and the effects of localization errors across varying object counts were studied. Among the options, YOLOv8 detection confidence proved most effective at separating true from false positive detections, and it was selected for use in the adaptive planner. The planner's optimal settings depended on the spatial arrangement of objects in the field. Uniformly spread objects required more low-altitude inspections than non-uniformly spread ones, resulting in a longer total path. The planner displayed strong resilience to localization uncertainty. A higher number of objects led to an increase in flight path length, most notably for uniformly distributed objects. With a non-uniform layout, the adaptive planner consistently generated a shorter route than a low-altitude coverage pattern, even at high object densities. In summary, the introduced adaptive planner located non-uniformly scattered objects in a field faster than a standard coverage planner while attaining a comparable level of detection accuracy. The planner's source code is available at [https://github.com/wur-abe/uav_adaptive_planner].

Keywords Object detection, Adaptive path planning, Detection certainty, Unmanned aerial vehicles, Drones

*Correspondence:
Chinedu Okafor
chinedu.okafor@unn.edu.ng
[1] Department of AI for Cassava Disease Detection and Yield Forecasting, Faculty of Agriculture, University of Nigeria, Nsukka, Nigeria
[2] Department of Smart Farming Data Analytics, Faculty of Agriculture, University of Ibadan, Ibadan, Nigeria
[3] Department of Precision Weed Mapping, Faculty of Agriculture, Federal University of Agriculture Abeokuta, Abeokuta, Nigeria

Introduction

The adoption of Unmanned Aerial Vehicles (UAVs) in agriculture has expanded markedly over recent years [1]. Tasks vary from yield estimation [2] to identifying weeds [3] and diseases [4, 5]. Conducting these applications involves pinpointing objects such as weeds or disease symptoms across extensive fields. Conventionally, the UAV executes a predefined low-altitude coverage route at a constant height [6]. While simple to execute and providing uniform spatial detail across the field, such full coverage is redundant for object detection specifically. Operating at a higher elevation broadens the UAV's field of view, shrinking both the total flight distance and its duration, but it sacrifices spatial detail and, consequently, detection accuracy. By descending for a closer look only when detections are uncertain, it becomes possible to cut flight time while preserving the accuracy of a full low-altitude survey. Optimising UAV battery usage hinges

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

on reducing path length without sacrificing detection quality. Curbing flight time is especially critical when UAVs perform autonomous, repeated field monitoring, as it directly expands the area one UAV can oversee.

Adaptive path planning, also known as informative path planning, seeks to design a trajectory that maximises the collection of relevant information while respecting resource constraints, such as battery life [7]. This is accomplished by making the route responsive to live data and informing subsequent decisions about the UAV's movement using UAV-based adaptive planning. This approach has been employed for tasks such as anomaly spotting in ocean topography [8], weed scouting in agriculture [9], and multi-resolution semantic segmentation [10]. Popović *et al*. [9] found that an adaptive planner reduced map uncertainty (quantified by entropy) by 50% over the same period for active weed classification (simulated via printed markers) compared with a low-altitude coverage route. Stache *et al.* [10] designed a segmentation path planner utilising a modelled link between altitude and segmentation quality. It performs detailed inspections only at reduced altitudes when necessary, thereby shortening flight time. The demonstration involved segmenting soil, plants, and weeds via simulated flight over the WeedMap dataset, and features such as asphalt, buildings, and water via simulated flight over satellite imagery, achieving high segmentation accuracy with minimal flight time. van Essen *et al.* [11] describe a learning-driven adaptive planner that applies Reinforcement Learning to locate target objects in a field. The learnt policy navigates the UAV by fusing uncertain field priors with detections inside the current field of view. Simulation tests proved the learnt route could find 80% of objects in under half the flight time required by a coverage planner. Nevertheless, all these methods sidestep the certainty estimates directly offered by detection or segmentation networks, despite their potential usefulness for guiding adaptive planning decisions.

Raising the UAV's altitude widens its field of view but coarsens the image's spatial resolution. With fewer pixels depicting an object, both false-positive and false-negative rates increase, harming accuracy. An adaptive planner can exploit this altitude–accuracy trade-off by capturing detailed low-altitude imagery only on demand. For example, at a high altitude, the network might flag an item as a potential weed. The attached uncertainty score allows a choice: accept the finding as solid or trigger a new, lower-altitude image.

Detection networks usually issue a confidence score alongside each detection as a proxy for certainty. When that confidence dips, an adaptive planner might arrange a dedicated low-level inspection. Yet neural networks are known to miscalibrate confidence estimates [12], rejecting valid detections while keeping erroneous ones. Several strategies seek to refine these uncertainty estimates. Implementing Monte Carlo Dropout (MCD) at inference is a common choice due to its straightforward adoption [12]. Myojin *et al.* [13] noted a 1%–2% accuracy gain when employing MCD-based uncertainty measures for spotting lunar craters with an adapted YOLOv3. Within agriculture, MCD has supported uncertainty quantification for active learning in broccoli identification [14], fish species identification [15], obstacle recognition at solar farms [16], and evaluation of chicken plumage condition [17]. These works confirm MCD's value for gauging network uncertainty.

This paper aims to develop and test an adaptive path planner that discovers objects in agricultural fields more rapidly than a low-altitude coverage planner while maintaining detection performance. The method follows a high-altitude coverage route and exploits detection certainty to command low-altitude inspections when confidence is lacking. Unlike prior literature, the designed planner integrates uncertainty outputs directly from an object detection network, broadening its adaptability for other uses. Aside from field borders, no prior field knowledge is necessary, and the system's performance was assessed on genuine data with a real detection network. We have also publicly released the code at [https://github.com/wur-abe/uav_adaptive_planner]. Concretely, we (1) probed the effect of different certainty estimation techniques computed via MCD, (2) optimised the planner's flight altitude and certainty thresholds, (3) scrutinised the influence of UAV localisation errors, and finally, (4) studied the planner's efficiency relative to the number of objects in the field.

Materials and Methods

© 2022 The Author(s). | Published by Agro Intelligence Research Press

This section outlines the adaptive path planner (section 2.1), the techniques for estimating detection uncertainty (section 2.2), the data gathering process (section 2.3), the detection network training regime (section 2.4), the offline simulation setup (section 2.5), the reference baseline (section 2.6), the evaluation metrics (section 2.7), and lastly the experiments (section 2.8).

Adaptive path planner

Figure 1 illustrates the logic of the proposed adaptive path planner. The workflow begins with an operator manually outlining the flight zone the UAV needs to survey. From this designated zone, a high-altitude sweep route is computed (section 2.1.1) and subsequently flown. At every waypoint along this route, a downward-facing (nadir) image is taken, and a detection network processes it on the fly to locate objects (section 2.1.2). The objects spotted are then assigned geographic coordinates (section 2.1.3) and recorded on a dynamic map (section 2.1.4). Once the high-altitude sweep concludes, a secondary inspection route is generated, placing waypoints at a lower altitude above locations where object detections lacked confidence (section 2.1.5). Redundant, overlapping waypoints are filtered out to keep the inspection list manageable. The inspection route is carried out following the same operational pattern as the high-altitude sweep. As the inspection flight unfolds, the map is continuously refined: existing entries are updated, and newly encountered objects are added as needed. Ultimately, any object location stored in the map (whether originating from the high-altitude or low-altitude phase) with a certainty value greater than ceval is retained for final assessment.

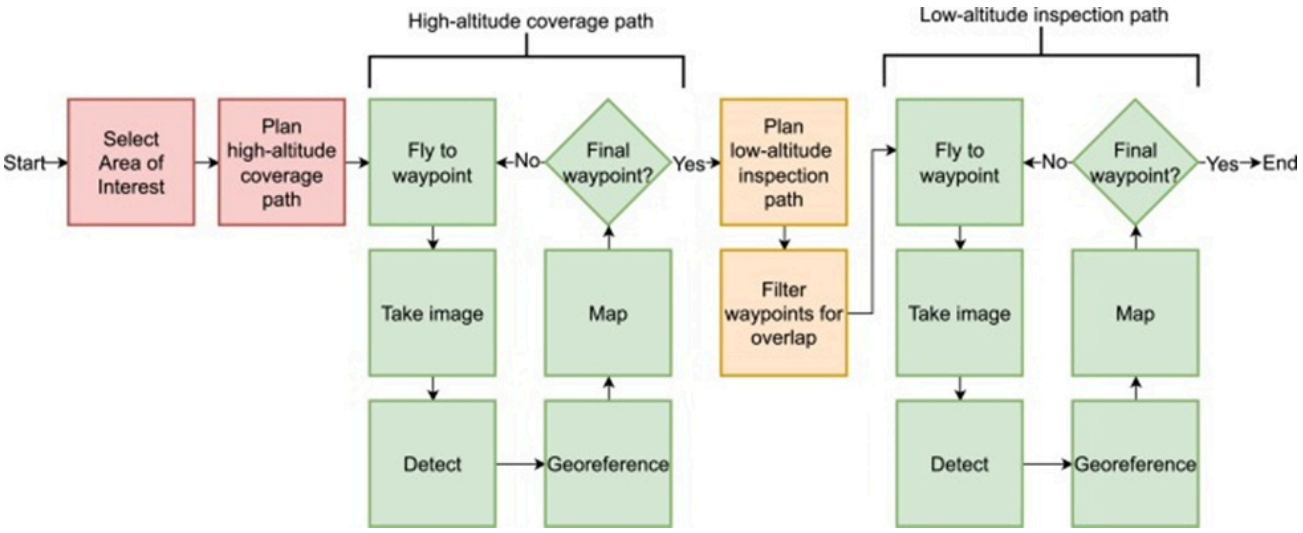


Figure 1. Block diagram of the adaptive path planner's main components: generating the high-altitude sweep route (red, section 2.1.1), generating the low-altitude inspection route (orange, section 2.1.5), and performing the flight routes (green).

Planning of high-altitude coverage path

The process starts with a manual step: drawing a polygon on a satellite image to mark the field's precise boundaries. Inside this polygon, a coverage flight path is produced at a height of hcov using the Fields2Cover library [18]. The field-of-view, FOV(h) (m), corresponding to a given altitude h, is worked out as follows:

$$\mathrm{FOV}(h) = \frac{\mathbf{c} \odot h}{f}, \quad (1)$$

© 2022 The Author(s). | Published by Agro Intelligence Research Press

Where ⊙ stands for element-wise multiplication, $\boldsymbol{c} = [cw, ch]^T$ captures the sensor's physical width and height (mm), and f is the lens focal length (mm). In practice, UAV localization is never perfect; relying on the raw field-of-view as the row spacing would risk skipping strips of the field. To build some tolerance against these positioning errors, the row width is reduced to 90% of the field of view, effectively introducing a 5% field-of-view buffer on either side. In the same vein, the forward spacing between consecutive waypoints along the track is also clamped to 90% of the field of view. **Figure 2** presents example coverage patterns for altitudes of 12 m, 24 m, 36 m, and 48 m.

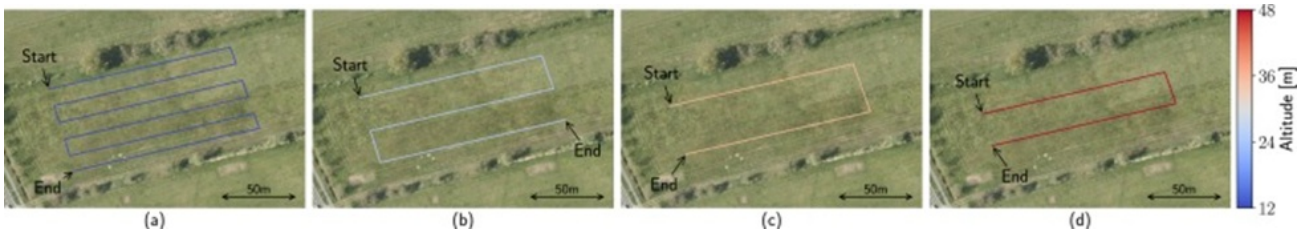


Figure 2. Sample coverage flight paths computed for altitudes of 12 m (a), 24 m (b), 36 m (c), and 48 m (d).

Detection

At every waypoint, the UAV captures a single nadir image. Objects within this image are identified by YOLOv8 [19], the eighth generation of the YOLO (You Only Look Once) family of fully convolutional object detectors, which predicts bounding boxes and class probabilities together in one evaluation. The nano configuration of YOLOv8 is selected to keep the computational footprint light, allowing the network to run directly on UAV hardware. During model inference, images are supplied at 2048x2048 pixel resolution, and the lowest permissible confidence is set to creject.

Georeferencing

Georeferencing of detected objects happens immediately, driven by the UAV's real-time position and heading. The first step is recalculating the field-of-view, FOV(h), at the UAV's current height via **Eq. 1**. The pixel coordinates of a detected object o, denoted $[uo, vo]^T$, are converted into local camera-frame coordinates, $\left[x_o^{local}, y_o^{local}\right]^T$ (m), through:

$$\begin{bmatrix} x_o^{\text{local}} \\ y_o^{\text{local}} \end{bmatrix} = \left(\begin{bmatrix} u_o \\ v_o \end{bmatrix} \odot \begin{bmatrix} \frac{1}{I_w} \\ \frac{1}{I_h} \end{bmatrix} - \frac{1}{2} \right) \odot \text{FOV}(h), \tag{2}$$

Where Iw and Ih represent the image's width and height, respectively. Afterwards, these local coordinates are transformed into Universal Transverse Mercator (UTM) world coordinates, $\left[x_o^{utm}, y_o^{utm}\right]^T$ (representing Easting and Northing), via an affine rotation:

$$\begin{bmatrix} x_o^{\text{utm}} \\ y_o^{\text{utm}} \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \cdot \begin{bmatrix} x_o^{\text{local}} \\ -y_o^{\text{local}} \end{bmatrix} + \begin{bmatrix} x_{\text{uav}}^{\text{utm}} \\ y_{\text{uav}}^{\text{utm}} \end{bmatrix}, \tag{3}$$

© 2022 The Author(s). | Published by Agro Intelligence Research Press

Here, ψ is the counter-clockwise heading of the drone measured from north, and $[x_{uav}^{utm}, y_{uav}^{utm}]^T$ provides the UAV's own UTM position. A sign inversion is applied to y_o^{local} because UTM is oriented with north upwards.

Mapping

Throughout the mission, the GNSS coordinates associated with each detection are stored in a persistent map. Because consecutive waypoints overlap and the UAV revisits areas at different altitudes, the same physical object may generate several observations. To fuse these, any detections whose Euclidean separation falls below ddist are treated as belonging to a single object. The default value ddist = 0.35 m was chosen based on the smallest physical gap between the placed objects in the field.

If a previously mapped object does not register in the current field-of-view when it should have been visible, the system flags it as a false positive and deletes it from the map, provided both of the following hold:

• The UAV's current altitude is lower than the minimum altitude among all past sightings of that object,

• And the highest certainty ever attached to that object is less than caccept.

When merging a newly arrived detection into the map, the rules are:

• If the UAV altitude linked to the new detection is lower than the minimum altitude of any prior observation of that mapped object, the object's stored certainty, class label, and coordinates are fully replaced by the new detection's values,

• Otherwise, if the new detection's certainty exceeds the certainty already stored for that object, the stored certainty, class label, and coordinates are overwritten by the new detection's values.

Planning of low-altitude inspection flight path

After the high-altitude sweep is complete, an inspection flight is organized to provide a closer, higher-resolution view of map objects whose certainty remains unconvincing (as defined in section 2.2). For each mapped object i, a decision Di among three possible actions—accept without further checks, reject entirely, or schedule a close-up inspection—is taken according to:

$$\mathscr{D}_i = \begin{cases} \text{accept} & c_i > c_{\text{accept}} \\ \text{reject} & c_i < c_{\text{reject}} \\ \text{inspect} & \text{otherwise} \end{cases}, \quad (4)$$

Where ci is the certainty assigned to the object (section 2.2), caccept acts as the acceptance bar, and creject acts as the rejection bar. Waypoints for an inspection route at altitude hinspect are then compiled for all objects whose Di outcome is 'inspect'. This unordered collection of waypoints is arranged into the shortest possible visiting sequence via iterative stochastic local search, minimizing the total route length [20]. Each waypoint is assigned a heading pointing to the next. To further condense the route, the planner iteratively prunes waypoints: a waypoint is removed if the object it targets is forecast to already lie within the field-of-view of the preceding waypoint. **Figure 3** gives an example of an inspection route constructed in this way.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

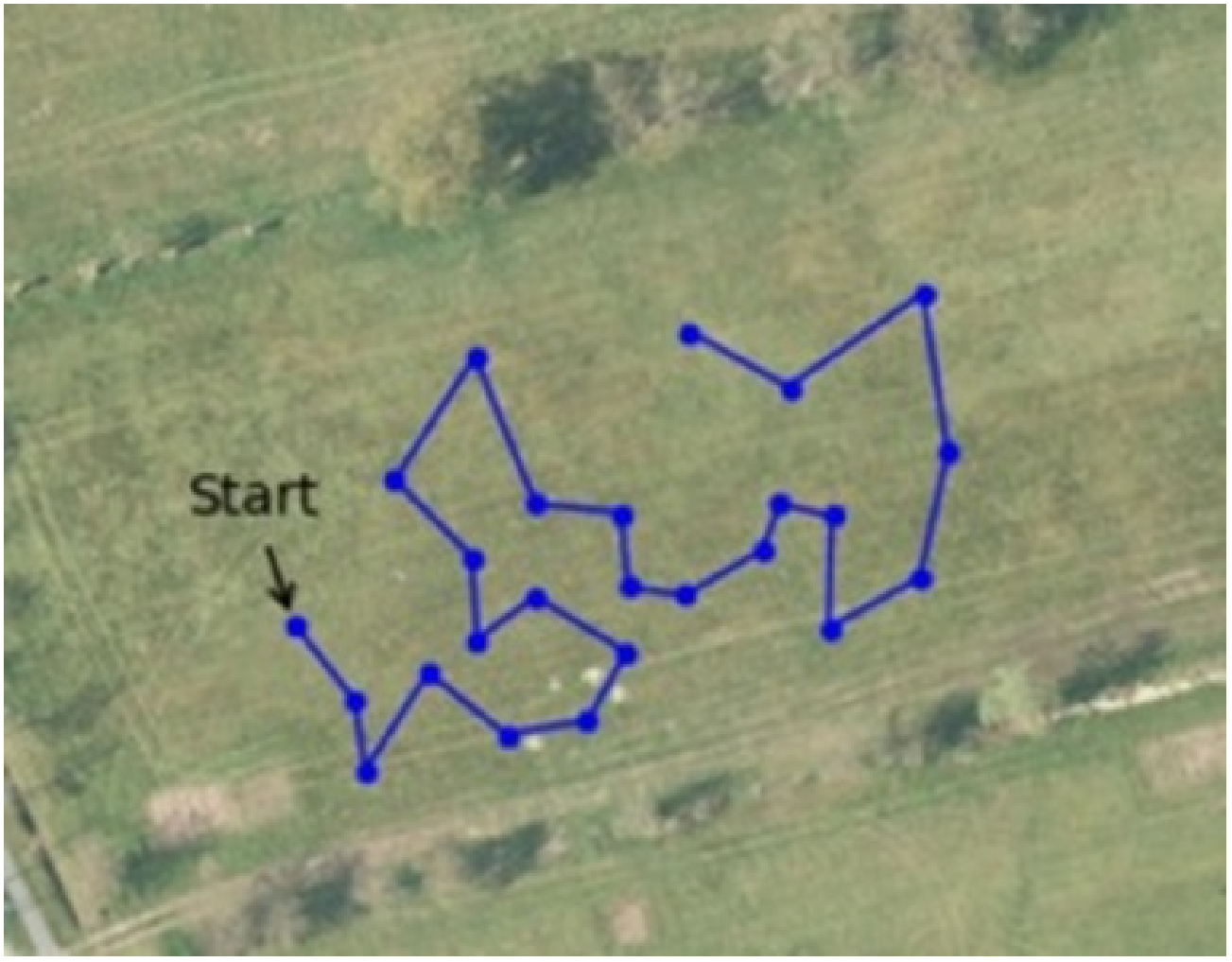


Figure 3. Illustration of an inspection route that visits uncertain detections to obtain imagery with greater spatial detail at the marked waypoints (blue dots). The computed low-altitude inspection route is shown in blue.

Detection certainty measures

The adaptive path planner detailed above makes its routing decisions by consulting the certainty associated with each detection, directing the UAV to re-examine predictions with low certainty that might constitute errors. This operational logic hinges on the existence of a genuine statistical relationship between certainty scores and network mistakes, since the system treats certainty as an indicator of whether a real object occupies that location. In our study, six distinct formulations for quantifying certainty are benchmarked against one another. The most straightforward candidate is the raw detection confidence that the network assigns to its highest-probability class, denoted as cyolo. Internally, this scalar combines the model's confidence in bounding box placement and size with its confidence in the object's identity. A widely documented limitation, however, is that neural network confidence estimates tend to suffer from systematic miscalibration [12]. Consequently, we derive five additional certainty metrics based on Monte Carlo dropout (MCD). Producing these MCD-based certainties involves two sequential phases: injecting MCD into the YOLOv8 inference pipeline (section 2.2.1), followed by the computation of the actual certainty values (section 2.2.2).

Monte-Carlo dropout

Dropout functions as a stochastic regularization technique that randomly prunes inter-neuron connections and is conventionally activated during training to suppress overfitting by forcing the network to develop redundant internal representations that reach identical conclusions through different routes [21]. Gal and Ghahramani [22] later proposed

© 2022 The Author(s). | Published by Agro Intelligence Research Press

retaining dropout during inference and showed that this Monte Carlo dropout procedure can serve as a window into the model's epistemic uncertainty. For any connection governed by dropout, the probability of removal equals q, and the entire stochastic process repeats over N independent forward evaluations. Near-identical outputs emerging from all N passes, even as connections are arbitrarily disabled, imply the network holds a firm belief in its output. On the flip side, major discrepancies between runs betray a model that lacks a clear interpretation of the input.

Figure 4 presents a condensed view of the YOLOv8 design, delineating the backbone, neck, and head segments, along with the pathways that relay information from earlier stages. The marked crosses signify the positions where dropout is inserted. By confining dropout to the interface between the neck and the head, the backbone and neck activations need only be calculated once, yielding computational savings during the repeated stochastic passes.

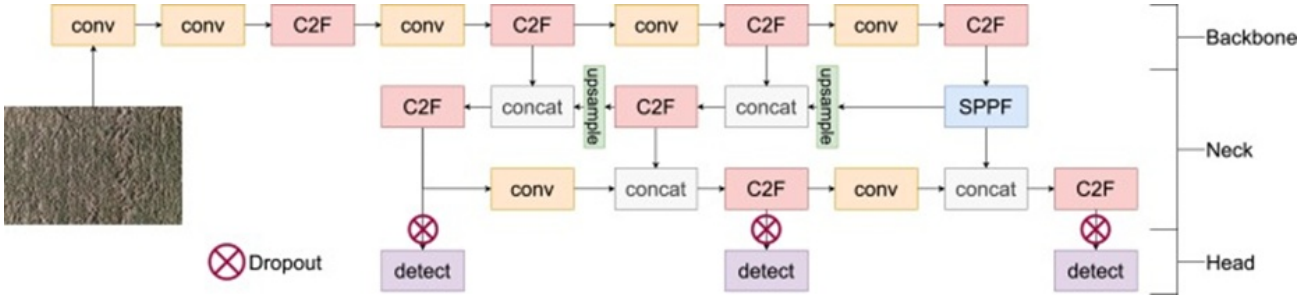


Figure 4. Condensed structural overview of YOLOv8, pinpointing the injection sites for Monte-Carlo dropout between the neck and head.

A standard 'conv' module packages a 2D convolution, batch normalization, and sigmoid nonlinearity. 'C2F' labels a Cross-Stage Partial unit housing two convolutional layers. 'SPPF' indicates a Spatial Pyramid Pooling module, and 'concat' refers to a feature concatenation step. The 'detect' module chains two 'conv' units, followed by a final 2D convolution that produces the detection list. See Jocher *et al.* [19] for an exhaustive account.

When processing a specific image i, a single conventional forward pass without any dropout yields the reference set of predictions Drefi, while N=20 additional passes run with dropout active at probability q=0.25. Their outputs are aggregated into a detection set Si:

$$S_i = \{D_{1,i}, D_{2,i}, \ldots, D_{N,i}\}, \quad (5)$$

Where Dj,i describes what the j-th MCD run predicts for object i. The record Dj,i captures bounding-box geometry, class-probability vectors, and a scalar confidence. To establish correspondence between these stochastic detections and the reference detection D_i^{ref} we evaluate the intersection-over-union (IoU) between the reference box and every MCD box $D_{j,i}^{\mathrm{mc}}$ for pass j:

$$\mathrm{IoU}(D_i^{\mathrm{ref}}, D_{j,i}^{\mathrm{mc}}) = \frac{|D_i^{\mathrm{ref}} \cap D_{j,i}^{\mathrm{mc}}|}{|D_i^{\mathrm{ref}} \cup D_{j,i}^{\mathrm{mc}}|}. \quad (6)$$

Any MCD detection D_(j,i)^mc whose IoU with the reference surpasses 0.5 earns membership in Si. Crucially, this matching step ignores the class label entirely and operates solely on spatial overlap.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

Certainty calculation

Beyond the native detection confidence, five extra per-detection certainty metrics are extracted from the detection sets: the averaged YOLO confidence, occurrence certainty, localization certainty, class certainty, and a multiplicative combined certainty.

Averaged YOLO certainty:

The averaged YOLO certainty c_i^{yolo} tied to reference detection i is the arithmetic mean of the YOLO confidence figures populating its detection set Si:

$$c_i^{\text{yolo}} = \frac{1}{\|S_i\|} \cdot \sum_{j=1}^{\|S_i\|} c_{j,i}^{\text{yolo}}, \quad (7)$$

With $c_{j,i}^{\text{yolo}}$ representing the YOLOv8 confidence attributed to the j-th element of that set. In the case where Si contains no entries, c_i^{yolo} takes the value 0.0.

Occurrence certainty:

The occurrence certainty c_i^{occ} assigned to reference detection i is given by:

$$c_i^{\text{occ}} = \frac{\|S_i\|}{N}. \quad (8)$$

Location certainty:

The location certainty c_i^{loc} for reference detection, i equates to the mean IoU computed between the reference box and all members of Si:

$$c_i^{\text{loc}} = \frac{1}{\|S_i\|} \cdot \sum_{D_{j,i}^{\text{mc}} \in S_i} \text{IoU}(D_i^{\text{ref}}, D_{j,i}^{\text{mc}}). \quad (9)$$

This formulation draws on the methodology of Blok *et al.* [14]. For an empty Si, c_i^{loc} is defined as 0.0.

Class certainty:

Deriving the class certainty for reference detection i begins with computing the entropy Hi:

© 2022 The Author(s). | Published by Agro Intelligence Research Press

$$H_i(k) = -\sum_{j=1}^{\|S_i\|} P_{j,i}(k) \cdot \ln P_{j,i}(k), \quad (10)$$

Where Pj,i(k) specifies the probability allocated to class k∈K by detection j residing in Si. These entropy figures are subsequently rescaled to fall within [0,1] through the softmax mapping σi(k):

$$\sigma_i(k) = \frac{e^{H_i(k)}}{\sum_{k_z \in K} e^{H_i(k_z)}}. \quad (11)$$

The class certainty c_i^{cls} is then read off as the peak value of this normalized distribution across the set of classes:

$$c_i^{\text{cls}} = \max_{k \in K} \left(\sigma_i(k) \right). \quad (12)$$

The per-class probabilities P(k) originate from running the softmax function (Equation (11)) over the raw class activation outputs taken from the network's last layer.

Combined certainty:

The combined certainty c_i^{comb} for reference detection i is constructed by forming the product of the three preceding components—occurrence, location, and class:

$$c_i^{\text{comb}} = c_i^{\text{occ}} \cdot c_i^{\text{loc}} \cdot c_i^{\text{cls}}. \quad (13)$$

A detection that simultaneously exhibits high occurrence, precise localization, and consistent class assignment attains a higher combined score. At the same time, degradation in any of the three factors lowers the combined score accordingly.

Data collection

Aerial imagery was captured over a grassy plot of roughly 0.75 hectares at the Wageningen University and Research experimental station, in the vicinity of Wageningen, The Netherlands. To gather repeatable observations spanning the entire growing season, we mimicked a weed scouting use case by distributing two different artificial plant types on the grass surface **(Figure 5)**. Viewed from directly overhead, each plant spans approximately 10 cm by 10 cm, and no two plants were placed closer than 1 m apart. Data acquisitions took place on four separate occasions (13 February 2024, 23 April 2024, 18 July 2024, and 1 August 2024) with the explicit goal of strengthening the detection network's ability to generalize. Because certain sections of the field could not be accessed on every date, subtle differences in field perimeter and total surface area exist from one dataset to the next.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

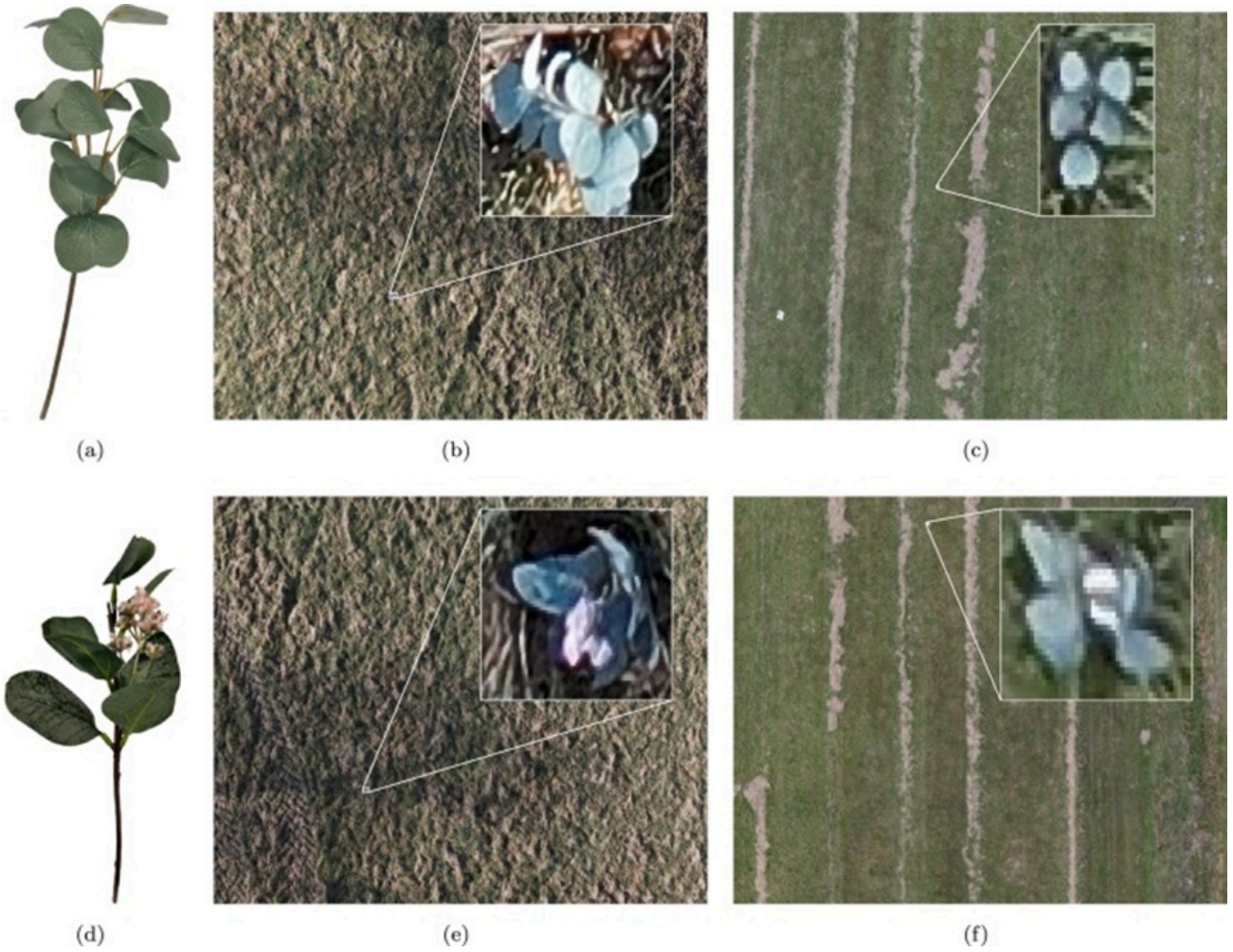
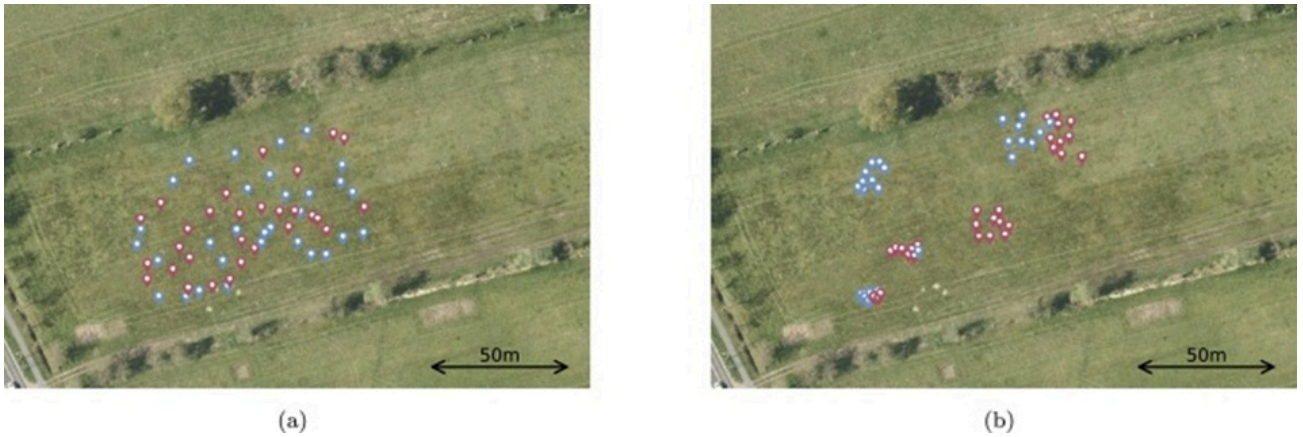


Figure 5. Sample appearances of weed type I (a–c) and weed type II (d–f) recorded at 12 m above ground (b, e) and 32 m above ground (c, f), drawn from the 13 February 2024 collection (b, c) and the 1 August 2024 collection (e, f). The white blossom characteristic of weed type II is apparent.

Given that real agricultural weeds seldom distribute evenly across a field [23]. Since that spatial configuration can dramatically shape an adaptive planner's performance [11], we explicitly studied the consequences of different spatial arrangements. During each of the four field campaigns, two layouts were used: one with the artificial plants uniformly positioned, and another with them grouped into clusters. **Figure 6** shows the layouts for the 13 February 2024 campaign.



© 2022 The Author(s). | Published by Agro Intelligence Research Press

Figure 6. RTK-surveyed coordinates of the artificial plants set out in a uniform arrangement (a) and a clustered arrangement (b) during the 13 February 2024 data collection. Different colors encode the two weed types.

All images were acquired using a Matrice 300 RTK UAV carrying a Zenmuse P1 camera (DJI, Shenzhen, China) fitted with a 35 mm lens. Image acquisition parameters specified 80% forward overlap and 70% side overlap, with each raw frame measuring 8192×5460 pixels. The UAV flew at three different altitudes: 12 m, 24 m, and 32 m. In addition, eight ground control points (GCPs) were distributed throughout the field to enhance the georeferencing fidelity of the orthomosaics, as described further in section 2.5. To produce ground truth, the locations of both the artificial plants and the GCPs were surveyed with a Topcon HiPer SR RTK-GNSS receiver (Topcon Positioning Systems, Inc., Tokyo, Japan), achieving an accuracy of approximately 1.0–1.5 cm.

Training of the detection network

For each unique pairing of date and spatial distribution, a randomly selected 10% portion of the plants (equating to 6 plants per date per distribution) was earmarked for the training partition. In comparison, 5% (3 plants per date per distribution) was held back for validation. By design, the large majority of plants were never exposed to the network during training—an essential precondition for fairly testing the adaptive planner. Because of heavy image overlap, each plant was visible in multiple frames, so every UAV image (covering the 12 m, 24 m, and 32 m flight levels) that showed at least one training or validation plant was extracted and annotated by hand. To erect a strict barrier against data leakage, any plants that appeared in training images but did not belong to the training partition had their bounding boxes filled with solid green pixels, effectively masking them from the network. The same green-mask procedure was applied to the validation images for plants outside the validation partition. Training was conducted jointly on the full dataset, aggregating all dates, altitudes, and distribution types to maximize generalization. The pooled training corpus contained 1618 images with 2000 annotations, while the pooled validation corpus contained 889 images with 977 annotations.

The network ingested inputs at a resolution of 2048×2048 pixels, and all other hyperparameters were kept at their factory defaults. Optimization ran for 250 epochs using a batch size of 16 on a machine configured with an AMD Ryzen 5950X CPU and an NVIDIA RTX 3090 GPU.

Offline simulation environment

To support extensive experimentation and methodical parameter sweeping with the adaptive planner, the collected aerial images were stitched into georeferenced orthomosaics, yielding a high-fidelity offline simulation sandbox. For every combination of date and plant distribution, a dedicated orthomosaic was reconstructed from the 12 m altitude captures with Agisoft Metashape v2.1.0 (Agisoft LLC, St. Petersburg, Russia). The eight GCPs were integrated into the reconstruction workflow to sharpen georeferencing precision. Altogether, eight separate orthomosaics were generated, each processed using the software’s high-quality preset.

When emulating a specific waypoint, the UAV’s instantaneous field of view was first computed. The orthomosaic was rotated to align with the UAV’s heading, and the footprint corresponding to that field of view was cropped and centered on the UAV’s GNSS position. The cropped patch was then resampled to match the original capture dimensions via linear interpolation. The resulting synthetic frames bore a close visual resemblance to the genuine photographs originally acquired by the UAV.

Baseline

We contrast the proposed adaptive scheme against a standard full-coverage planner constructed via Fields2Cover [18], which traverses the entire field while holding altitude constant. Identical to the settings applied in the adaptive planner, the spacing between adjacent rows and the gap separating consecutive waypoints along the track were both set to 90% of the sensor’s field of view.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

Performance metrics

Two criteria were used to judge both planners: the quality of the resulting detections and the overall length of the flown route. Any mapped object whose associated certainty lies above c_{eval} = 0.5 was checked against the field-surveyed GNSS ground truth. A mapped object counted as a true positive (TP) provided its predicted class matched the true class, and the Euclidean offset from the surveyed coordinate did not exceed 0.35 m. When a mapped object had no counterpart in the ground truth set, it was labeled a false positive (FP); ground truth objects missing from the map were scored as false negatives (FN). The F1-score encapsulates detection quality.

Route length was obtained by summing the Euclidean distances between consecutive waypoints. Given that the various datasets differed in both outline and surface area, a normalized distance metric r_k^{diff} was introduced to relate the adaptive route to its coverage baseline:

$$r_k^{\text{diff}} = \frac{l_k}{l_k^{\text{cov}}}, \quad (14)$$

Here, lk stands for the length of the adaptive route on dataset k and l_k^{cov} for the length of the corresponding 12 m altitude coverage baseline. Values of r_k^{diff} A value above 1.0 indicates the adaptive path exceeds the baseline in length; a value below 1.0 indicates it is shorter.

Experiments

Experiment 1: relationship between detection certainty and errors

The certainty figure attached to a detection constitutes a cornerstone of the adaptive planning logic, as it dictates whether a low-level re-inspection will be scheduled. Consequently, for the decision mechanism to function sensibly, the certainty metric must afford a genuine separation between accurate and inaccurate predictions. To examine how effectively the different certainty formulations (laid out in section 2.2.2) discriminate TPs from FPs at varying heights above ground, we scored them on the original aerial images taken at 12 m, 24 m, and 32 m across all 8 datasets. Any detection whose positional error fell within 0.35 m of the image perimeter was discarded from the analysis; in such border cases, the actual object might reside just outside the frame due to slight mismatches between the image's GNSS tag and the surveyed object's coordinate, which would skew the evaluation. For every candidate certainty metric, we plot how its values are distributed among TPs and FPs at the three altitudes. A wide separation between the two groups indicates that the metric can reliably distinguish reliable detections from questionable ones. To formally gauge this degree of separation, we run Welch's t-test [24] for the one-sided alternative hypothesis that the mean certainty of TPs, μ_{tp}, exceeds that of FPs, μ_{fp}:

$$t = \frac{\mu_{\text{tp}} - \mu_{\text{fp}}}{\sqrt{\left(\frac{\sigma_{\text{tp}}}{\sqrt{N_{\text{tp}}}}\right)^2 + \left(\frac{\sigma_{\text{fp}}}{\sqrt{N_{\text{fp}}}}\right)^2}}, \quad (15)$$

With N denoting the sample count and σ the standard deviation. A larger t-statistic indicates that the measure is better at distinguishing true from false positives. At each altitude, the metric delivering the highest t-value was selected for use in experiments 2, 3, and 4.

Experiment 2: parameter optimization

© 2022 The Author(s). | Published by Agro Intelligence Research Press

The behavior of the adaptive planner pivots on three key parameters: the altitude assigned to the coverage sweep (hcov), the threshold for outright acceptance (caccept), and the threshold for outright rejection (creject). These quantities are interdependent—for example, operating the coverage phase at a higher altitude is most effective when paired with a relatively low rejection threshold, because detection confidence naturally decreases with increasing altitude. A strict rejection threshold would then discard a large share of detections, thereby inflating the false-negative count. To map out how these parameters jointly shape the planner's performance, we swept hcov from 12 m through 48 m in 12 m steps, caccept from 1.0 down to 0.4 in 0.2 increments, and tested creject at 0.05, 0.2, and 0.4. Every combination was examined, except for caccept = 0.4 and creject = 0.4, as this pairing could never trigger the decision to inspect (refer to section 2.1.5). The F1-score and the relative difference in route length compared to the baseline coverage path were computed for each combination by running the planner on the clustered and uniform orthomosaics. The parameter configuration that concurrently offered strong detection results and a comparatively compact route was carried forward to experiments 3 and 4. Throughout all trials, the inspection height, hinspect, was set to 12 m, corresponding to the lowest altitude used during field data collection. For reference, we also assessed the F1-score and relative route length of the coverage planner at the same coverage altitudes (hcov).

Experiment 3: localization errors

Flaws in estimating the UAV's position and orientation propagate directly into the georeferencing of the objects spotted in the imagery. These georeferencing discrepancies can stem from GNSS-derived errors in horizontal position and altitude, as well as from IMU-based errors in roll, pitch, and heading estimates. To probe how localization inaccuracies influence the F1-score, five graded levels of error were tested. **Table 1** lists the maximum allowable offsets in position, altitude, gimbal roll, gimbal pitch, and UAV heading for each level, sampled from uniform distributions. The DJI M300 platform used for data gathering exhibited roughly 1.5 cm of horizontal and vertical positional error within the test field (per the manufacturer's specifications) and approximately 0.5 degrees of error in roll, pitch, and heading (the mean georeferencing residual recorded during orthomosaic construction), which aligns with the level designated 'good' in **Table 1**. These positioning and orientation errors were emulated by perturbing the crop location and heading within the offline simulation framework. Because larger errors disrupt the mapping process and cause separate map entries to be spawned for what is in reality a single object (section 2.1.4), we relaxed the merging distance ddist to 0.5 m, 0.9 m, 1.8 m, and 4.7 m for the uncertainty levels good, decent, poor, and very poor, respectively. The perfect level retained the original 0.35 m value. These adjusted distances correspond to the worst-case spatial offset that results when all maximum errors in **Table 1** occur simultaneously. Critically, these looser merging thresholds were applied solely during the mapping stage; the evaluation itself continued to use the original 0.35 m limit for matching mapped objects to ground truth, guaranteeing a level playing field. For each error tier, the adaptive planner's F1-score was compared against that of a coverage path flown at 12 m, 24 m, 36 m, and 48 m altitude, separately for the clustered and uniform orthomosaics.

Table 1. Localization uncertainty tiers employed in experiment 3, specified as offsets in UAV position, altitude, heading, and camera gimbal roll and pitch. The listed values give the lower and upper bounds of the uniform distribution from which each offset is drawn.

Level	Heading	Pitch	Roll	Altitude	Position
Perfect	± 0.0°	± 0.0°	± 0.0°	± 0.000 m	± 0.000 m
Good	± 0.5°	± 0.5°	± 0.5°	± 0.015 m	± 0.015 m
Decent	± 1.0°	± 1.0°	± 1.0°	± 0.030 m	± 0.030 m
Poor	± 2.0°	± 2.0°	± 2.0°	± 0.100 m	± 0.100 m
Very poor	± 5.0°	± 5.0°	± 5.0°	± 0.200 m	± 0.200 m

Experiment 4: number of objects

© 2022 The Author(s). | Published by Agro Intelligence Research Press

To explore how the total count of target objects within the field affects the length of the resulting flight path, the object population was systematically modified. For each dataset, an additional 140 object instances were synthesized from annotated plant crops in the training partition, bringing the total to 200 objects. These cropped patches were rotated at random and composited into the orthomosaic to create plausible scenes for the detection network. The strategy for placing the supplementary objects depended on whether the original orthomosaic followed a clustered or a uniform pattern:

Clustered distribution: the number of objects assigned to a given cluster, k, was drawn from a normal distribution:

$$k = \mathrm{round}(x), x \sim \mathcal{N}(\mu = 8, \sigma = 3), \quad (16)$$

Where μ is the mean, and σ is the standard deviation. Object positions within each cluster were generated from a multivariate Gaussian distribution whose center was randomly located within the field boundaries and whose covariance matrix Σloc was a randomly constructed positive semidefinite matrix:

$$\Sigma = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, a_{ij} \sim \mathcal{N}(\mu = 5, \sigma = 2), \quad (17)$$

$$\Sigma_{\mathrm{loc}} = \Sigma \cdot \Sigma^{T}. \quad (18)$$

A minimum distance of 15.0 m was enforced between the centers of any two clusters, and a 1.0 m minimum spacing was maintained between all individual objects; positions were resampled if these constraints were not satisfied.

Uniform distribution: the extra objects were scattered uniformly at random across the field area, again subject to a 1.0 m minimum pairwise distance.

The influence of object count was examined by varying the number from 0 up to 200 in steps of 20. Since the unmodified orthomosaics each contained 60 objects, scenarios requiring fewer than 60 were realized by obscuring the surplus object locations—replacing their pixel values with green fill—and scenarios requiring more than 60 objects were drawn from the pool of additional synthetic instances. For the clustered arrangement, object locations were ordered by cluster affiliation, so that objects were inserted into or removed from an existing cluster before another cluster was expanded or contracted. **Figure 7** shows illustrative examples of the original and added object positions in the 13 February 2024 dataset for both spatial layouts. For each object count, we plotted object density (objects per hectare) against the relative difference in route length between the adaptive planner and the baseline. This analysis was carried out separately for the orthomosaics exhibiting clustered and uniform object distributions.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

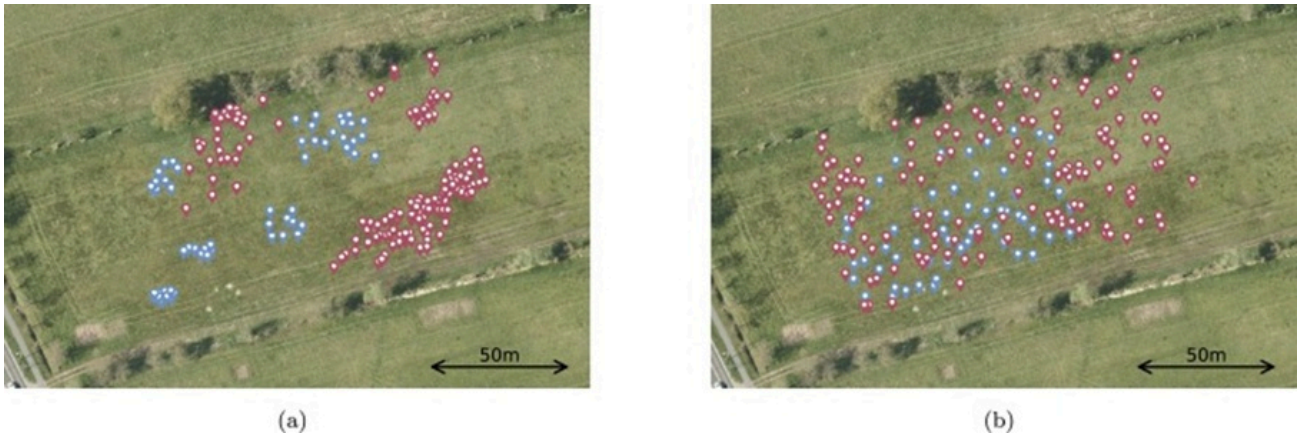
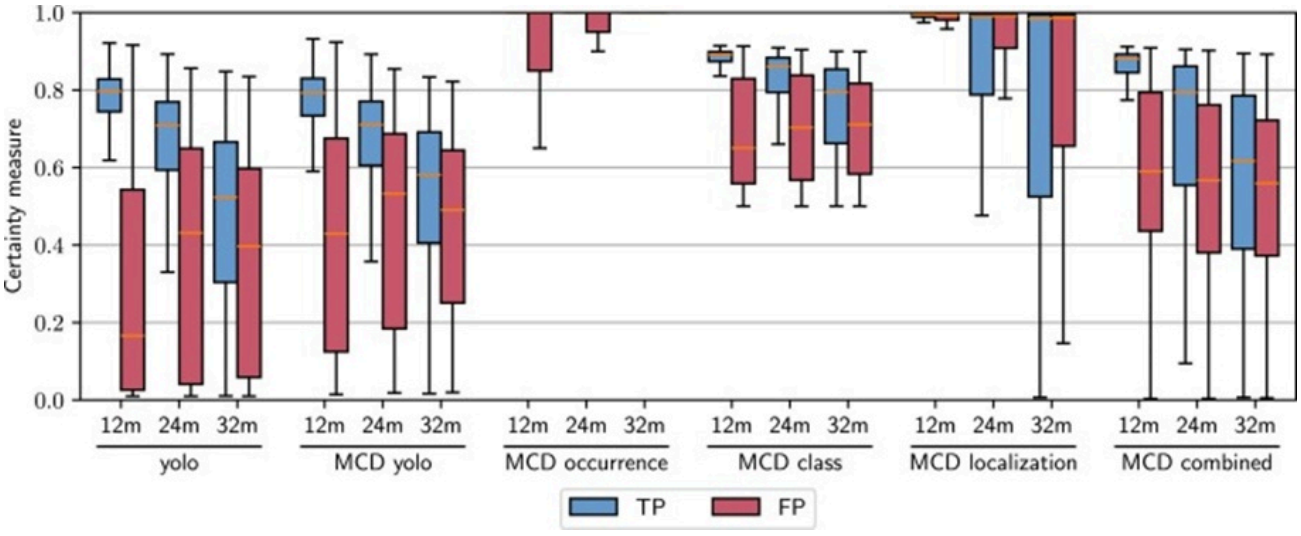


Figure 7. Illustration of the original (blue) and added (red) object placements on the 13 February 2024 dataset under the clustered distribution (a) and the uniform distribution (b).

Results and Discussion

Experiment 1: relationship between detection certainty and errors

To assess the capacity of the various certainty metrics to distinguish TP from FP detections, **Figure 8** plots the distribution of the native YOLOv8 confidence and each MCD-derived certainty measure for TPs and FPs across the different altitudes. A desirable certainty metric would concentrate TP detections at the high end of the scale and FP detections at the low end. Across the board, TPs tended to receive higher certainty scores than FPs, though this pattern did not hold for localization certainty at 24 m and 32 m. For all metrics examined, increasing altitude caused the TP and FP distributions to overlap more. The native YOLOv8 confidence cleanly separated TPs from FPs at 12 m, but this gap narrowed considerably with increasing height. The MCD-based occurrence measure collapsed to 1.0 in most cases, rendering it unable to discriminate TPs from FPs. The MCD class uncertainty measure successfully separated TPs and FPs at 12 m altitude. Localization certainty likewise produced a clear split at 12 m, though its distributions blended substantially at the higher altitudes. The combined MCD certainty did not surpass the native YOLOv8 confidence in distinguishing correct from incorrect detections.



© 2022 The Author(s). | Published by Agro Intelligence Research Press

Figure 8. Boxplots showing how true positive (TP) and false positive (FP) detections are distributed for the standard YOLOv8 certainty and the certainty measures derived via Monte-Carlo Dropout (MCD) at several altitudes.

Table 2 reports the t-values from Welch's t-test, testing the alternative hypothesis that the TP mean confidence is greater than the FP mean confidence. Larger t-values denote stronger separability. As shown in Figure 8, higher altitudes reduced discriminability, making it harder to distinguish between correct and erroneous detections. The standard YOLOv8 certainty metric ranked first at both 12 m and 24 m. At 32 m, the MCD class metric took the lead, though the native YOLOv8 confidence followed closely as the second-best performer.

Table 2. T-values from Welch's t-test for the standard YOLOv8 certainty and the Monte-Carlo Dropout (MCD) certainty methods, evaluating the alternative hypothesis that the mean confidence of true positives is higher than that of false positives. Asterisks (") denote a statistically significant difference ($\alpha = 0.05$) between the TP and FP distributions; the top t-value at each altitude is shown in bold.

Certainty method	32 m	24 m	12 m
yolo	18.0*	**48.3***	**101.1***
MCD yolo	15.6*	40.7*	77.8*
MCD occurrence	11.1*	25.6*	41.6*
MCD class	**18.3***	42.5*	78.7*
MCD localization	-5.1	-3.2	3.4*
MCD combined	5.7*	23.1*	61.4*

When MCD was activated during inference, the mean total processing duration per image—spanning preprocessing, inference, and postprocessing—climbed to 105.2 ± 116.1 ms, compared with 22.1 ± 2.9 ms for a standard YOLOv8 forward pass without dropout.

Because the MCD-derived certainty measures offered no tangible improvement over the native YOLOv8 confidence for separating correct from incorrect detections at 12 m and 24 m, and given that the advantage of MCD class over YOLOv8 confidence at 32 m was slight, and bearing in mind that MCD inflated processing time roughly fivefold, all subsequent experiments rely on the standard YOLOv8 confidence as the certainty measure. Averaged over all images in the dataset and using the YOLOv8 confidence score, the F1-scores stood at 0.83 for 12 m, 0.70 for 24 m, and 0.44 for 32 m altitude.

Experiment 2: parameter optimization

Figure 9 presents the impact of the coverage altitude (hcov), acceptance confidence threshold (caccept), and rejection confidence threshold (creject) on the F1-Scores delivered by the adaptive planner and the coverage planner, split by whether the plants were clustered or uniformly distributed. For the coverage planner, the F1-score dropped steadily as altitude rose. The adaptive planner also suffered performance losses at higher altitudes, but the decline was cushioned, with the degree of mitigation depending on the rejection threshold. Raising the rejection threshold pulled the F1-score downward for both distribution types, and this effect was amplified when combined with a high coverage altitude. F1-scores generally came out higher on the clustered layouts than on the uniform ones; nonetheless, the directional influence of the parameters remained consistent across both distribution types.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

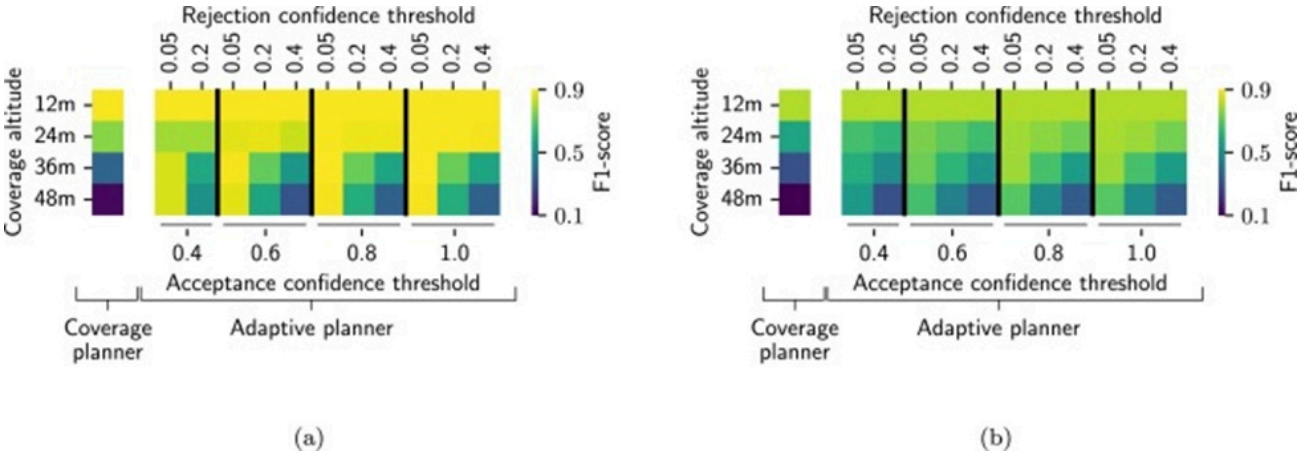


Figure 9. Heatmaps depicting the F1-score as a function of coverage altitude, acceptance confidence threshold, and rejection confidence threshold, for the coverage planner and the adaptive planner evaluated on the clustered distribution (a) and the uniform distribution (b).

Figure 10 shows how the three parameters influence the relative path length (relative to the 12 m coverage baseline) for both planners. Operating a coverage path at 48 m altitude slashed the path length by roughly 70%. The adaptive planner produced routes longer than the 48 m coverage path owing to the appended inspection segments; even so, it remained about 40% shorter than the 12 m coverage reference. Imposing a higher rejection confidence trimmed the flight path, since fewer objects were queued for low-level inspection. This effect was more pronounced under the uniform distribution, where objects were, on average, farther apart. Raising the acceptance threshold increased the path length, as a larger fraction of objects was deemed to require low-altitude verification. However, this effect manifested visibly only at the 24 m coverage altitude.

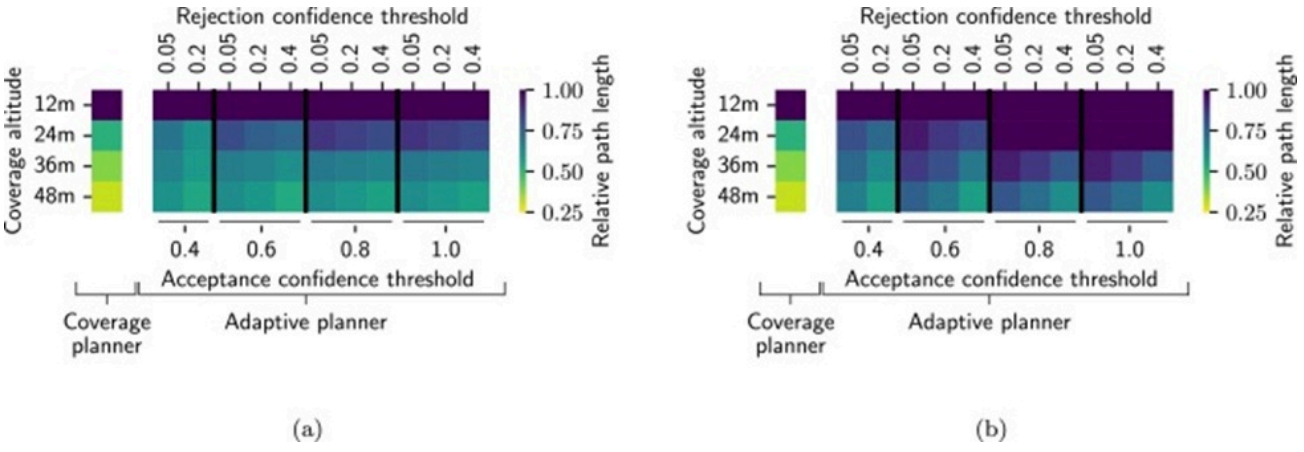


Figure 10. Heatmaps depicting the relative flight path length as a function of coverage altitude, acceptance confidence threshold, and rejection confidence threshold, for the coverage planner and the adaptive planner evaluated on the clustered distribution (a) and the uniform distribution (b).

When examined together, Figures 9 and 10 expose the compromise that must be struck between detection fidelity and route compactness. Flying the coverage path at 12 m yielded the peak F1-score while simultaneously incurring the longest travel distance. A strict rejection threshold, coupled with a 48 m coverage altitude, led to many false negatives and thus very few low-altitude inspections, resulting in a short flight path at the expense of detection performance. Conversely, a lenient rejection threshold bolstered detection performance but lengthened the route.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

Table 3 presents the parameter choices that achieved both respectable detection performance and a relatively short path. These settings were adopted for experiments 3 and 4. The table also lists the adaptive planner's F1-score, the F1-score of a coverage planner flying at the same altitude, and the normalized path length. The coverage path outperformed the adaptive path in F1-score for both the clustered and uniform distributions, but at the cost of greater travel distance. For the clustered distribution, trading away 2 percentage points of F1-score bought a 37% reduction in path length. The benefit was narrower for the uniform distribution: a 2-point F1-score sacrifice yielded just a 6% path length reduction. **Figure 11** shows the adaptive flight paths obtained with these optimized parameters across the four data-collection dates.

Table 3. Optimized parameter values and the corresponding F1-score and normalized flight path length for the clustered and uniform distributions, chosen to achieve a high F1-score and a short path length jointly.

Parameter	Uniform distribution	Clustered distribution
Coverage altitude	36 m	48 m
Acceptance threshold	0.8	0.6
Rejection threshold	0.05	0.05
F1-score adaptive planner	0.784 ± 0.075	0.867 ± 0.024
F1-score coverage planner	0.807 ± 0.039	0.891 ± 0.021
Normalized flight path length	0.938 ± 0.116	0.631 ± 0.035

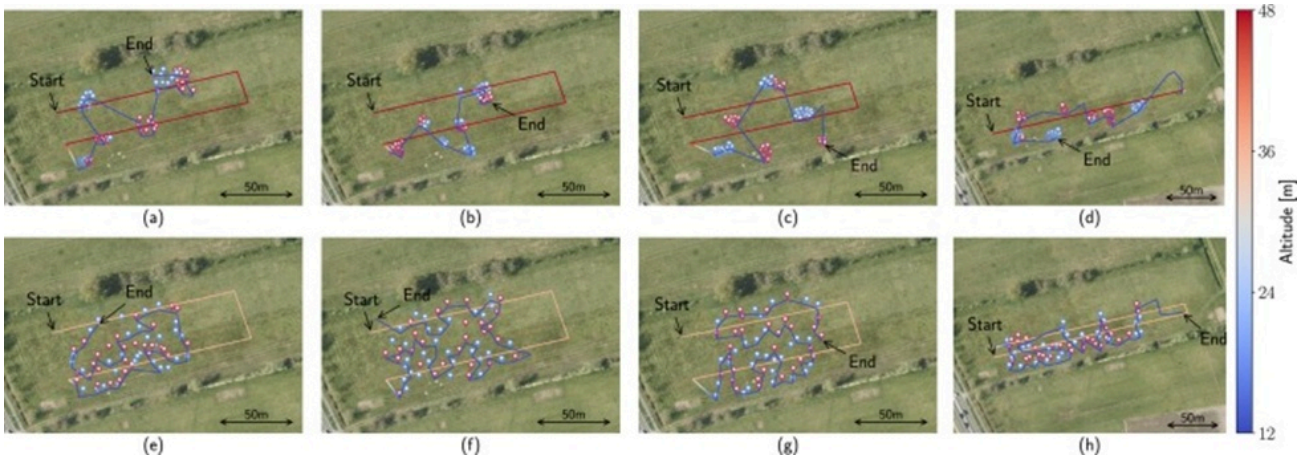


Figure 11. Flight paths produced by the adaptive planner under the optimized parameter settings for the datasets of 13 February 2024 (a, e), 23 April 2024 (b, f), 18 July 2024 (c, g), and 1 August 2024 (d, h), with clustered (a–d) and uniform (e–h) distributions of plants. Arrows indicate the UAV's start and end points; red and blue markers mark the ground truth positions of weed types I and II, respectively.

Experiment 3: localization errors

Figure 12 illustrates the consequences of localization errors for the F1-score attained by the adaptive planner and the low-altitude coverage planner, shown separately for clustered and uniform object arrangements. Five error grades were evaluated, ranging from 'perfect' to 'very poor' (detailed in section 2.8.3). Greater localization inaccuracy pulled down F1-scores for both planning approaches. Crucially, because absolute georeferencing error shrinks at lower flight altitudes, the adaptive planner held an edge over the coverage planner at 24 m, 36 m, and 48 m by virtue of the low-altitude inspections it scheduled, which improved the precision of object coordinates. A steep performance decline occurred

© 2022 The Author(s). | Published by Agro Intelligence Research Press

between the ‘decent’ and ‘poor’ tiers, suggesting a maximum acceptable error envelope of 0.03 m in position and altitude, along with 1-degree errors in heading, gimbal roll, and gimbal pitch. Under more severe error conditions, the depressed F1-scores can be explained by a proliferation of false positives: separate sightings of the same plant were not consolidated, resulting in duplicate entries in the map. No noteworthy disparity in F1-score emerged between the clustered and uniform distributions.

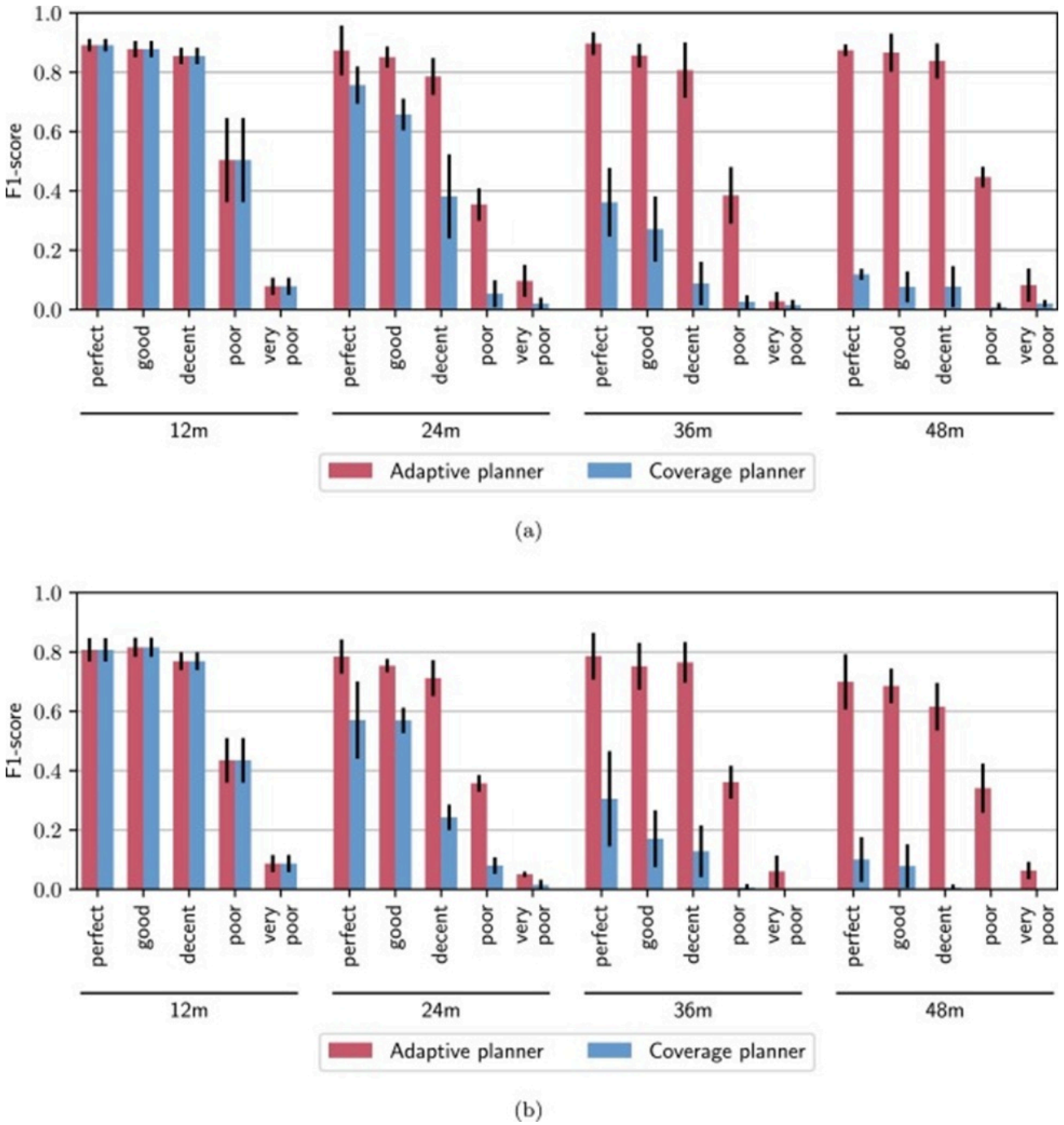


Figure 12. Influence of localization error on the F1-score for the adaptive planner and the coverage path planner at multiple altitudes, shown for the clustered distribution (a) and the uniform distribution (b) of plants. Error bars represent the standard deviation around the mean.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

Experiment 4: number of objects

The relationship between field object population and relative route length is captured in **Figure 13**, with results broken out by clustered versus uniform placement. Where plants were grouped into clusters, the adaptive strategy consistently yielded a more compact route than what a low-altitude full-coverage sweep would require—this held for every object count examined. The gap, however, narrowed considerably as density increased, with the two approaches becoming nearly indistinguishable beyond approximately 250 objects per hectare. The uniform configuration told a different story entirely: once the field contained more than roughly 80–120 objects per hectare, the adaptive approach actually produced a longer route than simply executing a complete low-altitude pass. The underlying reason is straightforward—when targets are evenly distributed, the adaptive method is compelled to investigate a large fraction of them at low altitude, and the overhead of planning and flying a piecemeal inspection route outweighs the efficiency of a systematic full-coverage sweep.

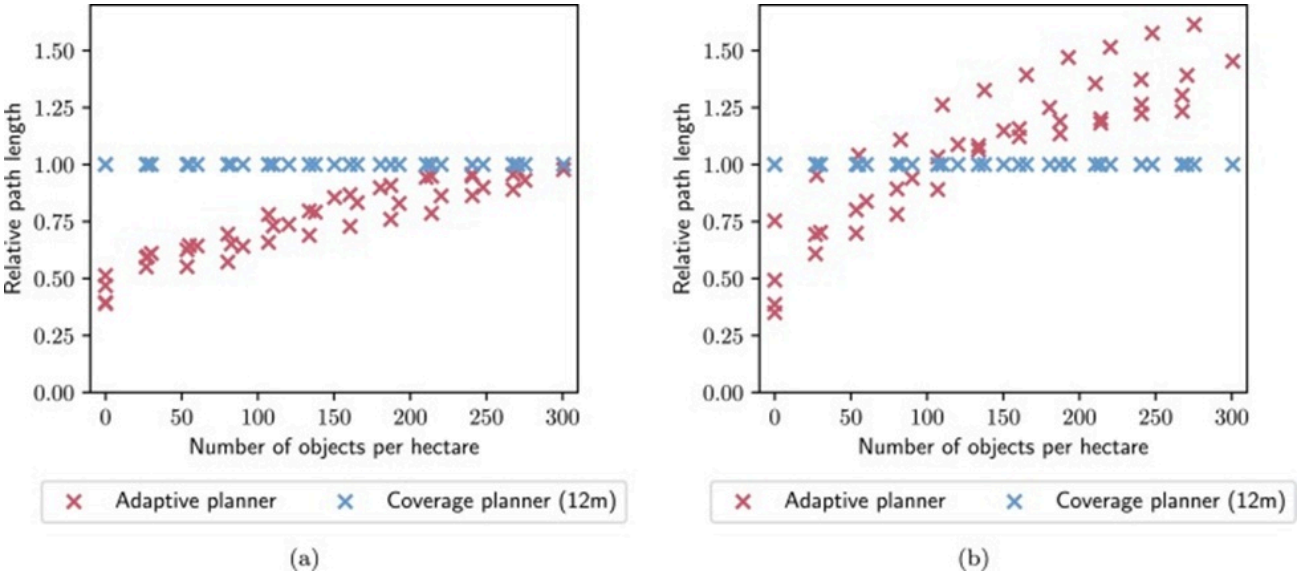


Figure 13. Normalized flight path length plotted against object count for both the adaptive planner and the coverage planner, evaluated separately on the clustered distribution (a) and the uniform distribution (b) of plants.

Figure 14 offers a visual comparison of the routes generated by the two planning paradigms on the modified datasets from each of the four collection dates, with object counts set at 20, 80, 140, and 200. The uniform-distribution scenarios forced the UAV to crisscross the field visiting inspection points spread more or less everywhere, which multiplied the number of waypoints and thereby stretched the total travel distance well beyond what was needed for the clustered layouts.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

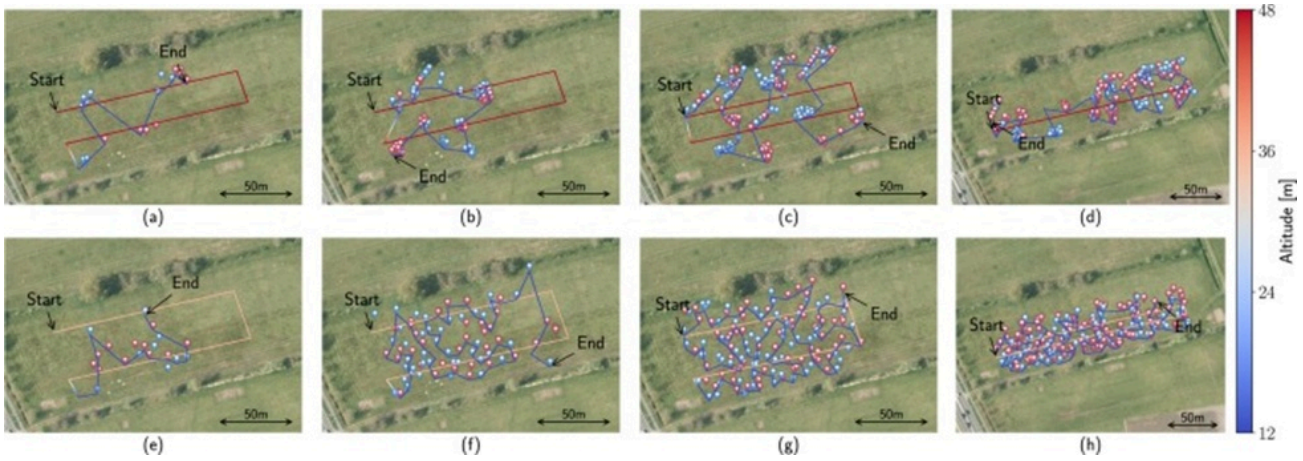


Figure 14. Routes produced by the adaptive planner on the datasets of 13 February 2024 with 20 objects (a, e), 23 April 2024 with 80 objects (b, f), 18 July 2024 with 140 objects (c, g), and 1 August 2024 with 200 objects (d, h), under clustered (a–d) and uniform (e–h) plant arrangements. Arrows denote the UAV's takeoff and landing positions; red and blue markers correspond to the surveyed locations of weed types I and II.

What emerges most clearly from the data is the considerable benefit an adaptive planning strategy can offer over a blanket coverage approach—provided the objects of interest are spatially aggregated rather than spread uniformly. With the tuned parameter values listed in **Table 3**, our adaptive planner trimmed the total flight distance by 37% while incurring an F1-score penalty of only 2 percentage points relative to the coverage baseline, evaluated on the clustered-distribution datasets. A further advantage became apparent in the localization error trials: the adaptive planner proved markedly less sensitive to positioning inaccuracies than the coverage planner **(Figure 12)**. Raising the number of objects under a clustered arrangement showed that the adaptive method held its edge over a 12 m coverage flight up to densities of around 300 objects per hectare **(Figure 13a)**. The uniform case flipped this result—once object counts climbed past roughly 100, the coverage planner emerged as the more efficient choice **(Figure 13b)**. The discussion that follows situates these findings within the existing literature (section 4.1), takes a closer look at the certainty estimation approach (section 4.2), unpacks the consequences of detection errors for the adaptive planner (section 4.3), considers the influence of localization inaccuracies (section 4.4), assesses the realism of the offline simulation pipeline (section 4.5), and reflects on what would be required to deploy the planner in real operational settings (section 4.6).

Comparison with literature

The benchmark we adopted was a full-coverage flight—the standard operating procedure in most real-world applications—and against this reference, our method delivered a 37% shorter route for 2 F1-score points. For perspective, the learning-based adaptive planner introduced by van Essen *et al.* [11] achieved a far more dramatic 74% reduction in path length, although detection accuracy suffered a more substantial 12-point drop. In the work of Popović *et al.* [9], their planner found 85% of all objects using only half the flight time required by a coverage approach and located the complete set of objects within roughly the same duration as the coverage baseline. Each of these prior contributions illuminates the inevitable tension between saving flight time and preserving detection fidelity. Compared with those systems, our planner lands at a different point on that trade-off curve: the sacrifice in detection quality is comparatively small, but the time savings are also less dramatic. Experiment 2 confirmed that this operating point is adjustable. By modifying the coverage altitude and the two confidence thresholds, one can push the balance further toward route compactness if the application permits a larger accuracy penalty.

It also became clear that adopting an adaptive planning strategy only made sense when the spatial pattern of the targets departed from uniformity. This conclusion echoes that of Popović *et al.* [9] and van Essen *et al.* [11], who each reported that route efficiency deteriorated once objects were evenly distributed across the field.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

Detection certainty estimation

Relying on Monte-Carlo Dropout to produce certainty scores did not translate into better separation of true and false positives in the first experiment, with the sole exception of the 32 m altitude condition. Prior work in active learning has demonstrated that MCD-based certainty can effectively identify the most ambiguous image for human labeling [14, 15], and a positive correlation between MCD-derived certainty and detection accuracy has been reported. Our experiment 1 data are not at odds with this general pattern: true positives did, on aggregate, receive higher certainty assignments than false positives. Myojin *et al.* [13], in contrast, observed a small F1-score gain when MCD uncertainty was incorporated into a customized YOLOv3 pipeline designed to detect lunar craters. The gain they reported, however, was tiny. Since YOLOv8 represents a more capable detector than YOLOv3 in terms of raw accuracy [25], the confidence values it natively outputs tend to be better calibrated to begin with.

Perhaps the most serious practical obstacle to using MCD for uncertainty estimation in our context is the sheer increase in computation time. Our measurements indicated that turning on MCD—even while caching the backbone and neck feature maps across passes—caused inference to run five times slower than a standard forward pass. For an adaptive planner intended to operate with on-board processing under limited power and thermal envelopes, this slowdown could substantially erode the very time savings the planner is designed to achieve. A worthwhile direction for subsequent research is to investigate computationally cheaper uncertainty quantification approaches, with the Laplacian approximation being a prominent candidate [12].

Impact of detection errors

Pushing the coverage phase to higher altitudes raised the number of objects the detection network failed to spot, which mechanically lowered the F1-score **(Figure 9)**. A false negative during the high-altitude sweep is particularly damaging because it precludes any possibility of a follow-up low-altitude inspection—the object is never revisited. This dynamic explains the finding from experiment 2 that the rejection confidence threshold that worked best was the smallest tested, 0.05, because it minimized the false-negative tally. The trade-off was a larger pool of false positives, but these were later filtered out during the inspection phase. The damage inflicted by false negatives depends heavily on the spatial arrangement of the targets. In a clustered setting, if at least some members of a cluster are detected with low certainty, the resulting low-altitude inspection will typically bring the rest of the cluster into view—including individuals that were entirely missed at high altitude. This safety net does not exist in the uniform case. Experiment 2 captured this asymmetry by finding that the optimal coverage altitude was lower for uniform distributions, thereby directly reducing the false negative rate during the initial sweep.

Impact of localization errors

The third experiment showed that the adaptive planner possesses greater inherent tolerance to localization errors than the coverage baseline, at least within a band extending to roughly 3 cm of positional error and 1 degree of attitude error. The mechanism behind this robustness is straightforward: since localization uncertainty scales down with flight altitude, the low-altitude inspection passes naturally improve the georeferencing accuracy of the final map. An airframe equipped with RTK-GNSS should easily stay within this error budget, indicating the practical feasibility of the approach. The level of localization precision actually needed also depends on the downstream task. Consider a weed control use case where the resulting map serves to direct a removal robot: here, the accuracy requirement is less stringent because the map only needs to get the robot close enough that the weeds fall within its on-board sensing range.

Validity of the offline simulation method

Our evaluation methodology relied on simulated flights conducted over orthomosaics stitched from real drone imagery. Orthomosaics are inherently perspective-corrected, which could, in principle, remove some of the geometric complexity present in raw aerial photographs. In practice, however, because the artificial plants occupied a footprint far smaller than the UAV's altitude, cropping small regions from the orthomosaic produced image patches whose viewpoint geometry

© 2022 The Author(s). | Published by Agro Intelligence Research Press

closely approximated that of a genuine nadir capture. Corroborating this fidelity, the F1-scores recorded on real imagery at 12 m in experiment 1 aligned well with the F1-scores obtained from the simulated coverage flights at 12 m in experiments 2, 3, and 4. Furthermore, published object detection results in agricultural domains typically fall within an F1-score window of 0.7 to 0.9 [26-28], a range that neatly brackets the performance figures observed throughout our experiments. These consistencies support the conclusion that orthomosaic-driven simulation offers a reliable stand-in for actual flight. The logical next step is to validate the planner during live field trials.

Operational opportunities for an adaptive path planner

The practical value an adaptive path planner delivers is tightly coupled to the nature of the task it serves. In settings where objects are scattered uniformly across a field, executing a standard low-altitude coverage sweep generally makes more sense, except when object density is below roughly 70 individuals per hectare. Once the spatial arrangement shifts toward clustering and patchiness, however, the adaptive scheme begins to pay off—it shortens the time spent in the air while maintaining detection accuracy close to that of a full-coverage baseline. Weed mapping stands out as a candidate application that could capitalize on this behavior, since field observations confirm that many weed species form discrete, well-defined patches rather than dispersing evenly [23, 29].

An additional lever for magnifying the planner's advantage is the ability to fly faster at higher altitudes. Because the sensor footprint expands with altitude and airspeed limits are primarily governed by the need to avoid motion blur, climbing higher permits a corresponding increase in flight velocity [30]. Exploiting this relationship would widen the performance gap between the adaptive approach and a low-altitude coverage flight beyond the margins documented in this paper.

Another design choice that could enhance the method's effectiveness is to incorporate multiple complementary sensors, rather than restricting both the high- and low-altitude phases to a single RGB camera. In our trials, a single RGB imager served both roles. Yet certain plant conditions—infection-related biotic stress, as one illustration—manifest themselves in the thermal portion of the spectrum [31], even though thermal cameras characteristically produce imagery of limited spatial detail, which curtails their ground coverage per frame. An adaptive workflow that fuses a detailed RGB camera with a thermal sensor could operate as follows: the RGB feed collected from altitude would pinpoint suspicious areas, which would then be revisited for a close-range thermal scan. This arrangement focuses the thermal camera's narrow field of view precisely on the locations where its diagnostic capability is needed most.

Conclusion

The work reported here presents an adaptive path-planning framework that couples a high-altitude survey pass with low-altitude verification sweeps, launched only for detections about which the network expresses low confidence. Evaluation was conducted on georeferenced orthomosaics derived from actual UAV-based RGB imagery. Among the several certainty quantification methods examined, the raw confidence score output by YOLOv8 proved most effective at distinguishing correct detections from erroneous ones. It was consequently adopted as the criterion for triggering closer inspection of uncertain plant candidates. In trials where plants were arranged in clusters, the adaptive planner slashed total flight distance by 37% compared with a low-altitude full-coverage route, at a modest expense of 2 percentage points in F1-score. Under uniform plant placement, the same approach achieved a 6% reduction in path length, paired with a 2-point decrement in F1-score. The adaptive planner also handled localization inaccuracies better than a coverage planner, a resilience conferred by the low-altitude inspection passes. For uniform distributions, a coverage path maintained an efficiency edge except when object density was below 80–120 objects per hectare. Yet wherever the objects were non-uniformly distributed, the adaptive planner remained the faster option up to densities of 250–300 objects per hectare. In sum, the proposed adaptive path planner enables quicker localization of non-uniformly distributed objects within a field

© 2022 The Author(s). | Published by Agro Intelligence Research Press

than a conventional coverage path planner, all while achieving detection accuracy that closely matches the coverage baseline.

Acknowledgements

None

Conflict of interest

None

Financial support

This research is part of the Synergia research program; funding was obtained from the Dutch Research Council (NWO grant 17626), IMEC-One Planet, and other private parties.

Ethics statement

None

Received: 08 Sep 2021 Revised: 12 Nov 2021 Accepted: 29 Nov 2021
Published online: 25 May 2022

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

1. Rejeb A, Abdollahi A, Rejeb K, Treiblmaier H. Drones in agriculture: a review and bibliometric analysis. Comput Electron Agric. 2022;198:107017. https://doi.org/10.1016/j.compag.2022.107017.

2. Yuan J, Zhang Y, Zheng Z, Yao W, Wang W, Guo L. Grain crop yield prediction using machine learning based on UAV remote sensing: a systematic literature review. Drones. 2024;8:559. https://doi.org/10.3390/drones8100559.

3. Murad NY, Mahmood T, Forkan ARM, Morshed A, Jayaraman PP, Siddiqui MS. Weed detection using deep learning: a systematic literature review. Sensors. 2023;23:3670. https://doi.org/10.3390/s23073670..

4. Shahi TB, Xu CY, Neupane A, Guo W. Recent advances in crop disease detection using UAV and deep learning techniques. Remote Sens. 2023;15:2450.

https://doi.org/10.3390/rs15092450.

5. Kouadio L, El Jarroudi M, Belabess Z, Laasli SE, Roni MZK, Amine IDI, et al. A review on UAV-based applications for plant disease detection and monitoring. Remote Sens. 2023;15:4273. https://doi.org/10.3390/rs15174273.

6. Cabreira T, Brisolara L, Ferreira PR Jr. Survey on coverage path planning with unmanned aerial vehicles. Drones. 2019;3:4. https://doi.org/10.3390/drones3010004.

7. Popović M, Ott J, Rückin J, Kochenderfer MJ. Learning-based methods for adaptive informative path planning. Robot Auton Syst. 2024;179:104727. https://doi.org/10.1016/j.robot.2024.104727.

8. Blanchard A, Sapsis T. Informative path planning for anomaly detection in environment exploration and monitoring. Ocean Eng. 2022;243:110242. https://doi.org/10.1016/j.oceaneng.2021.110242.

9. Popović M, Hitz G, Nieto J, Sa I, Siegwart R, Galceran E. Online Informative Path Planning for Active Classification Using UAVs. Proceedings - IEEE International Conference on Robotics and Automation, 9781509046331, Institute of Electrical and Electronics Engineers Inc. 2017:5753-8. https://doi.org/10.1109/ICRA.2017.7989676.

10. Stache F, Westheider J, Magistri F, Stachniss C, Popović M. Adaptive path planning for UAVs for multi-resolution semantic segmentation. Robot Auton Syst. 2023;159:104288. https://doi.org/10.1016/j.robot.2022.104288.

11. van Essen R, van Henten E, Kootstra G. Learning UAV-based path planning for efficient localization of objects using prior knowledge. arXiv:2412.11717 [cs]. 2024.

12. Gawlikowski J, Tassi CRN, Ali M, Lee J, Humt M, Feng J, et al. A survey of uncertainty in deep neural networks. Artif Intell Rev. 2023;56:1513-89. https://doi.org/10.1007/s10462-023-10562-9..

13. Myojin T, Hashimoto S, Mori K, Sugawara K, Ishihama N. Improving reliability of object detection for lunar craters using Monte Carlo dropout. In: Tetko IV, Kůrková V, Karpov P, Theis F, editors. Artificial Neural Networks and Machine Learning – ICANN 2019: Image Processing. Springer International Publishing, Munich, Germany. 2019:68-80. https://doi.org/10.1007/978-3-030-30508-6_6.

14. Blok PM, Kootstra G, Elghor HE, Diallo B, Van Evert FK, Van Henten EJ. Active learning with MaskAL reduces annotation effort for training Mask R-CNN on a broccoli dataset with visually similar classes. Comput Electron Agric. 2022;197:106917. https://doi.org/10.1016/j.compag.2022.106917.

15. Sokolova M, Blok PM, Mencarelli A, Vroegop A, Helmond AV, Kootstra G. Improved detection of discarded fish species through BoxAL active learning. http://arxiv.org/abs/2410.04880. arXiv:2410.04880 [cs]. 2024

16. Rodriguez-Vazquez J, Prieto-Centeno I, Fernandez-Cortizas M, Perez-Saura D, Molina M, Campoy P. Real-time object detection for autonomous solar farm inspection via UAVs. Sensors. 2024;24:777. https://doi.org/10.3390/s24030777.

17. Lamping C, Kootstra G, Derks M. Uncertainty estimation for deep neural networks to improve the assessment of plumage conditions of chickens. Smart Agric Technol. 2023;5:100308. https://doi.org/10.1016/j.atech.2023.100308.

18. Mier G, Valente J, de Bruin S. Fields2Cover: an open-source coverage path planning library for unmanned agricultural vehicles. IEEE Robot Autom Lett. 2023.

© 2022 The Author(s). | Published by Agro Intelligence Research Press

https://doi.org/10.1109/LRA.2023.3248439.

19. Jocher G, Chaurasia A, Qiu J. Ultralytics YOLOv8. https://github.com/ultralytics/ultralytics. 2023

20. Mulvad SH. fast-tsp: a fast TSP solver with Python bindings. https://github.com/shmulvad/fast-tsp. 2022.

21. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014;15:1929-58.

22. Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. arXiv:1506.02142 [stat]. 2016.
https://doi.org/10.48550/arXiv.1506.02142..

23. Cardina J, Johnson GA, Sparrow DH. The Nature and Consequence of Weed Spatial Distribution. Weed Sci. 1997;45:364-73.

24. Welch BL. The generalization of “Student’s” problem when several different population variances are involved. Biometrika. 1947;34:28-35.

25. Jegham N, Koh CY, Abdelatti M, Hendawi A. Evaluating the evolution of YOLO (you only look once) models: a comprehensive benchmark study of YOLO11 and its predecessors. http://arxiv.org/abs/2411.00201. 2024. arXiv:2411.00201 [cs].
https://doi.org/10.48550/arXiv.2411.00201..

26. Rai N, Zhang Y, Ram BG, Schumacher L, Yellavajjala RK, Bajwa S, et al. Applications of deep learning in precision weed management: a review. Comput Electron Agric. 2023;206:107698.
https://doi.org/10.1016/j.compag.2023.107698.

27. Ruigrok T, Van Henten EJ, Kootstra G. Improved generalization of a plant-detection model for precision weed control. Comput Electron Agric. 2023;204:107554.
https://doi.org/10.1016/j.compag.2022.107554.

28. Rehman MU, Eesaar H, Abbas Z, Seneviratne L, Hussain I, Chong KT. Advanced drone-based weed detection using feature-enriched deep learning approach. Knowl Based Syst. 2024;305:112655.
https://doi.org/10.1016/j.knosys.2024.112655.

29. Xu B, Fan J, Chao J, Arsenijevic N, Werle R, Zhang Z. Instance segmentation method for weed detection using UAV imagery in soybean fields. Comput Electron Agric. 2023;211:107994.
https://doi.org/10.1016/j.compag.2023.107994.

30. Seifert E, Seifert S, Vogt H, Drew D, Van Aardt J, Kunneke A, et al. Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images. Remote Sens. 2019;11:1252.
https://doi.org/10.3390/rs11101252.

31. Pineda M, Barón M, Pérez-Bueno ML. Thermal imaging for plant stress detection and phenotyping. Remote Sens. 2020;13:68.
https://doi.org/10.3390/rs13010068..

© 2022 The Author(s). | Published by Agro Intelligence Research Press