Leveraging foundation models for data-limited ecological applications

Kyle Doherty[1], **Max Gurinas**[2], **Erik Samsoe**[1], **Charles Casper**[1],
Beau Larkin[1], Philip Ramsey[1], Brandon Trabucco[3], Ruslan Salakhutdinov[3]
[1]MPG Ranch, [2]University Of Chicago Laboratory Schools, [3]Carnegie Mellon University
kdoherty@mpgranch.com

Abstract

Human-driven change in natural ecosystems is a global challenge relevant to human and natural communities alike. Yet, ecological data (species presence/absence or abundance), the bedrock of global change impact monitoring, are difficult to gather due to challenging field conditions, few subject experts (e.g., botanists), and brief monitoring windows. Therefore, the default condition of ecological data analysis is one of data limitation. As the generalization abilities of large foundation models grow, we might leverage these models to derive ecological insights from few data. And because ecological data are inherently more rare, they also offer the machine learning community an opportunity to better study the out-of-distribution performance of foundation models in few-shot contexts. To illustrate these principles, we gathered a field-validated dataset of presence and absence of leafy spurge (*Euphorbia esula*), a weed that invades natural areas and displaces native species in North America. We then surveyed these areas with a consumer-grade drone and extracted images from ground truth locations. We fine-tuned a convolutional neural network and a state-of-the-art vision transformer on these data, then contrasted few-shot performance with that of off-the-shelf GPT-4 checkpoints. While we achieved state-of-the-art classification performance on the full dataset with the fined-tuned DINOv2 vision transformer (0.85 test accuracy), GPT-4o nearly matched prior SOTA performance (0.75 test accuracy) from in-context learning when shown only 8 examples per class. Furthermore, we observed a 10% test accuracy improvement between best GPT-4-turbo and GPT-4o results, illustrating the rapid advances in recent months. Our findings demonstrate the mutual benefit of pairing ecological data with the study of the generalization abilities of foundation models. We release the Leafy Spurge Dataset for further few-shot experiments and evaluation. Please find our code and data at our website: https://leafy-spurge-dataset.github.io (Creative Commons Attribution 4.0 International license).

1 Introduction

Gathering ecological data requires expertise in species identification, spatial planning, and haste to capture ephemeral patterns. Consequently, scaling ecological monitoring efforts across larger areas is difficult without supporting technologies such as satellite or drone-based imaging (Turner et al., 2003). Agriculture has pioneered these remote sensing approaches to detect and map changes to crop health (Mulla, 2013) and weed plant occurrence (Lamb and Brown, 2001). A number of benchmark datasets exist in the agricultural domain with target objectives such as detecting weeds and assessing crop health in high resolution drone imagery (dos Santos Ferreira et al., 2017; Chiu et al., 2020; Krestenitis et al., 2022; Genze et al., 2022; Wildeboer, 2023). Increasingly, remote sensing tools are

Foundation Models for Science Workshop,38th Conference on Neural Information Processing Systems (NeurIPS 2024).

Figure 1: A sagebrush community invaded by leafy spurge (left) and a closeup of a leafy spurge inflorescence (right).

applied in wildland settings with similar goals: to monitor change in plant communities and mount a management response (Pettorelli et al., 2014). Yet, there are few publicly available benchmark datasets containing high resolution imagery of wildlands, and these are biased toward forested areas (Weinstein et al., 2021; arura uav, 2023; Mowla et al., 2024), which represent a minority of terrestrial biomes (FAO, 2020). While the challenges posed by large-scale ecological monitoring are formidable, this effort is critical to tracking change in natural ecosystems, and the dearth of benchmark datasets in wildlands represents a notable gap.

There are stark differences between remote sensing applications in agricultural and wildland contexts. The diversity of species present in wildland systems is often far greater than those in agricultural contexts (Phalan et al., 2011). Thus, for tasks such as image classification, discerning target plants from background species can be more difficult, as the background domain is more varied. Furthermore, identifying diagnostic features of target plants, such as flower and leaf morphology, is challenging, requiring botanical expertise in the field. It may be impossible to resolve such features in coarse-grained images, necessitating the use of drone-based platforms that can fly lower and gather fine-grained information (Gallmann et al., 2022; Amputu et al., 2023).The terrain of wildlands is also more complex than in agricultural sites, which are situated in flat areas to facilitate mechanized tillage, seeding, and pest management (Özkan et al., 2020). As a consequence, gathering data in wildlands is time consuming and costly due to steep terrain, lack of roads, and poor connectivity for navigation systems. Additionally, complex terrain generates varied lighting conditions (Corripio, 2003), which may pose challenges for classifiers (Rodriguez-Galiano and Chica-Olmo, 2012). Therefore, applying machine learning solutions to remote sensing of wildland phenomena is inherently more difficult than in agricultural domains because gathering data is costly and the content of images more diverse.

One important use-case of remote sensing in wildlands is that of weed plant detection. When an invasive plant invades a natural area, it can cause harm by a variety of mechanisms, including competition with native plants for space and resources (Maron and Marler, 2008), disruption of pollinator services (Pearson et al., 2012), catalyzing catastrophic wildfire regimes (Balch et al., 2013; Bradley et al., 2018), and others. Once established, invasive plants are difficult to remove, requiring expensive monitoring and treatment. Leafy spurge (*Euphorbia esula*; **Fig. 1**) is an example case introduced to North America in the late 19th century that quickly spread from agricultural areas to wildlands. This noxious weed is avoided by cattle and wild grazing mammals, resulting in economic losses exceeding $100 million in the northern Great Plains (Leistritz et al., 2004). Extensive control methods have been employed to manage its spread (Gaskin et al., 2021), and efforts to monitor expansion are of growing interest. A recent study utilizing 4 m satellite imagery achieved a field-validated accuracy of 0.59 (Mattillo et al., 2023), while a prior drone imaging study utilizing 3 cm data achieved an accuracy of 0.78 for flowering spurge plants occupying more than 10% ground cover (Yang et al., 2021). Therefore, current evidence suggests higher-resolution imagery can enhance classifier performance, but the task of remotely sensing leafy spurge is not solved, and could prove challenging.

We gathered high resolution drone imagery of grasslands undergoing ecological restoration in western Montana, USA where leafy spurge has established and is targeted for removal. In parallel we collected ground truth of spurge presence and absence in the field with precision GPS systems throughout the study region. Given the costly nature of acquiring ecological data from wildlands, we assert that there is a natural pairing with few-shot research and the study of generalization capacity of foundation models. Therefore, we focus our experiments on performance in the context of data limitation. Specifically, we evaluated classifiers tasked with identifying leafy spurge when ablating training data quantity, testing both proven CNN and state-of-the-art vision transformer architectures. Additionally, we explored the few-shot performance of large multi-modal models to evaluate their generalization capabilities when applied to our datasets, which is out of the training set distribution. We release and describe these data here for the purpose of advancing spurge detection and management as well as furnishing the machine learning research community with a unique, real-world dataset.

2 Image acquisition and post-processing

We surveyed the study area on June 12, 2023, during a 4-hour window (from 11:11 to 15:11) with a DJI Mavic 3M drone. The drone captured 8241 images at 50 m above ground level across an area of 118 hectares (**Fig. 2**). During the survey there was light wind and sparse cloud cover at 3700 m. We programmed the drone flight such that images overlapped to improve performance of the feature matching algorithm during post-processing, which merges raw images into a single spatially contiguous and georeferenced image, or orthomosaic. The side overlap ratio was 70% and a front overlap ratio was 80% between adjacent images. This process generated an estimated ground sampling distance (GSD) of 1.27 cm per pixel in the resultant orthomosaic. A key benefit of orthosmosaicing imagery is to georeference each pixel (as opposed to only centroids of single images), which enables predictive mapping of the target weed with downstream classifiers.

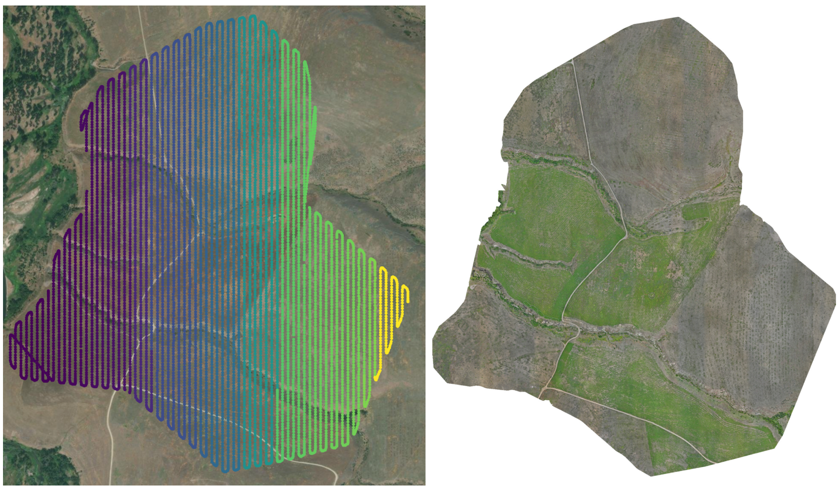

Figure 2: Left: A map of the drone survey flight plan where each point represents the coordinates of a photo in the survey, colored by time of photo. Right: These overlapping images are processed into a large, georeferenced orthomosaic (right).

We performed all post-processing with the DroneDeploy software suite (`https://www.dronedeploy.com`). This involved feature matching of overlapping images, correction of geometric distortions, and the generation of a georeferenced orthomosaic. Prior to surveying, we established 32 ground control points (GCPs; points for which positions were verified with a precision GPS) across the study area. These GCPs were used during post-processing to further minimize georeferencing error of pixels in the orthomosaic product. The Root Mean Squared Error (RMSE) of GCP position was 7.32 cm after post-processing.

3 Ground truth acquisition

After surveying our study area botanists visited sites within to gather ground truth of spurge presence and absence. Upon visiting a site, botanists conducted random walks to gather coordinates of spurge presence and absence using an Emlid RS2 GPS, capable of reporting positions with sub-centimeter accuracy. When a technician encountered a target plant, they would record its position. Technicians also gathered coordinates for spurge absences in this manner. For spurge absence cases, ground truth indicates that no spurge plants were detected in a 0.5 x 0.5m box centered on the coordinates. During the walk technicians gathered data until they acquired 50 presences and 50 absences at each site. We visited a total of 10 sites (**Fig. 3**; top left panel), sampling in this manner, and accumulated 500 examples per presence/absence class (**Fig. 3**, bottom panels). While the majority of sites were geographically separated, two sites overlapped due to weather-related time constraints. These were assigned to the same training split (strategy described below) to ensure no data leakage to the test set.

We extracted training images from crops of our orthoimage corresponding to ground truth bounding boxes as well as a larger size (described below) to use in downstream classification. Each instance of ground truth corresponds to 0.5 x 0.5m, or 39 x 39 pixels (**Fig. 3**; top right panel).

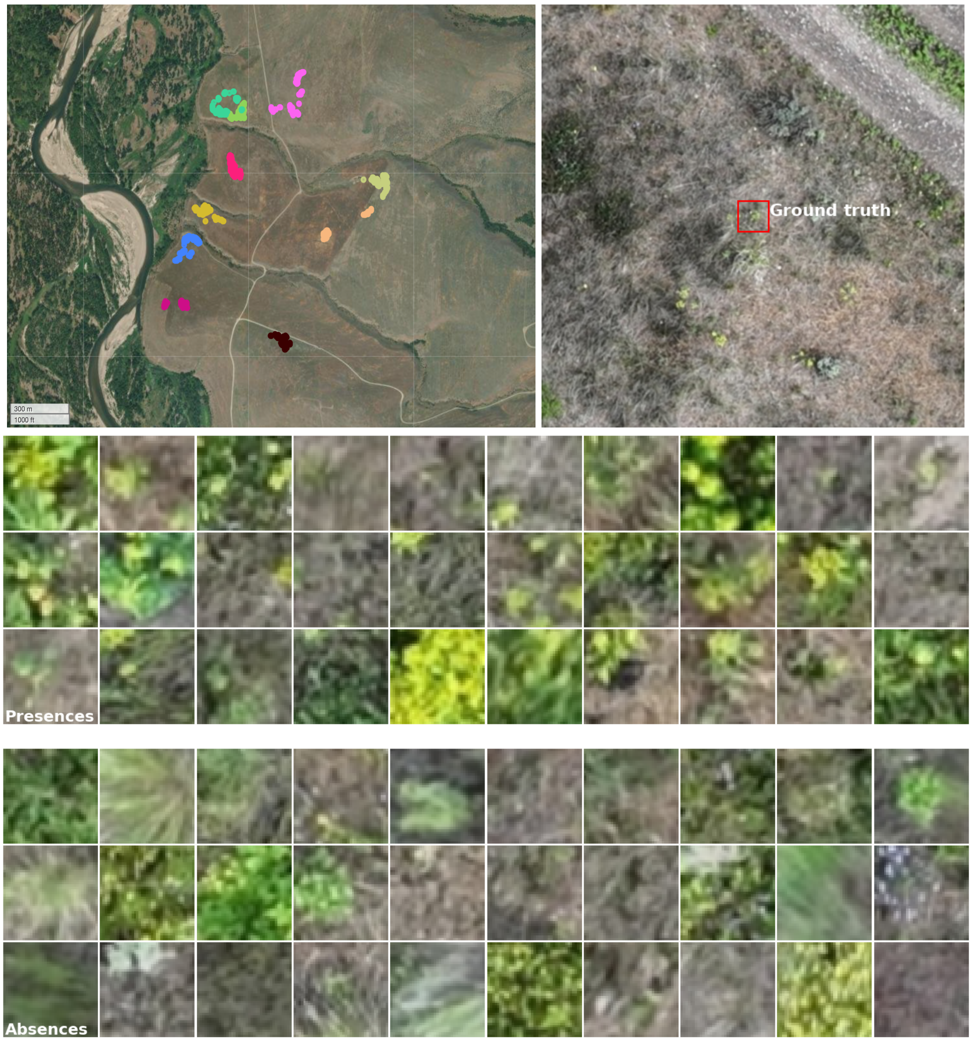

Figure 3: Top Left: Locations of leafy spurge ground truth samples. Point colors indicate sampling site identity. Top Right: The red box indicates the dimensions of a leafy spurge ground truth sample (0.5 x 0.5m). Bottom: Examples from presence and absence classes in the dataset.

4 Benchmarks of Leafy Spurge Classifiers

4.1 Assignment of Data Splits

We used the geographic separation of our sampling sites as the basis for splitting data for evaluation. We selected eight sites (800 image/label pairs) for the training set and selected the two remaining sites (200 image/label pairs) for test sets. The data from one of these test sites we release with this publication, while we reserve the data from the second site for evaluation of progress at a later date. The intent of establishing test data spatially is to simulate performance on new data gathered from recently invaded areas.

4.2 Conventional Vision Architectures

We evaluated two computer vision architectures for the image classification task: ResNet50 (He et al., 2016) and DINOv2 (Oquab et al., 2023). ResNet50 is a widely adopted convolutional neural network, while DINOv2 is a more recent vision transformer-based model. We used pre-trained weights from their respective checkpoints (`facebook/DINOv2-base` for DINOv2) for initialization. We preprocessed the images by normalizing with ImageNet statistics and resizing to 224x224 pixels. We conducted experiments with two image sizes from the dataset: 39x39 pixels ('crop' revision) and 1024x1024 pixels ('context' revision). The intent of training with larger, 1024 pixel images was to explore if broader context around the ground truth could aid classifier performance, as is reported for larger vision transformer models (McKinzie et al., 2024). To enhance model generalization and mitigate overfitting, we applied the following data augmentation techniques during training with a probability of 0.5: ColorJitter (brightness=0.8, contrast=0.7, saturation=0, hue=0), RandomHorizontalFlip, RandomVerticalFlip, and RandomRotation (degrees=90).

We trained for 50 epochs using the Adam optimizer (Kingma and Ba, 2017) with a learning rate of 0.0001 and batch size of 32 for both ResNet50 and DINOv2. For the DINOv2 training we applied Low-rank Adaptation (Hu et al., 2021) with rank and alpha parameters set to 8. In addition to testing performance on the full dataset, we conducted few-shot experiments, randomly sampling without replacement 1, 2, 4, 8, 16, 32, 64, 128, and 256 examples per class. For each experiment (combination of model and image size, dataset revision, and examples per clas), we tested 8 unique seeds to account for variability. For each seed, we split the training set into 80% for training and 20% for validation. We evaluated model performance on both the validation and test sets during training. For our performance metric we calculated the 95% confidence interval of the proportion of correctly classified samples. Total compute for these experiments was 248 hours on an internal cluster of 40 2080ti GPUs. The longest period of training for a single seed was one hour and eighteen minutes, observed with DINOv2 architecture. This experiment is representative of the time it would take to train a production model, and could be replicated with publicly available services, such as GPU-equipped Google Colab runtimes, for less than $10 USD at this time.

4.3 Dialogue and Few-shot Experiments with GPT-4-turbo and GPT-4o

We conducted a qualitative dialogue with the large multi-modal model, GPT-4o (OpenAI et al. (2024); **Box 5.2**). Our objective was to explore its zero-shot performance when shown a single 1024x1024 pixel image of areas invaded by leafy spurge. An ecologist familiar with the study area drafted a series of prompts to assess GPT-4o's capacity to identify the target with increasing levels of context. Initially, we prompted the model to identify what is in the image. We then asked it to identify the location of the image. Finally, we provided GPT-4o with the region and asked it to identify which weeds were present.

We also assessed GPT-4-turbo and GPT-4o models in few-shot contexts. We provided a system prompt stating that the model is an expert in ecology and plant species, and that it would be viewing aerial imagery. We then prompted the models with images and labels for 1, 2, 4, 8, and 16 examples per class from the training set, then selected a test set image at random, asking it to indicate if spurge was present or absent. We sampled 32 examples per class for test set evaluation in this manner. We evaluated eight random seeds. We sampled the test set to mitigate API costs. The images were resized to 256 pixels without additional augmentation. Total compute for these experiments was 28 hours on our internal cluster of 40 2080ti GPUs.

5 Results

5.1 Performance of Conventional Vision Architectures

We found both DINOv2 and ResNet50 architectures suitable for detection of the target plant, leafy spurge, though performance was contingent on image size (**Table 1**, **Fig. 4**). Performance of each model type was similar when trained on the full dataset of smaller images, but ResNet50 could not satisfactorily classify spurge in larger images (1024x1024 pixels). Notably, models trained on smaller images (39x39) whose dimensions correspond directly to the 0.5x0.5m bounds of ground truth performed better with fewer examples per class than those trained on larger images (1024x1024). Both the full dataset and few-shot results diverged from our expectation that the vision transformer model, DINOv2, would benefit from higher resolution images. Our results indicate that, given sufficient data, DINOv2 can isolate salient information even when supplied with a broader context, though this is not the case with ResNet50. The DINOv2 confidence interval was wider for the larger images than for smaller ones, suggesting that this capacity to identify the target amidst broader context was dataset-dependent.

Table 1: Test accuracy when trained on 350 examples per class with ResNet50 and DINOv2 architectures and two image sizes. We report 95% confidence intervals of accuracy metrics in parentheses.

Architecture	Image Size	Test Accuracy
DINOv2	39x39	0.85 (0.82, 0.87)
ResNet50	39x39	0.84 (0.79, 0.88)
DINOv2	1024x1024	0.83 (0.77, 0.92)
ResNet50	1024x1024	0.65 (0.60, 0.75)

5.2 Few-shot Experiments and Dialogue with GPT-4-turbo and GPT-4o

We found that GPT-4 checkpoints Turbo and Omni had some capacity to classify test data when shown a handful of example images of leafy spurge presence and absence from the training set. We observed best performance with Omni, which achieved a test accuracy of 0.75 when shown eight examples per class of the 39x39 pixel images (**Fig. 4**). Turbo was less accurate (0.68), with best performance observed when shown four examples per class of 39x39 pixel images. Both models were unable to classify the larger image sizes. Notably, performance was not monotonically related to training set size, where, e.g., Omni performance decreased from 8 to 16 examples per class, though it is unclear what might be driving this pattern.

In dialogues with GPT-4o the model stated that our prompt image (**Fig. 5**; **Box 5.2**) was taken from the air. Furthermore, it suggested correctly that the image could be of a semi-arid grassland. When prompted with the region where the image was taken and asked to identify which weed species may be present, GPT-4o listed leafy spurge among candidate weeds and selected it among the two most likely species based upon the yellow flowers observed in the image.

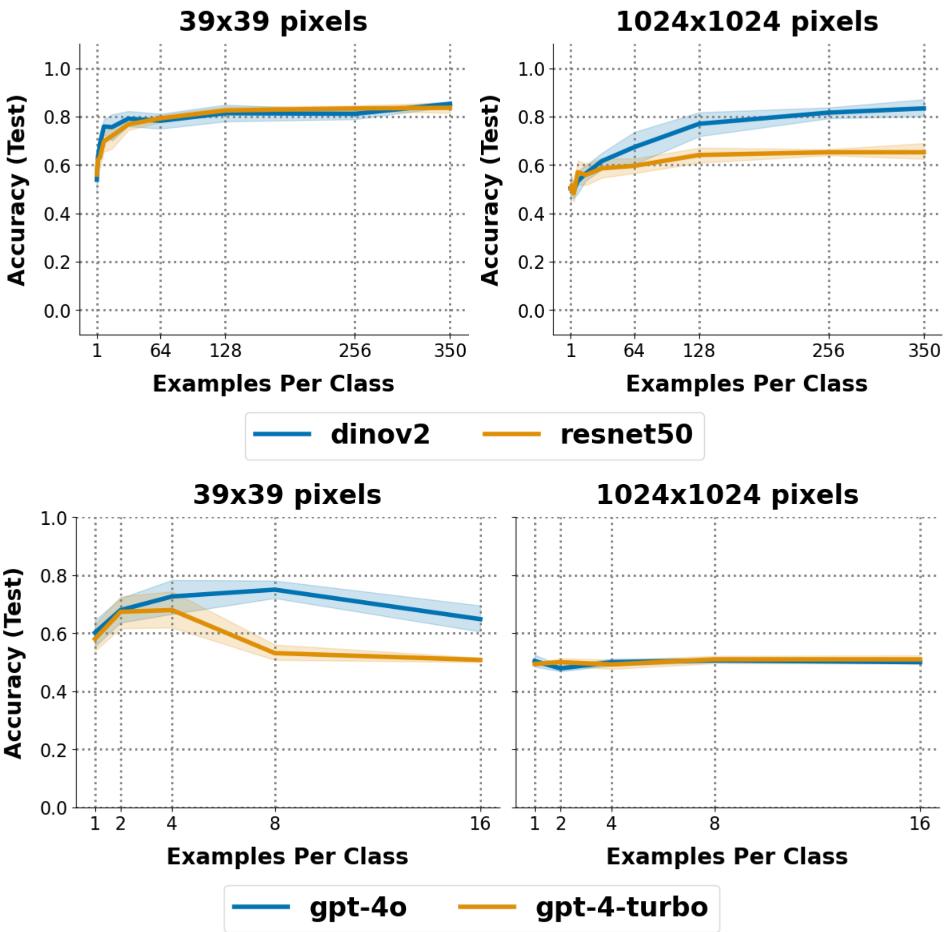

Figure 4: We present classification results on the Leafy Spurge Dataset using different architectures and models. The top panel shows results for ResNet50 (orange) and DINOv2 (blue) when training on 39x39 pixel (left panel) and 1024x1024 pixel (right panel) images. The bottom panel presents few-shot classification results of GPT-4-turbo (orange) and GPT-4o (blue) foundation models on the same dataset with similar image sizes. For both panels, we report test set accuracy across resampled dataset sizes (1 through 256 examples per class in the top panel and 1 through 16 examples per class in the bottom panel) and the full dataset in the top panel (350 examples per class). Cross-validated means (eight seeds) are represented as lines, with 95% confidence intervals as bands.

Figure 5: A sample image from Leafy Spurge Dataset included with text prompts in a dialogue with GPT-4o (**Box 5.2**).

GPT-4o Dialogue Box

RESEARCHER: What is in this image?

GPT-4O: The image appears to be an aerial view of a patch of land. The ground is mostly covered with sparse vegetation, including some small green plants and weeds. The area looks dry and there are some patches of bare soil. There are a few clusters of green plants and some yellow-flowered plants scattered across the area.

RESEARCHER: In what region was this image taken?

GPT-4O: Based on the image alone, it is difficult to definitively determine the specific region where it was taken. The vegetation and dry soil could be indicative of many different regions around the world, such as arid or semi-arid areas, grasslands, or even certain temperate zones. Without additional context or information such as geographical markers, specific types of flora, or metadata, it's not possible to accurately identify the exact location.

RESEARCHER: The image was taken in the Bitterroot Valley, Montana, USA. What weeds are in this image?

GPT-4O: In the Bitterroot Valley of Montana, there are several types of weeds that commonly grow. Based on the image, which shows some scattered green vegetation and flowering plants, here are a few common weeds you might find:

1. Spotted Knapweed (*Centaurea stoebe*) Typically produces purple flowers. The plant has deeply lobed leaves and can form dense stands.
2. Dalmatian Toadflax (*Linaria dalmatica*) Characterized by yellow, snapdragon-like flowers. It has waxy leaves and can grow tall.
3. Houndstongue (*Cynoglossum officinale*) Produces reddish-purple flowers. The leaves are large, rough, and hairy.
4. Leafy Spurge (*Euphorbia esula*) Known for its yellow-green bracts. The plant has narrow leaves and produces a milky sap when broken.
5. Canada Thistle (*Cirsium arvense*) Features small purple or pink flower heads. The leaves are spiny and it often forms dense patches.

Given the yellow flowering plants visible in the image, there is a possibility that some of them could be Leafy Spurge or Dalmatian Toadflax. However, identifying plants positively based only on an aerial image can be challenging. For accurate identification, closer, more detailed photos or in-person examination would be needed. Consulting a local botanist or using plant identification guides specific to the area can also be helpful.

6 Relevance to the Study of Foundation Models

We demonstrated how the novelty of Leafy Spurge Dataset can be leveraged to fairly evaluate the generalization abilities of large multi-modal models such as GPT-4o (OpenAI et al., 2024). When tasked with classification of leafy spurge from few example cases, this model dramatically outperformed GPT-4-turbo, illustrating the rapid progress in recent months. Yet, fine-tuning conventional vision models (DINOv2 and Resnet50) with the full dataset offered best overall performance. Upon initial consideration, our dataset may appear niche, however its novelty is an asset in evaluating models trained on internet-scale data. Zero and few-shot researchers currently struggle to source data that are truly outside the domain of the training set of large models. For example, common semantic concepts, such as mammal species, are well represented in the corpus of the text-to-image diffusion model, such as Stable Diffusion (Rombach and Ommer, 2021), and naive evaluations of zero and few-shot generations of these concepts may not be fair due to data leakage (Trabucco et al., 2023). Datasets of wildlands phenomena are under-represented, generally, due to the myriad challenges of data collection we describe earlier. We hope our work highlights their importance to the machine learning research community and spurs future initiatives to develop ecological datasets that can rigorously test the boundaries of model generalization.

7 Applied Considerations for Future Study

We hope that those exploring our data will tailor their work to benefit the land management community whose objective is to contain and remove leafy spurge. One primary consideration is that of the spatial scale of spurge predictions. At present, leafy spurge plants are treated with herbicide by a human applicator, who is able to target individual plants or groups of plants. For this tasks, knowledge of plant presence on the landscape at 0.5 m scale (the scale of a single plant) would be more than adequate for successful treatment. In contrast, pixel-level inference (leaf-scale) offers no practical benefit, as applicators can not spray at 1.3 cm resolution. Therefore, mapping spurge extents by tiling out the orthomosaic into thumbnails of this size, conducting inference on tiles, and reconstituting the products into a mosaic would indeed be useful. In addition to labelled images we serve the full unlabelled orthomosaic, excluding test regions. Unsupervised learning on these data might enhance classifier performance and a successful application of this technique could benefit the broader field of remote sensing where vast amounts of unlabelled aerial images are a common condition.

8 Limitations and Ethical Considerations

Leafy spurge has spread throughout regions of North America, though we chose to focus specifically on an Inter-mountain West grassland, an ecosystem estimated to cover 100,000 hectares but once thought to be 11.7 million hectares (Belovsky and Slade, 2020). Prior estimates of wildflower and grass species counts in this grassland type are 69 and 27, respectively (Mueggler and Stewart, 1980; Pokorny et al., 2004); values representing aggregations from many sites. Among these grassland sites, our survey area is notably diverse, harboring 74 wildflower species and 27 grass species. Therefore, classifiers trained on or data could be expected to learn a diversity of plant forms extensible throughout Inter-mountain West grasslands. It is unclear how classifiers would perform outside of our grassland type, where background vegetation would be dramatically different. Thus, deploying models trained on our data is unlikely to *solve* spurge detection at continental scales, however, given the dramatic reduction in extent of Inter-mountain West grasslands, such models would greatly benefit conservation and restoration efforts of these threatened, high-value areas.

We do not provide segmentation masks that might be used to localize the aerial canopy of plants. This is because, as discussed above, pixel-level inference offers no practical benefit. Additionally, collecting ground truth in 1.3 cm imagery is not possible, as map georeferencing error is greater than a pixel, and leaf position changes with the wind. False-negative rates would be high where spurge is patchily distributed and harbors thousands of plants lacking flowers, which could elude even trained botanists. For this reason, post-hoc annotation of imagery would also be unreliable, because plants not showing flowers would evade detection. Additionally, even those with flowers could be confused with non-target plants with yellow flowers (**Fig.** 3).

Efforts such as ours to automate plant identification highlight a complex outlook for the job prospects of natural resource specialists, such as botanists, but enumerable future monitoring applications will require human expertise to assemble ground truth for efforts that share our aims.

9 Conclusions

The Leafy Spurge Dataset offers the machine learning community an opportunity to study foundation model generalization performance on an out-of-distribution real world problem, while advancing the fields of ecology and remote sensing in parallel. In our initial analyses, we found that the task of leafy spurge presence/absence classification is tractable, but not solved. In future work we hope to explore performance benefits when incorporating diverse modalities of data offered by drone surveys, including structure-from-motion elevation products and multi-spectral images, which could be paired with standard three-channel data.

10 Acknowledgments

We would like to thank the staff at MPG Ranch for assistance in gathering ground truth and aerial imagery, as well as the owners of MPG Ranch for funding this work and for their ongoing support of applied ecological, conservation, and machine learning research.

References

Amputu, V., Knox, N., Braun, A., Heshmati, S., Retzlaff, R., Röder, A., and Tielbörger, K. (2023). Unmanned aerial systems accurately map rangeland condition indicators in a dryland savannah. *Ecological Informatics*, 75:102007.

arura uav (2023). Uav tree identification - new dataset. `https://universe.roboflow.com/arura-uav/uav-tree-identification-new`. visited on 2024-08-16.

Balch, J., Bradley, B., D'Antonio, C., and Gomez-Dans, J. (2013). Introduced annual grass increases regional fire activity across the arid western usa (1980–2009). *Global Change Biology*, 19:173–183.

Belovsky, G. E. and Slade, J. B. (2020). Climate change and primary production: Forty years in a bunchgrass prairie. *PLoS ONE*, 15(12):e0243496.

Bradley, B., Curtis, C., Fusco, E., Abatzoglou, J., Balch, J., Dadashi, S., and Tuanmu, M. (2018). Cheatgrass (bromus tectorum) distribution in the intermountain western united states and its relationship to fire frequency, seasonality, and ignitions. *Biological invasions*, 20:1493–1506.

Chiu, M. T., Xu, X., Wei, Y., Huang, Z., Schwing, A., Brunner, R., Khachatrian, H., Karapetyan, H., Dozier, I., Rose, G., Wilson, D., Tudor, A., Hovakimyan, N., Huang, T. S., and Shi, H. (2020). Agriculture-vision: A large aerial image database for agricultural pattern analysis. *arXiv preprint arXiv:2001.01306*. cs.CV.

Corripio, J. G. (2003). Vectorial algebra algorithms for calculating terrain parameters from dems and the position of the sun for solar radiation modelling in mountainous terrain. *International Journal of Geographical Information Science*, 17(1):1–23.

dos Santos Ferreira, A., Pistori, H., Freitas, D. M., and da Silva, G. G. (2017). Data for: Weed Detection in Soybean Crops Using ConvNets. (available at: https://doi.org/10.17632/3fmjm7ncc6.2).

FAO (2020). Global Forest Resources Assessment 2020. (available at: https://doi.org/10.4060/ca8753en).

Gallmann, J., Schüpbach, B., Jacot, K., Albrecht, M., Winizki, J., Kirchgessner, N., and Aasen, H. (2022). Flower mapping in grasslands with drones and deep learning. *Frontiers in Plant Science*, 12:774965. This article is part of the Research Topic "Plant Biodiversity Science in the Era of Artificial Intelligence".

Gaskin, J., Espeland, E., Johnson, C., Larson, D., Mangold, J., McGee, R., and Tekiela, D. (2021). Managing invasive plants on great plains grasslands: A discussion of current challenges. *Rangeland Ecology & Management*, 78:235–249.

Genze, N., Ajekwe, R., Grieb, M., and Grimm, D. (2022). Deep learning-based early weed segmentation using motion blurred UAV images of sorghum fields. (available at: https://doi.org/10.17632/4hh45vkp38.4).

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., and Chen, W. (2021). Lora: Low-rank adaptation of large language models.

Kingma, D. P. and Ba, J. (2017). Adam: A method for stochastic optimization.

Krestenitis, M., Raptis, E. K., Kapoutsis, A. C., Ioannidis, K., Kosmatopoulos, E. B., Vrochidis, S., and Kompatsiaris, I. (2022). Cofly-weeddb: A uav image dataset for weed detection and species identification. *Data in Brief*, 45:108575. Open access under a Creative Commons license.

Lamb, D. and Brown, R. (2001). Using remote sensing to identify, map, and monitor the distribution of weeds. *Journal of Agricultural Engineering Research*, 78(2):117–125.

Leistritz, F., Bangsund, D., and Hodur, N. (2004). Assessing the economic impact of invasive weeds: the case of leafy spurge (euphorbia esula). *Weed Technology*, pages 1392–1395.

Maron, J. and Marler, M. (2008). Field-based competitive impacts between invaders and natives at varying resource supply. *Journal of Ecology*, 96(6):1187–1197.

Mattillo, C. M., Tekiela, D. R., and Norton, U. (2023). Remote mapping of leafy spurge (*Euphorbia esula*, l.) in northwestern colorado. *Frontiers in Remote Sensing*, 4. Part of the Research Topic: Women in Remote Sensing: 2022.

McKinzie, B., Gan, Z., Biard, J.-P. F., Dodge, S., Dufter, P., Zhang, B., Shah, D., Du, X., Peng, F., Zhang, H., Weers, F., Belyi, A., Singh, K., Kang, D., Jain, A., He, H., Schwarzer, M., Gunter, T., Kong, X., Zhang, A., Wang, J., Wang, C., Du, N., Lei, T., Wiseman, S., Lee, M., Wang, Z., Pang, R., Grasch, P., Toshev, A., and Yang, Y. (2024). Mm1: Methods, analysis & insights from multimodal llm pre-training. *arXiv preprint arXiv:2403.09611*.

Mowla, M. N., Asadi, D., Tekeoglu, K. N., Masum, S., and Rabie, K. (2024). Uavs-ffdb: A high-resolution dataset for advancing forest fire detection and monitoring using unmanned aerial vehicles (uavs). *Data in Brief*, 55:110706. Under a Creative Commons license.

Mueggler, W. and Stewart, W. (1980). *Grassland and shrubland habitat types of western Montana*. Intermountain Forest and Range Experiment Station, Ogden, UT.

Mulla, D. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114(4):358–371.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., Bello, I., Berdine, J., Bernadett-Shapiro, G., Berner, C., Bogdonoff, L., Boiko, O., Boyd, M., Brakman, A.-L., Brockman, G., Brooks, T., Brundage, M., Button, K., Cai, T., Campbell, R., Cann, A., Carey, B., Carlson, C., Carmichael, R., Chan, B., Chang, C., Chantzis, F., Chen, D., Chen, S., Chen, R., Chen, J., Chen, M., Chess, B., Cho, C., Chu, C., Chung, H. W., Cummings, D., Currier, J., Dai, Y., Decareaux, C., Degry, T., Deutsch, N., Deville, D., Dhar, A., Dohan, D., Dowling, S., Dunning, S., Ecoffet, A., Eleti, A., Eloundou, T., Farhi, D., Fedus, L., Felix, N., Fishman, S. P., Forte, J., Fulford, I., Gao, L., Georges, E., Gibson, C., Goel, V., Gogineni, T., Goh, G., Gontijo-Lopes, R., Gordon, J., Grafstein, M., Gray, S., Greene, R., Gross, J., Gu, S. S., Guo, Y., Hallacy, C., Han, J., Harris, J., He, Y., Heaton, M., Heidecke, J., Hesse, C., Hickey, A., Hickey, W., Hoeschele, P., Houghton, B., Hsu, K., Hu, S., Hu, X., Huizinga, J., Jain, S., Jain, S., Jang, J., Jiang, A., Jiang, R., Jin, H., Jin, D., Jomoto, S., Jonn, B., Jun, H., Kaftan, T., Łukasz Kaiser, Kamali, A., Kanitscheider, I., Keskar, N. S., Khan, T., Kilpatrick, L., Kim, J. W., Kim, C., Kim, Y., Kirchner, J. H., Kiros, J., Knight, M., Kokotajlo, D., Łukasz Kondraciuk, Kondrich, A., Konstantinidis, A., Kosic, K., Krueger, G., Kuo, V., Lampe, M., Lan, I., Lee, T., Leike, J., Leung, J., Levy, D., Li, C. M., Lim, R., Lin, M., Lin, S., Litwin, M., Lopez, T., Lowe, R., Lue, P., Makanju, A., Malfacini, K., Manning, S., Markov, T., Markovski, Y., Martin, B., Mayer, K., Mayne, A., McGrew, B., McKinney, S. M., McLeavey, C., McMillan, P., McNeil, J., Medina, D., Mehta, A., Menick, J., Metz, L., Mishchenko, A., Mishkin, P., Monaco, V., Morikawa, E., Mossing, D., Mu, T., Murati, M., Murk, O., Mély, D., Nair, A., Nakano, R., Nayak, R., Neelakantan, A., Ngo, R., Noh, H., Ouyang, L., O'Keefe, C., Pachocki, J., Paino, A., Palermo, J., Pantuliano, A., Parascandolo, G., Parish, J., Parparita, E., Passos, A., Pavlov, M., Peng, A., Perelman, A., de Avila Belbute Peres, F., Petrov, M., de Oliveira Pinto, H. P., Michael, Pokorny, Pokrass, M., Pong, V. H., Powell, T., Power, A., Power, B., Proehl, E., Puri, R., Radford, A., Rae, J., Ramesh, A., Raymond, C., Real, F., Rimbach, K., Ross, C., Rotsted, B., Roussez, H., Ryder, N., Saltarelli, M., Sanders, T., Santurkar, S., Sastry, G., Schmidt, H., Schnurr, D., Schulman, J., Selsam, D., Sheppard, K., Sherbakov, T., Shieh, J., Shoker, S., Shyam, P., Sidor, S., Sigler, E., Simens, M., Sitkin, J., Slama, K., Sohl, I., Sokolowsky, B., Song, Y., Staudacher, N., Such, F. P., Summers, N., Sutskever, I., Tang, J., Tezak, N., Thompson, M. B., Tillet, P., Tootoonchian, A., Tseng, E., Tuggle, P., Turley, N., Tworek, J., Uribe, J. F. C., Vallone, A., Vijayvergiya, A., Voss, C., Wainwright, C., Wang, J. J., Wang, A., Wang, B., Ward, J., Wei, J., Weinmann, C., Welihinda, A., Welinder, P., Weng, J., Weng, L., Wiethoff, M., Willner, D., Winter, C., Wolrich, S., Wong, H., Workman, L., Wu, S., Wu, J., Wu, M., Xiao, K., Xu, T., Yoo, S., Yu, K., Yuan, Q., Zaremba, W., Zellers, R., Zhang, C., Zhang, M., Zhao, S., Zheng, T., Zhuang, J., Zhuk, W., and Zoph, B. (2024). Gpt-4 technical report.

Oquab, M., Darcet, T., Moutakanni, T., Vo, H. V., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Howes, R., Huang, P.-Y., Xu, H., Sharma, V., Li, S.-W., Galuba, W., Rabbat, M., Assran, M., Ballas, N., Synnaeve, G., Misra, I., Jegou, H., Mairal, J., Labatut, P., Joulin, A., and Bojanowski, P. (2023). Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*.

Özkan, B., Dengiz, O., and Turan, I. D. (2020). Site suitability analysis for potential agricultural land with spatial fuzzy multi-criteria decision analysis in regional scale under semi-arid terrestrial ecosystem. *Scientific Reports*, 10(1):22074.

Pearson, D., Ortega, Y., and Sears, S. (2012). Darwin's naturalization hypothesis up-close: Intermountain grassland invaders differ morphologically and phenologically from native community dominants. *Biological Invasions*, 14:901–913.

Pettorelli, N., Laurance, W., O'Brien, T., Wegmann, M., Nagendra, H., and Turner, W. (2014). Satellite remote sensing for applied ecologists: opportunities and challenges. *Journal of Applied Ecology*, 51(4):839–848.

Phalan, B., Balmford, A., Green, R., and Scharlemann, J. (2011). Minimising the harm to biodiversity of producing more food globally. *Food Policy*, 36:S62–S71.

Pokorny, M., Sheley, R., Svejcar, T., and Engel, R. (2004). Plant species diversity in a grassland plant community: evidence for forbs as a critical management consideration. *Western North American Naturalist*, 64(2):219–230.

Rodriguez-Galiano, V. and Chica-Olmo, M. (2012). Land cover change analysis of a mediterranean area in spain using different sources of data: Multi-seasonal landsat images, land surface temperature, digital terrain models and texture. *Applied Geography*, 35(1-2):208–218.

Rombach, R. and Ommer, B. (2021). High-resolution image synthesis with latent diffusion models. *arXiv preprint arXiv:2112.10752*.

Trabucco, B., Doherty, K., Gurinas, M., and Salakhutdinov, R. (2023). Effective data augmentation with diffusion models. *arXiv preprint arXiv:2302.07944*.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., and Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, 18(6):306–314.

Weinstein, B., Graves, S., Marconi, S., Singh, A., Zare, A., Stewart, D., Bohlman, S., and White, E. (2021). A benchmark dataset for canopy crown detection and delineation in co-registered airborne rgb, lidar and hyperspectral imagery from the national ecological observation network. *PLOS Computational Biology*, 17:e1009180.

Wildeboer, S. (2023). Megaweeds dataset. Consists of seven existing datasets, including WeedCrop, Chicory, Sesame, Sugar beet, Weed-Detection v2, Maize, and CottonWeed12 datasets.

Yang, X., Smith, A., Bourchier, R., Hodge, K., Ostrander, D., and Houston, B. (2021). Mapping flowering leafy spurge infestations in a heterogeneous landscape using unmanned aerial vehicle red-green-blue images and a hybrid classification method. *International Journal of Remote Sensing*, 42(23):8930–8951.

11 Appendix A - Data Access and Licensing

11.1 Website

Our website serves as a portal to access our Hugging Face dataset and GitHub code repositories for reproducing our analyses. The DOI for our dataset is: **doi:10.57967/hf/2508**. We will use the website as a venue for news and updates related to our work. We will ensure our dataset is maintained via Hugging Face's service.

11.2 Licensing & Author Responsibility

We authorize use of our data with a Creative Commons Attribution 4.0 International license. The authors bear all responsibility for violations of rights, and affirm that no rights were violated in the process of collecting data, nor the process of making the data accessible to the public. We confirm that all images in this work were captured on private property with the landowner's permission and drone flights were conducted in strict compliance with FAA regulations and relevant local laws. The views expressed in this paper are those of the authors and do not necessarily reflect the views of our respective institutions.

11.3 Dataset Access

Our data are hosted publicly as a Hugging Face Dataset (`https://www.huggingface.co`). We serve two image sizes corresponding to the 39x39 and 1024x1024 pixel images as configs "crop" and "context", respectively. Additionally, we serve the full unlabelled orthomosaic as config "unlabelled."

Listing 1: Python code for loading our dataset from Hugging Face.

```
from datasets import load_dataset
import requests

#39x39 pixel train and test data
crop_train = load_dataset('mpg-ranch/leafy_spurge',
    'crop', split='train')
crop_test = load_dataset('mpg-ranch/leafy_spurge',
    'crop', split='test')

#1024x1024 pixel train test data
context_train = load_dataset('mpg-ranch/leafy_spurge',
    'context', split='train')
context_test = load_dataset('mpg-ranch/leafy_spurge',
    'context', split='test')

#unlabelled orthomosaic
orthomosaic = load_dataset('mpg-ranch/leafy_spurge',
    'unlabelled', split='train')
```

11.4 Croissant Metadata Access

The Croissant metadata are accessible via the Hugging Face API as shown below.

Listing 2: Python code for loading Croissant Metadata from Hugging Face.

```
#access Croissant metadata via the Hugging Face API
headers = {"Authorization": f"Bearer {API_TOKEN}"}
API_URL = "https://huggingface.co/api/datasets/mpg-ranch/leafy_spurge/
    ↪ croissant"
def query():
    response = requests.get(API_URL, headers=headers)
    return response.json()
croissant_metadata = query()
```

12 Appendix B - Dataset Datasheet

Questions from the Datasheets for Datasets (`https://arxiv.org/abs/1803.09010`) paper, v7.

Jump to section:

- **Motivation** 12.1
- **Composition** 12.2

- **Collection process** 12.3
- **Preprocessing/cleaning/labeling** 12.4
- **Uses** 12.5
- **Distribution** 12.6
- **Maintenance** 12.7

12.1 Motivation

The dataset was created to promote research of ecological monitoring of problem weeds with aerial drones in a natural setting. Developing robust classifiers from this dataset will benefit the commmuntiy of ecologists who hope to control leafy spurge (*Euphorbia esula*) a plant that harms the ecological integrity of many areas in North America. Because of its novelty, the dataset offers zero and few-shot researchers an opportunity to explore data that lie outside the training set of large generative models.

For what purpose was the dataset created? We lack agile and robust systems to monitor and map the invasion of weed plants into high-value natural areas. We hope this dataset will enable the training of powerful image classifiers for this purpose.

Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)? This dataset was created by staff scientists at MPG Ranch, a conservation group in western Montana, USA, with the guidance of collaborators at the Machine Learning Department of Carnegie Mellon University.

Who funded the creation of the dataset This work was funded by the owners of MPG Ranch.

Any other comments? None.

12.2 Composition

What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)? Each instance of data is a top-down aerial image paired with a binary label: leafy spurge present or absent. Botanists visited the coordinates at the center of each image to ground truth plant presence. The ground truth corresponds to a 0.5 x 0.5m area. Two versions of the data are presented: one in which only the ground truth pixels (39 x 39) are visible, and one in which the context surrounding the ground turth is included (1024x1024). The images were extracted from a large orthomosaic produced from a drone survey with a DJI Mavic 3M drone. Additionally, we provide coordinates of image centers, elevation, time of ground truth, and ground truth site/cluster (one of ten) corresponding to the instance.

How many instances are there in total (of each type, if appropriate)? In the training set there are 406 instances of presences and 394 instances of absences. In the test set there are 51 instances of presences and 49 instances of absences. We reserve an additional 100 instances for assessment of research progress at a future date.

Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? The dataset is a sample of instances drawn from a landscape that varies in background plant community composition. The sampling strategy was targeted to represent the breadth of plant community types of land management concern at MPG Ranch, a conservation property in western Montana, USA.

What data does each instance consist of? Each instances is a labelled image with metadata described above.

Is there a label or target associated with each instance? The binary label is leafy spurge present or absent.

Is any information missing from individual instances? We provide a label only. More granular information, such as segmentation masks, are not available.

Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)? Geographic location of instances we report in the metadata in both geographic (EPSG Code 4326) and projected (EPSG Code 32611) coordinate systems. The density of instances varies across the landscape, but each are minimally separated by 1 meter. Additionally, the cluster feature in the metadata represents a distinct sampling campaign. E.G., a 1-hour search where the composition of background plants was similar across instances.

Are there recommended data splits (e.g., training, development/validation, testing)? Yes, we provide training (800 instances) and test splits (100 instances). The test instances are geographically isolated from the training instances. The intent of splitting in this way was to simulate classifier performance in recently invaded areas without prior weed occupation.

Are there any errors, sources of noise, or redundancies in the dataset? Both the onboard GPS of our survey drone and the GPS used to ground truth plants in the field were susceptible to error, where the positional error of the ground truth is not more than 1 cm and the positional error of the drone is as much as 7.32 cm. The consequence of this is that a target plant may be offset with respect to the image center. Some of the images may be blurred as a byproduct of the orthomosaic process, which attempts to stitch together many images in a drone survey, and may suffer in quality when insufficient features are matched across images.

Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)? The dataset is self-contained.

Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? No.

Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? No.

Does the dataset relate to people? No.

Does the dataset identify any subpopulations (e.g., by age, gender)? No.

Is it possible to identify individuals (i.e., one or more natural persons), either directly or indirectly (i.e., in combination with other data) from the dataset? No.

Does the dataset contain data that might be considered sensitive in any way (e.g., data that reveals racial or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? No.

Any other comments? None.

12.3 Collection process

How was the data associated with each instance acquired? The data were accquired by trained botanists and drone operators. First, the drone operator surveyed an 150 hectare area known to harbor leafy spurge. During this survey the drone gathered many overlapping images. After the survey these images were stitched together into a single large image, referred to as an orthomosaic, using Drone Deploy software. In this process, each pixel in the image was georeferenced using information about drone GPS position, camera angle, and lens physical properties. Shortly after, a botanist visited ten sites within the survey and gathered coordinates of spurge presence and absence. Afterwards we cropped square image tiles from the larger orthomosaic using these coordinates.

What mechanisms or procedures were used to collect the data (e.g., hardware apparatus or sensor, manual human curation, software program, software API)? The drone used to survey the area was a DJI Mavic 3M equipped with a real-time kinematic GPS system. The GPS used to gather coordinates of leafy spurge presence and absence was an Emlid RS2 rover unit.

If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)? The botanist conducted a random walk in areas differing in background vegetation species.

Who was involved in the data collection process (e.g., students, crowdworkers, contractors) and how were they compensated (e.g., how much were crowdworkers paid)? Staff scientist of MPG Ranch and one student from Chicago Labs Schools gathered the field data and processed it. Staff were compensated with salaries while the student received technical training in exchange for their services.

Over what timeframe was the data collected? The data were gathered over the course of a week in June 2023.

Were any ethical review processes conducted (e.g., by an institutional review board)? Not applicable.

Does the dataset relate to people? No.

Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (e.g., websites)? We gathered the data directly from the landscape.

Were the individuals in question notified about the data collection? Not applicable.

Did the individuals in question consent to the collection and use of their data? Not applicable.

If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? Not applicable.

Has an analysis of the potential impact of the dataset and its use on data subjects (e.g., a data protection impact analysis) been conducted? Not applicable.

Any other comments? None.

12.4 Preprocessing/cleaning/labeling

Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? We assembled the orthomosaic and geoferenced it using Drone Deploy's services. We provide crops of two sizes, 39x39 and 1024x1024 pixels (crop and context configurations on `https://www.huggingface.co`, at the locations of presence and absence ground truth gathered in the field by botanists.

Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (e.g., to support unanticipated future uses)? We also serve the full orthomosaic (unlabelled configuration on `https://www.huggingface.co`).

Is the software used to preprocess/clean/label the instances available? Drone Deploy's paid services are available to the public, though the underlying process for orthomosaic generation is closed-source.

Any other comments? None.

12.5 Uses

Has the dataset been used for any tasks already? We have established benchmark classifier performance with ResNet50 and DINOv2 architectures. We have conducted a brief dialogue with GPT-4o to explore its knowledge of drone imagery, ecology, and the target weed.

Is there a repository that links to any or all papers or systems that use the dataset? Yes, please find our website at `https://leafy-spurge-dataset.github.io`.

What (other) tasks could the dataset be used for? The full orhtomosaic might also be studied for unsupervised learning techniques. Multi-modal experiments, pairing elevation metadata with imagery, could also prove fruitful.

Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? No.

Are there tasks for which the dataset should not be used? We invite all uses.

Any other comments? None.

12.6 Distribution

Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created? We host the dataset freely on `https://www.huggingface.co`.

How will the dataset will be distributed (e.g., tarball on website, API, GitHub)? We host the dataset freely on `https://huggingface.co/datasets/mpg-ranch/leafy_spurge`.

When will the dataset be distributed? The dataset is currently available to the public.

Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? We authorize use of our data with a Creative Commons Attribution 4.0 International license.

Have any third parties imposed IP-based or other restrictions on the data associated with the instances? No.

Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? No.

Any other comments? None.

12.7 Maintenance

Who is supporting/hosting/maintaining the dataset? MPG Ranch staff are providing all support for this dataset.

How can the owner/curator/manager of the dataset be contacted (e.g., email address)? Email Kyle Doherty, the curator at kdoherty@mpgranch.com

Is there an erratum? Not at this time.

Will the dataset be updated (e.g., to correct labeling errors, add new instances, delete instances)? We will release a second test split at a later date to evaluate progress in task performance. All changes will be tracked via Hugging Face's version control systems.

If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (e.g., were individuals in question told that their data would be retained for a fixed period of time and then deleted)? Not applicable.

Will older versions of the dataset continue to be supported/hosted/maintained? These will be accessible through prior commits on Hugging Face.

If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? Please contact Kyle Doherty at kdoherty@mpgranch.com for extensions of this dataset.

Any other comments? None.

13 Appendix C - Drone and Sensor Specifications

The Mavic 3M is equipped with Real-Time Kinematic (RTK) positioning to enhance the GPS accuracy and ensure centimeter-level precision in the camera position. The drone features a 4/3 CMOS image sensor, with a resolution of 20 Megapixels (MP) and operated within an RGB color space. The lens provided a field of view (FOV) of 84° and an equivalent focal length of 24 mm. The ISO range was set between 100 and 6400, with a median shutter speed of 1/640 s. Additional specifications include an actual focal length of 12.29 mm, an aperture of f/2.8, and a minimum exposure time of 1/2,000 s. Images, with dimensions of 5280 X 3956 pixels, varied in size from 9.4 MB to 12.6 MB.