Data-Centric Mapping for Autonomous Broadacre Farming

Mojtaba Hedayatpour[1], Shing-Yan Loo, Parash Rahman, Sean Scheideman, Ganesh Tata

***Abstract*— Mojow is using data-centric AI to automate tractors and other farm vehicles for broad-acre agriculture. Our main product, EYEBOX, maps the immediate environment to locally optimize the path and to update the larger global map of the entire farm. In accurately mapping the field, we can confidently perform all vehicle-based operations on the farm including autonomous coverage over the entire field, autonomous mapping the field boundaries, autonomous mapping of farm yards, and various other tasks. In order to accomplish a detailed map we require excellent 360 degree perception, which is why the EYEBOX system is equipped with cameras, LiDAR, RADAR and GNSS. We train machine learning models with data to filter our sensors for important information. Our success in using this data-centric approach has prompted us to create a wide variety of datasets for the many vehicle-based operations that constitute the majority of efforts in farming. As our datasets expand in diversity, our machine-learning models become more robust to previously unseen situations.**

I. INTRODUCTION

With our product, the EYEBOX (Figure 1), we aim to significantly decrease the manpower involved in vehicle-based operations. Broad-acre farmers use large machines to traverse enormous distances to prepare, grow, maintain, and harvest their crops. They require many man-hours to accomplish their goals, and it has only gotten more difficult to fulfill the demand for skilled farmhands. EYEBOX is a retrofit kit for tractors and other farm vehicles that enables autonomous mapping of the region immediately around the vehicle, *the local map*. This accurate local map can then be in turn used to update a large map of the entire farm, known as *the global map*. The global map can be used to facilitate long distance path planning so the farmer can send tractors long distance autonomously. Most importantly, the global map facilitates path planning over an entire field so every part of the field can be covered by specialized vehicles.

Fig. 1. An image of an example EYEBOX (left) with its 3 cameras and LiDAR sensor that can be mounted on a tractor. The right image shows a tractor with two EYEBOXs on its roof and a GNSS on its hood.

[1] Mojow Inc., Edmonton, AB, Canada `mojtaba@mojow.ai`

Autonomous navigation around a farm requires a high level of environmental awareness. This is a difficult challenge as farms vary greatly in their appearance from farm to farm and throughout the year - unexpected falling trees, uncovered rocks, and wet patches to get stuck in are common. Finally, one can find many objects to avoid running into on farms, such as unharvested crops, crop rows, hay bales, rock piles, people, animals, sloughs, dilapidated equipment, etc. In addition, hilly areas and rough areas should be detected and mapped to plan out path plans for those regions on a global and local scale.

To autonomously create our local and global maps, we use the crucial information from the EYEBOX sensors: cameras, LiDAR, RADAR, and GNSS. Vision, although not the only component, is key in farming as there are a lot of challenges when using GPS-only systems, as is traditionally done for autosteering systems, and there are challenges with LiDAR-only systems. GPS-only systems lack real-time updates of close-by obstacles. We and others [3] have found LiDAR-only systems have great difficulty in the dusty environments ever-present in some fields and with the long grass easily mistakable for an elevated surface. Vision models can provide key information about the surrounding region, even going as far as replacing or working with LiDAR for depth estimation [4]. GNSS is required for updating the global map with the local maps generated by the other sensors.

We extract key information about the surrounding environment for mapping with machine learning models. These models are key to ensuring the maps are accurate locally, and therefore globally, too. To train and improve the machine learning models, Mojow continuously collects and annotates datasets. We have found that many challenges in the diversity of farms can be overcome by collecting and training machine learning models with the appropriate data. However, there is a distinct lack of ground-level perception datasets in farming for the many vehicle-based applications we target. There are a few datasets related to farming, but they are far from capturing the diverse perspectives and situations we need to have reliable local awareness in the fields, roads, and farm yards. Often, the existing datasets are related to aerial views of the farm for aerial analysis, which is the more common application. Therefore, throughout the year, we have collected diverse and interesting datasets to aid in mapping for our target applications, focusing on data where our models struggle.

II. DATASETS FOR MAPPING

Our perception system relies heavily on our high-quality, well-maintained datasets for training accurate models for local and global mapping. Accurate local maps require time synchronized sensors, precisely annotated data, and the extrinsics information available between the sensors. The extrinsics between the sensors and the GNSS is then required for the global mapping of the entire farm area.

We use various annotation types for our datasets. The annotations can be pixel-wise semantic segmentation image annotations (Figure 2) where each pixel in the image is allocated an important class, such as field, road, power pole, sky, person etc. The annotations can be bounding boxes if noting a class's nearby presence is more important than the shape of the obstacle (Figure 2). We also have LiDAR point clouds annotated with pointwise labels and 3D bounding boxes for the parts of the map that require them.

It is also worth noting that some annotations are time-expensive to complete. Therefore, to better prioritize our annotations, we add text labels to our data with the interesting aspects to easier prioritize difficult or interesting scenarios. Some examples of our text labels include the following: *clear skies*, *person on the road*, *tractor hood in view*, *glare*, and *overexposed lighting*.

Below are some examples of applications that use our mapping and the kind of datasets required for them.

1) Implement folding/unfolding: Tractors are multipurpose vehicles. They are used to tow *implements* behind them. Implements are large tools that help perform farm operations like tilling, weeding, rolling, and seeding. These tools are often so large that special driving maneuvers are required to fold and unfold them. We train our machine learning model to detect the pose of the implement to map out the implements and perform autonomous implement folding and unfolding. In addition, we map out the implement during path following to prevent collisions and to plan out paths around obstacles.

2) Field boundary detection: Mapping out a field's boundary is key to plan autonomous coverage paths in broad-acre farming. Traditionally, farmers have to drive carefully around the field boundary manually to map their field boundaries. Our field boundary models detect the field boundary from camera images with pixel-wise accuracy (Figure 2 so the field boundary can be autonomously mapped with precise GPS coordinates. Our models are trained with field boundary datasets that contain various field boundaries, including wooden poles, ditches, thin strips of green grass, fences, fence poles, roads, and long-dried grass. We have greatly improved our models' generalization as we add more data and expect stronger performance as our datasets continue to increase.

3) Unharvested crop detection: During harvest, there are multiple vehicles that are expected to traverse around unharvested crops. To ensure these unharvested crops are not damaged, we train machine learning models to map their boundary coordinates and avoid them (Figure 2.

4) Road boundary and farm yard navigation: Rural roads and farm yards are driven through to move vehicles from one field to another, to maintain/refuel the vehicle, or to store the vehicles. Our current road models are trained on a diverse set of roads and farm yards from across Canada to handle the detection of roads in an image and map their boundary coordinates locally. Note that there are more publicly available datasets about off-road driving, including rural roads [1], [5], [6], [8]. However, it is important to continually increase the diversity and magnitude of our own datasets to handle all situationsm, and it is crucial to have GNSS data to update the global map.

5) Rock detection: Rocks are a great nuisance for farmers as they can damage farm equipment. Our rock detection datasets for rocks of different sizes can be used to train models to detect them and map them in the field for later pickup.

6) Person detection: As a final example, for safety reasons, we have a person detection dataset (Figure 2). This dataset can be complemented by other publicly available person detection datasets on the farm. One of its main purposes is to test our safety systems, which stop operations whenever a person is too close.

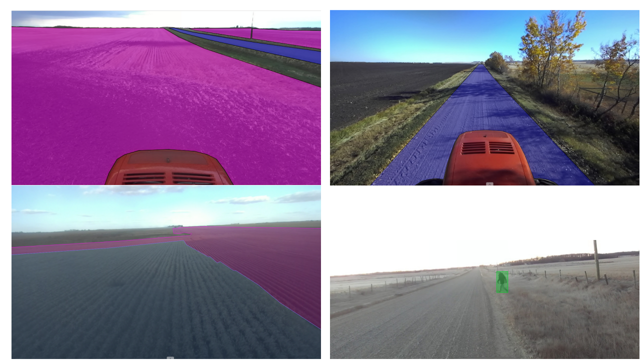

Fig. 2. Examples of annotations: fields (top left), road (top right), unharvested crop (bottom left), and person in view (bottom right).

References

[1] P. Jiang, P. Osteen, M. Wigness, and S. Saripalli. "Rellis-3d dataset: Data, benchmarks and analysis." arXiv preprint:2011.12954, 2020.

[2] T. Guan, D. Kothandaraman, R. Chandra, A. J. Sathyamoorthy, K. Weerakoon, and D. Manocha. "GA-Nav: Efficient Terrain Segmentation for Robot Navigation in Unstructured Outdoor Environments." IEEE Robotics and Automation Letters Volume 7(3), pp. 8138-8145, 2022.

[3] D. Maturana, P-W. Chou, M. Uenoyama and S. Scherer. "Real-time Semantic Mapping for Autonomous Off-Road Navigation". Proceedings of 11th International Conference on Field and Service Robotics (FSR '17), pp. 335-350, 2017.

[4] A. Masoumian, H. A. Rashwan, J. Cristiano, M. S. Asif, D. Puig. "Monocular Depth Estimation Using Deep Learning: A Review." Sensors 22, pg. 5353, 2022.

[5] P. Mortimer1, R. Hagmanns, M. Granero, T. Luettel, J. Petereit, and H-J. Wuensche1. "The GOOSE Dataset for Perception in Unstructured Environments." IEEE International Conference on Robotics and Automation (ICRA) 2024.

[6] P. Neigel, J. R. Rambach, and D. Stricker. "OFFSED: Off-Road Semantic Segmentation Dataset." Proceedings of the 16th Int. Conference on Computer Vision Theory and Applications Volume 4: VISAPP, pp. 552-557, 2021.

[7] M. Sivaprakasam, P. Maheshwari, M. Guaman Castro, S. Triest, M. Nye, S. Willits, A. Saba, W. Wang, and S. A. Scherer, TartanDrive 2.0: More Modalities and Better Infrastructure to Further Self-Supervised Learning Research in Off-Road Driving Tasks. CoRR abs/2402.01913, 2024.

[8] M. Wigness, S. Eum, J. G. Rogers, D. Han, and H. Kwon, A RUGD Dataset for Autonomous Navigation and Visual Perception in Unstructured Outdoor Environments. International Conference on Intelligent Robots and Systems (IROS) 2019.