



HAL
open science

An anomaly detection approach to monitor the structure-based navigation in agricultural robotics

Hassan Nehme, Clément Aubry, Romain Rossi, Rémi Boutteau

► **To cite this version:**

Hassan Nehme, Clément Aubry, Romain Rossi, Rémi Boutteau. An anomaly detection approach to monitor the structure-based navigation in agricultural robotics. IEEE International Conference on Automation Science and Engineering (CASE), Aug 2021, Lyon, France. 10.1109/CASE49439.2021.9551568 . hal-03255483

HAL Id: hal-03255483

<https://hal.science/hal-03255483v1>

Submitted on 11 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Anomaly Detection Approach to Monitor the Structured-Based Navigation in Agricultural Robotics

Hassan Nehme^{1,2}, Clément Aubry¹, Romain Rossi² and Rémi Boutteau³

Abstract—Local perception navigation methods allow agricultural robots to accurately track crop row structures while performing automated farming tasks. The integration of these methods as a part of a fully autonomous navigation solution requires continuous assessment of their reliability since they rely solely on sensor data in a changing and unpredictable environment. This paper presents a data-driven monitoring approach for the task of structure-based navigation in agriculture. The proposed method applies semi-supervised anomaly detection, aiming to learn a model of *normal* scene geometry that characterizes a domain of reliable execution of the considered task. To this end, a convolutional neural network was trained in one-class classification fashion on Hough representations of LiDAR point clouds. In experimentation, the learned *normal* model was used to derive a confidence measure for a LiDAR-based tracking algorithm allowing its integration as a part of a hybrid navigation solution in vineyards for a commercial robotic platform.

I. INTRODUCTION

Autonomous robots are essential for improving farming productivity, with autonomous navigation being a key feature in their design for precision agriculture applications. Automated farming tasks are generally performed while accurately tracking parallel crop rows. A current research interest is to develop hybrid navigation methods [1] that integrate local perception systems (e.g., LiDAR, vision) along with Global Navigation Satellite Systems (GNSS). In this setup, local perception exploits geometric patterns and provides accurate tracking with respect to the perceived structure in the field e.g. crop rows. Therefore, a hybrid navigation method reduces the accuracy requirements of global navigation systems that will manage global aspects of the task such as transitioning from one row to another, while leaving accurate structure tracking for local perception.

Local perception navigation methods are interesting since they eliminate the effort of creating a predefined path while providing local awareness and responsiveness with respect to the actual structure. However, this increased flexibility means that the structure-based navigation function will rely solely on sensor data acquired in an unpredictable environment under uncontrolled conditions. This raises concerns about the viability of local perception methods as a whole solution for near production robots since a perception algorithm is as reliable as its input data. A way to respond to these

concerns is by providing an assessment of input data quality through data-driven monitoring. Such an assessment avoids failures due to the application of a perception algorithm in anomalous situations: ranging from *low level* anomalies such as corrupted data due to poor acquisition conditions (e.g., low light, shadows, a leaf covering the sensor) to *high level* anomalies caused by the perceived scene not complying with the base assumptions of the tracking algorithm. Moreover, a quantitative assessment providing a confidence measure defines guidelines for integrating the perception algorithm as a part of a complex or hybrid solution with deliberation and planning functions.

This work proposes a monitoring method for the general local structure-based navigation task in agriculture. As a main contribution, we show that the Hough transform representation of LiDAR point clouds can be exploited as an abstraction of the field’s geometry that allows a convolutional neural network model to capture a normal¹ criterion about the traversability of the structure. The monitoring is done through anomaly detection on the Hough representation of input data. A fully convolutional one-class classifier is trained on data acquired during normal execution of the task. The goal of the classifier is to provide an anomaly score reflecting the conformity of a test observation with a learned model of normal execution. The anomaly score can be treated as a confidence measure to be used to decide whether to rely on the execution of the monitored task or to switch to another behavior and report an error.

The proposed method was illustrated in experimentation where monitoring the task of structure-based navigation of an agricultural robot in vineyards was considered. The monitoring provided a fusion rule defining the integration guideline of a LiDAR-based structure tracking method as a part of a hybrid navigation solution. The derived fusion rule defined the domain of application of the LiDAR-based method based on pure assessment of input data quality. Moreover, the fusion rule provided a smooth transition to GNSS-based U-turns near the end of each vine row where less information about the structure is available for the perception algorithm impacting the quality of its tracking signals.

The remainder of the paper is organized as follows: section II provides preliminaries on the structure tracking task and reviews related anomaly detection methods, section III states the problem and explains the steps of the proposed method, an evaluation of the proposed method in an experimental case study is provided in section IV and finally section V

¹ SITIA, 7 rue de l’halbrane, 44340 Bouguenais, France f.last@sitia.fr

² Normandie Univ, UNIROUEN, ESIGELEC, IRSEEM, 76000 Rouen, France first.last@esigelec.fr

³ Normandie Univ, UNIROUEN, UNILEHAVRE, INSA Rouen, LITIS, 76000 Rouen, France. first.last@univ-rouen.fr

¹Throughout the paper, normal means not anomalous.

concludes the paper.

II. PRELIMINARIES AND RELATED WORK

This work proposes a monitoring approach for the general structure-based navigation task in agriculture. We are interested in the viability of methods and how they fit in the overall solution rather than the specific technique implemented. The questions to be asked are: how reliable is the perception-based algorithm and how well does it fit in a complex solution ready for commercialization (e.g. integration in a hybrid navigation solution). Not enough attention is given to these issues in the literature since the main concern is usually to develop a new tracking technique or algorithm. Yet, they are important when deciding which technique to deploy in a production robot. To this end, integrating a structure tracking algorithm as a navigation solution in a system requires defining a confidence measure whether it is implemented as a sole solution or as part of a hybrid solution (based on sensor fusion). Prior attempts to define the confidence measure used algorithm-specific [2] or sensor-specific information [3]. In this work, we follow a data-driven monitoring paradigm [4], that is performing anomaly detection on a set of data without the use of explicit details or knowledge about the navigation algorithm.

This section provides context about the task to be monitored, i.e. structure-based navigation (which is an input not an element of the solution) and the considered paradigm considered for the monitoring, i.e. anomaly detection. First, we review structure tracking using Hough transform since it is considered as the data representation for the proposed monitoring approach. Then, we review the anomaly detection problem in the context of mobile robot navigation with a focus on one-classification methods.

a) Structure-based navigation and Hough transform: Local structure-based navigation is the task of tracking crop rows using local perception sensors (LiDAR scanners and cameras). The core of the solutions proposed in the literature is line fitting techniques. Hough transform [5] is one of the most common line fitting techniques employed for the task of identifying crop rows; it was applied to identify crop rows from vision data in [6], [7] and from LiDAR data in [8], [9], [10]. Hough transform maps input data from their feature space (image space or 3d metric space for LiDAR scans) to a parameter space called Hough space. Parametric models, e.g. a line, can be identified in the parameter space through a voting procedure. For the task of crop rows detection, the goal is to search for a pattern of parallel lines in the Hough space. To facilitate the search, additional hypotheses about the field have been made such as row spacing [11] or more generally exploiting row parallelism criterion [7], [12].

The above mentioned hypotheses were made after observing *invariants* in the Hough representation of input data. Therefore, the presence of these *invariants* is conditioned to the presence of a parallel pattern in the perceived scene which is the base assumption of any structure tracking algorithm. The work presented in this paper is a generalization of this reflection: we assume that these *invariants* capture

a normal criterion about the perceived structure that can be modeled through unsupervised representation learning without the need of explicit hypotheses.

b) Anomaly detection: Anomaly detection is the task of identifying samples in data that do not conform to a well-defined notion of normal behavior [13]. Depending on the task, the notion of normal behavior can be defined by gathering anomaly-free data. In this context, an anomaly detection algorithm consists of three major steps: first, modeling an anomaly-free background model from a representative dataset. Then, defining an anomaly score as a distance or a measure of deviation from the background model. Finally, defining an appropriate decision rule based on the anomaly score for the final binary decision (anomaly or not). This formalization was synthesized in a recent survey on image anomalies [14] where anomaly detection methods were classified based on the assumptions made in each of the three aforementioned steps. An illustration of a generalized anomaly detection pipeline is provided in Figure 1.

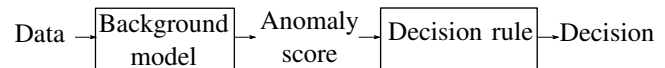


Fig. 1: The anomaly detection pipeline considered in this work.

For applications in mobile robot navigation, a semi-supervised anomaly detection approach is often considered. In such an approach, the background model is learned from a set of normal data only. This makes semi-supervised anomaly detection a practical solution since it avoids gathering data representing all the anomalies that can occur which can be a dangerous and costly process. This approach was applied for traversability assessment [15] and rare obstacle detection in outdoor environments [16] where normal data were fairly accessible and were gathered from a safe execution of the navigation task.

Semi-supervised anomaly detection is regarded as one-class classification. One-class Support-Vector Machine (OC-SVM) [17] is a classical method that defines the background model as a domain in feature space where normal data lie. OC-SVM does not scale well for high dimensional data (such as images) and large training sets. [18] introduced a deep one-class classification method called Deep Support Vector Data Description (Deep SVDD). Deep SVDD is an adaptation of the domain-based training criterion to deep learning that jointly learns to extract features that are enclosed in a hypersphere of minimal volume in feature space.

This work adopts a semi-supervised anomaly detection approach with deep SVDD to monitor the normal execution of the task of structure-based navigation in agriculture. Hough transform is considered as the representation of input data. The assumption is that a background model learned from Hough space captures specific features defining a normal criterion of the structure.

III. MATERIALS AND METHODS

The monitoring method presented in this paper is developed as a part of the navigation solution for *Trektor*², a multi-purpose agricultural robot developed by SITIA. The robot, shown in two vineyard configurations in Figure 2, is designed to intervene on different crops with their variety of features; ranging from tiny plants (e.g. market gardening) to dense small trees (e.g. viticulture). A parallel structure is common for all of these fields whether it is provided by crop rows or man-made structures (e.g. wheel tracks). Thus, it is worthy to study the integration of a structure tracking algorithm in the overall navigation solution. This section explains the motivation and development of the anomaly detection method to monitor an existing structure tracking algorithm.



(a) Narrow vineyard.

(b) Wide vineyard.

Fig. 2: SITIA's *Trektor*: the robotic platform used in the experimentation.

A. Navigation framework

One of the main applications of our robotic platform is autonomous navigation in vineyards. The task requires global planning in order to cover the entire field: i.e. a global GNSS path defining the rows that need to be worked and in what direction and order. Hence, the robot is equipped with an accurate global navigation system based on the fusion of RTK-GNSS and high accuracy inertial navigation systems. Leveraging this accuracy in practice requires the global path to be in high resolution which requires a significant effort in the task preparation phase. Hence comes a compromise with local perception that significantly reduces the resolution requirements of the global path by providing accurate tracking with respect to the perceived crop rows. In this setting, only the start and end points of the desired rows need to be georeferenced in order to create the global path that handles the transition between rows while leaving accurate row tracking for the local perception algorithm.

Figure 4 provides an illustration of the navigation task in a vineyard. The local perception for this task is provided by a 16-channel 3D LiDAR sensor (Ouster OS1-16) mounted on top of the robot, as shown in Figure 3, which allows the perception of multiple rows. As the robot approaches the end of a row, a less portion of the rows is perceived by the LiDAR sensor making its tracking less reliable; until a location where no structure information is available (noted as a red cross in Figure 4), here the transition to global navigation should be

triggered in order to handle the U-turn and enter the next row. This is one of the main scenarios that motivate a data-driven monitoring of the structure tracking algorithm. By doing so, a perception-based algorithm becomes more viable by providing an indicator of its reliability.



Fig. 3: LiDAR-sensor's position on top of the robot.

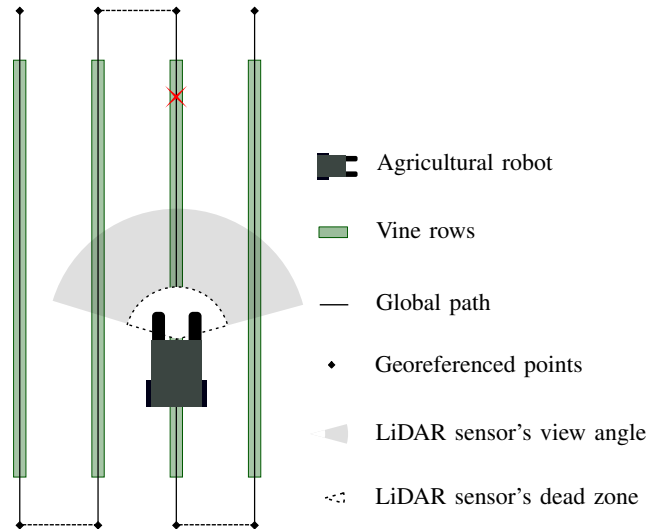


Fig. 4: Illustration of the navigation task as a sequence of LiDAR-based structure tracking and GNSS-based U-turns. Less structure information is perceived by the LiDAR sensor at row ends where the transition to global navigation should be triggered (near the red cross).

B. Monitoring via anomaly detection

The goal is to provide an indicator of the reliability of a structure tracking algorithm based solely on its input. This can be achieved following an anomaly detection paradigm (Figure 1) with three major elements: background model, anomaly score and decision rule. In this setting, the learned background model will provide an anomaly score reflecting a deviation of the inputs from its domain of reliable execution. Learning a background model is followed by an empirical evaluation of the produced anomaly score's statistics on a validation dataset with the goal of deriving the final decision rule.

As it has been stated in section II, the background model is conditioned to a notion of normal behavior that is defined by a set of data. Therefore, it is essential to carefully consider

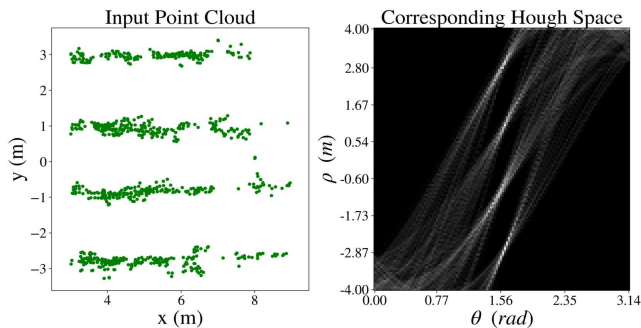
²www.sitia.fr/en/innovation-2/trektor/

beforehand an assumption about the definition of the desired normal behavior and how to represent it with a dataset since the background model will learn patterns that illustrate this assumption. The following subsections describe the different elements of the proposed solution.

1) *Data representation*: The task to be monitored is the structure tracking algorithm. This algorithm applies pattern analysis on the Hough representation of point clouds in order to identify a pattern of parallel lines. Hough transform maps input points from their 2D metric space to the parameter space of the model to be identified; e.g. a line in the considered application. In details, Hough transform defines a line using the polar parameters (ρ, θ) where $\theta \in [0, \pi]$ is the normal angle and $\rho \in \mathbb{R}$ is the signed distance from the origin [5]. A point (x_0, y_0) in the input point cloud maps to the sinusoid:

$$\rho = x_0 \cos\theta + y_0 \sin\theta \quad (1)$$

in the polar parameter space that defines all the possible lines that pass through it. Thus, each point *votes* for the parameter of candidate lines. After considering all points, the Hough representation of the input cloud is a heatmap in the 2D parameter space where the coordinates of high heat define the most likely model candidates.



(a) Sample input point cloud. (b) Corresponding Hough space.

Fig. 5: A sample input point cloud and the corresponding Hough space.

Figure 5 shows a sample point cloud acquisition of a vineyard canopy along with the corresponding Hough transform. Four dense sinusoid intersections can be noticed in the Hough space. These intersections define the parameters of the four vine rows in the input cloud. This observation has been studied and it is exploited by most Hough-based crop rows detection methods [7], [10].

While the raw inputs of the structure tracking algorithm are point clouds, further preprocessing followed by the application of Hough transform are performed prior to pattern analysis. Hence, the Hough representation of an input point cloud is the direct input to pattern analysis. Since most crop rows detection algorithms exploit different properties in the Hough space in order to detect a pattern of parallel rows, we consider that Hough transform abstracts defining geometric features of crop row structures, and thus, it is considered as the representation of input data to the proposed anomaly

detection method. This consideration is intended to reduce the capacity requirements of the subsequent background model by exploiting an already-computed feature extraction. This results in a light monitoring module that is compatible with a wide range of Hough-based tracking algorithms.

2) *Defining an anomaly - Data collection*: The aim is to learn a model of typical scene structures that characterizes a known and reliable structure tracking. With Hough transform as input representation, we collect structure data samples from previous safe navigation where we ensured that the field is suitable for structure tracking; meaning that no obstacles were present and the overall structure was visually consistent. We collected data from two different vineyards, each of which has its unique geometric characteristics (e.g., row spacing, height, weed pressure). The considered fields cover two vineyard configurations: narrow vineyard where the robot is centred on a vine row (Figure 2a) and wide vineyard where the robot passes between two vine rows (Figure 2b). This ensures that the collected data samples are diverse and cover the most likely scenarios that the robot will face.

The desired normal behavior is when the perceived scene allows the structure algorithm to provide pertinent results. Therefore, we define the normal behavior with samples collected in the in-row portion of the field (as opposed to the U-turn portion where no structure information is available). We used location information along with the global path (Figure 4) to annotate the collected data. The normal data zones were defined by shifting the geo-referenced end-points by 5 meters; this ensured that normal data do not include samples from the transition between structure tracking and U-turns as these samples contain little to no structure information.

Data samples are collected as 128×128 Hough images. Figure 6 shows examples of normal and anomalous data samples. The normal sample was collected in the in-row portion of the field and it shows clearly the sinusoid intersections pattern discussed in subsection III-B.1. Whereas the anomalous sample collected during a U-turn does not show this pattern, and, consequently, it does not contain structure information. The collected dataset is split into a training set that contains 14K normal samples, and a validation set that contains equally 4K of normal and anomalous samples. The aim of the balanced validation set is to monitor the evolution of the learned model using binary classification metrics such as the area under the receiver operating characteristic curve (ROC-AUC). The validation set will also be used to derive statistics about the learned model that are useful to define a decision rule in the consequent step of the anomaly detection pipeline.

3) *Background model*: After collecting training data, the next step in the anomaly detection pipeline (Figure 1) is learning a background model that captures the notion of normal behavior. We employ the background model as a one-class classifier to be learned from the set of normal Hough images. The considered classifier follows the approach proposed by Ruff et al. [18]. In this approach, the one-class classifier is considered to be a neural network $\phi(\cdot; \omega)$ with

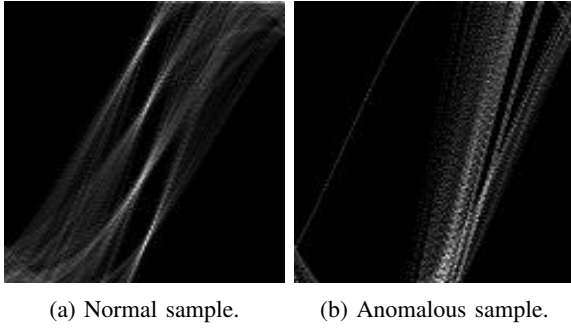


Fig. 6: Hough space for normal and anomalous samples from the collected dataset.

weights ω that transforms input images to feature vectors. The model is trained to produce feature vectors that are enclosed in a hypersphere of minimal volume in the feature space. The volume of the hypersphere is jointly minimized along with learning the neural transformation weights ω . We adopt the One-Class Deep SVDD training loss proposed in the original paper [18]:

$$L = \|\phi(X; \omega) - \mathbf{c}\|^2 + \lambda \|\omega\|^2 \quad (2)$$

where X is an input image, \mathbf{c} is the hypersphere center and ω represents the network weights. The second term in the training loss is a regularizer on network weights which is controlled by a hyperparameter λ . This training criterion forces the learned feature vectors of normal samples to be enclosed in a hypersphere of minimal volume. The hypersphere center \mathbf{c} is set arbitrarily from an initial forward pass of the untrained network and it is fixed during training. This initialization technique follows the recommendation of Ruff et al. [18] to choose an arbitrarily center $\mathbf{c} \neq \mathbf{0}$ in order to avoid trivial solutions (in which, the network weights collapse to zero).

The proposed method is designed to monitor a real-time navigation application. Moreover, the representation of input clouds as Hough images reduces the capacity requirements of the background model $\phi(\cdot; \omega)$. These reasons motivate the choice of a light convolutional network. To this end, we propose a light fully convolutional network architecture with seven downsampling convolutional blocks. The proposed architecture is illustrated in Figure 7; it produces a $1 \times 1 \times 128$ feature vector as an output for a $128 \times 128 \times 1$ input image. Each downsampling block consists of a convolutional layer followed by batch normalization and leaky ReLU activation function (with leakiness $\alpha = 0.2$). All the layers are defined with no bias term since it was shown in [18] that networks with bias term collapse to a constant function mapping when trained to minimize the deep SVDD loss.

4) *Anomaly score:* As an anomaly score for the model, we use the distance $\|\phi_\omega(X_{\text{test}}) - \mathbf{c}\|^2$ between the transformed features of a test sample X_{test} and the hypersphere center \mathbf{c} of the learned model. Since the deep SVDD training criterion forces the learned feature vectors of normal samples to be enclosed in a hypersphere, anomalous samples will be

mapped to a location in the feature space that is at a larger distance from the hypersphere center than the mapping of normal samples.

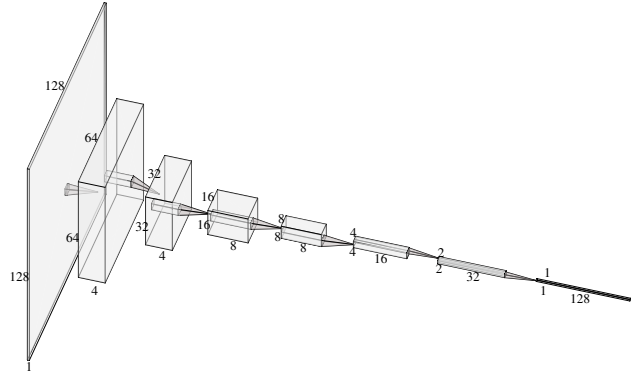


Fig. 7: The light fully convolutional network architecture adopted for the background model.

5) *Decision rule:* After training, the learned background model will produce a raw anomaly score whose scale is not easily interpreted. This motivates the derivation of a decision rule, the final part of the anomaly detection pipeline. The decision rule aims to interpret raw anomaly scores as meaningful scalars that meet the requirements of the considered application. Hence, the decision rule is application specific. In the following, we explain the derivation of a decision rule for the application of monitoring the LiDAR-based structure tracking algorithm; allowing its integration, along with GNSS-based navigation, in the hybrid navigation setup introduced in Figure 4.

In practice, each of the two navigation algorithms eventually guide the robot by delivering tracking signals to a control law as lateral deviation Δl and angular deviation $\Delta \gamma$. The anomaly score produced by the learned background model can be used to define a fusion rule of the tracking signals by attributing a confidence measure $\delta \in [0, 1]$ to the structure algorithm. Hence, the fusion rule for the lateral deviation is:

$$\Delta l = \delta \Delta l_{\text{LiDAR}} + (1 - \delta) \Delta l_{\text{GNSS}} \quad (3)$$

and the same applies for the angular deviation. A confidence measure of 1 means that the tracking relies solely on the LiDAR-based algorithm. This is the desired case when the robot is between crop rows; the opposite is true when the robot is in the overhead performing U-turn where the tracking should rely solely on the GNSS. Therefore, the confidence measure is an inverse of the notion of anomaly assumed by the proposed anomaly detection algorithm. To this end, the decision rule should define the transformation of raw anomaly scores into the desired confidence measure δ .

A simple, yet effective, decision rule consists of defining a suitable threshold th for the anomaly score. This can be achieved empirically from the ROC curve evaluation of the learned background model on the validation set. However, this rule results in hard transitions between the final binary decision for samples with an anomaly score

varying around the threshold, which is not suitable for the considered application. To overcome the hard transitioning issue, we propose a soft decision rule as a generalization of binary thresholding. The desired output is a normalized anomaly score $s \in [0, 1]$ where the value of s for samples with an anomaly score near the threshold th should be around 0.5. Figure 8 illustrates the definition of the soft decision rule that takes into account the standard-deviations σ_n and σ_a of the distribution³ of raw anomaly scores on normal and anomalous validation samples. The soft decision rule is a saturation function, the limits of saturation and the function slopes are defined empirically using the standard-deviations σ_n and σ_a . Specifically, the saturation limits are set to $th - 0.5\sigma_n$ and $th + 0.5\sigma_a$ respectively and the function slope is set to $\frac{1}{\sigma_n}$ for $s < th$ and to $\frac{1}{\sigma_a}$ for $s > th$.

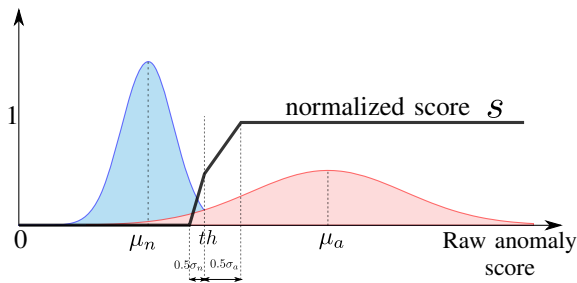


Fig. 8: Deriving an anomaly decision rule empirically. The rule is a saturation function that transforms a raw anomaly score into a meaningful decision metric $s \in [0, 1]$. Saturation limits and the function slopes depend on the standard-deviations σ_n and σ_a of raw anomaly scores on normal (blue distribution) and anomalous (red distribution) validation samples.

Finally, the confidence measure δ can be defined as $1 - s$ where $s \in [0, 1]$ is the normalized anomaly score produced by the soft decision rule.

IV. EXPERIMENTAL EVALUATION

As a case study, we consider evaluating the proposed monitoring approach in the hybrid navigation setup, introduced in Figure 4, that combines LiDAR-based structure tracking along with GNSS-based navigation. The goal is to deploy the proposed anomaly detection method to handle the transition between the two navigation algorithms based on the assessment of the Hough representation of input point clouds. To this end, the background model was trained on the collected dataset of normal Hough images and validated on the validation set using ROC-AUC binary classification metric.

The fully convolutional network (Section III-B.3) was trained from randomly initialized weights to minimize the deep SVDD training loss using the Adam optimizer with a learning rate of 10^{-5} . The training was implemented in Tensorflow and we trained for 100 epochs. Figure 9 shows

³ for illustration, gaussian distributions are assumed which is not necessarily the case in practice, Nevertheless, this assumption holds true since we only use standard-deviations from empirical evaluation

the results of the training as the evolution of the training loss and the evaluation of the ROC-AUC metric on the validation set per epoch. The ROC curve is obtained by evaluating the true positive and false positive rates for different thresholds on the anomaly score; an anomaly score above a given threshold means that the sample is classified as an anomaly (i.e. positive in the notation of binary classification). The network achieved a final AUC of 0.988 on the validation dataset, this value indicates that the validation samples are highly separable (between normal and anomalous) thanks to the anomaly score of the learned model.

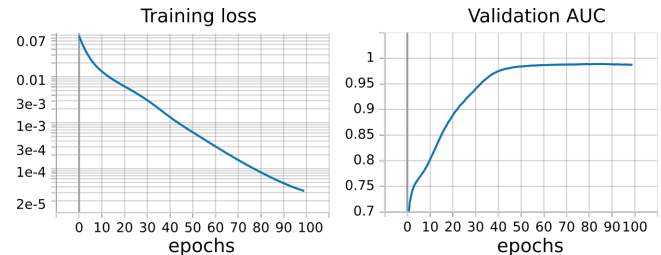


Fig. 9: The evolution of the SVDD loss on the training set and the ROC-AUC metric on the validation set during 100 training epochs.

Raw anomaly scores, on the validation set, of the learned background model were used to derive the decision rule as discussed in Section III-B.5. Finally, the complete anomaly detection method delivered the desired monitoring via the confidence measure δ that defines the fusion rule of Equation 3.

We tested this confidence measure during the navigation on a test vine row that was not covered by the training nor the validation datasets. The background model was deployed to provide estimations in real-time (the model ran at the same 10 Hz frequency as the LiDAR sensor). The localization module of the robot provided a reference for the evaluation as a flag indicating when the robot enters the row. Figure 10 shows the results of this evaluation as the evolution of the confidence measure, the in-row flag and the raw anomaly score during the navigation. Before entering the row, the LiDAR perceived a crop row structure that is relatively far ahead, this resulted in a zero confidence measure that gradually increased as the robot approached the row and perceived a more complete structure that complies with the learned background model. The confidence measure was then saturated at 1 a few seconds before the in-row flag, this is logical since the in-row flag is based on the position of the robot's center knowing that the LiDAR sensor is mounted at the front meaning that it will perceive a complete structure before entering the row. The same logic applies near the row-end where the confidence measure transitions before the in-row flag; since at this position, structure perception begins to be incomplete. The figure also shows the evolution of the raw anomaly score (scaled up for visualization) which is consistent with the previous explanation and shows that the learned anomaly detection method performs well in real experimental scenarios allowing reliable integration of a

structure-based navigation algorithm.

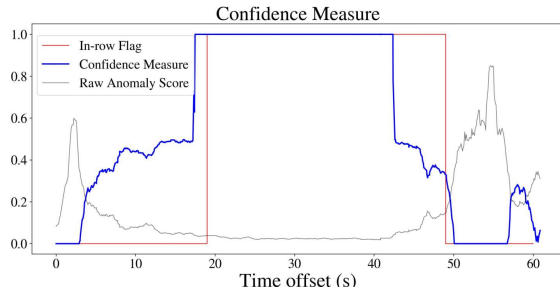


Fig. 10: The results of the evaluation in a hybrid navigation scenario as the evolution of the raw anomaly score (scaled up for visualization), the derived confidence measure of the LiDAR-based algorithm and the in-row flag provided by the GNSS system during the navigation on a vine row. The produced confidence measure saturates a few seconds before entering and exiting the row which correlates with positioning of the LiDAR sensor on the robot.

V. CONCLUSION

This paper presented a data-driven monitoring approach for the task of structure-based robot navigation in agricultural fields. The proposed method follows a semi-supervised anomaly detection paradigm in order to provide a quantitative assessment of the reliability of a LiDAR-based structure tracking algorithm. Hough transform was shown to be an effective abstraction of the field’s geometry, from which a light fully convolutional background model capturing a normal criterion about the structure was learned. The background model was trained on a dataset of normal Hough images collected during previous safe navigation in vineyard fields.

Field experimentations considered the case of hybrid robot navigation in vineyards. The learned background model was used to derive a fusion rule that handles the transition between the structure-based and GNSS-based algorithms. A decision rule was derived empirically by considering the statistics of anomaly scores produced by the background model on a validation dataset. The overall anomaly detection method was validated in a real navigation scenario where it was shown that it defines guidelines for the integration of a perception-based navigation algorithm as a viable part of a complete autonomous navigation solution.

Future work would consider studying the transfer of the proposed anomaly detection algorithm to different agricultural structures such as wheel tracks for market gardening. Another long-term goal is to consider modeling a more complex definition of anomalies that includes obstacles. This requires the investigation of an anomaly detection approach on different input modalities such as images, which motivates background modeling with generative adversarial networks.

ACKNOWLEDGMENT

This work is partially supported by the *Agence Nationale de la Recherche* ANR through the ROSE challenge, ANR-17-ROSE-0002-01 and by the *Association Nationale de la*

Recherche et de la Technologie ANRT through the CIFRE PhD, ANRT CIFRE N° 2018/0792. This work was performed in part on computing resources provided by CRIANN (Centre Régional Informatique et d’Applications Numériques de Normandie, Normandy, France)

REFERENCES

- [1] S. Kanagasingham, M. Ekpanyapong, and R. Chaihan, “Integrating machine vision-based row guidance with GPS and compass-based routing to achieve autonomous navigation for a rice field weeding robot,” *Precision Agriculture*, vol. 21, no. 4, pp. 831–855, Aug. 2020.
- [2] A. English, P. Ross, D. Ball, B. Upcroft, and P. Corke, “Learning crop models for vision-based guidance of agricultural robots,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Hamburg, Germany: IEEE, Sept. 2015, pp. 1158–1163.
- [3] F. Rovira-Mas, S. Han, J. Wei, and J. F. Reid, “Fuzzy logic model for sensor fusion of machine vision and gps in autonomous navigation,” in *2005 ASAE Annual Meeting*. American Society of Agricultural and Biological Engineers, 2005, p. 1.
- [4] O. Pettersson, “Execution monitoring in robotics: A survey,” *Robotics and Autonomous Systems*, vol. 53, no. 2, pp. 73–88, 2005.
- [5] R. O. Duda and P. E. Hart, “Use of the hough transformation to detect lines and curves in pictures,” *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [6] F. Rovira-Más, Q. Zhang, J. F. Reid, and J. D. Will, “Hough-transform-based vision algorithm for crop row detection of an automated agricultural vehicle,” *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 219, no. 8, pp. 999–1010, Aug. 2005.
- [7] C. Gée, J. Bossu, G. Jones, and F. Truchetet, “Crop/weed discrimination in perspective agronomic images,” *Computers and Electronics in Agriculture*, vol. 60, no. 1, pp. 49–59, 2008.
- [8] B. Hamner, S. Singh, and M. Bergerman, “Improving orchard efficiency with autonomous utility vehicles,” in *2010 Pittsburgh, Pennsylvania, June 20-June 23, 2010*. American Society of Agricultural and Biological Engineers, 2010, p. 1.
- [9] D. Reiser, G. Miguel, M. V. Arellano, H. W. Griepentrog, and D. S. Paraforos, “Crop row detection in maize for developing navigation algorithms under changing plant growth stages,” in *Robot 2015: Second Iberian Robotics Conference*. Springer, 2016, pp. 371–382.
- [10] W. Winterhalter, F. V. Fleckenstein, C. Dornhege, and W. Burgard, “Crop Row Detection on Tiny Plants With the Pattern Hough Transform,” *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3394–3401, Oct. 2018.
- [11] B. Åstrand and A.-J. Baerveldt, “A vision based row-following system for agricultural field machinery,” *Mechatronics*, vol. 15, no. 2, pp. 251–269, 2005.
- [12] G. Jones, C. Gée, and F. Truchetet, “Modelling agronomic images for weed detection and comparison of crop/weed discrimination algorithm performance,” *Precision Agriculture*, vol. 10, no. 1, pp. 1–15, 2009.
- [13] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.
- [14] T. Ehret, A. Davy, J.-M. Morel, and M. Delbracio, “Image anomalies: A review and synthesis of detection methods,” *Journal of Mathematical Imaging and Vision*, vol. 61, no. 5, pp. 710–743, 2019.
- [15] N. Hirose, A. Sadeghian, M. Vázquez, P. Goebel, and S. Savarese, “Gonet: A semi-supervised deep learning approach for traversability estimation,” in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 3044–3051.
- [16] L. Wellhausen, R. Ranftl, and M. Hutter, “Safe robot navigation via multi-modal anomaly detection,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1326–1333, 2020.
- [17] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, “Estimating the support of a high-dimensional distribution,” *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.
- [18] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Müller, and M. Kloft, “Deep one-class classification,” in *International conference on machine learning*. PMLR, 2018, pp. 4393–4402.