Proceedings of the 20th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 399–404 DOI: 10.15439/2025F3412 ISSN 2300-5963 ACSIS, Vol. 43

Anomaly Detection for Unmanned Surface Vehicles Based on a Multi-Modal Bayesian Generative Model

Micheale H. Tekulu*, Ali Krayani*, Pamela Zontone*, Lucio Marcenaro*,
Francesco Caprile‡, Antonino Masaracchia†, and Carlo Regazzoni*
*DITEN, University of Genoa, Italy
†School of Electronic Engineering and Computer Science, Queen Mary University of London, UK
‡Leonardo Sistemi Integrati, Italy

***Abstract*—In this paper, we propose a novel method for abnormality detection in Unmanned Surface Vehicles (USVs) based on a Multi-Modal Bayesian generative model to enhance safety and monitoring. During the training phase, we use a Null Force Filter and an unsupervised clustering algorithm on multimodal data collected from Global Positioning System (GPS) and motor current sensors. In the testing phase, we use a Coupled Modified Markov Jump Particle Filter (CM-MJPF) to infer the GPS position and motor current of the USV, as well as to detect abnormalities in both modalities. Due to the coupled methodology, the system is able to learn the statistical similarity between the evolving GPS and motor current data. As a result, the causality of defects is inherently captured within the dynamical inference, making the proposed approach explainable.**

***Index Terms*—Bayesian generative model, anomaly detection, unmanned surface vehicle, unsupervised clustering, explainable AI.**

I. Introduction

IN Intelligent Transportation Systems (ITS), the goal of improving transport safety has led to groundbreaking developments such as autonomous surface vehicles. Since unmanned surface vehicles are autonomous, they are prone to mechanical or electrical faults [1]. If such faults are not detected accurately and in a timely manner, they may not only cause malfunction of the agent but also cause critical hazard. Therefore, to minimize potential risks and ensure robust functionality, the agent might relay on its exteroceptive and proprioceptive sensors, like motor current and GPS, to monitor its internal states while interacting with the external environment.

An anomaly is an unexpected event or behavior that could indicate a problem or fault (deviation from normality) [2], [3]. Recently, as part of machine learning, Bayesian Networks have proven well suited for anomaly detection because they support both discrete and continuous variables, model uncertainty, and allow for representation of time series data [4]. Sensors mounted on the agent help it sense and perceive the environment, and the signal measured by the sensors can reflect the agent condition. Relaying only on one sensor for the detection of operational failures can lead to inaccuracies due to inadequate information. Similarly, considering multiple sensors independently also provides limited information. Therefore, it is important to have a comprehensive system that integrates multiple interacting sensors [5].

In this work, we propose an interactive Multi-Modal Bayesian generative model based on time series data, which can detect USV operational faults. Unlike studies in literature, our approach utilizes the interaction of multiple sensors using the Bayesian learning mechanism. Additionally, our method is able to differentiate the main source of the defect by exploiting the causal behavior of Bayesian networks. The main contributions of this paper are the following.

- A novel data-driven Bayesian generative model for USVs is proposed.
- The proposed method uses an interactive multi-modal approach, where we combine both GPS and motor current for learning the model.
- The proposed framework can trace the cause of the abnormality in an explainable way.
- The proposed method can be generalized to any kind of multi-modal sensors.

The remainder of the paper is organized as follows: Section 2 presents a review of related work on probabilistic inference approaches for abnormality detection in autonomous vehicles. Section 3 describes the proposed method in detail. After this, the simulation and experimental results obtained from the proposed approach, along with a discussion, are presented in Section 4. Finally, Section 5 concludes the paper and outlines directions for future work.

II. Related Work

Model-based fault diagnosis uses physics-based or mathematical models to detect and identify system faults. These models describe the expected behavior of vehicle components under normal conditions, allowing for comparison with real-time sensor data. In [6], the authors propose a steering wheel fault diagnosis method for autonomous vehicles, using a model-based residual generator with an SVM (Support Vector Machine) classifier. Similarly, in [7], they use Position Estimation Models, along with an Extended Kalman Filter (EKF), to detect faults by computing residuals (Mahalanobis distance) between estimated positions from sensor pairs. Abci

B. et al. [8] also formulate a model-based fault detection and isolation (FDI) system for autonomous mobile robots that can detect and isolate sensor and actuator faults using an informational framework. Although model-based fault diagnosis applications are useful, they have some drawbacks, such as inaccuracies in modeling complex scenarios.

Due to the improvement of machine learning and computational techniques, there has been growing interest in data-driven fault diagnosis. This paper, therefore, takes advantage of data-based abnormality detection methods focused on autonomous vehicles. While only a few studies in the literature directly address the operational fault diagnosis for USVs, as examined in this paper, several related works share similar characteristics and offer relevant insights. In [10], the authors use a data-driven Neural network to detect and isolate faults, based on the differences between the measured and estimated outputs (residuals). Furthermore, in [9], Gaussian process regression to generate the residuals is used.

The methods mentioned above focus on a single data source. However, as systems become more complex, relying on data from just one source is not sufficient for effective monitoring. To address this limitation, fault diagnosis systems now employ multi-sensor technology to capture a wider range of characteristic data. In [13], the authors propose convolutional neural network (CNN) architectures for the fault diagnosis of an induction motor, using different combinations of multi-sensory signals such as vibration, current, voltage, and speed. Similarly, Ö. Gültekin et al. [5] utilize time–frequency representations of time series data obtained from both vibration and sound sensors, and use a CNN model for the classification process to diagnose the operational faults of autonomous transfer vehicles. Shaoxuan Xia et al. [14] applied a hierarchical attention-based multi-source data fusion method for fault diagnosis in Autonomous Underwater Vehicles (AUVs), addressing challenges like multi-source data heterogeneity and strong coupling. They used a Bidirectional Long Short-Term Memory (BiLSTM) network for feature extraction and a Multilayer perceptron (MLP) for fault detection. Although multisensory fusion provides more information compared to using a single type of sensor, previous studies in the literature often consider each sensor independently. Moreover, they did not explain the cause of a defect. Unlike most of the papers available in the literature, our method is an interactive and multi-modal approach, and is able to explain the cause of a defect.

III. Methodology

A. Experimental Setup

We use a telemetry system to collect the necessary data for our experiment, as shown in fig. 1. Data from the inverter are gathered via the Controller Area Network bus (CANbus) communication protocol. This is achieved by sending a request to the inverter for a specific parameter reading, after which the inverter responds with the corresponding measurement. Moreover, the GPS module communicate via universal asynchronous receiver / transmitter (UART). Data provided by

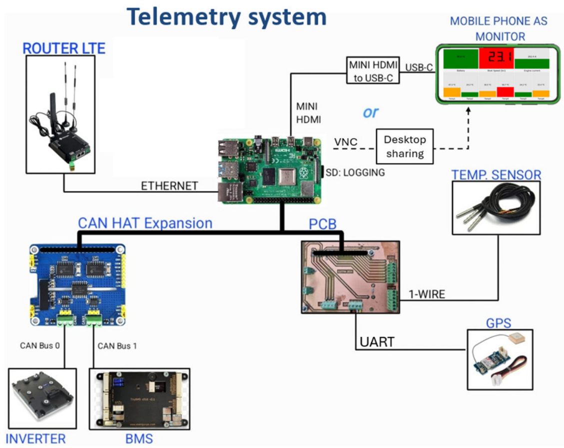


Fig. 1: Experimental setup-Telemetry system

GPS are encoded according to the NMEA (National Marine Electronics Association) standard.

B. Preliminaries

- Generative models: There are two types of models in machine learning, i.e., discriminative and generative. Discriminative models focus on learning the probability of a specific class label given an input, which is the approach commonly used in classification tasks. Conversely, generative models learn the underlying data distribution, enabling them to produce new data samples that follow the same pattern.
- Bayesian Networks (BNs): They are a type of generative model that allow us to model causal relationship between random variables.
- Dynamic Bayesian networks (DBNs): They are a particular type of BN which can describe dynamic processes that evolve over time.
  The main advantage of a DBN is its capability to formulate causal relationships between semantic information (high level) and sensory data (low level) in a hierarchical level. DBNs with hierarchical structure are called Hierarchical Dynamic Bayesian networks (HDBNs). Fig. 2 shows a DBN representation for three hierarchical variable levels. The nodes represent the random variables for a continuous level state (X_t and Z_t) or a discrete level state (S_t). The links represent conditional dependencies between nodes.

C. Proposed DBN for Learning Interaction

1) Training phase: Initially, the USV perceives its surroundings under the static assumption that the environmental state does not change. Therefore, the USV predicts its position (GPS) and internal state (motor current) using a null force filter with the following model:

$$X_t^n = X_{t-1}^n + W_t^n, \tag{1}$$

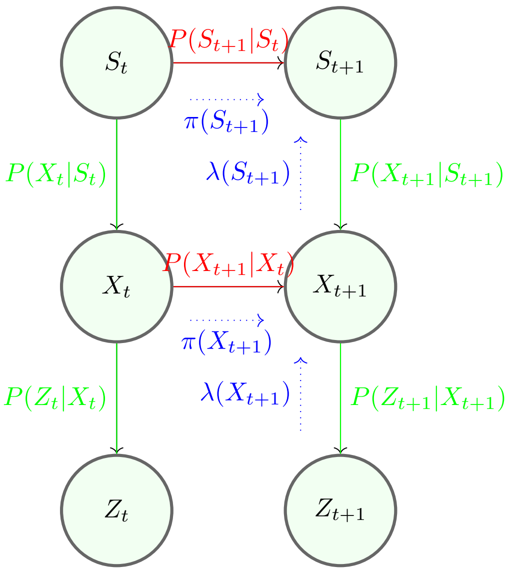


Fig. 2: Three level DBN. The links highlighted in red are inter-slice links, and the ones in green are intra-slice links at different times.

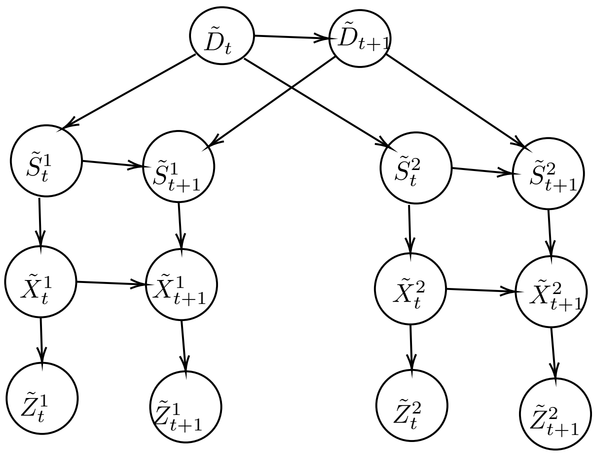


Fig. 3: Learning interactions among GDBN Models for both modalities (1 for motor current and 2 is for GPS), where $\tilde{D}$ is the Interaction Discrete level, $\tilde{S}$ is the Superstate Discrete level, $\tilde{X}$ is the continuous level and $\tilde{Z}$ is the measurement level

where W_t^n represents the noise accounting for the measurement uncertainty of the n-th sensor (with n=1 for motor current and n=2 for GPS) at time t.

In order to explore the dynamic rule through time in the environment, a new rule should be created by exploiting the generalized error, which is the difference between prediction and observation. Generalized errors $\tilde{X}_t$ are expressed as:

$$\tilde{X}_t = [X_t, \dot{X}_t]^T \tag{2}$$

where X_t is calculated using equation (1), and $\dot{X}_t$ can be considered as the innovation of the null force filter calculated

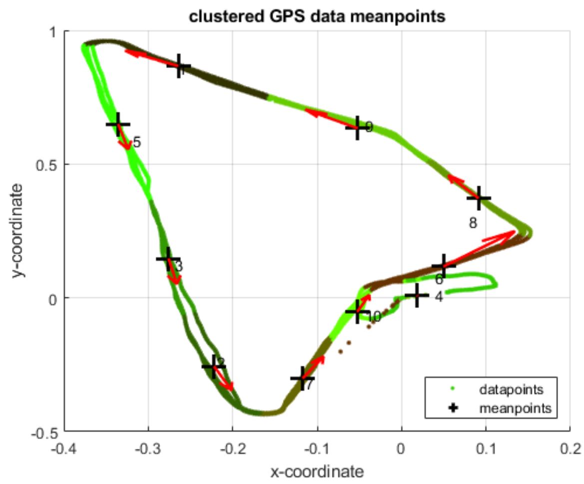


Fig. 4: Clustered GPS data

using this equation:

$$\dot{X}_t = H^{-1}Z_t - X_t \tag{3}$$

where Z_t is the measurement acquired from the sensors at time t and H is matrix that maps the observation to the state X.

Consequently, the Generalized errors collected from previous experience are used as input to an unsupervised clustering algorithm (i.e., Growing Neural Gas (GNG)), which encodes them to discrete variables or neurons ($\tilde{S}^n$) called super states (letters):

$$\tilde{S}^n = \tilde{S}_1^n, \tilde{S}_2^n, ..., \tilde{S}_M^n \tag{4}$$

where M is the total number of superstates.

Each generalized superstate $\tilde{S}_t^n(\tilde{S}_t^n \in \tilde{S}^n)$ is associated with its corresponding statistical properties mean $(M^{\tilde{S}_t^n})$ and covariance matrix$(Q^{\tilde{S}_t^n})$, as shown in figure 4.

Following the clustering of each sensor's generalized errors, a word or meta cluster ($\tilde{D}_t$) is formed, containing the superstate clusters (letters) of the two sensors, motor current and GPS ($\tilde{S}_t^1$ and $\tilde{S}_t^2$), at time t.

After meta-clustering, a $P \times P$ word transition matrix (Π) is defined as:

$$\Pi = \begin{bmatrix} \pi(\tilde{D}_t = \tilde{D}_1) \\ \vdots \\ \pi(\tilde{D}_t = \tilde{D}_p) \end{bmatrix} = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1P} \\ \vdots & \ddots & \vdots \\ \pi_{P1} & \cdots & \pi_{pp} \end{bmatrix} \tag{5}$$

where P is the number of unique words in the dictionary. This matrix is learned by estimating the transition probabilities $\pi_{ij} = P(\tilde{D}_t = j|\tilde{D}_{t-1} = i), j, i \in \tilde{D}$ over time.

2) Testing phase:

- *Real time Joint Position and Motor current inference:* In order to perform inference at different hierarchical levels, an interactive modified Markov Jump Particle Filter (IM-MJPF) has been employed, as illustrated in figure 3, to extract the knowledge embodied in the

coupled GDBN (C-GDBN). The MJPF is a switching model that uses a combination of a particle filter (PF), to predict discrete superstates, and Kalman Filters (KF) for continuous states prediction and estimation. Such a switching behavior, between the dynamic transitions at discrete/continuous levels and observations, enables the updating of the belief about the motor current state and its corresponding position by passing local messages in simultaneous inference modes, namely the predictive or causal inference (top-down) and the diagnostic inference (bottom-up). The top down messages from $\tilde{D}$ to $\tilde{X}$ depend on the clustered values, including super states with their corresponding mean and covariance statistics. The temporal predictive messages ($\pi(\tilde{D}_t) = P(\tilde{D}_t|\tilde{D}_{t-1})$) and ($\pi(\tilde{S}_t{}^n) = P(\tilde{S}_t{}^n|\tilde{S}_{t-1}^n), \pi(\tilde{X}_t{}^n) = P(\tilde{X}_t{}^n|\tilde{X}_{t-1}^n)$) for both modalities depend on the dynamic rule stored in the model (intra-slice messages) as shown in figure 2.
The particle filter is employed to predict the discrete values (or words) relied in the word transition matrix encoded in the dynamic model as a proposal distribution. Once the word is predicted, each superstate of the two modalities is inherently contained within the word. Then, for each particle, a Kalman filter is employed to predict continuous states (position and motor current) of the agent using this equation: (6).

$$\tilde{X}_t^n = A\tilde{X}_{t-1}^n + BU^{\tilde{S}_t^n} + W_t^n \tag{6}$$

where A and B are the dynamic model matrix and the control model matrix, respectively. And $U^{\tilde{S}_t^n}$ is a control vector which represents the dynamic rules of the signal temporal evolution encoded in the superstate.
Once the evidence is observed, a message is backward-propagated from the bottom level towards the higher levels (λ messages as shown in figure 2). This message (reasoning) is based on likelihood models consisting of messages ($\lambda(\tilde{D}_t), \lambda(\tilde{S}_t{}^n), \lambda(\tilde{X}_t{}^n)$), in which all continuous nodes are Gaussian-distributed variables and the two GDBN models are conditionally independent given the interaction node:

$$\lambda(\tilde{D}_t) = \prod_{i=1}^{n} \lambda(\tilde{S}_t{}^n) \tag{7}$$

$$\lambda(\tilde{S}_t{}^n) = \lambda(\tilde{X}_t{}^n) P(\tilde{X}_t{}^n|\tilde{S}_t{}^n) \tag{8}$$

$$\lambda(\tilde{X}_t{}^n) = P(\tilde{Z}_t{}^n|\tilde{X}_t{}^n) \approx N(\mu_{\tilde{Z}_t{}^n}, \Sigma_{\tilde{Z}_t{}^n}) \tag{9}$$

Consequently, each posterior distribution ($belief = \pi(.) * \lambda(.)$) at word, superstate, and state level is updated according to $\lambda(\tilde{D}_t), \lambda(\tilde{S}_t{}^n), \lambda(\tilde{X}_t{}^n)$, respectively. Furthermore, these beliefs will represent the initial states on the next time instance.

- *Multi-modal Abnormality Measurements:*
During the journey, the USV checks whether it is in an optimal state or experiencing an abnormality by predicting the motor current and GPS signal. When new data arrive, the trained model uses the GDBN to make sense of it and spot different types or levels of unusual patterns. To measure the motor current or GPS deviation we use an abnormality indicator at superstate level defined as a distance between the predictive message ($\pi(\tilde{S}_t{}^1)$) and the diagnostic message ($\lambda(\tilde{S}_t{}^1)$) entering the node $\tilde{S}_t{}^1$.
We use the symmetric Kullback-Leibler Divergence (KLDA) to measure the similarity between the two discrete probability distributions ($\pi(\tilde{S}_t{}^1)$ and $\lambda(\tilde{S}_t{}^1)$) defined by:

$$\begin{aligned} KLDA = D_{KL} &\left(\pi(\tilde{S}_t{}^1) \parallel \lambda(\tilde{S}_t{}^1)\right) \\ &+ D_{KL}\left(\lambda(\tilde{S}_t{}^1) \parallel \pi(\tilde{S}_t{}^1)\right) \end{aligned} \tag{10}$$

where $\parallel$ identifies the divergence.
Similarly, we use the Bhattacharyya distance (BD) to measure the probabilistic distance between the two contínous probability distributions ($\pi(\tilde{X}_t{}^1)$ and $\lambda(\tilde{X}_t{}^1)$) defined by:

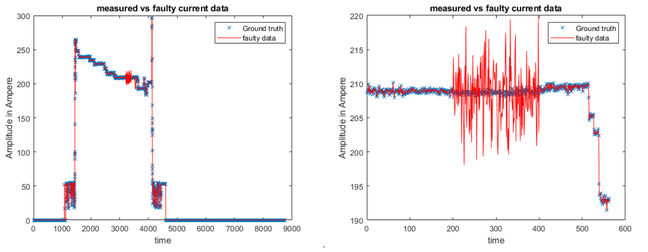
$$BD(p,q) = -ln\left(\sum_{i=1}^{k} \sqrt{p_i * q_i}\right) \tag{11}$$

where $p = \pi(\tilde{X}_t{}^1)$, $q = \lambda(\tilde{X}_t{}^1)$, and k represents the number of elements considered in the computation. Furthermore, this can be calculated for all other possible sensors attached to the agent including GPS.
- *Evaluation matrices:*
To assess the performance of our multi-modal abnormality detection approach, we use the Receiver Operating Characteristic (ROC) curve. The ROC curve visually shows how a model performs at different classification thresholds. It is created by plotting the true positive rate (TPR) against the false positive rate (FPR) for all possible threshold values. The AUC, or Area Under the Curve, summarizes this graph into a single number, indicating the model's ability to distinguish between classes. An AUC of 1.0 reflects perfect classification, whereas an AUC of 0.5 suggests the model is no better than making random guesses.

IV. Simulation and Experimental Results

A permanent magnet synchronous motor (PMSM) is a motor known for its outstanding dynamic performance and high reliability. It is extensively applied in areas such as electric vehicles, rail transportation, smart manufacturing, and more [11]. Our USV uses PMSM as its core engine.

An inter-turn short circuit (ITSC) fault in a stator typically happens when the insulation between the windings of the same phase breaks down. In this section, we perform the simulation of the ITSC fault to test our abnormality detection framework. ITSC in the stator results in a huge eddy current in the short circuit [12]. This will increase the RMS (Root Mean Square) current amplitude in the motor. This is simulated by adding Gaussian noise to the real motor current data, as shown in figure 5.



(a) Normal current data vs faulty (b) Chunked version-Visualization

Fig. 5: Normal vs Faulty RMS Current signal

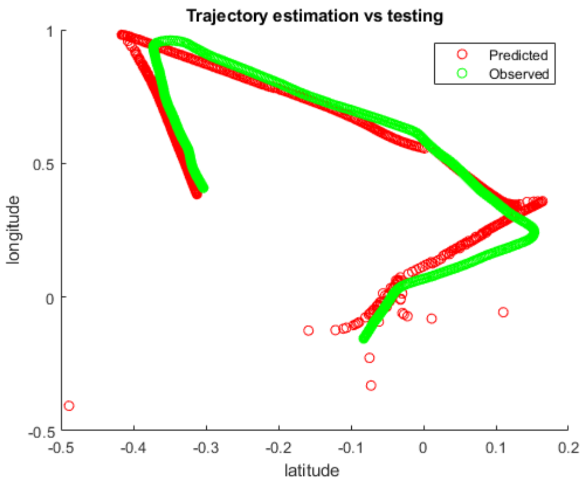


Fig. 6: Trajectory (chunk of the GPS data) estimation vs observed

The multi-sensor tracking capability of the USV is demonstrated by the spatio-temporal prediction of the trajectories, as in figure 6. Consequently, the motor current can be estimated based on the interaction between the two sensors as shown in figure 7.

To enable scalable, efficient, and stochastic learning of the posterior, we divided our dataset into 8 examples (mini-batches) during the training of the Bayesian generative model. During testing, we use one of the batches containing a simulated motor fault as the test set.

The decision making process of the autonomous agent is illustrated in figures 8 and 9, in which the abnormality is detected if it passes the predefined threshold (set during training). This demonstrates the effectiveness of our framework in detecting abnormalities. Moreover, the cause of the abnormality can be traced to either a GPS error or a motor current fault. This makes our framework explainable.

In addition, we carried out a performance test for the continuous state (motor current) abnormality detection. The ROC curves shown in both figures 10 and 11 indicate that our framework achieves good abnormality detection efficiency.

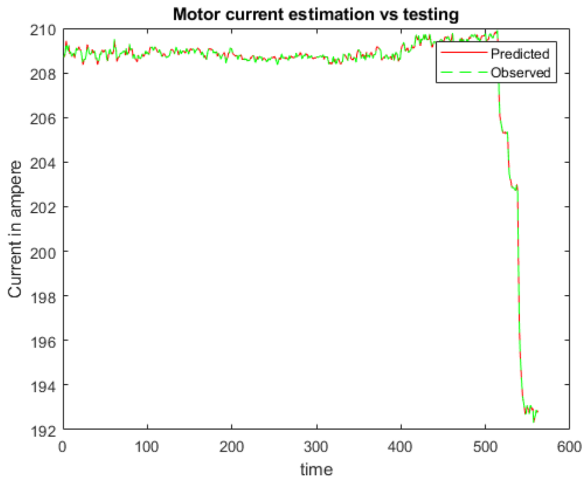


Fig. 7: Motor current estimation vs observed (chunk of motor current data cross ponding to GPS shown in figure 6)

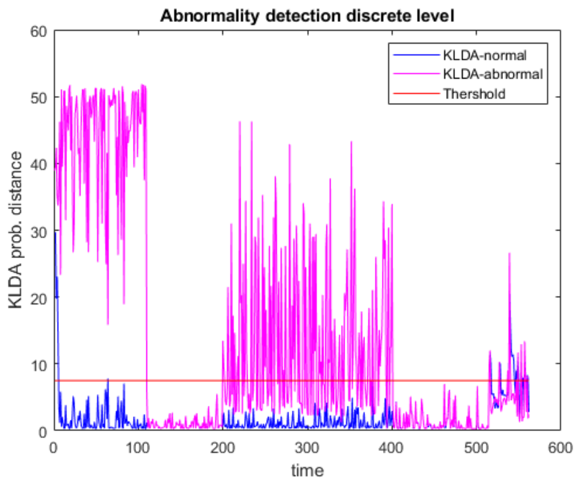


Fig. 8: Tracking abnormality at cluster level (Superstate level)

V. Conclusion and Future Work

This paper proposes a multi-modal Bayesian generative model for fault detection and probabilistic inference of system states in an Unmanned Surface Vehicle (USV). The proposed approach uses coupled Dynamic Bayesian Networks (DBNs) to assimilate different observations over time, enabling incremental learning and inherent explainability. The method predicts future states based on past observations, providing an interpretable explanation of the model's reasoning.

The performance of the proposed method is evaluated using real-world data and simulated faults that mimic actual motor malfunctions. The results indicate that the approach can predict and detect faults in a USV with high accuracy. Furthermore, it is capable of identifying the root cause of the defect.

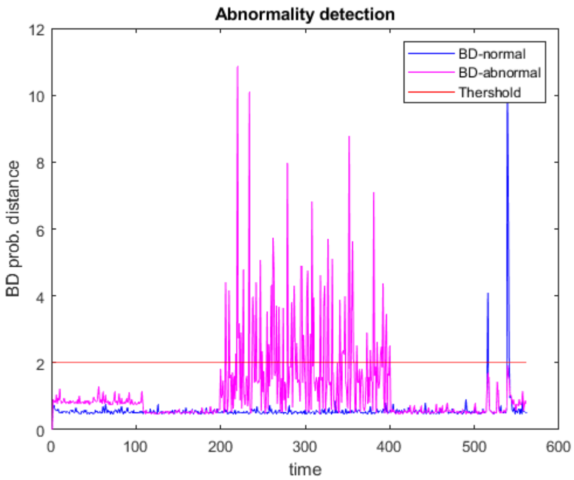


Fig. 9: Tracking abnormality at continuous level

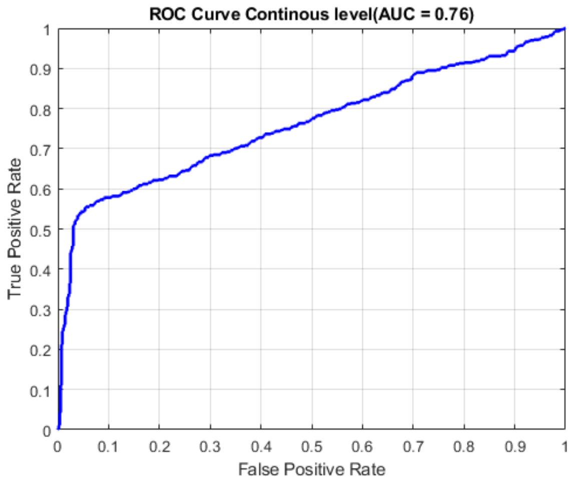


Fig. 11: ROC (BD) curve for motor abnormality detection

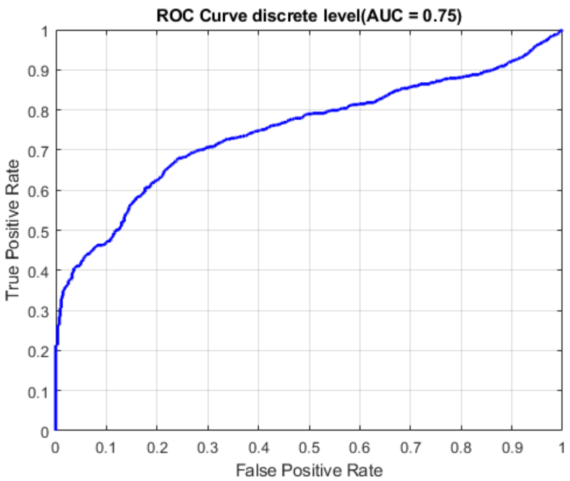


Fig. 10: ROC (KLDA) curve for motor abnormality detection

In our future work, we are planning to calculate meta-cluster level (word level) abnormality to detect possible anomalies on the agent and to learn the time evolution synchronization between the multi-modal sensors used in our approach.

REFERENCES

[1] S. Baccari, M. Hadded, H. Ghazzai, H. Touati and M. Elhadef, "Anomaly Detection in Connected and Autonomous Vehicles: A Survey, Analysis, and Research Challenges," in IEEE Access, vol. 12, pp. 19250-19276, 2024, doi: https://doi.org/10.1109/ACCESS.2024.3361829.

[2] G. Slavic, M. Bracco, L. Marcenaro, D. M. Gómez, C. Regazzoni, and P. Zontone, "Joint Data-Driven Analysis of Visual-Odometric Anomaly Signals in Generative Ai-Based Agents," 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Seoul, Korea, Republic of, 2024, pp. 264-268, doi: 10.1109/ICASSPW62465.2024.10626634.

[3] S. Memon, A. Krayani, P. Zontone, L. Marcenaro, D. M. Gomez, and C. Regazzoni, "Learning 3D LiDAR Perception Models for Self-Aware Autonomous Systems," 2024 27th International Conference on Information Fusion (FUSION), Venice, Italy, 2024, pp. 1-8, doi: 10.23919/FUSION59988.2024.10706364.

[4] https://www.bayesserver.com/docs/techniques/anomaly-detection/

[5] Ö. Gültekin, E. Cinar, K. Özkan, and A. Yazıcı, "Multisensory data fusion-based deep learning approach for fault diagnosis of an industrial autonomous transfer vehicle," Expert Systems with Applications, vol. 200, p. 117055, Aug. 15, 2022, doi: https://doi.org/10.1016/j.eswa.2022.117055.

[6] Q. Shi and H. Zhang, "Fault Diagnosis of an Autonomous Vehicle With an Improved SVM Algorithm Subject to Unbalanced Datasets," in IEEE Transactions on Industrial Electronics, vol. 68, no. 7, pp. 6248-6256, July 2021, doi: https://doi.org/10.1109/TIE.2020.2994868.

[7] Pratama, P.S., Gulakari, A.V., Setiawan, Y.D. et al. Trajectory tracking and fault detection algorithm for automatic guided vehicle based on multiple positioning modules. Int. J. Control Autom. Syst. 14, 400–410 (2016), doi: https://doi.org/10.1007/s12555-014-0294-y

[8] Abci, B., El Badaoui El Najjar, M., Cocquempot, V. et al. An informational approach for sensor and actuator fault diagnosis for autonomous mobile robots. J Intell Robot Syst 99, 387–406 (2020), doi: https://doi.org/10.1007/s10846-019-01099-7

[9] P. Castaldi, S. Farsoni, M. Menghini and S. Simani, "Data–Driven Fault Detection and Isolation of the Actuators of an Autonomous Underwater Vehicle," 2021 5th International Conference on Control and Fault-Tolerant Systems (SysTol), Saint-Raphael, France, 2021, pp. 139-144, doi: https://doi.org/10.1109/SysTol52990.2021.9595605.

[10] Changxin Zhang and etal (2022). A data-driven fault detection and diagnosis method via just-in-time learning for unmanned ground vehicles. Automatika, 64(2), 277–290, https://doi.org/10.1080/00051144.2022.2142924

[11] Chen, Z.; Liang, K.; Peng, T.; Wang, Y. Multi-Condition PMSM Fault Diagnosis Based on Convolutional Neural network Phase Tracker. Symmetry 2022, 14, 295, doi:https://doi.org/10.3390/sym14020295

[12] Chen, X.; Qin, P.; Chen, Y.; Zhao, J.; Li, W.; Mao, Y.; Zhao, T. Inter-Turn Short Circuit Fault Diagnosis of PMSM. Electronics 2022, 11, 1576, doi: https://doi.org/10.3390/electronics11101576

[13] Ayankoso, S., He, Y., Gu, F., Ball, A. (2024). Multisensory Data-Based Fault Diagnosis of Induction Motors Using 1D and 2D Convolutional Neural Networks. In: Ball, A.D., Ouyang, H., Sinha, J.K., Wang, Z. (eds) Proceedings of the UNIfied Conference of DAMAS, IncoME and TEPEN Conferences (UNIfied 2023). TEPEN IncoME-V DAMAS 2023 2023 2023. Mechanisms and Machine Science, vol 152. Springer, Cham, doi: https://doi.org/10.1007/978-3-031-49421-5_92

[14] Shaoxuan Xia, Xiaofeng Zhou, Haibo Shi, Shuai Li, Chunhui Xu, A fault diagnosis method with multi-source data fusion based on hierarchical attention for AUV, Ocean Engineering, Volume 266, Part 1,2022,112595, ISSN 0029-8018, doi: https://doi.org/10.1016/j.oceaneng.2022.112595.