Learning 3D LiDAR Perception Models for Self-Aware Autonomous Systems

Saleemullah Memon[1,2], Ali Krayani[1], Pamela Zontone[1], Lucio Marcenaro[1], David Martin Gomez[2], and Carlo Regazzoni[1]
[1]*Department of Engineering and Naval Architecture (DITEN), University of Genova, Italy*
[2]*Intelligent Systems Lab, University Carlos III de Madrid, Spain*
emails: saleemmemon1320@gmail.com, ali.krayani@edu.unige.it,
{pamela.zontone, lucio.marcenaro, carlo.regazzoni}@unige.it, dmgomez@ing.uc3m.es

***Abstract*—Intelligent transportation systems (ITSs) provide a paradigm change in perceiving and interacting with transportation networks, leading to enhanced levels of safety, sustainability, and efficiency. Vehicular-to-everything (V2X) communication is the core component in the ITSs. The proprioceptive and exteroceptive sensors allow these vehicles to be aware of the surrounding environment and respond to emergencies by utilizing their abilities to reach a high level of self-awareness. In this paper, we propose a self-awareness approach to learn a generative dynamic Bayesian network (G-DBN) from the real-time LiDAR perception. Without reducing the dimensionality, we perform offline training and online testing phases on the three-dimensional (3D) point clouds. In the offline training phase, initially, the raw point clouds are preprocessed using a joint probabilistic data association filter (JPDAF) to obtain the 3D tracks of the multiple vehicles in space. Then, we perform an unsupervised clustering on all the generalized states (GSs) containing positions and velocities (a 6D vector) by considering the growing neural gas (GNG) technique, thus achieving a trained model from the 3D LiDAR point clouds. In the online testing phase, the high-dimensional Markov jump particle filter (HD-MJPF) utilizes the G-DBN's probabilistic information to predict the positions of multiple vehicles and to detect the abnormalities at the discrete and continuous levels in normal and abnormal scenarios. Our proposed approach is useful for learning high-dimensional generative models and provides a way to meet the current curse of dimensionality challenges, that machine learning models are suffering.**

***Index Terms*—Generative Dynamic Bayesian Network (G-DBN), 3D LiDAR, Self-Awareness, Intelligent Transportation Systems (ITSs)**

I. Introduction

Intelligent transportation systems (ITSs) effectively integrate innovative and advanced technologies containing artificial intelligence (AI), self-awareness, internet of things, sensors, and automatic controls. ITSs lead to higher levels of safety, sustainability, and efficiency for smart cities, environmental monitoring, and industrial automation systems [1]. For efficient ITSs, vehicle-to-everything (V2X) communication allows vehicles to engage with various external elements, such as other vehicles (V2V), infrastructure (V2I), networks (V2N), and even pedestrians (V2P). The various proprioceptive and exteroceptive sensors including, LiDAR, RaDAR, cameras, and communication devices equipped with these vehicles allow them to sense the surrounding environment and themselves. These co-existing data make it feasible to enhance effectiveness and safety through real-time traffic monitoring, incident detection, blockage prediction, collision avoidance, emergency braking, and adaptive traffic signal control for self-aware autonomous systems. Similarly, in manual driving scenarios, systems that use proprioceptive sensors to record drivers' physiological signals for analyzing their emotional state [2]–[5] could integrate these signals with video, LiDAR, or other relevant data (coming from exteroceptive sensors) to enhance safety.

The three-dimensional (3D) LiDAR sensors capture high-resolution and densely clouded data, which provide more efficient spatial information and greater precision as compared to cameras [6]. Experts have recently introduced some techniques for combining cameras and LiDAR to improve the overall performance of autonomous systems [6]–[8]. Low-dimensional data such as the angle of steering and the position of the vehicle can easily be modeled by probabilistic reasoning and machine learning approaches [9]. The raw data collected from the real world is often diverse, high-dimensional, and complex [10], hence, these learning approaches often find it challenging to manage a limited number of high-dimensional data. Some well-known 3D models have been introduced in the literature, such as 3D convolutional neural networks (CNNs) [11], EfficientLO-Net [12], and Multiview CNNs [13]. However, these techniques have several drawbacks, such as sparse volumes, limited data, high computational costs, and complex comprehension scenarios. Also, these investigations do not concentrate on dynamic abnormality detection and tracking relevant characteristics from the LiDAR point clouds in the framework of self-aware autonomous systems. Therefore, in this paper, we introduce a self-aware approach for learning a generative dynamic Bayesian model from the high-dimensional 3D LiDAR point cloud perception. Without reducing the dimensionality of the point clouds, we perform offline training and online testing phases on the real-time 3D LiDAR sensor data. The main contributions of this paper are:

- To simultaneously detect and track multiple vehicles from a stationary 3D LiDAR sensor observations, by using a joint probabilistic data association filter (JPDAF).
- To learn a 3D LiDAR perception-based generative dynamic Bayesian network (G-DBN) model from the gen-

eralized states (GSs) of the tracked objects containing a 6D vector of positions and velocities.
- To detect the abnormalities at discrete and continuous levels of our proposed model from a high-dimensional Markov jump particle filter (HD-MJPF) and compare it with the unseen situations.

The remaining organization of this paper is described as follows. Section II explains the basics of 3D LiDAR point clouds and gives an overview of dynamic Bayesian networks (DBNs). Section III defines the system model. Section IV describes the proposed methodology. Section V explains the simulation outcomes of the HD-MJPF. Section VI finally draws some conclusions and possible future work.

II. Background and Related Work

The background technologies that define the basics of our proposed approach are described in this section. It includes an overview of 3D LiDAR sensors and DBNs studied by various researchers in the literature.

A. Light Detection and Ranging (LiDAR) Point Clouds

LiDAR systems are considered the eyes of modern autonomous driving vehicles, which accurately perceive the surroundings and produce high-definition and high-resolution point clouds. They provide a 360-degree view of the whole environment and use distance and angle measurements within their range of view. 3D LiDAR sensors record such measurements for multiple locations in space to produce a 3D map of the surrounding environment. A standard 3D LiDAR point cloud contains the X, Y, and Z coordinates of each point as well as extra properties like intensity, return type, and even color information. The large number of points and the multiple attributes associated with each point make 3D point clouds highly dimensional. The point clouds have very high resolution, showing the sensor's ability to detect even tiny objects present near the agent. Plenty of research has been done by considering 2D LiDARs [11], [14], and LiDAR with additional sensors [15], [16], to improve the overall performance of autonomous driving. With several advantages, the LiDAR systems also have some limitations to perform accurately under adverse and harsh weather conditions such as rain, snow, wind, and fog by absorbing or scattering laser beams [17]. In this situation, the LiDAR produces inconsistent and anomalous point clouds. Consequently, the precision as well as the reliability of the point clouds are reduced in outdoor environments [18], [19].

B. Dynamic Bayesian Networks (DBNs)

DBNs are probabilistic graphical models (PGMs) that combine the concepts of both the probabilistic and graph theories. A DBN model contains various nodes that represent random variables, and directed links between these nodes indicate a conditional dependency between the linked nodes evolving over time. The DBN models are helpful in a variety of domains where time-dependent representations need to be described as they offer a powerful framework for explaining and reasoning the temporal evolution of complex systems. An example of a 3-level, 2-slice temporal DBN is shown in Fig. 1, where the model captures the temporal evolution of the system from the present time t to the future time t+1. Two types of links can be found in the model, i.e., the inter-slice and intra-slice directed links, which connect the nodes across different and at the same time instant, respectively. Three level random variables are shown at each time instant: the low-level node ($\tilde{Z}_t$) corresponds to the observations that come from the sensors, and two high-level nodes ($\tilde{X}_t$) and ($\tilde{S}_t$) correspond to the continuous and discrete hidden level, respectively. Due to the hierarchical nature, which encodes the semantic relationships between low-level and high-level random variables, these causal models have been extensively used in the literature [20]–[22].

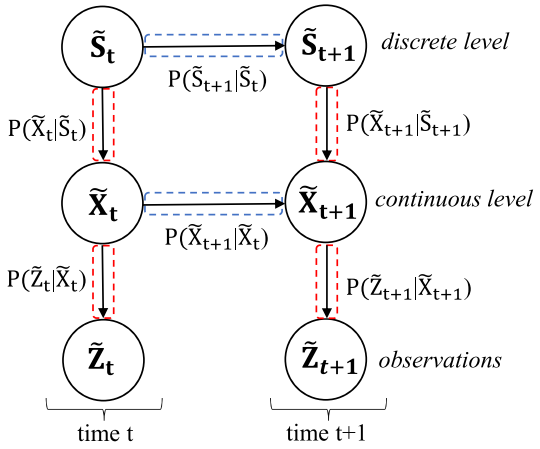


Fig. 1: An example of a 3-level 2-slice DBN. The blue dotted lines represent the inter-slice links and the red dotted lines represent the intra-slice links. The brackets at the bottom depict the variable nodes at present time t and future time t + 1.

III. System Model

Our proposed system model contains a high-frequency mmWave base station (BS), which acts as a wireless roadside receiver. Autonomous vehicles (AVs), which move near the BS, act as wireless transmitters. The fixed BS is equipped with a synchronized 3D LiDAR sensor which provides valuable awareness about the surrounding environment, as shown in Fig. 2. The high-resolution ability of the sensor provides accurate mobility patterns, locations of the AVs, pedestrians, and other static and dynamic objects present in the scenario. With the help of this environmental awareness, the BS is considered as a self-aware agent. It leverages this sensing information and enhances the network performance and connectivity by predicting the future positions of these dynamic objects. The generative models are suitable in this case to learn incrementally from the time series LiDAR point cloud observations. The learned model can adapt to unforeseen situations in the environment, by updating its belief about the

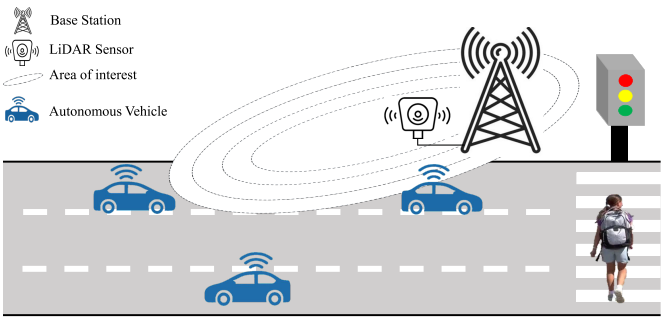


Fig. 2: An illustration of the system model.

hidden states as soon as new observations from the LiDAR sensor appear.

IV. Proposed Methodology

In this section, we will begin by introducing the real-world dataset, the initial preprocessing steps of the 3D LiDAR point clouds, and the detection and tracking of multiple dynamic vehicles through the JPDAF approach. Later, the proposed methodology, which comprises two phases is described, i.e., the offline training phase and the online testing phase. The whole process is illustrated in Fig. 3.

A. Dataset: Overview & Preprocessing

1) Deep sense 6G dataset: For both the offline training and the online testing phases, we consider the deep sense 6G dataset [23], which is based on real-world observations and is suitable for our scenario.

The street-level vehicle-to-infrastructure (V2I) communication scenario (i.e., scenario n. 31 in the dataset) collects data in an outdoor wireless environment. The deep sense 6G dataset is collected in a two-way city street on McAllister Avenue in Arizona, USA. Here, vehicles of various sizes travel at different speeds in both directions, with a street width of 10.6 meters, and a vehicle speed limit of 25mph (40.6 km per hour). To perceive the physical environment, only a time series of point cloud data, captured by the exteroceptive 3D LiDAR sensor is taken into account for learning a G-DBN model. The LiDAR is fixed on the vehicle (which is static), its range is 100 meters, and the maximum motor spin frequency is 20Hz. A total of 7012 samples of 3D point clouds are captured during the daytime scenario.

2) Preprocessing of 3D point clouds: Initially, the raw point clouds are preprocessed using theLiDAR viewer toolbox (MATLAB 2024.a). The ground removal filter is applied to remove the ground plane up to 0.25 meters, which contains unwanted information about the road and other small plants. The cropping method is chosen to consider only the region of interest where multiple vehicles are appearing. We chose X-limits and Y-limits from -60 to 50 meters, and Z-limits from -1.5 to 12 meters. The denoising neighbor filter is then chosen to remove the outliers from the space. In this way, each 3D point cloud is carefully retrieved from the region

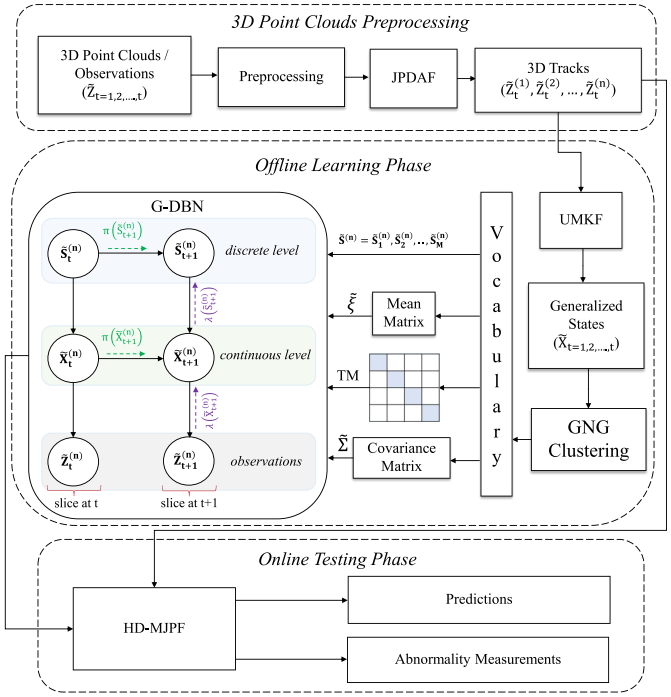


Fig. 3: An illustration of the proposed block scheme. The top part depicts the preprocessing step of 3D point cloud data, the middle part shows the offline training phase, and the bottom part shows the online testing phase.

to achieve meaningful information about the dynamic objects (vehicles) in our region of interest, as shown in Fig. 4. In each point cloud, almost 10000 points are extracted from the 16474 total points. These points are still considered as uncertain observations for the next prepossessing step, as we are interested in tracking the vehicles and their corresponding positions appearing at different time instants inside the scene.

3) Joint Probabilistic Data Association Filter (JPDAF): The uncertain and noisy 3D LiDAR point cloud observations $\tilde{Z}_{t=1,2,\ldots,t}$ at different times t, lead to uncertain predictions of the paths that the vehicles are taking. As in the Kalman filter, the predictions for the appearance of a single object can be corrected by considering the uncertain measurements. A single Kalman filter fails to perform this task of predicting the future positions of multiple objects by considering uncertain measurements. For sparsely distributed observations, the global nearest neighbor (GNN) approach better tracks the objects, where the nearest observation to the track is taken for the updating process. The approach is simple with low computational cost, but not suitable for highly dense 3D point clouds because of the uncertainty in assigning the proper point clouds to the corresponding objects [24]. The challenge also comes when a single point cloud observation belongs to different objects. This problem is commonly referred to as the data association problem. To deal with these issues, we consider a JPDAF approach, which offers an optimal solution when multiple targets are present in the environment [25], [26].

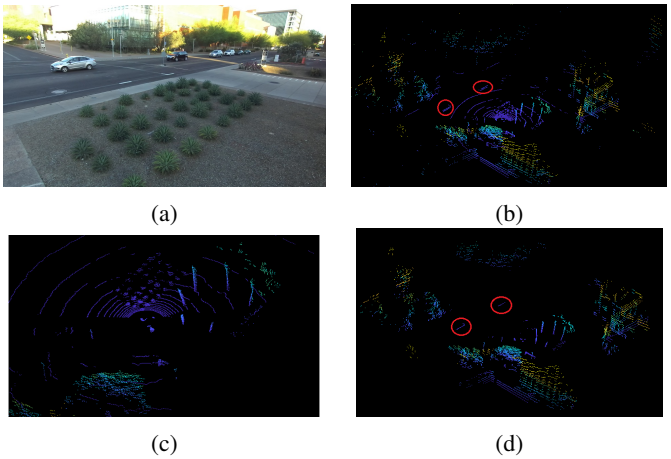


Fig. 4: Extracting features from 3D point clouds: (a) Real-time scenario captured by the camera sensor, (b) 3D LiDAR's raw point clouds, where red circles show the vehicles, (c) Zoom-in view showing the ground plane and other small plants that are unwanted for our case, (d) Required features from the point clouds after applying ground removal, cropping, and denoising neighbor filtering methods.

JPDAF is an extension of the probabilistic data association filter which adds a logic layer to ensure the tracks are correctly distributed over the observations $\tilde{Z}_{t=1,2,\ldots,t}$. For detection, we have considered the bounding box detector which processes the point clouds within the defined spatial limits, segments it into clusters, and generates bounding boxes around the detected vehicles. Each observation within a gated region enclosing a track is considered as a potential update for the target. This detection and tracking of multiple vehicles over time is taken as an important pre-processing step. The workflow of this whole process is shown in Fig. 5.

Initially, we created the tentative tracks by considering all the detected vehicles from the point clouds. Then, we

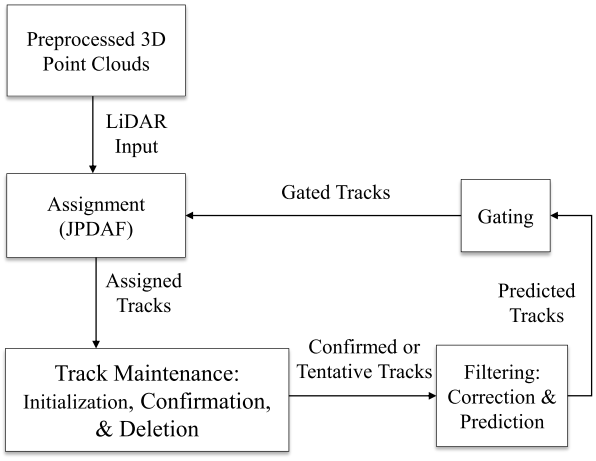


Fig. 5: A multiple vehicle tracking approach, where the JPDAF is used in the assignment step for the data association problem.

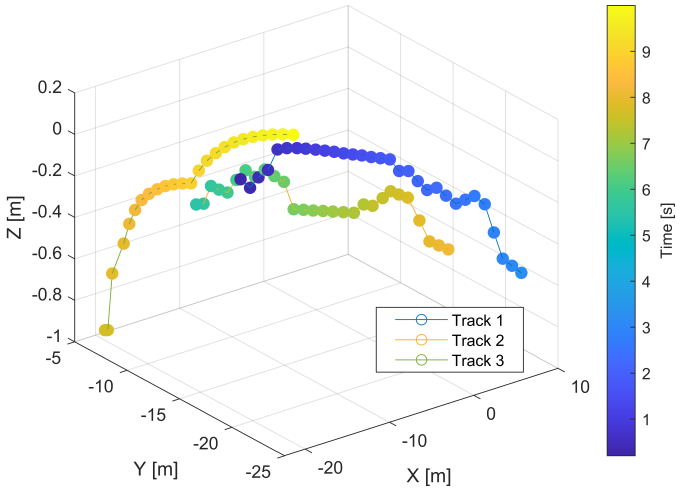


Fig. 6: The output of the JPDAF approach, where three vehicles are detected from 100 point clouds of LiDAR along with their corresponding appearance times.

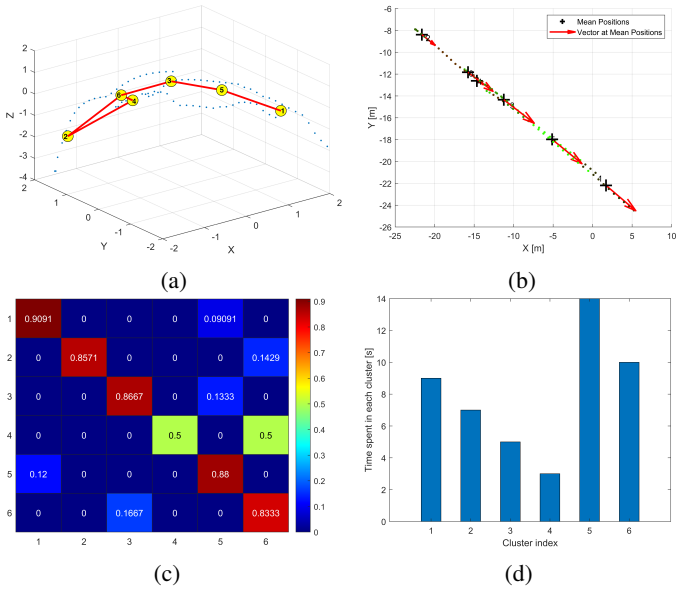
only considered the longest tracks, which appeared for long intervals of time, i.e., we confirmed those tracks that were detected at least 3 times in the last 3 updates, otherwise we deleted the tentative tracks, as we are interested in learning the time series observations for our proposed G-DBN. This way, we performed the JPDAF on 3D LiDAR point cloud observations $\tilde{Z}_{t=1,2,\ldots,t}$ at each time instant, and obtained three longest tracks i.e., track 1 ($\tilde{Z}_t^{(1)}$), track 2 ($\tilde{Z}_t^{(2)}$), and track 3 ($\tilde{Z}_t^{(3)}$) inside the region as depicted in Fig. 6. The 3D visualization of these dynamic tracks shows the significance of this filter in tracking multiple vehicles in the 3D space, where the z-coordinate shows the height of the tracked objects. The ranges of all the axes depend on the limits defined in our region of interest. Further, the figure clearly shows that the first longest track i.e., track 1 appears for 4 seconds (from 1 to 4). Similarly, track 2 and track 3 are the second and third longest tracks appearing for 3 seconds i.e., 8 to 10 and 6 to 8 seconds, respectively.

B. Offline Training Phase

1) Unmotivated Kalman Filter (UMKF): We utilized an UMKF, or null force Kalman filter, which assumes that the tracked vehicles are not moving, i.e., they are not affected by any external force, having unmodified speed with respect to the previous time instants. The filter allows the extraction of the filtered positions and velocities of the dynamic tracks, by performing a linear relationship as:

$$Z_t^{(n)} = HX_t^{(n)} + n_t \tag{1}$$

where H represents the identity matrix and n_t represents zero-mean Gaussian distribution with covariance $\tilde{R}$. Thus, the GSs $\tilde{X}_t^{(n)}$ containing the states and their derivatives are obtained at the G-DBN's continuous level for each time instant [27], [28].



(a) (b) (c) (d)

Fig. 7: GNG clustering outputs: (a) Clustering of the GSs, (b) Mean (centroid) and number of points each cluster contains, (c) State transition matrix containing transitioning probabilities for each cluster, (d) Maximum time spent in each cluster.

Here, we consider the first-order derivatives of the states as:

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

2) Growing Neural Network (GNG) Clustering: Clustering, often regarded as the primary unsupervised learning, can reveal complex data patterns and identify similarities within complex structures. The GNG process is based on a self-organizing neural network that can learn the structure of incoming data incrementally, dynamically, and adaptively without knowing the number of clusters. It adds additional units during training, resulting in a network that reflects the data's topology, where each node represents a neuron that mapped input data. We preferred this learning algorithm among others because it facilitates an adaptive growth structure, is suitable for our 3D dynamic environment, does not require a priory fixed number of clusters, and allows for a better adaptation to the new observations without retraining the model. The GSs (output of UMKF) containing a 6D vector of positions and velocities of the vehicles are given as input to the GNG algorithm [29], [30]. During this training process, we carefully defined certain key parameters inside the GNG to obtain meaningful clusters and optimize network learning and growth. The growth rate is chosen to be 50 which controls the rate of adding new nodes to the network. The alpha and beta parameters are defined as the weights of velocities and positions, respectively. We set alpha to 0.99 and beta to 0.96. Assigning higher weights to alpha results in the robust transition matrices which consequently gives lower abnormalities at the superstate level. By defining these parameters, we obtain 6 clusters on the GSs as shown in Fig. 7(a). Each cluster represents a discrete variable (i.e., discrete level of G-DBN). The mean (center) of each cluster, the number of points each cluster contains, and the arrow direction associated with the corresponding mean (the velocity) are shown in Fig. 7(b). The transition probabilities of moving from one cluster to another cluster over time are shown in Fig. 7(c). This transition matrix (TM) shows that the probabilities of staying in the same cluster, represented by the diagonal elements, are high, whereas the probabilities of changing clusters, represented by off-diagonal elements, are low. During this learning period, the maximum time spent in each cluster, which is important for the generation of temporal TMs, is also visualized in Fig. 7(d). We extracted all these features from each M cluster, referring to them as an output of GNG clustering, and formed a complete vocabulary, containing a set of temporal TMs for each time instant, an overall TM, the maximum time spent in each cluster, the cluster's mean (centroid), and the cluster's covariance matrices.

C. Online Testing Phase

During the online testing phase, the 3D LiDAR point clouds are initially pre-processed to extract the required features inside the region of interest, the JPDAF is then utilized to get the vehicle's tracks. Similarly to what has been done during the training phase, the 3D tracks of the vehicles are obtained to test the trained model. Here, for the testing phase, we utilize a high-dimensional version of MJPF [32]. It uses the probabilistic information learned by the G-DBN to infer future positions and identify the abnormalities at the hidden levels in both normal and abnormal scenarios.

1) Abnormality Measurements: Here, the abnormality is defined as the difference in measurement between the predictions performed by the learned G-DBN model and the future observations coming from the LiDAR sensor. We have observed the abnormalities at the hidden levels, i.e., the abnormality at the continuous level (CLA) and the abnormality at the discrete level (KLDA). The CLA measures the difference between the LiDAR's future observations $\lambda(\tilde{X}_{t+1}^{(n)})$ and the predicted GSs $\pi(\tilde{X}_{t+1}^{(n)})$, where $\pi(\tilde{X}_{t+1}^{(n)})$ and $\lambda(\tilde{X}_{t+1}^{(n)})$ are defined as the forward and backward belief propagation parameters of the G-DBN, respectively, as shown in Fig. 3. The CLA abnormality is calculated by using Bhattacharyya distance [31], [32]. The KLDA instead computes the difference between the predicted transition messages at superstate level $\pi(\tilde{S}_{t+1}^{(n)})$ and the likelihood messages $\lambda(\tilde{S}_{t+1}^{(n)})$. The KLDA abnormality is calculated by using Kullback-Leibler divergence [33].

2) Testing Configurations: The abnormality detection performance of our proposed model is examined under normal and abnormal situations, which are described below.

a) Normal Situation: In the normal case, we employ the same set of tracks for testing purposes to evaluate both the prediction and abnormality detection capabilities of the learned G-DBN model. This situation allows us to analyze the behavior of the model based on the 3D point cloud data on which it is trained. This analysis also facilitates a way to compare the performance of our model when new abnormal situations occur.

b) Anomalous Point Clouds (New Situation): For the abnormal situation, our model is tested on the anomalous 3D point clouds. This situation appears when the LiDAR struggles to perform well under adverse and harsh weather conditions. We consider this situation when the laser beams emitted by the LiDAR are scattered and absorbed by water droplets in foggy weather. This scattering phenomenon lowers the resolution and range of the sensor, which results in yielding sparser, less detailed, and missed point cloud observations. This new situation allows us to analyze the performance and behavior of our proposed model when anomalous point clouds appear in the scene.

V. Simulation Results & Discussion

This section discusses the performance of our proposed G-DBN model. In particular, we investigate our model's self-aware ability in terms of predictions and abnormalities detection from 3D LiDAR point clouds. The proposed G-DBN model, which is trained to create the vocabularies, is performing good predictions on all the learned tracks under normal situations (see Fig. 8). The time series observations coming from the LiDAR point clouds match the predictions performed by the model at each time instant (the LiDAR collects 10 point cloud frames per second). These correct predictions for the vehicles' future positions show the significance and learning ability of our proposed model when considering the 3D point clouds.

Fig. 9 shows the predicted and observed curves for three different tracks which were not observed by the model before. The figures show that some of the LiDAR frames become missed or less detailed, because of the model not performing correct predictions. We further analyzed that this anomalous situation (new situation) occurred for a very short period in which some point cloud frames became anomalous. However,

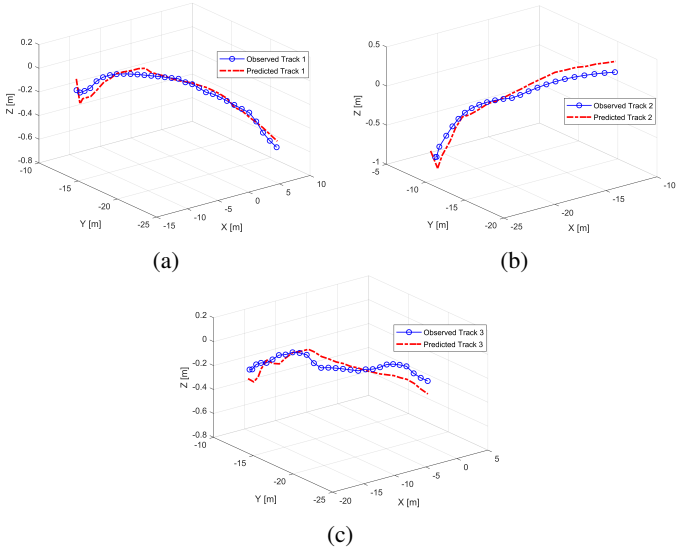


Fig. 8: Observed and predicted tracks in a normal situation: (a) Track 1, (b) Track 2, (c) Track 3.

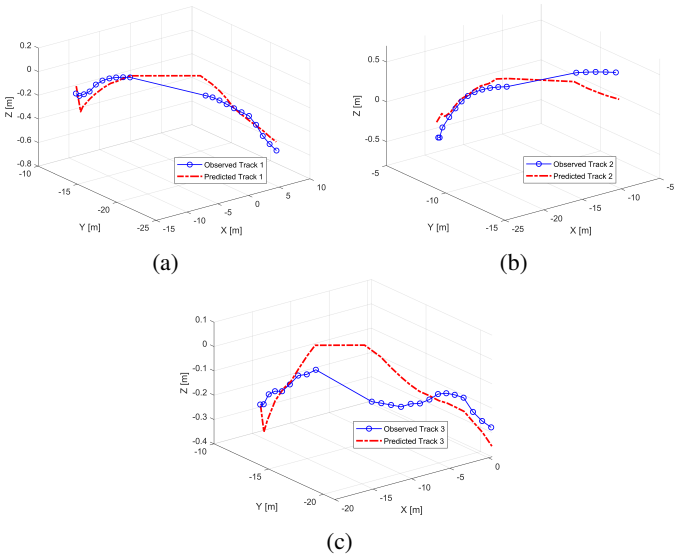


Fig. 9: Observed and predicted tracks in an abnormal situation: (a) Track 1, (b) Track 2, (c) Track 3.

as soon as the clear and detailed LiDAR point cloud observations appear, the model gives the correct predictions for all the tracks. We have also analyzed the abnormal performance of our model at hidden levels to test if our model is coherent with the new situations. For all the testing configurations, the abnormalities are shown in Fig. 10. The normalized CLA and KLDA abnormalities in the normal case for all three 3 tracks are shown at the top of the figure (first six subplots), while the last six subplots show CLA and KLDA abnormalities when the new situation appears. The thresholds of 0.8 and 0.2 are set for CLA and KLDA, respectively. For the normal case, the higher CLA values appearing in all the 3 tracks are due to the vehicle's continuous maneuvering behavior in the 3D space.

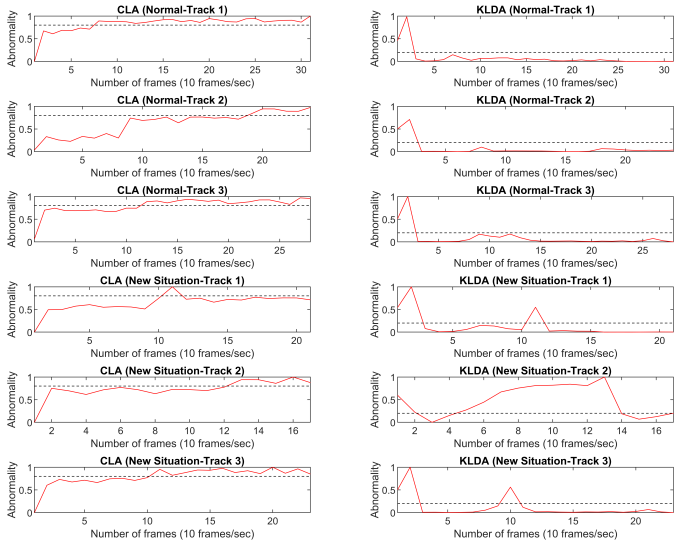


Fig. 10: CLA and KLDA abnormalities in all the tracks for both normal and abnormal situations.

This can also be seen from the normal 3D track observation, as shown in Fig. 8. This type of behavior also depends on the direction of the laser beam directed by the LiDAR sensor, from which the point cloud frame is captured. Hence, the larger maneuvering of the vehicles inside the scenario leads to higher CLA abnormalities. In this normal situation, KLDA values are lower than the defined thresholds for all three tracks which show the significant performance of the learned model. The initial peaks in all the KLDAs are due to the far initial guess, which is common in Bayesian filtering. The other peaks appear because of the transitions learned between the clustered GSs, which are less deviated from the predicted transitions of the HD-MJPF. In the abnormal case (new situation), the CLA and KLDA for all the tracks are higher for some observed point cloud frames, compared to the normal situation, as illustrated in Fig. 10. This is due to the new situation when the sensor gives anomalous point clouds. Here, the point cloud frames from 10 to 12 (for track 1), 5 to 13 (for track 2), and 9 to 11 (for track 3) are anomalous.

VI. Conclusion & Future Work

In this paper, we have proposed a data-driven approach based on a G-DBN which is learned from the high-dimensional 3D LiDAR point clouds. This novel approach is useful for self-aware autonomous systems, where artificial agents learn the 3D environment. To learn this G-DBN model, initially, we employed the JPDAF filter to solve the data association problem from the dense point clouds and extract the 3D dynamic tracks of the vehicles. Then, we perform an unsupervised clustering on all the GSs containing positions and velocities (a 6D vector) by considering the GNG technique, which represents a trained model from the 3D LiDAR point clouds. In the online testing phase, the HD-MJPF utilizes the G-DBN's probabilistic information to predict the positions of multiple vehicles and to detect abnormalities at hidden levels. The results show that our proposed G-DBN model performs well, although not considering a large number of training point cloud sequences, and this also gives us low abnormalities at superstate levels due to the effective learning. Moreover, it also shows a better understanding of both normal and abnormal real-world situations, which is important for self-aware autonomous systems.

The work could be extended by learning an interactive generalized DBN model to perform an interaction between the multiple vehicles separately, and using their corresponding mmWave signals received by the base station [33]. This interaction will be very helpful for self-aware agents to predict future blockages in high-frequency communication systems. The dataset used in this work is also suitable for blockage identification, blockage prediction, and beam prediction. Hence, considering LiDAR with mmWave signals to jointly learn the generative data-driven approaches will enhance the reliability and reduce the latency in the framework of integrated sensing and communication. Furthermore, for the interactive G-DBN, future work may also focus on associating the extracted tracks of each vehicle to the model it has learned, which could be challenging for self-aware agents in a dynamic environment.

Acknowledgment

This research was partially funded by the European Union's Horizon Europe research and innovation programme under the Grant Agreement No. 10112113, and by the European Union - NextGenerationEU and by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.5, project "RAISE - Robotics and AI for Socio-economic Empowerment" (ECS00000035).

References

[1] M. Yu, "Construction of Regional Intelligent Transportation System in Smart City Road Network via 5G Network," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 2, pp. 2208-2216, 2023.

[2] A. Najafi, T., A. Affanni, R. Rinaldo, and P. Zontone, "Driver attention assessment using physiological measures from EEG, ECG, and EDA signals," Sensors, vol. 23, no. 4, 2023.

[3] P. Zontone, A. Affanni, A. Piras, and R. Rinaldo, "Stress recognition in a simulated city environment using Skin Potential Response (SPR) signals," 2021 IEEE International Workshop on Metrology for Automotive (MetroAutomotive), pp. 135–140, 2021.

[4] P. Zontone, A. Affanni, R. Bernardini, A. Piras, and R. Rinaldo, "Low-complexity classification algorithm to identify drivers' stress using electrodermal activity (EDA) measurements," 2020 Springer International Publishing [Biomedical Engineering and Computational Intelligence: Proceedings of The World Thematic Conference—Biomedical Engineering and Computational Intelligence (BIOCOM), pp. 25-33, 2018].

[5] S. Memon, et al., "The design of wireless portable electrocardiograph monitoring system based on ZigBee," EAI Endorsed Transactions on Scalable Information Systems, vol. 7, no. 28, pp. e6–e6, 2020.

[6] L. Zhao, H. Zhou, X. Zhu, X. Song, H. Li and W. Tao, "LIF-Seg: LiDAR and Camera Image Fusion for 3D LiDAR Semantic Segmentation," IEEE Transactions on Multimedia, vol. 26, pp. 1158–1168, 2024.

[7] L.-H. Wen and K.-H. Jo, "Fast and accurate 3D object detection for LiDAR-camera-based autonomous vehicles using one shared voxelbased backbone," IEEE Access, vol. 9, pp. 22080–22089, 2021.

[8] Y. Sun, J. Li, Y. Wang, X. Xu, X. Yang and Z. Sun, "ATOP: An Attention-to-Optimization Approach for Automatic LiDAR-Camera Calibration via Cross-Modal Object Matching," IEEE Transactions on Intelligent Vehicles, vol. 8, no. 1, pp. 696–708, 2023.

[9] D. Campo, et al., "Learning Probabilistic Awareness Models for Detecting Abnormalities in Vehicle Motions," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 3, pp. 1308–1320, 2020.

[10] Gao, L., Song, J., Liu, X., Shao, J., Liu, J. and Shao, J., "Learning in high-dimensional multimedia data: the state of the art," Multimedia Systems, vol. 23, pp. 303–313, 2017.

[11] S. Wu, C. Chakrabarti and A. Alkhateeb, "LiDAR-Aided Mobile Blockage Prediction in Real-World Millimeter Wave Systems," 2022 IEEE Wireless Communications and Networking Conference (WCNC), pp. 2631–2636, 2022.

[12] G. Wang, X. Wu, S. Jiang, Z. Liu and H. Wang, "Efficient 3D Deep LiDAR Odometry," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 5, pp. 5749–5765, 2023.

[13] J. Deng, W. Zhou, Y. Zhang and H. Li, "From Multi-View to Hollow-3D: Hallucinated Hollow-3D R-CNN for 3D Object Detection," IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 12, pp. 4722–4734, 2021.

[14] S. Jiang, G. Charan and A. Alkhateeb, "LiDAR Aided Future Beam Prediction in Real-World Millimeter Wave V2I Communications," IEEE Wireless Communications Letters, vol. 12, no. 2, pp. 212–216, 2023.

[15] H. Yin, Y. Lu, J. Lin, M. Schratter and D. Watzenig, "Multi-Object Tracking with Object Candidate Fusion for Camera and LiDAR Data," 2023 26th IEEE International Conference on Intelligent Transportation Systems (ITSC), pp. 2965–2970, 2023.

[16] S. Pang, D. Morris and H. Radha, "Fast-CLOCs: Fast Camera-LiDAR Object Candidates Fusion for 3D Object Detection," 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 3747-3756, 2022.

[17] Z. Chiyu, et al., "Detecting the Anomalies in LiDAR Pointcloud," ArXiv abs/2308.00187, 2023.

[18] K. Montalban, C. Reymann, D. Atchuthan, P.-E. Dupouy, N. Riviere, and S. Lacroix, "A quantitative analysis of point clouds from automotive lidars exposed to artificial rain and fog," Atmosphere, vol. 12, no. 6, p. 738, 2021.

[19] D. Bogdoll, M. Nitsche and J. Zollner, "Anomaly Detection in Autonomous Driving: A Survey," 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 4487–4498, 2022.

[20] A. Krayani, G. Barabino, L. Marcenaro and C. Regazzoni, "Integrated Sensing and Communication for Joint GPS Spoofing and Jamming Detection in Vehicular V2X Networks," 2023 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–7, 2023.

[21] S. Nozari, A. Krayani, P. Marin, L. Marcenaro, D. Martin and C. Regazzoni, "Adapting Exploratory Behaviour in Active Inference for Autonomous Driving," ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5, 2023.

[22] A.S. Alemaw, G. Slavic, H. Iqbal, L. Marcenaro, D.M. Gomez and C. Regazzoni, "A Data-Driven Approach for the Localization of Interacting Agents via a Multi-Modal Dynamic Bayesian Network Framework," 2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–8, 2022.

[23] A. Alkhateeb, et al., "DeepSense 6G: A Large-Scale Real-World Multi-Modal Sensing and Communication Dataset," IEEE Communications Magazine, vol. 61, no. 9, pp. 122–128, 2023.

[24] D. Smith and S. Singh, "Approaches to Multisensor Data Fusion in Target Tracking: A Survey," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 12, pp. 1696–1710, 2006.

[25] A. Mazzù, S. Chiappino, L. Marcenaro and C. S. Regazzoni, "A track-before-detect algorithm using joint probabilistic data association filter and interacting multiple models," 2014 IEEE International Conference on Image Processing (ICIP), pp. 4947–4951, 2014.

[26] I.J. Cox, "A review of statistical data association techniques for motion correspondence," International Journal of Computer Vision, vol. 10, no. 1, pp. 53–66, 1993.

[27] G. Slavic, P. M. Plaza, L. Marcenaro, D. M. Gómez and C. Regazzoni, "Simultaneous Localization and Anomaly Detection from First-Person Video Data through a Coupled Dynamic Bayesian Network Model," 2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–8, 2022.

[28] S. Nozari, A. Krayani, L. Marcenaro, D. Martin and C. Regazzoni, "Incremental Learning through Probabilistic Behavior Prediction," 2022 30th European Signal Processing Conference (EUSIPCO), pp. 1502–1506, 2022.

[29] H. Iqbal, D. Campo, M. Baydoun, L. Marcenaro, D.M. Gomez, and C. Regazzoni, "Clustering optimization for abnormality detection in semi-autonomous systems," 1st International Workshop on Multimodal Understanding and Learning for Embodied Applications, pp. 33-–41, 2019.

[30] G. Slavic, A.S. Alemaw, L. Marcenaro, D.M Gomez and C. Regazzoni, "A Kalman Variational Autoencoder Model Assisted by Odometric Clustering for Video Frame Prediction and Anomaly Detection," IEEE Transactions on Image Processing, vol. 32, pp. 415–429, 2023.

[31] A. Krayani, A. S. Alam, L. Marcenaro, A. Nallanathan and C. Regazzoni, "Automatic Jamming Signal Classification in Cognitive UAV Radios," IEEE Transactions on Vehicular Technology, vol. 71, no. 12, pp. 12972-12988, Dec. 2022.

[32] H. Iqbal, D. Campo, P. Marin-Plaza, L. Marcenaro, D. M. Gómez and C. Regazzoni, "Modeling Perception in Autonomous Vehicles via 3D Convolutional Representations on LiDAR," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 9, pp. 14608–14619, 2022.

[33] A. Krayani, M. Baydoun, L. Marcenaro, A.S. Alam and C. Regazzoni, "Self-Learning Bayesian Generative Models for Jammer Detection in Cognitive-UAV-Radios," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, pp. 1–7, 2020.