INCREMENTAL LEARNING OF ABNORMALITIES IN AUTONOMOUS SYSTEMS

Hassan Zaal*, Hafsa Iqbal*, Damian Campo, Lucio Marcenaro, and Carlo S. Regazzoni

Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN),
University of Genova, Italy

{hassan.zaal, hafsa.iqbal, damian.campo}@ginevra.dibe.unige.it
{carlo.regazzoni, lucio.marcenaro}@unige.it

Abstract

In autonomous systems, self-awareness capabilities are useful to allow artificial agents to detect abnormal situations based on previous experiences. This paper presents a method that facilitates the incremental learning of new models by an agent. Available learned models can dynamically generate probabilistic predictions as well as evaluate their mismatch from current observations. Observed mismatches are grouped through an unsupervised learning strategy into different classes, each of them corresponding to a dynamic model in a given region of the state space. Such clusters define switching Dynamic Bayesian Networks (DBNs) employed for predicting future instances and detect anomalies. Inferences generated by several DBNs that use different sensorial data are compared quantitatively. For testing the proposed approach, it is considered the multi-sensorial data generated by a robot performing various tasks in a controlled environment and a real autonomous vehicle moving at a University Campus.

1. Introduction

Dynamical models that allow autonomous systems to predict and make decisions are fundamental for the next generation of smart devices and services. The combined use of machine learning and signal processing techniques has demonstrated to facilitate the designing of predictive models that recognize sensory information as part of a given context where multiple entities interact [1, 2, 3]. A self-aware autonomous system is characterized by its ability to recognize own states, possible actions and their effects on both, the environment (system's surroundings) and own body [4]. Such characteristics can be seen as a result of learning different predictive models based on multisensory data, facilitating the understanding of a phenomenon from various viewpoints. Currently, self-awareness is an essential topic in the field of robotics [5]. Autonomous mobile robots are designed to support humans by cooperating with them for accomplishing a given task. Accordingly, the next generation of smart robots is required to be able to learn from their own experiences and interact dynamically with human users. Current advancements in robotics are focused on increasing the robots' autonomy by employing data fusion strategies for predicting future instances of multisensory data. Recent research developments in autonomous systems also include modern transfer learning techniques for mapping learned knowledge among diverse robot architectures [6]. Multi-agent Transfer learning and learning in dynamic environments have recently examined [7]. The aforementioned research raises the necessity of learning new experiences incrementally in an automatic fashion.

Developing a system that can learn incrementally from the environment is the most difficult task, which falls in the category of computational intelligence. Its difficulty lies on the lack of necessary information to handle unpredictable situations at the design time. Applications of such task include autonomous robots and self-driving vehicles [8, 9].

The world is dynamic and objects can move, appear or disappear. Every situation that an autonomous system may face, cannot be predicted and there is no complete model which can train the system for every predictable situation. That's why it has an incomplete knowledge of its own way of working. In a lot of existing approaches, this is done offline by trying to take into account a large amount of possible situations. These systems have impressive results in particular environments and situations; nonetheless, they can be totally disoriented in other situations[3, 4, 10]. Machine learning is one of the promising research areas that try to solve this kind of problems.

The identification of new situations by using an initial training the model and the inclusion of such such novelties

---

*Both authors contributed equally to this research.

as part of the system's knowledge, i.e., updating the initial training model, it is defined as incremental learning. Many machine learning algorithms support incremental learning, which facilitates a faster classification and forecasting times [11].This paper introduces the application of machine learning techniques for incremental learning. The adaptation of new situations learned incrementally and the internal representation should be updated by keeping the previous knowledge.

The most important researched parameter of incremental learning is the complexity of the system [12]. If the system is complex at the initial training phase, it cannot be implemented in real time devices because its complexity will increase as it experiences new situations. Reduction in complexity is understood by a concept of *free energy principle*. The generic relation between free energy, complexity, and accuracy of the model is shown in Eq. (1),

$$Free\ energy = complexity - accuracy, \qquad (1)$$

this shows that by minimizing the free energy, the accuracy of model increases and its complexity decreases, leading to the optimization of the model.

This paper presents a method that analyzes internal and external data sources from a robot and autonomous vehicle while performing specific tasks in a controlled environment. Abnormalities from each module are obtained and the accuracy for each of them is calculated. Our approach shows how anomalies can be learned incrementally from an autonomous system and used for prediction and decision-making purposes.

Detailed contribution. The novelties of this paper include:
i) the detection of novel situations,
ii) the comparison of performances for detecting abnormalities between different modules of a moving robot,
iii) the usage of an anomaly measurement based on the Hellinger distance.
iv) the minimization of free energy principle,
v) the understanding of new situations by learning incrementally and extend the knowledge.

The rest of the paper is organized as follows: Section 1.1 describes the motivation and importance of our work. Section 2 explains the training and testing dataset from the robot and autonomous vehicle and the proposed methodology for incremental learning. Section 3 illustrates the experimental results of the proposed method for incremental learning and compare the performances of the different modules. Section 4 concludes the paper and presents future paths of the research.

1.1. Motivation

Learning from data distributions that could change over time is a challenging problem. A free energy principle

[13, 14] illustrates how self-organizing agents prevent learning to disorganize. This can be performed by minimizing the free energy, i.e, minimize the difference between the system's state and observation, which leads to improving the system model. The following subsections present the free energy principle, the motivation of using it and its importance for incremental learning.

1.1.1 Free energy Principle

The free energy [13, 14] is a function that represents the prediction error between sensory input and a given generative model, i.e., a model estimates the future states of sensory samples and explain their causes. Eq. (2) represents the free energy function, which is the expected energy under a density that links the system's state to the environment $\psi(\phi)$. $p(\mathbf{X}, \phi)$ is a generative density function, where $\mathbf{X}$ denotes the generalized state and ϕ refers to the causes of the sensory input. The generative density function factorizes into a likelihood $p(\mathbf{X}|\phi)$ and prior $p(\phi)$. The $< \cdot >_\psi$ denotes the expectation under density ψ.

$$F = - \int \psi(\phi) \ln \frac{p(\mathbf{X}, \phi)}{\psi(\phi)} d\phi$$
$$= - < \ln(p(\mathbf{X}, \phi)) >_\psi + < \ln(\psi(\phi)) >_\psi \qquad (2)$$

where $p(\mathbf{X}, \phi) = p(\mathbf{X}|\phi)p(\phi)$

The Free energy principle illustrates how the autonomous system tries to minimize the variations of free energy. This minimization can be performed in two ways:

- Modify the model representation of the system to optimize the bound.

- Apply actions, i.e., *energy*, to minimize a bound on surprise based on the deviations between sensory input and the generative model.

In this paper, we focus on modifying the model representation to minimize free energy. As we explained in the introduction, the minimization of free energy leads to the reduction of the system's complexity. Such reduction in complexity is obtained when we group our data in an optimal number of clusters and in return we have a small and useful vocabulary which makes our system less complex. Accordingly, optimization is the point where we get maximum accuracy for the detection of abnormalities by utilizing the minimum (optimal) number of clusters.

1.1.2 Incremental learning

Knowledge acquisition is essential for both humans and artificial machines. As data becomes available over time while facing changeable environments, algorithms that can

process and understand new concepts from such data are required. This leads to incremental learning concepts, allowing to acquire new knowledge while sustaining the previous one [15, 16]. Many machine learning researchers in different domains tried to tackle the problem of incrementally learning without forgetting previous knowledge [17].In this paper, we also tackle this problem by proposing an approach based on the free energy principle to incrementally learn without losing previous knowledge.

2. Proposed Methodology

The whole process is shown in Figure 1. According to this, the initial data is obtained using null-force filter to train the model. Clustering the data in the optimal number of groups, this optimization is obtained by the minimization of the free energy principle. After that detect whether the situation is new or already learned if abnormalities are detected then incremental learning performed on that abnormal data and learned it to extend the knowledge.
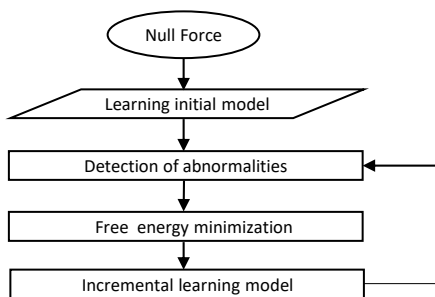

Figure 1. Model of autonomous adaptive system using incremental learning.

2.1. Experimental data-set

(a)   (b)

Figure 2. Autonomous vehicle "iCab" and LegoEV3 Robot used for collecting data-set.

Experiments are performed on the data-set obtained from autonomous vehicle iCAB and LegoEV3 robot as shown in Figure 2[4].

In the case of *iCAB* three different tasks perform on autonomous vehicle i.e., perimeter monitoring, u-turn and obstacle avoidance. Perimeter monitoring is used for training

purpose and the other two are used for testing and learned the abnormalities of these tasks.

In the case of *robot* two different experiments are conducted i.e., perimeter monitoring and obstacle avoidance. Three types of sensors are used for monitoring the robot's state, which in turn correspond to three modules of information. They are: *(i) Odometry*, consisting of the two encoders' data of the motors. *(ii) Sonar*, composed of the two distance measurements from ultrasonic sensors. *(iii) External camera*, which consists of the robot's positions extracted in each frame from a surveillance camera placed at the top of the scene as shown in Figure 4(a,b).

For the acquisition of the robot's location from an external camera, we perform the tracking of a big red circle placed on the robot's top, see figure 2(b) and consider the center as its position. For noise removal and adjustment of illumination, median filters [18] and gamma correction [19] techniques are utilized. Measurements from the three modules are aligned temporally, such that signals can be analyzed contemporaneously.
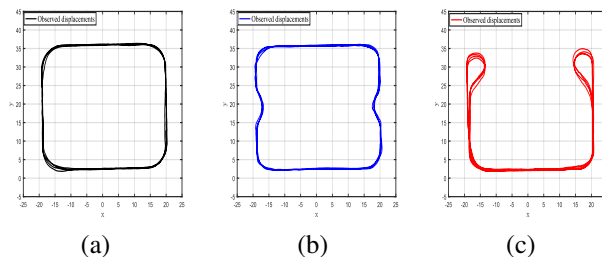

(a)   (b)   (c)

Figure 3. Three different action scenarios for iCAB data-set: (a) Perimeter monitoring under normal situation, (b,c) Obstacle avoidance and U turn respectively under abnormal situation.
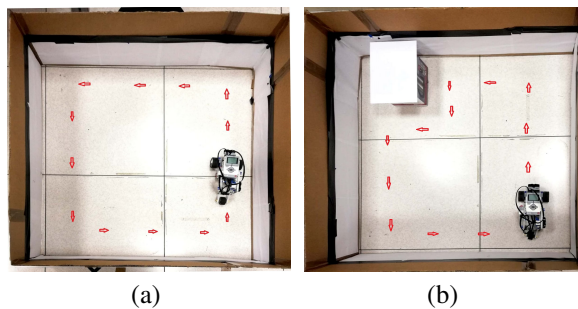

(a)   (b)

Figure 4. Two different scenarios for robot data-set (a) perimeter monitoring under normal situation, (b) perimeter monitoring in the presence of an abnormality due to an obstacle.

Perimeter monitoring (PM) task. The entity follows a rectangular path in the proposed environment as shown in Figure 3(a) and in Figure 4(a). The multisensory information collected from the PM task is used as training data from which predictive models are learned and embedded into the links of Bayesian network architecture. The PM task is con-

sidered as the initial training task from which abnormalities are detected and compared and then learned from these abnormalities.

Obstacle avoidance task. The entity performs perimeter monitoring until it encounters an obstacle, the abnormality is a standing pedestrian and obstacle in case of the robot between the perimeter maneuvering. When the entity encounters the obstacle, it performs an avoidance maneuver by surrounding the obstacle and then continuing its rectangular path as shown in Figure 3(b) and in Figure 4(b). Multisensory information collected from this task is employed as testing data to detect anomalies with respect to the PM task.

U-turn. The autonomous vehicle follows perimeter monitoring and performs u-turn when encountered with an obstacle and then continues perimeter monitoring in the opposite direction as shown in Figure 3(c).

2.2. Generalized states

Let Z_k^m be the measurements of the module m in a time instant k. Additionally, let X_k^m be the states associated to the measurements Z_k^m; such that $Z_k^m = g(X_k^m) + \omega_k$. $g(\cdot)$ is a function that maps states into observations and ω_k encodes the sensors' noise. As explained in [20, 21], states' time derivatives allow predictive models to make inferences dynamically even when some unknown quantities are constant. The generalized states of a module m is defined as:

$$\boldsymbol{X}_k^m = [X_k^m \ \dot{X}_k^m \ \ddot{X}_k^m \cdots X_k^{(L)m}]^\mathsf{T}, \qquad (3)$$

where (L) indexes the L-th time derivative of the vector state. The l-th time derivative in a given module m at the time k is approximated as:

$$X_k^{(l)m} = \frac{X_k^{(l-1)m} - X_{k-1}^{(l-1)m}}{\Delta k}, \qquad (4)$$

where $X_k^{(0)m} = X_k^m$ and Δk is the uniform sampling time for all multisensory data.

2.3. Vocabulary and label generation

A clustering approach based on the application of several GNGs facilitates to group generalized states into a set of regions that encode the dynamics of observed experiences. GNG presents some advantages in comparison with other clustering algorithms like K-means or SOM. In GNG, the learning is continuous, and the addition/removal of nodes is automated [22].

For each module, a total number of $L + 1$ GNGs are trained, where L is the maximum time derivative order of states. Each GNG encodes data related to a specific derivative of order (l), where $l \in [0, L]$. Such GNG training strategy guarantees the compactness in clusters that carry the same type of information concerning time derivative orders.

The output of each GNG consists of a set of nodes that encode constant behaviors for a specific module's time derivative order. Nodes inside each GNG can be seen as a set of *letters* containing the main behaviors of generalized states. The set of nodes encoding derivatives of order l-th in a module m is defined as follows:

$$V_l^m = \{\bar{X}_1^{(l),m}, \bar{X}_2^{(l),m}, \cdots \bar{X}_{N_l^m}^{(l),m}\}, \qquad (5)$$

where N_l^m is the number of nodes in the GNG associated to the l-th order derivative of data in module m. The node n is defined as the centroid values $\bar{X}_n^{(l),m}$, which have the same dimension and form of states X_k^m. Such centroids are the codifications of state time derivatives in each module. Accordingly, V_l^m can be seen a *vocabulary* which contains a set codified versions, i.e., nodes (centroids), of the state time derivative of order l-th for the module m.

By considering the possible combinations among centroids (letters) containing different derivative orders in a module m, it is possible to obtain a set of *words*, which define generalized states in an entirely semantic way. Such words are defined as:

$$W^m = \{\varphi_m, \dot{\varphi}_m, \cdots \varphi_m^{(L)}\}, \qquad (6)$$

where $\varphi_m^{(l)} \in V_l^m$. W^m contains all possible combinations of discrete generalized states. W^m is a high-level hierarchy variable that explains the module states from a semantic viewpoint.

2.4. State estimation

The creation of modules' vocabularies, see equation (5); and words, see equation (6); is performed based on measurements belonging to training data, i.e., previously observed data. Accordingly, by observing how such measurements activate series of words/letters through time, it is possible to obtain transition models that facilitate the estimation of future words, i.e., $p(W_{k+1}^m|W_k^m, t_k)$; and letters, i.e., $p(V_{l,k+1}^m|V_{l,k}^m, t_k)$, depending on the time t_k spent in the current word W_k^m.

Estimations of future letters carry information about the dynamics of generalized states $\boldsymbol{X}^m$. In this paper, due to the high sample rate of multisensory data, only the first order time derivative of states is considered, such that $\boldsymbol{X}_k^m = [X_k^m \ \dot{X}_k^m]^\mathsf{T}$. For such a case, it is possible to consider models of constant velocity for tracking generalized states' dynamics, such that:

$$\boldsymbol{X}_{k+1}^m = A\boldsymbol{X}_k^m + BU_k + w_k, \qquad (7)$$

where,

$$A = \begin{bmatrix} I_j & 0_{j,j} \\ 0_{j,j} & 0_{j,j} \end{bmatrix} ; \ B = \begin{bmatrix} 0_{j,j} \\ I_j\Delta k \end{bmatrix}.$$

The variable j indexes the number of states in the module in question. I_j is an identity matrix of dimension j. $0_{j,j}$

is a zero $j \times j$ matrix. $w_k \sim \mathcal{N}(0, \sigma)$, encodes the model noise. U_k is a control vector that defines the states' dynamics. $U_k = E(V_{1,k+1}^m | V_{1,k}^m, t_k)$, where $E(\cdot)$ is the expectation operator. As can be seen, higher hierarchy levels defines the dynamics of generalized states.

Proposed discrete and continuous information can be described as a DBN, see Figure 5. For inference purposes, A PF algorithm is employed for estimating future modules' discrete information, i.e., words and letters. To each particle in the PF, it is attached a KF that uses equation (7) as a dynamical model. The joint usage of continuous and discrete probabilities in a DBN hierarchy facilitates the inference of future states in each module by the MJPF [3].
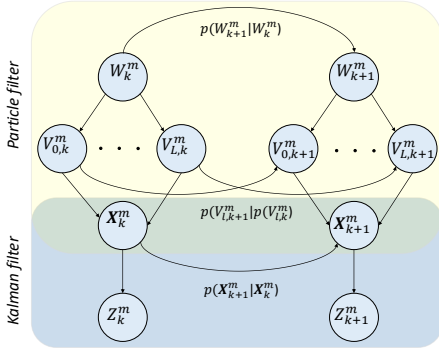

Figure 5. Proposed DBN

2.5. Abnormality measurement

As an abnormality measurement, it is used the Hellinger distance [23] between predicted generalized states, i.e., $p(\mathbf{X}_k^m | \mathbf{X}_{k-1}^m)$ and measurement evidence, i.e., $p(Z_k | \mathbf{X}_k^m)$, such that:

$$\theta_k^m = \sqrt{1 - \lambda_k^m}, \tag{8}$$

where λ_k^m is defined as the Bhattacharyya coefficient [24], such that:

$$\lambda_k^m = \int \sqrt{p(\mathbf{X}_k^m | \mathbf{X}_{k-1}^m) p(Z_k | \mathbf{X}_k^m)} \, d\mathbf{X}_k^m. \tag{9}$$

As shown in equation (8), the proposed abnormal measurement is defined for each module. Such consideration makes possible to compare abnormal behaviors from multiple sensory perspectives. The variable $\theta_k^m \in [0, 1]$, where values close to 0 indicate that measurements match with predictions; whereas values close to 1 reveal the presence of an abnormality.

2.6. Incremental update of vocabulary

An initial DBN model (we call it a "null force" model) is learned by using an unmotivated Kalman filter (UMKF). In the UMKF, the model predicts that the system will stay in the same state in the next time instance, while, the next

observation shows a change in the system's state. By using this way, a generative model will be learned. Figure 6 shows these deviations.

Figure 7 shows the surprise (deviations) between the sensory samples from the generative model, which is generated from density distribution over perimeter monitoring data and sensory input, which is a pedestrian avoidance scenario. These deviations represent the free energy that we want to minimize. The abnormal signals are employed for learning new models incrementally in an automatic way to minimize free energy. A threshold is applied to use the high abnormal signals to generate a new DBN model with new vocabularies that are associated with high deviations as it is shown in figures 10 and 11. This threshold is the boundary of each node which is selected on the basis of the following equation,

$$\xi_{N_l} = E(d_{N_l}) + 3 * \sqrt{(V(d_{N_l}))} \tag{10}$$

where d_{N_l} is the data inside N_l node, E denotes the mean and V represents the variance. This threshold acts as decision boundary in abnormality measurement such that if the new observations fall inside that boundary it will be normal and we will not experience any abnormality but if it will fall outside the boundary then it will be considered as abnormal which gives high peaks. As mentioned before, both the initial DBN models for each system and each sensor module and the new DBN models are learned based-on free energy principle. In this case, the system tends to minimize the free energy by modifying the model representation of the system, i.e., adjusts the distribution that was learned from previous scenarios to also cover the new observations, which produces a new generative model that gives no abnormalities when it tested with previously seen scenarios. In this paper, for us minimization of the error between predictions and observations is equivalent to minimization the free energy, i.e., the free energy is minimized when the error tends to zero.

3. Experimental Results

This section explains the results obtained from the proposed methodology. First, we select the one module from the three modules of the robot having a higher performance by evaluating the abnormality signals. Then use that module for learning the new situations incrementally which are considered an abnormality with respect to the training model.

3.1. Evaluation of abnormality signals among different modules

As mentioned before, the PM data generated by the robot is employed for training probabilistic models that facilitate the prediction of the state of multisensory data. Predicting models are then used for estimating the states of multisensory information in case of abnormalities by introducing an
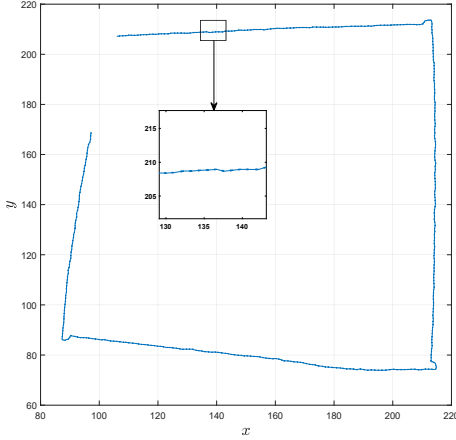
Figure 6. Free Energy (deviations) for initial DBN model in perimeter monitoring scenario. An initial generative model is learned
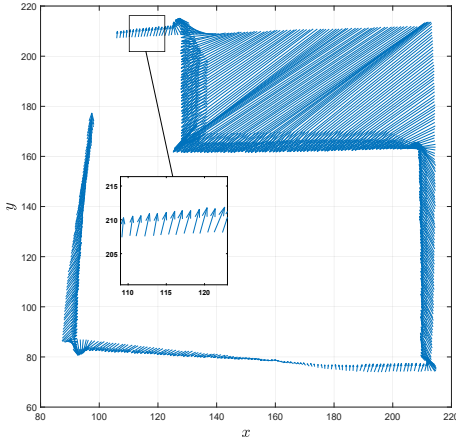

Figure 7. Free Energy (deviations) between the generative model (over perimeter monitoring data distribution) and measured input (obstacle avoidance data) in the robot environment

avoidance maneuver. Results obtained in each module are described and compared as follows:

The proposed method is applied to learn predictive models that describe the PM tasks in each module of the robot. Figure 8 shows the abnormality signals produced by our method when the robot performs the obstacle avoidance task. The red line represents the signal generated by the sonar module. It can be seen that in the initial time instances (from $k = 0$ to 35), as the front sensor is detecting the presence of the obstacle, an abnormality (before the actual abnormal motion) is obtained. In the case of the external camera and odometry modules, blue and black lines respectively, the initial period of the robot's task is similar to the training phase, which leads to low abnormality signals. When the robot starts the avoidance maneuver (from $k = 36$ to 170), the sonar module still detects an abnormal situation.

In addition, the signals from the visual and odometry module start detecting anomalies with respect to the regular PM. Additionally, while the robot finishes the avoidance maneuver (from $k = 100$ to 170), the sonar module perceives previously learned distances from the PM task. After finishing the avoidance maneuver, the measurements return to the normal case, except for some false alarms from the sonars' signals and odometry measurements. This can be explained by the noise produced by the sensors/actuators and the slight differences of curving with respect to the training data. The abnormalities detected by the sonar modules could lead to understanding the causality between modalities. We can expect that abnormalities in visual and odometry measurements are a consequence of previously anomalies detected by the sonar module.

Figure 9 presents the ROC curve for the testing cases mentioned previously. Overall, it can be noticed that the external camera's module is the most robust approach for detecting abnormalities. Even in the case of noisy data, the visual information offers high performance at detecting anomalies correctly. From the curves, it is possible to see that other modules of the robot present the lowest accuracy. This is due to the noise of the sensor data and in the case of the sonar module, its low performance is also due to the earlier detection of anomalies discussed previously. Because of the higher accuracy in the external camera, that module is used for incremental learning instead of odometry and sonar module.

3.2. Incremental learning results

Robot case. Figure 12 shows the abnormality signals produced by different DBN models. The first one is trained with perimeter monitoring task, while the second one is trained fully trained by the proposed method to learn incrementally. The red line represents the obtained abnormality measurements for a DBN model is trained with a single scenario (PM) and tested with obstacle avoidance (OA). It can be seen in the time instances (from k = 0 to 35) that there is abnormality as the robot in the testing task performs an avoidance maneuver. While the blue line represents the obtained abnormality measurements for a full DBN model (we call it a full model because it is incrementally trained with all scenarios) tested with obstacle avoidance (OA). The results show no abnormality in the time instances (from k = 0 to 35) where the robot performs an avoidance maneuver, which means the new DBN is successfully learned incrementally.

Vehicle case. Figure 13 presents a comparison for abnormality measurements between a DBN model has already trained with perimeter monitoring (PM) scenario and a new DBN model is incrementally learned based on the proposed approach. In this figure, the testing scenario is U-turn. Results in green color represents the measurements of test-
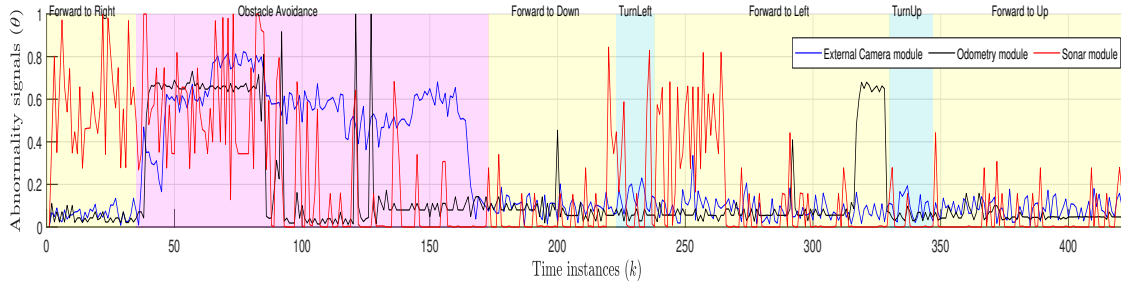
Figure 8. Abnormality Measurements in case of an obstacle avoidance for the robot
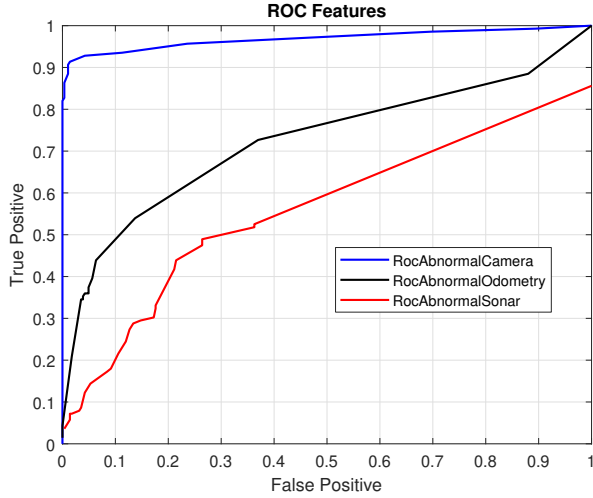

Figure 9. ROC curve

ing with U-turn where the DBN model is trained with PM. While the results in blue color represents the measurements of testing with U-turn where the DBN model is incrementally learned all scenarios.

4. Conclusion

We proposed an algorithm to employ abnormal signals for learning new models incrementally in an automatic way. The initial model for each DBN and the newly learned models are based on the minimization of the free energy principle. Results show that the new models of DBNs incrementally learned to produce no abnormalities when tested on sequences that generated deviations from which the models were created. As future work, the proposed approach will be employed for multi-agent transfer learning or to a new environment.

References

[1] X. Nie, J. Feng, J. Xing, S. Xiao, and S. Yan. Hierarchical contextual refinement networks for human pose estimation. *IEEE Transactions on Image Processing*, 28(2):924–936, 2019.

[2] Z. Zafar, S.H. Paplu, and K. Berns. Real-time recognition of extroversion-introversion trait in context of human-robot interaction. *Mechanisms and Machine Science*, 67:63–70, 2019.

[3] M Baydoun, D Campo, V Sanguineti, L Marcenaro, A Cavallaro, and C Regazzoni. Learning switching models for abnormality detection for autonomous driving. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 2606–2613. IEEE, 2018.

[4] Mahdyar Ravanbakhsh, Mohamad Baydoun, Damian Campo, Pablo Marin, David Martin, Lucio Marcenaro, and Carlo S Regazzoni. Learning multi-modal self-awareness models for autonomous vehicles from human driving. *arXiv preprint arXiv:1806.02609*, 2018.

[5] Mordechai Ben-Ari and Francesco Mondada. Robots and their applications. In *Elements of Robotics*, pages 1–20. Springer, 2018.

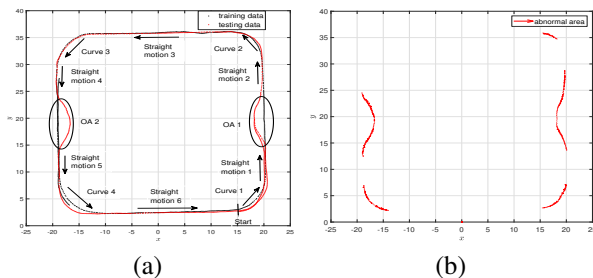[6] Mordechai Ben-Ari and Francesco Mondada. *Elements of Robotics*. Springer, 2018.

(a)   (b)

Figure 10. Observation of data related to the obstacle avoidance with parameter monitoring for iCAB and (b) extract the abnormal part to learn.
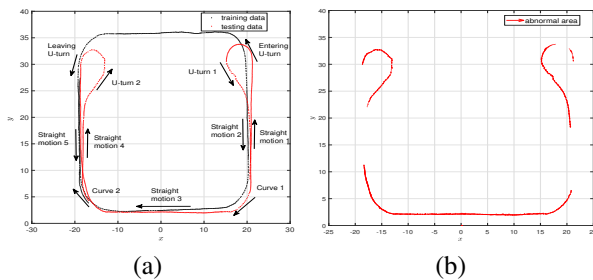

(a)   (b)

Figure 11. (a)Observation of data related to U-turn with perimeter monitoring for iCAB and (b) extract the abnormal part to learn.
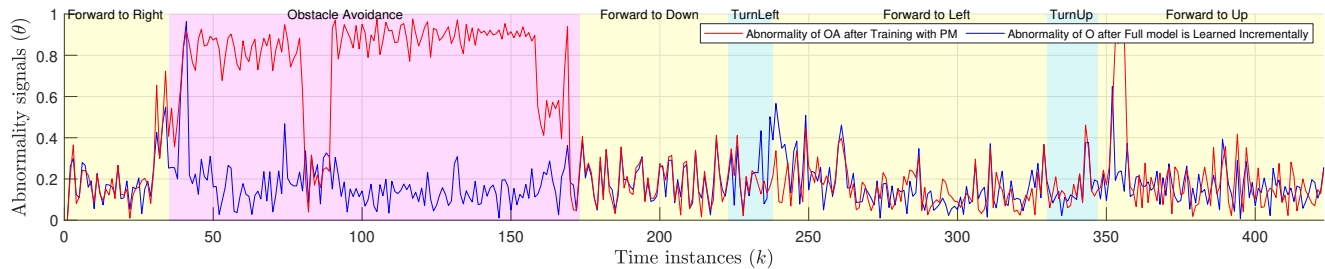
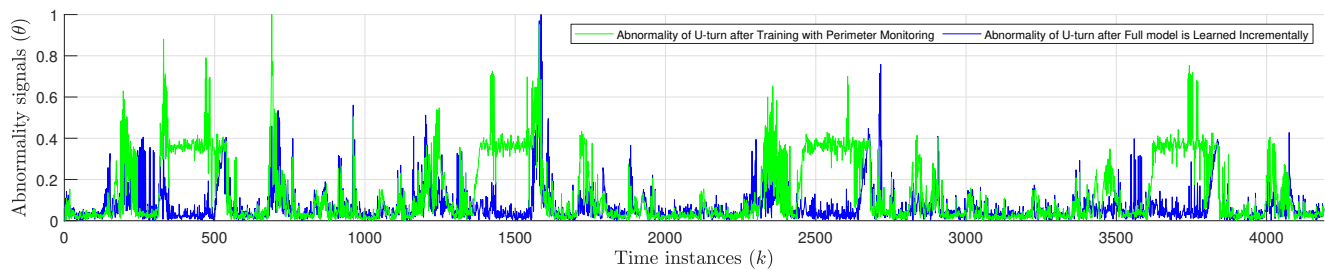Figure 12. Abnormality Measurements related to obstacle avoidance for the robot where we use video information

Figure 13. Abnormality Measurements related to the U-turn for iCAB

[7] Felipe Leno Da Silva and Anna Helena Reali Costa. A survey on transfer learning for multiagent reinforcement learning systems. *Journal of Artificial Intelligence Research*, 64:645–703, 2019.

[8] Reid G Simmons. Structured control for autonomous robots. *IEEE transactions on robotics and automation*, 10(1):34–43, 1994.

[9] Emilio Frazzoli, Munther A Dahleh, and Eric Feron. Real-time motion planning for agile autonomous vehicles. *Journal of guidance, control, and dynamics*, 25(1):116–129, 2002.

[10] Mohamad Baydoun, Mahdyar Ravanbakhsh, Damian Campo, Pablo Marin, David Martin, Lucio Marcenaro, Andrea Cavallaro, and Carlo S Regazzoni. A multi-perspective approach to anomaly detection for self-aware embodied agents. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6598–6602. IEEE, 2018.

[11] Patrick Reignier, Volker Hansen, and James L Crowley. Incremental supervised learning for mobile robot reactive control. *Robotics and autonomous systems*, 19(3-4):247–257, 1997.

[12] Arjan Gijsberts and Giorgio Metta. Incremental learning of robot dynamics using random features. In *2011 IEEE International Conference on Robotics and Automation*, pages 951–956. IEEE, 2011.

[13] Karl Friston, James Kilner, and Lee Harrison. A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3):70–87, 2006.

[14] Karl Friston. The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences*, 13(7):293–301, 2009.

[15] Ryan Elwell and Robi Polikar. Incremental learning of concept drift in nonstationary environments. *IEEE Transactions on Neural Networks*, 22(10):1517–1531, 2011.

[16] Zhi Wang, Chunlin Chen, Hanxiong Li, Daoyi Dong, and Tzyh Jong Tarn. Incremental reinforcement learning with prioritized sweeping for dynamic environments. *IEEE/ASME Transactions on Mechatronics*, 2019.

[17] Ian J Goodfellow, Mehdi Mirza, Da Xiao, Aaron Courville, and Yoshua Bengio. An empirical investigation of catastrophic forgetting in gradient-based neural networks. *arXiv preprint arXiv:1312.6211*, 2013.

[18] Jae S Lim. Two-dimensional signal and image processing. *Englewood Cliffs, NJ, Prentice Hall, 1990, 710 p.*, 1990.

[19] Hany Farid. Blind inverse gamma correction. *IEEE Transactions on Image Processing*, 10(10):1428–1433, 2001.

[20] Karl Friston and Stefan Kiebel. Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1521):1211–1221, 2009.

[21] B. Balaji and K. Friston. Bayesian state estimation using generalized coordinates. In *Signal Processing, Sensor Fusion, and Target Recognition XX*, volume 8050, page 80501Y, June 2011.

[22] Bernd Fritzke. A growing neural gas network learns topologies. In *Proceedings of the 7th International Conference on Neural Information Processing Systems*, NIPS'94, pages 625–632, Cambridge, MA, USA, 1994. MIT Press.

[23] Rodolfo Lourenzutti and Renato A. Krohling. The hellinger distance in multicriteria decision making: An illustration to the topsis and todim methods. *Expert Syst. Appl.*, 41(9):4414–4421, July 2014.

[24] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, 35:99–109, 1943.