A MULTI-PERSPECTIVE APPROACH TO ANOMALY DETECTION FOR SELF-AWARE EMBODIED AGENTS

M. Baydoun[1,2], M. Ravanbakhsh[1], D. Campo[1], P. Marin[3], D. Martin[3], **L. Marcenaro**[1], A. Cavallaro[2], C. Regazzoni[1,3]

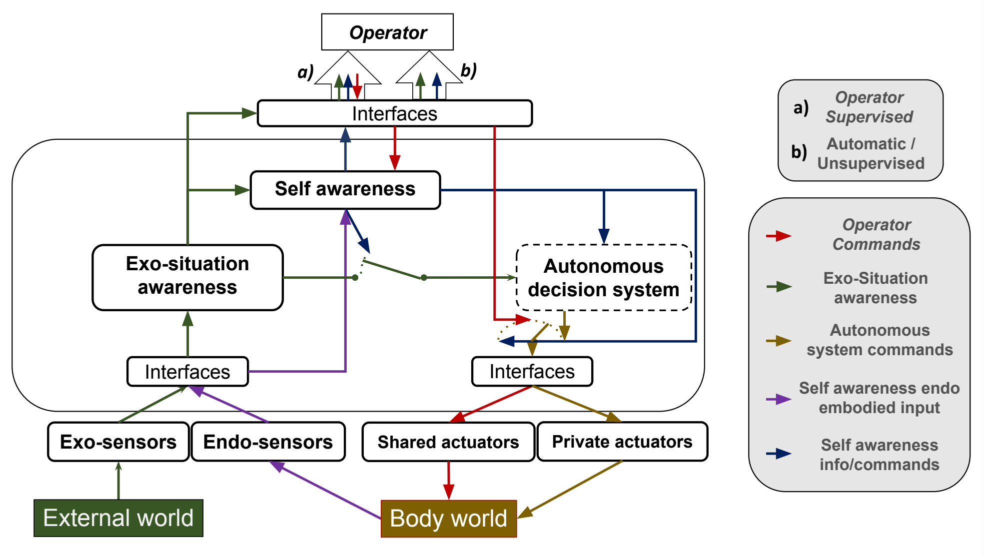
43rd IEEE International Conference on Acoustics, Speech and Signal Processing April 15-20 2018, Calgary, Alberta, Canada
Tuesday 17th April, 2018

[1]DITEN, University of Genoa, Italy.
[2]CIS, Queen Mary University of London, UK.
[3]Carlos III University of Madrid, Spain.

Outline

Introduction

Generic self-awareness scheme



Motivation

- An autonomous system need perception to navigate through scenes and recognize objects in real environments [1].
- The capability of detecting abnormal situations based on self-awareness is an important task that allows autonomous systems to increase their situational awareness and the effectiveness of the decision making submodules [2].

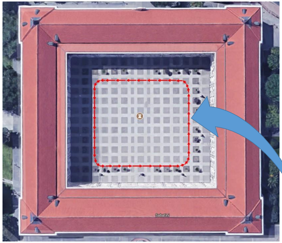
[1] D. Ramík, C. Sabourin, R. Moreno, and K. Madani, "A machine learning based intelligent vision system for autonomous object detection and recognition," *Applied Intelligence*, vol. 40, no. 2, pp. 358–375, 2014

[2] V. Bastani, L. Marcenaro, and C. S. Regazzoni, "Online nonparametric bayesian activity mining and analysis from surveillance video," *IEEE Transactions on Image Processing*, vol. 25, pp. 2089–2102, May 2016

Objectives

- We focus on multi-sensor anomaly detection for moving cognitive agents using both external and private first-person visual observations.
- The observation types are used to characterize agents motion in a given environment.
- The proposed method provides two levels:
  i) A Shared Level (SL) self-awareness from external viewpoint.
  ii) A Private level (PL) self-awareness from first person viewpoint.

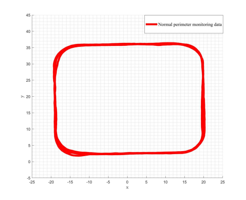
Problem definition



Task is perimeter monitoring by doing turn inside the environment

Normal situation

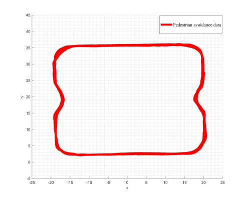
SL



PL

Abnormal situation

SL

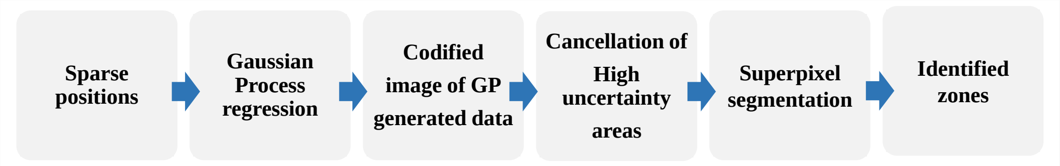


PL

Shared Level of self-awareness

Representation of observed dynamic motion

- Sparse positions represents the Locations of the entity take from input video or sensor.

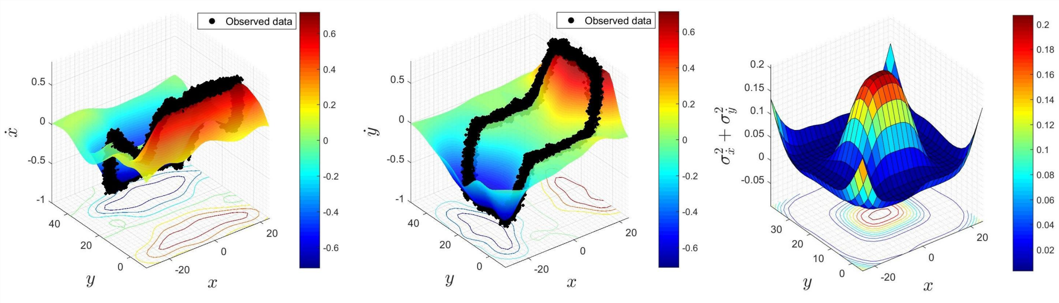


Representation of observed dynamic motion: Gaussian Process approach

- It is proposed to use a GP approach [3], such that:

$$\dot{\tilde{X}} = g(\tilde{X}) + v, \tag{1}$$

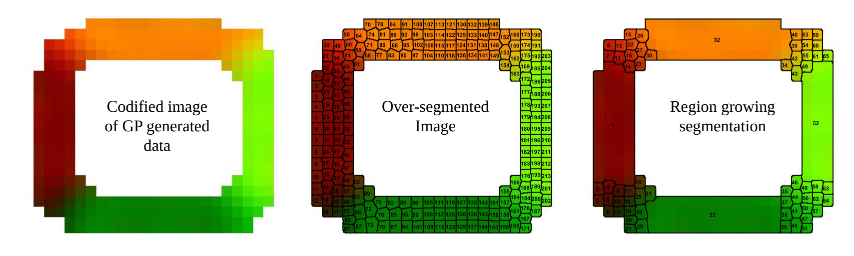
Where $\dot{\tilde{X}}$ represents an estimation of velocity, $g(\cdot)$ takes location information and estimates the expected motion (action) at such position for a given activity.



[3] K. Kim, D. Lee, and I. Essa, "Gaussian process regression flow for analysis of motion trajectories," pp. 1164–1171, 2011

Representation of observed dynamic motion: Superpixel algorithm

- Using a superpixel algorithm[4] to discretize the image plane into N zones:



- Linear dynamic model:

$$X_{k+1} = X_k + \Delta k U_{n,k} + w_m, \tag{2}$$

where $U_{n,k} = [\dot{x}_n, \dot{y}_n]^T$, is a control input that encodes the action (motivation) of the agent.

[4]Z. Li and J. Chen, "Superpixel segmentation using linear spectral clustering," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1356–1363, June 2015

Abnormality detection: Kalman filter method

- Building a set of Kalman Filters (KFs) based on the built dynamical models given the N zones.
- KFs' innovations can be used to express abnormalities since they quantify the deviations from normal learned models in the environment:

$$\epsilon_{k,n} = Z_k - \hat{X}^n_{k|k-1}, \tag{3}$$

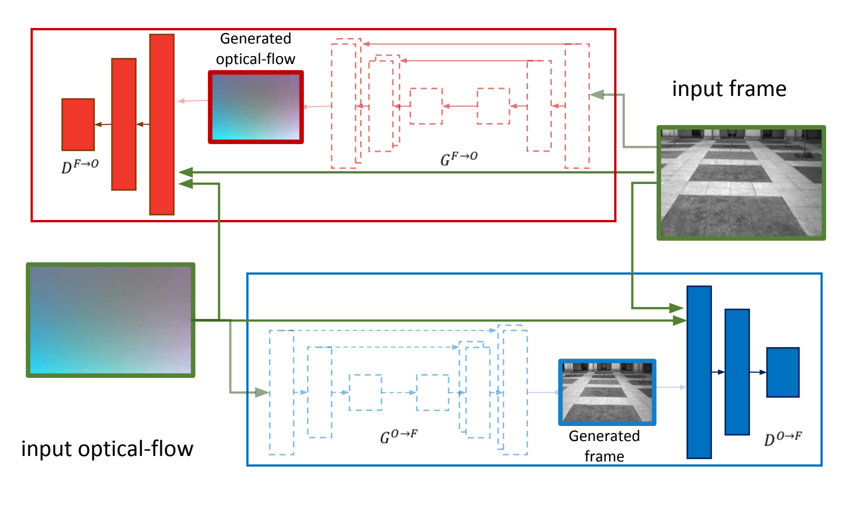
  where $\epsilon_{k,n}$ is the innovation generated in the zone n where the agent is located. Z_k represents observed spatial data and $\hat{X}^n_{k|k-1}$ is the KF estimation of the agent's location at the future time k calculated in the time instant $k-1$ (2).
- Innovation vectors are composed of two components, the magnitude of those vectors can be considered as a final measure of abnormality, ξ:

$$\xi_k = ||\epsilon_{k,n}||_2,$$

Private Layer of self-awareness

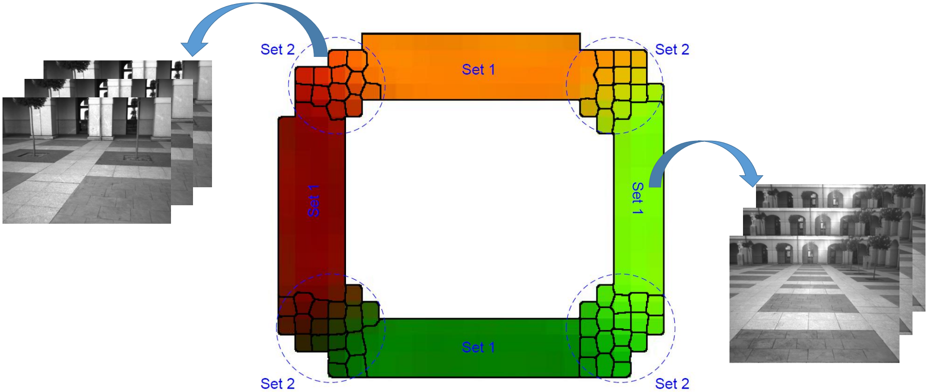
Learning the normal pattern of the observed scene

- Two networks (GANs[5]) structure are used to learn the normal pattern of the observed scene.



Learning the normal pattern of the observed scene

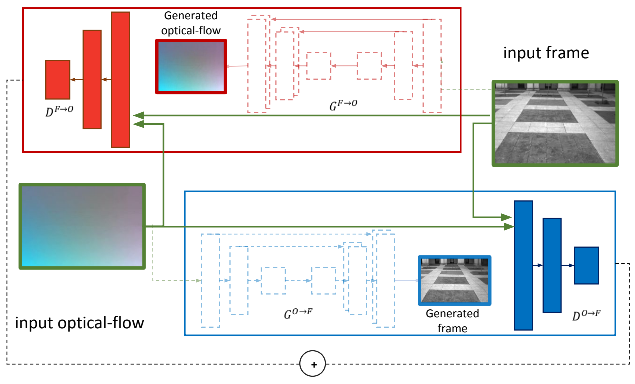
- Frames (F) and corresponding optical-flow images (O) are collected from the *normal* scenario.
- Constructing a *Bank of Discriminators* on the GP identified zones grouping into two sets:
  i) $Set1$: which is trained on a straight path.
  ii) $Set2$: that is trained over the curves.



Anomaly detection by using discriminators of GANs

1- Given a test frame F and its corresponding optical-flow image O, we first produce the reconstructed p_O and p_F using $G^{F\to O}$ and $G^{O\to F}$, respectively.

2- The pairs of patch-based discriminators $\hat{D}^{F\to O}$ and $\hat{D}^{O\to F}$ are applied respectively to the first and second tasks.



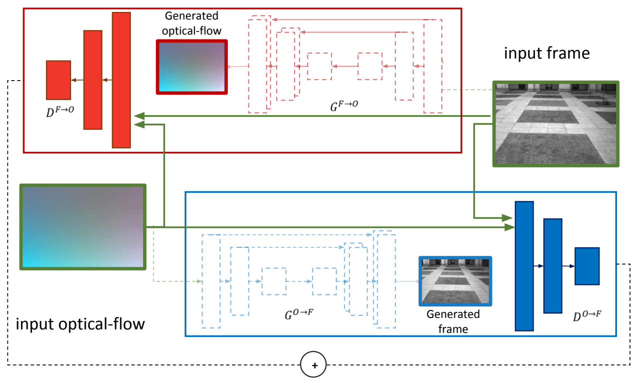
Anomaly detection by using discriminators of GANs

3- Computing scores for the ground truth: S^O and S^F, and the prediction: S^{p_O} and S^{p_F}.

4- Define abnormality as innovation w.r.t the Discriminators scores:

i) The two scores are summed: $S_{observation} = S^O + S^F$ and $S_{prediction} = S^{p_O} + S^{p_F}$

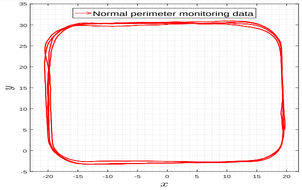
ii) Innovation: $\tilde{Y} = S_{observation} - S_{prediction}$



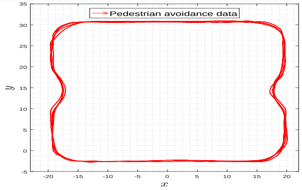
Results

Experimental setup

Proposed approach is validated with data acquired from a real vehicle 'iCab' during a perimeter monitoring task.



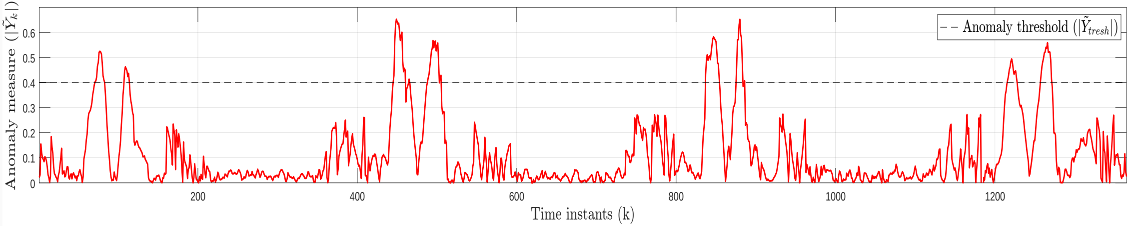
Perimeter monitoring



Pedestrian avoidance

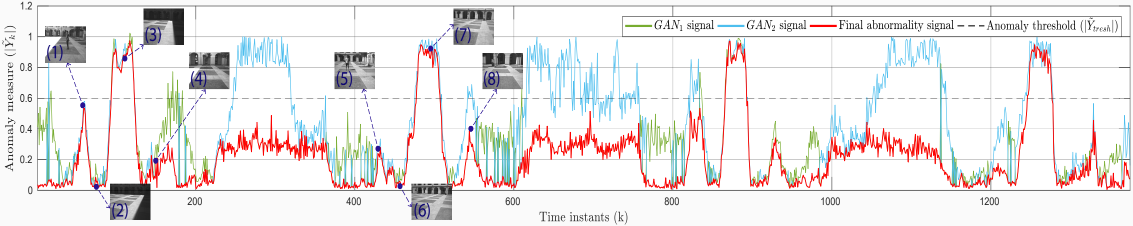
Pedestrian avoidance from 'iCab' on board camera

SL anomaly measurements: perimeter control activity by GP through time with avoidance of static pedestrians.



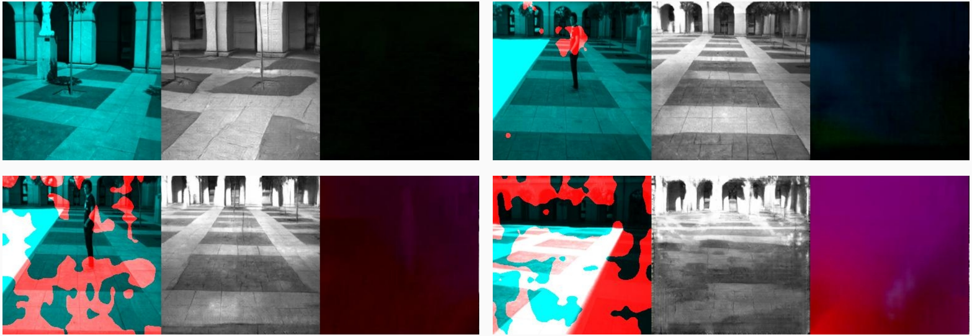
Private Level Self Awareness abnormality detection

PL anomaly measurements: the distances between the observations and predictions by GANs during the time.



Private Level Self Awareness abnormality detection

Visualization of local abnormality in first-person vision



Conclusions

Conclusions

- Self-awareness in autonomous system
- Shared and private layers for self-awareness
- Methodology based on multi-perspective approach to detect anomalies for moving agents
- SL and PL learned models are used to predict the dynamics of a vehicle performing a task

Thank you!

D. Ramík, C. Sabourin, R. Moreno, and K. Madani, "A machine learning based intelligent vision system for autonomous object detection and recognition," *Applied Intelligence*, vol. 40, no. 2, pp. 358–375, 2014.

V. Bastani, L. Marcenaro, and C. S. Regazzoni, "Online nonparametric bayesian activity mining and analysis from surveillance video," *IEEE Transactions on Image Processing*, vol. 25, pp. 2089–2102, May 2016.

K. Kim, D. Lee, and I. Essa, "Gaussian process regression flow for analysis of motion trajectories," pp. 1164–1171, 2011.

Z. Li and J. Chen, "Superpixel segmentation using linear spectral clustering," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1356–1363, June 2015.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems 27* (Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, eds.), pp. 2672–2680, Curran Associates, Inc., 2014.