HIERARCHY OF GANS FOR LEARNING EMBODIED SELF-AWARENESS MODEL

Mahdyar Ravanbakhsh **Mohamad Baydoun** **Damian Campo** **Pablo Marin** **David Martin** **Lucio Marcenaro** **Carlo S. Regazzoni**

DITEN, University of Genova, Italy Carlos III University of Madrid, Spain

Introduction

Recently, several architectures have been proposed for learning embodied agents complex **self-awareness (SA)** models [1,2,3].
The main goal is that the autonomous system **learn a model of itself** while doing a certain task when driven by a human by looking at.
Accordingly, by transferring the learned model from human driver to the autonomous system develop a **self-awareness** model.

- We propose the **dynamic incremental SA models .**
- Experiences done by an agent modeled in a **hierarchical fashion**, starting from more simple situations to more structured ones.

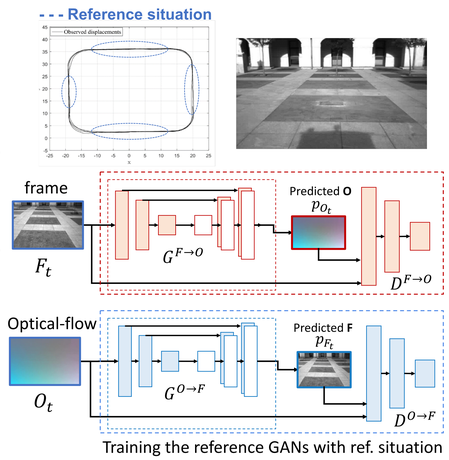
- A **cross-modal Generative Adversarial Networks (GANs)** used to process high dimensional visual data.
- Different levels of the GANs are detected in a **weakly-supervised** manner using **GANs discriminators** decision boundaries.

Training hierarchy of GANs

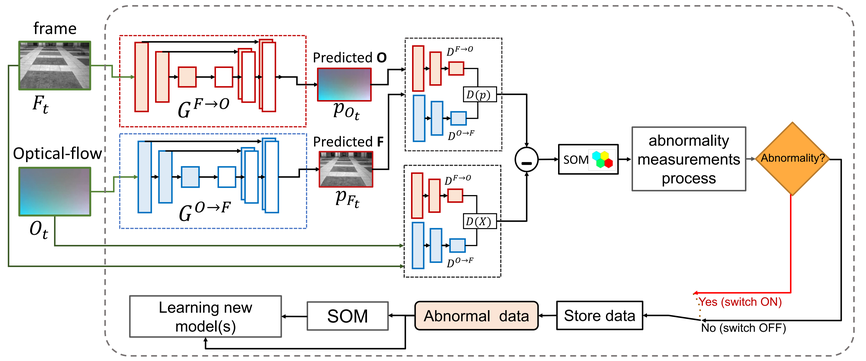
- **Goal**: learning a complex distribution.

- **Idea**: Split the distribution into smaller sets and estimate it as a **mixture** of multiple **small distributions**.

- **Procedure**: Learning starts by *normal* scenarios and from a simple situation, when the agent moves straight.
- Using the **discriminator scores** as a measure to **detect** new situations, and train new models **incrementally**.



Training the reference GANs with ref. situation



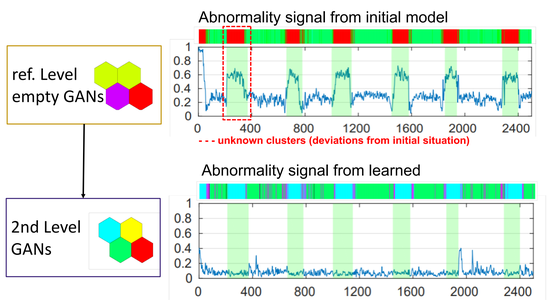
The incremental adaptive process to learn GANs hierarchy with normal training data

- The **discriminator scores** correspond to the error/**innovations** with respect to the other models that already learned.

Training output

a) The reference GANs: detects straight movements, fails on curves

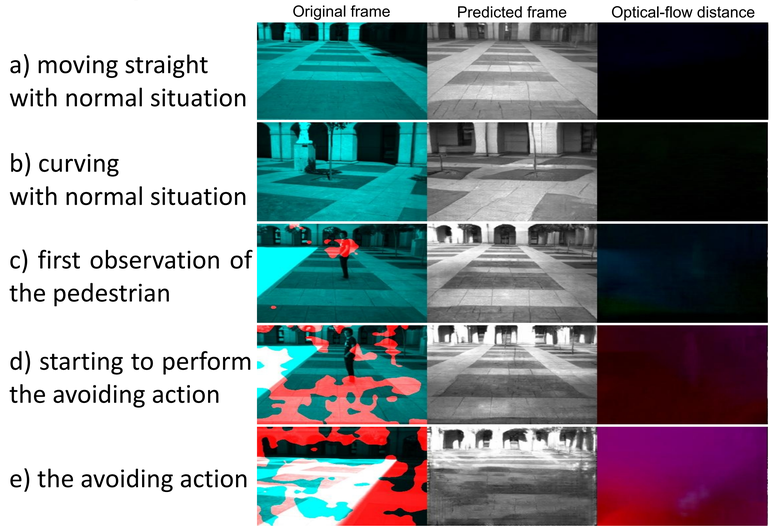
b) Detected new situation: training the next level of GANs.



Online testing

Test scenario is an **abnormal** situation (presence of a pedestrian), where it **never observed** by the autonomous agent before.

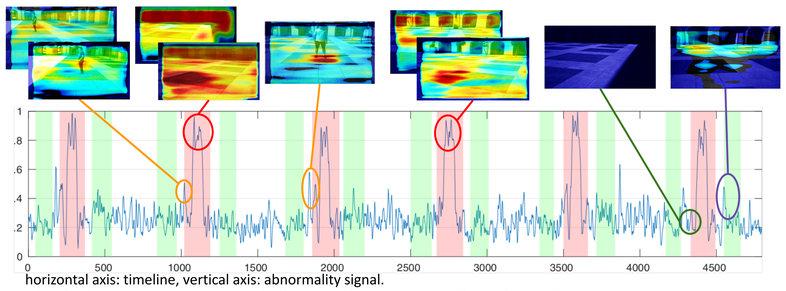
Abnormality visualization



Red blobs are the possible abnormal areas

Results analysis

- Abnormality heatmap shows the possible abnormal areas



horizontal axis: timeline, vertical axis: abnormality signal.
white background: moving on a straight line, green bars: curving, red bars: abnormal events

References

[1] M. Baydoun, D. Campo, V. Sanguineti, L. Marcenaro, A. Cavallaro, and C. Regazzoni, "**Learning switching models for abnormality detection for autonomous driving**", FUSION, 2018.
[2] M. Ravanbakhsh, M. Baydoun, D. Campo, L. Marcenaro, and C. Regazzoni, "**Learning multi-modal self-awareness models for autonomous vehicles from human driving**", FUSION, 2018.
[3] P. R. Lewis, M. Platzner, B. Rinner, J. Tørresen, andX. Yao, "**Self-aware Computing Systems**", 2016.

Universidad Carlos III de Madrid
UNIVERSITÀ DEGLI STUDI DI GENOVA