CAI4DSA: Collaborative Explainable neuro-symbolic AI for Decision Support Assistant

University of Florence, DINFO Dept.

Paolo Nesi, DISIT lab, UNIFI

10 March 2025

Aula Magna, Polo Fibonacci, University of Pisa

Future Artificial Intelligence Research

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB | SNAP4CITY

Aim

- **CAI4DSA aims at exploiting AI for creating decision support systems** (DSS) which can evolve with the collaboration among and with humans.
- The **NGDSS** should be capable to react to changes in the context, interact with humans, and learn from the occurrence.
- **Symbolic and neural models** could be also shared in communities to create a global knowledge and understanding of the problems and solutions.

outcomes

- **New Algorithms or Models**: Development of novel AI algorithms or models that can enhance decision-making processes.
- **NGDSS Theories**: Contribution to the theoretical understanding of how AI can be effectively integrated into decision support systems.
- **Early Prototype** of the NGDSS, providing a proof of concept for how AI can enhance decision-making in a number of target domains

Main objectives

Obj.1. Continuous incremental **reinforced learning** to address the possibility of learning taking into account new facts and implications.

Obj.2. **human readable multidomain XAI** representations to facilitate human understanding and trust by providing explanations of the produced decisions/suggestions for a given scenario with a particular care to **explain seasonalities** of variables/conditions, contextual aspects, and the effects of inputs with respect to outputs and KPI.

Obj.3. **extract and formalize new knowledge from XAI**, results/KPIs as facts, relationships, constraints, models (causals, maths, etc.) also deciding if: they can actually drive the learning and/or decision process, for shortening the learning process, reducing features, reducing temporal span in time series analysis, increasing precision.

Obj.4. **Human NewGenDSS interaction which should be capable to work reconcile with humans**: (i) collecting information the relevance/confidence of the produced results is not satisfactory, (ii) providing motivations (in connection with KPI-Driven XAI) behind the lack of knowledge and/or data, (iii) collecting human suggestions and present some counterindications or simply demonstrate that cannot be verified, (iv) provide general learning process indicators which should give the evidence that the general knowledge / understanding is improving and its coverage, (v) provide instruments for defining scenarios.

Obj.5. **Validation of NewGenDSS in critical infrastructure** domains: mobility and transport, environment, medicine, security, complex manufacturing, etc.

Main Domains vs research Groups

- **Mobility and Transport**: identification of optimal solutions (macro scale scenarios and parameters) to solve critical traffic conditions coming from the occurrence of natural and non-natural disasters, impacting on traffic flow, traffic direction, regulation of traffic controllers. **(DISIT lab, P. Nesi)**
- **Medicine**: identification ablation targets in atrial fibrillation (AF): **(AI Lab, P. Frasconi).**
- **Complex manufacturing**: simulation of curation processed depending on the product and placement of material in autoclaves for reducing costs and optimizing time. The model is a complex partial differential equation in turbulent conditions, which the state-of-the-art approaches use finite differences to solve in unacceptable execution time, without addressing the variation of parameters. **(DISIT Lab, P. Nesi)**
- **Security and computer vision:** continual learning can aid in updating DCNN models with new data inevitably alters the latent space environmental observations, surveillance. **(MICC, P. Pala)**

main team members

DISIT LAB

- ***Paolo Nesi,*** *PO: Monthly commitment: 0,166 (project PI and CV attached)*
- *Pierfrancesco Bellini, PA: Monthly commitment 0,2 (CV attached)*

AI Lab:

- ***Paolo Frasconi,*** *PO: Monthly commitment 0,133 person/month. (CV attached)*
- *Simone Marinai, PA: Monthly commitment 0,133 person/month*
- *Marco Lippi, PA: Monthly commitment 0,133 person/month*

MICC:

- *PO: Monthly commitment 0,2 (CV attached)*
- *Lorenzo* ***Pietro Pala,*** *Seidenari (PA): Monthly commitment 0,0866*
- *Andrew Bagdanov (PA): Monthly commitment 0,0866 (CV attached)*

Mobility and Transport:

- **Comparing Techniques for Temporal Explainable Artificial Intelligence**. Canti, E., Collini, E., Ipsaro Palesi, L. A., & Nesi, P. (2024, July). In *2024 IEEE 10th International Conference on Big Data Computing Service and Machine Learning Applications (BigDataService)* (pp. 87-91). IEEE.
  - https://ieeexplore.ieee.org/document/10730341

- **Neuro-Symbolic Optimization of Mobility Infrastructure via Deep-RL-GNN**, E. Collini, L. A. Ipsaro Palesi, C. Lucchesi, P. Nesi, submitted to IEEE Access
  - **Evolution of: Mobility Infrastructure Optimization via Stochastic Relaxation the Traffic Flow Reconstruction**. Bilotta, S., Collini, E., Fanfani, M., & Nesi, P. (2024).
  - https://assets-eu.researchsquare.com/files/rs-5093651/v1_covered_1b2f1826-1464-4a57-9514-7810680a57c1.pdf

- **New version of the Km4City ontology** integrating detailed digital twin aspects and results of neuro-symbolic deductions.
  - In progress
- **LLM for DSS solution generation** in the context of Smart City, mobility advisors
  - In progress

Comparing Techniques for Temporal Explainable Artificial Intelligence

Objective

Evaluate **XAI methodologies** on the time series analysis in the AI models, focus is on **global XAI aspects**

Time series
- With seasonality
- Without short-term seasonality
- With outliers

eXplainability methodologies
- Shapley Values
- Integrated Gradients
- LIME

Explainable AI Methodologies

SHAP

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f_{S\cup\{i\}}(x_{S\cup\{i\}}) - f_S(x_S)]$$

LIME

$$IG_i^{approx}(x) := (x_i - x_i') \times \sum_{k=1}^{m} \frac{\partial F(x' + \frac{k}{m} \times (x - x'))}{\partial x_i} \times \frac{1}{m}$$

IG

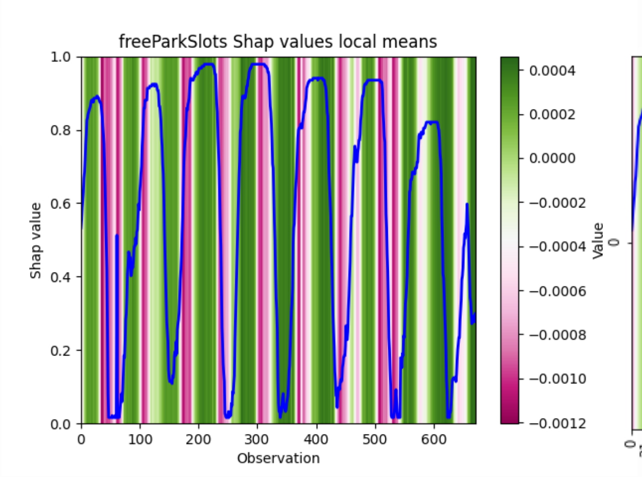
$$\xi(x) = \underset{g \in G}{\operatorname{arg}min} \mathcal{L}(f, g, \pi_x) + \Omega(g)$$

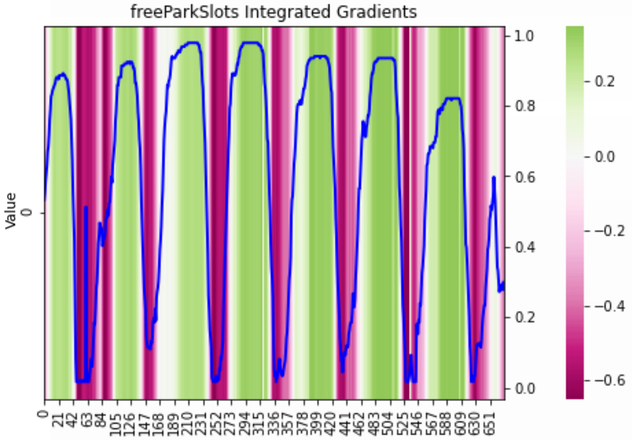
$$\mathcal{L}(f, g, \pi_x) = \sum_{z, z' \in Z} \pi_x(z)(f(z) - g(z'))^2$$

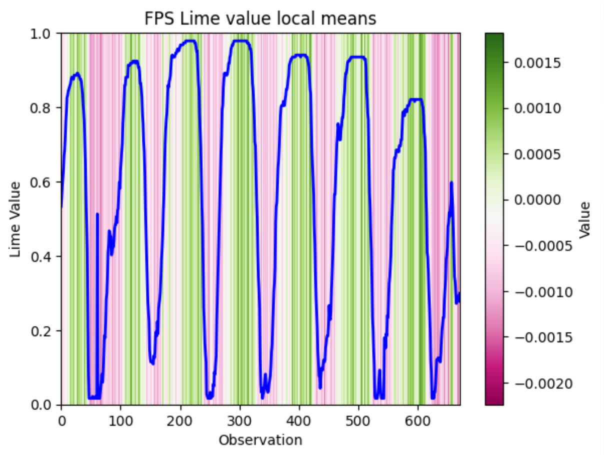
Comparing Techniques for Temporal Explainable Artificial Intelligence

Experimental result

Time series with seasonality



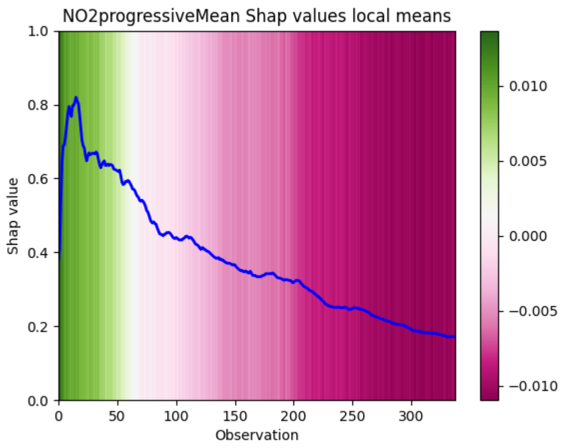


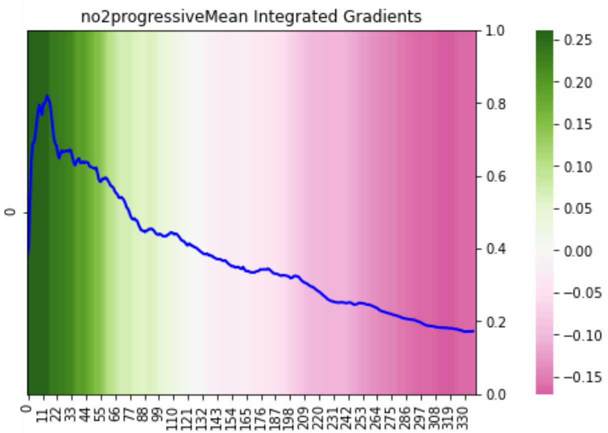


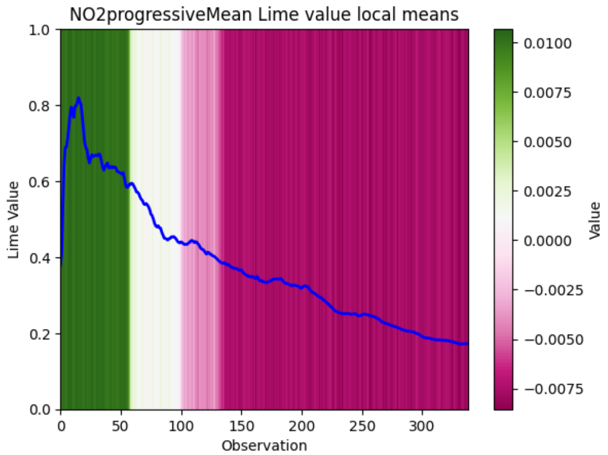
Comparing Techniques for Temporal Explainable Artificial Intelligence

Experimental result

Time series without short-term seasonality



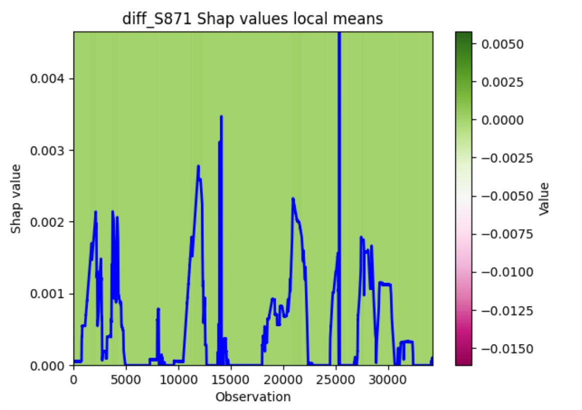


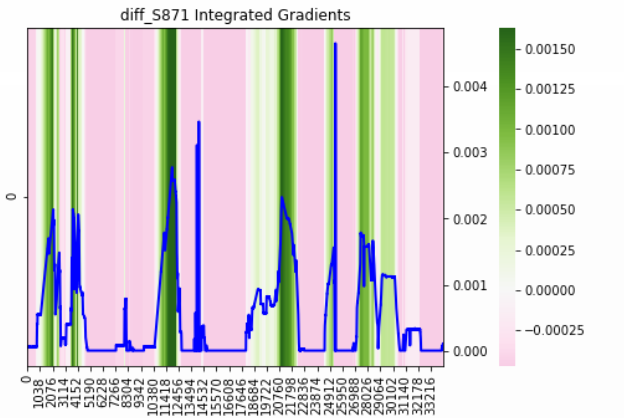


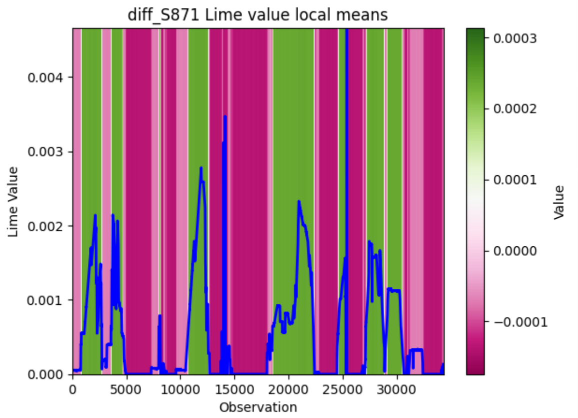
Comparing Techniques for Temporal Explainable Artificial Intelligence

Experimental result

Time series with outliers







Comparing Techniques for Temporal Explainable Artificial Intelligence

Results of: comparing XAI for time series

Time series

- With seasonality
- Without short-term seasonality
- With outliers

With seasonality, Without short-term seasonality: coherent behavior from the XAI methodologies (SHAP, IG, and LIME) compared

With outliers: SHAP was not able to provide insights on the target variable compared to LIME and IG. Between these two IG was able to provide more coherent explanations with respect to LIME

EXECUTIONTIME
IG resulted faster than SHAP and LIME

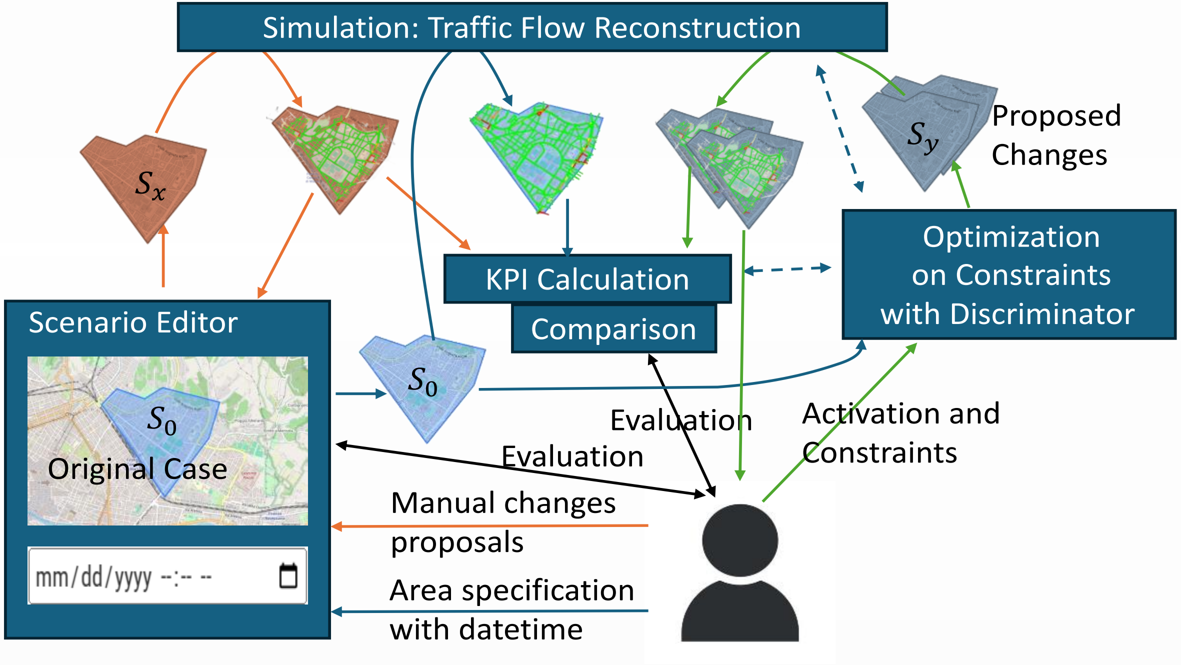
Comparing Techniques for Temporal Explainable Artificial Intelligence

Neuro-Symbolic Optimization of Mobility Infrastructure via Deep-RL-GNN

Objectives

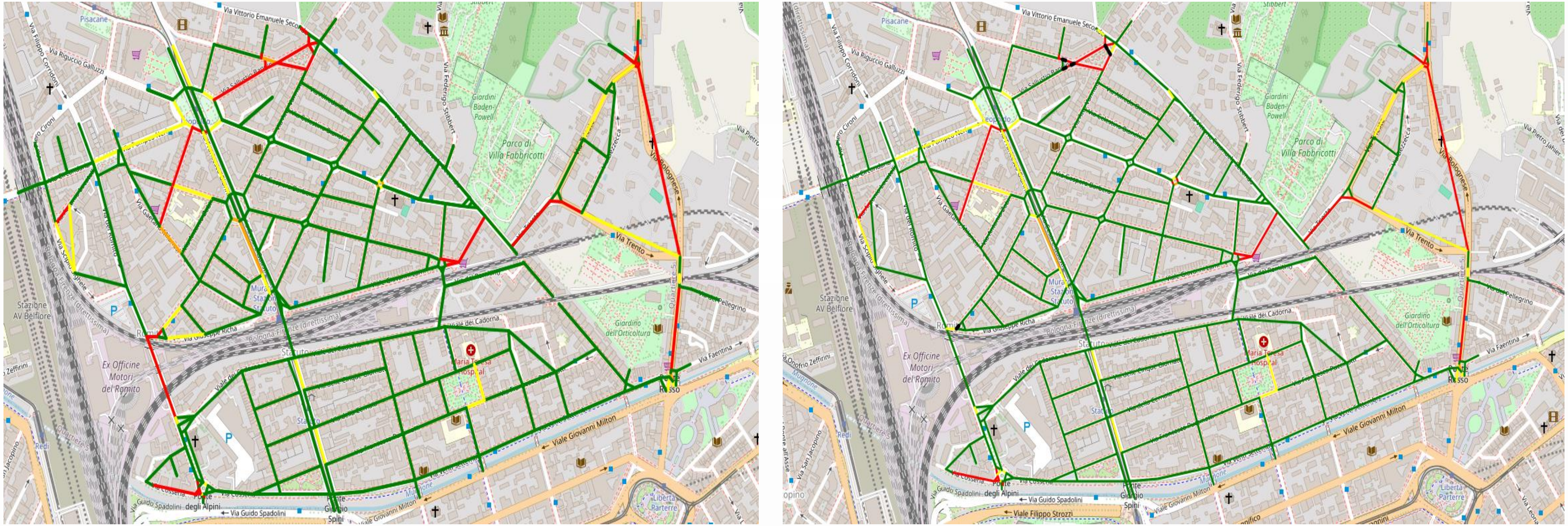
- **Optimization with respect to various KPIs on traffic conditions** including congestion levels, fuel consumption, CO_2 emissions based on **Deep-RL-GNN**.
  - **Taking into account limitations in generating solutions**
  - **Taking into account constraints such as: (i) not modifiable road segments, (ii) traffic regulations**
  - **Generating viable road lane changes**.
- **Comparing the solution with respect to stochastic relaxation optimization approach.** The proposed solution provides
  - **better performance** with respect to stochastic relaxation optimization
  - **reduction of time and costs with respect to the state-of-the-art solution.**
- **Proposing and assessing the AI model transferability** on different (i) urban scenarios, and (ii) traffic workload conditions.

Data Flow of the solution



Neuro-Symbolic Optimization of Mobility Infrastructure via Deep-RL-GNN

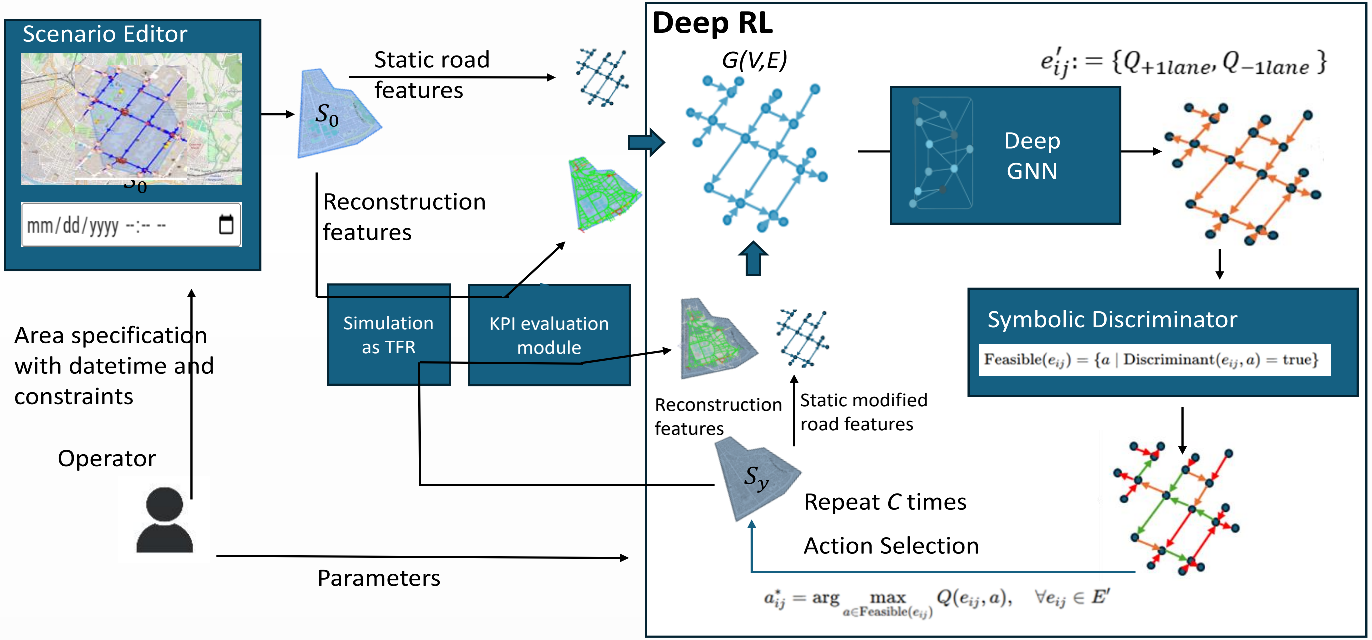
Original and optimized (with changes)



Neuro-Symbolic Optimization of Mobility Infrastructure via Deep-RL-GNN

Traffic flow density represented

Deep-RL-GNN

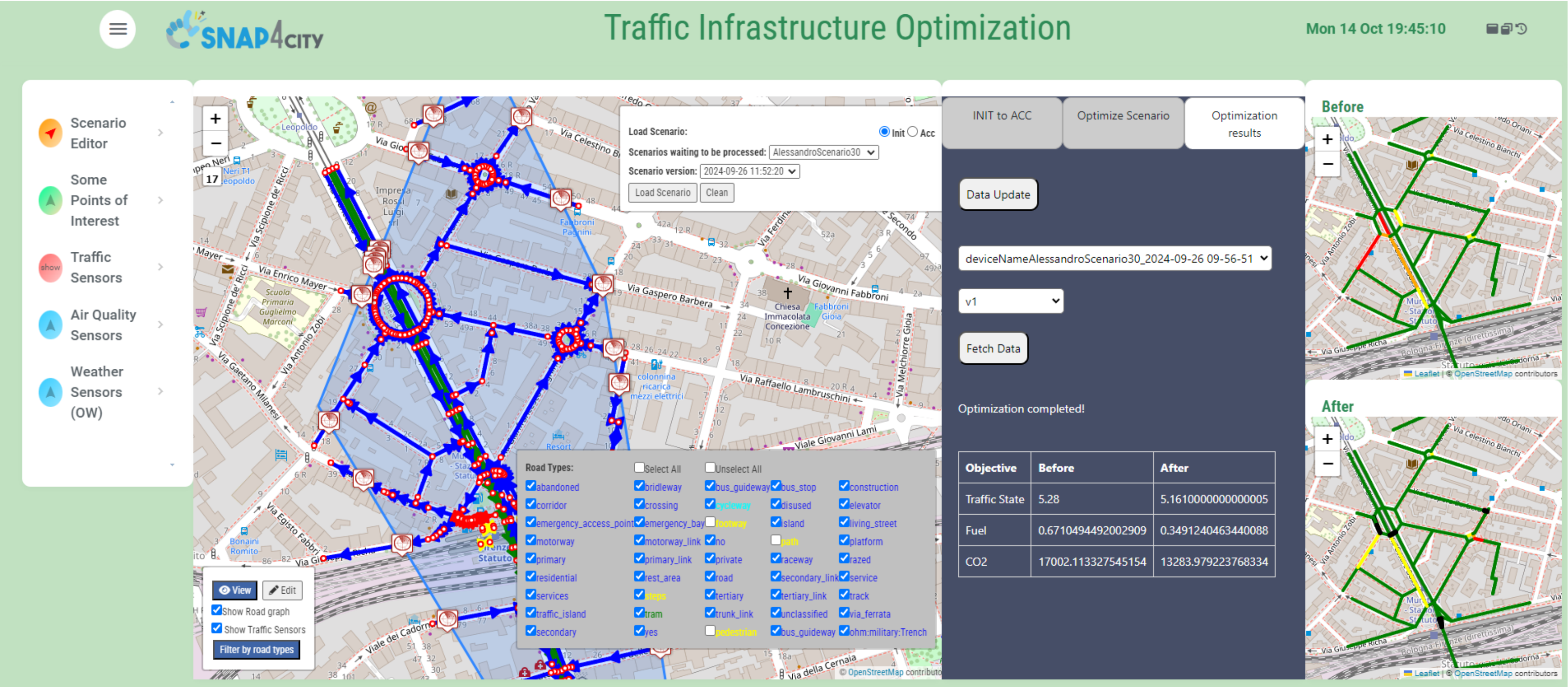


Neuro-Symbolic Optimization of Mobility Infrastructure via Deep-RL-GNN

Future Artificial Intelligence Research

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB | SNAP4CITY

Traffic Infrastructure Optimization



Results of Neuro-Symbolic Optimization

The results reported on a limited number of **4 changes** improvements in a new urban scenario improving **fuel consumption and CO2 emissions** in both scenarios.

The proposed approach **outperforms Stochastic Relaxation** in all cases.

- Solution has validated the generalization and transferability capabilities of the Deep-RL-GNN, across different urban areas and at various traffic conditions.
- Compared to **Transfer Learning on**
  - traffic load achieves a 9.82% mean improvement with respect to the initial scenario,
  - area change improves KPIs by 25.58% with respect to the initial scenario.

Complex Manufacturing

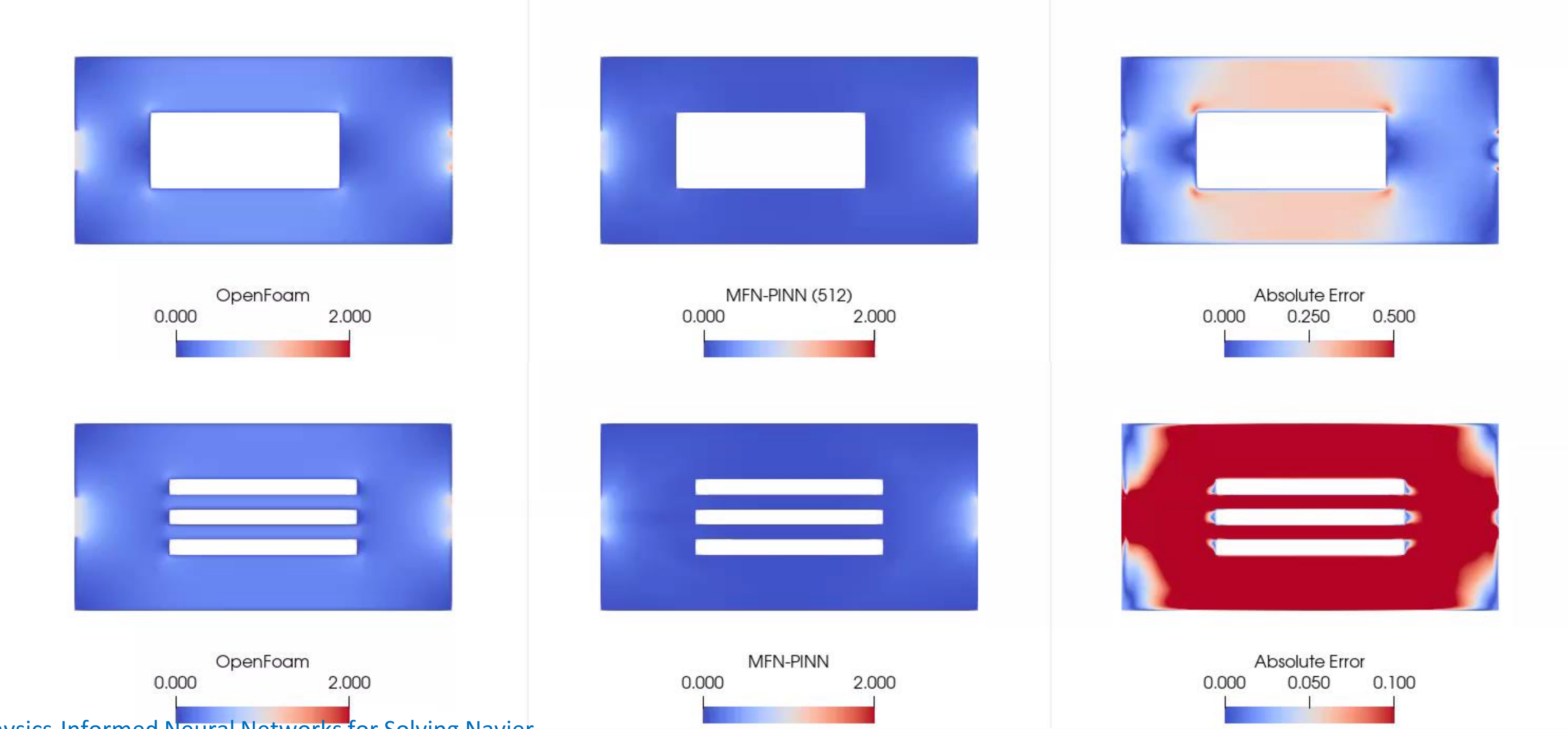
- T. Botarelli, M. Fanfani, P. Nesi, L. Pinelli, "**Using Physics-Informed Neural Networks for Solving Navier-Stokes Equations in Fluid Dynamic Complex Scenarios**",
  - Engineering Applications of Artificial Intelligence, Elsevier, 2025.
  - https://www.sciencedirect.com/science/article/pii/S0952197625003471

Using Physics-Informed Neural Networks for Solving Navier-Stokes Equations in Fluid Dynamic Complex Scenarios

- **Usage of PINN for solving PDEs for fluid flow** on complex scenarios
  - Constrained pipes as autoclaves
  - Multiple shapes insides
- **Definition of learning processes and methods**
- **Training strategies** to reduce solution time:
  - Solutions valid for multiple shapes
  - Transfer learning
- **Validation wrt Open Foam**
- **Precision on steady and transitory cases**
- **Videos on https://www.snap4city.org/1010**

Using Physics-Informed Neural Networks for Solving Navier-Stokes Equations in Fluid Dynamic Complex Scenarios

Comparison of PINN vs OpenFoam and error



Using Physics-Informed Neural Networks for Solving Navier-Stokes Equations in Fluid Dynamic Complex Scenarios

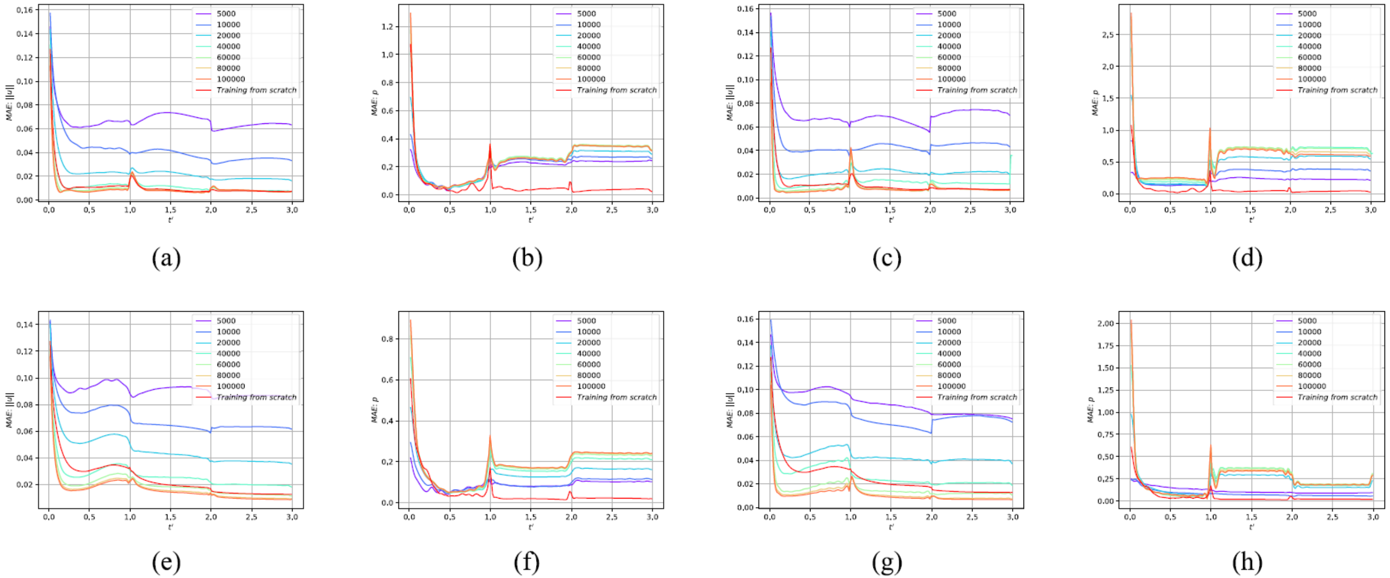


Fig. 13. Plots of $MAE(t')$ for the fine-tuning tests starting from a model trained on the single box geometry (Section 7.1, Table 8). In (a) and (b) velocity magnitude $\|\mathbf{u}\|$ and pressure p errors for the circular cylinder using $lr = 10^{-3}$; in (c) and (d) the same test using $lr = 10^{-2}$. Similarly, in (e) and (f) velocity magnitude $\|\mathbf{u}\|$ and pressure p errors for the three-boxes using $lr = 10^{-3}$; in (g) and (h) using $lr = 10^{-2}$. In all the plots errors related to the estimates obtained with a model trained from scratch are reported in red. By augmenting the number of fine-tuning epochs more accurate results are achieved in all the cases. Moreover, 40000 epochs are sufficient to obtain pretty accurate results, while after 60000 epochs results are practically equals to those obtained by training the model from scratch. (Best viewed in color). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Using Physics-Informed Neural Networks for Solving Navier-Stokes Equations in Fluid Dynamic Complex Scenarios

Security and Computer Vision

Tasks

- Leverage Spiking Neural Networks for efficient and robust adaptation of Convolutional NN architectures to the domain of neuromorphic camera streams
- Efficient adaptation of cross-modal Vision Language Models for intra-modal tasks
- The Probability Simplex Leads to Compatible Representations (under review)

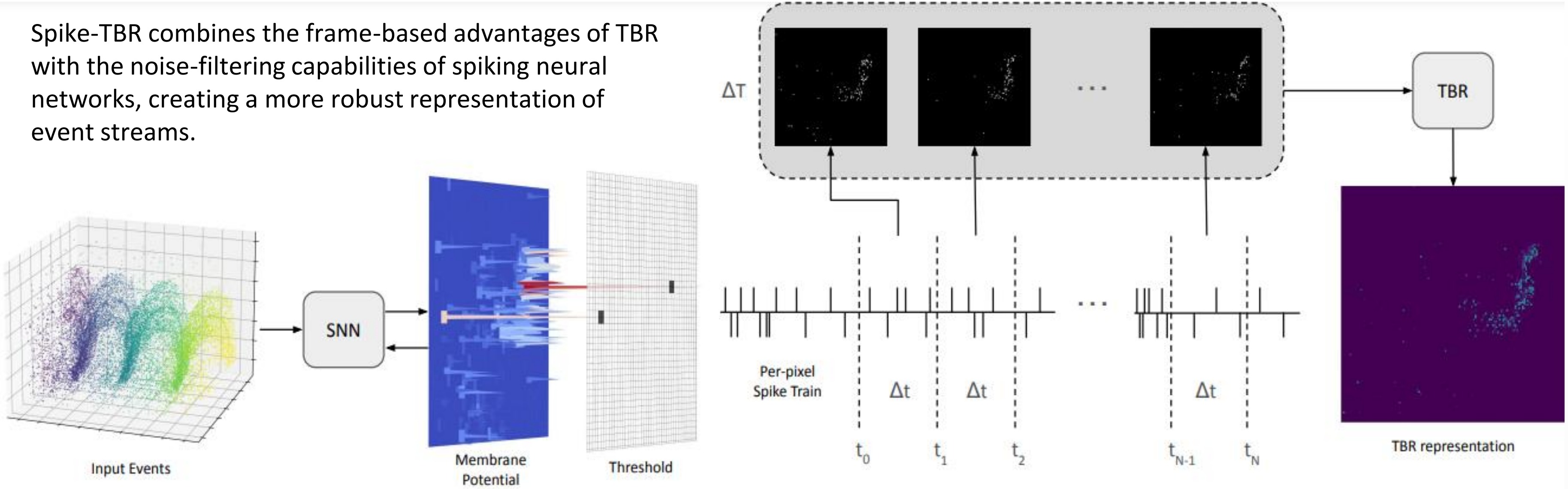
Publications:

- Cross the Gap: Inter-modal CLIP Representations Are Superior for Intra-modal Tasks. Int. Conf. on Learning Representations, ICLR, Singapore, April 2025.

Spiking NN for Neuromorphic Vision

Neuromorphic cameras, also referred to as event cameras, offer significant advantages over traditional frame-based sensors, including higher temporal resolution, lower latency and dynamic range. However, efficiently converting event streams into formats compatible with standard computer vision pipelines remains a challenging problem, particularly in the presence of noise. We propose Spike-TBR, a novel event-based encoding strategy based on Temporal Binary Representation (TBR), addressing its vulnerability to noise by integrating spiking neurons.

Spike-TBR combines the frame-based advantages of TBR with the noise-filtering capabilities of spiking neural networks, creating a more robust representation of event streams.

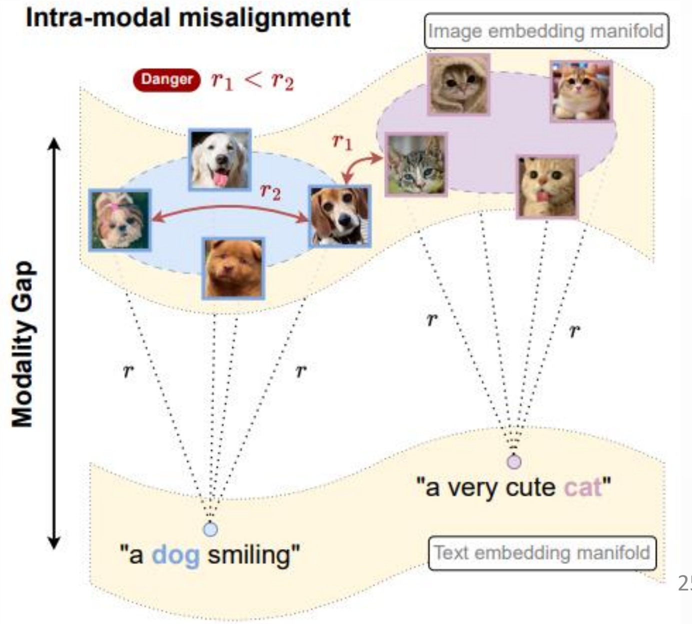


Exploiting inter-modal representations for intra-modal tasks

Pre-trained multi-modal Vision Language Models like CLIP are widely used for a variety of applications. Previous work has shown that, due to contrastive pre-training, there is a modality gap between the text and image feature embedding spaces.

We show that the common practice of individually exploiting the text or image encoders of these powerful multimodal models is highly suboptimal for intra-modal tasks like image-to-image retrieval. We argue that this is inherently due to the inter-modal contrastive loss commonly used to train CLIP models.

We empirically show in multiple settings (image retrieval, text retrieval, and zero-shot image classification) that approaching these tasks inter-modally significantly improves performance with respect to intra-modal baselines on more than fifteen datasets.

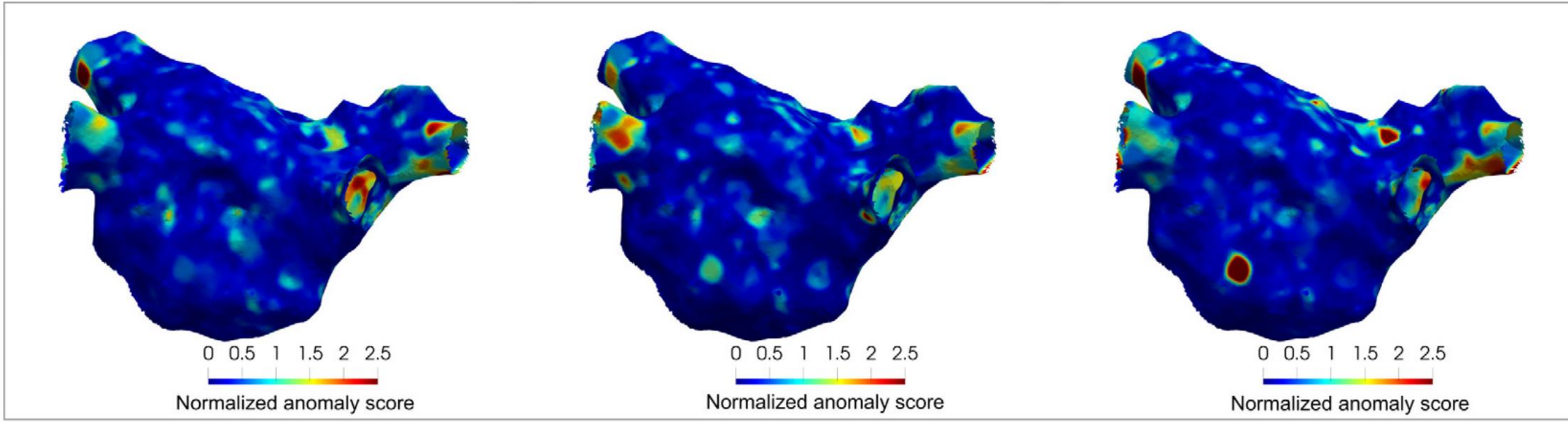


Cross the Gap: Inter-modal CLIP Representations Are Superior for Intra-modal Tasks.
Int. Conf. on Learning Representations, Singapore, April 2025.

AI Lab: Medicine vs LLM

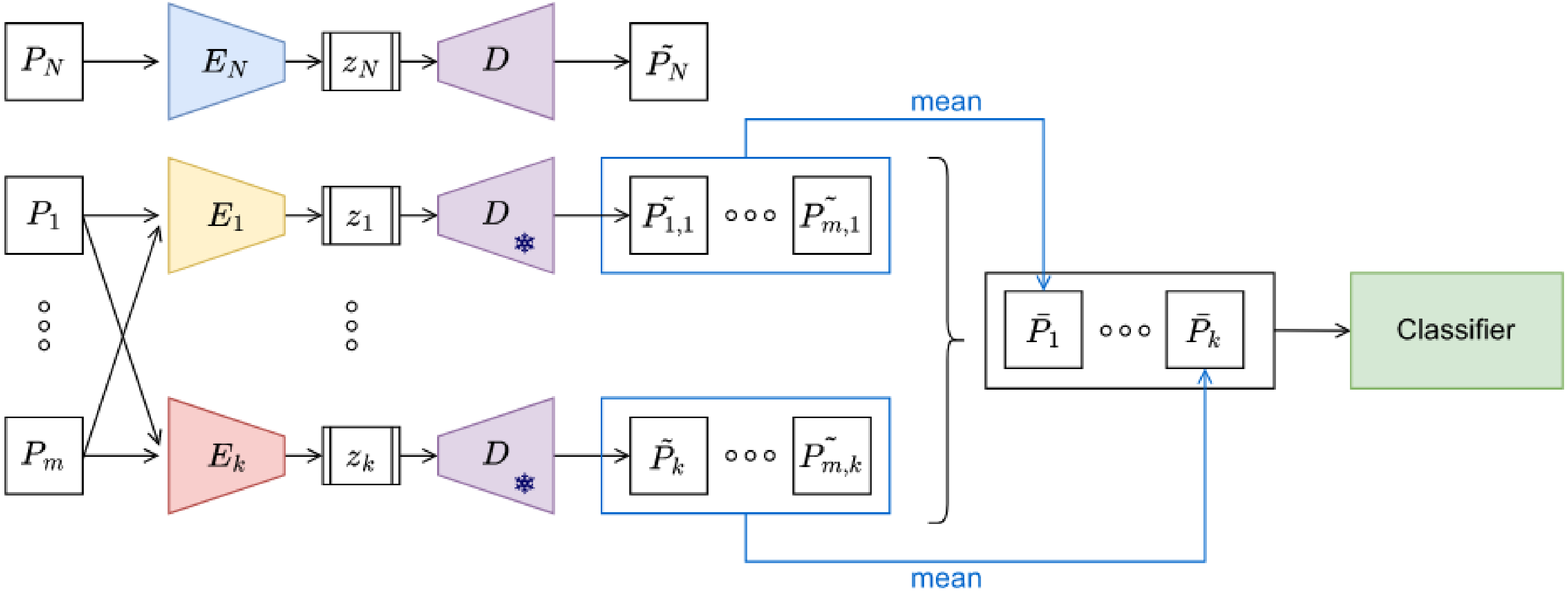
- **A comprehensive tool that pursues the identification, characterization, and linking of visual elements to semantic and context data, leveraging large language models for interoperability.**
  - The tool supports users to interact with scholarly documents, enabling multi-layered exploration and offering deeper insights (Massai & Marinai 2025).
- **Automatic analysis of privacy policies with Large Language Models**, in collaboration with the University of Bologna (Lagioia et al. 2025).
- **A system for spotting anomalous electrocardiograms** during ablation procedures as atrial fibrillation therapy (Bindini et al. 2024a).
  - This is a first step towards interpreting specific endocardium tissue as potential ablation targets in a human-AI interaction process.
- **An algorithm for solving the attribution problem in the few-shot learning setting** (Bindini et al. 2024b).
  - Methods of this kind are important to aid human forensic experts to explain which generative model created a certain deep fake in spite of limited data availability.
- **A large scale study to investigate the replacement of traditional stereological approaches to cell counting with AI-based** cell localizers.
  - Positive results have been achieved on 3D human brain images (Checcucci et al. 2024).

A system for spotting anomalous electrocardiograms during ablation procedures as atrial fibrillation therapy (Bindini et al. 2024a).



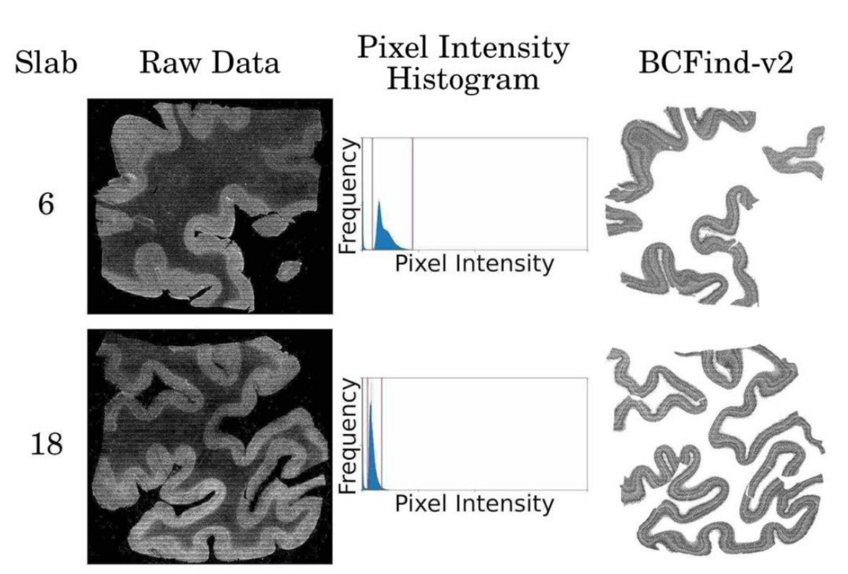
An algorithm for solving the attribution problem in the few-shot learning setting (Bindini et al. 2024b).

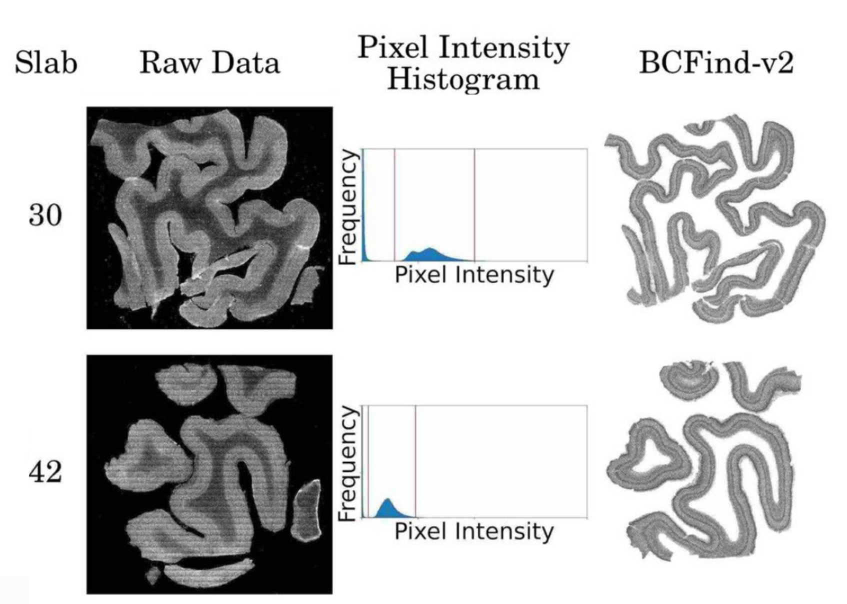
AI Lab



AI Lab

A large scale study to investigate the replacement of traditional stereological approaches to cell counting with AI-based cell localizers.





Future Artificial Intelligence Research

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

AI Lab

Medicine tasks references

- Lagioia, F., Liepina, R., Lippi, M., Palka, P., Sartor, G., "**Make Privacy Policies Longer and Appoint LLM Readers**",
  - accepted for publication in the AI&Law journal, 2025
- L. Massai, S. Marinai, "**An integrated system for interacting with multi-page scholarly documents**"
  - to appear: Proc. IRCDL conference, February 20-21 2025.
- Nardoni, V., Lippi, M., Hyeraci, G., Maccari, M., Tarazjani, A. D., Virgili, G., Gini, R., Marinai, S., **Towards Automatically Filling Questionnaires from Clinical Records with Large Language Models**, Eighth Workshop on Natural Language for Artificial Intelligence (NL4AI), November 26-27th, 2024, Bolzano, Italy
  - https://ceur-ws.org/Vol-3877/paper18.pdf
- Bindini, L., Pagani, S., Bernardini, A., Grossi, B., Giomi, A., Frontera, A., & Frasconi, P. (2024a). **All-in-one electrical atrial substrate indicators with deep anomaly detection**. Biomedical Signal Processing and Control, 98, 106737.
  - https://www.sciencedirect.com/science/article/pii/S174680942400795X
- Bindini, L., Bertazzini, G., Baracchi, D., Shullani, D., Frasconi, P., & Piva, A. (2024b). **Tiny Autoencoders are Effective Few-Shot Generative Model Detectors**. 2024 IEEE International Workshop on Information Forensics and Security (WIFS), 1–6.
  - https://ieeexplore.ieee.org/document/10810686
- Checcucci, C., Wicinski, B., Mazzamuto, G., Scardigli, M., Ramazzotti, J., Brady, N., Pavone, F. S., Hof, P. R., Costantini, I., & Frasconi, P. (2024). **Deep learning-based localization algorithms on fluorescence human brain 3D reconstruction**: A comparative study using stereology as a reference. Scientific Reports, 14(1), 14629.
  - https://www.nature.com/articles/s41598-024-65092-3

CAI4DSA: Next steps

- **Continuing the work on:**
  - DISIT Lab: DSS, XAI, PINN, Deep RL GNN, LLM, neuro-symbolic
  - MICC: Security by computer vision
  - AI Lab: Medical applications/LLM, anomaly detection
- **Converging on exploiting the solution on NewGenDSS**
  - Exploiting results into Symbolic description and ontology to play the role of expert system
  - New version of the Km4City ontology integrating more detailed digital twin aspects.
  - LLM for DSS solution generation in the context of Smart City advisor