Thesis and Stage Opportunities AY23/24

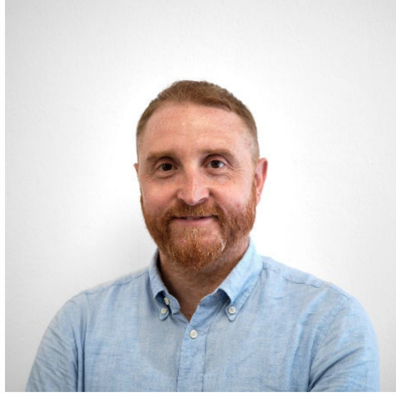
Giacomo Boracchi, Politecnico di Milano, DEIB

https://boracchi.faculty.polimi.it/

Updated on September 2023

The Team

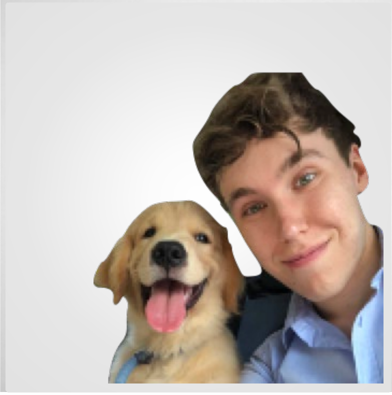
We are 3 faculties, 7 PhD students, 2 Research Assistants (Assegnista di Ricerca)


giacomo.boracchi@polimi.it


luca.magri@polimi.it


federica.arrigoni@polimi.it


filippo.leveni@polimi.it


andrea.porfiridalcin@polimi.it


antoninomaria.rizzo@mail.polimi.it


loris.giulivi@polimi.it


michele.craighero@polimi.it


giuseppe.bertolini@polimi.it


riccardo.margheritti@polimi.it


edoardo.peretti@polimi.it


roberto.basla@polimi.it

Research Topics

Deep Learning & Image Processing

- Detection and Visual Recognition
- Neural Networks for 3D vision
- Adversarial Attacks

Computer Vision

- Object detection / template matching
- Multi-view geometry
- Robust Model Fitting
- 3D reconstuction & Structure from Motion
- Quantum Computer Vision

These slides…

…provide both:

- A short overview of the research background (**Background** slides) together with some references that you should check
- Possible research directions along which to develop a thesis under our supervision (**Research directions** slides)
- Research institutions / companies involved in these research activities

Thesis Types:

- **Research thesis:** for those of you willing to investigate some new problem, discover something new. *Enthusiasm, talent, sense of initiative, willingness to learn and try.*
- **Thesis on industrial research projects:** for those of you looking for a direct and full involvement in a reserch project with the industry. You will work under our strict supervision. *Dependability, talent, rigour.*
- **Internships:** for those of you looking for an experience in a R&D company. Still, we regularly collaborate with these companies.

What do we offer

- Our supervision + co-supervision by a PhD student / colleague of ours
- General guidance on «how to address thesis work». Meetings with your collegues concerning:
    - Thesis management
    - Related work
    - Problem Formulation
    - Thesis / Executive Summary Preparation
    - Thesis Presentation

- Every now and then we organize also meetings to illustrate research activities in the group.

If you feel like a brilliant candidate…

… and you are considering pursuing a PhD related to our research topics…

… and you have a wonderful idea / some special interest for a research topic of ours…

Just drop me an email and we can arrange a short meeting to discuss whether there is room for an additional research thesis

giacomo.boracchi@polimi.it

How to apply for a thesis

Fill in the form

https://forms.gle/XkvosV64kngwUiy38

Remember to tell us:

- Which thesis are you interested in and which are **«first options»**
- What is your background, which courses you have passed that are inherent to the thesis you selected.
- Mention also:
    - any information you deem relevant for working on that specific thesis.
    - successful projects and your programming expertise.
- A short motivation paragraph.

Deadline to apply for a thesis

We will start thesis assignment on Wednesday September 13th

To be eligible for a thesis, submit your proposal by that time, even though we might leave the form opened until the end of that week.

You will receive an email from us to schedule a meeting (possibly in person, but if not possible online) right after the deadline.

Before the deadline, you are welcome to write us email for enquiring

Deep Learning and Image Processing

DL1: Multi-view object detection CNN

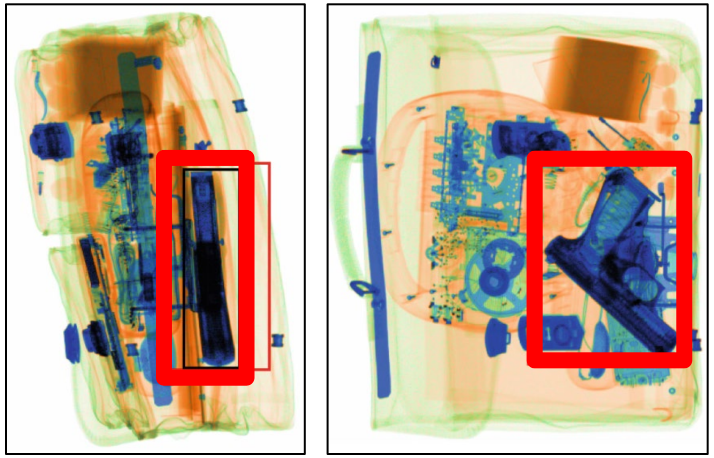
Background

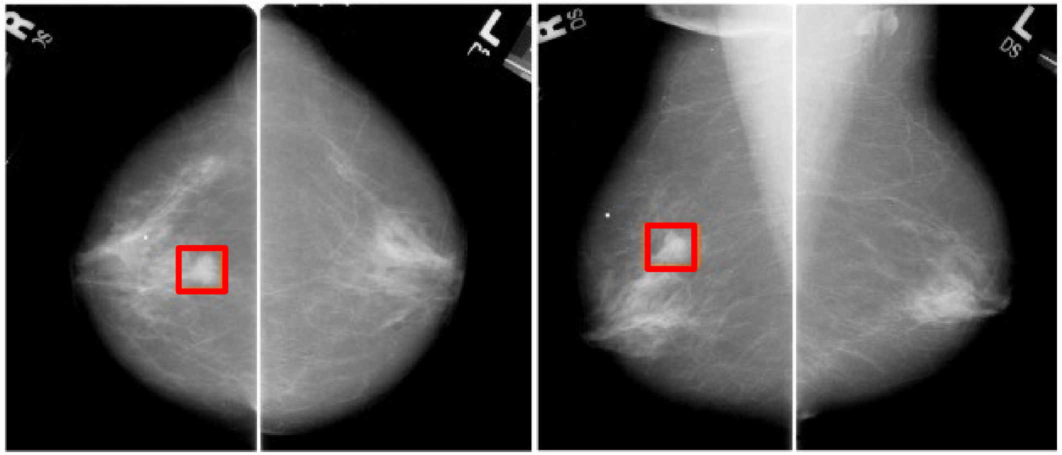
Detection systems usually take only a single input image. In some applications, such as

- Baggage inspection in airports
- Tumor detection in X-ray images

you have multiple views of the same objects.

Exploiting multiple-views can increase detection power.





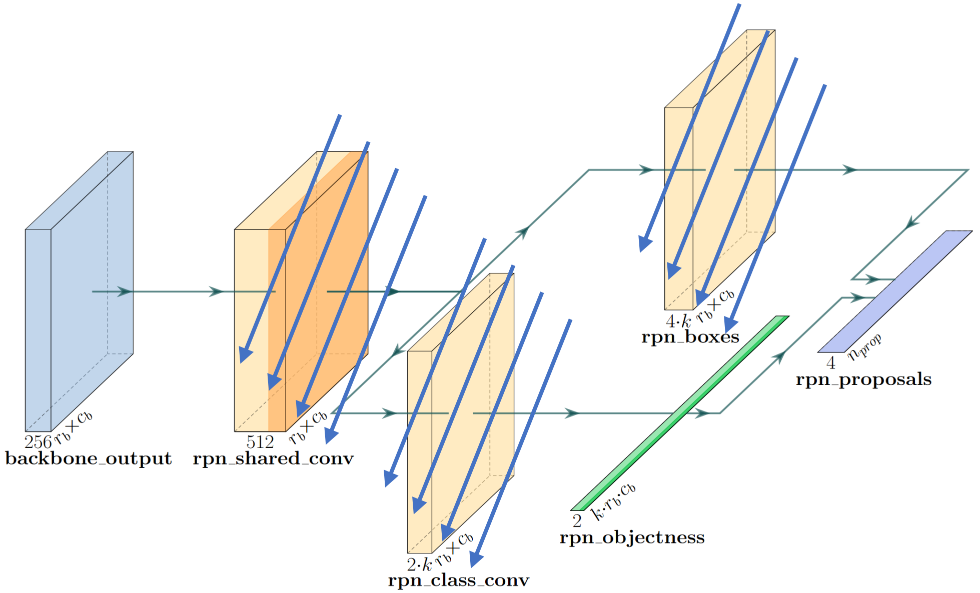
[1] Steitz et al., "Multi-view x-ray r-cnn." (2018)
[2] Wimmer et al. "Multi-task fusion for improving mammography screening data classification" (2021)

Research Directions

DL1: Multi-view object detection CNN

Goal

- Define a novel CNN architecture to jointly analyze multiple views containing the same target to be detected
- Train the multi-view CNN to detect tumors in mammography images



Requirements: AN2DL, Python programming

Materials and Methods: Databases for Screening Mammography, access to GPUs and Faster-RCNN framework.
Possible collaboration with Niguarda Radiography Division.

Start: asap **Duration:** 9 – 12 months

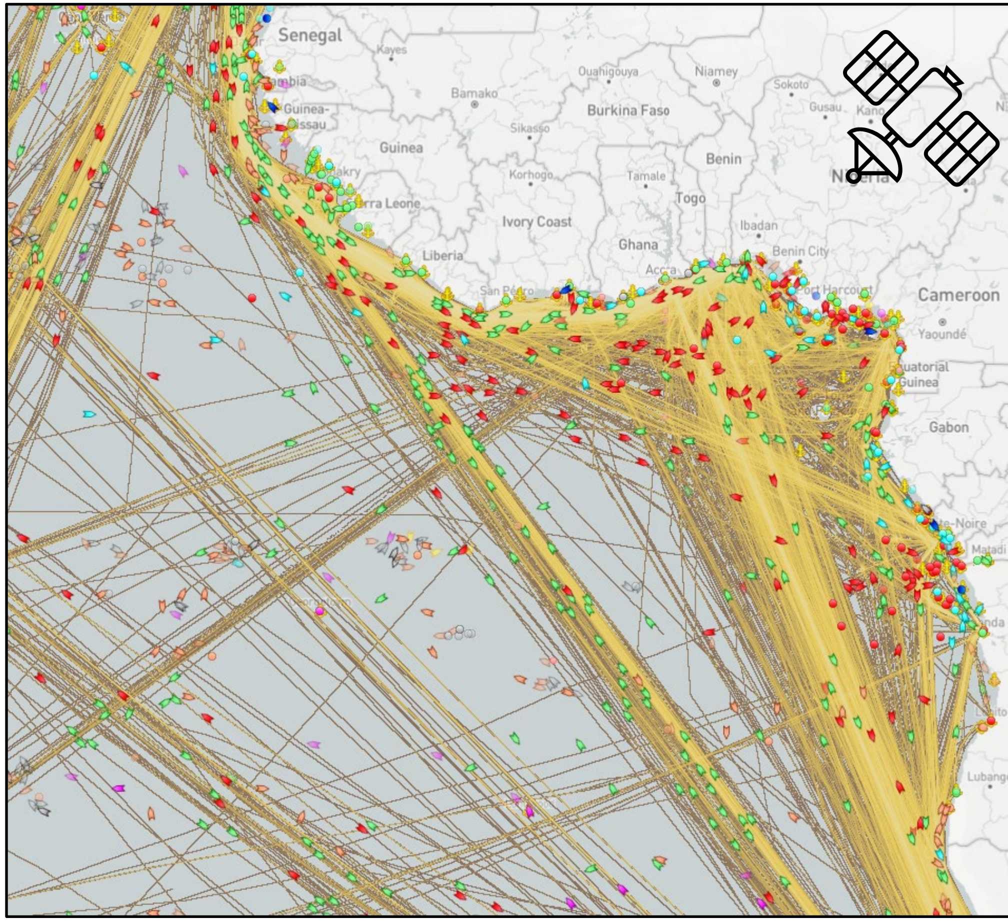
Basla, Arrigoni, Magri, Boracchi

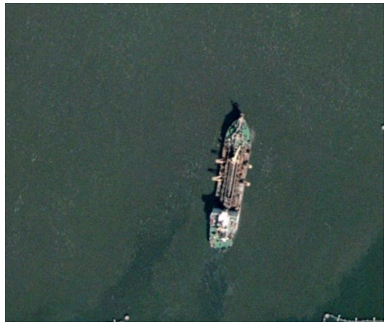
DL2: Dark Vessels Detection and Tracking in Satellite Images

Background

Transnational trafficking of e-waste is an environmental crime where cargo ships deliver containers of e-waste to an illegal dumpsite in poor countries.

These ships often turn off their transponder to avoid being tracked, becoming "*dark vessels*" that cannot be tracked.





PERIVALLON

EU Project

[1] Ciocarlan et al., "Ship Detection in Sentinel 2 Multi-Spectral Images with Self-Supervised Learning" (2018)
[2] Kizilkaya et al. "VHRShips: An Extensive Benchmark Dataset for Scalable Deep Learning-Based Ship Detection Applications" (2022)

Basla, Boracchi

Research Directions

DL2: Ship segmentation and course prediction in satellite images

Idea: combine vessel route prediction and satellite route query to infer when the vessel gets visible by a satellite. Detect vessels in satellite images and identify these by estimating the area

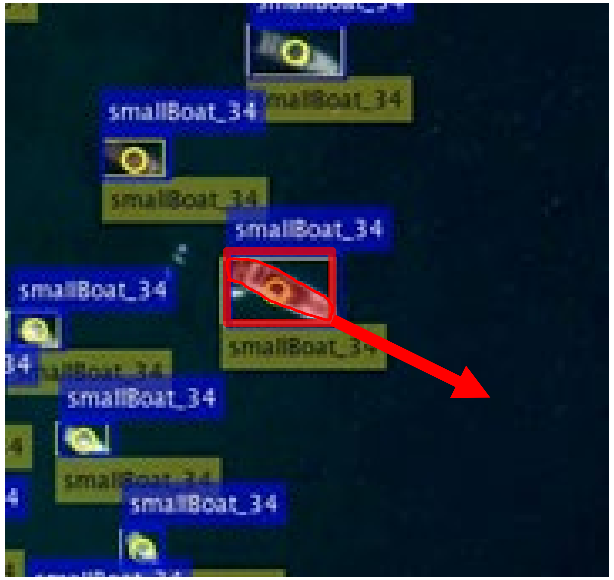
Goal

- Train a CNN to detect ships in satellite images and estimate their area via instance segmentation / pixel clustering
- Predict their trajectories and intersect them with a Sentinel satellite pass

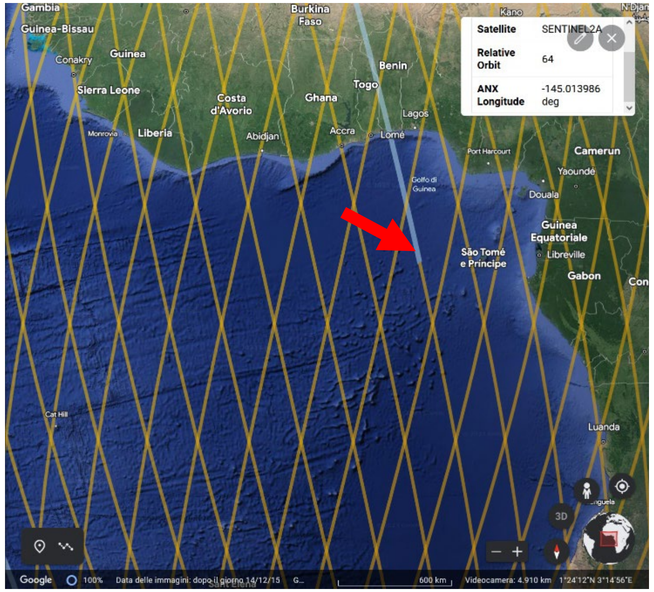
Requirements: AN2DL, Python programming

Materials and Methods: Satellite imagery datasets, access to GPUs

Start: now **Duration:** 9 – 12 months



Trajectory forecasting



Satellite orbits

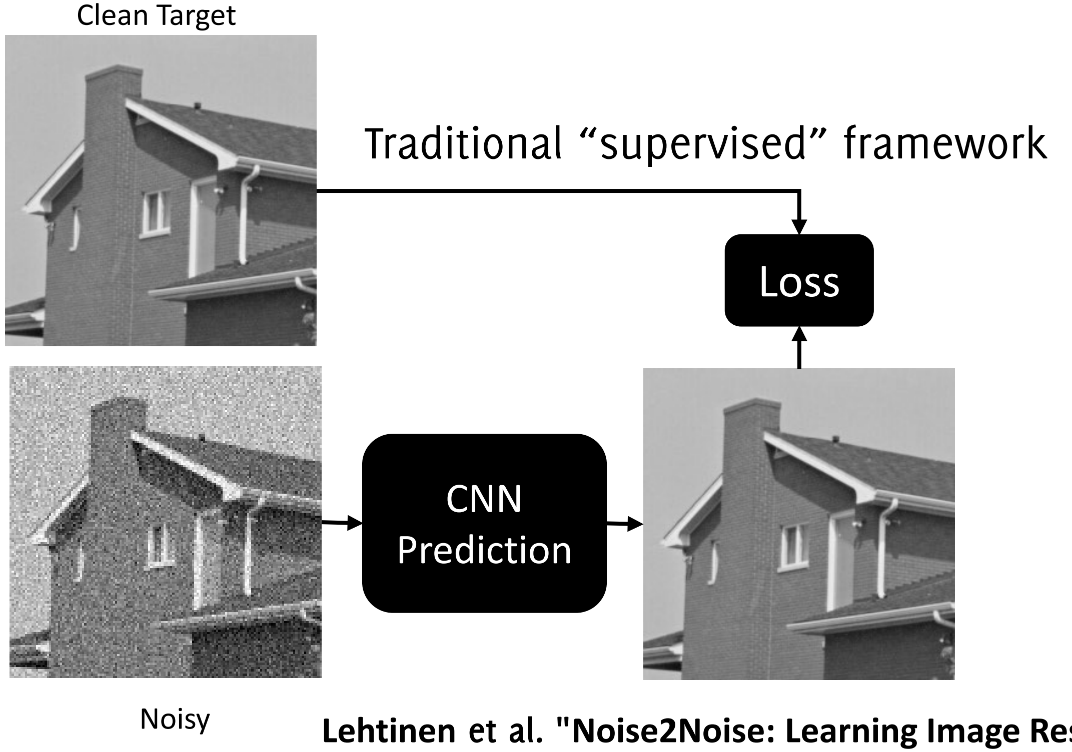
Basla, Boracchi

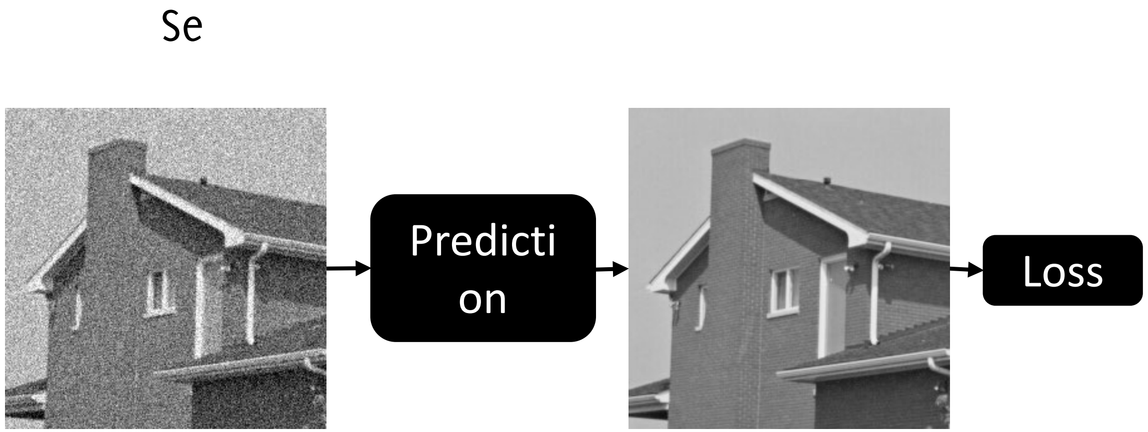
DL3 Non-local Self* Denoising Networks.

Background

Image denoising require **many clean-noisy image pairs for training.** In practice, when the noise characteristics are complex, collecting such pairs might be unfeasible.

The challenge is to train denoisers from a **single** image, as in many cases **clean-noisy pairs are not available** (e.g., microscopy).





Lehtinen et al. "Noise2Noise: Learning Image Restoration without Clean Data" (ICML 2018)

DL3 Non-local Self* Denoising Networks.

Self-similarity in Natural Images

Research Directions

Goal: Design a novel framework for training a deep image denoiser for real images that is

- self-supervised
- promotes non-local self-similarity

Requirements: AN2DL, Python programming

Materials and Methods: access to GPUs

Start: asap **Duration:** 9 – 12 months

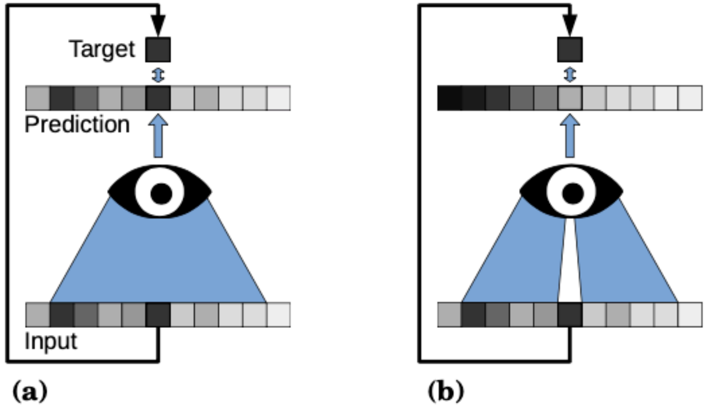


Figure 2: A conventional network versus our proposed blind-spot network. **(a)** In the conventional network the pre-

Krull et al. "Noise2void-learning denoising from single noisy images" (CVPR 2019)
Batson et al. "Noise2self: Blind denoising by self-supervision" (CVPR 2019)
Quan et al. "Self2self with dropout: Learning self-supervised denoising from single image" (CVPR 2020)

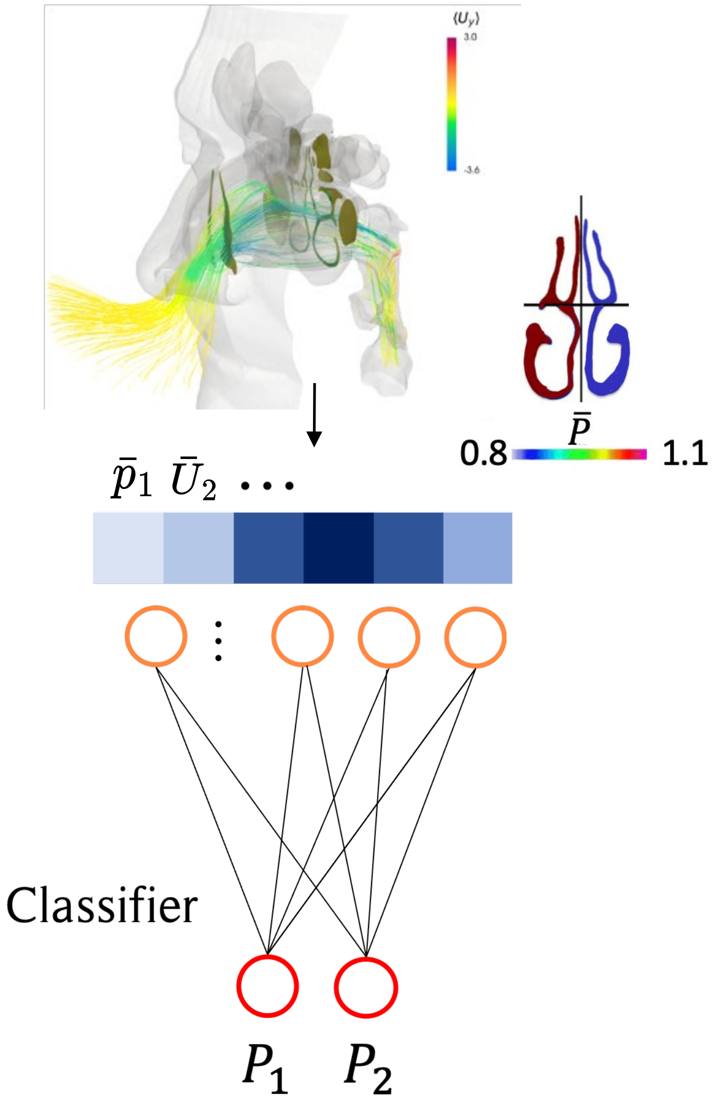
Peretti, Boracchi

DL4. Pathology identification from CFD data

Background

Computational Fluid Dynamics (CFD) provides unique information for diagnosing complex systems. Our work, in collaboration with DAER, consists in classifying pathologies in human upper airways, by CFD flow.

Annotated CFD data are difficult to gather, and the large size of CFD flow data usually requires a **feature extraction process** to reduce input dimensionality before feeding it into a NN.



[1] Schillaci; Quadrio; Pipolo; Restelli; Boracchi «*Inferring functional properties from fluid dynamics features*». 25th International Conference on Pattern Recognition (ICPR), 2020.

Margheritti, Boracchi

Research Directions

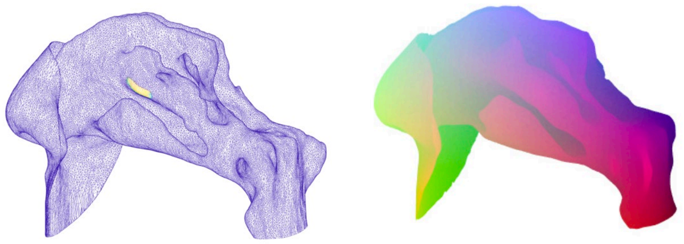
DL4. Pathology identification from CFD data

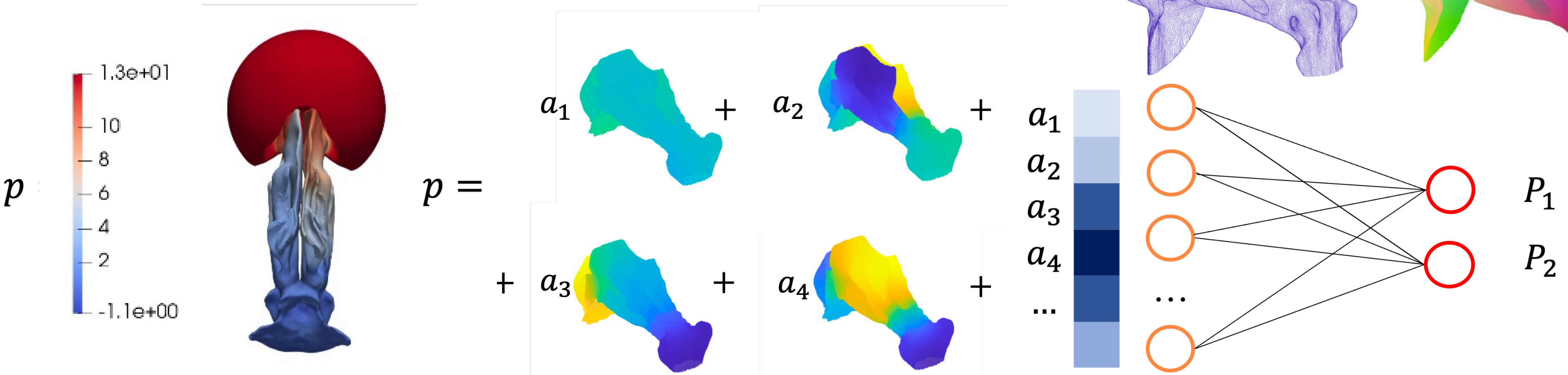
Goal: Explore **DL techniques** to deal with surface data, using 3D NN architectures and **self-supervised learning**. Leverage **functional mapping** to decompose CFD surface data in spectral coefficent to be fed to the NN.

Requirements: Python/Matlab programming

Materials and Methods: access to GPUs/ CINECA, flow fields database.

Start: asap **Duration:** 9 – 12 months





[2] M. Ovsjanikov, M. Ben Chen, J. Solomon, A. Butscher, and L. Guibas. Functional maps: A flexible representation of maps between shapes. ACM Transactions on Graphics, 2012.

Margheritti, Boracchi

DL5 Instance Segmentation and Uncertainty Assessment in Medical Images

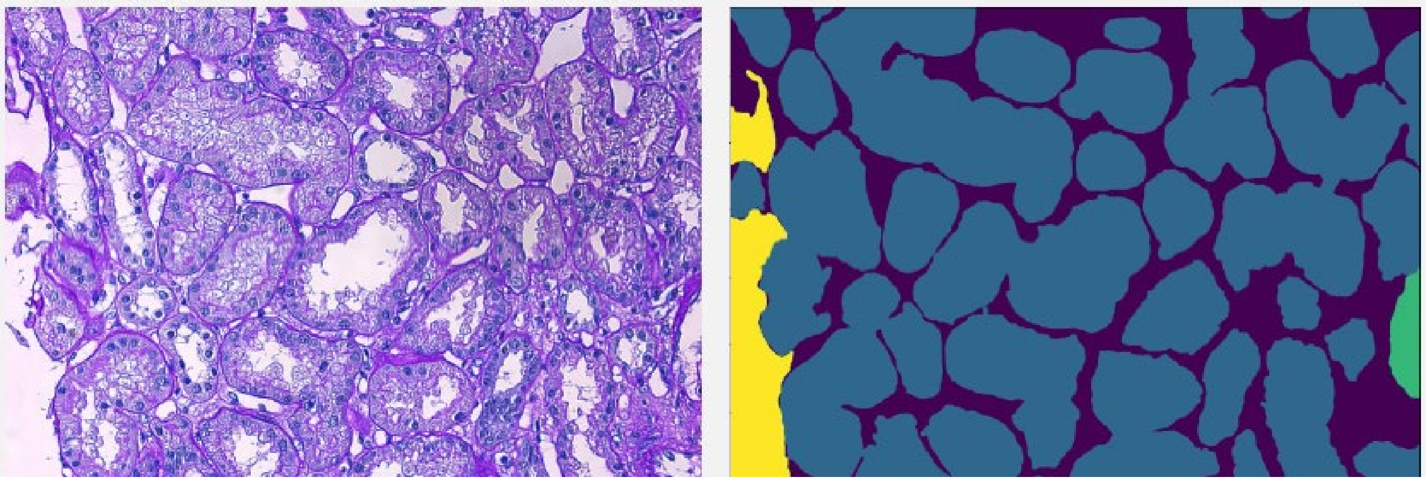
Background

Instance segmentation networks are key in biomedical images to assess the state of a pathology in tissues/organs

The major limitations are:

- **annotations are expensive** and required for each type of diseases
- **predictions are not trustworthy.**

In collaboration with Mario Negri designed methods for training a segmentation network using scribble-like annotations



Localize regions belonging to cells or to interstitium in kidney's tissue. Dense annotations for training a Segmentation Network

UNIVERSITÀ DEGLI STUDI DI BERGAMO

ISTITUTO DI RICERCHE FARMACOLOGICHE MARIO NEGRI · IRCCS

Boracchi

Research Directions

DL5 Instance Segmentation and Uncertainty Assessment in Medical Images

List of potential goals:

- Train SoTA segmentation NN for histological image using sparse annotations
- Quantitatively assess the uncertainty of each pixel using Bayesian NN
- Interact with **experts to design realistic** and efficient data-augmentation procedures
- Use uncertainty and sparse annotations in an "active learning" framework, where you query annotations for selective images where uncertainty is large

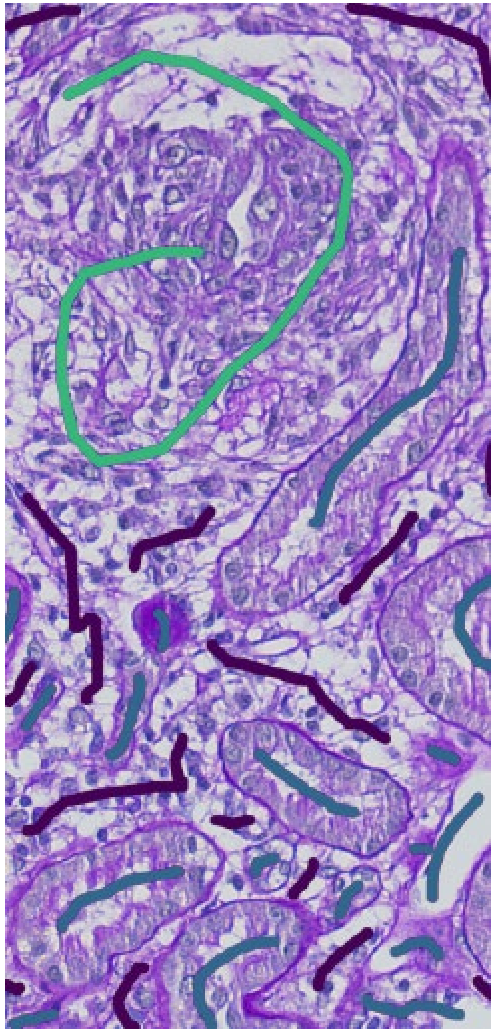
Materials and Methods:

- Annotated histological datasets and demo codes
- Access to a server mounting GPUs for training networks once stable.

Required background: AN2DL, Python

Thesis start: ASAP

Scribble annotations, fast and easy to gather, good to fine tune models

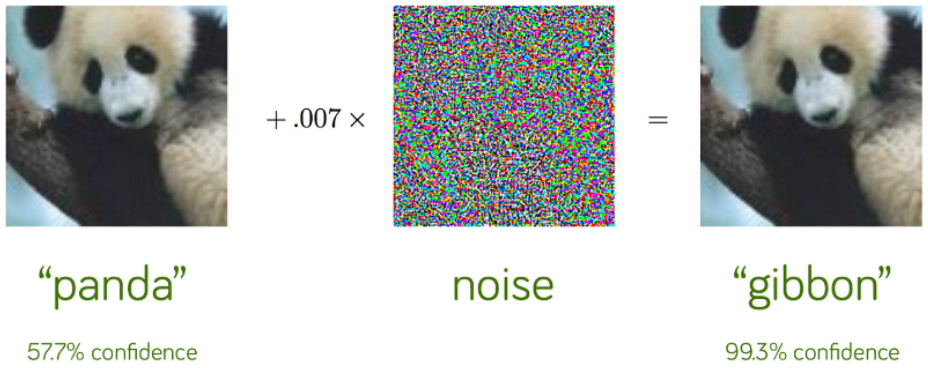


Boracchi

DL6 Adversarial Clothes

Background

- Adversarial attacks are small perturbations that change the model's prediction.
- Most attacks are constrained to **digital images**, and as such can't be physically deployed.
- We have recently proposed Adversarial Scratches [1], as a new attack that instead is designed to be deployable.



Example adversarial attack: A small perturbation can make a big difference

[1] Giulivi, L., Jere, M., Rossi, L., Koushanfar, F., Ciocarlie, G., Hitaj, B., & Boracchi, G. [Adversarial Scratches: Deployable Attacks to CNN Classifiers](#) Pattern Recognition 2023

Giulivi, Boracchi

Research Directions

DL6 Adversarial Clothes

Privacy as an application of Adv. Attacks. Adversarial clothes can be designed to deceive object localization and person re-identification systems.

Goals:

- Study efficient parametric attacks and optimization strategies that can be deployed to garments.
- Design garments to prevent human detection and also manipulation using generative models

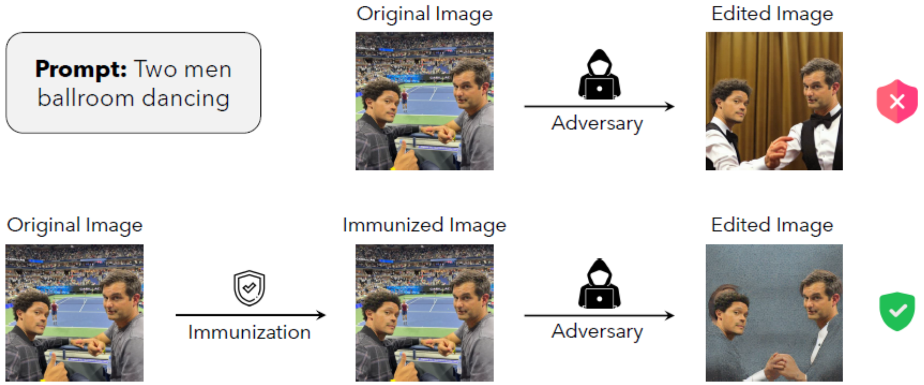
Thesis is in collaboration with [cap_able design].

Materials and Resources

You will have access to GPU powered machines to run experiments.

A flexible framework for generation and evaluation of adversarial attacks will be provided.

An example of an adversarial sweater, which is able to fool a TinyYolo model in a real-world setting.



Salman, Hadi, et al. "Raising the cost of malicious ai-powered image editing." *arXiv preprint arXiv:2302.06588 and* ICML 2023
Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

Giulivi, Boracchi

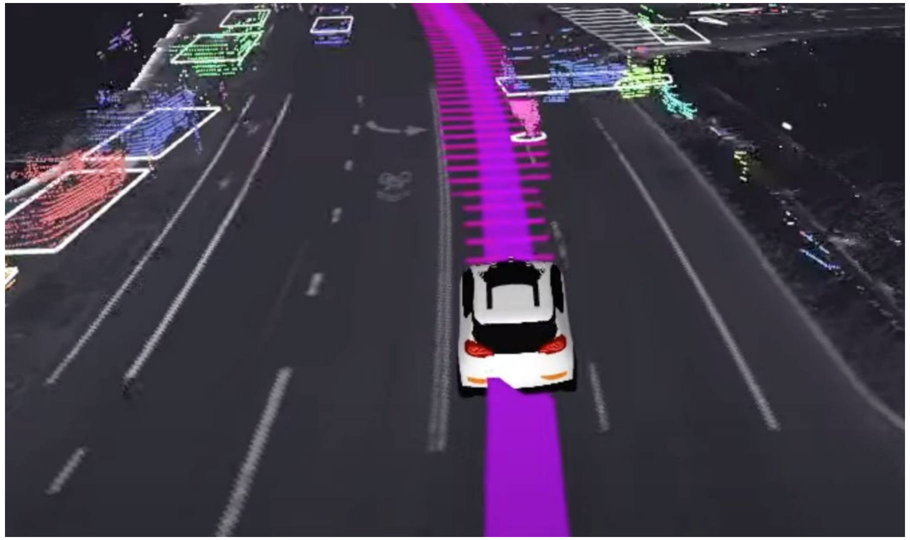
DL7. Non-Rigid 3D Reconstruction

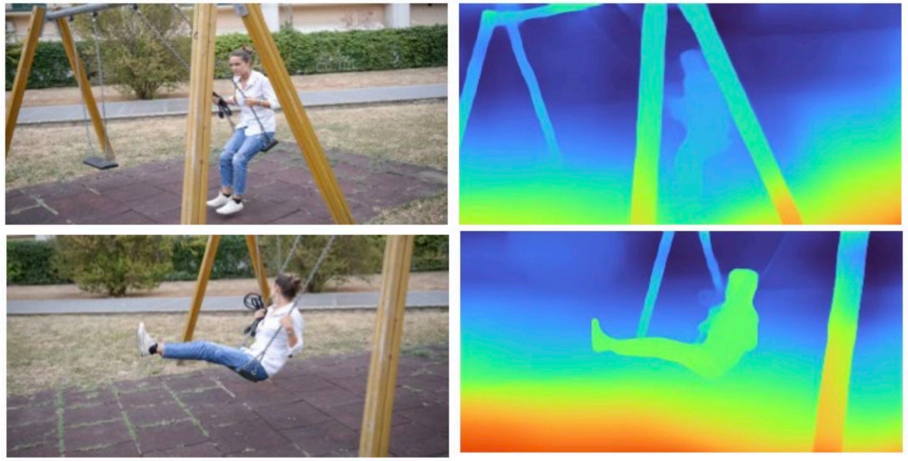
Background

Recent progress in Deep Learning and Computer Vision has made it easy to **reconstruct static environments**, but the world is hardly static in practice.

Moving and deformable objects are ubiquitous and need to be reconstructed in 3D to allow robots and autonomous vehicles to operate safely and reliably.

Also, many Augmented Reality applications **may not work due to 3D reconstruction algorithms not supporting moving objects.**





Dal Cin, Magri, Boracchi

DL7. Non-Rigid 3D Reconstruction

Goal

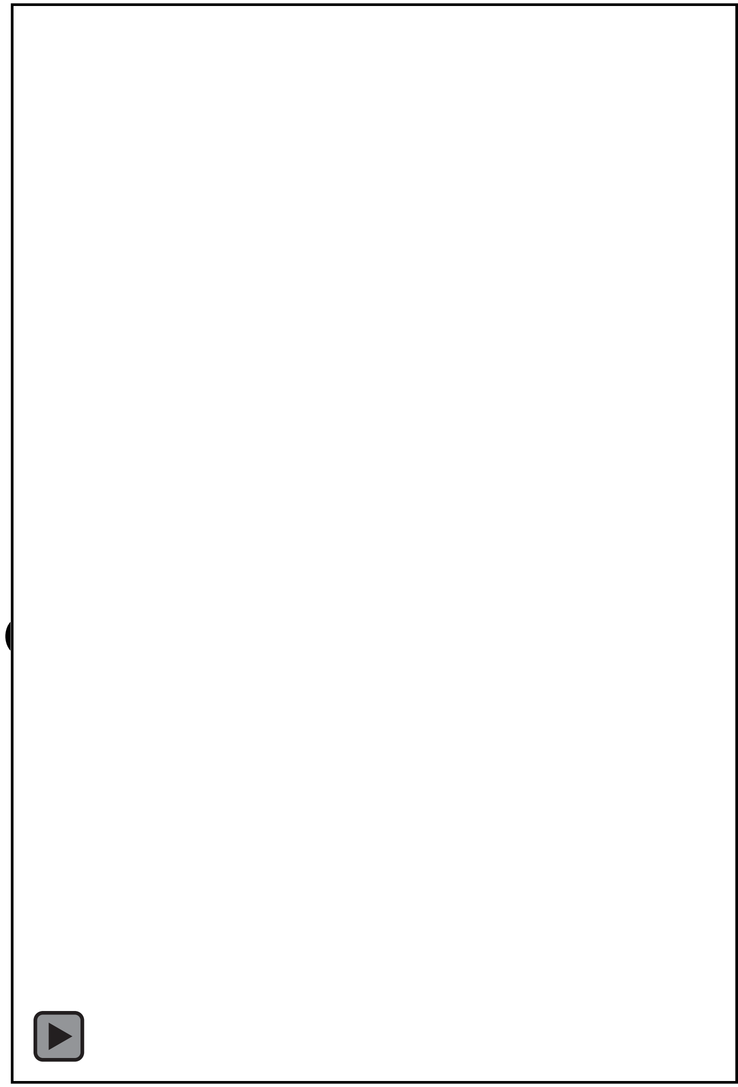
- Investigate existing approaches to non-rigid 3D reconstruction and identify limitations.
- Develop novel non-rigid 3D reconstruction algorithms combining Deep Learning and geometric multi-view constraints.

Requirements: Knowledge of Deep Learning (AN2DL), Pytho PyTorch or TensorFlow.

Preferable but not mandatory: 3D Computer Vision (IACV).

Materials and Methods: Multi-GPU access. Dataset. Depth cameras available

Start: now **Duration:** 9 – 12 months



Background

DL8. Video Object Segmentation (VOS)

The task of video object segmentation is to produce accurate and temporally consistent pixel-wise masks of **moving objects** of interest in video sequences.

[1] Zhou et al. "A survey on deep learning technique for video segmentation". IEEE TPAMI 2023

Research Directions

DL8. Video Object Segmentation (VOS)

Goal: Derive new solutions for video object segmentation, addressing the open challenges such as

- Strong viewpoint changes (deformations)
- Appearance changes (illumination)
- Motion blur and scale changes
- Occlusions and interacting objects
- Real-time requirements

Requirements: AN2DL, Python programming

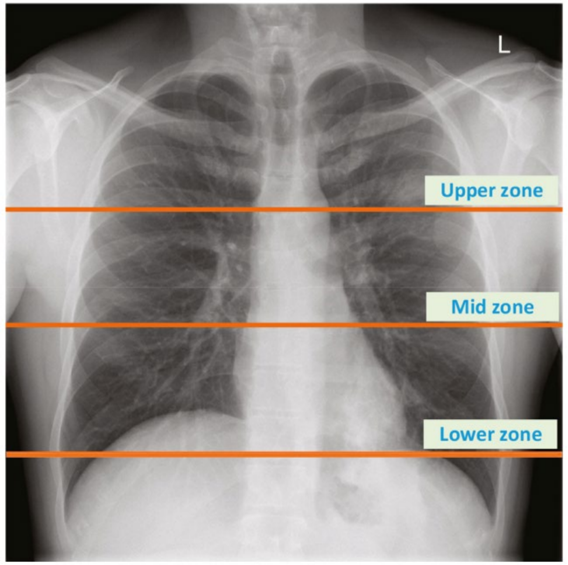
Materials and Methods: Access to GPUs

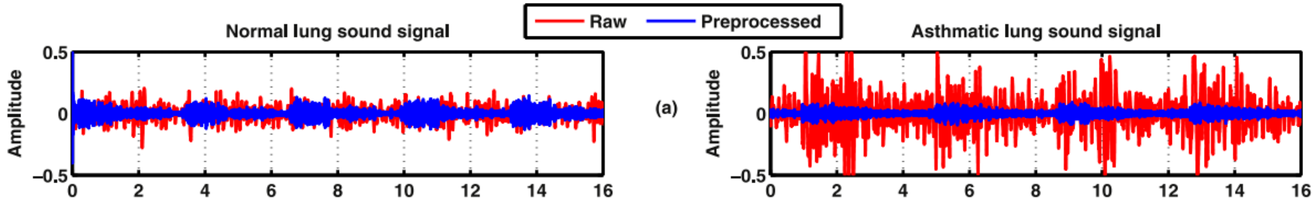
Rizzo, Arrigoni, Magri, Boracchi

DL9. Asthma detection from lung sound signals

Background

- The automatic analysis of respiratory sounds has been a field of great research interest during the last decades.
- However, DL techniques have not yet been applied systematically to automatically analyze lung sound signals





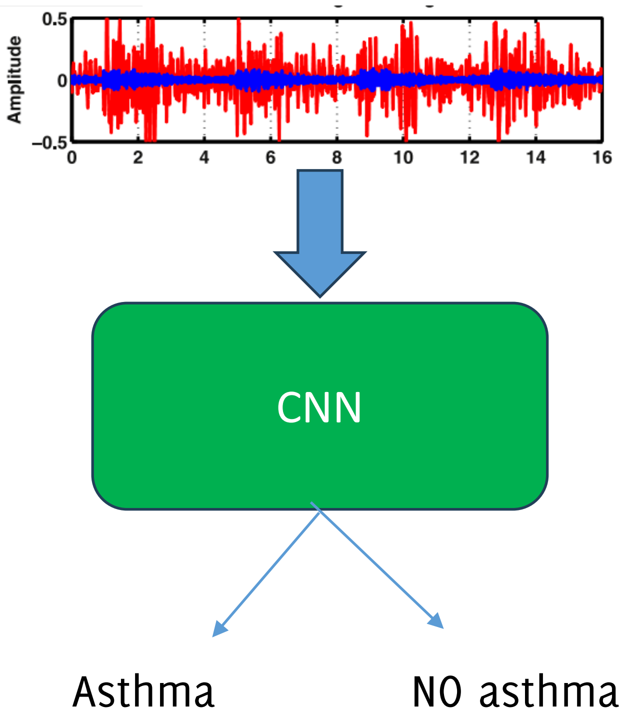
Craighero, Boracchi

DL. Asthma detection from lung sound signals

Research Directions

Goal: Explore DL techniques to automatically analyse lung sound signals to detect an episode of asthma.

- Address this as a classification problem
- Address localization/explainability of the network outcomes



Requirements: AN2DL, Python programming

Materials and Methods: access to GPUs, public datasets [1, 2], possible contact with a doctor and an asthmatic patient

Start: now **Duration:** 9 – 12 months

[1] https://data.mendeley.com/datasets/jwyy9np4gv/3
[2] https://www.kaggle.com/datasets/vbookshelf/respiratory-sound-database

Craighero, Boracchi

Computer Vision
& Pattern Recognition

CV1. Geometric Aware Camera Pose Estimation

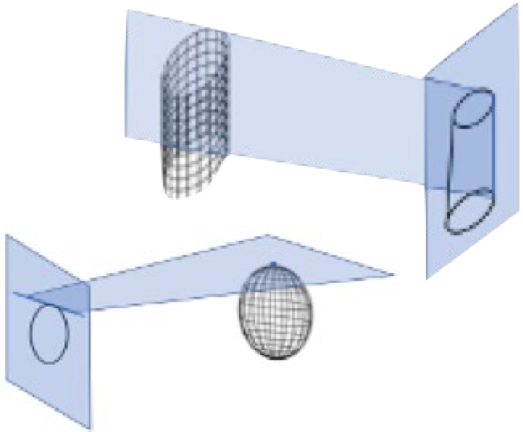
Background

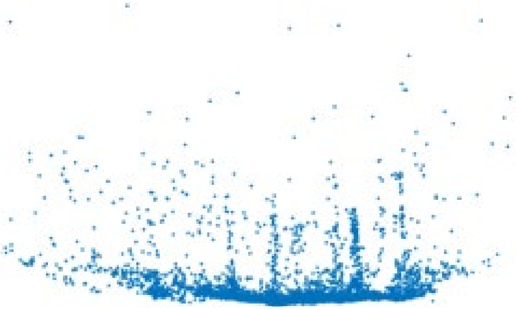
A fundamental problem in 3D Vision is the estimation of the location and orientation of a camera in a complex 3D environment.

Traditional solutions leverage **3D points** in the scene, which are onerous to triangulate, nosiy and difficult to track.

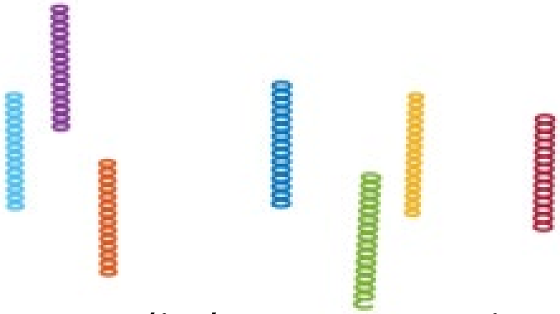
The challenges are to

- use **other geometric primitives** that provide richer cues
- make the processing very **efficient** to met the typical real-time requirements of Augmented Reality applications





Point-wise reconstruction



Cylinder reconstruction

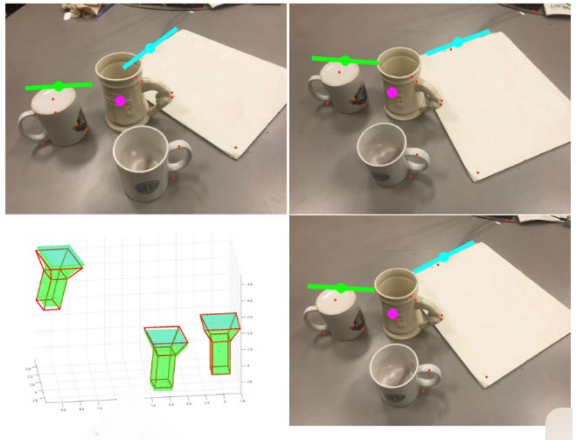
Dal Cin, Magri, Boracchi

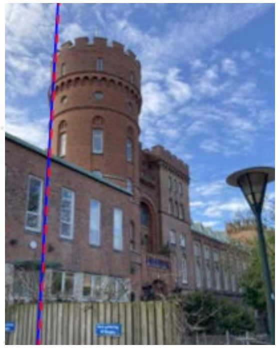
CV1. Geometric Aware Camera Pose Estimation

Research Directions

Goal

- Design and compare Camera Pose Estimation algorithm that exploit geometric primitives (2D lines conic, 3D planes quadrics)
- Implement efficient primitive fitting and tracking algorithm. Memory requirement is crucial for AR applications.
- Deploy algorithms on GPU hardware and (if interested).





Preferable but not mandatory: 3D Computer Vision (IACV).

Materials and Methods: Multi-GPU access. Depth cameras available.

Start: October **Duration:** 9 – 12 months

[1] "TRPLP. Trifocal Relative Pose from Lines at Points " CVPR 2020.
[2] Gummeson, Anna et al. "Fast and efficient minimal solvers for quadric based camera pose estimation." *ICPR 2022*

Dal Cin, Magri, Boracchi

CV2. Shape Abstraction via primitive decomposition

Background

Challenges in 3D Data Analysis:

- Impractical representation
  - Meshes and point clouds are inefficient in terms of storage
  - Lack of standard grid/template representation that fits well within deep learning architectures
- Lack of annotated data
- Variability and diversity in real data (shapes, sizes, orientations, viewpoints, occlusions, noise, outliers..)

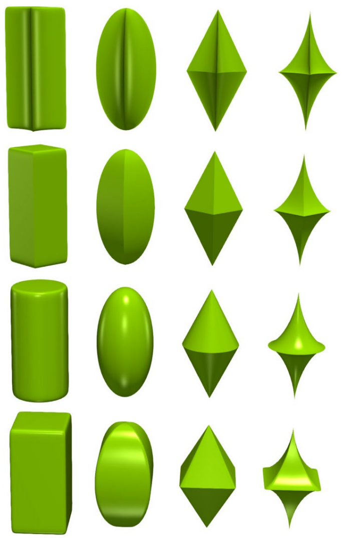
There is a need for a more compact and abstract representation for tackling high level tasks as object retrieval/recognition and classification

Bertolini, Magri

Research Directions

CV2. Shape Abstraction via primitive decomposition

superquadrics



Goal: - Design superquadric fitting algorithms to provide compact representation of noisy data

- Test the primitive decomposition in applications (e.g., shape recognition, retrieval, registration, collision detection)

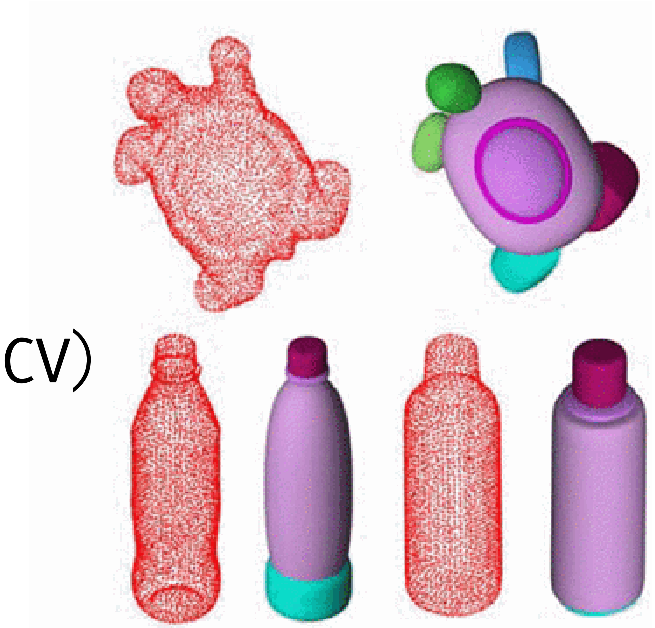
Materials and methods:

- Large real datasets contaminated by noise and artifacts
- Depth cameras to acquire 3D point clouds are available.

Requirements: Python programming

Prefferred but not mandatory; Basics of Computer Vision (IACV)

Start: now, **Duration:** 9 – 12 months.



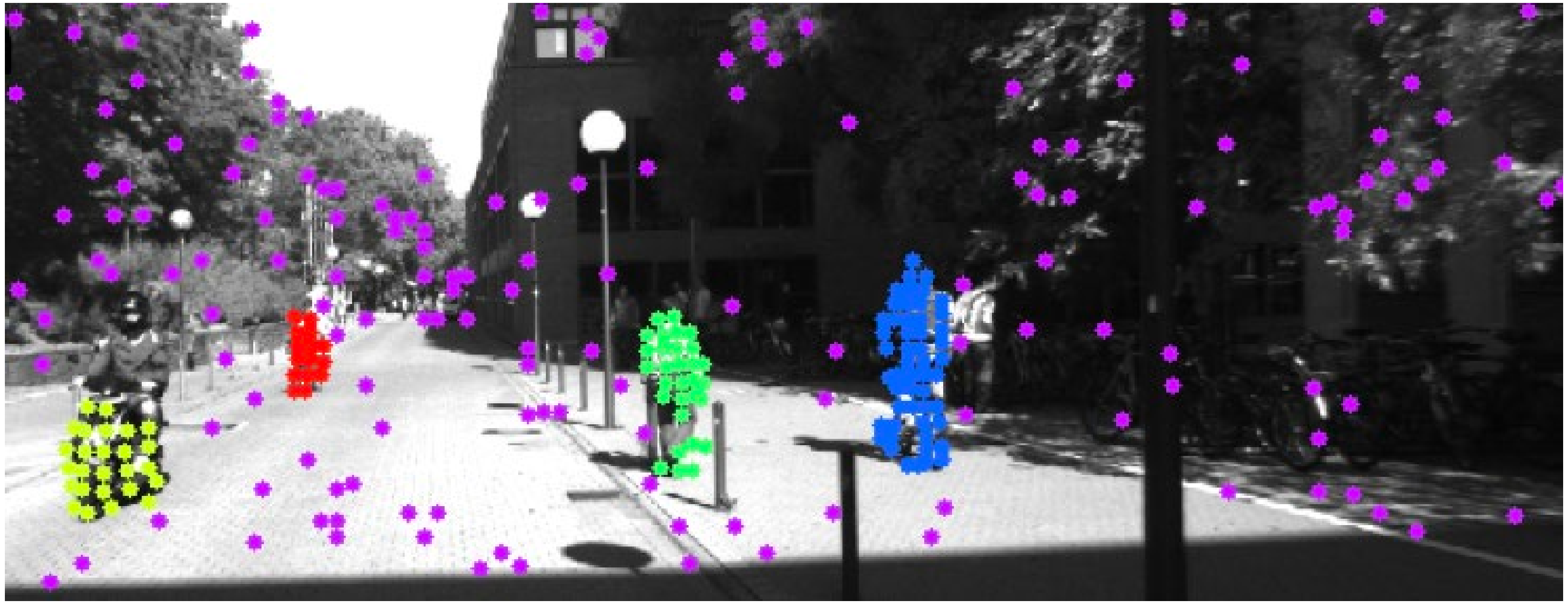
Bertolini, Magri

[1] Liu, Weixiao, et al. "Robust and Accurate Superquadric Recovery: a Probabilistic Approach"
[2] Liu, Weixiao, et al. "Marching-Primitives: Shape Abstraction from Signed Distance Function."

CV3. Geometric Motion Segmentation

Background

Motion segmentation refers to the problem of identifying multiple moving objects in a collection of images. This is typically cast to a **clustering problem** where sparse correspondences belonging to the same moving object should be grouped together.



Arrigoni, Magri

CV3. Geometric Motion Segmentation

Research Directions

Goal: deal with the challenges of dynamic scenes ubiquitous in real applications (e.g. autonomous driving)

- **degenerate motions**: pure translational or rotational motions
- **outliers**: mismatches in the motion trajectories

Materials & Methods: geometric models

Preferable but not mandatory: Knowledge of 3D Computer Vision (IACV).

Research thesis: 9-12 months

Start: now

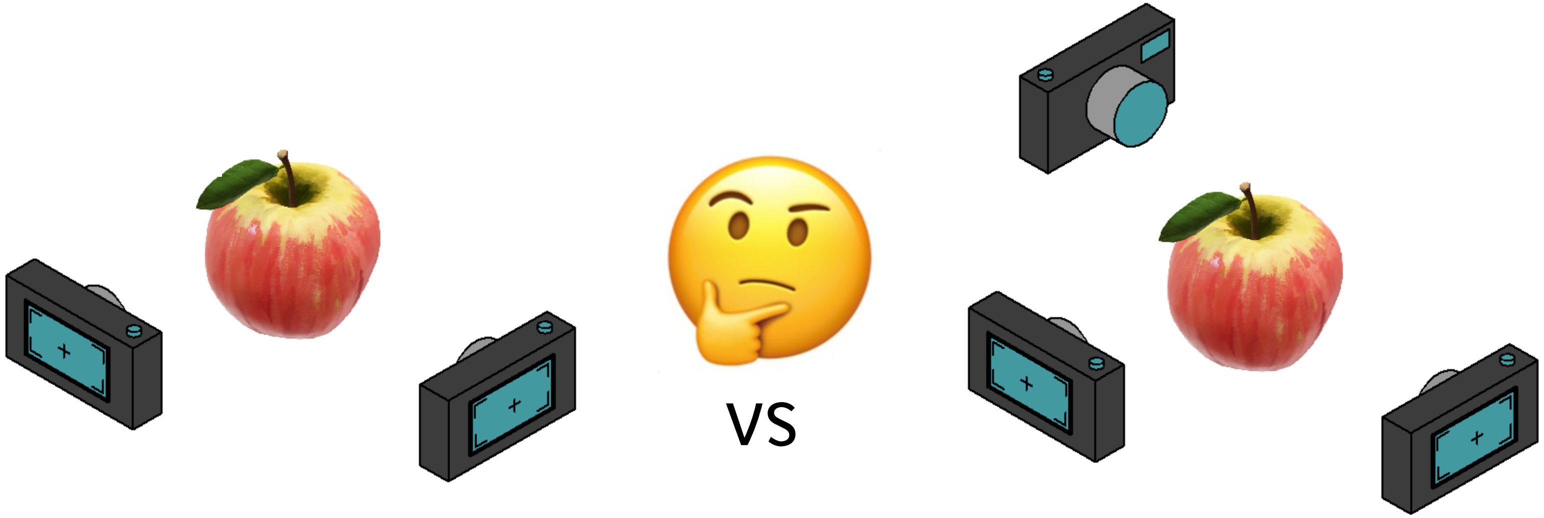
[1] Xu, Cheong & Li. "Motion Segmentation by Exploiting complementary geometric models". CVPR 2018, https://alex-xun-xu.github.io/ProjectPage/CVPR_18/
[2] Tron & Vidal. “A Benchmark for the comparison of motion segmentation algorithms”. CVPR 2007, http://www.vision.jhu.edu/data/

Arrigoni, Magri

CV4. Three views or two views?

Background

The **3D reconstruction** of a scene can be performed starting from multiple two-view reconstructions. However, when we have **3 views**, more stringent geometric constraints can be imposed to reconstruct the scene.



Commonly used in practice
Ambiguity in the correspondences

Not used in practice
Unambigous

[1] Hartley & Zissermann. Multiple view geometry in Computer Vision. Cambridge University Press, second edition, 2004

Arrigoni, Magri

Research Directions

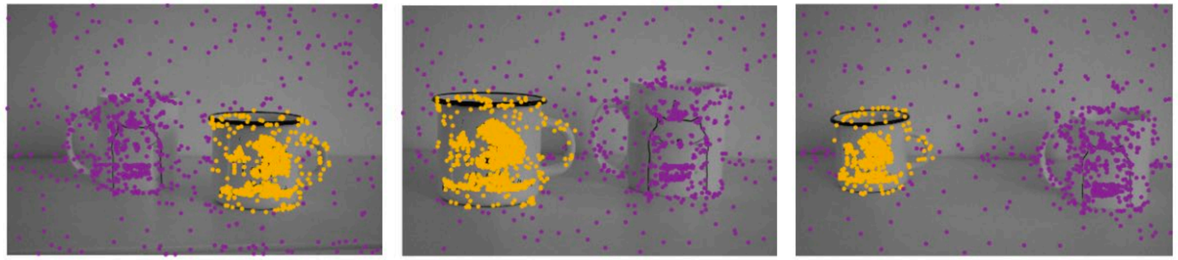
CV4. Three views or two views?

The goal of the thesis is to to establish whether in practical applications (pose estimation, motion segmentation) it is ***better*** to use a solver based on 2 views or 3 views.

Research thesis: 9-12 months

Requirements:

- Basics of Computer Vision
- Basics of Python/Matlab programming



Start: now

In collaboration with Czech Technical University in Prague

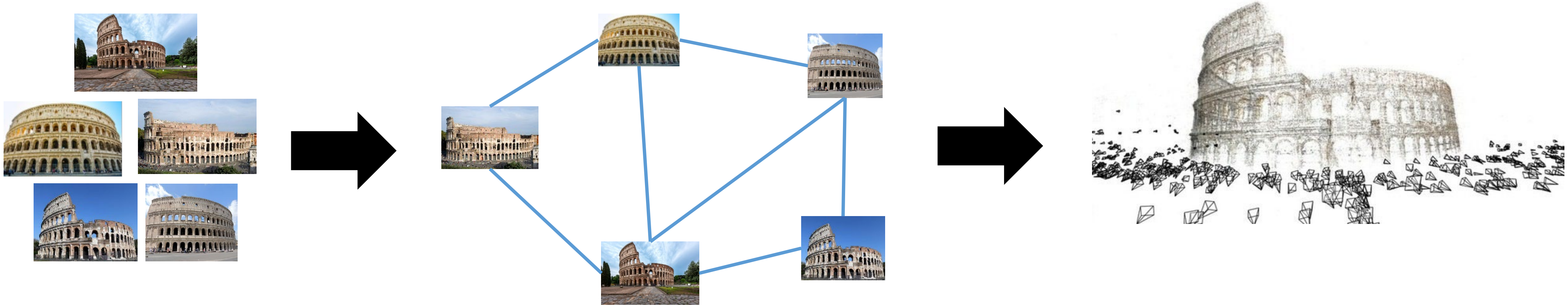
CTU
CZECH TECHNICAL UNIVERSITY IN PRAGUE

Arrigoni, Magri

CV5. Graphs in Multi-view Geometry

Background

Representing cameras/images and their pairwise relations as vertices and edges of a **graph**, it is a powerful tool to study and solve relevant problems in multi-view geometry.



[1] Arrigoni & Fusiello "Synchronization problems in computer vision with closed-form solutions". IJCV 2020
[2] Arrigoni, Fusiello, Ricci & Pajdla. "Viewing graph solvability via cycle consistency." ICCV 2021

Arrigoni

CV5. Graphs in Multi-view Geometry

Research Directions

Goal: develop new methods and/or theory to solve relevant multi-view problems by exploiting a graph representation:

- *Solvability:* detect solvable/unsolvable 3D reconstruction problems
- *Synchronization*: enforce global consistency to address relevant applications (3D reconstruction, 3D registration, image matching, …)

Materials & Methods: geometric models, possibly algebraic geometry

Requirements: Knowledge of 3D Computer Vision (IACV).

Research thesis: 9-12 months

In collaboration with University of Udine

UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura

Arrigoni

CV6. Explainable Clustering

Background

Machine learning models are mostly “black box”, but these days, are entering in fields where performance is not the only objective (e.g healthcare, transportation, or law).

Can you imagine blindly performing a surgery on a patient just because a computer said so? It is essential to provide insight into what parts of the data the algorithm used to make its prediction.

The aim of this thesis is to develop explainable clustering algorithms.

Leveni, Magri, Boracchi

CV6. Explainable Clustering

Research Directions

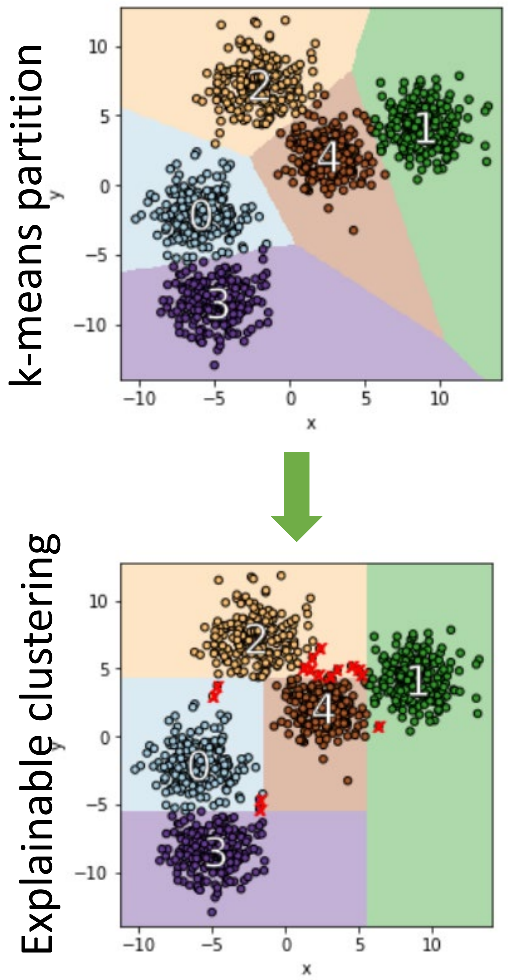
Goal: Study new clustering algorithms whose output are easily interpreted by experts:

- Characterize normal data by injecting expert knowledge
- Derive automatic methods for feature selection
- Analyze the tradeoff between explainability and clustering accuracy

Materials and Methods:

- Preliminary study available,
  server mounting **GPUs will be provided**

Start: now, **Duration:** 9 – 12 months.



[1] Frost et Al. ExKMC: Expanding Explainable k-Means Clustering
[2] Moshkovitz et Al. Connecting Interpretability and Robustness in Decision Trees
[3] [Introduction to Explainable K-means clustering on yt]

Leveni, Magri, Boracchi

Thesis with Companies and Internship AY 23/24

Research opportunities

Important Disclaimers

- These thesis are activated within existing research collaborations with companies / research institutions and typically lead to a full, reseach-oriented thesis
- All these companies are sponsoring PhD Scolarship and/or Industrial Research Projects
- You might be collaborating with other MsC/PhD students currently involved in the same research project
- Internship opportunities are often available and not always mandatory. It is also possible to work to a thesis in academia addressing these problem with the project team.

Antares Vision

Via del Ferro, 16, 25039 Travagliato BS

R&D Group

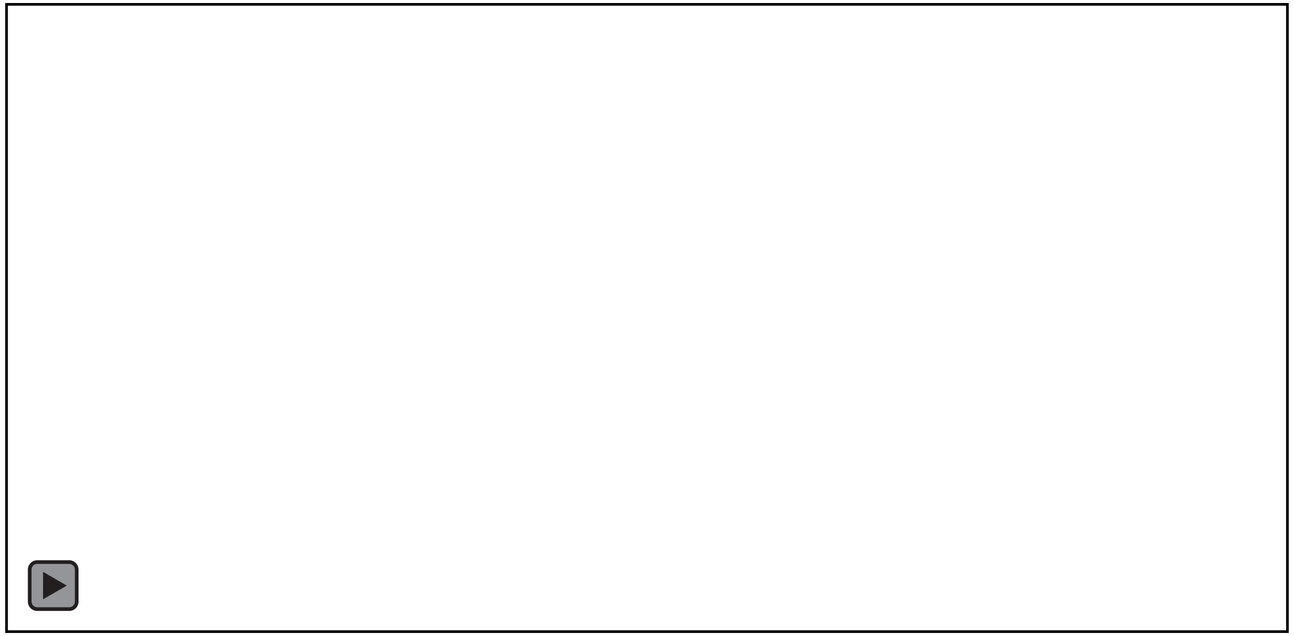
AV1. Calibration of Liquid Lens Cameras

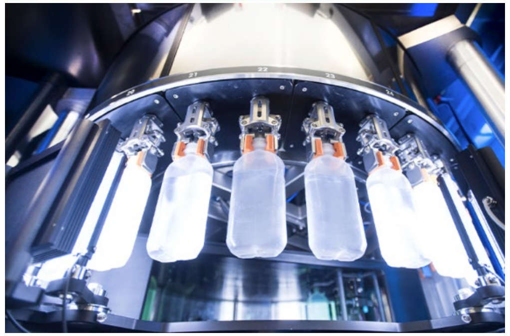
Background

Quality Inspection machines require very accurate measuraments from camera arrays in very short times. Typically, objects of different sizes require different carefully tailored optical settings.

Liquid lens camera is a novel technology that improves the focusing speed and the adjustment of focal length.

In principle, this technology would allow to quickly inspect object of different sizes or at different distances, but it is necessary to accurately estimate the focal lenght of these cameras.





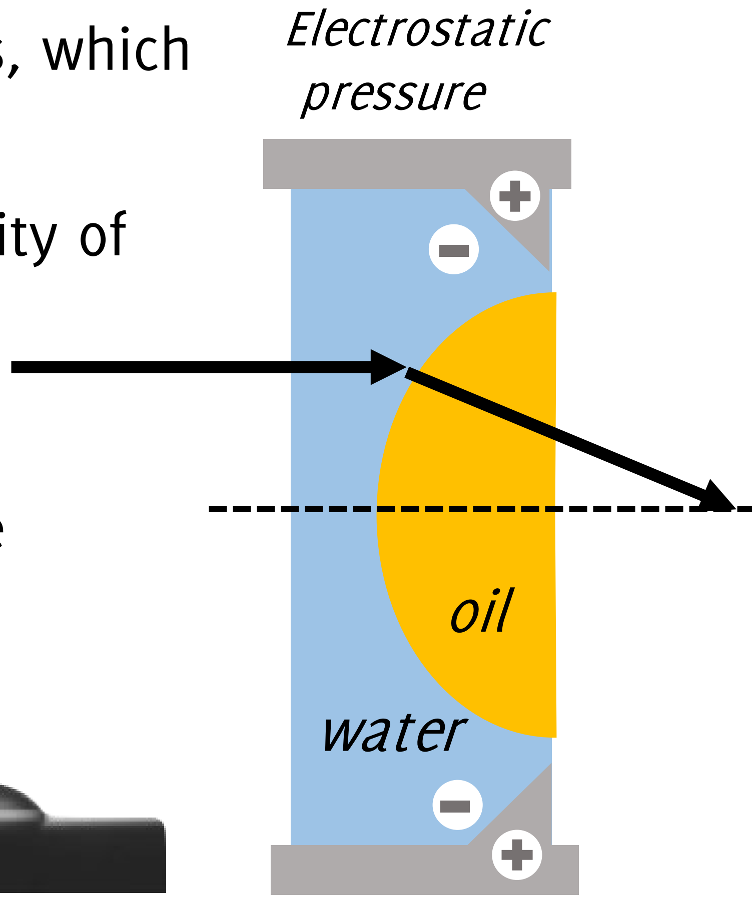
Dal Cin, Magri, Boracchi

Research Directions

AV 1: Calibration of Liquid Lens Cameras

Goals

- Understand the principles underpinning liquid lens cameras, which differ from the traditional pinhole model.
- Develop novel calibration algorithms to improve the reliability of measuraments for quality inspection.



Materials and Methods

- Access to the Luiquid Lens cameras, hardware and software provided by Antares Vision

Requirements: IACV, Python

Start: now

Duration: 9-12 months

In collaboration with

ANTARES VISION GROUP

Thesis/Internships @STMicroelectronics

Agrate Brianza

Applied Math Team – Software Platforms

SRA-System Research and Application

life.augmented

Internship Info

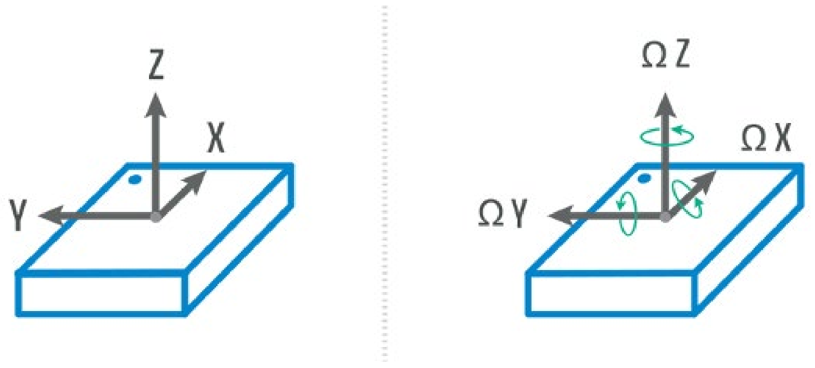
- The research thesis will be carried out under the supervision of the Applied Math Team in the STMicroelectronics Labs (Agrate Brianza, Centro Direzionale Colleoni + Smart Working).
- Internship lasts 6 months and can be extended up to other 6 months (typically, internships last 9 months).
- Start to be agreed after September 2023.
- The net salary is 800 €/month (graduate students: 1000 €/month)
- Canteen + transportation from Gobba / Lambrate included
- Contact: Pasqualina Fragneto pasqualina.fragneto@st.com

ST1. MEMS Sensors Calibration

Background

MEMS sensors as Inertial Measurement Units IMUs are affected by systematic errors (imprecise **scaling factors**, **axes misalignments** and **offsets**) that decrease accuracy in the position and attitudes estimation. Sensor **calibration** is critical to ensure the output of the sensor is accurate and repeatable within the same operating conditions.

This problem can be tackled by optimization techniques (time-consuming and not automated) or deep learning (data-driven) techniques.



Accelerometer   Gyroscope

Industry   Computing

Personal Electronics   Automotive

ST1. MEMS Sensors Calibration

Research Directions

Internship Goals:

- Design and implement algorithms for automatically calibrate MEMS sensors (at least gyroscopes), and possibly recalibrate them on the field;
- Optimize the algorithms to run on a microcontroller;
- Run exhaustive tests on real data.

Materials and Methods:

- Public datasets and acquisitions from ST sensors will be provided;
- Access to servers mounting GPU

Prerequisites: Machine Learning, Optimization, Signal Processing, Python and C programming.