3rd INRIA-DFKI EUROPEAN SUMMER SCHOOL ON ARTIFICIAL INTELLIGENCE

SOPHIA ANTIPOLIS, SEPTEMBER 4-8, 2023

PROGRAMME AND PRACTICAL INFORMATION

A warm welcome to IDESSAI 2023

If you are reading these lines, it means you are participating in the third edition of the joint Summer School by DFKI and Inria. This year's Summer School takes place in-person in Sophia Antipolis, at the Inria center.

This event, after the enriching joint presentations at Vivatech, is another milestone this year within the Inria-DFKI partnership. Our priority is to establish Europe at the top of AI research and innovation worldwide and therefore to build up a community of young researchers that both deeply understand the science of AI in all its different facets but also consider its impact on humans and our society.

IDESSAI 2023 is a prime example of networking and exchange among researchers. Many young PhD students come together to present and discuss their research and interact with each other as well as with experts in the field. The focus of this year's IDESSAI is on "Simulation & AI" and "AI for Agriculture & the Environment". These two areas are at the forefront of socio-economic issues related to AI.

We are excited to have you join these discussions and form a new bond across borders.

We wish you a fruitful week!

Pierre Alliez
Scientific coordinator of the Inria-DFKI partnership

Philipp Slusallek
Executive Director, DFKI Saarbrücken

Inria

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH

Special thanks also granted to our sponsors:

Inria

Université franco-allemande Deutsch-Französische Hochschule

3iA Côte d'Azur
Institut interdisciplinaire d'intelligence artificielle

Getting here

By plane
The closest international airport is Nice-Côte d'Azur (20km).
Regular shuttles from the airport (Terminal 1 and 2) to Sophia Antipolis by the Zou bus service:

- Line 630 (previously 230) – Bus stop "Inria".

By car
GPS 43.61619N (43°36'58") 7.06786E (07°04'04")
Access map

- From Marseille-Aix by the A8 motorway, take the exit Antibes-Grasse-Sophia Antipolis, direction Sophia Antipolis.
- From Nice by the A8 motorway, take the exit Antibes-Sophia Antipolis, direction Sophia Antipolis (the road on the right just after the toll).

By train
The closest railway station is located in Antibes. Regular shuttles from Antibes SNCF railway station to Sophia Antipolis by bus with Envibus:

- Line B and 12 – Bus stop "Inria";
- Line A and 09 – Bus stop "St. Philippe" (+7 minute walk).

Fare ticket: 1.50 € (or on mobile/at ticket machines: 1.00 €)

You can also use the Zou map webapp to help plan your journey:
https://services-zou.maregionsud.fr/en/journey

By taxi
About 20 minutes, depending on traffic congestion (approx. 55€)

- Transfert Service (English spoken) +33 (0) 6 09 50 92 53
- Taxi Sophia +33 (0) 6 27 51 01 51
- Centrale Orange Taxi +33 (0) 820 906 960
- Motorbike transport +33 (0) 6 58 79 81 31
  see also: http://www.cote-azur.com.fr

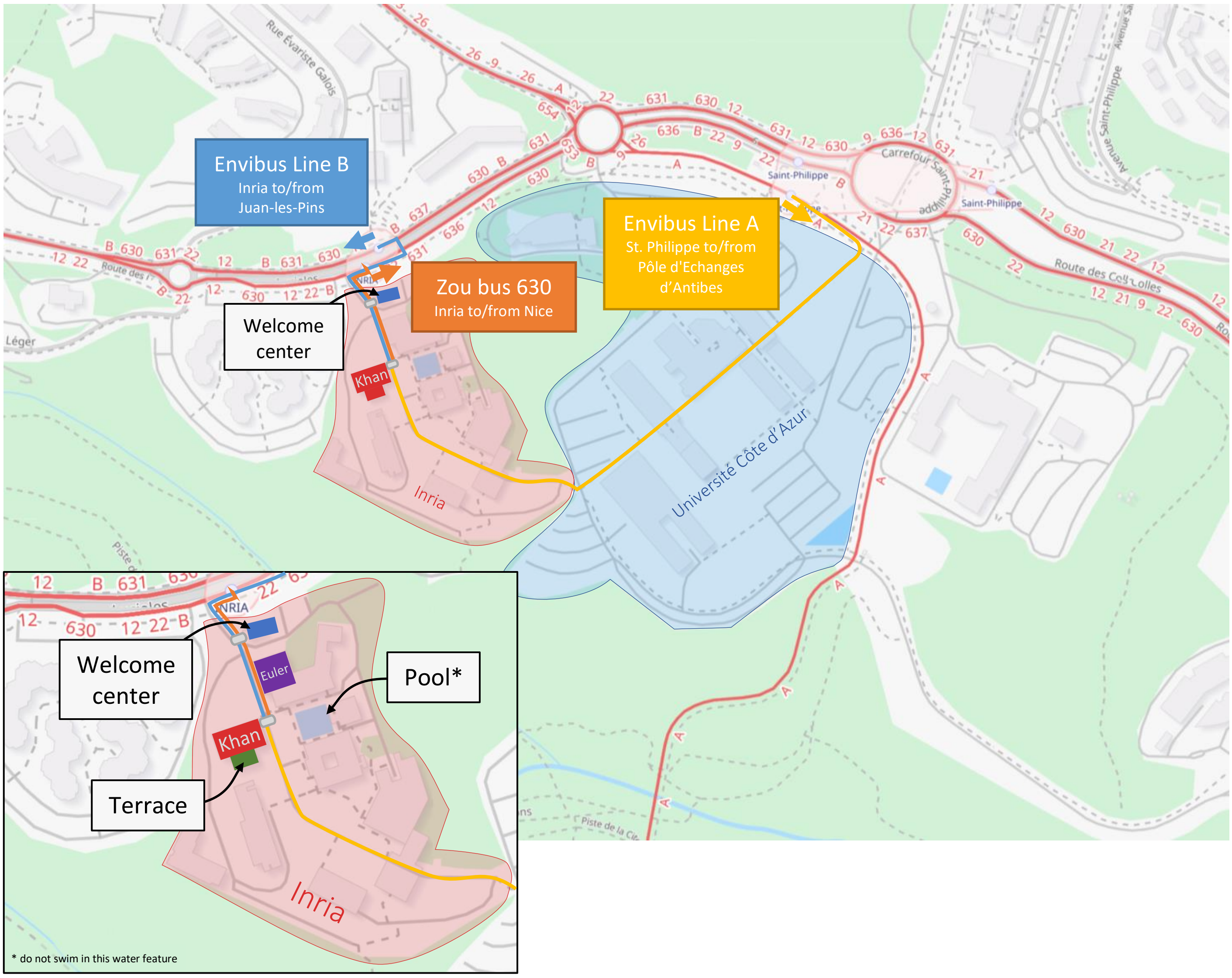
Car rental
Nice Cote d'Azur Airport offers a choice of several car rental companies (Terminal 2).

Routes to access the Inria center on foot are marked on the map on the next page.

Venue and building

The Summer School will take place at Inria center at Université Côte d'Azur, 2004 route des Lucioles, Sophia Antipolis. Tel: +33 (0)4 92 38 77 77

Upon your arrival, please come to the check-in desk located in the foyer of the Khan building. If you need help finding the building, please go to the welcome office (blue). For IDESSAI attendees everything (lectures, lunches, welcome party, etc.) will happen in the Kahn building (red) and its terrace (green). GAIN workshop attendees will also have activities in the Euler building (purple).



Wi-Fi network with access control on site

You should have received your credentials by email automatically a few days before the conference. If not, please do not hesitate to contact us on site.

Inria is also involved in Eduroam project.

If your institution provides you with an Eduroam account, you are encouraged to initialize it on your laptop properly (help in English / aide pour les français).

Schedule

The schedule for Track A and Track B is expected to be as follows:

N.B. The GAIN workshop will be running in parallel in the Euler Violet room

Track A: Simulation & AI

	Mon, Sep 04	Tue, Sep 05	Wed, Sep 06	Thu, Sep 07	Fri, Sep 08
9:00-10:00		Keynote 2	Keynote 3	Keynote 4	Keynote 5
10:00-10:30		Coffee break	Coffee break	Coffee break	Coffee break
10:30-11:30		Course 2	Course 4	Talk 1	Keynote 6
11:30-11:45		Coffee break	Coffee break	Coffee break	Coffee break
11:45-13:15		Course 2	Course 4	Course 6	Collaborative wrap-up
13:15-14:30	Opening ceremony	Lunch	Lunch	Lunch	Lunch
14:30-15:30	Keynote 1	Participant posters and demos	Participant posters and demos	Participant posters and demos	
15:30-15:45	Coffee break	Coffee break	Coffee break	Coffee break	
15:45-16:45	Course 1	Course 3	Course 5	Course 6	
16:45-17:00	Coffee break	Coffee break	Coffee break	Coffee break	
17:00-18:00	Course 1	Course 3	Course 5	Talk 2	
18:00-18:30	Welcome	〃	〃	〃	
18:30- open	party		Special night		

Locations

Opening ceremony	Amphitheatre, Khan building
Keynotes	Amphitheatre, Khan building
Courses	Amphitheatre, Khan building
Talks	Amphitheatre, Khan building
Participant posters and demos	Foyer, Khan building
Welcome party	By the pool
Special night	By the pool
Collaborative wrap-up	Amphitheatre, Khan building
Coffee breaks	Foyer, Khan building
Lunch	Terrace adjacent to Khan building

Track A program

See the schedule for Track B further below.

Mon, Sep 04, 2023

- 13:15-14:30 — Opening ceremony
- 14:30-15:30 — Keynote 1: Petra Gospodnetić — Pandora’s box of synthetic images for AI
- 15:45-18:00 — Course 1: Matthieu Lecce — Generating high-quality synthetic datasets for computer vision deep learning models

Tue, Sep 05, 2023

- 09:00-10:00 — Keynote 2: Matthias Nachtmann — The beauty of data – core ingredient for scalable, sustainable crop production
- 10:30-13:15 — Course 2: Régis Duvigneau — Physics-informed neural networks for simulation
- 14:30-15:30 — Participant posters & demos
    - Demo: Timothée Stassin — "Training AI models for forest land use classification using the FAO FRA RSS dataset" in Room 1-2-3, Khan building.
- 15:45-18:30 — Course 3: Shai Machnes & Anurag Saha Roy — Qruise: First steps on the path to an ML Physicist

Wed, Sep 06, 2023

- 09:00-10:00 — Keynote 3: Bertrand Le Saux — Next-Generation Machine Learning for Earth Observation
- 10:30-13:15 — Course 4: Victor Michel-Dansac — Numerical schemes for hyperbolic equations enhanced by Scientific Machine Learning
- 14:30-15:30 — Participant posters & demos
    - Demo: Antoine Thébault — "Traffic Management Learning - Railways Systems" in Room 1-2-3, Khan building.
- 15:45-18:30 — Course 5: Katja Schladitz & Tim Dahmen — Surrogate model for Monte-Carlo simulation of electron matter interaction

Thu, Sep 07, 2023

- 09:00-10:00 — Keynote 4: Stefania Fresca — Deep learning-based reduced order models for scientific applications
- 10:30-11:30 — Talk 1: Guillaume Cordonnier — Deep learning for fast and efficient glacier modeling, with applications in glacial erosion
- 11:45-13:15 — Course 6: Gaétan Bahl & Laurent Pilati — Applications of AI inference at the edge with NXP processors
- 14:30-15:30 — Participant posters & demos
- 15:45-16:45 — Course 6 cont.: Gaétan Bahl & Laurent Pilati — Applications of AI inference at the edge with NXP processors
- 17:00-18:30 — Talk 2: Mathieu Desbrun — Super-resolution via AI for fluid simulation

Fri, Sep 08, 2023

- 09:00-10:00 — Keynote 5: Roland Lenain — Robotics for agriculture
- 10:30-11:30 — Keynote 6: Siddhartha Mishra — Learning Operators
- 11:45-13:15 — Collaborative wrap-up

Track B: AI for Agriculture and the Environment

	Mon, Sep 04	Tue, Sep 05	Wed, Sep 06	Thu, Sep 07	Fri, Sep 08
9:00-10:00		Keynote 2	Keynote 3	Keynote 4	Keynote 5
10:00-10:30		Coffee break	Coffee break	Coffee break	Coffee break
10:30-11:30		Course 2	Course 4	Course 6	Keynote 6
11:30-11:45		Coffee break	Coffee break	Coffee break	Coffee break
11:45-13:15		Course 2	Course 4	Course 6	Collaborative wrap-up
13:15-14:30	Opening ceremony	Lunch	Lunch	Lunch	Lunch
14:30-15:30	Keynote 1	Participant posters and demos	Participant posters and demos	Participant posters and demos	
15:30-15:45	Coffee break	Coffee break	Coffee break	Coffee break	
15:45-16:45	Course 1	Course 3	Talk 5.1	Course 7	
16:45-17:00	Coffee break	Coffee break	Coffee break	Coffee break	
17:00-18:00	Course 1	Course 3		Course 7	
18:00-18:30	Welcome	”		”	
18:30- open	party		Special night		

Locations

Opening ceremony	Amphitheatre, Khan building
Keynotes	Amphitheatre, Khan building
Courses	Room 1-2-3, Khan building
Talks	Room 1-2-3, Khan building
Participant posters and demos	Foyer, Khan building
Welcome party	By the pool
Special night	By the pool
Collaborative wrap-up	Amphitheatre, Khan building
Coffee breaks	Foyer, Khan building
Lunch	Terrace adjacent to Khan building

Track B program

See the abstracts for Track A above.

Mon, Sep 04, 2023

- 13:15-14:30 — Opening ceremony
- 14:30-15:30 — Keynote 1: Petra Gospodnetić — Pandora's box of synthetic images for AI
- 15:45-18:00 — Course 1: Diego Marcos & Dino Ienco — AI for Earth surface monitoring through satellite image time series data

Tue, Sep 05, 2023

- 09:00-10:00 — Keynote 2: Matthias Nachtmann — The beauty of data – core ingredient for scalable, sustainable crop production
- 10:30-13:15 — Course 2: Olivier Bernard — Hybrid modelling of artificial microbial ecosystems for bioenergy production and waste water treatment
- 14:30-15:30 — Participant posters & demos

  - Demo: Timothée Stassin — "Training AI models for forest land use classification using the FAO FRA RSS dataset" in Room 1-2-3, Khan building.
- 15:45-18:30 — Course 3: Max Polzin — Co-Developing Robotics and Agrifood Systems for a Sustainable Future

Wed, Sep 06, 2023

- 09:00-10:00 — Keynote 3: Bertrand Le Saux — Next-Generation Machine Learning for Earth Observation
- 10:30-13:15 — Course 4: Patrick Armengaud — Decision support tools in agriculture: machine learning, process-based modelling, digitalization – chosen examples and live practical use
- 14:30-15:30 — Participant posters & demos
  - Demo: Antoine Thébault — "Traffic Management Learning - Railways Systems" in Room 1-2-3, Khan building.
- 15:45-16:45 — Talk 5.1: François Bremond & Vincent Calcagno — Video analytics for monitoring biocontrol agents

Thu, Sep 07, 2023

- 09:00-10:00 — Keynote 4: Stefania Fresca — Deep learning-based reduced order models for scientific applications
- 10:30-13:15 — Course 6: Alexis Joly & Mathias Chouet — Cooperative learning for biodiversity monitoring: how does the Pl@ntNet platform work and how can you use it in your own application?
- 14:30-15:30 — Participant posters & demos
- 15:45-18:30 — Course 7: Odalric-Ambrym Maillard — Reinforcement learning and application to sustainable gardening

Fri, Sep 08, 2023

- 09:00-10:00 — Keynote 5: Roland Lenain — Robotics for agriculture
- 10:30-11:30 — Keynote 6: Siddhartha Mishra — Learning Operators
- 11:45-13:15 — Collaborative wrap-up

Abstracts

Track A abstracts

See the schedule for Track B further below.

Mon, Sep 04, 2023

13:15-14:30 — Opening ceremony

Outline

Pierre Alliez & Roberto M. Dyke	Welcome, schedule and practical information
Cécile Vigouroux	The center's director of international relationships on the Inria-DFKI partnership
Tim Dahmen & Pierre Alliez	Scientific program of the summer school
Charles Bouveyron	3iA Côte d'Azur
Mohamed Behery	CLAIRE's Rising Researcher Network
Maureen Clerc	Words from the center's director of research

14:30-15:30 — Keynote 1: Petra Gospodnetić — Pandora's box of synthetic images for AI

Chaired by: Tim Dahmen

We are surrounded by virtual worlds with different levels of realism. Just about every movie today has stunning computer graphics footage which is nearly impossible for us to distinguish from real life footage. There are apps allowing us to place virtual objects in real environments to help us envision what changes in our environment could look like. We have tools which allow us to digitally plan, design and recreate scenes and objects which will later on be produced according to the design. At the same time, we are developing extremely powerful AI models whose performance is heavily dependent on the real data we train them on. Given that the amount and balance of the data is often the main source of problems when developing a model, it is only logical that we ask "Hey, why don't we create synthetic data instead and feed that to the network?". With that seemingly innocent and completely justified question, we have uncovered a whole new world of challenges such as when is "realistic" realistic enough? Which features of the image must be realistic, and what can we *get away with*? Which methods should we use? Do we really want pixel-perfect annotations? How do we measure dataset balance? And finally, often unmentioned, what does the ideal team for simulation-based machine learning look like? Which expertise is needed?

15:45-18:00 — Course 1: Matthieu Lecce — Generating high-quality synthetic datasets for computer vision deep learning models

Chaired by: Abir Affane

In the evolving landscape of machine learning, the trend has clearly shifted towards a data-centric approach, especially in computer vision. With the growing complexity of models comes an increasing need for diverse and high-quality data. Yet, sourcing and annotating real-world data present notable challenges in terms of cost, scalability, and precision. Enter synthetic dataset generation through computer graphics—a promising solution.
This talk will shed light on the computer graphics-based methodology for data creation, emphasizing its merits in crafting realistic environments and human representations. We'll underline their critical role in areas like object detection, human pose estimation, and activity recognition. A spotlight will be given to home robotics and automation, where the right data can drive significant performance improvements.

Balancing the need for realistic images with practical challenges, we'll address the obstacles faced in making synthetic data believable to AI models. What does it take to make a synthetic image "believable" to an AI model?

To conclude, we'll share a demonstration of AI Verse's tool, which allows users to procedurally construct scenes and capture annotated images. After setting up desired objects, activities, and scene constraints, the tool introduces random lighting and camera placements. This provides both the required variability and the capability to focus on specific use cases essential for the user's application. We believe this approach offers a meaningful step towards generating more relevant datasets for machine learning endeavors.

Tue, Sep 05, 2023

09:00-10:00 — Keynote 2: Matthias Nachtmann — The beauty of data – core ingredient for scalable, sustainable crop production

Chaired by: Diego Marcos

Today, we enjoy the best quality and availability of food. At the same time, looking at weather developments, logistics issues, volatile wheat prices and a growing population, we feel this might change in the future. This keynote will focus on two questions: “How to increase sustainability and productivity of food production?” and “What opportunities data and data science provide?”. Discussions will include real business cases, reference the latest scientific publications and examine lots of data examples. The speech tries to motivate attendees to consider agriculture and food as their future scope of engagement.

10:30-13:15 — Course 2: Régis Duvigneau — Physics-informed neural networks for simulation

Chaired by: Tim Dahmen

Physics-Informed Neural Networks (PINNs) have recently emerged as a new learning method that can model various phenomena, by mixing both data observations and physical laws, such as partial differential equations.
In this course, we first introduce the underlying concepts of PINNs and demonstrate how this approach can be efficiently used for forward analysis, parametric modeling or inverse problem solving. Then, the reasons for possible failures are discussed and advanced techniques to overcome these difficulties are presented. Illustrations are provided for different problems in mechanics, heat transfer and flow dynamics.

14:30-15:30 — Participant posters & demos

Demo by Timothée Stassin on “Training AI models for forest land use classification using the FAO FRA RSS dataset” in Room 1-2-3, Khan building.

15:45-18:30 — Course 3: Shai Machnes & Anurag Saha Roy — Qruise: First steps on the path to an ML Physicist

Chaired by: Tim Dahmen

What does it mean to build an ML Physicist? At Qruise we believe the answer is a 3-stage moonshot, comprising of ML approaches for numerics, analytics and LLMs. We’ll present the progress we have made building the first stage, combining extremely detailed differentiable digital twins, reinforcement learning, adversarial approaches, SGD and more, and a glimpse of how the system is currently calibrating superconducting quantum processors at Forschungszentrum Jülich. You will also be able to simulate and optimize quantum gates on a virtual quantum device in the Qruise cloud. Finally, we will sketch our plans going forward, with the 2nd stage analytics and 3rd stage LLMs, and our unique view of how these need to be trained.

Wed, Sep 06, 2023

09:00-10:00 — Keynote 3: Bertrand Le Saux — Next-Generation Machine Learning for Earth Observation

Chaired by: Dino Ienco

The last decade has witnessed a profound transformation in Earth Observation (EO) through the deep learning revolution, with Artificial Intelligence (AI) being increasingly employed to tackle various EO challenges. However, the unique characteristics of EO data, which are in fine geospatial measurements from diverse sensors, and the intricate nature of underlying physical processes (originating in the Earth system, weather and climate, or human activities), pose significant obstacles. As a result, progress in adopting machine learning techniques varies across applications, and the full potential of these methods remains untapped. This presentation will provide an overview of the past decade's advancements, current status, and challenges in using machine learning for EO, followed by insights into the next stage of the ongoing machine learning revolution, including foundational models and their implications.

It will also delve into the future of EO, exploring the emergence of quantum machine learning and quantum computing for EO, as well as the potential of EO processing in modular high-performance computing (HPC) environments.

10:30-13:15 — Course 4: Victor Michel-Dansac — Numerical schemes for hyperbolic equations enhanced by Scientific Machine Learning

Chaired by: Roberto Dyke

The goal of this course is to first present, then implement, a modification of the bases in the classical Discontinuous Galerkin (DG) method applied to hyperbolic systems of balance laws. To that end, we first introduce the physical context of hyperbolic systems of conservation laws and balance laws, which are systems of Partial Differential Equations (PDEs) able to model complex physical phenomena (e.g., in gas dynamics, fluid mechanics,…). Then, we provide a gentle introduction to the challenges arising when discretizing such systems, including preservation of steady solutions and high-order accuracy. Then, we introduce the DG method, as well as the strategy we use to enhance it based on some given prior. The remainder of the lecture is devoted to the generation of such a prior, for which we elect to use a Physics-Informed Neural Network (PINN). PINNs are introduced in a general manner, before being applied to the problem at hand. Validation numerical experiments are finally presented.

Following the lecture, a practical computer session is proposed. The goal of this session is to train a PINN, and then implement it as a modification of a DG basis. Trainees will be provided with a python code with missing functions, to be implemented during the course of the computer session. Trainees will need access to a computer with the PyTorch (≥2.0), SciPy, NumPy, Matplotlib and iPyKernel packages. One can also run the code from Google Colab.

14:30-15:30 — Participant posters & demos

Demo by Antoine Thébault on “Traffic Management Learning - Railways Systems” in Room 1-2-3, Khan building.

15:45-18:30 — Course 5: Katja Schladitz & Tim Dahmen — Surrogate model for Monte-Carlo simulation of electron matter interaction

Chaired by: Roberto Dyke

We demonstrate a simulation-based approach to create synthetic training images for the segmentation of electron microscopy images of porous materials. This practically highly relevant class of materials is difficult to image and segment on the nanoscale, as the behavior of the electron beam interacting with pores can be quite erratic and creates strong imaging artifacts.

To train a segmentation model, we show how statistical geometry models can be used to create synthetic microstructures. The electron microscopy imaging of the microstructures can in principle be simulated using Monte-Carlo simulations, but the computational cost for mass data generation is prohibitive. We therefore show how neural networks can be used as an approximate surrogate model for the Monte-Carlo simulation, allowing several orders of magnitude of speedup. As data representation is key to achieve good results with the surrogate models, we also investigate several aspects of representing and converting synthetic microstructures.

Thu, Sep 07, 2023

09:00-10:00 — Keynote 4: Stefania Fresca — Deep learning-based reduced order models for scientific applications

Chaired by: Stephane Lanteri

The solution of differential problems by means of full order models (FOMs), such as, e.g., the finite element method, entails prohibitive computational costs when it comes to real-time simulations and multi-query routines. The purpose of reduced order modeling is to replace FOMs with reduced order models (ROMs) characterized by much lower complexity but still able to express the physical features of the system under investigation.
Conventional ROMs anchored to the assumption of modal linear superimposition, such as proper orthogonal decomposition (POD), may reveal inefficiencies when dealing with nonlinear time-dependent parametrized PDEs, especially for problems featuring coherent structures propagating over time. To overcome these difficulties, we propose an alternative approach based on deep learning (DL) algorithms, where tools such as convolutional neural networks (CNNs) are used to build an efficient nonlinear surrogate. In the resulting DL-ROM, both the nonlinear trial manifold and the nonlinear reduced dynamics are learned in a non-intrusive way by relying on DL models trained on a set of FOM snapshots, obtained for different parameter values [1, 2]. Accuracy and efficiency of the DL-ROM technique are assessed in several applications, ranging from cardiac electrophysiology [3] to fluid dynamics [4], showing that new queries to the DL-ROM can be computed in real-time.

Finally, with the aim of moving towards a rigorous justification on the mathematical foundations of DL-ROMs, error bounds are derived for the approximation of nonlinear operators by means of CNNs. The resulting error estimates provide a clear interpretation on the role played by the hyperparameters of dense and convolutional layers [5].

[1] S. Fresca, A. Manzoni, L. Dede' 2021 A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs. Journal of Scientific Computing, 87(2):1-36.
[2] S. Fresca, A. Manzoni 2022 POD-DL-ROM: enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition. Computer Methods in Applied Mechanics and Engineering, 388, 114181.
[3] S. Fresca, A. Manzoni L. Dede', A. Quarteroni 2021 POD-enhanced deep learning-based reduced order models for the real-time simulation of cardiac electrophysiology in the left atrium. Frontiers in Physiology, 12, 1431.
[4] S. Fresca, A. Manzoni 2021 Real-time simulation of parameter-dependent fluid flows through deep learning-based reduced order models. Fluids, 6(7), 259.
[5] N. R. Franco, S. Fresca, A. Manzoni, P. Zunino 2023 Approximation bounds for convolutional neural networks in operator learning. Neural Networks, 161, 129-141.

10:30-11:30 — Talk 1: Guillaume Cordonnier — Deep learning for fast and efficient glacier modeling, with applications in glacial erosion

Chaired by: Pierre Alliez

Climate change destabilizes the fragile equilibrium around mountain glaciers and threatens the surrounding population. A better understanding of the glacier dynamics and their potential future is therefore critical, which drives the development of numerical models for the simulation of the evolution of glaciers and their impact on the landscape. We will discuss the challenges of glacial modeling and how we use deep learning to enable a model that is accurate and several orders of magnitude faster than previous works. We will introduce an ice-flow emulator and different training strategies, either data-based through a collection of runs from existing physical models or data-free thanks to a loss that trains the model to solve the underlying physical problem. We will then show a few applications, from inverse problems common in glaciology to recent advances in glacial erosion.

11:45-13:15 & 15:45-16:45 — Course 6: Gaétan Bahl & Laurent Pilati — Applications of AI inference at the edge with NXP processors

Chaired by: Armand Zampieri

NXP Semiconductors is building the processors that will enable the future in intelligent embedded applications. We aim to do so in a safe and secure way by offering complete, fast and reliable solutions to our clients. Thanks to NXP, autonomous vehicles, robots, drones and mobile devices of the future will be equipped with neural accelerators (such as our eIQ Neutron NPU) that allow fast machine learning inference on the edge. However, the hardware is only the first part of the complete solution. The NXP Technology Engineering Centers are in charge of the software enabling efficient video and audio processing pipelines on NXP processors. In order to push the boundaries of what we can achieve with our hardware, we develop efficient in-house AI architectures specifically tailored to our processors and our customers' use cases.
In this talk, we will present our portfolio of processors for machine learning applications, as well as our in-house tools. Then, we will showcase various vision and audio applications, such as stereo vision and speaker identification. Finally, a lab will allow you to familiarize yourself with our tools in order to use them for your future projects.

14:30-15:30 — Participant posters & demos

17:00-18:30 — Talk 2: Mathieu Desbrun — Super-resolution via AI for fluid simulation

Chaired by: Pierre Alliez

Simulating turbulent flows with fine details is computationally intensive. While we have seen great strides towards AI-aided simulation which promises to be faster than traditional direct numerical simulation, they still cannot provide fine simulations of arbitrary initial conditions with accurate large-scale behaviors. Performing, instead, a low-resolution numerical simulation to which physically-consistent high-frequency details are added automatically to induce a higher effective resolution is an attractive alternative. However, predicting the fine and intricate details of a turbulent flow field in both space and time from a coarse input (i.e., spatial/temporal upsampling through super-resolution) remains a major challenge despite the availability of modern machine learning tools. We demonstrate that a neural network approach can reproduce the visual complexity of turbulent flows from spatially and temporally coarse velocity fields even when using a generic training set. By generating finer spatial and/or temporal details through embarrassingly-parallel upsampling of small local patches, an ML-based approach can efficiently predict high-resolution turbulence details across a variety of grid resolutions — offering a whole range of applications varying from fluid flow upsampling to fluid data compression. We will be reviewing an approach that turns an input coarse animation into a sparse linear combination of small velocity patches present in a pretrained over-complete dictionary. We will discuss the efficiency and generalizability of this method for synthesizing turbulent flows on a series of complex examples, highlighting dramatically better results in spatio-temporal upsampling and flow data compression than existing methods as assessed by both qualitative and quantitative comparisons.

Fri, Sep 08, 2023

09:00-10:00 — Keynote 5: Roland Lenain — Robotics for agriculture

Chaired by: Ezio Malis

Climate change, caused by human activities, has become a reality. Conventional agriculture practices play an important role in this process due to the dissemination of chemicals that have been used to feed an increasing worldwide population. The rise of new practices, such as Precision Agriculture, Organic Farming, or more generally agroecological principles may permit to reduce the environmental impacts on food production. However, such an alternative way of farming requires frequent actions, with an increased level of accuracy to act at a plant scale in a discriminated way. Such work is hardly tractable by human operators, as tasks to be done, such as mechanical weeding, appear to be difficult and time consuming. As a result, robotics appears as a promising way to implement agroecological practices. Such robots however need to adapt their behavior to the diversity of missions and context of evolution, that are hardly achievable by conventional automation approaches. In this talk, after a short state of the art of current agriculture robots, we propose to investigate the contributions of advanced control algorithms and AI to enhance the capabilities of robots acting in an evolution environment, in interaction with vegetation and ensuring safety. These conditions are mandatory to develop robotics tools for farmers, allowing the proposition of new practices to feed the population without damaging the environment.

10:30-11:30 — Keynote 6: Siddhartha Mishra — Learning Operators

Chaired by: Stephane Lanteri

Operator learning is a fast-emerging field of machine learning where the objective is to learn operators from data. The applications are primarily for the solutions of both forward and inverse problems for partial differential equations. In this talk, we will provide a survey of this field through the prism of representation equivalence and examine how existing architectures (DeepONet, FNO etc.) fit into this paradigm. We will also introduce convolutional neural operators. Finally, we will briefly describe neural inverse operators, a state-of-the-art model for learning solutions of inverse problems associated with PDEs.

11:45-13:15 — Collaborative wrap-up

Track B abstracts

See the abstracts for Track A above.

Mon, Sep 04, 2023

13:15-14:30 — Opening ceremony

Outline

Pierre Alliez & Roberto M. Dyke	Welcome, schedule and practical information
Cécile Vigouroux	The center's director of international relationships on the Inria-DFKI partnership
Tim Dahmen & Pierre Alliez	Scientific program of the summer school
Charles Bouveyron	3iA Côte d'Azur
Mohamed Behery	CLAIRE's Rising Researcher Network
Maureen Clerc	Words from the center's director of research

14:30-15:30 — Keynote 1: Petra Gospodnetić — Pandora's box of synthetic images for AI

Chaired by: Tim Dahmen

We are surrounded by virtual worlds with different levels of realism. Just about every movie today has stunning computer graphics footage which is nearly impossible for us to distinguish from real life footage. There are apps allowing us to place virtual objects in real environments to help us envision what changes in our environment could look like. We have tools which allow us to digitally plan, design and recreate scenes and objects which will later on be produced according to the design. At the same time, we are developing extremely powerful AI models whose performance is heavily dependent on the real data we train them on. Given that the amount and balance of the data is often the main source of problems when developing a model, it is only logical that we ask "Hey, why don't we create synthetic data instead and feed that to the network?". With that seemingly innocent and completely justified question, we have uncovered a whole new world of challenges such as when is "realistic" realistic enough? Which features of the image must be realistic, and what can we *get away with*? Which methods should we use? Do we really want pixel-perfect annotations? How do we measure dataset balance? And finally, often unmentioned, what does the ideal team for simulation-based machine learning look like? Which expertise is needed?

15:45-18:00 — Course 1: Diego Marcos & Dino Ienco — AI for Earth surface monitoring through satellite image time series data

Chaired by: Armand Zampieri

In recent years, modern Earth observation programs are collecting a large amount of remote sensing data to take the pulse of our planet. A well-known example are the Sentinel missions, in the framework of the Copernicus program, which enabled the acquisition of satellite images in the same spatial area on a weekly basis. Such a stream of images can be organized as satellite image time series (SITS) that provide a dynamic picture of Earth and can be used to monitor and describe the processes occurring on its surface. As opposed to situations where we deal with single-scene images, SITS often provide as much information via the temporal dynamics as from spatial context. In this lecture we will provide an overview about the opportunities offered by modern Earth Observation programs, how the collected Earth Observation data can be leveraged to support the Sustainable Development Goals, how Deep Learning approaches are being leveraged to manage such heterogeneous and complex sources of information and several examples related to the exploitation of SITS data for environmental and agricultural applications.

Tue, Sep 05, 2023

09:00-10:00 — Keynote 2: Matthias Nachtmann — The beauty of data – core ingredient for scalable, sustainable crop production

Chaired by: Diego Marcos

Today, we enjoy the best quality and availability of food. At the same time, looking at weather developments, logistics issues, volatile wheat prices and a growing population, we feel this might change in the future. This keynote will focus on two questions: “How to increase sustainability and productivity of food production?” and “What opportunities data and data science provide?”. Discussions will include real business cases, reference the latest scientific publications and examine lots of data examples. The speech tries to motivate attendees to consider agriculture and food as their future scope of engagement.

10:30-13:15 — Course 2: Olivier Bernard — Hybrid modelling of artificial microbial ecosystems for bioenergy production and waste water treatment

Chaired by: Diego Marcos

Microbial ecosystems are more and more used for wastewater remediation, resource recovery and bioenergy production. The most famous processes are anaerobic digestion or activated sludge, but an increasing number of microorganisms are used, such as microalgae or yeasts with new promises and a high potential for targeting environmental issues. However, several bottlenecks must still be faced for reaching higher efficiency and lower costs at industrial-scale. Mathematical models turned out to be very powerful tools both for unraveling the complex nonlinear interactions among microbial populations and also for optimizing the process operations. So far, mechanistic models (MM) consisting of ordinary differential equations representing mass balance-based systems, were largely used for simulating biological processes. However, the complexity of these models prevents calibrating them in a realistic context—accounting for environmental fluctuations—and their for efficient process control and optimization. The recent progress in Artificial Neural Networks (ANN) can significantly contribute to improve modeling of environmental processes. However, they do not guarantee that fundamental properties such as mass conservation or variable positivity are respected. Hybrid models combine the best of two worlds: on the one hand, they respect physical laws, such as conservation of mass, and allow training of neural networks to fit different data sets. Also, these models can overcome some limitations of the neural networks which require a large amount of data, whereas a hybrid strategy is more data efficient. The aim of this tutorial is to detail in a practical case study an approach to combine hybridizing MM and ANN. Since stoichiometry of mass balance is less affected by the local environmental conditions, the neural network structure can focus on the biological kinetics and catch the process specificity and the hidden information in the data.

The example of the hybridization approach carried out with the high-fidelity ALBA model will be presented. This model which is constituted by different modules to simulate the biological and physic-chemical nonlinear dynamics of algae-bacteria systems was calibrated and validated on three different pilot-scale applications (over 600 days). The structure of the hybrid kinetics NN will be chosen in order that the model trajectories respect constraints such as state positivity or mass conservation. Then a two-step approach will be presented for the model training. First, the neural network is pre-trained to reproduce the ALBA kinetics (target) generated by the mechanistic model. This leads to a first set of parameters of the NN (weight and biases). This makes a starting point for the second step, where a back propagation approach is used, closing the gap between the model and the real data.

14:30-15:30 — Participant posters & demos

Demo by Timothée Stassin on “Training AI models for forest land use classification using the FAO FRA RSS dataset” in Room 1-2-3, Khan building.

15:45-18:30 — Course 3: Max Polzin — Co-Developing Robotics and Agrifood Systems for a Sustainable Future

Chaired by: Ezio Malis

The rapid advancement of robotics and their integration into our agrifood systems has the potential to fundamentally reshape the way we perceive, handle, produce and consume our food. With a growing world population and food being a main contributor to the global greenhouse gas emissions, novel technology must not only increase productivity but simultaneously reduce emissions.

At EPFL's CREATE Lab, our research focuses on the development and integration of novel robotics systems which integrate into farming processes, addressing its environmental challenges and deepening our understanding of the natural environment. This talk will be split into three parts:

1. Agricultural Robotics at CREATE Lab:
   We will present an overview of the robots developed at the CREATE Lab and their applications in the agrifood chain, from efficient farming practices to understanding how food is consumed.
2. Sensor Technology in Robotics:
   A deep dive into the sensors and measurement tools used in modern robotics. We will explore the potential of integrating these technologies into AI-driven agriculture, highlighting areas where AI and robotics converge to offer novel solutions in agritech.
3. Brainstorming the Future of Agritech Robotics:
   Participants will collaboratively explore challenges and solutions in agritech robotics and AI. Through a structured discussion, multiple groups will present their findings, which open questions could be answered by robotics and AI and what such robots might look like.

Join us for a systematic exploration of the current state, challenges, and future prospects of robotics in the agrifood chain system, emphasizing both practical applications and research frontiers.

Wed, Sep 06, 2023

09:00-10:00 — Keynote 3: Bertrand Le Saux — Next-Generation Machine Learning for Earth Observation

Chaired by: Dino Ienco

The last decade has witnessed a profound transformation in Earth Observation (EO) through the deep learning revolution, with Artificial Intelligence (AI) being increasingly employed to tackle various EO challenges. However, the unique characteristics of EO data, which are in fine geospatial measurements from diverse sensors, and the intricate nature of underlying physical processes (originating in the Earth system, weather and climate, or human activities), pose significant obstacles. As a result, progress in adopting machine learning techniques varies across applications, and the full potential of these methods remains untapped. This presentation will provide an overview of the past decade's advancements, current status, and challenges in using machine learning for EO, followed by insights into the next stage of the ongoing machine learning revolution, including foundational models and their implications.
It will also delve into the future of EO, exploring the emergence of quantum machine learning and quantum computing for EO, as well as the potential of EO processing in modular high-performance computing (HPC) environments.

10:30-13:15 — Course 4: Patrick Armengaud — Decision support tools in agriculture: machine learning, process-based modelling, digitalization – chosen examples and live practical use

Chaired by: Dino Ienco

Farming activities are strongly dependent on a wide range of interacting environmental factors determining crop productivity, livestock management, farm profitability and sustainability. The use of sensors (internet of things), process-based modeling, and machine learning algorithms embedded in the appropriate digital infrastructure enable the creation of robust digital twins of farms to better understand and decide on the most relevant and impactful intervention to engage by farmers and agricultural advisors. From simple decision-making data aggregation to integrated models, decision support tools are supporting farmer practices in optimizing e.g., irrigation, fertilization, phytoprotection; monitoring animal nutrition, health, welfare; and promoting sustainable but profitable agricultural transitions.

After presenting a general background of agricultural challenges and the journey of data from the research lab to a reliable and continuously improved building block, a set of different commercial decision support tools will be practically presented for cattle management, cereals productivity, tomato weekly production forecast, grapevine management and support of small holder farming in Southern countries.

14:30-15:30 — Participant posters & demos

Demo by Antoine Thébault on “Traffic Management Learning - Railways Systems” in Room 1-2-3, Khan building.

15:45-16:45 — Talk 5.1: François Bremond & Vincent Calcagno — Video analytics for monitoring biocontrol agents

Chaired by: Dino Ienco

Trichogramma wasp behaviors are studied extensively due to their effectiveness as biological control agents across the globe. However, the field of intra/inter-species Trichogramma behavior is yet to be explored thoroughly. To study these behaviors, it is crucial to identify and track Trichogramma individuals over a long period in a lab setup. The current research-front in this area consists of using Artificial Intelligence (mostly Deep Learning methods) to improve the performance of more classical tracking technologies. For this, we propose a robust tracking pipeline named TrichTrack. Due to the unavailability of labeled data, we train our detector using an iterative weakly supervised method. We also use a weakly supervised method to train a Re-Identification (ReID) network by leveraging noisy tracklet sampling. This enables us to distinguish Trichogramma individuals that are indistinguishable from human eyes. We also develop a two-staged tracking module that filters out the easy association to improve its efficiency. Our method outperforms existing insect trackers on most of the MOT metrics.

Thu, Sep 07, 2023

09:00-10:00 — Keynote 4: Stefania Fresca — Deep learning-based reduced order models for scientific applications

Chaired by: Stephane Lanteri

The solution of differential problems by means of full order models (FOMs), such as, e.g., the finite element method, entails prohibitive computational costs when it comes to real-time simulations and multi-query routines. The purpose of reduced order modeling is to replace FOMs with reduced order models (ROMs) characterized by much lower complexity but still able to express the physical features of the system under investigation.
Conventional ROMs anchored to the assumption of modal linear superimposition, such as proper orthogonal decomposition (POD), may reveal inefficiencies when dealing with nonlinear time-dependent parametrized PDEs, especially for problems featuring coherent structures propagating over time. To overcome these difficulties, we propose an alternative approach based on deep learning (DL) algorithms, where tools such as convolutional neural networks (CNNs) are used to build an efficient nonlinear surrogate. In the resulting DL-ROM, both the nonlinear trial manifold and the nonlinear reduced dynamics are learned in a non-intrusive way by relying on DL models trained on a set of FOM snapshots, obtained for different parameter values [1, 2]. Accuracy and efficiency of the DL-ROM technique are assessed in several applications, ranging from cardiac electrophysiology [3] to fluid dynamics [4], showing that new queries to the DL-ROM can be computed in real-time.

Finally, with the aim of moving towards a rigorous justification on the mathematical foundations of DL-ROMs, error bounds are derived for the approximation of nonlinear operators by means of CNNs. The resulting error estimates provide a clear interpretation on the role played by the hyperparameters of dense and convolutional layers [5].

[1] S. Fresca, A. Manzoni, L. Dede' 2021 A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs. Journal of Scientific Computing, 87(2):1-36.
[2] S. Fresca, A. Manzoni 2022 POD-DL-ROM: enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition. Computer Methods in Applied Mechanics and Engineering, 388, 114181.
[3] S. Fresca, A. Manzoni L. Dede', A. Quarteroni 2021 POD-enhanced deep learning-based reduced order models for the real-time simulation of cardiac electrophysiology in the left atrium. Frontiers in Physiology, 12, 1431.
[4] S. Fresca, A. Manzoni 2021 Real-time simulation of parameter-dependent fluid flows through deep learning-based reduced order models. Fluids, 6(7), 259.
[5] N. R. Franco, S. Fresca, A. Manzoni, P. Zunino 2023 Approximation bounds for convolutional neural networks in operator learning. Neural Networks, 161, 129-141.

10:30-13:15 — Course 6: Alexis Joly & Mathias Chouet — Cooperative learning for biodiversity monitoring: how does the Pl@ntNet platform work and how can you use it in your own application?

Chaired by: Florent Lafarge

Pl@ntNet is a citizen science platform that relies on Artificial Intelligence to facilitate the identification and inventory of plant species using mobile phones. It is one of the largest biodiversity observatories in the world, with several million contributors in over 200 countries. The key principle of Pl@ntNet is cooperative learning: users generate a large number of observations of plants in the field, which are automatically identified by an artificial intelligence algorithm and revised by the community itself, with a weighting principle based on the user's expertise. Observations that reach a sufficient degree of confidence are then added to the training set of the AI model, which is enriched and progresses over time. In addition, the data produced is used to model species distribution and map biodiversity indicators at very high spatial resolution. In the first part of this course, we'll take a detailed look at how the Pl@ntNet platform works, and the research we're currently carrying out. We'll then move on to a more practical part, with a short field test of the mobile application, followed by a programming session based on the Pl@ntNet API.

14:30-15:30 — Participant posters & demos

15:45-18:30 — Course 7: Odalric-Ambrym Maillard — Reinforcement learning and application to sustainable gardening

Chaired by: Mohamed Behery

In this lecture, I will first recall some strong links between agronomy and statistics, and show how statistical Reinforcement Learning may help answer the modern goals of Agroecology. Then, after summarizing classical models used in Reinforcement Learning (bandits, MDPs), we will explore the challenges of making RL indeed applicable to experimental sciences, in particular from the perspective of designing decision companion algorithms for practitioners in the field. We will see that applying RL for agroecology triggers many novel and non-trivial research questions, and then show how we can handle some of them by adapting bandit strategies and exploring alternative models to MDPs. I will finally present two simulation domains (a realistic one, and a gamified one) both compatible with the Gymnasium API, designed to challenge the field and test your favorite ideas. This lecture is expected to give more questions than answers, and invites people to join the research effort.

Fri, Sep 08, 2023

09:00-10:00 — Keynote 5: Roland Lenain — Robotics for agriculture

Chaired by: Ezio Malis

Climate change, caused by human activities, has become a reality. Conventional agriculture practices play an important role in this process due to the dissemination of chemicals that have been used to feed an increasing worldwide population. The rise of new practices, such as Precision Agriculture, Organic Farming, or more generally agroecological principles may permit to reduce the environmental impacts on food production. However, such an alternative way of farming requires frequent actions, with an increased level of accuracy to act at a plant scale in a discriminated way. Such work is hardly tractable by human operators, as tasks to be done, such as mechanical weeding, appear to be difficult and time consuming. As a result, robotics appears as a promising way to implement agroecological practices. Such robots however need to adapt their behavior to the diversity of missions and context of evolution, that are hardly achievable by conventional automation approaches. In this talk, after a short state of the art of current agriculture robots, we propose to investigate the contributions of advanced control algorithms and AI to enhance the capabilities of robots acting in an evolution environment, in interaction with vegetation and ensuring safety. These conditions are mandatory to develop robotics tools for farmers, allowing the proposition of new practices to feed the population without damaging the environment.

10:30-11:30 — Keynote 6: Siddhartha Mishra — Learning Operators

Chaired by: Stephane Lanteri

Operator learning is a fast-emerging field of machine learning where the objective is to learn operators from data. The applications are primarily for the solutions of both forward and inverse problems for partial differential equations. In this talk, we will provide a survey of this field through the prism of representation equivalence and examine how existing architectures (DeepONet, FNO etc.) fit into this paradigm. We will also introduce convolutional neural operators. Finally, we will briefly describe neural inverse operators, a state-of-the-art model for learning solutions of inverse problems associated with PDEs.

11:45-13:15 — Collaborative wrap-up