CALL FOR PROPOSALS

EXCHANGE SCHEMES – CATALOGUE OF CHALLENGES

ENFIELD: EUROPEAN LIGHTHOUSE TO MANIFEST TRUSTWORTHY AND GREEN AI

Co-funded by
the European Union

TABLE OF CONTENTS

INTRODUCTION

This is the catalogue of challenges available to the first out of four open calls for individual researchers exchange under the ENFIELD[1] (European Lighthouse to Manifest Trustworthy and Green AI) project, co-funded by the European Union. Through the ENFIELD Exchange Scheme open calls and the Financial Support to Third Parties (FSTP) mechanism, the project aims to attract the top-level researchers to conduct foundational research activities related to specific scientific/technological challenges in artificial intelligence, contributing to ENFIELD network creation and expansion to European AI labs.

[1] Grant Agreement nº 101120657, funded by the European Union.

Pillar Green-AI

G-AI. 1 Advancing Green AI on The Edge: Innovations for Sustainable, Efficient, And Continual Learning In Edge Computing
Keywords: Quantization and Pruning; Hardware Aware Architecture Search; On-Device Learning; Continual Learning (CL)
STATE-OF-THE-ART
The state-of-the-art in Green AI on the edge focuses on minimizing power consumption and maximizing computational efficiency. Techniques like model quantization, pruning, and data distillation are pivotal for deploying lightweight yet powerful AI models on edge devices. Such models are optimized for performance without the traditional energy drain, making use of innovative optimization methods tailored for the edge. Information-theoretic analyses enhance neural network training by identifying and eliminating redundancies, thereby reducing the environmental footprint. Continual learning emerges as a key strategy, adapting to new data with minimal retraining, crucial for applications with limited computational resources. This evolving paradigm requires the reimagining of industrial systems to embed Green AI principles fundamentally, promoting sustainability and resource efficiency from the edge to the cloud. The push towards Green AI integrates cutting-edge research with the urgent need for environmentally conscious technology deployment.
SCIENTIFIC CHALLENGES
• **Data Efficiency:** Innovating training data reduction techniques while maintaining model accuracy. • **Quantization & Pruning:** Creating methods to streamline model size and computational needs, balancing efficiency and effectiveness. • **Hardware-aware Differential Neural Architecture Search:** Developing a framework that automatically searches for efficient deep AI model on different hardware types, considering constraints related to computing, energy and memory capacities. • **Continual Learning:** Advancing continual learning strategies for energy and resource efficiency on the edge, in particular Continual On-device Learning with limited hardware resources. • **Communication Overhead:** Engineering IoT protocols to cut data transfer volume and improve latency, upholding security. • **Intermittent Computing:** Probing AI stability on edge devices with irregular power availability. • **Model Optimization:** Tailoring neural network optimization for industrial edge applications within operational constraints. • **Information-Theoretic Analyses:** Utilizing information theory for efficient neural network pruning and accelerated generalization. • **Adaptable CL Methodologies:** Developing versatile continual learning on-device methods that require less memory and processing. • **Sparse Architectures:** Researching lean architectures to lower the costs of training and inference.
RESEARCH ACTIVITIES
i) Development of resource-efficient neural networks for edge computing: (1) Implementing advanced quantization, pruning, and compression techniques; (2) Developing a differential architecture research that automatically searches for efficient deep AI model on different hardware types, considering constraints related to computing, energy and memory capacities; (3) Aiming to reduce data and power consumption without compromising AI performance; (4) Addressing the challenge of maintaining model accuracy alongside increased efficiency. **ii) Information-theoretic approaches for enhanced AI learning:** (1) Utilizing information theory and advanced techniques for analysing and improving neural network training; (2) Innovating in the relatively unexplored domain of information-theoretic pruning; (3) Potential to achieve significant reductions in energy usage. **iii) Continual Learning (CL) for dynamic adaptation:** (1) Developing CL systems that learn incrementally and require minimal retraining, in particular, Continual On-device Learning with limited hardware resources; (2) Adapting to new data streams, emphasizing computer science and cognitive science collaboration. **iv) Green AI adoption in industrial systems:** (1) Driving a cultural and technological shift towards sustainable AI practices; (2) Designing new software protocols for energy-efficient AI in edge-to-cloud continuum.
EXPECTED RESULTS
Anticipated scientific advancements include crafting energy-savvy, high-performing AI models suited for intermittent resources and systems that learn continuously to cut down on retraining needs, energy, and computational load. It will test rehearsal and sparse architectures for their adaptability and memory efficiency. Success will be measured by reduced energy use, improved IoT efficiencies, and lowered operation latencies, moving towards industrial robustness compatible with ecological sustainability. Also, the beneficiary is expected to have one peer-reviewed publication (conference, workshop, journal).
POSSIBLE HOST ORGANISATIONS
• IMT, Télécom Paris / Institut Mines-Télécom (https://www.telecom-paris.fr/en/home) • KNOW Center, Methods & Algorithms for AI Group (https://www.know-center.at/en/research/research-at-the-know-center/methods-algorithms-for-ai/) • POLIMI, Politecnico di Milano, Department of Management, Economics and Industrial Engineering (https://www.som.polimi.it/en/the-school/about-us/dig/) • SINTEF, Department of Sustainable Communication Technologies (https://www.sintef.no/en/digital/departments-new/department-of-sustainable-communication-technologies/) • TU/e, Department of Mathematics and Computer Science (https://www.tue.nl/en/our-university/departments/mathematics-and-computer-science)

Pillar Green-AI

G-AI.2 Optimizing Green AI in The Edge-To-Cloud Continuum
Keywords: Distributed AI; Edge-to-Cloud Orchestration; Lifecycle Assessment (LCA); Hybrid AI Models; Continual Learning Adaptation.
STATE-OF-THE-ART
Green AI research is steering towards sustainability, focusing on energy-efficient protocols for edge-cloud AI tasks to lower carbon emissions. Addressing the key challenges of task distribution and coordination, data transfer optimization, and low-latency model distribution and adaptation are crucial, given the environmental concerns accentuated by the shift from cloud to edge processing. Life cycle assessment (LCA) gains importance to understand the ecological implications of diverse hardware systems. Innovations lie in hybrid AI models that blend symbolic and data-driven approaches for better optimization, and continual learning technologies for self-improving AI systems.
SCIENTIFIC CHALLENGES
• Balancing energy efficiency with computational performance across the Edge-to-Cloud continuum. • Optimizing the allocation of distributed AI tasks between edge and cloud to reduce energy consumption and carbon emissions. • Addressing the challenge of lifecycle assessments due to the hardware heterogeneity in Edge-to-Cloud systems. • Developing energy-efficient communication protocols suited for AI tasks in these distributed systems. • Creating intelligent task orchestration strategies to optimize energy use across the continuum. • Innovating hybrid AI models that merge data-driven methods with symbolic AI or prior domain knowledge for better optimization. • Implementing continual learning mechanisms for AI systems to adapt to changing conditions while minimizing environmental impact. • Investigating the environmental implications of deploying distributed AI on edge-cloud continuum with specific energy consumption behaviours. • Guiding architectural decisions in system design to lower the environmental footprint of Edge-to-Cloud systems. • Integrating Green AI principles into industrial systems to promote sustainability in practical applications.
RESEARCH ACTIVITIES
• Developing energy-efficient protocols for AI edge-to-cloud operations. • Optimizing AI task distribution and coordination to reduce energy use in data transfers. • Adapting distributed AI models for efficiency across diverse network nodes. • Conducting LCA to gauge environmental impacts of Edge-to-Cloud AI systems. • Crafting hybrid AI models that merge theoretical and data-driven methods for lower energy consumption. • Integrating Green AI in industrial practices for sustainable system development. • Aiming for distributed AI systems that continually learn and self-optimize for environmental efficiency.
EXPECTED RESULTS
The ENFIELD project is poised to gain from Green AI research, with innovations such as energy-efficient protocols and AI-optimized algorithms reducing energy use and emissions. These advancements promise sustainable AI across edge-to-cloud platforms, with distributed intelligent task management and hybrid AI models enhancing problem-solving and lowering environmental impact. Lifecycle assessments will inform greener architectural choices, and continual learning will yield AI systems that improve performance while minimizing carbon footprint, marking ENFIELD's commitment to eco-friendly tech progress. The beneficiary is expected to have one peer-reviewed publication (conference, workshop, journal).
POSSIBLE HOST ORGANISATIONS
• IMT, Mines Saint-Etienne / Institut Mines-Télécom *(https://www.mines-stetienne.fr/en/)* • KNOW Center, *Methods & Algorithms for AI Group (https://www.know-center.at/en/research/research-at-the-know-center/methods-algorithms-for-ai/)* • SINTEF, Department of Sustainable Communication Technologies (https://www.sintef.no/en/digital/departments-new/department-of-sustainable-communication-technologies/)

Pillar Green-AI

G-AI.3 Green AI Metrics Initiative

Keywords: Standardization of Green-AI Metrics; Energy-Efficient Architectures; Lifecycle Environmental Impact; Computational Efficiency; Cross-Disciplinary Collaboration.

STATE-OF-THE-ART

State of the art in monitoring Green AI metrics is focused on developing standardized, accurate metrics to measure the environmental impact of AI throughout its lifecycle. These metrics aim to evaluate AI architectures not only for performance accuracy but also for energy efficiency and reduced carbon emissions, including hardware manufacturing impacts. Challenges include standardizing these metrics, data scarcity for environmental impact assessment, and the need for collaboration across disciplines. Innovations in this field could lead to new measurement methods, tools, and principles for energy-efficient AI, providing a competitive edge while benefiting the environment. Significant work is also directed at estimating the computational efficiency, like floating-point operations (FLOPs), for various AI models, facilitating comparisons under fixed computational budgets, crucial for SMEs with limited resources. Furthermore, sectors like telecommunications are exploring dynamic management of network capacities using AI to reduce energy consumption. In industry, there's a movement towards integrating Green AI principles into system development to promote efficiency and robustness without relying solely on the latest hardware advances.

SCIENTIFIC CHALLENGES

One of the primary scientific challenges in monitoring Green AI metrics lies in the establishment and standardization of these metrics across varied AI system architectures. This includes not only the computation of the efficiency and accuracy of algorithms but also accounting for the environmental impact throughout the AI lifecycle, from hardware production to operational deployment.
Researchers must grapple with the dearth of universally accepted metrics and the limited availability of comprehensive data needed to assess the full environmental footprint of AI technologies. The goal is to create a suite of standardized Green AI metrics that balances performance with energy efficiency, guiding the design of AI systems that are both robust and sustainable. For a comprehensive proposal, researchers would require access to current AI models, energy consumption data, and cross-sectoral environmental impact assessments, alongside the tools for the simulation and evaluation of AI architectures under these new metrics.

RESEARCH ACTIVITIES

1) Crafting universal Green AI metrics for every stage of the AI system lifecycle; 2) Evaluating AI architectures to ensure they are both energy-efficient and performant; 3) Assessing the full environmental impact of AI, from hardware creation to usage; 4) Innovating actionable Green AI metrics tailored for practical application; 5) Optimizing AI systems for high performance with minimal energy consumption; 6) Benchmarking AI embedding models within set computational limits, aiding SMEs; 7) Refining energy metrics for greener telecommunication networks; 8) Designing and validating methods for dynamic network capacity management; 9) Integrating Green AI principles into industrial development processes; 10) Sharing best practices and shifting industrial mindsets toward Green AI; 11) Developing and implementing new tools for measuring and minimizing AI's environmental impact.

EXPECTED RESULTS

The ENFIELD project is poised to benefit significantly from the focused research activities on monitoring green-AI metrics. The expected scientific progress includes the development of standardized, universally applicable green AI metrics. These will facilitate the evaluation of AI architectures not just on accuracy, but also on energy efficiency and carbon emissions, addressing the entire AI lifecycle including hardware manufacturing. Scientific results will encompass actionable recommendations for designing AI systems that balance energy efficiency with high performance. Moreover, the project will yield a suite of new tools and methodologies that enhance the environmental sustainability of AI applications, from edge devices to large-scale networks. Evaluation measures will be grounded in real-world scenarios, such as the energy-efficient operation of communication networks and the integration of green AI principles in industrial systems. By benchmarking AI architectures under standardized green metrics, the ENFIELD project will not only drive innovation but also lead to more sustainable AI practices across various sectors, setting new precedents for responsible and environmentally conscious technology development.
The beneficiary is expected to produce a technical report describing the methodology or tool being developed during the exchange. This technical report may lead to a peer-reviewed publication (conference, workshop, journal).

POSSIBLE HOST ORGANISATIONS

- IMT, Mines Saint-Etienne / Institut Mines-Télécom *(https://www.mines-stetienne.fr/en/)*
- TELENOR, Advanced Analytics & AI Department (https://www.telenor.com/innovation/research/analytics-and-ai/)
- SINTEF, Department of Sustainable Communication Technologies (https://www.sintef.no/en/digital/departments-new/department-of-sustainable-communication-technologies/)

Pillar Adaptative-AI

A-AI.1 Approaches to Incremental Learning Robustness and Trustworthiness

Keywords: Incremental learning; Evolving systems; Concept drifts; Change adaptation; Robustness and Trust

STATE-OF-THE-ART

Existing AI systems have difficulty adapting to dynamic environments. Adaptive AI systems are therefore needed in situations where rapid changes in the external environment or evolving corporate objectives demand an optimized response. The state-of-the-art AI systems currently in use have significant shortcomings, such as excessive data requirements, vulnerability to adversarial attacks, lack of robustness in the face of disruption, inability to talk about change, and so on. One of the main shortcomings of the de facto standard for deployed AI systems is that they assume that data is static, whereas it evolves. This requires the development of continual Adaptive AI systems that are versatile and robust to changes in the external environment, with a graceful degradation of consolidated knowledge. This requires both adaptive AI algorithms capable of learning on the fly with data/sample efficiency and streamlined AI infrastructure and engineering that enables on-device learning.

SCIENTIFIC CHALLENGES

- **Robust and Secure Continuous learning:** Continuously learning from data requires the inclusion of specific constraints to avoid undesirable behaviours such as catastrophic forgetting, incorporating noisy data or inaccurate labels. The multi-faceted boundaries of what is learned can be achieved through incremental procedures with normative control frameworks, which can explicitly encode the boundaries that the adaptive AI system can operate and learn on-line. Limiting continual learning therefore build trust and empowers humans to drive the learning process.
- **Adaptive** AI characteristics including dependability (reliability), robustness, versatility, and adaptability (graceful degradation).
- **Realization and identifiability of evolutionary models**, contrastive learning, or regulation-based approaches. Metrics to assess forgetting will also need to be addressed (typically for complexity, memory, accuracy, timeliness.
- **Hybrid-AI Systems (Event and Continuous dynamics):** Integration of arbitrary Knowledge into AI systems. Adaptive AI model has to handle high variation and volume of (most of the time unknown) objects and defects in recycling, refurbishment, and remanufacturing applications.
- **Cautious Classification in Dynamic Setting:** handling imprecision taking advantage of prior knowledge, particularly in sequential decision-making.
- **Adaptive Real-time Voice Conversion and Synthesis**: The primary challenge is to develop a real-time adaptive AI system capable of dynamically adjusting voice conversion techniques while maintaining voice consistency and quality.

RESEARCH ACTIVITIES

i) Development of Hybrid AI Systems: (1) Collection of knowledge bases on the Web, suitable for pretraining.; (2) Development of an approach (e.g. inspired by the transformer architecture) to pretrain knowledge models on large knowledge bases (e.g. following the RDF2Vec method); (3) Application and evaluation of the approach on the knowledge bases previously identified; (4) Adaptive model sensitive to concept drift and able to learn evolving processes; (5) Quantization of stochastic and epistemic uncertainties; (6) Synthetic data generation; **ii) Design of a new framework for Adaptive modelling:** (1) Learn new knowledge on the fly; (2) Interact with users in thrust; (3) Capable of dynamically adjusting signal and image techniques for human-centric AI; (4) Capable of enhancing interactions with humans by anticipation of behaviours and movements changes; **iii) Leveraging multiple modalities to enhance adaptability:** (1) Explore different modalities in deep learning; (2) Explore foundation models and ways to leverage the pre-trained big models; (3) Explore different learning objectives; **iv) Designing Cautious Classification in Dynamic Setting:** (1) Evaluation of existing approaches in different contexts; (2) Proposal and study of new approaches; **v) Implementing Adaptive Real-time Voice Conversion and Synthesis**: (1) Develop real-time adaptive AI models for voice conversation; (2) Collect different voice data for training and testing; (3) Implement a feedback mechanism to capture user preferences; (4) Evaluate system performance through user studies; (5) Refine the model interactivity.

EXPECTED RESULTS

The ENFIELD project aims to foster contributions on data-driven and knowledge-based methods for enabling adaptations in learning and to collaborate around shared frameworks that integrate concept drift, incremental learning, continual learning, uncertainty estimation, stochastic and epistemic model robustness. As examples, Brain-Inspired Adaptive AI model, Hybrid AI Systems, multiple modalities tracking, Adaptive Deep Reinforcement Learning, designing Cautious Classification in Dynamic Setting and implementing Adaptive Human, interaction systems.
The beneficiary is expected to have one scientific publication (ideally in a Q1 journal or A*/A rank conference).

POSSIBLE HOST ORGANISATIONS

- IMT, Mines Alès / Institut Mines-Télécom, Euromov DHM Lab (https://dhm.euromov.eu/)
- TU/e, Department of Mathematics and Computer Science (https://www.tue.nl/en/our-university/departments/mathematics-and-computer-science)
- UPB, National University of Science and Technology POLITEHNICA Bucharest, Artificial Intelligence and Multi-Agent Systems Laboratory (https://aimas.cs.pub.ro/)

A-AI.2 Advancing Adaptive AI on The Edge: Innovations for Sustainable, Efficient, and Continual Learning in Edge Computing
Keywords: Continual Learning (CL); On-Device Learning; Hardware-aware AI compression; Adaptive Deep Reinforcement Learning.
STATE-OF-THE-ART
Edge devices deployed in real-world scenarios are faced with constantly changing environments with non-stationary live-streaming data. These devices must therefore be Adaptive learners to cope with the changes around them. This requires the development of continual Adaptive AI systems that are versatile and robust to changes in the external environment, with a graceful degradation of consolidated knowledge. This requires both adaptive AI algorithms capable of learning on the fly with data/sample efficiency and streamlined AI infrastructure and engineering that enables on-device learning.
SCIENTIFIC CHALLENGES
• **Adaptive Continual On-Device Learning:** Developing adaptive continual on-device learning algorithms with limited hardware resources (memory, computing, energy, etc). • **Hardware-aware AI Model Compression:** Developing a framework that optimally compress and accelerate the AI model on different hardware types, considering hardware constraints related to computing, energy and memory capacities. • **Hardware-aware Adaptive Differential Neural Architecture Search:** Developing a framework that automatically searches for adaptive AI models on different hardware types, considering hardware constraints related to computing, energy, and memory capacities. • **Adaptive Deep Reinforcement Learning (DRL):** Control policies learned using DRL are specific to the learning environment in particular non-stationary environments (e.g. a robot will learn how to walk on a specific virtual terrain but fail to do so on another terrain). Efficient methods for adapting trained policies to new environments could help to use them in more versatile contexts. Strong links with sim2Real, i.e. training policies to be used in real settings, using virtual environments. • **Designing sparsification techniques (activation, representation, and gradient sparsity).** At the same time, the artificial backlog of maintaining coding for sparse memories need to be reduced.
RESEARCH ACTIVITIES
Development of Adaptive Continual On-Device Learning for edge computing: • Implementing adaptive AI model compression techniques for inference on different hardware types, considering hardware constraints related to computing, energy, and memory capacities. • Developing an energy-aware model selection strategy. • Developing adaptive differential neural architecture research that automatically searches for efficient deep AI model on different hardware types, considering constraints related to computing, energy and memory capacities. • Developing acceleration technique for training by reducing the FLOPs and memory usage. • Developing continual learning technique requirement minimal retraining, in particular Continual On-device Learning with limited hardware resources.
EXPECTED RESULTS
The ENFIELD project aims to revolutionize sustainable industry practices through Adaptive-AI at the edge. Development of Adaptive Continual On-Device Learning for edge computing including hardware-aware adaptive compression techniques, hardware-aware neural architecture research, it seeks to reduce the energy consumption of edge devices. It enables continuous learning in order to reduce retraining requirements and computing load. The beneficiary is expected to have one scientific publication (ideally in a Q1 journal or A*/A rank conference).
POSSIBLE HOST ORGANISATIONS
• BME, Speech Technology and Smart Interactions Laboratory (https://www.tmit.bme.hu/speechlab?language=en) • IMT, Télécom Paris / Institut Mines-Télécom (https://www.telecom-paris.fr/en/home) • NRS, Norsk Regnesentral Norwegian Computing Center, DART department (https://nr.no/en/areas/ict/) • SINTEF, Department of Sustainable Communication Technologies (https://www.sintef.no/en/digital/departments-new/department-of-sustainable-communication-technologies/) • TU/e, Department of Mathematics and Computer Science (https://www.tue.nl/en/our-university/departments/mathematics-and-computer-science)

Pillar Adaptative-AI

A-AI.3 Neuroscience Inspired Adaptive AI

Keywords: Continual Learning, Lifelong Learning, Brain Inspired AI, Multimodal Learning, Sparsity

STATE OF THE ART

Most of the existing Continual Learning methods are computationally expensive and ineffective. They fail to mimic the intricacies of the learning mechanisms and the interactions of multiple memory systems in the human brain which might account for the gap in the learning capabilities of existing AI and the human brain. More and more recent methods have drawn inspiration from the brain e.g. experience replay, synaptic consolidation and multiple memory systems which have shown promise and makes a strong case for further work in this promising direction.

SCIENTIFIC CHALLENGES

Identify the gaps between Humans and Existing AI: Our enhanced understanding of the brain from neuroscience studies and the advancement in AI presents a unique opportunity to revisit the design of DNNs to enhance their CL capabilities and generalization. The goal is to identify the key components of the learning machinery of the brain that enables it to excel at lifelong learning which are missing in existing approaches.

Design Brain Inspired CL Methods: Once we have Identified the critical components of the learning machinery and design of the brain that enables efficient CL in the brain, bring these insights to the design of DNNs.

Design Learning Mechanisms: The major challenges lie in faithfully designing the learning mechanisms and error-based learning that can maintain the delicate balance between the stability and plasticity of the model.

Design Multiple Memory Systems based on the Complementary Learning Systems Theory: implement multiple memory systems that aggregate information at different time scales such that information is effectively consolidated, and we can disentangle learning and remembering.

Representation Learning: Explore different representations (e.g. sparse, dense, orthogonal etc), learning objectives in each memory and their interactions.

Leverage Multiple Modalities: A salient feature of the brain that may play a critical role in enhancing its lifelong learning capabilities is that it processes and integrates information from multiple modalities. Hence, we aim to combine information from different modalities which allows the models to develop a more comprehensive understanding of the environment as it receives multiple views of the object, leading to a more accurate and robust representation, which is less sensitive to modality- specific regularities and shift in distributions.

Efficient Continual Learning using Sparse Predictive Coding: Reduce interference through the lens of sparsity. The overarching objective is to lead the charge in developing cutting-edge sparsification techniques—encompassing activation, representation, and gradient sparsity—meticulously crafted to enable efficient continual learning on resource constrained devices.

RESEARCH ACTIVITIES

Brain Inspired Architectures: i) Design Neuroscience inspired Deep Learning architectures that incorporates multiple brain inspired components in a complementary manner; ii) Incorporate components like context dependent processing of information, synaptic consolidation, replay, population coding, multiple memory systems etc.; iii) Explore different learning objectives and cognitive biases to enhance continual learning.

Complementary Learning Systems Theory: i) Explore different approaches to designing the short term and long-term memories inspired by the complementary learning systems theory in the brain; ii) Explore different representations (e.g. sparse, dense, orthogonal, etc.) in each memory and different interactions between the two memories. Inspired by the interactions between Hippocampus and Neocortex

Sparse Coding: Explore different approaches to incorporate sparsity in learning. This includes sparse representations, connections, and gradients. Take inspiration from sparse coding in the brain.

Multimodal Continual Learning: i) Explore different modalities in deep learning and approaches to leverage the complementary information in each modality to learn a more holistic and robust representation of the objects; ii) Explore foundation models and ways to leverage the pre-trained big models.

EXPECTED RESULTS

The ENFIELD project aims to bridge the gap between the capabilities of humans and existing AI. By leveraging insights from our enhanced understanding of the brain, it aims to design the next generation of brain inspired models that enable efficient and effective continual learning in deep neural networks. The goal is to designs AI models that can be deployed in our dynamic environment and meet the ever-changing industrial requirements.

The beneficiary is expected to have one scientific publication (ideally in a Q1 journal or A*/A rank conference).

POSSIBLE HOST ORGANISATIONS

- IMT, Télécom Paris / Institut Mines-Télécom (https://www.telecom-paris.fr/en/home)
- TU/e, Department of Mathematics and Computer Science (https://www.tue.nl/en/our-university/departments/mathematics-and-computer-science)

HC-AI.1 Evolving Symbolic Models for Decision-Making
Keywords: Symbolic AI; Reinforcement learning; Learning; Data-driven; Evolving.
STATE-OF-THE-ART
Neuro-symbolic learning uses context-free grammar (from automata theory) as a symbolic representation and learns from an oracle (i.e., artificial neural network trained with reinforcement learning) in a supervised learning setting (imitation learning)[2]. As far as we know, the only publication that does not use an oracle is[3], which conducts a program architecture search on top of a continuous relaxation of the architecture space defined by programming language grammar rules. Another research direction is iterative machine learning, where humans are part of the learning process and can tune hyperparameters of a meta-heuristic optimiser[4]. Genetic programming for symbolic regression is also an alternative for learning symbolic models from data but uses a tree-based representation for the knowledge that can be ineffective for complex symbolic structures. An interesting work is[5], where evolutionary search with a list of 65 basic mathematical operations is used to automatically discover ML algorithms from scratch with minimal human intervention.
SCIENTIFIC CHALLENGES
Grow symbolic models from data based on the interaction between the AI-decision system and the environment, where reinforcement learning can be used for constructing the model. The human can define a template for the symbolic model and/or participate in the learning phase (e.g., change hyperparameters, modify intermediary solutions).
RESEARCH ACTIVITIES
• Study different symbolic representations for control/decision problems (e.g., from automata theory). This representation might be domain-specific, which means having a domain-specific language for each use case. • Development of symbolic search methods based on well-established search-based algorithms such as simulated annealing or Monte Carlo Tree search. • Application of the developed approach in verticals (energy and/or healthcare).
EXPECTED RESULTS
The expected scientific progress includes the development of a new method for symbolic AI that is capable of learning from data within a reinforcement learning framework. The method can be used to augment existing expert systems in different domains, e.g., use the existing expert system as a template or starting point for the learning, or find new structures and symbolic representations for those systems. It should offer higher interpretability to humans since they are part of the learning process in three possible stages: (1) design of the model's template/structure, (2) modify or improve solutions during the learning phase (iterative learning), and (3) analyze and modify the final solution. The beneficiary is expected to have one peer-reviewed publication (preferably in a top-tier journal or conference) and the code published in open-source (GitHub).
POSSIBLE HOST ORGANISATIONS
• INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, Center for Power and Energy Systems (https://www.inesctec.pt/en/centres/cpes) • TU/e, Department of Industrial Engineering and Innovation Sciences (https://www.tue.nl/en/research/research-groups/innovation-sciences/human-technology-interaction or https://www.tue.nl/en/research/research-groups/industrial-engineering/information-systems-ieis)

[2] Verma, A., Murali, V., Singh, R., Kohli, P., Chaudhuri, S. (2018, July). Programmatically interpretable reinforcement learning. In International Conference on Machine Learning (pp. 5045-5054). PMLR.
[3] Qiu, W., & Zhu, H. (2021, October). Programmatic reinforcement learning without oracles. In International Conference on Learning Representations.
[4] Holzinger, A., Plass, M., Kickmeier-Rust, M., et al. (2019). Interactive machine learning: experimental evidence for the human in the algorithmic loop: A case study on Ant Colony Optimization. Applied Intelligence, 49, 2401-2414.
[5] Real, E., Liang, C., So, D., Le, Q. (2020, November). Automl-zero: Evolving machine learning algorithms from scratch. In International conference on machine learning (pp. 8007-8019). PMLR.

Pillar Human-Centric-AI

HC-AI.2 Novel Explainable AI Methods for Decision Making

Keywords: Explainability; Spatio-temporal Models; Decision making; Healthcare; AAL

STATE-OF-THE-ART

Models such as spatio-temporal graph neural networks or Visual Transformers are powerful tools for modelling spatio-temporal dependencies in multi-modal and spatio-temporal contextual relationships. Despite the many approaches proposed in the literature, Human Action Recognition (HAR) from video sequences is still a challenging research task, especially when performed in a real-world environment, for example in Healthcare, Ambient Assistive Living applications or Manufacturing. Different DNN architectures, such as GNN, TCN, or Spatial Temporal GCN, have been proposed for solving the HAR problem. Recently, Vision Transformers have emerged as a promising approach for HAR, with models such as Video action Transformer Network, ConvTransformer Network, or Spatio-Temporal Attention Network. However, their black box nature limits their interpretability for trustworthy decision making. There are still few approaches that try to explain how the network arrived at a decision or, most important, why it failed to predict the correct result, and, as far as we know, none such approach for ViT-based HAR.

SCIENTIFIC CHALLENGES

- Understand user's requirements for explainability and interpretability of DNN models for HAR.
- XAI methods for Spatial Temporal GCN predictions.
- XAI methods for Vision Tranformers predictions.
- Design a high-level representation structure to capture the spatial and temporal dimension of human action, as a basis for linking DNN explanations to human decision making.
- Conceive interpretable explanations linked to high level representations that are able to capture changes in the environment of an application.

RESEARCH ACTIVITIES

- Design and implement explainable algorithms for Spatial Temporal based models and validate them on different HAR contexts.
- Explore model specific and model agnostic XAI methods to capture the flow of information and dynamics in spatio-temporal structures.
- Explore different approaches to link explanations to symbolic structures.
- Design a symbolic framework to allow explanation and interpretation of decision-making.
- Evaluate from a qualitative and from a quantitative point of view the quality of explanations.

EXPECTED RESULTS

The expected scientific progress includes the development of high-level XAI models for decision-making applications based on spatio-temporal models. These models may be used in different applications, for example Human Action Recognition in Healthcare, AAL or Manufacturing (Human-AI collaboration). The beneficiary is expected to develop demos for these models in one or more of the above-mentioned application domains and to produce one peer-reviewed publication (preferably in a top-tier journal or conference)

POSSIBLE HOST ORGANISATIONS

- INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, Center for Power and Energy Systems (https://www.inesctec.pt/en/centres/cpes)
- TU/e, Department of Industrial Engineering and Innovation Sciences (https://www.tue.nl/en/research/research-groups/innovation-sciences/human-technology-interaction)
- UPB – National University of Science and Technology POLITEHNICA Bucharest, Artificial Intelligence and Multi-Agent Systems Laboratory (https://aimas.cs.pub.ro/)

HC-AI.3 Interpretable Data-Driven Decision Support Systems
Keywords: Interpretable decision making; Automatic decisions; Collaborative human decisions; Integrated collaborated environment; Medical domain
State-of-the-art
Decision making in the medical domain is a challenging task, especially when different specialists and factors must contribute to the decisions. AI based tools for helping the decision are invaluable in this respect. However, collaboration between doctors and such tools is currently difficult because of different reasons. An integrated collaborated environment in which humans and AI tools collaborate is an opportunity to alleviate these difficulties.
Scientific Challenges
• Trustworthy collective human decisions in healthcare supported by AI tools. • Human-oriented explanations of automatic decisions in diagnostic based on medical data. • Explore different metrics to evaluate explainability and interpretability of medical decisions supported by AI tools.
Research activities
• Design and implement an explainable decision support system for diagnosis based on medical images (CT or MRI). • Develop an approach to support medical decisions performed by different specialists contributing to the decision (based on medical data, human factors, context, history of the patient). • Design and implement algorithms to enhance the explainability of automatic decisions based on different medical data to support collaborative human decisions. • Evaluate the effectiveness of the approach for different use cases.
Expected results
The expected scientific progress includes the development of a supporting environment in which AI tools and humans (from different specialities) are able to collaborate to make decisions. The beneficiary is expected to develop a demo for such a decision-making support system environment and to produce one peer-reviewed publication (preferably in a top-tier journal or conference).
Possible Host Organisations
• TUC, Chemnitz University of Technology, Distributed and Self-organizing Systems Group (https://vsr.informatik.tu-chemnitz.de/research/) • TU/e, Department of Industrial Engineering and Innovation Sciences (https://www.tue.nl/en/research/research-groups/innovation-sciences/human-technology-interaction or https://www.tue.nl/en/research/research-groups/industrial-engineering/information-systems-ieis) • UPB, National University of Science and Technology POLITEHNICA Bucharest, Artificial Intelligence and Multi-Agent Systems Laboratory (https://aimas.cs.pub.ro/)

T-AI.1 Modelling Trust in Distributed AI System Architectures

Keywords: Trustworthy AI; Distributed Systems; Trust Modelling; Software Architecture; Method Engineering

STATE-OF-THE-ART

The rapid adoption of large language models (LLMs) has created a side effect of using AI-generated content in different application domain. The unregulated use of such approaches, however, can potentially lead to malicious consequences such as plagiarism, generating fake news, spamming, identity theft/spoofing, etc. Furthermore, among the major cons of AI-generated content are the lack of trust of the final outcome. Therefore, reliable detection of AI-generated text can be critical to ensure the responsible use of LLMs.

SCIENTIFIC CHALLENGES

AI components are used in different parts of complex distributed or even federated systems. This raises the challenge of modeling trust in distributed AI system architectures due to the additional interaction between AI components communicating with each other and with other non-AI components of the system. Trust modeling in this context needs to balance modeling architectural complexity of distr. AI systems with a representation useful for non-IT stakeholders while supporting automatic analyses. There is no established theoretical foundation of trust aspects in distributed AI system architectures yet, which is required as basis for the creation of a suitable modeling method. To design a method that can be employed by practitioners in industry contexts, not only the analytical algorithmic aspects are important, but also the visual representation and the effort for model creation and management. To address the challenge of modeling trust in distributed AI systems, interdisciplinary collaboration of researchers from different fields such as software architecture, security, or human-computer interaction is required.

RESEARCH ACTIVITIES

- Creation of a taxonomy of trust in distributed AI system architectures
- Specification of a suitable visual modelling language
- Development of infrastructure supporting the modelling
- Design of algorithms for automatic analyses
- Evaluation of trust modelling in distributed AI systems

EXPECTED RESULTS

We expect the exchange to provide valuable contributions to the long-term goal of designing a method for architectural trust modelling in complex distributed AI systems. The method will facilitate the creation of distributed “trustworthy by design” AI systems by enabling system architects to document and analyse trust in their architectural system blueprints. For researchers, the results will contribute to establishing a common vocabulary and representation of trust in distributed AI systems as a first step to consolidate the body of knowledge in this relatively young field and facilitate the communication and thus collaboration. These results will be published in at least 1 conference publication for discussion with experts in the same area. The exchange also aims at fostering knowledge transfer and networking with other groups working in related fields such as information systems, distributed systems and software engineering.

POSSIBLE HOST ORGANISATIONS

- ICCS, Institute of Communication and Computer Systems, Computer Networks Laboratory (https://www.cn.ntua.gr/)
- NTNU, Norwegian University of Science and Technology, Critical Infrastructure Security and Resilience group (https://www.ntnu.edu/iik/cisar)
- TUC, Chemnitz University of Technology, Distributed and Self-organizing Systems Group (https://vsr.informatik.tu-chemnitz.de/research/)

Pillar Trustworthy -AI

T-AI.2 Detection of AI Generated Content
Keywords: AI content; Generative AI; LLM; Trust; Big data
STATE-OF-THE-ART
The rapid adoption of large language models (LLMs) has created a side effect of using AI-generated content in different application domain. The unregulated use of such approaches, however, can potentially lead to malicious consequences such as plagiarism, generating fake news, spamming, identity theft/spoofing, etc. Furthermore, among the major cons of AI-generated content are the lack of trust of the final outcome. Therefore, reliable detection of AI-generated text can be critical to ensure the responsible use of LLMs.
SCIENTIFIC CHALLENGES
The main research challenges relate to (1) watermarking content, (2) retrieval-based defenses, (3) paraphrasing attacks (4) spoofing attacks, (5) measuring randomness in LLMs, (6) data quality and fairness, (7) computational power when big data sets are used and analyzed, (8) sensitive personally identifiable information, (9) NLP models with complex reasoning abilities and interpretability, (10) transparency of large language models.
RESEARCH ACTIVITIES
Research on the detectability of the output of currently used LLM models. Evaluate the effectiveness of known defence techniques. Test current attack approaches. Methodology and metrics to evaluate LLM randomness. Report on existing LLMs defences. Novel attack techniques. To ensure the trustworthiness of the generative content, methods and tools need to be investigated able to capture aspects such as intended tone, flow, and context.
EXPECTED RESULTS
The ENFIELD will leverage novel scientific results to increase the trustworthiness AI. By leveraging the results from this topic directions and guidelines towards the development of a trustworthy AI framework for EU will be facilitated. In addition to that, the involved partners and research will collaborate, exchange knowledge, and expertise to further develop their research activities and future collaborations. The beneficiary is expected to have one peer-reviewed publication (conference, workshop, journal).
POSSIBLE HOST ORGANISATIONS
• BME, Speech Technology and Smart Interactions Laboratory (https://www.tmit.bme.hu/speechlab?language=en) • NTNU, Critical Infrastructure Security and Resilience group, (https://www.ntnu.edu/iik/cisar) • UoN, School of Computer Science (https://www.nottingham.ac.uk/computerscience/research/index.aspx)

T-AI.3 Secure Voice Biometrics with Fake Voice Detection

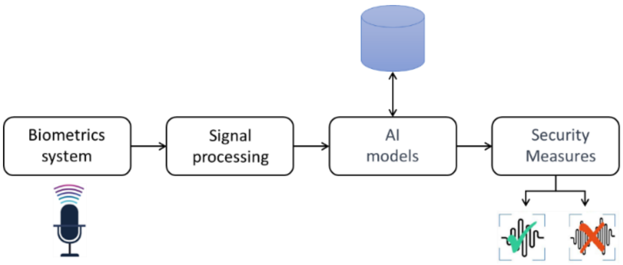
Keywords: Voice spoofing; Biometric security; Speech signal processing; Robust authentication; Acoustic analysis.

STATE-OF-THE-ART

Current voice biometric systems are at the central of biometric authentication, but they face increasing concerns related to data privacy and security. The rise of fake voice generation technologies presents a significant challenge to the integrity of voice biometrics. State-of-the-art solutions in this field are actively addressing the need to develop robust defences against not only traditional security risks but also the voice spoofing and deepfake technologies. As the use of voice biometrics continues to expand in applications like access control, financial transactions, and identity verification, it is essential to address these scientific challenges and opportunities to ensure the trustworthiness and reliability of voice-based authentication methods.

SCIENTIFIC CHALLENGES

- Developing AI models to enhance the security and trustworthiness of voice biometrics is a complicated task. This requires the creation of superior algorithm capable of distinguishing real from synthetic voices. This challenge requires the combination of cutting-edge deep learning techniques with voice recognition to continuously adapt and secure against emerging threats. This involves dealing with various accent and language variations, background noise, and voice quality issues.
- Protecting sensitive voice data is another crucial aspect. This challenge involves developing mechanisms that safeguard stored and transmitted voice samples. It needs a deep understanding of data encryption and secure communication protocols designed to voice biometrics.
- Addressing voice spoofing is important because it directly impacts the reliability and security of voice biometric systems. As voice authentication becomes more common in various sectors, including finance, healthcare, and access control, the threat of voice spoofing presents a significant risk. Developing robust anti-spoofing techniques is necessary to ensure the trustworthiness and integrity of voice-based security measures, maintaining user confidence in the technology and safeguarding sensitive information against fraudulent activities. This challenge needs advanced signal processing, machine learning, and behavioural analysis methods.



- Advancing techniques for the detection of fake voice samples requires exploring speech characteristics and signal analysis. This challenge requires not only identifying synthetic voice attributes but also understanding how these attributes differ from natural human speech. This challenge needs deep learning, feature engineering, and acoustic analysis to design more accurate and reliable fake voice detection methods.

Research activities

- Develop AI models for secure voice biometrics, integrating encryption, privacy-preserving methods, and fake voice detection.
- Investigate methods for detecting and preventing voice spoofing, as well as the generation of fake voice samples.
- Perform accurate testing to ensure the system's trustworthiness, reliability, and fake voice detection capabilities.

Expected results

- At least one scientific publication (conference paper)
- Expected results involve the investigation of novel methods to detect and prevent voice spoofing, ensuring the system's strength against manipulated voice samples. Moreover, accurate testing will be conducted to verify the system's trustworthiness, reliability, and its capabilities in detecting fake voices.

Possible Host Organisations

- BME, Speech Technology and Smart Interactions Laboratory (https://www.tmit.bme.hu/speechlab?language=en)
- ICCS, Computer Networks Laboratory (https://www.iccs.gr/)
- NTNU, Critical Infrastructure Security and Resilience group, (https://www.ntnu.edu/iik/cisar)
- UoN, School of Computer Science (https://www.nottingham.ac.uk/computerscience/research/index.aspx

Vertical Energy

VE.1 Methods of Explainable Machine Learning applied to LiDAR Scan Analysis

Keywords: LiDAR (Light Detection And Ranging); Green AI; Explainable AI; overhead lines.

State-of-the-art

Electrical grids' careful inspection an important and challenging problem. Often, it is based on LiDAR large-scale point clouds with high-point density, no sparsity and small object occlusion. The captured point clouds are quite extensive and mostly composed of arboreal areas which make the task of transparently detect objects such as power grid poles a hard task. One plausible way to approach this problem is to employ 3D semantic segmentation methodologies. State-of-the-art represents 3D scenes as volumetric grids (Wu et al. 2015; Maturana and Scherer 2015) and as 3D point clouds (Qi et al. 2017a,b; Thomas et al. 2019). Volumetric methods allow for the use of global feature descriptors, such as 3D convolutions but are restricted in terms of resolution due to the cubic growth of computational complexity and memory footprint. They also introduce challenges to ML models, namely heterogeneous density, lack of structure and permutation invariance. Most proposals are tailored to boost performance in urban settings (e.g., Semantic3D (Hackel et al. 2017), SensatUrban (Hu et al. 2021) and SemanticKITTI (Behley et al. 2019)), where data are sparse, objects are often occluded and may demonstrate anisotropy w.r.t. density – this is not the case of the challenge we propose to overcome.

Scientific Challenges

We propose SCENE-Net, an intrinsically interpretable 3D point cloud semantic segmentation framework identifying signature geometric shapes via group equivariant non-expansive operators (GENEOs) that allow for fast training even with a small amount data, and robustness both to labelling noise and strong imbalance.
We expect to reduce the training time on a regular laptop below 1 and half hour for 40 000 km of overhead linesand inference time to around 20 ms.
We propose to use a SCENE-Net composed of just 11 trainable geometrical parameters (like the radius of a ball or the height of a Cylinder), reaching a Precision gain of 24% against a comparable CNN with more than 2000 uninterpretable parameters.

Research activities

Within the PhD thesis work, between ENFIELD beneficiary CNET and externals: EDP LABELEC, Portugal, Nova-FCT, Portugal and Univ. Milano, Italy, the student plans to work on the following items:
(1) interpretability of the model.
(2) improvement of accuracy and precision of the model.
(3) robustness to noisy labels.
(4) training and inference time improvement.
(5) inference performance to high resolution from train with low resolution voxel-grids.
We will benchmark our model at least against a traditional CNN with similar architecture. Further details about the dataset, the samples used, the training protocol, and ablation studies are expected to be shared, discussed, and presented in public events.

Expected results

- 2 scientific publications, one in conference and another one in scientific journal.
- Software application for object automatic georeferencing and classification for industrial use.
- Green AI benchmark and scale-up impact analysis on energy used for processing and correspondent CO_2 reduction.

Possible Host Organisations

- EDP CNET (https://www.edp.com/en/innovation/NEW)
- EDP LABELEC (https://www.edp.com/pt-pt/inovacao/labelec#sobre-a-labelec)

Vertical Energy

VE.2 Coordinated Edge Control of Electric Vehicles Charging at Low Voltage Grid (or Microgrids)

Keywords: Electric vehicles; Edge intelligence; Optimization; Renewable energy; Microgrid.

STATE-OF-THE-ART

Simultaneous charging of multiple electric vehicles (EVs) poses a potential challenge to the low voltage local grid (microgrid), diminishing its hosting capacity and creating a bottleneck in the pursuit of decarbonizing the mobility sector[6]. This issue necessitates a coordinated approach with other resources tied to the electrical grid, such as local photovoltaic (PV) panels and small-scale storage. Intelligent EV charging strategies become imperative to effectively manage charging rates and schedules, utilizing local data from EV Supply Equipment (EVSE) and taking into account grid operating conditions, electricity tariffs, and the expectations of EV drivers.

While commercial EVSEs commonly adhere to standards like ISO 15118, offering features such as secure communication and smart charging[7], the control of smart charging is typically centralized due to high computational demands associated with executing functions at the edge (i.e., at the EVSE). This centralized approach requires a predictive model to accommodate drivers' preferences, namely target state-of-charge and departure hour. Additionally, forecasting the charging requirements of electric vehicles proves challenging due to irregular time series data, requiring the prediction of multiple parameters to accurately quantify flexibility in EV charging[8].

SCIENTIFIC CHALLENGES

The main scientific challenge is to develop frugal (or green) data-driven control solutions (e.g., based on EV forecasting, and grid operating conditions) that enable distributed intelligence and local control at the EVSE level, and coordinated control between multiple EVSE (and other resources like PV and storage). Ensuring data privacy and security is a fundamental requirement, but the coordinated control inteligence should be have low computational requirements to facilitate wide adoption from EVSE manufacturers, e.g., a blend of a rule-based and data-driven approach with capacity to adapt to new operating conditions.

RESEARCH ACTIVITIES

Apply AI techniques (see 'Brief description of the scientific challenge') to smart electric vehicles charging, considering that the functions should run locally at the edge.

At the core of the testing capabilities is the Smart Grids and Electric Vehicles Laboratory (SGEVL) at INESC TEC premises. SGEVL has two configurable physical microgrids with islanding capabilities, which can be extended to the virtual domain using a Power-Hardware-in-the-Loop (PHIL) setup based on an OPAL-RT real-time digital simulator located in SGEVL. Associated with it, there are numerous physical assets such as loads, energy storage, PV generators, and a smart metering infrastructure. These installations integrate an EV charging infrastructure with commercial EV chargers and built-in-house EV charging prototypes with edge computing capabilities. Using the Iskraemeco Edge modules (and their charge controller that hosts the Edge module) for the distributed AI/ML part is also an option to explore in this work.

EXPECTED RESULTS

- Expected impacts are increase network hosting capacity of EVs (postponing network reinforcement) and promotion of the use of renewable energy for EV charging.
- 1 scientific publication (ideally in a journal) .
- 1 new method (in open source) for optimizing the electric vehicle charging considering local grid constraints and driver's requirements.

POSSIBLE HOST ORGANISATIONS

INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, Center for Power and Energy Systems (https://www.inesctec.pt/en/centres/cpes), in collaboration with Iskraemeco (https://iskraemeco.com/) Iskraemeco (https://iskraemeco.com/)

---

[6] Lopes, J. A. P., Soares, F. J., Almeida, P. M. R. (2010). Integration of electric vehicles in the electric power system. Proceedings of the IEEE, 99(1), 168-183.

[7] Schmutzler, J., Wietfeld, C., Andersen, C. A. (2012, October). Distributed energy resource management for electric vehicles using IEC 61850 and ISO/IEC 15118. In 2012 IEEE Vehicle Power and Propulsion Conference (pp. 1457-1462). IEEE.

[8] Bessa, R. J., Matos, M. A. (2013). Global against divided optimization for the participation of an EV aggregator in the day-ahead electricity market. Part I: Theory. Electric Power Systems Research, 95, 309-318.

Vertical Space

VS.1 Fast and Accurate Atmospheric RT Simulations for Satellite Microwave Instruments

Keywords: Physics; Radiative Transfer; Satellites; Weather; Efficiency

STATE-OF-THE-ART

Atmospheric radiative transfer models (RTMs) are necessary for some satellite- and ground-based measurements. The fundamental component of RTMs is a partial integro-differential equation, the radiative transfer equation. Models aimed at time-critical operational applications, such as weather forecasting, use physical simplifications and coarse discretization to obtain solutions efficiently at the expense of accuracy. The use of deep has been suggested as an alternative to the manual simplifications[9]. However, the complexity of the interaction between our atmosphere and the electromagnetic spectrum formulates a significant challenge that requires detailed inspection. In particular, at the microwave region, where several upcoming satellite instruments will measure hydrometeor properties characterized by non-spherical shapes, which are usually neglected to reduce the computational time, machine learning approaches are yet to be explored.

SCIENTIFIC CHALLENGES

Using a reference RTM[10] the challenge is to develop a system through machine learning which is cable to find accurate and fast solutions to the radiative transfer equation for the microwave and submillimetre region under realistic all-sky conditions. This system should be compared to an operational RTM [11,12] to assess any improvement by machine learning. The researcher will be provided with the required RTM simulations. Satellite data or climatological data can be downloaded on the fly.

RESEARCH ACTIVITIES

- Understanding the differences between fast and accurate RTMs, including the implications of the simplified physics from the fast RTMs.
- Selection and compilation of the observational data available online required for assessing or developing the machine learning system.
- Survey of physics-informed machine learning literature and related works.
- Development of the described ML system with either pure supervised learning or more elaborate techniques, such as physics-informed machine learning[13]

Assessment of any advantage of the resulting system over conventional manual simplifications.

EXPECTED RESULTS

The expected scientific progress includes, among others, the development and publication of a prototype of an AI tool, which can be used by researchers and meteorological agencies to improve weather forecasting or climatological studies and, consequently, decision making. The capabilities of this prototype and scientific findings are expected to be disseminated with one conference presentation and one scientific publication.

POSSIBLE HOST ORGANISATIONS

- CHALMERS, Chalmers University of Technology, Department of Space, Earth and Environment, division of Geoscience and Remote Sensing (https://www.chalmers.se/en/departments/see/research/geo/)

---

[9] Stegmann et al., 2022. "A deep learning approach to fast radiative transfer", J. Quant. Spectrosc. Radiat. Transf., https://doi.org/10.1016/j.jqsrt.2022.108088.
[10] Buehler et al., 2018. "ARTS, the atmospheric radiative transfer simulator — version 2.2, the planetary toolbox edition", Geosci. Model Dev., https://doi.org/10.5194/gmd-11-1537-2018.
[11] Barlakas et al., 2022. "On the accuracy of RTTOV-SCATT for radiative transfer at all-sky microwave and submillimeter frequencies", J. Quant. Spectrosc. Radiat. Transf., https://doi.org/10.1016/j.jqsrt.2022.108137.
[12] Saunders et al., 2018. "An update on the RTTOV fast radiative transfer model (currently at version 12)", Geosci. Model Dev., https://doi.org/10.5194/gmd-11-2717-2018.
[13] Mishra et al., 2021. "Physics informed neural networks for simulating radiative transfer", J. Quant. Spectrosc. Radiat. Transf., https://doi.org/10.1016/j.jqsrt.2021.107705

Co-funded by the European Union
INOVA+ RESEARCH & TECHNOLOGY INNOVATION FOR GROWTH

Vertical Space

VS.2 Generative Models for 3D Cloud Fields

Keywords: Generative Modelling; Cloud Properties; Atmosphere; Reconstruction Algorithms; Satellites

STATE-OF-THE-ART

The CloudSat satellite has been the gold standard for obtaining cloud vertical structures on a global scale. However, limitations in the satellite orbit and instrument hinder the use of CloudSat data for satellite data simulators that need 2D or 3D atmospheric input data. Leinonen et al.[14] studied the reconstruction of cloud fields using a conditional GANs trained against CloudSat data and conditioned on MODIS observations and auxiliary data, e.g., forecast data. They highlight limitations of their approach.
GANs catalysed the interest of the scientific community to push the development and application of generative models where, for example, diffusion models have arisen as a popular alternative to GANs. Generative models are yet to be exploited for generating atmospheric cloud fields.

SCIENTIFIC CHALLENGES

Development and evaluation of two generative models: an unconditional generative model and a conditional one. The emphasis would be in investigating diffusion models given its recent success, but any generative model, including GANs, can be considered. The models would be trained using CloudSat as a reference data, and the conditional model should only use public hyperspectral data. The models should, ideally, be able to generate 3D cloud fields, as opposed to the 2D cloud fields observed by CloudSat. The main challenge is, thus, to investigate if advances in the field of generative modelling can offer a better performance than the model presented by Leinonen et al.[9] as well as joint stochastic 3D retrievals of atmospheric cloud fields. Any satellite data to be used is publicly available.

RESEARCH ACTIVITIES

- Survey of generative models for atmospheric applications and development of the two models described above.
- Survey of conventional cloud field reconstruction algorithms to be used for future satellite missions[15].
- Selection of the satellite data to be used, publicly available.
- Ensure physical realism of the machine learning models.
- Comparison of the developed conditional generative model with a discriminative model which offers marginal distributions[16]

EXPECTED RESULTS

The expected scientific progress includes, among others, the development and publication of a prototype of an AI tool, which can be used by researchers to simulate cloud fields, thus mitigating the need for costly and limited real world observations. The capabilities of this prototype and scientific findings are expected to be disseminated with one conference presentation and one scientific publication.

POSSIBLE HOST ORGANISATIONS

- CHALMERS, Chalmers University of Technology, Department of Space, Earth and Environment, division of Geoscience and Remote Sensing (https://www.chalmers.se/en/departments/see/research/geo/)

[14] Leinonen et al., 2019. "Reconstruction of Cloud Vertical Structure With a Generative Adversarial Network", Geophys. Res. Lett., https://doi.org/10.1029/2019GL082532.
[15] Barker et al., 2011. "A 3D cloud-construction algorithm for the EarthCARE satellite mission", Q. J. R. Meteorol. Soc., https://doi.org/10.1002/qj.824
[16] Amell et al., 2023. "The Chalmers Cloud Ice Climatology: Retrieval implementation and validation", under review, https://github.com/SEE-GEO/ccic.

VS.3 Cost-Effective Precipitation Retrievals

Keywords: Rain; Uncertainty Quantification; Convnets; Retrieval; Geostationary Satellite.

STATE-OF-THE-ART

ML approaches have several advantages to the retrieval of precipitation, i.e., historical estimation of precipitation rates, over conventional approaches when considering satellite imagery[17]. Pfreundschuh et al. presented a method for atmospheric retrievals, which is an alternative to flexible but expensive statistical approaches[18], that can describe the uncertainty in the retrieval due to data variability, eliminating the need for ensemble predictions. Using this method, Amell et al. presented in a similar approach to but where only the European geostationary satellite was used[11]. In this case, they focused only on the effectivity of the neural-network retrieval but not its efficiency, thus requires non-minimal computational resources to run inference for, e.g., the whole African continent. Furthermore, they solely targeted Africa despite the satellite covers a larger area and did not use any time dimension.

SCIENTIFIC CHALLENGES

The challenge consists of analysing the shortcomings presented by Amell et al.[11] with a focus on exploring an inexpensive neural network architecture that offers at least a similar performance as the CNN used in and which offers a case-specific retrieval error. Africa should be included in the area supported for the retrievals, with the possibility to extend them to the full disc. A training dataset will be assembled consisting of Meteosat infrared observations labelled with the latest precipitation rate estimates from the Global Precipitation Measurement Core Observatory.

RESEARCH ACTIVITIES

• Compilation of neural network architectures that have a small computational footprint.
• Training and comparison of the performance of such neural network models.
• Assessment of the difficulty to incorporate the temporal domain in the retrievals through, e.g., autoregressive models, RNNs or semi-supervised learning, while maintaining a small computational footprint.
• Formal evaluation of the distributions retrieved with the method presented by Pfreundschuh et al.[13] to assess whether alternative methods, e.g., predicting distribution parameters from a family of distributions, offers a more efficient retrieval without degrading performance.
• Validation of the retrievals against independent datasets, coming from other retrieval schemes, accumulated rain, or sub-hourly rain rate estimates. The latter can be difficult to accomplish due to the availability of suitable data.

EXPECTED RESULTS

The expected scientific progress includes the development and publication of a ML-based, open-source retrieval model which can be used as an operational tool, as a benchmark, for example, for approaches that target local areas, or for helping to develop upcoming satellites retrieval schemes. This tool and any scientific findings are expected to be disseminated with one conference presentation and one scientific publication.

POSSIBLE HOST ORGANISATIONS

- CHALMERS, Chalmers University of Technology, Department of Space, Earth and Environment, division of Geoscience and Remote Sensing (https://www.chalmers.se/en/departments/see/research/geo/)

[17] Pfreundschuh et al., 2022. "An improved near-real-time precipitation retrieval for Brazil", Atmos. Meas. Tech., https://doi.org/10.5194/amt-15-6907-2022

[18] Pfreundschuh et al., 2018. "A neural network approach to estimating a posteriori distributions of Bayesian retrieval problems", Atmos. Meas. Tech., https://doi.org/10.5194/amt-11-4627-2018.

VM.1 Automatic and Efficient Identification of Batch Production Patterns for Tool Machines

Keywords: Condition-based maintenance, tool machines, Health management, time series

STATE-OF-THE-ART

The development of condition-based maintenance practices for tool machines is leading to a better understanding of wear and failure phenomena, and their characterization has been identified as a key factor in optimizing unit shutdowns and improving plant safety. Once the data has been collected and pre-processed, it is necessary to label the time series to identify the different batch production patterns to characterize the "health condition" of the equipment and optimize maintenance processes, considering spatial and temporal distortion and given that only very few patterns are labeled. Methods usually used include expert systems, e.g., rule-based or closed-form control methods, DTW algorithms.

SCIENTIFIC CHALLENGES

IA algorithms for edge to cloud continuum and AI embedded systems. AI based methods to improve data sample efficiency, number of learning parameters through effective regularization schemes. Adaptive IA on the edge exploiting historical data to perform pattern recognition e.g reinforcement learning, brain-inspired algorithms in continual learning, learning under noisy labels, automated transfer training to address the problem of very few or no labelled data.

RESEARCH ACTIVITIES

State-of-the art of technics and methods used for data labelling in the manufacturing field. Using data made available by Predict, development of innovative IA approaches for automatic labelling of batch patterns using historical sets of data. Apply green IA on the edge-to-cloud continuum to perform data pre-processing, data fusion for different sources and AI-big data analytics at the edge of the network and enabling optimization of cloud computing. Cloud platforms then perform further enrichment, aggregation and running complex analytics on the filtered data such as classification.

EXPECTED RESULTS

- 1 scientific publication (conference, workshops).
- An IA-based solution to improve the characterization of the phase's labels and better efficiency of the recognition process.
- IA algorithm embedded to perform all or part of the processing on the edge minimizing hardware power consumption, IoT protocol efficiency, communication overhead, data storage energy usage.

POSSIBLE HOST ORGANISATIONS

- POLIMI, Politecnico di Milano, Department of Management, Economics and Industrial Engineering (https://www.som.polimi.it/en/the-school/about-us/dig/)
- PREDICT R&T department(https://www.predict.fr/)

Vertical Manufacturing

VM.2 Self-X Integration in manufacturing domain
Keywords: manufacturing, self-X, autonomous computing, MAPE-K
STATE-OF-THE-ART
One of the most experienced solutions to these issues is the iteration of the training phases of algorithms, but this mitigation presents either high resource consumption or can lead to catastrophic interferences, which constitute a severe risk for the performances and for addressing responsibilities. In recent years, the practitioners' community resumed however frameworks and requirements from the Control domain (namely MAPE-K and Self-X) to grant the controlled system, the capability to self-adapt to unpredictable events.
SCIENTIFIC CHALLENGES
The recent advancements in data production and analysis in manufacturing allowed several Machine Learning techniques to be exploited in the production domain in several aspects (maintenance, scheduling, decision making...). One of the limits of these approaches, however, sits in the lack of resilience of the aforementioned techniques towards unpredicted and unpredictable events, which drift the performances of these algorithms far from the region they were trained/designed for (e.g., a new production recipe is introduced, and the algorithms are not able to classify the system status).
RESEARCH ACTIVITIES
Starting from a defined and centralised software architecture, the proposed solution is supposed to be able to make AI pipelines able to deal implement self-X capabilities. The solution is supposed to be tailored onto a lab-scale production environment and to deal with non-PLC signals (e.g., energy consumption) clustering the production in new defined classes. Alternatively, real-like industrial datasets are also available.
EXPECTED RESULTS
• At least one journal scientific publication (Scopus-indexed). • Development of software modules implementing self-X functionalities. • Experimentation in laboratory environment and/or from existing datasets.
POSSIBLE HOST ORGANISATIONS
• POLIMI, Politecnico di Milano, Department of Management, Economics and Industrial Engineering (https://www.som.polimi.it/en/the-school/about-us/dig/)