THE Φ-LAB EXPLORE OFFICE

Learning from Earth Observation for a more sustainable planet

Φ-lab team (2020) Φ-lab Explore Office Research & Innovation (R&I) Activities, Focus and Strategy

esa

FOREWORD

Could AI for Space be the next big race? Today, it sounds like a dream but in a few years it will be reality. Exploring this is a challenge, but also a great opportunity for Europe to drive innovation, science, and businesses. EO and AI are a perfect match. Europe has one of the best Earth Observation capabilities world-wide based on data from Sentinels, Earth Explorers, meteorological or commercial missions. They produce large continuous streams of unique data, taking the pulse of our planet. Classical analysis methods have reached their limits. AI offers a completely new view but also automates and accelerates the processing chain. To fully benefit from the wealth of these data, AI and EO will need to combine their strengths. A new way to combine Machine Learning with geo-physical modelling will help our scientists to build a "Digital Twin Earth". This should help citizens and decision-makers to preserve our planet.

This new frontier at the intersection of AI and space is certainly a very exciting space to watch, but also to unleash your imagination and turn it into innovation.

Josef Aschbacher
Director of ESA Earth Observation Programmes

Earth Observation (EO) deals with acquiring and processing information about our planet and takes advantage of all advances in sensing and computing technologies. Recently, computing – or, more widely, digital – technologies have undergone a revolution, affecting especially the EO downstream sector, for example applications such as big data analytics, as well as the upstream sector, with new approaches to small satellite constellations, in addition it is impacting the intersection of these sectors in the direction of a seamless sensing system-of-systems. One small example of such new technology is represented by new types of low-cost sensors capable of automatic re-programming through machine learning techniques. This fast pace of change calls for us to revisit our perspective, returning to the basics of “why” we realise today’s space systems as we do and calling for a more end-to-end but agile approach to shaping the new generation of missions, which respond to rapidly evolving user needs, even those yet to be articulated. The Φ-Department aims to bring this end-to-end and insightful perspective to the future of EO. Its Φ-lab provides the means to experiment in that vision, by objectively assessing both the strengths and weaknesses of new digital technologies such as AI. This document invites you to explore some of the new scientific and societal applications of AI applied to EO (AI4EO). It describes new tools to mine large amount of data and reviews some of the results of the research team, which are just a few elements of a much wider community effort. I wish you an inspiring journey in the rapidly growing field of AI4EO!

Pierluigi Silvestrin
Head of the Future Systems Department

→ *Illustration of the beautiful Fibonacci spiral whose growth factor is donoted by Φ, the “Golden Ratio”. The name of the Φ-lab has been chosen to capture the state of flux and harmony needed to implement sustained change. © Shutterstock*

EMBRACING EXTERNAL TRENDS

Powerful trends in digital and sensing technologies are rapidly changing our world, thereby also transforming the world of Earth Observation (EO).

In particular, extraordinary developments in Information and Communication Technologies (ICT), including the Internet, Cloud computing and Artificial Intelligence (AI) are giving rise to radically new ways of storing, distributing and analysing big data about our planet. This "digital" revolution is also accompanied by a "sensing" revolution that is delivering unprecedented amounts of data on the state of our planet and its changes.

Europe is leading this sensing revolution in space through the Copernicus initiative, and the corresponding development of the Sentinel missions, which monitor our planet on an operational and sustained basis. In addition, an emerging trend, referred to as New Space in the US or Space 4.0 in Europe, is rapidly appearing through the expanding commoditisation and commercialisation of space. In particular, the increased capabilities and rapidly declining costs of building and launching small satellites is allowing new EO actors – including start-ups, ICT giants and other kind of actors – to enter the space business. Consequently, innovative constellations of standardised small satellites are delivering new data on our planet with high spatial resolution and high temporal frequency.

These different global data sets from space lead to a far more comprehensive picture of our planet, which is complemented by data derived from millions of smart sensors connected to the Internet (referred to as the Internet of Things, or IoT) and from Unmanned Aerial Vehicles systems. Such streams of data offer alternative possibilities not only for science but also for entrepreneurs, who are turning big data into new types of information-based services and businesses.

> *If you fight powerful external trends, you're probably fighting the future. Embrace them and you have a tailwind.*
>
> **Jeff Bezos**
> *How to Shape the Future, letter to stakeholders (2018)*

However, these opportunities bring new challenges for scientists, business, data and software providers seeking to exploit the vast and diverse amount of data fully by capitalising on innovative big data analytics techniques such as AI.

The Φ-Department and its associated Φ-lab have been set up by the Director of ESA EO programmes to explore how EO can benefit from the latest transformative technologies to help shape the future EO capabilities. The Φ-lab is one small part of a much wider ESA ecosystem of activities looking at these applications.

This document focuses only on the activities and use cases of the Φ-lab Explore Office, and in particular on AI4EO projects.

TRANSFORMATIVE TECHNOLOGIES

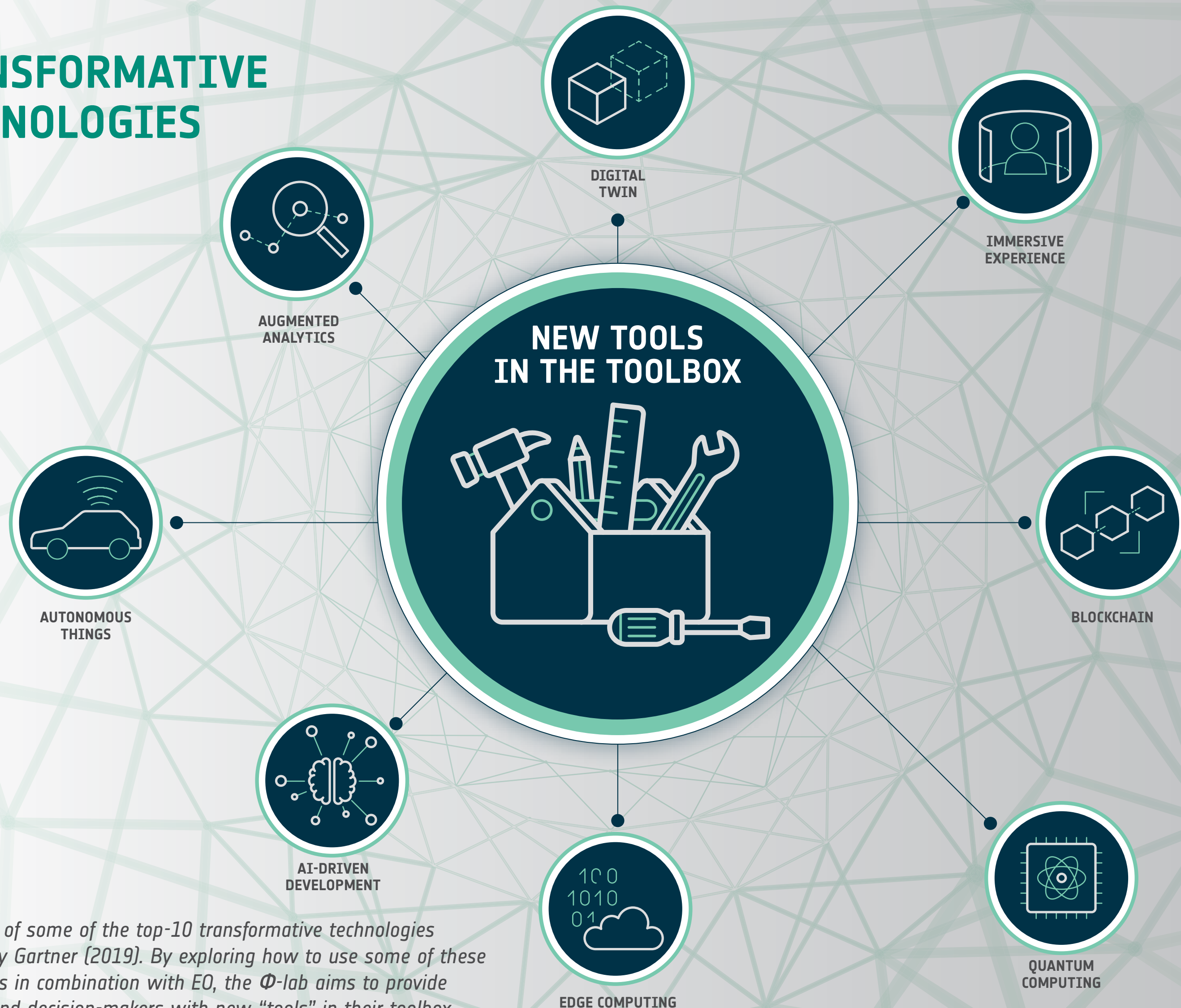


Illustration of some of the top-10 transformative technologies identified by Gartner (2019). By exploring how to use some of these technologies in combination with EO, the Φ-lab aims to provide scientists and decision-makers with new "tools" in their toolbox.

> 1N73LL1G3NC3 15 7H3 4B1L17Y 70 4D4P7 70 CH4NG3
>
> **573PH3N H4WK1NG**

TABLE OF CONTENTS

THE Φ-LAB EXPLORE OFFICE

The Φ-lab is a division of the ESA's Future Systems Department (EOP-Φ). It aims to accelerate the future of Earth Observation (EO) by operating as a "**catalyst**" for EO innovative ideas and transformative technologies and as a "**bridge**" to the ecosystem of new players, connecting internally across ESA and externally to emerging New Space and ICT actors.

The Φ-lab is composed of the Invest Office and the Explore Office. The purpose of the Explore Office is to explore the potential of so-called transformative technologies for innovative EO solutions.

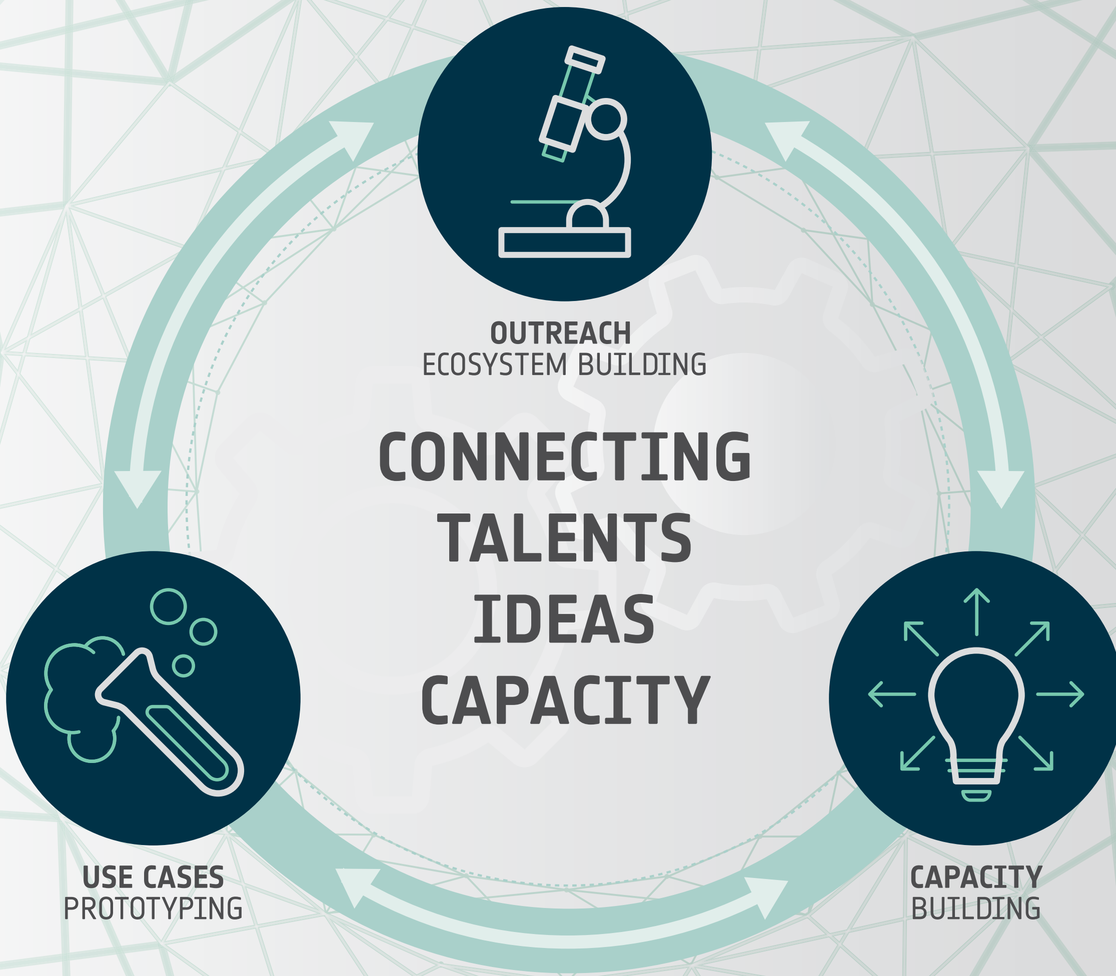
This approach cultivates open innovation through collaboration, both internally across ESA and externally by connecting to new players in the industrial, ICT, New Space, startups ecosystem and research world.

Supported by a small core team of staff in charge of industrial activities and a diverse team of researchers – including ESA research fellows, Young Graduate Trainees, visiting scientists from both academia and industry – these new players will collaborate on a suite of pilot projects focusing on transformative data-driven technologies.

For the moment, the main focus is concentrated on Artificial Intelligence (AI) for computer vision and big data analytics. The Φ-lab also operates as an open space or collaborative platform hosting visitors from academia and industry to investigate new ideas together, providing ESA with a kind of "**Think Tank**" to explore new activities.

*"Our goal is to enable a **connected network** that brings together expertise and ideas from researchers, industry, ICT players, innovators and start-ups to foster scalable learning from EO data, science and AI for a more sustainable planet."*

Pierre-Philippe Mathieu
Head of Φ-lab Explore Office


OUTREACH
ECOSYSTEM BUILDING
CONNECTING
TALENTS
IDEAS
CAPACITY
USE CASES
PROTOTYPING
CAPACITY
BUILDING

MAIN FOCUS ON AI4EO

The AI and Deep Learning Revolution

AI is a general-purpose technology already transforming the global economy and many business models across industry sectors, including Space. In this context we believe AI has a huge, but still largely untapped potential for EO technology.

Described as the "**new electricity**" fuelling the fourth Industrial revolution in the "AI for the Earth" report presented at the World Economic Forum in 2018, AI research is a real driver of innovation, future growth and competitiveness for societies and industries worldwide that can bring enormous opportunities.

Currently, AI is in the midst of a true "**renaissance**", driven by Moore's Law and now super-fed by Big Data. It is deeply entrenched in our society and routinely used in everyday life in applications ranging from recommendation engines, language services (e.g. translation, speech recognition), face recognition, virtual assistants (e.g. Siri, Alexa) and autonomous vehicles (e.g. drones, self-driving cars) – transforming the way we work and live.

What is more, over the last decade, Machine Learning (ML) has undergone a major revolution driven by the unique convergence of large-scale computing capability (e.g. Cloud computing, Graphics Processing Unit (GPU) architectures, High Performance Computing (HPC), easy access to large volumes of data through the Internet, and the availability of new algorithms enabling robust training of large-scale "deep" neural networks.

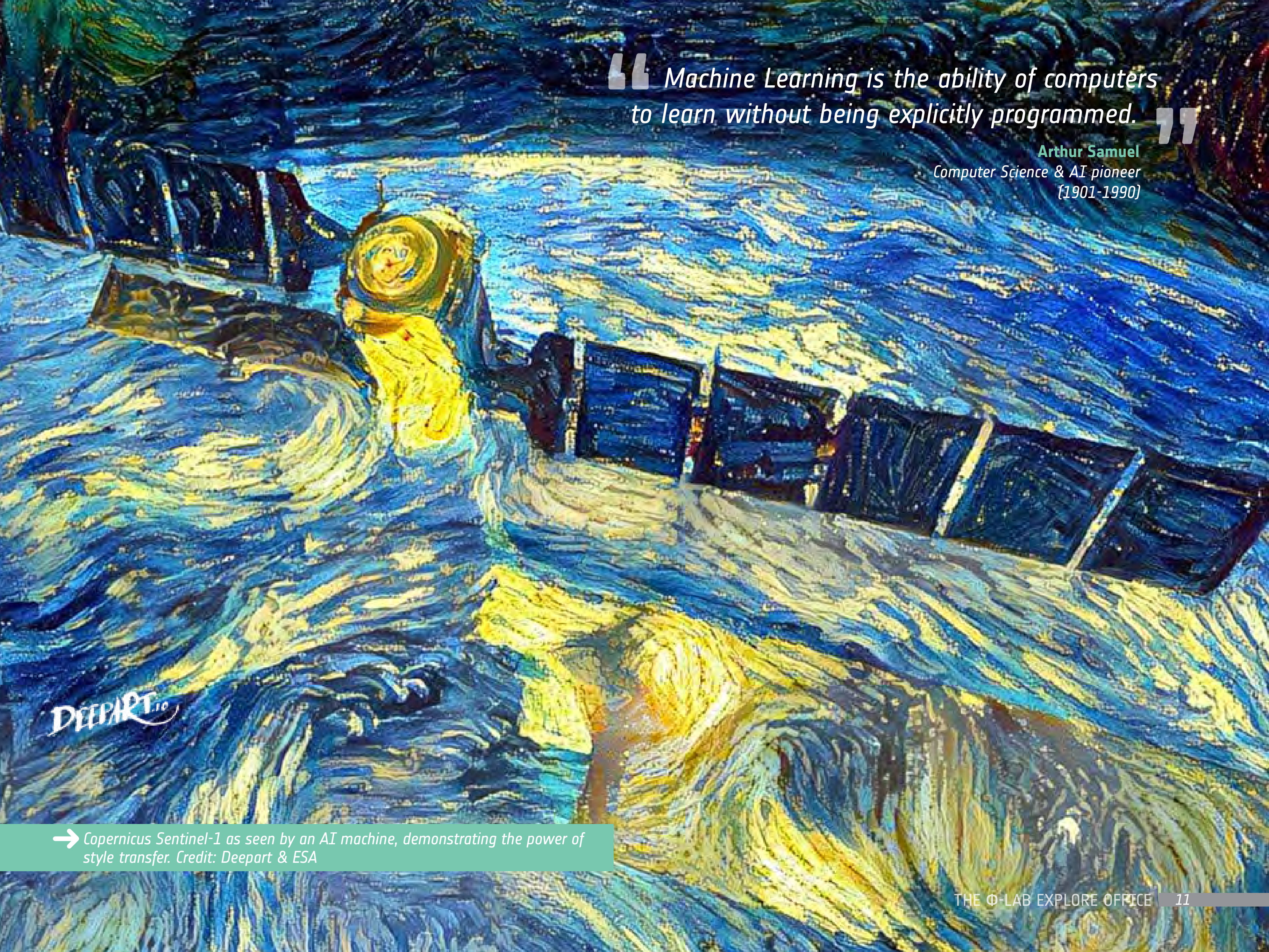
Even more recently, DL and emerging Reinforcement Learning (RL) capabilities are becoming the "workhorses" of AI. DL algorithms are indeed bringing dramatic improvements in the automatic recognition of objects, as first demonstrated in 2012 by results for some tasks within the ImageNet Large Scale Visualisation Challenge [image-net.org].

Specifically concerning EO, so far, AI applications have mainly involved Computer Vision (CV) to interpret and understand very high-resolution satellite imagery, but many other areas such as Earth Science, prediction and Big Data analytics could also benefit.

Therefore strengthening AI capabilities within ESA by leveraging European EO assets (including data and technical expertise) will help the science community and industry to realise the full potential of EO data in delivering socio-economic benefits and, at the same time, allow Europe to position itself in a rapidly changing AI landscape.

“Machine Learning is the ability of computers to learn without being explicitly programmed.”

Arthur Samuel
Computer Science & AI pioneer
(1901-1990)



→ Copernicus Sentinel-1 as seen by an AI machine, demonstrating the power of style transfer. Credit: Deepart & ESA

When AI meets big EO data

AI and EO is a marriage made in heaven, however with its own strengths as well as weaknesses. In a sense, we are now at an inflection point, at a kind of **crossroads of opportunities**, whereby on the one hand AI is becoming one of the most transformative technologies of the century, while on the other hand a growing European EO capability is delivering a totally unique, comprehensive and dynamic picture of the planet, thereby generating big open data sets to be explored by AI.

Due to the rapid increase in the volume and variety of EO data sources, AI techniques become increasingly necessary to analyse the data in an **automatic, flexible** and **scalable way**. Today, EO data remotely sensed from space are particularly suited – but at the same time challenging - for AI processing as they are:

→ **Big** in size and volume with terrabytes of data routinely streamed daily from space, which need to be transformed into "small" actionable information. For example, it would take an operator several hundred years to look at the **trillions of pixels** routinely acquired by Sentinel-2 on a weekly basis;

→ **Diverse** including data from a variety of sensors, from optical (multi-spectral and hyperspectral) to radar data. Up until now, AI has been mainly applied to optical imagery, in particular at very high resolution by use of traditional Computer Vision techniques (using mainly RGB bands). More work is needed to make full use of **all available spatial, temporal and spectral information** of EO data at the global scale. E.g. exploiting the full information of the complex nature of radar data within AI schemes, including information on the amplitude, frequency, phase or polarization of the collected radar echoes;

→ **Complex** and **physically-based** capturing dynamic features of a highly non-linear coupled Earth System. More meaningful data extraction requires integration of physical principles into the statistical approach, and goes well beyond mere automated feature recognition where a wide variety of training datasets are available.

Machines learning algorithms powered by AI are therefore critical to needed to **accelerate "insight"** into the data but should always be used in combination with **domain experts** vital to **properly interpret** the statistical correlations and data. The intersection of AI and EO remains an evolving, but a rapidly growing field. The application of ML to EO has developed rapidly over the last decade, but the emergence of DL has accelerated growth further, as illustrated by the increase in the number of publications. However, although very powerful, DL techniques suffer from their own inherent limitations; they are data hungry while lacking transparency and unable to distinguish causation. This calls for more research which connects AI expertise to address EO problems. An objective of the Φ-lab is to foster the development of a seamless "EO Data Engine" comprising of AI4EO use cases, data sets and pre-trained algorithms (referred as Apps) that can then be used in a modular way. These Apps can also be chained and regularly updated with new information, much the same as an app on your mobile phone. Developing such a data engine will open new opportunities for innovative business models to provide users with updated information. The Φ-lab, in collaboration across ESA, provides some elements of this huge and highly complex process.

AI4EO DATA ENGINE

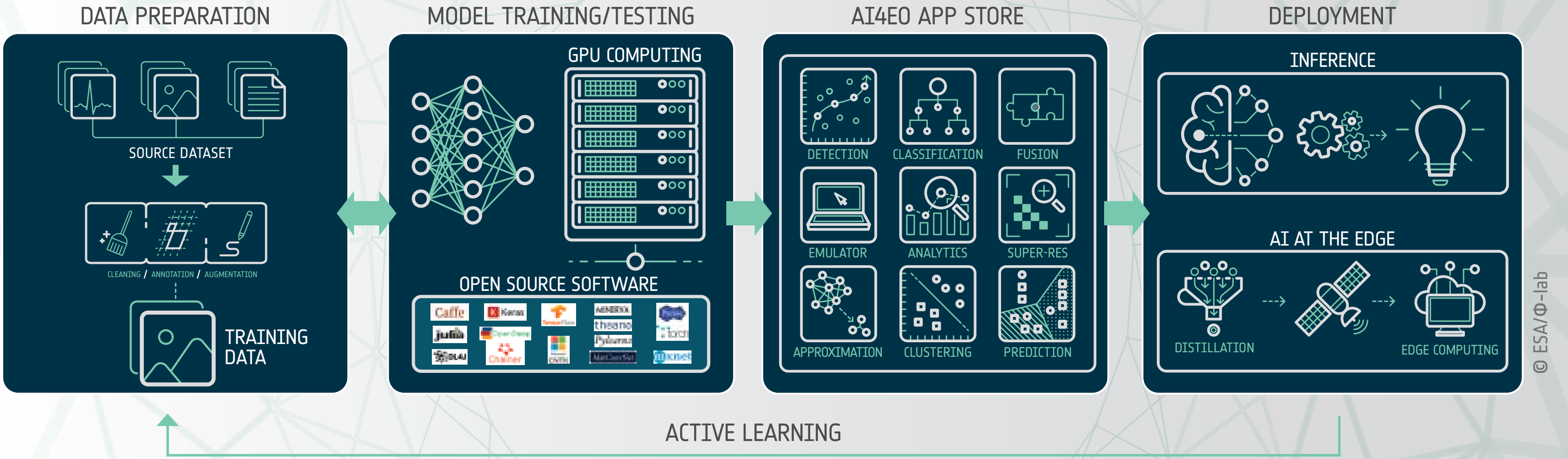


Illustration of an adaptative data engine for EO data based on Machine Learning. It is derived from a new type of value chain, whereby the algorithms are reverse engineered by learning from the data as opposed to the traditional way of explicitly coding knowledge. In AI-powered value chains, most of the work is made up-front (e.g. data preparation, learning) and the result comes from the application of a trained ML scheme (e.g. inference). The process of transforming raw data into information becomes thereby automatic, optimised and scalable, keeping in mind that the quality of the output is related to the quality of the input.

Let the data talk!

Building a Machine Learning software 2.0 stack to query our planet

The revolution of Machine Learning (ML) is also a revolution in developing software in a radically new way. While traditional software 1.0 relies on "encoding" rules, the new paradigm of software 2.0 relies on learning from data to "reverse engineer" the software. And given the rapid growth in data, this will lead to a transformative way in developing sophisticated algorithms to solve complex problems.

Our overall goal is to accelerate insight into big EO data by fostering the development of a searchable database of our planet using state-of-the-art AI techniques to query and detect "relevant" changes, in a way analogous to what you would do with a search engine for linked data.

Today, our current EO assets are data rich but label poor, and we are therefore only scratching the surface of what AI can do for EO. New Space companies, such as Planet, are actively working towards achieving such a vision of a searchable planet by combining high-resolution EO data with AI. A growing amount of data is available to fuel the ML revolution further, presenting exciting opportunities to exploring their potential to push the boundaries of AI in the field of EO.


AI
ICE MELTING
SHIP DETECTION
PERMAFROST ANOMALY
FOREST LOSS
OCEAN BLOOM
OIL SPILL
EL NIÑO IMPACT
INSIGHT
FIRE ALERT

MAIN FOCUS ON AI4EO

Connecting AI4EO innovators

There is a **pressing need** to increase **Europe's strength** and positioning in the area of AI research applied to space. In particular, EO big data presents a niche area of the data economy. Much more could be gained from these valuable data by using AI, thus increasing opportunities for a growing variety of applications and services. The investment, research and innovation efforts required to achieve such an ambitious goal are huge, and **cannot be achieved alone**, but only through teaming and partnership. ESA can play an important role in such endeavors along with other partners at European level.

In this context, and as part of a first element of a wider ESA response, the Φ-lab aims to **operate as a "catalyst" enabling a network of AI and EO labs across Europe** to advance our knowledge related to AI4EO and develop new innovative applications and solutions with EO. In particular, **our vision** is to become a kind of **"hub" connecting a growing ecosystem of AI talents and capabilities across Europe** to tackle a suite of challenges (see below).

To do so, the Φ-lab is implementing a **suite of collaborative schemes** supporting the exchange of ideas and talents and developing new prototype solutions, including:

- **Internal Research Fellowship** opportunities;
- **External Visiting Researchers**, whereby the Φ-lab welcomes short to medium-term visits from scientists at MSc, PhD and Post-doctoral level working on topics of relevance and being seconded from their host organisation;
- **Visiting Professors**. Over the next few years, starting in late 2020, a team of visiting professors will provide advice, support ESA efforts and refine the research agenda to be pursued by ESA in general and the lab in particular;
- **Industrial activities** supporting the development of new AI schemes and pilot demonstration projects.

If you are interested in these opportunities or have ideas to contribute please contact us at **phi-lab@esa.int**

Φ-lab

Tackling AI4EO challenges by connecting talents and capabilities across Europe

The power of AI for EO is today largely untapped. Many challenges still need to be tackled at scientific, applications and capability levels to **deliver the maximum value** from open EO data from satellites for our society and economy. Some challenges are presented below:

Frontier scientific problems – including fundamental and applied research, ranging from classification, detection, indexing, understanding, prediction, data fusion, inversion problems, up to automation, but also new generation of applications and services. A few examples (by no means an exhaustive list) of research questions of interest to our community are provided on the following page.

Bringing AI at the edge – including the development of autonomous systems performing AI inference on board as well as intelligent swarm of satellites powered by AI and guided by machine-to-machine communication.

Digital Twin Earth (DTE) – integrating various data sources (including EO), models (Earth simulations) and AI to develop an interactive "**digital replica**" of the entire planet, which would provide scientists and decision-makers with a unique tool to quantify our impact on the planet and check how to adjust our trajectory towards a more sustainable future.

Building capacity – ensuring that Europe supports the talents, skills, and resources to develop AI4EO.

It is critical to address these challenges at the **European level**, by leveraging ESA assets, and fostering partnership, to realise the full potential of EO and deliver its full socio-economic benefits. EO is capable of delivering a unique and comprehensive dataset about the planet which can be analysed using AI, in ways that are mutually transformative of both EO and AI. In that sense, EO data with its unique and complex features could also help drive some advances in AI techniques, by trying to make the full use of all spectral, temporal and spatial EO information.

Failing to capitalise on this would mean **missing major opportunities for Europe** to shape the future of a rapidly changing, global AI landscape currently dominated by the US and China.

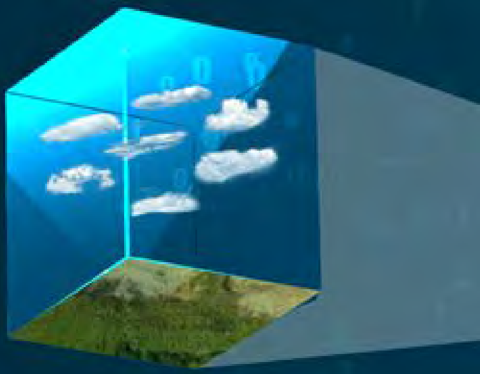
SOME AI4EO CHALLENGES

Towards scalable big data analytics

How to fuse and analyse data from various sources? How to turn big data into small actionable information? How to augment EO capability with AI? How to build "Virtual Sensors" synthetically reconstructing images? How to fuse optical, radar and hyperspectral information?

Towards trustworthy and explainable AI

How to trust AI? What is in the AI black box? How to quantify the error and uncertainty of the algorithm? How to make the AI decision making more transparent? How to resist against adversarial attacks?



Towards physics-aware AI

How to integrate "first principles" and "domain knowledge" into the AI statistical approach? How to bring together models, emulators and AI? How to satisfy physics by "design" ?

Towards self-learning AI

Can machines learn by themselves? How to build new "learning principles" in the AI? How to develop unsupervised learning for EO data without labels? Can we transfer learning to different domains to generalize?

Towards AI-based EO data fusion and prediction for Digital Twin Earth (DTE)

How to leverage the combination of EO, models and AI techniques to support high-resolution prediction and informed decision making with DTE? How to approximate the human-induced forcing and Earth System response function on our planet? How to improve prediction of the Earth System and develop "smart" satellites learning how to optimally target their observations to improve forecast?

> *All models are wrong but some are useful.*
>
> **George Box**
> *British Statistician (1919-2013)*

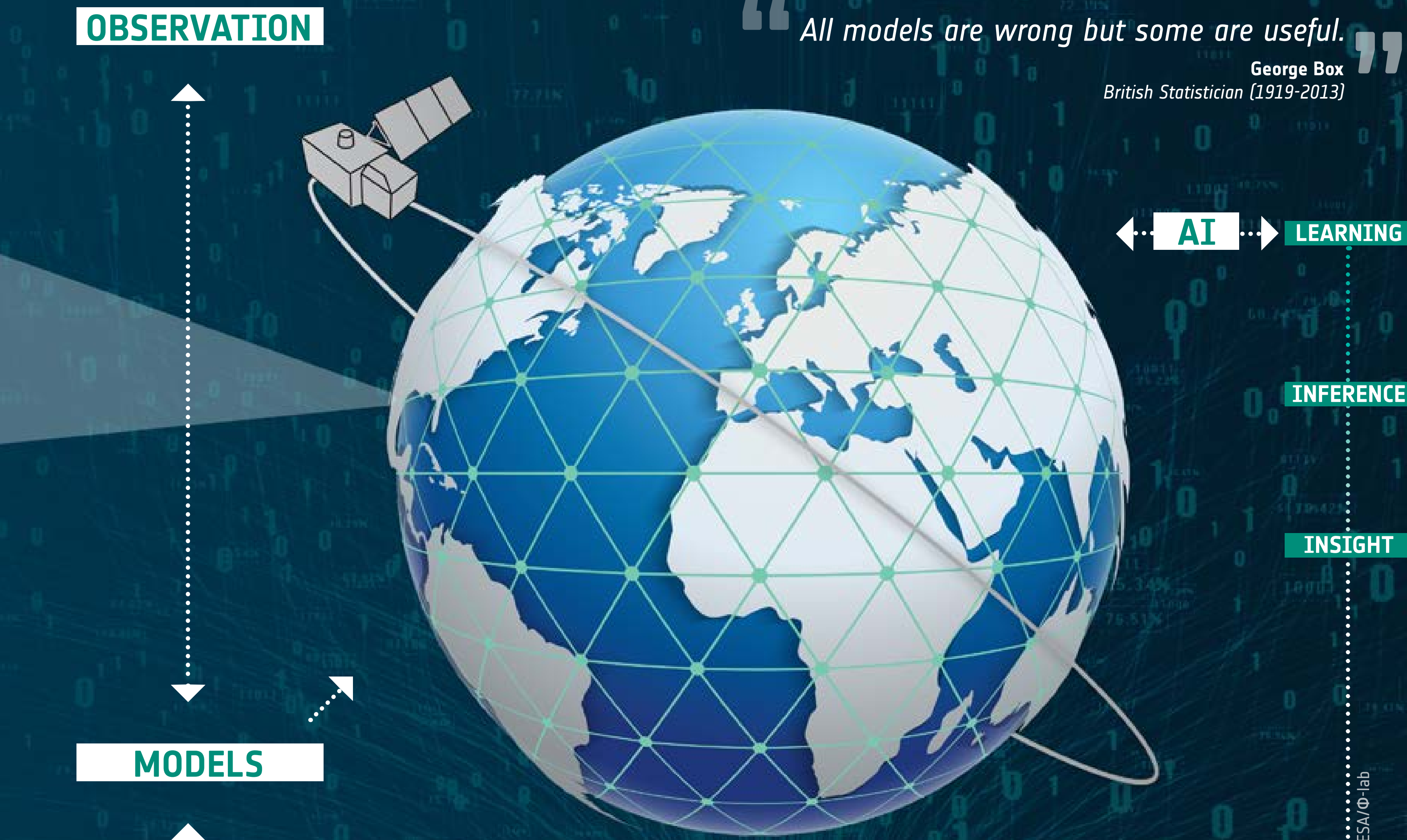


Illustration of dynamic coupling between Models, data and AI within a Digital Twin Earth (DTE) framework. In this schematic, the satellite can "learn" from the DTE how to target sensing to deliver necessary information to improve prediction.

THE TEAM

The Φ-lab Explore office offers a very exciting research and prototyping environment including a **core team of few staff**, a small team of **internal ESA researchers** complemented by a wider team of **external visiting researchers**, working in collaboration with the **Invest** office and across **ESA**.

THE CORE TEAM

The Φ-lab Explore office is managed by a small core team looking after research and industrial activities

Pierre-Philippe Mathieu

My passion is to connect the power of high-tech, data, ideas and people together to address global environmental challenges facing our planet, such as climate change. In particular, my current role as Head of the Φ-lab Explore Office is to explore the potential of new digital and transformative technologies, such as AI, in mining the large amount of data sets delivered routinely by EO satellites orbiting our planet. This role presents the unique opportunity to lead a small team of researchers at ESA in developing a suite of use cases, capacity and open tools with AI and EO – helping to shape the future of EO for the benefit of science and decision-makers.

My background is in Earth Sciences. I have a degree in mechanical engineering and an MSc from University of Liege (Belgium), a PhD in oceanography from the University of Louvain (Belgium), and a Management degree from the University of Reading Business School (UK). I bring an expertise of over 20 years of work in environmental monitoring and modelling, across disciplines from remote sensing, modelling up to weather risk management.

Sveinung Loekken

For many years, my focus has been on the development of new technologies, techniques and capabilities that can significantly boost the uptake and impact of Earth Observation data. My dream is to realise the full potential of remote sensing data by pairing the massive observational capability of Copernicus, national and new space missions with the similarly awesome capability of the IoT in the exploitation of the resulting data for the benefit of humanity.

My role in the Φ-lab Explore Office is to bring together people from research and industry across various disciplines to investigate the AI potential for new useful innovative EO techniques. This includes enabling the definition of research agendas, identifying gaps in capabilities and resources and defining and implementing the activities to fill them, as well as influencing decisions supporting the same objectives in ESA and beyond.

My background is in philosophy and applied mathematics/CS, and I hold a BSc from the University of Oslo and an MSc from Texas A&M University.

Bertrand Le Saux

My work is at the crossroads of AI, statistics and data science. I strongly believe that the convergence of Earth Observation and AI paves the way for building the "nervous system" of our planet, thus enabling new insights and solutions for environmental and population challenges. I am interested in tackling practical EO problems, such as understanding global environmental change, monitoring activity and natural hazards, or leveraging the huge data flows to build consistent models which help decision making.

My background is in Machine Learning applied to computer vision and EO. I hold an MSc in signal and image processing from Grenoble Institute of Technology, a PhD in computer science from Versailles University, on the topic of unsupervised data analysis, and an Habilitation degree from Paris-Saclay University.

Michela Calabritto

I believe change is an important driver of new understanding and opportunities and my approach is to greet it with both curiosity and enthusiasm. As the administration assistant within the Φ-lab, my role is to ensure the smooth running of administrative issues and infrastructure services which underpin our day-to-day work. Helping to improve the potential of the team through a balanced approach of professionalism and humanity is also a key part of my role, and I bring a strong set of interpersonal skills to the team. These skills have been acquired over more than 25 years of administrative work within the space sector and international customer assistance.

I also have a passion for the environment and I am fascinated by how cultural and linguistic diversity can enrich our every-day lives. I hold a BSc in environment and geography, which I gained while working at ESA.

THE ESA RESEARCH TEAM

The Φ-Lab Explore office Office is supporting a small team of innovative ESA researchers performing research on the use of AI4EO and fast prototyping

Christopher Stewart

As an enthusiast on the use of AI for new science and applications, helping to explore and accelerate the integration of disruptive technologies in EO workflows is a huge part of my task in the Φ-lab. Together with the EU Satellite Centre, I am prototyping new Deep Learning and cloud computing based methodologies for infrastructure mapping and monitoring with Copernicus Sentinel-1 in desert regions. In collaboration with the Italian National Geophysics Institute, I carry-out research in potential ionosphere-lithosphere coupling through a machine learning analysis of Swarm data. To facilitate the acquisition of training data to support machine learning analyses, I am managing the development of a crowdsourcing platform. Finally, I stimulate the assimilation of disruptive technologies, such as quantum computing, AI and virtual reality in EO through the organisation of workshops, trainings app camps and other events.

My background is in mathematics, and have specialised in EO since carrying-out an MSc in remote sensing and image interpretation at Edinburgh University. While working at ESA, I obtained a PhD in geoinformation at the University of Tor Vergata in Rome, focusing on SAR applications.

ESA Research Fellow from 01/04/18 to 31/03/21

Jennifer Adams

I am interested in the use of physics-based AI applied to EO data, in order to better understand the environmental and ecological interactions underlying the Earth system, particularly in light of environmental challenges such as climate change and ecosystem vulnerability. As an Internal Research Fellow at the Φ-lab, my main research focuses on applying AI to emulate physical radiative transfer models (RTMs). I am currently employing this method for Methane retrieval from Sentinel 5P's (S5P) TROPospheric Monitoring Instrument (TROPOMI), through using AI to learn the physics of costly RTMs within current retrieval schemes. I also work on building internal Φ-lab capacity for deploying AI at the edge, specifically on the Φ-sat-1 experiment to develop a cloud detection algorithm for hyperspectral data on-board a CubeSat, as well as providing support for hyperspectral data processing and parameter retrieval for forestry and agricultural applications. I have a degree in Geography and an MSc in Remote Sensing (MSc) from University College London (UCL), and a PhD from UCL, co-supervised and funded by the European Commission's Joint Research Centre, on the use of radiative transfer modelling to underpin uncertainty and traceability for EO products.

ESA Research Fellow from 01/04/18 to 31/09/20

Andreas Vollrath

I am a Geoinformatician graduated from the Friedrich Schiller University in Jena, Germany, where I familiarised with various types of EO analytics and relevant GIS tasks for environmental monitoring. I carried out a PhD in Earth Science at the University of Pavia, Italy, combining advanced DInSAR and GPS measurements for the three-dimensional retrieval of tiny surface deformation over tectonic faults. Working as a consultant at the UN-FAO, I integrated SAR processing routines for forest monitoring into the SEPAL platform.

As an Earth Observation Data Scientist with a strong background in radar remote sensing, I am investigating exploratory data preparation of Copernicus Sentinel-1 imagery and its impact on machine learning algorithms for land surface parameter retrieval. In this context, I am developing the Open SAR Toolkit (OST) for the automated production of high-level Analysis-Ready-Data (ARD) products. The aim is to provide a tool which will enable SAR newcomers to explore Sentinel-1 imagery across various application domains more easily. My thematic focus centres on supporting the UN's Sustainable Development Goals through the exploration of radar time-series, ranging from simple land cover classification to more advanced subjects such as forest damage assessment, the mapping of palm oil plantations, tree cover and biomass estimation as well as crop type detection.

ESA Research Fellow from 01/03/18 to 28/02/21

Artur Nowakowski

My main focus within the Φ-lab is on the application of computer vision techniques in Earth Observation science. I have proposed and now lead three case studies: (1) crop types mapping using drones, Copernicus Sentinel-2 and transfer learning from computer vision (in collaboration with World Food Program, UNICEF and Stanford University), (2) ML analysis of Swarm data (in collaboration with Istituto Nazionale di Geofisica e Vulcanologia and ESA Swarm team) and (3) "Seeing through the clouds" competition (in collaboration with CLAIRE and ESA Advance Concept Team).

My background is in computer science with a MSc and PhD obtained at Warsaw University of Technology. I started my professional research analysing computer vision images in the field of pattern recognition and artificial intelligence. Then, I enriched my experience with analysis of biomedical images oriented on stroke tissue recognition and renal cancer detection. In parallel, I have further developed my interests in EO optical data since joining the Space Research Centre of the Polish Academy of Science in 2011. I have developed new approaches for land cover classifications, features extraction and time series modelling for crop yield prediction within the projects funded by ESA, EU, national institutions and industry.

ESA Research Fellow from 1/04/18 to 31/03/21

James Wheeler

My primary research interests are in SAR remote sensing, land cover change mapping, and large dataset analysis. As I work with SAR imagery over large areas and long time periods, I also develop tools for processing and analysing these datasets.

My activities within the Φ-lab are a continuation of previous research into burned area detection, applying Deep Learning methods to support this, and the generation of SAR datasets and applications for other purposes, including VR. I also support training and SAR processing among colleagues and visitors to the Φ-lab.

I began my remote sensing career with a MSc at University College Cork, Ireland where I looked at subsidence monitoring in a peat bog using an Envisat-ASAR time series. I then completed a PhD in the Centre for Landscape and Climate Research at the University of Leicester, studying forest change in the Congo Basin and the discrimination of forest and flooded forest, using L-band SAR.

To date, I have worked on many different research topics, including the detection of small fires using Copernicus Sentinel-1 (for ESA Fire-CCI), soil moisture retrieval from Sentinel-1, and invasive species mapping using Sentinel-2.

ESA Research Fellow from 15/11/2019 to 14/11/2021

John Mrziglod ALUMNI

Deep Learning (DL) has already provided new and powerful solutions for common problems in the field of computer vision and I am confident that it has the potential to help us with traditional problems in Earth Observation too. In the Φ-lab, I work together with my colleagues and external partners on traditional use cases like crop-type mapping or cloud masking. We hope to find better and more efficient solutions by using new Deep Learning methods. At the same time, I investigate how we can capitalise both on the trend of smaller satellites in space and the new powerful AI-on-the-edge hardware to make new satellites smarter and more powerful.

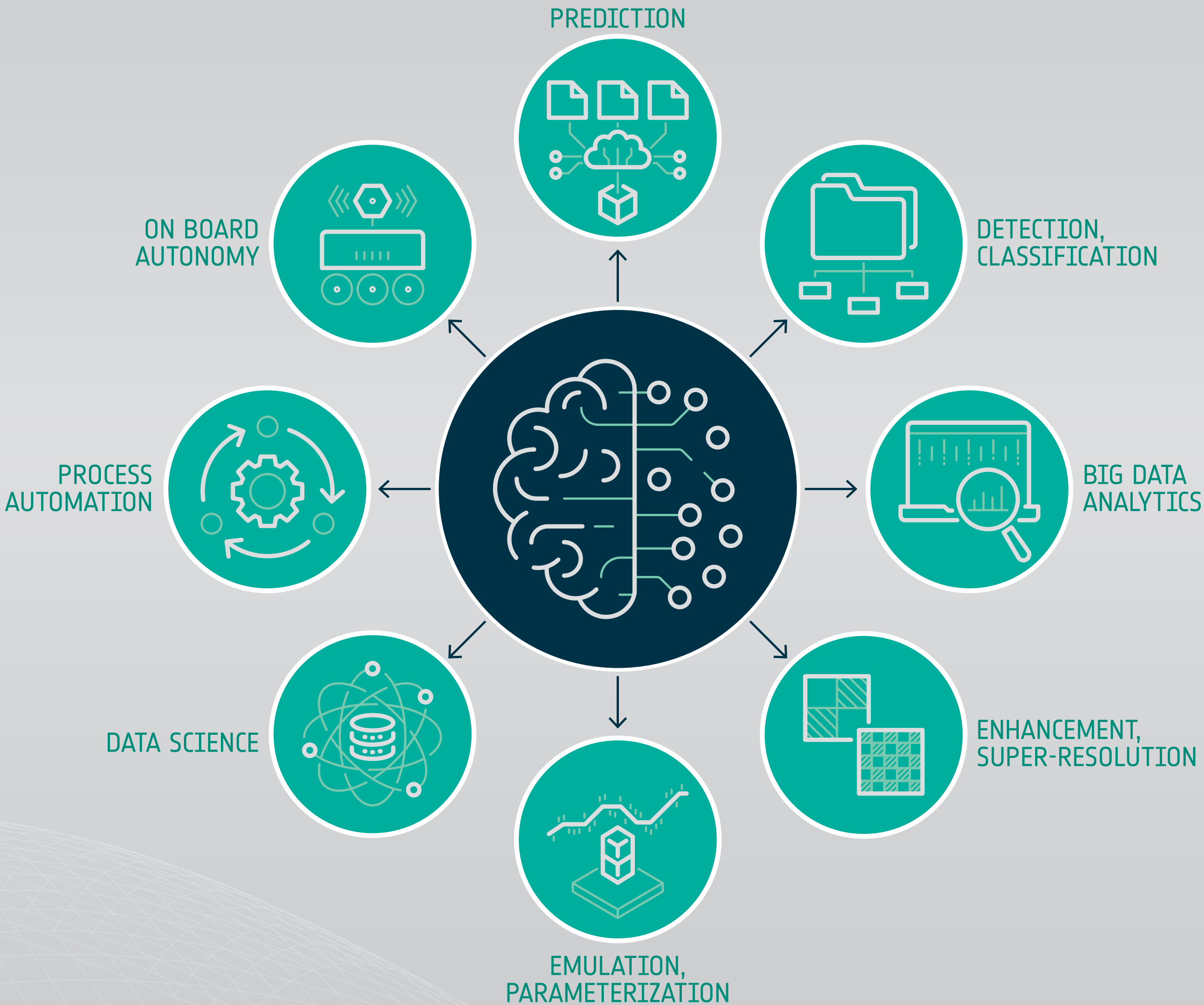
I have a BSc and a MSc in Meteorology from the University of Hamburg (Germany) and, because my passion is computer science, I have been studying programming in my leisure time since high school.

ESA Young Graduate Trainee (YGT) from 1/9/2018 to 28/2/2020

VISITING RESEARCHERS

The Φ-lab provides **young researchers** and **visiting scientists** the opportunity to join the lab team to perform their **own research** and applications in collaboration with the **ESA researchers.**

AI TURBO-CHARGING EO



VISITING RESEARCHERS

The Φ-lab Explore office is providing opportunities for young researchers, data scientists and innovators coming from industry, academia and start-ups to join the Lab for several months and perform their own research on AI4EO in collaboration with ESA data scientists.

Arnaud Dupeyrat

My passion is to connect machine learning, software best practices and available images in order to build useful applications for real-world challenges. In particular, my current role at Airbus is to explore unsupervised Machine Learning algorithms such as Deep Learning (DL), clustering methods and statistics hypothesis to detect changes through time series.

My background is in Computer science. I have a MSc from the University of INSA of Lyon (France), where I specialised myself in computer science and data analytics. Then, I worked as an R&D engineer focusing on automated machine learning and natural language processing. However, I have always been attracted by space which is why I decided to join Airbus and the Φ-lab, where I am working on remote sensing images.

Industry Fellow (Airbus) from 30/07/2019 to 30/06/2020

Mathis Lamarre

My passion is to use mathematics and computer science to solve relevant problems. As an Industry Fellow from Airbus Defence and Space in the Φ-lab, I can do just that. My mission is to apply Artificial Intelligence (AI) techniques to Earth Observation (EO) data, in particular to Synthetic Aperture Radar (SAR) imagery. With Deep Learning (DL), I can use the full spatiotemporal potential of time series of observations. My research activities include land cover segmentation and crop classification. As an industry visiting researcher, I also try to use datasets, tools both from ESA and Airbus: Copernicus Sentinel-1 and TerraSAR-X for example. I have a MSc in Computational Science and Engineering from ETH of Zürich, where I learned the theoretical basis of Machine Learning (ML) and Computer Vision (CV). I also carried concrete projects in DL, for example with a Berlin-based Augmented/Virtual Reality (AR/VR) start-up. I am interested in other aspects of ML as well, such as Natural Processing (NLP), and I wrote my MSc Thesis about word representation and text classification for chatbots at SAP Conversational AI in Paris. I am fascinated that similar models, or also medical ones, can be applied to monitoring agricultural topics from space.

Industry Fellow (Airbus) from 02/09/19 to 02/09/2020

VISITING RESEARCHERS

Andrea Radius

With a Master's thesis in SAR data processing and a PhD in Moving Target detection and Velocity estimation from SAR raw data, I acquired a strong knowledge of remote sensing and SAR and gained valuable experience working in international environments.

I strengthened my technical background during nine years spent in EDISOFT and METASENSING, where I focused on remote sensing and SAR data processing and gained experience in the scientific retrievals and user product validation starting from Earth Observation data, in particular for developing mathematical models and new algorithms based on signal processing and physics theory. While working for EUMETSAT, I acquired global visibility on the end-to-end processing chain for EPS-SG satellites.

Currently, I am employed in the processing team of ICEYE. My main task is to support processing algorithms definition and data calibration and validation, as well as collaborating with the analytic team in different applications, especially in the maritime domain. At the same time, I am the main contact point for all the activities of ICEYE that are developed in cooperation with ESA.

Industry Fellow (Iceye) from 12/07/19 to 12/07/20

Nikolaus Neugebauer *ALUMNI*

With my work I am trying to connect advanced technologies with real-world challenges. Agricultural insurance provides many opportunities to do so. There is a need for detailed field assessments on large geographical scales. This corresponds well with the capabilities of current remote sensing data. I use data processing, data analysis and Machine Learning to solve challenges such as the assessment of large scale drought events on crops like winter wheat. During my stay at the Φ-lab, I evaluated capabilities of different Sensors (Copernicus Sentinel-1 and 2) for detecting damage in Maize fields. The aim is to make these technologies and solutions available for the processes within the insurance sector as well as to the people actually working on the fields.

I studied Land and Infrastructure Management with a focus on geographic information systems at the University of Natural Resources and Life Sciences in Vienna, where I also worked after my graduation in the field of Remote Sensing. Since 2016, I work for the Austrian Hail Insurance where I am responsible for all Remote Sensing related activities.

Industry Fellow (Austrian Hail Insurance) from 05/02/2018 to 05/05/2018

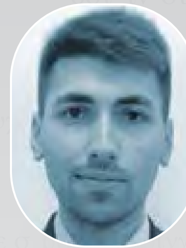
VISITING RESEARCHERS

Dario Spiller

I am a Post-Doc research fellow from ASI (Italian Space Agency) having guidance, navigation and control background and experience in artificial intelligence and metaheuristic algorithms. I have a MSc in astronautic engineering and a PhD in mechanical and aerospace engineering, both of them obtained at the University of Rome La Sapienza. During my PhD, I developed new techniques for optimal planning of trajectories based on artificial intelligence, especially on metaheuristic algorithms. Moreover, I developed novel strategies for star trackers concerning attitude determination and angular velocity estimation. I also spent almost a year in US for research at the Pennsylvania State University and the University of Florida. My current Post-Doc activity is focusing on applying artificial intelligence approaches for data mining of hyperspectral images (with a special focus on the PRISMA mission).

Joint ASI-ESA Research Fellow from 15/11/2019 to 14/11/2021

Alessandro Sebastianelli



My mission is to learn more about satellite data and processing techniques to contribute to the Earth monitoring for improved human security and health. Currently, I am working with the Φ-lab team as visiting scientist. I have been involved in different research activities, such as the use of Copernicus Sentinel-1 data and a series of Sentinel-2 data to remove the clouds from optical data, or the implementation of a speckle filter based on convolutional neural networks. I am enthusiastic about collaborating with Φ-lab team on projects related to my interest and passions, and I know that this will help me in the growth of my scientific career.

My background is in remote sensing through satellite data. I have a degree in Electronic Engineering for Automation and Telecommunications from University of Sannio (Italy) and recently I started my PhD.

Visiting Research Fellow (University of Sannio, IT)
from 01/07/2019 to 30/05/2020

Maria Pia Del Rosso

ALUMNI

I am a passionate about using technology to improve human life quality and believe that Remote Sensing with its capacity for global monitoring, provides an effective and unique tool. As a researcher, I focused my work in monitoring natural disasters, such as landslides and volcanic eruptions, because today's technologies provide new possibilities to reduce risks for human lives. During my experience as a visiting researcher in the Φ-lab, I found a way to use the latest methods of AI to mine information in satellite data for natural hazards assessment and, consequently, risk reduction.

My background is in Remote Sensing. I have a Master Degree in Electronic Engineering for Automation and Telecommunications from University of Sannio and I am candidate for a PhD in Information Technologies for Engineering.

Visiting Research Fellow (University of Sannio, IT)
from 01/07/2019 to 31/01/2020

Erika Puglisi

Working in a place where people from different countries and cultures cooperate to reach a common goal is illuminating. I am eager to use my skills for the development of tools and algorithms that can be part of a set of instruments used by scientists and third parties to study and understand how our planet evolves, how the climate changes and what we can do to prevent calamities, and also how to position the crop fields to optimise the food production.

I am working in the Φ-lab as an external researcher, and will be carrying out my Master's thesis taking part in several projects which concern the computer vision, such as crop and vegetation mapping and data analysis of satellite images. The main purpose is to find and develop algorithms that, with the use of artificial intelligence methods, can extract as much information as possible from the available data.

I have a BSc degree in Engineering in Computer Science from University of Rome La Sapienza and I am currently completing my studies with a MSc in Artificial Intelligence and Robotics Engineering from the same University.

Visiting Research Fellow (University La Sapienza, Rome, IT)
from 4/2/2020 to 30/9/2020

Alistair Francis ALUMNI

I implement Deep Learning techniques to solve segmentation and object detection problems in the observation of Earth, and of other planets. Specifically, my work whilst visiting the Φ-lab has focused on designing novel convolutional models that are able to handle images with arbitrary spectral responses as input. This sensor-independence is a means of achieving greater inter-operability between cloud masks for satellites to gain more value from existing labelled datasets that use images from a wide range of sensors. To this end, I am also interested in dataset harmonisation, and have developed a toolbox to help restructure multiple segmentation datasets into a single shared format.

My background is in physics and Space science. I obtained my BSc in physics from the University of Oxford, and have a MSc in Space Science from UCL. I am now pursuing a PhD in the Imaging Group at Mullard Space Science Laboratory, UCL.

Visiting Research Fellow (University College London, UK)
from 09/06/19 to 09/11/19

Mitra Baratchi ALUMNI

I am a computer scientist with a passion for using computational methods in addressing urban and environmental problems. My research focus is on the design of algorithms for extracting patterns from spatio-temporal data. The knowledge acquired from such data can be used to create a data-driven and system-level view of an urban/environmental phenomenon. This knowledge can lead to better decisions for addressing these problems. One of my goals is to make these algorithms more accessible to domain users by automating low-level tasks in the machine learning pipeline that are often time-consuming and tedious.

I am currently working as assistant professor at Leiden University. I had an amazing time as a Φ-lab visiting scholar for a month in October 2019, focusing on the design of new algorithms for enhancing physical spatio-temporal models using AI and machine learning. I received my PhD degree in computer science from the University of Twente in 2015, followed by a post-doc research period on spatio-temporal data analysis from 2015–2017. I joined Leiden University, respectively, in 2017. Currently, I am also affiliated with the Leiden Centre of Data Science (LCDS), where I collaborate with researchers in other disciplines (e.g., environmental scientist) with interest in spatio-temporal data modelling.

Visiting Research Fellow (University of Leiden, NL)
from 10/09/2019 to 18/10/2019

Beatrice Gottardi ALUMNI

As an entrepreneurial and curious grad student, I believe remote sensing has an increasingly important role in helping to address environmental problems. The availability of a wealth of free and open data from space provides a useful tool to help define resilience and adaptation strategies and monitor their effectiveness. In the Φ-lab, I investigated the application of AI on multi-temporal and multi-sensor (Copernicus Sentinel-1 and Sentinel-2) images over rural areas for crop mapping, with Google Earth Engine. Crop mapping helps land use planning monitoring, reducing food waste and enhancing water management. I had the chance to meet inspirational professionals and experts. From May 2019, I spent six months in the team improving my knowledge in the field and working on my master thesis project.

My background is Environmental Science. I have a degree in structure and dynamics of the atmosphere and a MSc from University of Bologna (Italy) with EIT Climate-KIC certificate. I have been working on developing my start-up and, thanks to this opportunity, I gave a new input to my business idea as well as sharing remote sensing capabilities with other researchers and decision-makers with whom I am collaborating.

Visiting Research Fellow (University of Bologna, IT) from 01/05/19 to 1/11/19

Beatrice Sabatella

My passion is the study of our planet and its evolution due to climate change. In particular, the focus of my research is the use of InSAR methods to better understand the fast changing areas of the cryosphere. I am fascinated by the Machine Learning potential in these fields and interested in evaluating how AI tools can improve and facilitate the comprehension of what we can observe today through satellite imagery. This interest grew from my earlier work at Φ-lab in 2018-2019, where I was involved in testing Machine Learning algorithms and in designing Earth Observation datasets for Artificial Intelligence exploitation. The goal was to develop a series of innovation events, including Hackathons, to foster the use of specific Airborne and Satellite data.

My background is in Aerospace Engineering, and I have focused on EO and geoscience since I started my first internship at the Italian National Institute for Nuclear Physics (INFN) in 2017. While pursuing my academic degree at University of Rome La Sapienza, I had the possibility to engage in two other internships, respectively at ESA ESRIN (Φ-lab) and NASA JPL, where I have been working on different Earth scenarios, from the Poles to the Tropics, focusing on SAR applications and performing environmental monitoring. My new visiting research period at the Φ-lab starts in March 2020.

Visiting Research Fellow (University La Sapienza, Rome, IT) from 15/01/2018 to 15/07/2018

VISITING PROFESSORS

The Φ-lab Explore will benefit from the unique guidance and support of a team of key experts in AI and EO who will visit the Lab from 2020 onwards.

Devis Tuia

Devis is Full Professor at the GeoInformation Science and Remote Sensing group of Wageningen University and Research (The Netherlands). His main research domain is in geospatial computer vision. He develops digital solutions to address problems of land planning and the environment, with a focus in urban recognition, land use modelling and analysis. Devis received his MSc and PhD degrees from the University of Lausanne (Switzerland), with a thesis studying kernel methods for VHR remote sensing of cities.

Sašo Džeroski

Sašo is the Head of the Department of Knowledge Technologies at the Jožef Stefan Institute, Ljubljana, Slovenia, and a full professor at the Jožef Stefan International Postgraduate School. His group develops Machine Learning and data mining methods for knowledge discovery (incl. methods for predicting structured outputs, e.g., multi-target prediction, and automated modelling of dynamical systems, e.g., discovery of differential equations). It also investigates the applications of such methods in a variety of domains, ranging from agriculture and ecology, through medicine and pharmacology, to Earth Observation and space operations. Sašo received his MSc and PhD degrees from the University of Ljubljana. He is a Fellow of EurAI (European Association for Artificial Intelligence), foreign member of the Macedonian Academy of Sciences and Arts, and member of Academia Europea (European Academy of Humanities, Letters and Sciences).

Jan van Rijn

Jan van Rijn obtained his PhD in Computer Science in 2016 at Leiden University. During his PhD, he made several funded research visits to the University of Waikato (New Zealand, three times) and University of Porto. After obtaining his PhD, he worked as a post-doc the Machine Learning lab in Freiburg, headed by Prof. Dr Frank Hutter, after which he moved to do a post-doc at Columbia University, in the City of New York. His research aim is to democratise the access to Machine Learning tools across all entities in society, and his research interests include fundamental Computer Science, Automated Machine Learning and Data Science.

Konrad Schindler

Konrad Schindler is Full Professor in the Department of Civil, Environmental and Geomatic Engineering at the Swiss Federal Institute of Technology (ETH) (Zurich). His research insterests are is photogrammetry, remote sensing, computer vision and image understanding.

Konrad received his MSc degree from Vienna University of Technology (Austria) and his PhD degree from Graz University of Technology (Austria).

VISITING PROFESSORS

Mihai Datcu

Mihai is Senior Scientist and Data Intelligence and Knowledge Discovery research group leader with the Remote Sensing Technology Institute (IMF) of the German Aerospace Center (DLR), Oberpfaffenhofen, and Professor with the Department of Applied Electronics and Information Engineering, Faculty of Electronics, Telecommunications and Information Technology, University Politehnica Bucharest, (UPB), Romania. His main research interests are in Data Science, Machine Learning and Artificial Intelligence, and Computational Imaging for space applications. Mihai received the MS and PhD degrees in electronics and telecommunications from UPB and the habilitation title in computer science from University Louis Pasteur, Strasbourg, France. He was awarded with 2017 Chaire d'excellence internationale Blaise Pascal for EO Data Science. He is a IEEE Fellow.

Xiaoxiang Zhu

Xiaoxiang is the Professor for Signal Processing in Earth Observation at Technical University of Munich (Germany), the head of the department EO Data Science at German Aerospace Center, the co-spokeswoman of the Munich Data Science Research School (MUDS), and the head of the Helmholtz Artificial Intelligence (HAICU) – Research Field "Aeronautics, Space and Transport". Her main research interests are remote sensing and Earth Observation, signal processing, machine learning and data science, with a special application focus on global urban mapping. Xiaoxiang was a visiting scientists in Italy, Japan and US. She received her MSc, PhD degrees and habilitation from TUM.

Gustau Camps Valls

Gustau is a Full Professor of electrical engineering and the Research Coordinator of the Image and Signal Processing (ISP) Group, in the Image Processing Laboratory, Universitat de València (Spain). His main research areas are machine learning for Earth Observation and geoscience data processing. He received his PhD degree in physics from the Universitat de València (Spain).

Holger Hoos

Holger is professor of Machine Learning at the University of Leiden (The Netherlands), adjunct professor of computer science at the University of British Columbia (Canada) and one of the co-founders of the Confederation of Laboratories for Artificial Intelligence Research in Europe (CLAIRE). His research interests are focused on empirical algorithmics with applications in artificial intelligence, bioinformatics and operations research with special interest for automated algorithm design and stochastic local search algorithms. He received his MSc and PhD degrees from the Darmstadt University of Technology (Germany).

Fabio del Frate

Fabio is an Associate Professor teaching courses on Remote Sensing and Applied Electromagnetism in various Master and PhD Programs, at the University of Rome "Tor Vergata" (Italy). His main research topic is focused on the use of Machine Learning approaches for image processing and for the retrieval of geo-physical parameters from EO data.

Fabio Del Frate received his MSc degree in Electronic Engineering and the PhD degree in Computer Science from the University of Rome "Tor Vergata" (Italy).

Begüm Demir

Begüm Demir is a Professor and Head of the Remote Sensing Image Analysis (RSiM) group at the Faculty of Electrical Engineering and Computer Science, Technische Universität Berlin, Germany. She received the PhD degree in 2010 in Electronic and Telecommunication Engineering from Kocaeli University, Turkey. She performs research in the field of processing and analysis of remote sensing images for Earth Observation with interdisciplinary approaches associated to remote sensing, machine learning and big data management. She was a recipient of a Starting Grant from the European Research Council with the project "BigEarth-Accurate and Scalable Processing of Big Data in Earth Observation" in 2017, and the "2018 Early Career Award" presented by the IEEE Geoscience and Remote Sensing Society. Dr Demir is a senior member of IEEE since 2016.

[...] Artificial Intelligence (AI) is about big data, data, data and again data. And we all know that the more data we have, the smarter our algorithms. This is a very simple equation. And therefore, it is so important to have access to data that are out there. This is why we want to give our businesses, but also our researchers, and the public services better access to data.

[...] Overall, today's message is that AI is a huge opportunity in Europe, for Europe. We do have a lot, but we have to unleash this potential that is out there. We want this innovation in Europe. We want to encourage our businesses, our researchers, the innovators, the entrepreneurs, to develop Artificial Intelligence. And we want to encourage our citizens to feel confident to use it. In Europe, basically we do have all what we need, but we have to unleash this potential.

Ursula von der Leyen
President of the European Commission
Press remarks on the EC strategy
on "Shaping Europe's Digital Future" (2020).

As soon as it works, no one calls it AI anymore.

John McCarthy
AI pioneer (1927-2011)

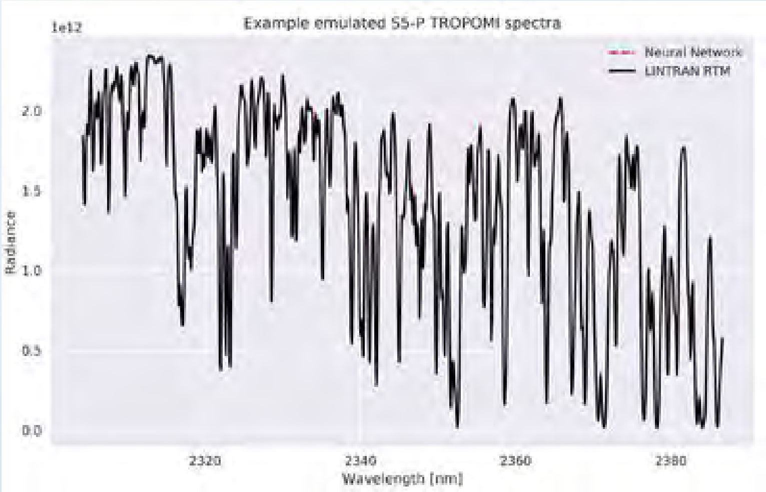
AI4EO

RESEARCH USE CASES

This section highlights some of the **research sprints performed by the team** within the lab in collaboration **across ESA and external partners.**

PHYSICS-BASED MACHINE LEARNING FOR COPERNICUS SENTINEL-5P METHANE RETRIEVAL

Jennifer Adams ESA Research Fellow



Example of a spectrum simulated by the LINTRAN RTM and the NN-based emulator, demonstrating that the NN can effectively emulate S5P TROPOMI SWIR spectra. Credits: Jennifer Adams, ESA.

IMPORTANCE OF THE WORK

A powerful Greenhouse Gas, Methane is one of the most significant drivers of global warming, and has impact on climate change 84 times greater than CO2 (per unit mass, over twenty years). It is also a precursor to the formation of tropospheric ozone, which can have a number of indirect effects on agriculture productivity, crop yields and human health. As a result, it is one of the key targets of Sentinel-5P's TROPospheric Monitoring Instrument (TROPOMI). Current retrieval algorithms are based on "optimal estimation methods", which have been proven to perform well, but at large computational costs. One of the key bottlenecks are the forward Radiative Transfer Model (RTM) simulations, which are a common problem in EO. Emulation of RTMs using ML are gaining interest recently, due to the computational and timing gains that can be achieved. Implementing an emulator of the RTM does not require large processing overheads during operational use, allowing faster and less computationally demanding retrievals.

> *Explainable and physics-aware AI for measuring Greenhouse Gases.*

OVERVIEW

This research project focuses on using AI to learn the physics of RTMs for Greenhouse Gases, in particular for Methane retrievals from S5P-TROPOMI. Applying physics-based ML approaches can significantly reduce processing overheads during operational use, allowing for 1) faster retrievals of Greenhouse Gases such as Methane, 2) an opportunity to use more advanced physics for trace gas retrievals, and 3) provide additional sources of S5P-TROPOMI products for validation of the currently provided products.

FINDINGS

A Neural Network (NN) was trained on a synthetic spectra database of TROPOMI Shortwave InfraRed (SWIR) spectra, based on simulations from the LINTRAN RTM, implemented within the current Methane retrieval algorithm. Different NN architectures were tested to find the best performing emulator. The resulting NN-based emulator can effectively replicate spectra simulated by LINTRAN, explaining 99% of the variance and to within ~3% uncertainty, with a speed up of up 10000 times faster than the RTM. Additional effort is placed on uncertainty characterisation of the emulator, which is a key requirement for S5P-TROPOMI Greenhouse Gas products.

CROP TYPES MAPPING USING DRONES, COPERNICUS SENTINEL-2 AND DAILY LIFE IMAGES

Artur Nowakowski, Dario Spiller ESA Research Fellow

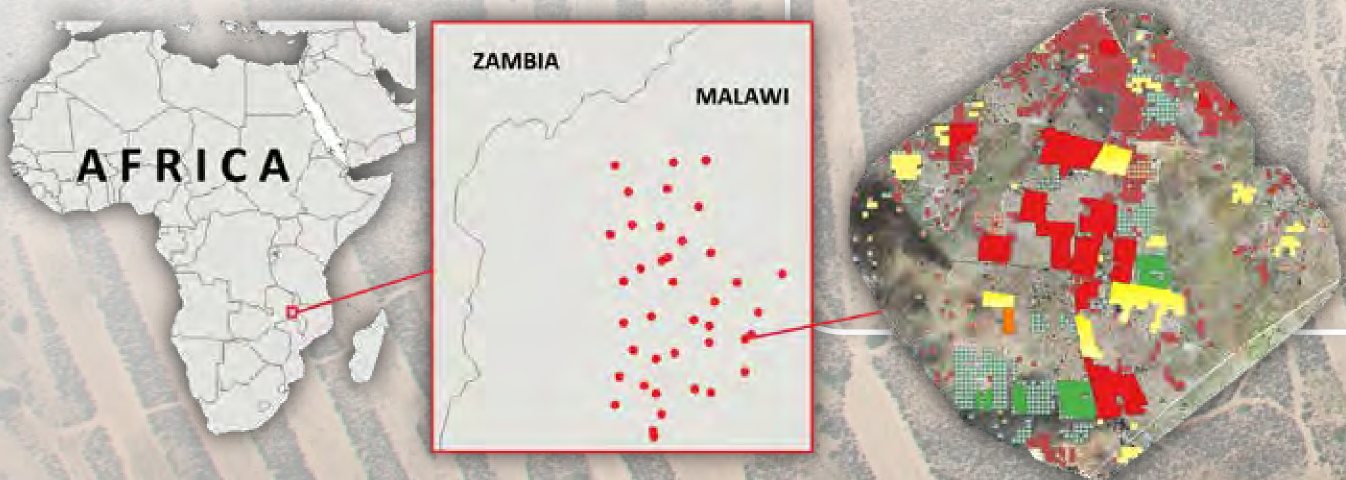
John Mrziglod ESA YGT

Rogerio Bonifacio WFP

Dohyung Kim Unicef

IMPORTANCE OF THE WORK

Crop mapping is important information for many applications, especially those supporting environmental, economic and policy issues. In particular, our partners, World Food Programme and UNICEF, need accurate crop type information to help manage food supplies and children's welfare in developing countries. Availability of new sensor platforms and latest development in transferring models between different image domains create a great opportunity to improve existing crop types mapping approaches.



Localisation of Malawi drone dataset with an example of results of crop type mapping. Credits: ESA

> *Smartphones working together with drones and satellites to solve problems in developing countries.*

OVERVIEW

This study explores a new opportunity to improve satellite-based information on crop types using drones and knowledge transferred from computer vision. The work is conducted in cooperation with World Food Programme, UNICEF and Stanford University. It is an unprecedented attempt not only within crop type mapping domain, but even within broader EO analysis, to fuse information extracted from images captured by satellites, drones and our cameras/smartphones.

FINDINGS

Using three datasets of drone images acquired in Malawi and Mozambique we found that transfer learning from computer vision could be successfully applied to crop types classification. It means that low-level visual information learned on millions of images containing objects like cats, dogs, etc. can help distinguish crop types in drone data for Sub-Saharan areas. Moreover, the classified drone data can be used to increase the number and diversity of input samples needed by Sentinel-2 based classification systems (e.g. Sen2Agri) which make use of multi-temporal information to increase the accuracy of crop types mapping. A key aspect is to provide a confidence level information together with the classification results to select only the most reliable samples.

INFRASTRUCTURE MONITORING IN DESERT REGIONS

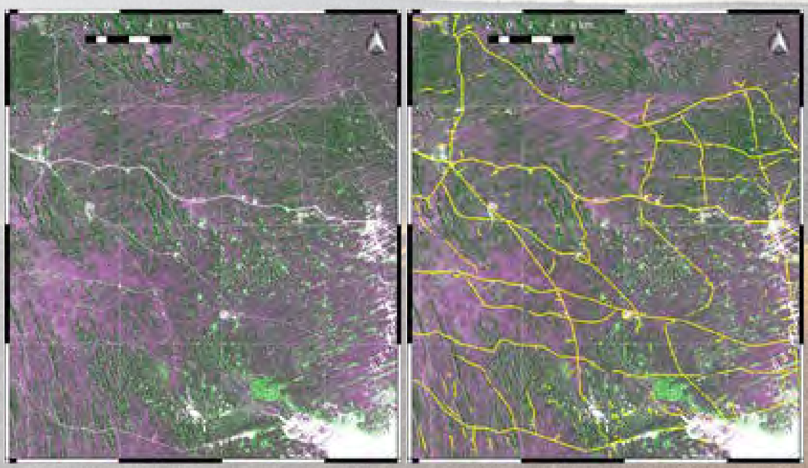
Christopher Stewart ESA Research Fellow

Adrian Luna, Michele Lazzarini EU SatCen

Alessandro Marin CGI

IMPORTANCE OF THE WORK

The systematic surveying of infrastructure in desert regions is a challenging problem, due to the extensive areas covered by many desert landscapes, and the often inhospitable terrain. A significant problem in many such areas is the burial of human infrastructure by mobile sand dunes and drift sand. This activity aims to address the need to detect human installations, and to monitor their state, through the prototyping of a scalable systematic infrastructure mapping and monitoring capability over desert regions. In particular, the methodology exploits new Deep Learning (DL) techniques that are revolutionising segmentation and object detection approaches using large EO datasets. The activity is carried out in collaboration with the European Union Satellite Centre, that supports the decision-making and actions of the European Union by providing Geospatial Intelligence products (e.g. critical infrastructure monitoring).



Part of North Sinai Desert shown in SAR multi-temporal intensity and coherence image. Infrastructure features extracted through Deep Learning overlaid in yellow on right. Credits: C. Stewart, ESA. Contains modified Copernicus Sentinel data (2019).

OVERVIEW

Since the end of 2018, a workflow has been implemented at ESA Φ-lab for the mapping of roads in desert areas. This uses the U-Net Deep Learning architecture for image segmentation, trained with ground truth data from Open Street Map. The input to the model comprises Copernicus Sentinel-1 average intensity and average coherence from consecutively acquired scenes in a two-monthly time series. This data is obtained and automatically processed from the Creodias cloud platform. The workflow has initially been implemented over the North Sinai Desert, but the intention is to scale-up to desert areas worldwide, and to include other types of infrastructure.

> *"AI revealing hidden infrastructures in remote desertic regions."*

FINDINGS

Results so far are positive, and have been validated against previous results obtained over the same area using mathematical morphology rather than Machine Learning (ML). These previous approaches have limitations in distinguishing natural from man-made features. The high contrast between radar backscatter over building materials, such as stone, cement and tarmac, and surrounding sandy areas, renders Sentinel-1 Synthetic Aperture Radar (SAR) highly efficient in the retrieval of anthropogenic structures in desert regions. However, ambiguities exist with natural features that also produce high backscatter, such as rock formations and sand dune ridges. Work is ongoing to improve the model before transferring it to other regions, with the intention of reducing local dependence and improving its robustness.

BRINGING CONVOLUTIONAL NEURAL NETWORKS AT THE EDGE

John Mrziglod ESA Young Graduate Trainee

Jennifer Adams ESA Research Fellow

IMPORTANCE OF THE WORK

Earth Observation satellites are becoming smaller but also more numerous. Improving their on-board data processing capabilities can both enhance their autonomy and reduce the amount of data that must be downloaded. Convolutional Neural Networks (CNN) have the potential to compress the data acquired on-board in a “smart way” enabling us to download only the data that we are interested in, e.g. cloud-free images or areas with wildfires.

Smart downstream transmission: download only the data that are useful, e.g. cloud-free images or areas with wildfires. Credits: John Mrziglod and Jennifer Adams, ESA.

OVERVIEW

In this experiment, ESA and its partners deploy a CNN on a nanosatellite with a hyperspectral instrument (FSSCat and HyperScout-2). Since CNNs are computationally very expensive and therefore not suitable for small satellites which just provide limited resources of power and processing, a specialised Vision Processing Unit (VPU) is shipped on-board. The work is conducted by a consortium which includes Cosine, University of Pisa, Sinergise and Ubotica. Cloud detection was chosen as a demo application to test the functionality of CNNs space.

> *“Satellites are getting smaller and possibly smarter.”*

FINDINGS

The availability of labelled cloud training data suitable for Deep Learning was a key focus of the project, since Hyperscout-2 is yet to acquire imagery following launch. Hyperscout-2 training data were simulated from Copernicus Sentinel-2 cloud labels generated by Synergise, including spatial, spectral (i.e. bandwidths) and radiometric (i.e. Signal-to-Noise-Ratio) transformations.

Following the work of the consortium closely, an internal capacity to port Deep Learning models on embedded devices has been built. CNNs for cloud classification and segmentation (86% accuracy, 15% FP) have been developed in the Φ-lab in order to gain a deeper understanding of the challenges of this technology.

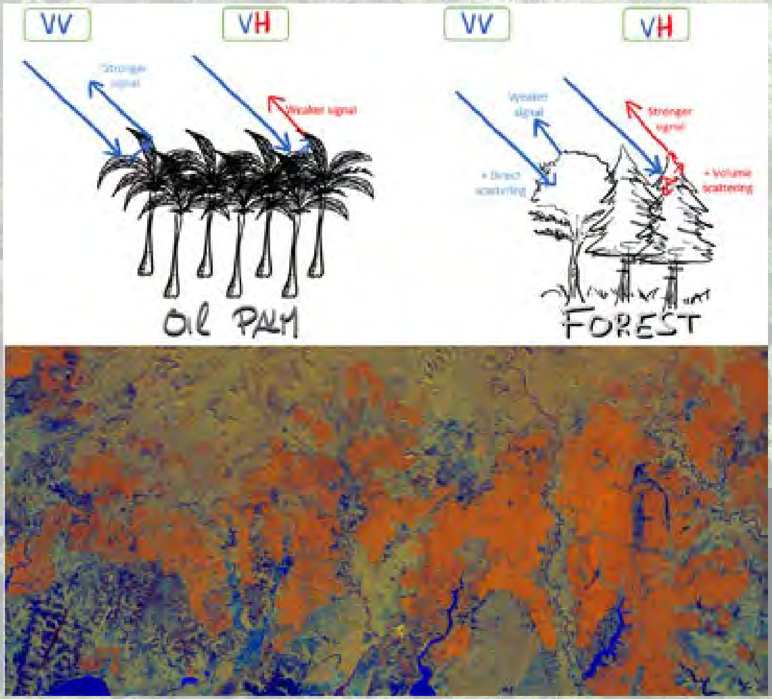
MAPPING PALM OIL WITH COPERNICUS SENTINEL-1

Andreas Vollrath ESA Research Fellow

Sara Aparício Solenix

IMPORTANCE OF THE WORK

Palm oil is the most commonly used vegetable oil in the world. As the demand grows, the expansion of palm oil plantations in the tropical regions will remain one of the main drivers of deforestation, with a considerable negative impact on biodiversity and the carbon cycle. Although the current hot spot of production is located in South-East Asia, future expansion is expected for African and South American tropical regions. The monitoring of the current extent and future expansion of palm oil plantations is crucial for validating the recent efforts towards a sustainable production on a global scale.



(Top): sketch of the differences in scattering mechanisms for C-Band VV and VH polarization between palm oil plantations (left) and natural forests (right). (Bottom): Image from Southwest Kalimantan, Indonesia. The orange image patches are palm oil plantations visible in a Sentinel-1 Timescan (Red: VV-median, Green: VH-median, Blue: Difference of VV- and VH median). Credits: Sara Aparício, Solenix and Andreas Vollrath, ESA

OVERVIEW

Global mapping of palm oil plantations is hindered by regular cloud coverage over tropical regions and the spectral similarity of palm oil plantations to natural forests in the optical domain. In this study, we overcome both limitations by using Copernicus Sentinel-1. The strength of our method capitalises mainly on the dual-polarised VV/VH C-Band SAR data, which is particularly well suited to distinguish palm trees from other forest types due to the structural differences of the canopies. In addition, the 12-day revisit cycle of Sentinel-1 over the tropics allows the creation of dense time-series. The temporal behaviour of the backscatter is then captured by using the timescan method, which depicts the usage of descriptive statistical parameters for every pixel in the full time-series. This helps in reducing the influence of environmental conditions to enable the improved detection of the plantations.

> *Identifying palm trees with radar imagery.*

FINDINGS

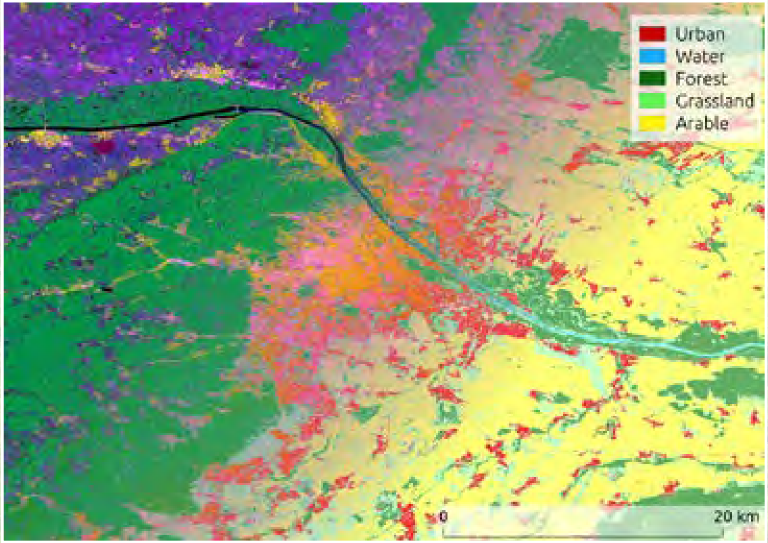
The backscatter signal of Sentinel-1 over trees is a function of the canopy structure. Palm oil trees are fundamentally different to natural forests in terms of canopy structure and are thus easy to spot using Sentinel-1 imagery. By combining advanced Analysis-Ready-Data products from the Open SAR Toolkit (OST) with cutting-edge machine-learning algorithms, this study demonstrated that highly accurate mapping of palm oil plantations over Ghana is feasible. Residual confusion of the classifier for palm oil plantations and urban areas are overcome by the combined use of backscatter and interferometric coherence. Since palm oil plantations are located in tropical regions, the independence from cloud cover qualifies the use of this methodology for global mapping.

ADVANCING DATA-DRIVEN LAND APPLICATIONS WITH COPERNICUS SENTINEL-1

Andreas Vollrath ESA Research Fellow

IMPORTANCE OF THE WORK

Copernicus Sentinel-1 provides reliable, repeated wide area monitoring with cloud penetrating capabilities. This allows for increased availability of dense time series, which are an important asset for measuring the pulse of our planet. The potential of turning high-level SAR products into relevant information is huge, considering that almost constant cloud cover affects large parts of the world. The timeliness of information coming from SAR data can therefore play a crucial role in helping to respond to many global pressures such as food security, emergency response, deforestation and urbanisation.



Sentinel-1 multi-parameter timescan (left part of the image, red: coherence average, green: VH-backscatter minimum, blue: Alpha average) blended into the resultant 5-class Land cover classification (right part of the image) with an overall accuracy of 85%. Credits: Andreas Vollrath, ESA

OVERVIEW

Copernicus Sentinel-1 is a unique SAR system in many ways. The globally systematic observation within an orbital tube of 100 meters allows for both the creation of dense time series of radar backscatter, as well as for the systematic generation of interferometric coherence. The combined use of both information layers has been shown to provide more accurate observations for a wide range of land surface parameters. Furthermore, dual-polarimetric H-A-Alpha decomposition parameters can be derived for each scene that characterise the scattering behaviour on the ground, based on the combined amplitude and phase measurements of the backscatter signal.

This activity explores the added information content from the combination of these high-level SAR products, including more advanced time-series and timescan products produced by the Open SAR Toolkit (see tools section), when used as distinct features in ML based EO applications. The respective data analysis includes different domains ranging from land cover to crop type mapping as well as tree height estimation and building footprint detection.

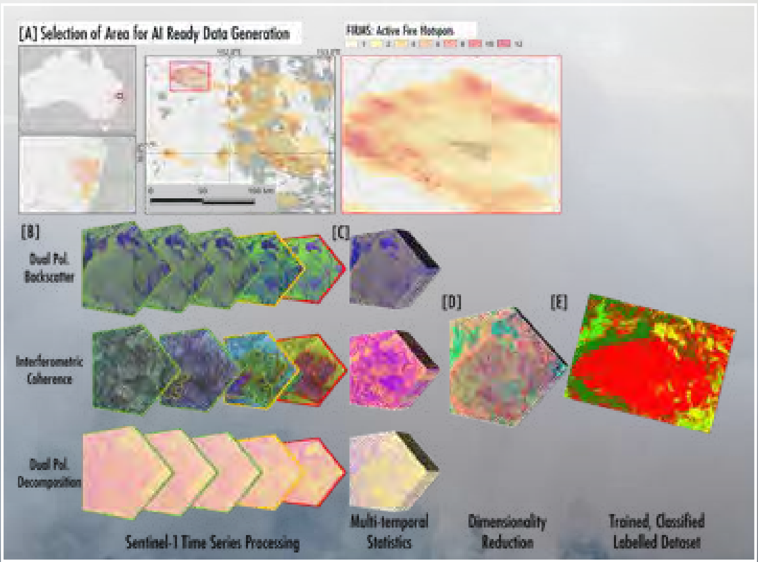
> *"Making the most out of Sentinel-1 for land monitoring and applications."*

FINDINGS

Preliminary results show that, in most cases, the combined use of backscatter, interferometric coherence and polarimetric decomposition parameters boosts the performance of ML algorithms used for qualitative as well as quantitative estimates of land surface parameters in general. While single imagery from Sentinel-1 are affected by distinct environmental conditions, multitemporal statistics calculated from dense time series single out those effects and allow for results comparable to optical data throughout various applications. In particular, the combination of multi-temporal information from the consecutive interferometric coherence and the radiometrically normalised radar backscatter unleash the full potential of SAR data, whereas the polarimetric decomposition parameters do not add significant added value in most cases.

BURNED AREA REPORTING FROM COPERNICUS SENTINEL-1 ANALYSIS-READY DATA

James Wheeler ESA Research Fellow



Generating AI ready labelled data: In [A] we acknowledge the use of data and imagery from LANCE FIRMS operated by NASA's Earth Science Data and Information System (ESDIS) with funding provided by NASA Headquarters. The base layer in [A] is copyright of OpenStreetMap® contributors (openstreetmap.org/copyright) with cartography licensed as CC BY-SA. Images from [B] to [E] contain modified Copernicus Sentinel data (2019), and are processed using the Open SAR Toolkit on the CreoDIAS platform; labelled data in [E] is generated using the Φ-lab IRIS tool.

IMPORTANCE OF THE WORK

Forest fires are increasing in frequency and severity on a global scale, causing serious human, environmental and economic damage. Information about burn scars and vegetation loss can be obscured from optical and infrared sensors by smoke, haze and clouds in the aftermath of fires. SAR sensors have the ability to image the surface through clouds, smoke and haze, and can provide information about structural changes in vegetation as well as changes in surface water content. This is particularly useful in areas of persistent cloud cover where a clear view from medium resolution optical images can take months to get after an incident and vegetation regrowth may make detection difficult. As well as the necessity to monitor fire outbreaks and predict spread in real time to coordinate fire-fighting efforts, measuring the contribution of emissions from fire events to atmospheric carbon dioxide is vital for climate modelling. A burned area labelled dataset from Copernicus Sentinel-1 constitutes a vital step towards the use of Deep Learning methods for rapid reporting of medium to high resolution burned areas in all weather conditions.

OVERVIEW

The aim of this research is to develop a novel method for generation of burned area measurements, using rapidly generated AI ready data, maximising the information available from a time series of Sentinel-1 SAR SLC products. This will generate labelled data from archive S-1 SLC data. The selection of training areas will be automated and based on global stratified sampling by fire incidence (using high confidence active fire hotspots from Sentinel-3 and MODIS), land cover type (using ESA Land Cover CCI or similar datasets), season, and Sentinel-1 availability. Pre-processing of S-1 SLC images will be fully automated and cloud based, on the Creodias platform.

> *"Using Sentinel-1 to enhance burned area monitoring."*

FINDINGS

Instruments on Copernicus satellites are used to produce images of smoke, flames and burn scars (Sentinel-2), monitor fire hotspots (Sentinel-3) and provide analysis of air pollution, such as carbon monoxide (Sentinel-5P). The C-band SAR sensor from the Sentinel-1 satellites may provide unique insight into the severity of burns as well as complementing the other Copernicus mission datasets. Initial priority is being given to generation of Analysis Ready Data (ARD) for gathering labelled training data from Sentinel-1 archives. The ARD generation aims to address the limitations of both datasets (e.g. flattening and masking geometric distortions in SAR, gap filling of clouds and shadow in optical), and will build on and contribute to existing efforts externally and in the Φ-lab.

SEEING THROUGH CLOUDS CHALLENGE

Artur Nowakowski ESA Research Fellow

Alessandro Sebastianelli University of Sannio, IT

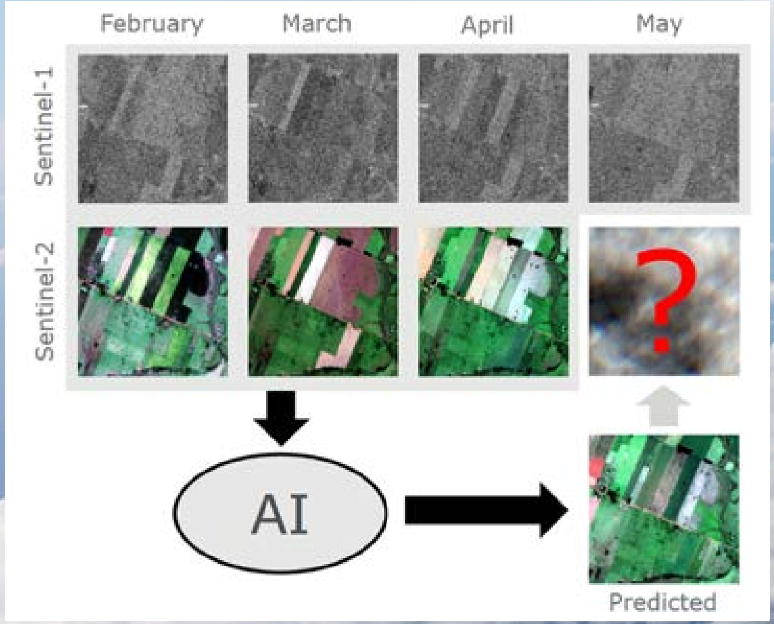
Erika Puglisi University La Sapienza, IT

Dario Izzo, Marcus Maertens, Dawa Derksten, Leopold Summerer ESA ACT

Mitra Baratchi, Jan van Rijn University of Leiden, NL

IMPORTANCE OF THE WORK

Gaps in time series caused by cloud cover are one of the main issues affecting optical imagery. Clouds are a serious obstacle in continuous land and water monitoring, which involves capturing the current status of the plant phenology. A reconstruction of areas below clouds offers a way to improve the accuracy of many applications based on optical images, especially those using Copernicus Sentinel-2 data. Recent developments in AI together with unprecedented amount of Sentinel-1 and Sentinel-2 data offer new opportunities for a solution to reconstruct land surface information normally hidden by the clouds. Some attempts already performed on pairs of radar and optical images are promising, however an exploitation of time series information can bring even more to the solution for the problem.



An example showing the idea of removing clouds from Sentinel-2 images. Credits: ESA, Contains modified Copernicus Sentinel-1 and -2 data.

OVERVIEW

This activity aims to design, implement and run a scientific competition on the reconstruction of land areas hidden by clouds on Sentinel-2 data using time series of Sentinel-1 and Sentinel-2 data. It is done in cooperation with ESA's Advanced Concepts Team, which has created ESA's Kelvins platform to run scientific competitions. This activity is also supported by CLAIRE (Confederation of Laboratories for Artificial Intelligence Research in Europe) having been identified as one of the most interesting topics for the pure AI community during the first ESA-CLAIRE workshop. In order to attract pure AI specialists participants, a minimum EO knowledge will be required. If successful, the competition could lead to a model to map the relationship between radar and optical data, which has not yet been established due to the different physical nature of these two types of measurements.

> *"Learning representation from radar data to mimick optical data."*

FINDINGS

The competition is now in a preparatory phase. An initial sample of a globally representative dataset containing Sentinel-2 and Sentinel-1 time series has been designed and prepared. Work is currently underway to develop a baseline solution to maximise the potential of the competition.

MACHINE LEARNING ANALYSIS OF SWARM DATA

Artur Nowakowski ESA Research Fellow

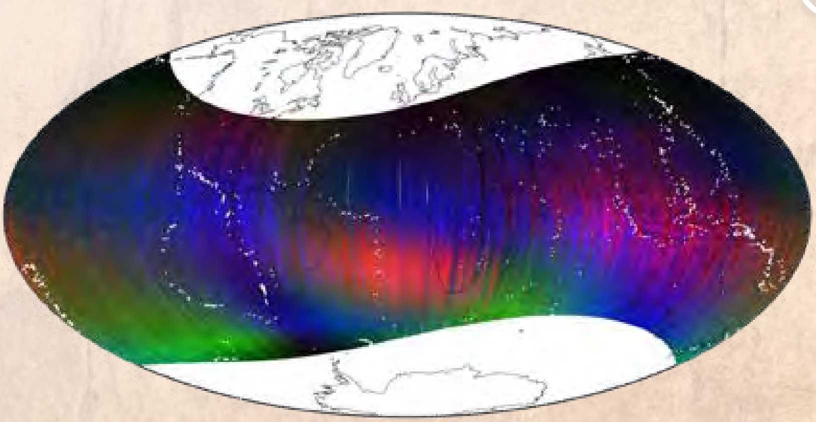
Christopher Stewart ESA Research Fellow

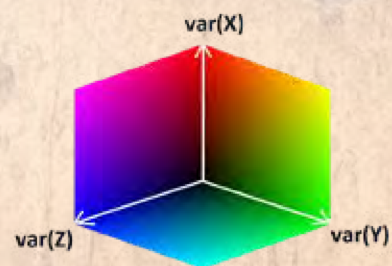
Lorenzo Trenchi Serco

Angelo De Santis, Gianfranco Cianchini, Alessandro Piscini INGV

IMPORTANCE OF THE WORK

Earthquakes are among the most severe of natural disasters. Global statistics from the United States Geological Survey (USGS) show that since the beginning of the 21st Century, the death toll resulting from earthquakes exceeds 700000, and according to estimates, by 2050 about 870 million people will be exposed to earthquakes. One activity of the ESA Φ-lab involves investigating potential coupling between the ionosphere (as measured by the Swarm constellation which monitor the Earth magnetic field) and earthquakes. The approach is entirely data driven, using Machine Learning and Deep Learning with global time series of Swarm and earthquake data to better understand the phenomenon. The activity is carried out in collaboration with domain experts from the Italian National Geophysics and Volcanology Institute (INGV) and within ESA.





Geomagnetic field statistics from five years of SWARM mission (in colour) and earthquake locations in corresponding time period (white dots). Credits: ESA.

OVERVIEW

Various data driven approaches have been carried out to determine a possible link between earthquake events in the lithosphere and magnetic signals in the ionosphere. Both potential co-seismic and earthquake precursor signals have been analysed making as little assumption as possible on the form that these signals may take. Both supervised and unsupervised Machine Learning and Deep Learning techniques have been applied, using the entire Swarm mission archive from all three satellites. While the approach is data-driven, domain expertise from INGV and ESA has helped guide the methodologies and interpretation of results.

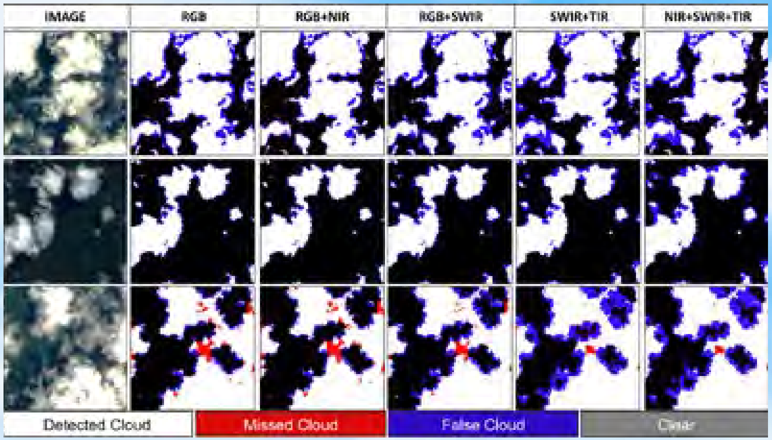
> *"New insight into Solid Earth science with AI."*

FINDINGS

The research carried out to date has led to the conclusion that any magnetic co-seismic anomalies in the ionosphere are likely to affect only a small percentage of earthquakes. Possible ionospheric precursor signals that may correlate with earthquakes are under investigation and are being analysed as a function of earthquake magnitude and Swarm data type. Statistical techniques are being used to assess the statistical significance of any potential correlation. Efforts are made to reduce any geographic or temporal bias.

SENSOR-INDEPENDENCE FOR CLOUD MASKING

Alistair Francis Visiting Research Fellow



Sensor-independent model prediction of images with different spectral band combinations used as input. This demonstrates that a single model can consistently predict cloud cover using different band combinations, allowing it to work on multiple sensors. In this case, we show results on a selection of Landsat 8 scenes. Credits: Alistair Francis, UCL.

IMPORTANCE OF THE WORK

In Earth Observation, the performance of ML algorithms are no longer limited by model design, but rather by dataset size for many problems. Labelled data is costly to produce, especially for segmentation problems, and the diversity of sensor design seen in satellites means that datasets are often not directly combinable. Therefore, models are left with small datasets and often struggle to profit from the wealth of labelled data already collected for other missions. Worse, this problem will only intensify as the number of satellites flown increases.

OVERVIEW

This research has focused on designing a CNN that is sensor-independent (usable across multiple sensors) for optical segmentation problems, specifically cloud masking. This novel sensor-independent model fuses the spatial information of the image, along with knowledge about the wavelengths used. It can ingest any number of spectral bands, and use knowledge learned on one sensor for prediction on another, allowing for simultaneous use across multiple sensors without the need for retraining.

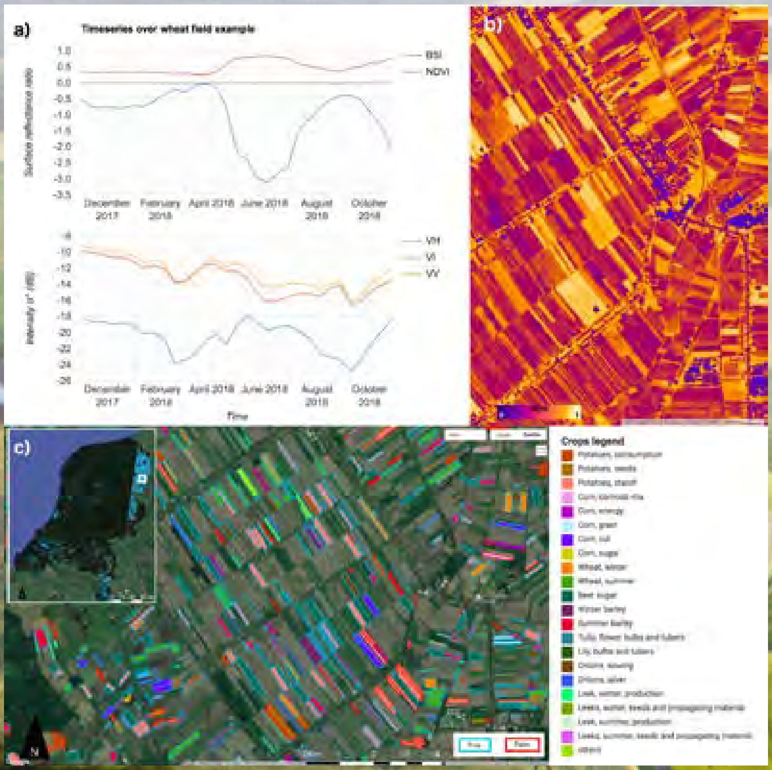
> *A single model for multiple satellites.*

FINDINGS

Initial results for cloud masking are promising. Although optimisation of the model design is still ongoing, we have found strong evidence to suggest that including training data from one satellite increases performance on another, even if the spectral bands are not the same. This suggests that combining many datasets from a diverse range of satellites will lead to even better results. Soon, this could generate a model that exhibits state-of-the-art performance across multiple sensors simultaneously.

CROP MAPPING WITH MULTI TEMPORAL AND MULTI-SENSOR IMAGES

Beatrice Gottardi Visiting Research Fellow



IMPORTANCE OF THE WORK

Agriculture is a fast-growing sector when it comes to the utilisation of EO-based products. It plays a crucial role in the global economy and is rapidly evolving particularly due to climate change and increasing production demand. As population and consumption increase, land, water and energy become progressively threatened resources. Hence, smart and more efficient land cover monitoring provides the means to address several other significant challenges. This study aims to exploit open and freely accessible cloud-computing tools and datasets for crop mapping, which have the potential of a knock-on disrupting added value in EO on worldwide users, from policy makers to commercial and private users.

(a) Example of moving mean profiles (upper panel) of Sentinel-2 Normalised Difference Vegetation Index (NDVI) and bare-soil index (BSI) and (lower panel) VV, VH and VI (vegetation index) σ^0 backscatter intensities. (b) Optical image of NDVI distribution on the Dutch territory (c) Crop types features extracted correctly through machine learning overlain in blue. Credits: Beatrice Gottardi, ESA. Contains modified Copernicus Sentinel data (2019).

OVERVIEW

The project constituted in developing a machine-learning Random Forest (RF) classifier that identifies major regional crop types from radar and multispectral images. It combines 12-day moving median composites of Copernicus Sentinel-1 and Sentinel-2 satellites through one agricultural year (2017–2018). Some commonly used vegetation indexes such as Normalized Difference Vegetation Index (NDVI) have been used in addition to images band intensities. The workflow has been applied to the territory of the Netherlands. The ground truth dataset comes from the national Basic Register of crop Plots. The study relies on a parcel-based classification, entirely developed on Google Earth Engine.

> *"Boost environmental monitoring with free and open accessible tools."*

FINDINGS

The combination of multi-sensors and multi-temporal images enhances the classification over areas with relevant cloud coverage through the year and gives more information about growth phenology. Accuracy assessment gives interesting results especially for some crop classes' detection (corn, wheat, sugar beet, onions and tulip), which range between 85% and 96% of accuracy. Among these, higher accuracy has been observed for winter crops of the same species. In fact, it depends on the observation period ingested in the model. A comparison with only Sentinel-1 imagery shows a small reduction of 1% of the overall accuracy of the classification. The model has been applied for the further year, however only few crop classes got good accuracy, so other approaches have to be investigated.

AN UNSUPERVISED SOLUTION TO DETECT URBAN CHANGES USING OPTICAL IMAGES

Arnaud Dupeyrat Industry Fellow, Airbus

IMPORTANCE OF THE WORK

Satellite sensors provide consistent and repeatable measurements that are capable of capturing the effects of many processes that cause change such as urban evolution, wildfire, and landslide. Despite the advanced technology providing satellite imagery, the detection and evaluation of this information is still to a large extent a human task. The exponential increase in image data has now overwhelmed the capacity of human inspection. The need for automated inspection has now become pressing.

Urban change detection map generated from Copernicus Sentinel-2 time series. Credits: Arnaud Dupeyrat, Airbus.

OVERVIEW

Despite having a tool facilitating the Machine Learning process, Airbus spends time and resources on the labelling task. The objective of this research at Airbus is to investigate unsupervised classification of time series from Copernicus Sentinel-2 data and deliver an algorithm to enable the detection of anomalies (or change) on Sentinel imagery without the need for massive tagging of imagery. In the first part of the activity, the emphasis will be on urban changes using an existing dataset, however the end goal is to be able to identify and classify several changes classes (urban construction, natural disaster, seasonal change, etc.)

> *"AI self-learns how cities are changing."*

FINDINGS

Unsupervised Learning is a Machine Learning technique that does not require supervision of the model, which is capable of working on its own to discover information without any indications.

Applying unsupervised techniques to satellite images is challenging. There are factors (the image quality, the intensity, the clouds, the cloud shadows...) which can vary considerably due to non-controllable parameters. The goal is first to be able to select clear/similar images through time without any training. This has been explored using Deep Learning and statistic methods.

When comparing similar images, the aim is to find where and when an (urban) change happened pixel-wise. In the first version of the project, the emphasis is on dimensional reduction; using an LSTM auto-encoder to capture important features in the spatial and the temporal space thus, allowing for the detection of the change location.

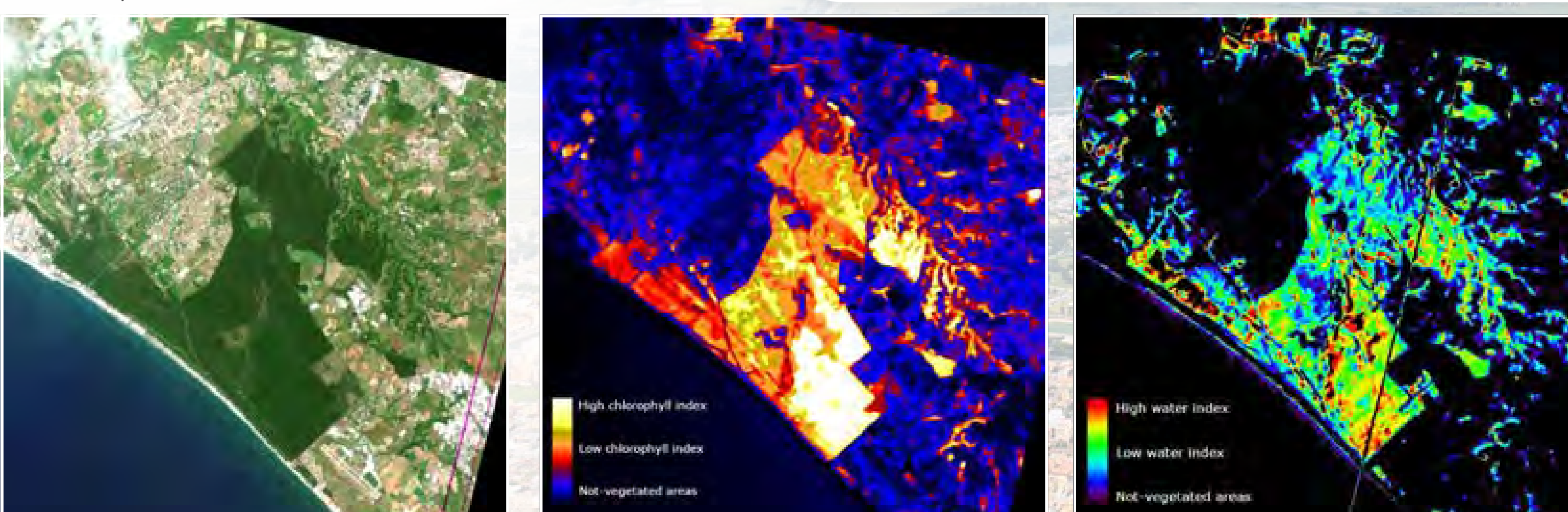
SUPPORT EO WITH HYPERSPECTRAL IMAGES

Dario Spiller Joint ASI-ESA Research Fellow

IMPORTANCE OF THE WORK

One of the current challenges in Earth Observation is the successful application of ML approaches for the extraction of relevant information from a large variety of sensors. The new Italian mission named PRISMA (PRecursore IperSpettrale della Missione Applicativa) will provide a new free dataset for research purposes. The application of latest development in AI provides the potential to more efficiently exploit the hyperspectral data and their unmatched spectral resolution.

Examples of post-processed images of Castel Fusano (Rome).
(Left): area in true colours, (Center): chlorophyll analysis, (Right): Water distribution.
Credits: Telespazio



OVERVIEW

This research activity is the result of a collaboration between ASI and ESA. The collaboration started in November 2019 and will focus mainly on the exploitation of the PRISMA data, although it is a potentially useful application in other new missions, such as the Φ sat-1.

The aim of the research activity is to exploit the hyperspectral data by means of artificial intelligence techniques. This can be carried out in two different ways: 1) by using optimisation methods to select the most relevant spectral bands for specific case studies (metaheuristic algorithms have already proven their ability in this tasks, according to the literature), and 2) by using Deep Learning to extract features from the whole collection of bands, which is more demanding from the computational point of view but does not discard any source of information.

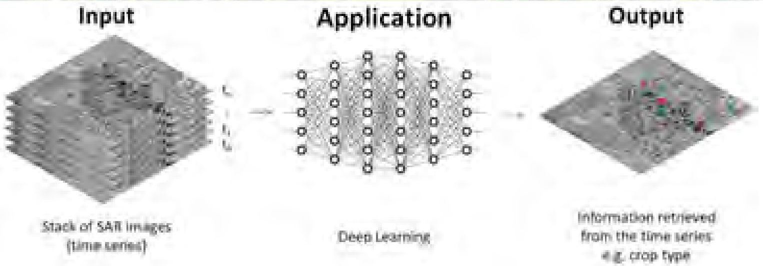
> *"Extracting information in hyperspectral data with AI."*

FINDINGS

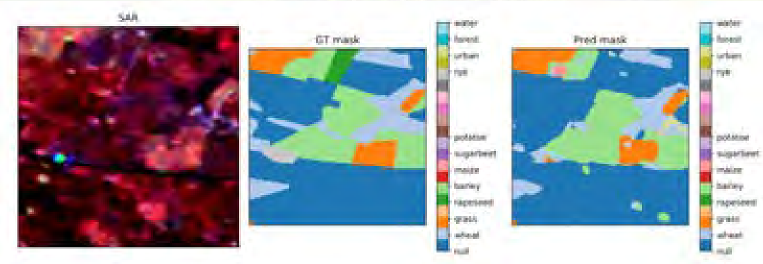
Hyperspectral images from PRISMA enabled the estimation of very important environmental variables (vegetation stress, water resources) using different post-processing methods. In particular, hyperspectral imaging demonstrated better discrimination than multispectral instruments, thanks to its fine and continuous spectral information. Both hyper- and multi-spectral sensors record radiance in the Visible to Near-Infrared (VNIR) and Short-Wave Infrared (SWIR) of the spectrum, VNIR spanning 400–1000 nm and SWIR 1000–2400 nm. Unlike multispectral sensors, which record in a fairly limited number of discrete spectral bands (4–20 bands), hyperspectral sensors include a very large number of contiguous and narrow spectral bands of 5–15 nm.

SPATIOTEMPORAL CROP TYPE CLASSIFICATION WITH DEEP LEARNING APPLIED TO SAR TIME SERIES

Mathis Lamarre Industry Fellow, Airbus



Overview of the project. Credits: Airbus DS



Example of 3D-CNN segmentation over patch of fields. SAR shows an arbitrary Red-Green-Blue composite over the area (actual input consists of 48 consecutive images). Ground-truth (GT) mask is compared to the predicted output (Pred mask). Credits: Mathis Lamarre, Airbus.

IMPORTANCE OF THE WORK

Land use classification is a critical task for agriculture monitoring. So far, most approaches based on remote sensing only exploit optical data. However, they are not always reliable as they depend heavily on clouds-free weather conditions. Radar on the other hand “sees through” clouds, which is very relevant in some areas.

Moreover, it captures additional information like texture or moisture. Besides, recent missions such as Copernicus Sentinel-1 or TerraSAR-X provide high-quality observations with a high revisit frequency. This enables the study of the growth cycle of the crops.

OVERVIEW

This work aims at using DL methods to analyse time series of SAR images for land cover monitoring and segmentation. The performance of different techniques is also compared, namely 2D-CNNs which consider spatial context and 3D-CNNs to capture temporal dependencies. It would also be interesting to use different SAR products and pre-processing algorithms as additional dimensions through polarimetry and interferometry.

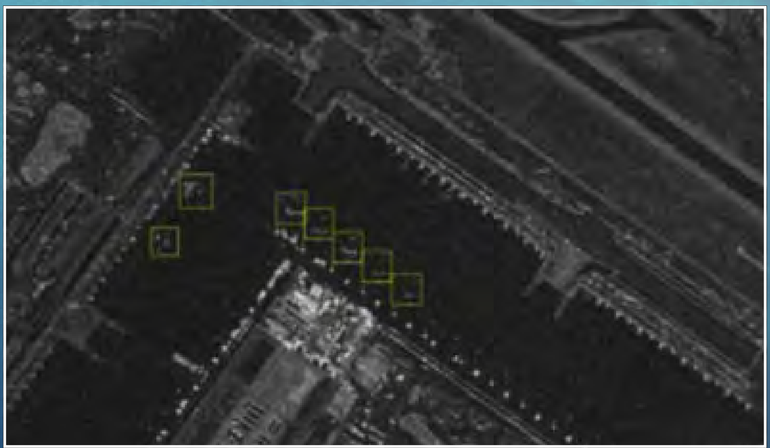
> *"From 2D to 3D, and towards 4D."*

FINDINGS

With a dataset limited in size, spatial information alone does not seem to be sufficient for accurate classification. A 2D-CNN trained on VV and VH backscatter does not generalise well on unseen fields. In order to incorporate the temporal dimension, the time stamps were stacked and used in a 3D-CNN. This type of network, originally developed for volumetric medical images, has also been used for hyperspectral image segmentation. This proved to yield good results for crop type classification. The field boundary detection is also quite accurate. The performance could still be improved for underrepresented classes and small fields. Next steps include refining the model architecture and incorporating other SAR outputs such as coherence.

USE OF AI FOR SAR IMAGE CLASSIFICATION ON-BOARD

Andrea Radius Industry Fellow, ICEYE



Plane detection using ML, from ICEYE. Credits: ICEYE.

IMPORTANCE OF THE WORK

The space environment is changing and new actors in terms of data providers complement the space industry with low cost micro-satellites with the purpose to increase EO data. The use of AI is the solution to manage these increasing volumes of data in the near future.

ICEYE is a new company that is building a constellation of micro-satellites equipped with SAR. Its activities in terms of data quality and AI techniques are focused to reinforce the synergy with ESA, increasing the collaboration, providing data validated through the ESA guidelines, sharing data for common projects as well as ideas on different applications.

OVERVIEW

The use of AI is particularly relevant when applied to the remote sensing area. The first challenge is the exploitation of a very large amount of ICEYE X-band EO data to maximise the informative content while minimising the processing time. In this context, a dedicated dataset is to be prepared with the needed ground truth and this will be shared with the Φ-lab researchers.

The second challenge is the use of AI for on-board processing. In perspective, AI could reduce the volume of data which is required to be downlinked in the process and could allow the on-board implementation of dedicated detection algorithms.

The aforementioned challenges are dependent on the calibration and validation process that is currently on-going on ICEYE satellites X2, X4 and X5.

> *“On-board AI to identify relevant change.”*

FINDINGS

With the amount of EO data drastically increasing and the appearance of new companies with new micro-satellites ICEYE is helping to foster the implementation of processing techniques based on AI to enable the rapid processing of large amounts of data, thus reducing the amount of unnecessary data transmission to the ground.

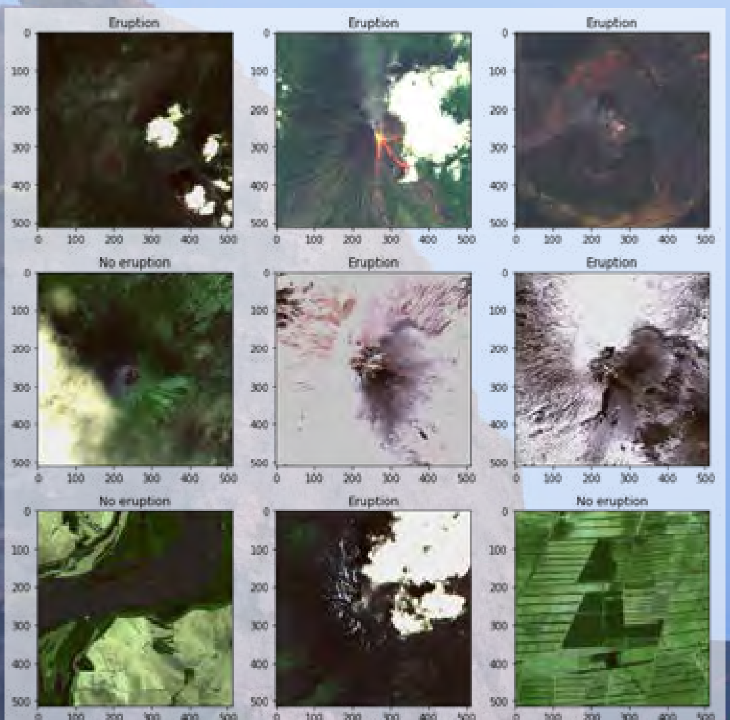
At the same time, it is important that each EO data provider characterises the quality of the data provided to the users after the calibration and validation phase.

ICEYE activities go in this direction, performing a quality assessment on the X-band SAR data and developing on board and on ground processing techniques based on AI.

In the context of the data processing on ground, ICEYE will prepare a dedicated dataset with the available ground truth for land classification and maritime monitoring using semantic segmentation with NN and will be shared with the Φ-lab researchers.

VOLCANIC ERUPTIONS DETECTION THROUGH CONVOLUTIONAL NEURAL NETWORKS

Maria Pia Del Rosso Visiting Research Fellow, University of Sannio



Test of the network on nine test images and table with the corresponding results in terms of prediction accuracy. Credit: Maria Pia Del Rosso, University of Sannio. Contains modified Copernicus Sentinel-2 data, 2019.

IMPORTANCE OF THE WORK

In the last few years, the increasing amount of available satellite data has facilitated the implementation of ML techniques to help people facing natural disasters. Among these natural disasters, volcanic eruptions continue to pose high risk, especially for people living in under-developed areas where networks of on-ground sensors are not installed. This project has been developed to show how satellite data can be used, and also mixed with other kinds of data to help improve earthquake prediction.

OVERVIEW

This research was initiated in July 2019 at ESA Φ-lab. After creating a dataset of optical satellite images containing thousands of samples both of eruptions and no eruptions in satellite data, an Artificial Neural Network was created to classify the images in the two categories required for the problem; the particular architecture for the network is the Convolutional Neural Network. This network has been trained firstly on the training dataset containing most of the samples, and then validated and tested on samples that the network has never seen before.

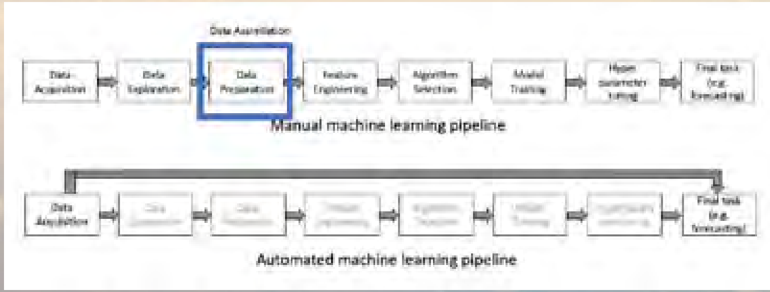
> *"AI and EO can help assessing risks."*

FINDINGS

The network achieved good results despite the low amount of images suitable for this purpose. An example of a test on nine images is presented left: the figure shows the images with their corresponding labels, while the table highlights the differences between the label values and the predicted ones. The network is almost perfect at recognising eruptions with lava flows and quite good at distinguishing eruption smokes from clouds. This works provides a preliminary classification of the images to better recognise the ones in which a volcanic eruption is present. The prediction step will be done using the images selected by the network. Additional types of data (radar, atmospheric composition data and non-satellite data) will also be used. A predictive machine learning model on eruptions would potentially avoid catastrophic consequences in terms damage and loss of lives.

AUTOMATING METHODS OF DATA ASSIMILATION

Mitra Baratchi Visiting Research Fellow, University of Leiden



Machine learning pipeline of spatio-temporal data. Automated machine learning focuses on finding automated solutions for tedious tasks in this pipeline. Credits: Mitra Baratchi, University of Leiden.

IMPORTANCE OF THE WORK

Physical domain knowledge of dynamical systems is normally encoded in the form of numerical models. Various applications make use of such models, for instance in acquiring better operational weather and ocean forecasts. There are different data assimilation techniques previously designed to take observational data acquired from noisy sensors and combine it with outputs of numerical models, to produce an optimal estimate of the evolving state of the system.

Data assimilation techniques have a several strengths in combining independent variables such as handling missing data, and implicit treatment of noise. However, the use of available data assimilation algorithms is difficult, mainly because there are many choices to be made by their user. Given a physical model, and a set of observational data, there are many data assimilation algorithms to choose from. Each of these algorithms has a different set of hyperparameters that need to be tuned. Additionally, models need to be parameterised before being used in combination with an algorithm. Identifying the optimal set of parameters and algorithms is purely a matter of trial and error and the choice depends on the dataset at hand. Therefore, users of these methods can benefit from a system that can identify the set of optimal choices in a computationally efficient way. This research focuses on automating the process of selection of the most suitable algorithm and set of parameters for each dataset.

OVERVIEW

This work fits within the context of a wider area of automated machine learning from spatio-temporal Earth Observation data. Before using data as an input to many machine learning algorithms many steps are performed. All these steps require difficult decisions to be made by a domain expert. Automating the process of making such decisions, allows for easier use of these techniques and simplifies acquiring machine learning models from data. Automated machine learning aims to find computational solutions for automating those task in the machine learning pipeline that purely depends on trial and error. Data assimilation can be considered as one of the techniques that can be performed as part of data preparation (e.g. by creating a denoised time series from raw observations). An automatic solution for data assimilation allows for easier use of these techniques while designing a machine learning pipelines, and thus better use of available domain knowledge encoded in physical models.

> *Exploring synergies between model and observations.*

FINDINGS

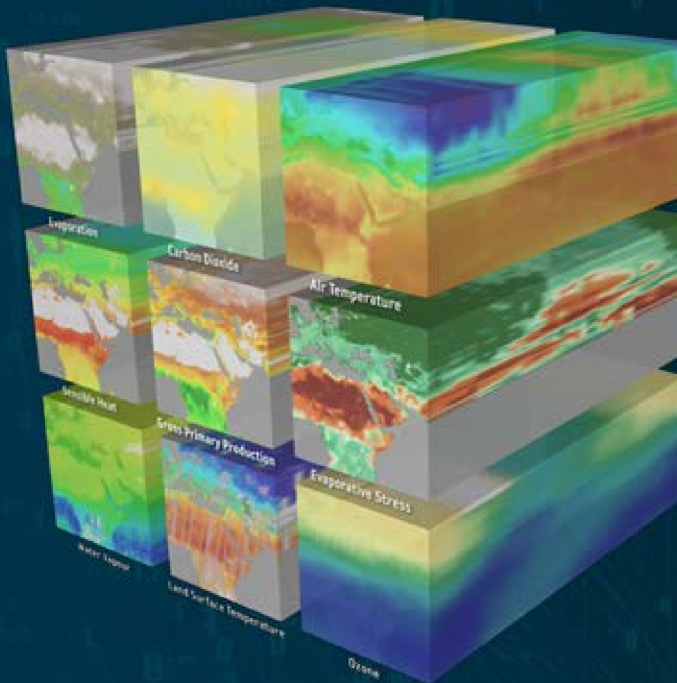
Recent advancements in the area of algorithm selection and hyperparameter optimisation have been employed to jointly select among the algorithms and also to parameterise them. Our evaluation results acquired from experiments on a data assimilation benchmark show that given a dataset, it is possible to automatically choose the best data assimilation algorithm, set the algorithm and model parameters. More importantly, the computational complexity of this task which is directly related to the number of evaluations needed in each data assimilation round is relatively low. This fact allows the applicability of automated data techniques within time-critical applications.

BUILDING CAPACITY

> *We have begun an Earth Science Grid-On-Demand service... Users now gain access to large volumes of Earth Observation data and can easily and quickly perform a variety of data reprocessing, including fine-tuning new algorithms to get optimal results.*

Luigi Fusco
Computer Science pioneer @ ESA

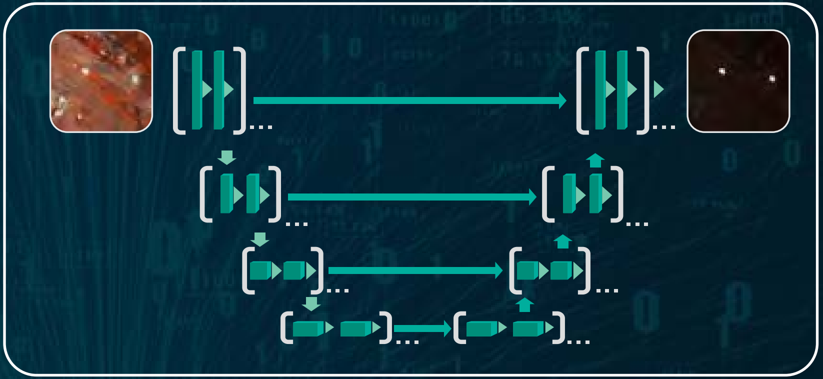
See **philab.esa.int** for more info on data and tools



Credits: Earth System Data Lab, ESA/MPI

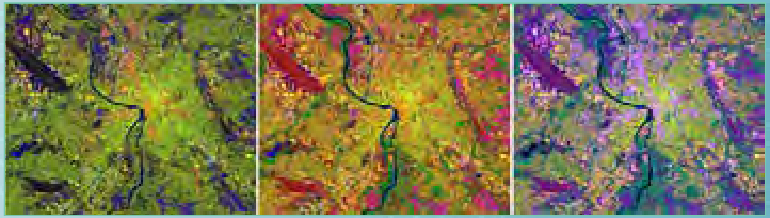
Preparing, cleaning and augmenting EO data for AI software

Exploring AI tools and architectures



THE OPEN SAR TOOLKIT FOR SENTINEL-1 ANALYSIS-READY DATA

Andreas Vollrath ESA Research Fellow



Different types of 10 m ARD Timescan products created by OST over Toulouse. Multi-temporal metrics are created from a full-year time-series of 2018 (60 images). (left) Backscatter only Timescan bands (R: VV-min, G: VH-min, B: VV-SD); (middle) Combined Coherence-Backscatter Timescan composite (R: Coherence-min, G: VH-min, B: VV-SD), (right) combined Coherence-Backscatter-Polarimetric RGB composite (R: Coherence-min, G: VH-min, B: Alpha-average). The increase in different colours is a sign of better ease of classification. Note for example the green areas in the coherence-backscatter composite (middle) along the river representing riparian vegetation, as compared to the backscatter composite on the left, where no clear distinction to other urban areas is possible

Copernicus Sentinel-1 is the first operational mission that systematically acquires SAR data on a global scale. The open and free data policy led to a massive rise of interest beyond traditional SAR experts and there is now the need of easy-to-use tools that accelerate the uptake of this valuable data source. The Open SAR Toolkit responds to the requirement by bundling the full workflow for the generation of Sentinel-1 Analysis-Ready-Data (ARD) for land in a single high-level python package. The concept of ARD is driven by the need for ready-made data that does not need further sensor-expert pre-processing, thus letting the user focus on the actual information extraction.

The toolkit can be considered as an end-to-end data preparation package that includes functionalities for data inventory and advanced sorting, as well as massive concurrent download from various data mirrors. The pre-processing routines are almost entirely based on ESA's Sentinel-1 toolbox and wrapped into a single function for the fully automated batch processing. Since, at the moment, there is no unique consensus on the specification of ARD products for SAR and the respective pre-processing steps involved, different types of ARD templates can be selected and customised.

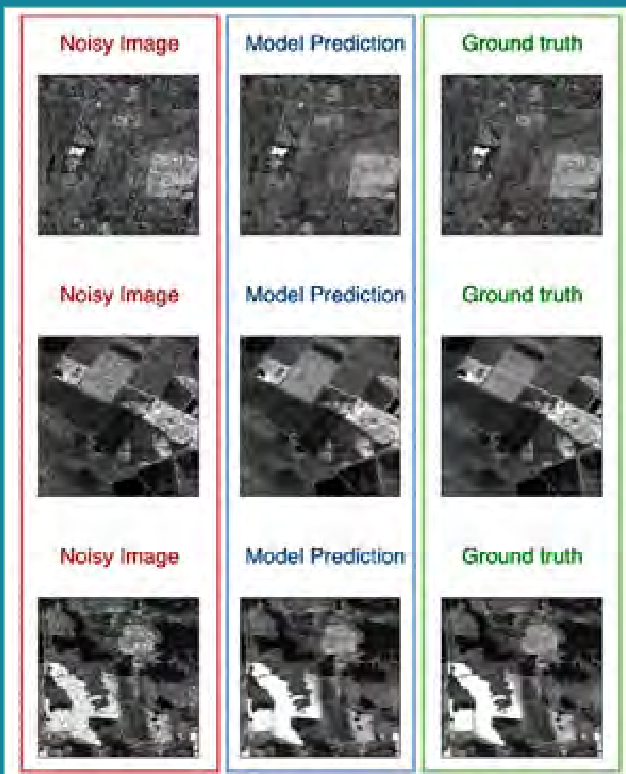
The toolkit does include advanced types of ARD such as the combined production of calibrated backscatter, interferometric coherence and the dual-polarimetric H-A-Alpha decomposition. Time-series and multi-temporal statistics (i.e. Timescans) can be produced for each of these layers. The generation of seamless large-scale mosaics over time is also possible. Jupyter notebooks are the main way to interact with the tools whereas tutorial notebooks are available to get started.

SPECKLE FILTERING THROUGH CONVOLUTIONAL NEURAL NETWORKS

Alessandro Sebastianelli Visiting researcher, University of Sannio

The idea is to develop an AI model capable of filtering speckle noise from Sentinel-1 data. After creating a dataset of Synthetic Aperture Radar (SAR) satellite images containing thousands of time series, we used it to get its speckle free version, by averaging over time (see figure below). The training set has been created by applying common statistical models representing noise to the speckle free dataset.

A Neural Network has been created to learn how to filter speckle from SAR images. An example of test on images is shown below: the figure shows the images with the corresponding labels while the graphs highlight the evaluation of the proposed model by using two metrics. This led to the understanding that the presence of speckle noise leads to bad results for the generative problem proposed for the "Seeing through clouds" project, and hence the reason for the development of this tool. This tool helps improve all the projects based on the Sentinel-1 images, in fact it can be used to pre-filter the data in a faster way.

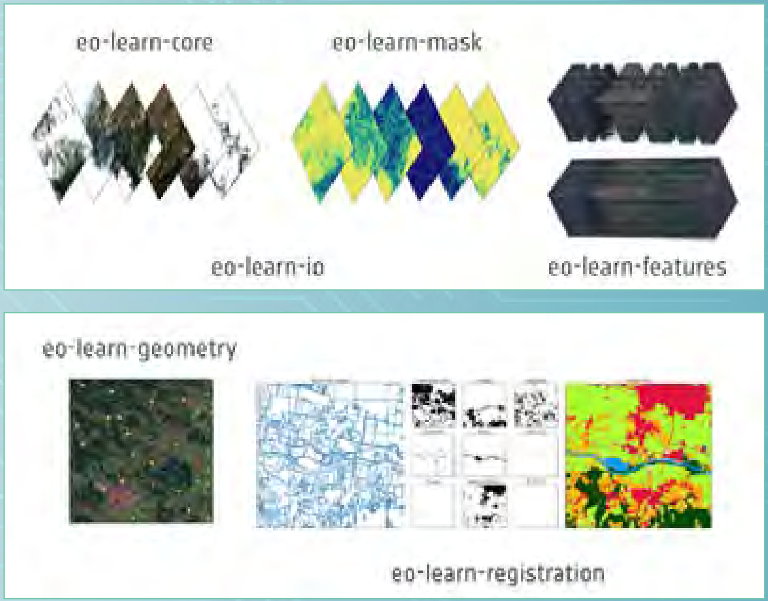


First column the noisy image, second column the AI model output and third column the ground truth.

Dataset creation

EO-LEARN OPEN-SOURCE TOOLKIT

Devis Peressutti Sinergise
Drew Bollinger Development Seed
Patrick Helber Vision Impulse
for QueryPlanet project



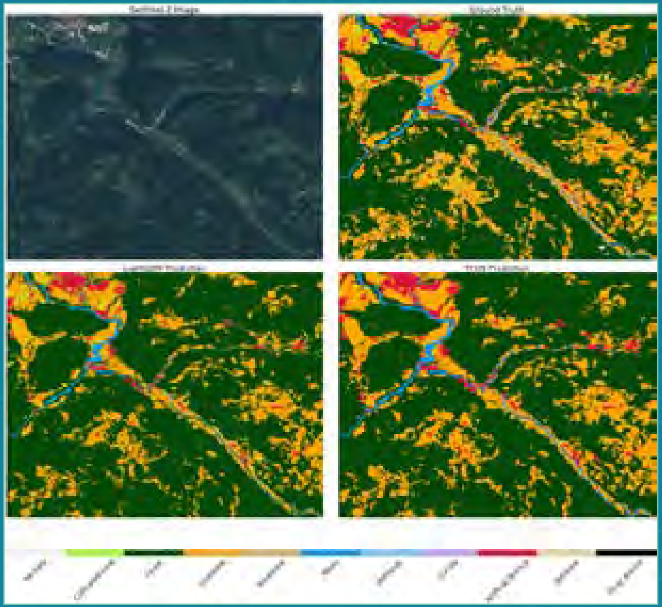
Example of ML pipeline for land cover classification.

The eo-learn is an open-source library bridging the gap between EO data and existing ML technologies. A collection of Python packages has been developed to seamlessly access and process spatio-temporal image sequences acquired by any satellite fleet in a timely and automatic manner. The design of eo-learn is modular and encourages collaboration – sharing and re-using of specific tasks in a typical EO-value-extraction workflow, such as cloud masking, image co-registration, feature extraction, classification, etc. Since it is open source, everyone is free to use any of the available tasks and is encouraged to improve the existing ones, develop new ones and share them with the rest of the community. A collection of Jupyter notebook examples was prepared showcasing how to couple eo-learn with the most popular ML frameworks, such as scikit-learn, Keras, PyTorch and fastai.

These examples allow users to quickly set up their environment using tools such as Docker and AWS SageMaker to ease the training and prediction procedure. Several AI-ready datasets are available, as for example automatic label retrieval from OpenStreetMap, high-resolution imagery from Mapbox and medium resolution from Sentinel, urban settlements in Europe, land cover datasets as well as many other open-source datasets.

Sample use-cases were demonstrated, readily available for re-use, covering water monitoring on global scale, land cover segmentation, mapping of urban settlements and others.

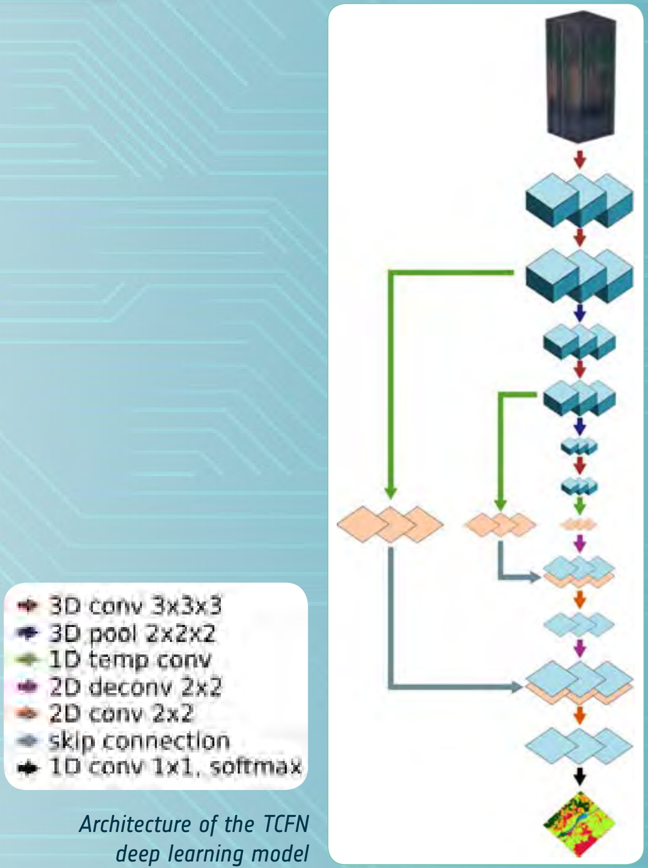
Blue Dot Water Observatory - global monitoring of water bodies.



Transfer Learning Example.

SPATIO-TEMPORAL DEEP LEARNING FOR LAND COVER CLASSIFICATION

Matic Lubej
Devis Peressutti
Sinergise
for QueryPlanet project



Architecture of the TCFN deep learning model

ML and, in particular DL methods have become the state-of-the-art in many vision, language, and signal processing tasks, due to their ability to extract patterns from complex high-dimensional input data. Classical ML methods, such as random forests and support vector machines have been used in many EO applications to analyse temporal series of remote sensing images. On the other hand, CNN have been employed to analyse the spatial correlations between neighbouring observations, however mainly in single temporal scene applications. This tool describes a deep learning architecture capable of simultaneously analysing the spatio-temporal relationships of satellite image series. The work is demonstrated on an application on the land cover classification of the Republic of Slovenia, using annual Copernicus Sentinel-2 satellite images for the year 2017. Temporal Fully-Convolutional Network (TFCN) extends the FCN architecture (e.g. U-Net), which is currently the state-of-the-art in single scene semantic segmentation. The architecture exploits spatio-temporal correlations to maximise the classification score, with the additional benefit of representing spatial relationships at different scales due to the encoding-decoding U-Net structure. The algorithm performs a 3D convolution in the spatial as well as the temporal dimension. By default, max-pooling is performed in the spatial domain only. As the target land cover labels are not-time dependent (i.e. one label per pixel is available for the entire time-series), 1D convolutions along the temporal dimension are performed in the decoding path of the architecture to linearly combine and reduce the temporal features. The schematic of the model architecture is shown in Figure. The output of the network results in a 2D label map which is compared to the ground-truth labels.

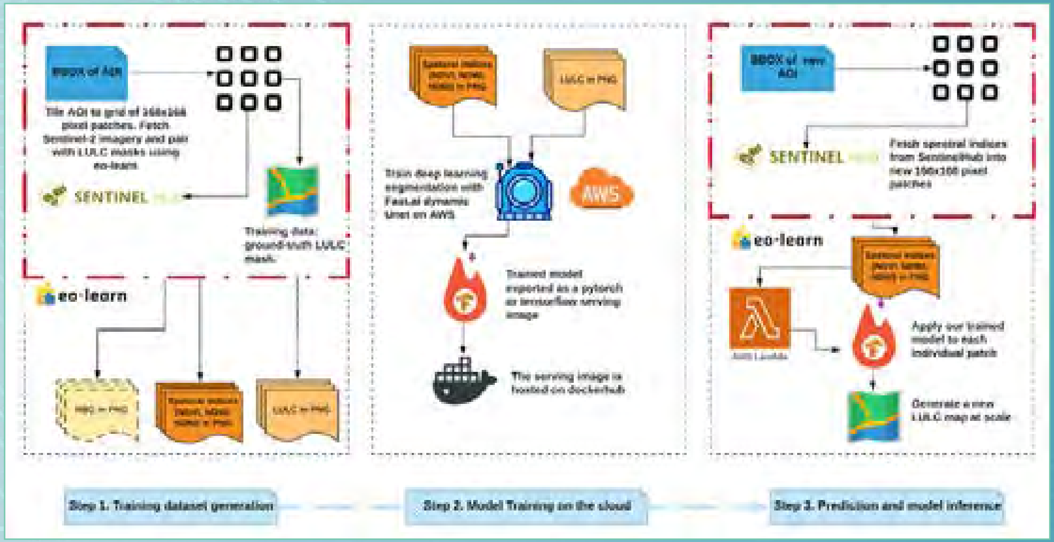
The architecture performed three encoding and decoding steps (i.e. three max-pooling and three deconvolution layers), with a bank of two convolution layers at each encoding and decoding scale. The number of convolutional features was set to 16 at the original scale, with a factor of 2 applied at each deeper level, with a kernel width of 3. The Adam optimiser (learning rate 0.001) is employed. The TFCN model is implemented in TensorFlow.

The framework was used to generate a land cover map of the Republic of Slovenia for the year 2017. The inputs to the framework are a shape-file defining the geometry of the AOI, the Sentinel-2 L1C images for the entire year, and a set of training labels. By avoiding to download and process entire tile products (e.g. Sentinel-2 granules), eo-learn provides flexibility and facilitates an automation of the processing pipelines. A pipeline is defined as a connected acyclic graph of well-specified tasks to be performed on the data. eo-learn supports parallelisation of operations, such that the same workflow (e.g. data preparation for land cover classification) can be run in parallel for the smaller patches constituting the AOI. Logging and reporting allow to monitor and debug the execution of the processing pipeline.

The trained model was used to predict the labels on the test sample and the obtained results were then validated against the ground-truth. An overall accuracy of 84.4% and a weighted F1 score of 85.4% were achieved. In general, poor prediction was obtained for under-represented classes such as wetlands and shrubland. These results represent preliminary work on a prototype architecture which was not optimised for the task at hand. Despite this, results in line with previously reported work were achieved. Optimisation of the architecture (e.g. number of features, depth of the network, number of convolutions) and of the hyper-parameters (e.g. learning rate, number of epochs, class weighting) could improve the results of TFCN even more.

DYNAMIC U-NET FOR TRACKING A RAPIDLY CHANGING PLANET

Drew Bollinger
Zhuangfang NaNa Yi
Development Seed | for QueryPlanet project



The deep learning pipeline that fetches and creates training data for LULC modelling on the cloud. It can be scaled up with our current open-source, cloud-based pipeline, chip-n-scale.

This tool demonstrates a method for rapid classification of Land Use and Land Cover (LULC) on Sentinel data by using Sentinel Hub, eo-learn, fastai and chip-n-scale prediction tools. To enable this workflow, a new utility was built, fastai-serving, that enables scaling predictions on the cloud. Tracking changes in land use is essential for understanding urban changes, deforestation, and animal habitats. Copernicus Sentinel-2 satellites, in particular, are useful for this task because of their relatively high revisit rate, resolution, and multiple agriculture-focused wavelengths. A single date land classification approach was designed, achieving very good results and is easy to quickly run and scale. This approach is often faster and cheaper than alternative approaches that use data from multiple points in time. Multitemporal approaches can better account for noise and clouds and tend to do much better at distinguishing between certain agricultural and vegetation areas.

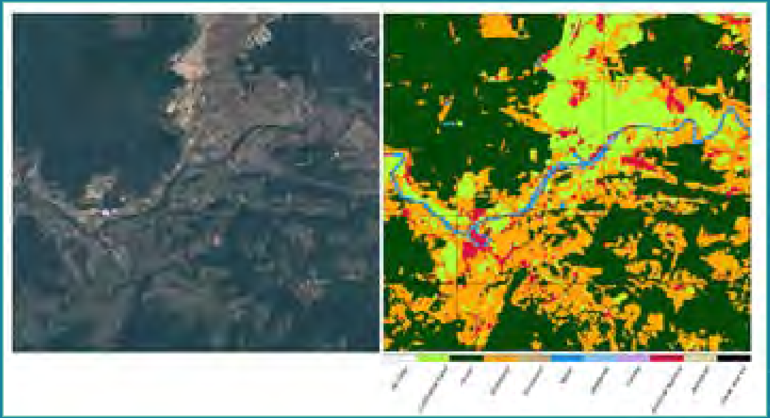
However, single observation methods can be valuable for certain applications. Sometimes it is valuable to trade a small amount of accuracy for a much faster and less expensive result. This might be the case when scaling over a very large area, when doing initial exploration and hypothesis testing, or when responding to certain rapid onset disaster situations. Using less data and limited hyperparameter tuning allows for faster iterations while developing a model. Predictions on a single date are also valuable for tracking land use change within a season. Urchn uses regular land use assessments as one input to flag areas where the map may be out of date.

The workflow optimises for speed and efficiency. It leverages fastai, a ML framework that packages many modern machine practices for easier algorithm development, and uses a pre-trained dynamic U-Net available in fastai to very quickly get the model running.

The pre-trained U-Net is designed for three-channel (RGB) imagery. To accommodate on Sentinel-2 imagery one needs to reduce the imagery input from the original thirteen bands without losing too much spectral information. The eo-learn is used to compute three-band combinations (NDVI, NDWI, and NDBI) that carry good signal for land use and substitute these in place of the RGB channels. The eo-learn framework also served as the primary data acquisition and transformation framework prior to the data being used in fastai.

Once the model was trained and tested over a sample area, one can use it to predict land use classification over a much larger area. In the case of the experiment, the model was run over all of Slovenia in order to compare with the previous models available for this area. The standard approach for this type of problem is to use “chip-n-scale”. Chip-n-scale relies on having the model as a TensorFlow Serving Image. In order to convert the model to this format, a new API was released for running fastai models with the same API as TensorFlow serving – fastai-serving.

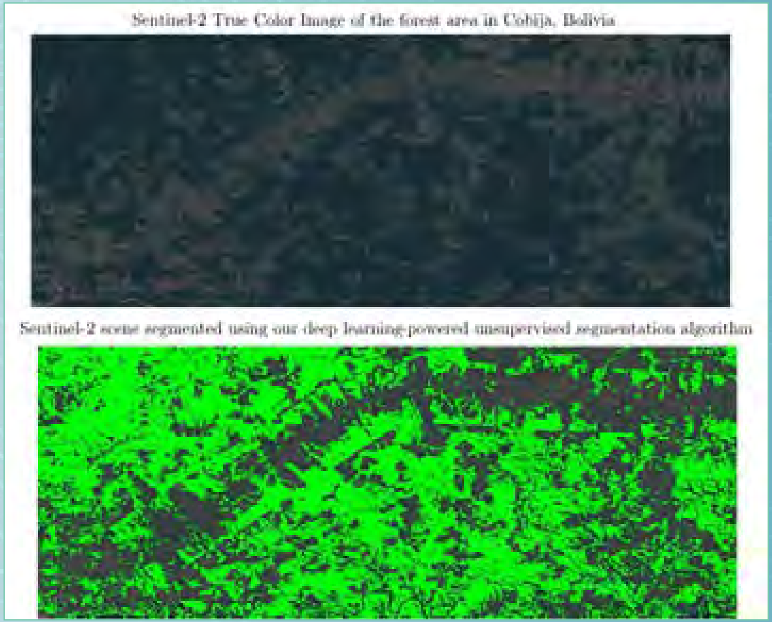
Using the framework, one is able to run inference over all of Slovenia (20 000 sq. km) in under 15 minutes. Because of the inherent variability using a single date, the model is more susceptible to noise or clouds, and has a lower overall accuracy score (85% vs 94% pixels correctly labelled). Still, the time and cost savings are a reasonable trade-off in certain situations.



A U-Net, applied ResNet50 as the encoder, was trained. The LULC prediction on the right, and the associated satellite imagery, in RGB, shown on the left.

MACHINE LEARNING TOOLBOX FOR HYPERSPECTRAL DATA

Jakub Nalepa | KP Labs, Silesian University of Technology | for Hypernet project

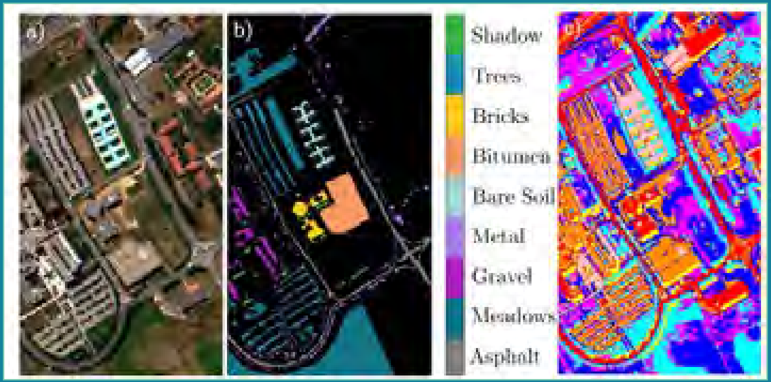


Classification and segmentation of multi/hyperspectral images have wide practical applications – automated segmentation of multi/hyperspectral imagery can help quantify the forest area in a fully reproducible way. Hence, it can be used to detect illegal logging. Credits: KP Labs.

HyperSpectral Imaging (HSI) has become a mature technology which brings exciting possibilities in various Earth Observation applications in a plethora of fields, including precision agriculture, forestry, event detection and tracking. This toolbox aims at addressing the most important challenges toward effective hyperspectral image analysis and at developing techniques and tools which will help extract value from such highly dimensional image data. Since the approaches are generic, they can be easily deployed in a range of real-life applications, therefore enable easier adoption of this revolutionising technology in practice. Tools include a comprehensive and ready-to-use software suite for hyperspectral image analysis. It does not only encompass classification/ segmentation algorithms, but also various techniques for data reduction, feature selection and extraction, visualisation, optimisation, batch analysis and processing, and much more. Also, it ensures that one can seamlessly load data using widely used hyperspectral data formats. The Jupyter notebooks make understanding of the software straightforward and allow a new user to quickly go through the internals of the package. The algorithms were verified and validated in the wild, and results were presented and published in several scientific journals.

Supervised classification and segmentation of hyperspectral images

Although the number of manually annotated ground-truth hyperspectral sets is still limited, supervised classification techniques are being actively developed in the literature. Here, the deep learning-powered algorithms have established the current state of the art in the field. Both state-of-the-art deep networks for hyperspectral image classification have been implemented (a spectral-spatial neural network alongside a convolutional neural network with multiple feature learning), and the attention-based convolutional neural networks (which allow to additionally determine the most informative bands during the training process), and various spectral models.



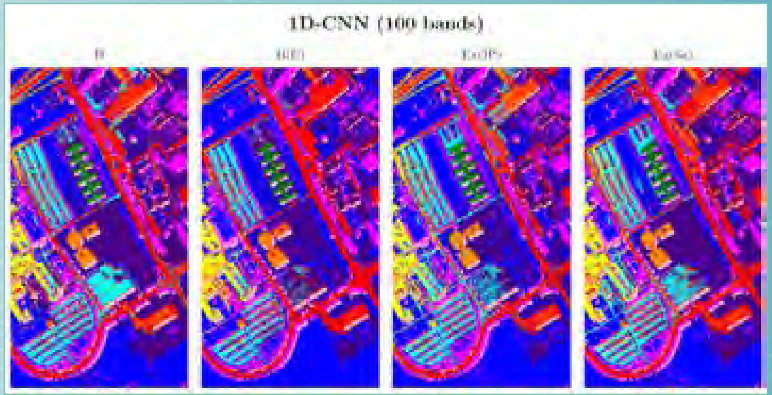
Unsupervised segmentation offers new possibilities of unrevealing information captured within newly acquired hyperspectral images and existent benchmarks. This example shows: 1) the Pavia University scene, 2) its ground truth (black colour is "unknown class"), and 3) the full 3D-CAE segmentation which is not only very detailed but also sheds new light on those 'unknown' objects.
Credits: KP Labs.

Unsupervised segmentation of hyperspectral images

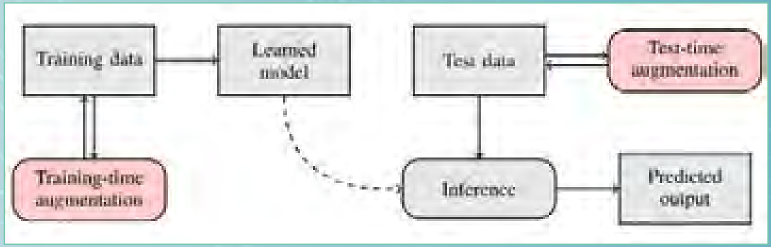
Although DL has established the state-of-the-art in the field, it remains challenging to train well-generalising models due to the lack of ground-truth data. This problem was tackled and an end-to-end approach is proposed to segment hyperspectral images in a fully unsupervised way. This introduces a new deep architecture, which couples 3D convolutional autoencoders (3D-CAE) with clustering and showed that it can be used to process any hyperspectral data without any prior class labels available.

Transfer learning for hyperspectral image classification

This shows how to effectively deal with a limited number and size of available hyperspectral ground-truth sets and apply transfer learning for building deep feature extractors in the supervised setting. Also, the spectral dimensionality reduction is exploited (by simulating wider spectral bands) to make the technique applicable over hyperspectral data acquired using different sensors, which may capture different numbers of hyperspectral bands. The experiments, performed over several benchmarks and backed up with statistical tests, indicated that this approach allows to effectively train well-generalising deep convolutional neural nets even using significantly reduced data.



Example visualisations of the segmentations obtained using the spectral convolutional neural network over spectrally reduced data (100 simulated bands) trained in various ways: using the balanced training set (B), reduced balanced data B(E), and with transfer learning, where feature extractors were trained over different source data – Ex(IP) and Ex(Sa). Credits: KP Labs.



Training-time augmentation increases the size and representativeness of a training set, whereas the test-time augmentation creates synthesised examples based on the incoming one to form a voting classification ensemble. Credits: KP Labs.

Training- and test-time augmentation of hyperspectral data

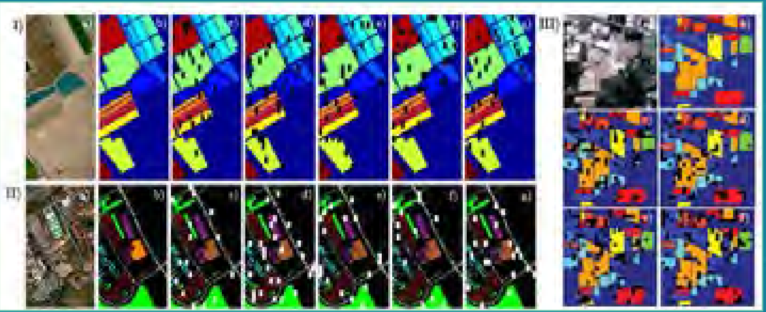
Data augmentation helps improving generalisation capabilities of deep neural networks when only limited ground-truth training data are available. This tool addresses test-time augmentation of hyperspectral data, which is executed during the inference rather than before the training of deep networks. Also, it introduces two augmentation techniques, which can be applied at both training time and test time and exploitation e.g., principal component analysis-based approaches. The experiments revealed that the augmentations boost generalisation of deep models and work in real time, and the test-time approach can be combined with training-time techniques to enhance the classification accuracy. Finally, the implementations include other state-of-the-art hyperspectral data augmentation algorithms, including generative adversarial networks and noise injection.

Validating hyperspectral image classification and segmentation

Validating hyperspectral image segmentation algorithms is a challenging task due to the limited number of manually annotated ground-truth sets. Practically all segmentation techniques have been tested using up to three benchmarks, with Salinas Valley, Pavia University and Indian Pines constituting the mainstream. A common approach is to extract training and test pixels from the very same hyperspectral scene, and almost all algorithms are being validated in the Monte-Carlo cross-validation setting. Such random selection of training and test sets may, however, lead to overoptimistic results (as the training-test information leak can occur), especially for spectral-spatial algorithms. To address this issue, a tool was developed for elaborating patch-based training-validation-test splits which helps quantify the classification performance of emerging hyperspectral classification algorithms without any information leakages. This tool has been used to elaborate the splits that are publicly available.

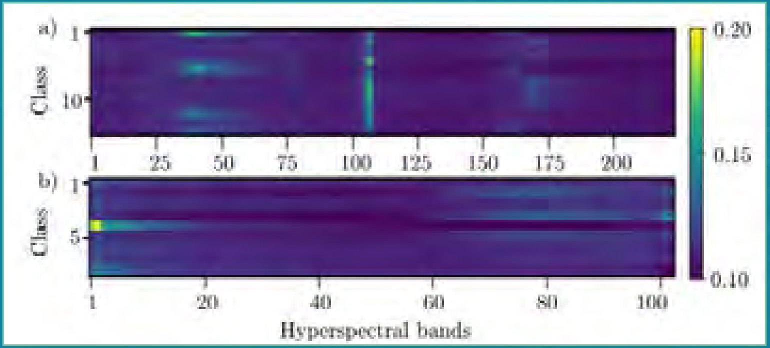
Attention-based CNN for hyperspectral band selection

To reduce the time (and ultimately cost) of transferring hyperspectral data from a satellite back to Earth, various band selection algorithms have been proposed. They are built upon the observation that for a vast number of applications only a subset of all bands conveys the important information about the underlying material, hence one can safely decrease the data dimensionality without deteriorating the performance of hyperspectral classification and segmentation techniques.



Our benchmark data generated over the (I) Salinas Valley (five non-overlapping folds), (II) Pavia University (five folds), and (III) Indian Pines (four folds). (a) True-colour composite, (b) ground-truth segmentation, (c)-(g) visualisation of all folds for Salinas and Pavia, (c)-(f) visualisation of all folds for Indian Pines. Black patches (white for Pavia University) indicate the training pixels, whereas the other pixels are used for testing.
Credits: KP Labs.

Here a novel algorithm is introduced for hyperspectral band selection that couples new attention-based convolutional neural networks used to weight the bands according to their importance with an anomaly detection technique which is exploited for selecting the most important bands. The proposed attention-based approach is data-driven, re-uses convolutional activations at different depths of a deep architecture, identifying the most informative regions of the spectrum. Additionally, other state-of-the-art band selection techniques were implemented, including both filter and wrapper approaches.



Example average attention-score heatmaps for a) Salinas Valley and b) Pavia University show that certain bands convey more information than the others (the brighter the regions are, the higher attention scores were obtained, hence these bands are more 'important'). Credits: KP Labs.

CROWDSOURCING PLATFORM

Christopher Stewart ESA Research Fellow

Example crowdsourcing activity for archaeological prospection.
Credits: Scifabric & ESA

A crowdsourcing tool has been developed for the Φ-lab to support activities where manual interpretation of EO data is required. Based on Pybossa, an Open Source framework for crowdsourcing, the tool allows users to freely publish EO related crowdsourcing projects, or contribute to already published activities.

An activity that has been released on this platform includes crowdsourcing for archaeological prospection. This forms part of a Φ-lab activity that attempts to prototype an automatic methodology for retrieving archaeological cropmarks through a combination of human interpretation and machine learning. Users are presented with tiles of high resolution optical EO data, and are asked to identify vegetation patterns as proxies of buried structures.

Results show that a redundancy of three independent interpretations of each tile is sufficient to yield accurate classifications, with many detected buried archaeological features including roads, buildings and even urban areas. Once a critical mass of training data is available, the intention is to use it to train a machine learning model to scale automatic detections over a wider area, which currently only includes the region surrounding the city of Rome.

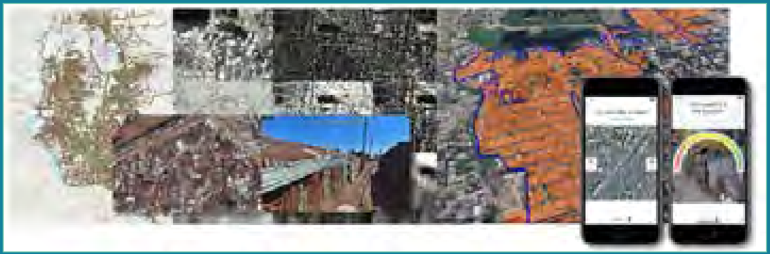
The pressure of development is putting the cultural heritage of many areas at risk. Results of this activity will demonstrate efficient alternatives to the conventional ground-based techniques of rescue archaeology.

GAMING APPROACHES FOR CROWDSOURCING URBAN INFORMATION

Tomas Soukup GISAT
Steffen Fritz IIASA
Christoph Perger Spatial Focus

There have been remarkable advancements in EO state-of-the-art information extraction techniques using ML/DL approaches recently, including urban domain. Still, local in-situ information for training of machine-learning algorithms, Quality Assurance/Quality Control and validation are often limiting factors, due to their sub-optimal availability and coverage. In general, there is a lack of appropriate in-situ data collections for most of urban monitoring themes since this is demanding both organisationally and financially. This is true particularly for remote or unsafe regions around the world and/or for complex thematic features. Therefore, there is still considerable room for improvement in organisation of such collections and thus a huge potential to further operationalise many urban mapping and monitoring services.

In this situation, a novel scalable framework for EO based information extraction, supported by crowdsourcing and gaming approaches has been developed within the framework of the ESA supported GAME.EO project.



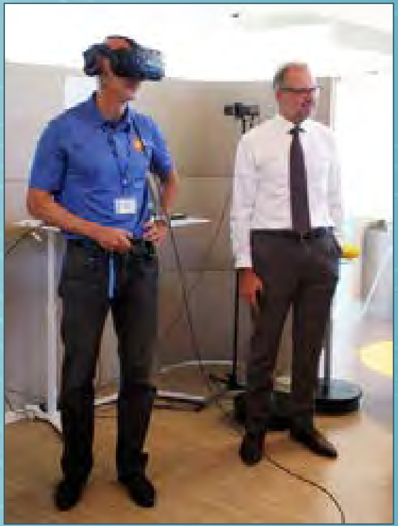
The approach is demonstrated using selected service cases for EO-based monitoring of slum areas (Sustainable Development Goal 11) in development support context. The tool aims for actionable use of crowdsourcing and gaming tools to enhance current machine-learning algorithms for the identification, delineation and further characterisation of slum areas. Developed framework and tools are tested in cooperation with World Bank Group users and stakeholders from the Water and Urban Global Practices.

Tools supporting crowdsourcing campaigns via gaming engagements can be used to mobilise and train volunteers to provide training data and to help extracting required information in a more timely and accurate manner and with lower operational costs than would be incurred using standard data collection services. It demonstrates a real potential and an added value of the synergies of crowdsourcing and EO-based information to support the development actors and cities research and operational activities globally.

Example of crowdsourcing and gaming tools supporting slum mapping in Dhaka, Bangladesh.
Credits: Gisat, GAME.EO

VIRTUAL REALITY (VR)

Paulo Sacramento Solenix



Immersive Experience was ranked at number 6 in Gartner's Top 10 Strategic Technology Trends for 2019. This category encompassed virtual reality (VR), augmented reality (AR) and mixed reality (MR) technologies that are changing the way in which people perceive the digital world. Gartner predicts that by 2022, 70% of enterprises will be experimenting with immersive technologies and 25% will be using them in production.

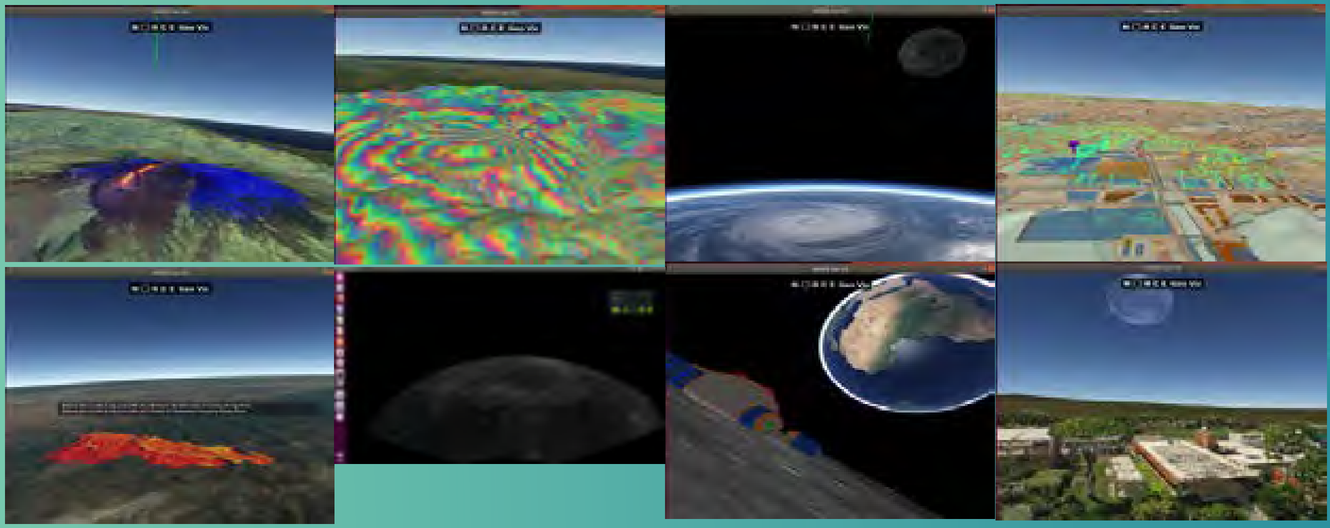
The Φ-lab hosts a permanent VR installation developed and deployed in partnership with NORCE, the Norwegian Research Centre. This installation utilises general purpose VR hardware, general purpose VR software and NORCE's GeoViz software.

The GeoViz software allows the user to navigate and explore a 3D globe of Earth as well as to fly to other planets/satellites (e.g. the Moon). ESA EO promotion and outreach materials, consisting of processed EO satellite imagery as well as infographics and videos can be explored via this virtual environment. Users can choose which locations to explore and, for each of those locations, which EO images/products to examine (typical products are Sentinel-2 band combinations, Sentinel-3 scenes and Sentinel-1 derived products such as interferograms, coherence and intensity). GeoViz is also able to display 3D models (e.g. of spacecraft,

Astronaut Luca Parmitano (left and right, with D/EOP Josef Aschbacher) and ESA DG Jan Wöerner during ESRIN 50 years event (bottom). Credits: ESA buildings), point clouds (e.g. of cities, sites, inSAR showing soil subsidence) and laser scans (e.g. of plantations). It is also possible to visualise hyperspectral data: hyperspectral datasets can be loaded and users can browse through 200 bands or more in a matter of seconds using convenient VR controller buttons.

The permanent Φ-lab VR installation is normally deployed in the Φ-lab's Archimedes mini-theatre, but it is also regularly used to support events, both inside and outside ESRIN, in several locations around Europe (ESRIN 50 years, Living Planet Symposium, Φ-Week, European Researcher's Night, EGU, SLUSH, Maker Faire).

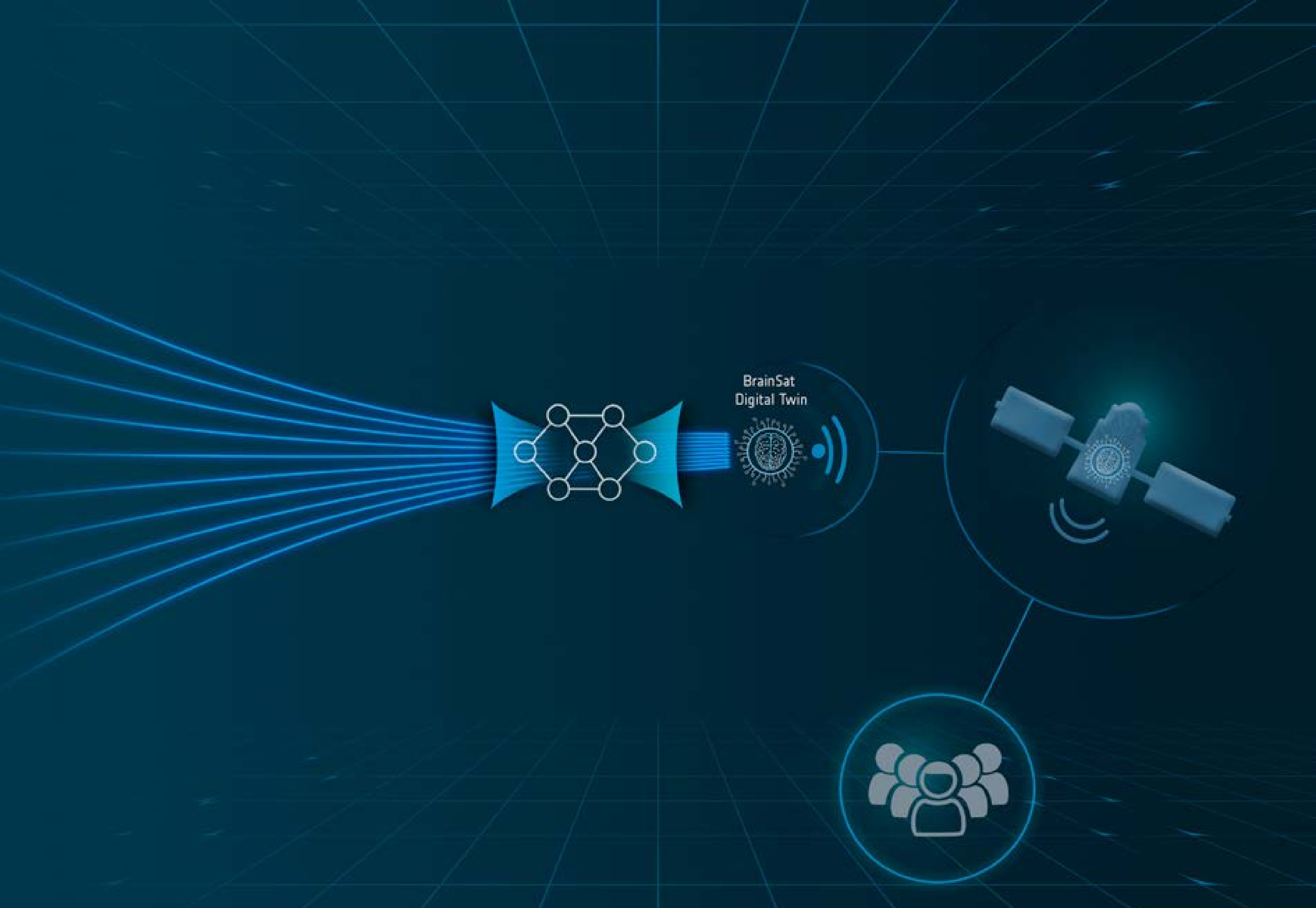
The Φ-lab engineering support team. Credits: ESA.



GeoViz showing hyperspectral AVIRIS band 96 before and after California King Fire 2014; the Moon; the Moon Village and perspective of the Earth from the Moon; ESRIN point-cloud model. Credits: ESA.

AI @ EDGE

> *What happens at the edge, stays at the edge.*
>
> **David Moloney**
> *Intel*


BrainSat
Digital Twin

Is the next big wave of AI happening at the edge?

Over the last decade, much of the progress in computer vision has been achieved through deep networks. Deep neural network models are computationally very demanding due to their large and growing size, already reaching several millions of parameters and neurons. Today, high-performance computers with dedicated architectures such as GPUs are needed to train and run these huge and complex models. This is why deploying DL models at the "**edge**", meaning running them **locally on devices** such as IoT sensors, mobile phones, drones or even spacecrafts, remains a big challenge due to the limited computing power and memory of these small edge devices.

However, today, a new model to perform AI@edge is rapidly emerging driven by the amazing advances in the new generation of AI accelerator chips, which can combine high-processing power (order Terraflops) with ultra-low-power consumption (1-10W) within a very small and light footprint (e.g. super-computer fitting into a hand). A global race for AI hardware is currently ongoing to build the most powerful, low-power and low-cost (typically below 100$) chips. This race is led by the big ICT companies such as for example NVIDIA with the Jetson Nano, Intel with the Myriad Vision Processing Unit, Google with Tensor Processing Unit.

Embedding the ML processing within the edge device and running it locally, i.e. bringing the computation as close as possible to the sensor, offers several advantages:

- **High Responsiveness.** The edge computing devices can rapidly process data in real-time and can take decisions autonomously without being connected. This is ideal for "time-critical decisions" such as self-driving or for managing disasters with EO data requiring low latency for information.
- **Low-data rate.** The AI-equipped sensor can extract in real-time the relevant information from the large amount of raw data and only streams back a small amount of highly actionable data. This is ideal when there are limitations on connectivity and bandwidth for streaming data.
- **Enhanced security.** Processing on the sensor prevents privacy leaks, as there is less streaming of data to the cloud. This is ideal in combination with blockchain to ensure secured and automatic machine-to-machine interactions (e.g. smart contracts) within a large network of connected sensors.
- **Versatility.** The edge device can be easily re-programmed for new functions by uploading new parameters, through active learning or other re-training techniques. This is ideal to reconfigure the sensor or extract more capability for a suite of versatile applications, as Tesla is doing when uploading new software on the car to increase autonomy and performance.

However, running ML at the edge also presents many challenges. One key issue is the need to "simplify" the "complex" DL algorithms to run on edge chips. A widely used approach to address this challenge is to cut some network links and neurons to reduce complexity and memory while trying to keep the accuracy. This technique called "pruning" is illustrated schematically below. In the lab we are exploring these simplifications for a variety of EO applications.

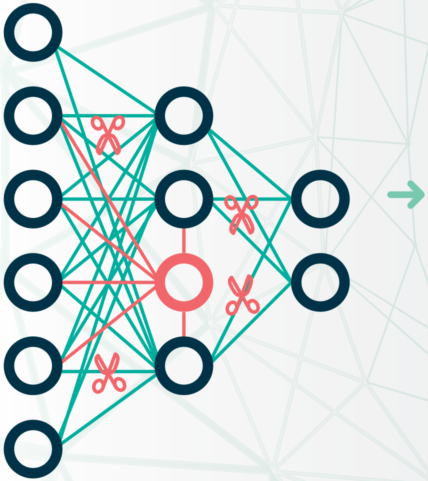


Illustration of the "pruning" of Neural Network, aiming to reduce the size of a "complex" network by removing unnecessary parameters or nodes to make it more "simple", while trying to maintain much of the performance and accuracy as the original.

Given this emerging high potential of AI@edge, ESA is now exploring the value to deploy AI at the ultimate edge: in Space. In this context, ESA will be launching a cubesat mission called FSSCat including a suite of sensors (see illustration) but also an experiment dedicated to AI on-board. The idea of the so-called Φ-sat-1 mission is to run AI on the Intel Myriad-2 chip to identify the relevant data of the Hyperscout-2 sensor to download (meaning without cloud cover).

This experiment could open new avenues for innovative use AI for space and AI in space, **whereby the sensors become "smarter" as they capture the "meaning" of the image while capturing light**. It could also explore new ways to upgrade the sensor capabilities by re-programming the brain of the satellite for versatile applications. This experiment is just the beginning of a family of Φ-sat missions. We are only scratching the surface of what AI can do for space and how it can shape it in the next decade.

“The value of satellite-based EO no longer grows with the ability to collect and transmit data back to Earth, it increasingly lies with the ability to transmit customer-relevant insight in real-time.”

Peter Platzer,
Spire, Φ-week 2019

→ *Illustration of the FSSCat mission consisting of two federated 6U Cubesats carrying a dual microwave payload (a GNSS-Reflectometer and a L-band radiometer with interference detection/mitigation) and a multi-spectral optical payload (including Thermal Infrared and hyperspectral sensor called HyperScout-2) to measure soil moisture, ice extent, and ice thickness, and to detect melting ponds over ice. It also includes a radio/optical inter-satellite link and an Iridium intersatellite link to test some of the techniques and technologies for upcoming satellite federations. A part of the mission referred to as "Φ-sat-1" aims to test the power of AI onboard to automatically classify suitable images to download (based on the cloud cover) by use of the Myriad Vision Processing Unit.*

Illustration of a vision of futureEO made by Josef Aschbacher during Φ-week 2019, whereby a network of "smart" and "connected" sensors (possibly powered by AI) in satellites, drones and IoT are able to talk to each other as part of an integrated observing system operating as a super-powerful adaptive "virtual telescope" monitoring our planet. The satellite layer is a unique European asset including a suite of research (Earth Explorers) and operational missions (e.g. meteorology and Copernicus).


aeolus
earthcare
biomass
flex
mtg-i
metop-sg-b
mtg-s
sentinel-4
metop-sg-a
sentinel-5
sentinel-5p
sentinel-6

OUTREACH & PARTNERSHIPS

Engaging with citizens, communities and users

> *If you want to build a ship, don't drum up people to collect wood and don't assign them tasks and work, but rather teach them to long for the endless immensity of the sea.*
>
> ***Antoine de Saint-Exupery*** *(1900-1944)*

THE Φ-WEEK

The Φ-week (phiweek.esa.int) is an interactive conference organised by EOP-S and the Φ-lab that concentrates on topics of EO Open Science and Future EO. This week, which gathered **600 people from 38 countries**, counting over **80 oral** presentations, **150 posters**, and over **40 exhibitors**, responds to the latest developments in EO Open Science and future trends. Beyond the classical oral and poster sessions the Φ-week included a variety of side events such as workshops, round tables, hackathons and start up tables.

Φ-week 2019 explored how satellite data coupled with new digital technologies can bring benefits to business, industry and science. It provided the opportunity for emerging space investors tech leaders, entrepreneurs, space scientists and Earth observation researchers to explore new opportunities for **cutting-edge technologies**. Over the course of the week, researchers presented an array of projects on how satellite data can be used with AI Machine Learning and Deep Learning to improve the understanding of our planet and our daily lives.

The Φ-week gave **researchers** the possibility to explore new tools being developed that offer insight into Earth as an interconnected system. While scientists and researchers work to improve the observations of our planet with new **technologies, entrepreneurs** focus on bringing these data products to the end user. The event offered the opportunity to solidify **partnerships** and create new opportunities between European research groups industry and space agencies. On the sidelines of the Φ-week multiple events were held on a variety of topics from **science communication** to **training sessions** on data processing tools, artificial intelligence applied to EO thematic **workshop** and even a **bootcamp** to develop innovative solutions to solve big industry challenges using Earth Observation data.

Participants at Φ-week side events: Φ-week bootcamp (top), Social Media and Science Communication Workshop, (bottom). *Credits: ESA*

CITIZEN SCIENCE EARTH OBSERVATION LAB (CSEOL) BOOTCAMP

Esrin-ESA hosted the **CSEOL Bootcamp and Pitch Day** from 9–11 July 2019. The bootcamp, organised by the ESA-funded **Citizen Science Earth Observation Lab** (cseol.eu), prepared nine innovative teams from Europe and beyond on the topic of Earth Observation and Citizen Science. The overall objective of the Citizen Science Earth Observation Lab (CSEOL) project is to generate, implement and validate a methodology to continuously feed new Citizen Science projects and ideas, engaging with a wide community of users of EO data and providers of information to foster the use of EO.

Group picture of the participants of CSEOL bootcamp. Credits: CSEOL.

About **30 participants** worked in teams to shape their proposed projects, guided by Citizen Science and Remote Sensing experts offering **one-to-one on-demand clinics** for each team in the two-days bootcamp.

During the **Pitch Day**, the nine projects teams delivered their pitch to a high-level judging panel for competitive funding. The winning CSEOL ideas were announced in the closing ceremony of 11 July.

CSEOL's Innovation methodology was based on three key phases – **Ideation, Greenhouse** and **Implementation**. The CSEOL bootcamp and Pitch day take place in the final part of the Greenhouse phase. Pilot projects selected and awarded during the Pitch day progress to the Implementation Phase, in which innovative ideas are developed and brought to concrete projects. The proposed projects entering the Implementation Phase to continue their experience in the CSEOL project were:

1. **Phenotandem: Harmonised phenology products from optical satellite and citizen science observations.** This project aims to create new phenology data combining in-situ observations and remote sensing products such as those provided by Sentinel-2. It will increase data quality for phenology at scale that has not been done before. The PhenoTandem team also participated in the Copernicus Roadshow and presented PhenoTandem in television programs in Spain.

(Left) Joan from Phenotandem in Frascati, (Center) A map showing the areas of interest to observe the flora, (Right) Phenotandem team observing the nature through CS. Credits: CSEOL.

2. **IceWatchApp: A mobile phone application for sea ice observations**. This project developed a data collection functionality, including attachment of photos from mobile devices, of its app for use by citizen scientists to map sea ice. These photographs will enhance satellite data from the Copernicus Sentinels and other missions where the lack of data from the Polar Regions hinders the development of automatic classification products. The IceWatch was the runner-up in the Nansen Poster Award, at the Arctic Frontiers Conference.

(Left) Alistair in Frascati, (Center): IceWatchApp – Sea ice observation from ice in motion cruise, (Right): IceWatchApp – data transfer completed from Univ Alaska to MET Norway. *Credits: CSEOL.*

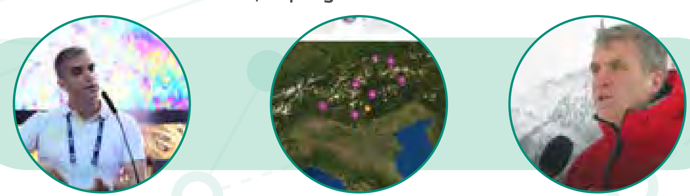
3. **Schools and satellites (SaS): a reliable rainfall product for West Africa.** This pilot project focused on the acquisition of ground and satellite data and interacted with citizen scientists on how to improve data collection as well as how to build the rain gauge from plastic bottles. It combined citizen science-based monitoring and a new satellite rainfall product which will use diverse sensors onboard Sentinel satellites. SaS also launched an initiative to collect smartphones that will be re-used by citizen scientists in Ghana.

(Left): SaS in Ghana testing the DISDROS, (Center): talking to the farmers learning about experiences and challenges, (Right): SaS visiting the High School. *Credits: CSEOL.*

4. **Sentinel Citizen: setting foundations for atmospheric commons.** The Sentinel Citizen pilot project contributes to enhancing air pollution forecasts through EO data in local communities in North Holland, with which they can influence policymakers to develop policy-oriented actions. For validation, data from the TROPOMI instrument and from the CAMS will be used. Sentinel Citizen delivered a project web page, presented the Sentinel Citizen project at different events and organised a workshop in the Netherlands.

(Left) Judith from Sentinel Citizen presenting in Frascati, (Right) Miha from Sentinel Citizen in Copernicus Roadshow. *Credits: CSEOL.*

5. **MySnowMaps: Dynamic mapping of snow conditions.** This pilot project displays convenient information in the form of maps of snow depth over large areas. It uses an innovative technology mixing physical models and satellite data and shares snow data for monitoring the water resources and improves the consciousness of the users. MySnowMaps has already developed the Android version of the MySnowMaps app which is now available in the Google Play Store. The team also organised a prize contest for engaging the citizen scientist community. MySnowMaps was featured on the Italian national television (RAI), a programme dedicated to innovation.



(Left): Matteo from Mysnowmaps in Frascati, (Center):Mysnowmaps – snow-meters measurements locations, (Right): Mysnowmaps being presented on TV. *Credits: CSEOL.*

SLUSH

SLUSH (slush.org) is a worldwide renowned innovation event that has start-up companies and investors as its major targets. The 2019 edition attracted some **25000 curious minds, including over 3500 start-ups and 2000 investors**.

In the last few years, in cooperation with the BICs (Business Incubation Centres), ESA Earth Observation has been present at SLUSH Helsinki with a stand. SLUSH is always an enthusiastic event, requiring a large team effort to prepare the stand, organise materials, experiences and demonstrations for the visitors. In 2019, the Φ-lab participated with its **Virtual Reality tour**, and also assisted visitors interactively, by presenting different ESA activities, results and answering questions.

It is always impressive to discover that the majority of participants (from all over the world) have little to no knowledge of ESA and its Earth Observation activities, so the stand also serves an important outreach and educational purpose. The collaboration with EAC allowed ESA astronaut Matthias Maurer to be present at the stand where he took the opportunity to go on a Virtual Reality Tour of the Moon. Some of the start-ups proposed ideas to be experimented by the astronaut in future missions (e.g. a glove interacting with a drone).

ESA stand at SLUSH Helsinki 2019. Astronaut Matthias Maurer experiencing Φ-lab VR with help of Paulo Sacramento.
Credits: ESA.

Right and left: General Views of the stand.

Emmanouil Lagoudakis, Paola Berretta and Paulo Sacramento at the interactive touch table.

Matthias Maurer and Tessa Nikander.

MAKER FAIRE

Maker Faire is a worldwide innovation event, originating in 2006 in the San Francisco Bay Area as a project of the editors of Make: magazine (makezine.com). Nowadays, it is organised on a larger scale in many cities around the world, including Rome, Berlin, Paris, Tokyo, and Shenzhen.

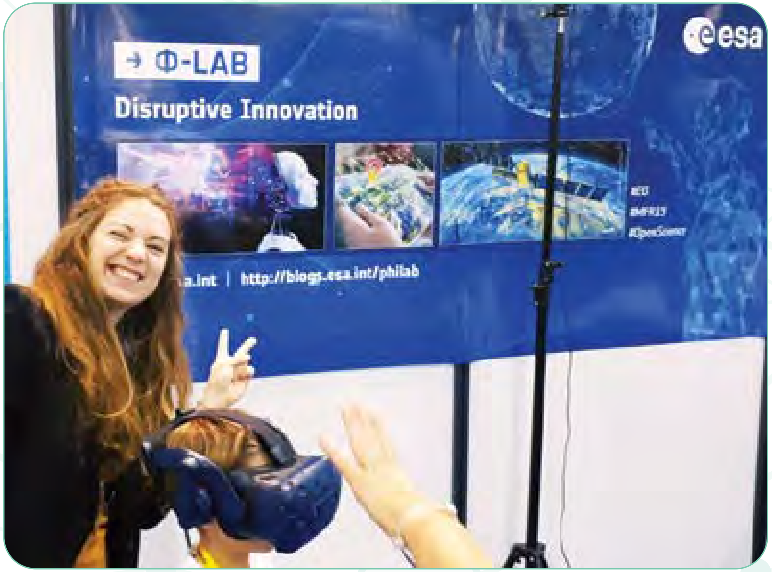
Maker Faire combines science, science fiction, technology, entertainment and business. It caters to curious participants of all ages wishing to experience first-hand the makers' fun and educational inventions. These inventions are the result of a desire to solve everyday problems, whether big or small. Maker Faire is not just a fair for field experts, there are inventions in the **fields of science and technology** (from 3D printers to wearables, through to drones, robots and digital manufacturing), and also new forms of art, entertainment, crafts, food experiments, etc.

Maker Faire Rome is organised by InnovaCamera, a Special Agency of the Rome Chamber of Commerce. **It is the most important Maker Faire in Europe** and the third worldwide after the "Bay Area" and "New York" ones. The Maker Faire was held in Rome for the first time in 2013. Since 2016 it has been hosted at Fiera di Roma, in order to accommodate the large number of visitors.

ESA and the Φ-lab were present for the first time at Maker Faire Rome in October 2019. More than **100000 visitors** attended over the three day event, with more than **27000 students present** on Friday morning, which is reserved for schools. There were a total of about 800 exhibitors/ stands, spread among seven pavilions of the Fiera di Roma.

ESA stand at Maker Faire 2019 with the Φ-lab VR set. Credits: ESA.

The ESA and Φ-lab stand was made possible through a partnership with the **La Sapienza University in Rome**, as part of the **CINI** (an Italian inter-university Consortium for Informatics) and particularly the **AIIS** lab (Italian national laboratory for AI and Intelligent Systems). Interactive tools, such as the **Φ-lab Virtual Reality (VR)** installation helped to attract hundreds of visitors to the stand, favouring **networking** and the establishment of dozens of contacts. Visitors were also provided a selection of ESA promotional materials (gadgets, bags, lanyards, posters, flyers of InCubed, EO, Copernicus, ESA mobile Apps.



AI FOR GOOD

AI 4 GOOD was organised on the 28–31 May 2019, in Geneva, Switzerland by ITU (itu.int) in partnership with sister UN agencies, the X-Prize Foundation (xprize.org), and the Association for Computing Machinery (acm.org). The "AI for Space" session hosted **inspiring talks** from top industry professionals, focusing on areas where AI techniques can be applied to space datasets in order to accelerate progress towards the **Sustainable Development Goals** (SDGs). It also covered the processes, infrastructure and ethical considerations when using this powerful technology to protect our planet and generate benefit for all humankind. The session was joined by Pierre-Philippe Mathieu, Head of Φ-lab Explore office, who presented **cutting edge use-cases,** methodologies and emerging best practices to delegates, including the fusion of ESA data sets to create enhanced products and ways of working with partners to build consortia to tackle complex challenges such, as the SGDs, using space technology. The discussion placed emphasis on the potential for AI and space to unlock a new era of planetary stewardship – leveraging AI's unprecedented capacities for prediction and rapid understanding of complexity on both a local and global scale. The session involved evaluations on the current state of play in AI and Space, barriers to deployment and real-time insight and equitable access to data. It identified projects related to SDGs where AI and Space fields can be combined to deliver a positive benefit to humanity and discussed opportunities for interdisciplinary cooperation to move forward in areas such as data accessibility, trust, algorithmic inequity, and accountability.

Participants at AI 4 Good Global Summit. Credits. AI 4 Good.

AI4EO CHALLENGE WITH FRONTIER DEVELOPMENT LAB (FDL)

The **Frontier Development Lab** (FDL) (frontierdevelopmentlab.org/) is an AI accelerator scheme based on a 8-week intensive **research sprint** addressing global challenges of our planet with AI and data. The initiative originated with NASA a few years ago and has now been extended with ESA, to Europe. FDL is an interdisciplinary PhD-postdoctoral level research programme that applies AI technologies to science. The scope is to push the frontiers of research and develop new tools to help solve some of the biggest challenges of our planet. These range from the effects of climate change to predicting space weather, from improving disaster response, to identifying meteorites which could hold the key to the history of our universe.

In particular, FDL Europe (fdleurope.org) is demonstrating the potential of ML for both scientific progress and discovery as well as new paradigms of engineering, from putting AI onboard spacecraft, to helping identify informal settlements. The central value is that subject experts - with deep knowledge of the problem domain - can develop AI enhanced workflows and solutions with peers from the data sciences. The challenges have been coached by EO and AI specialists.

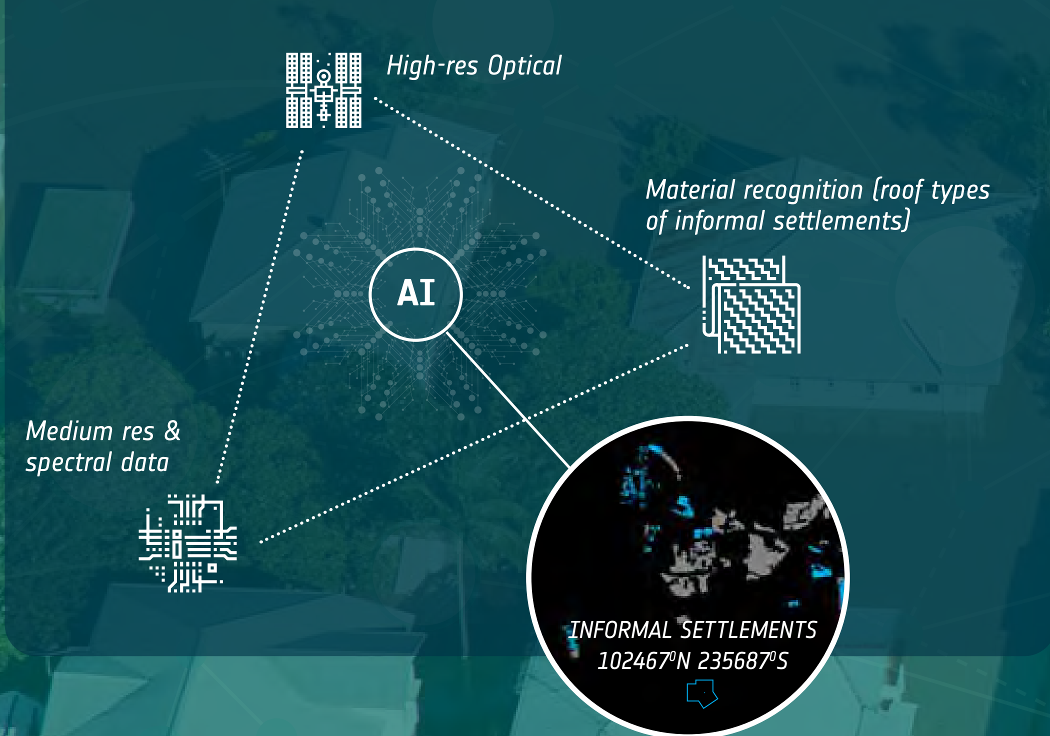
In 2018, FDL Europe in partnership with the European Space Agency, hosted a research sprint at the University of Oxford with a kick-off at Φ-lab premises. With the support of NVIDIA and the Satellite Applications catapults, two teams of researchers tackled challenges under the mission area **"Mission Control for Planet Earth"**. The research sprint was a great success and the outputs showed the great potential of Mission Control for Planet Earth. One team of researchers focused on identification and of mapping **informal settlements** to enable governments and aid agencies to better support these communities and a second team focused on the development of an AI technique, **Multi3Net**, to drastically reduce the time and effort required to create effective disaster impact maps to aid disaster response.

FDL Europe participants in ESRIN (ESA's Center for Earth Observation).
Credits: FDL Europe

In 2019 the challenges included on-board ML for Disaster Response (Floods) and Atmospheric Phenomena for climate variability. The Floods team produced numerous innovations in onboard ML including a whole earth-scale pre-processed and labelled training set for flood segmentation called **WorldFloods**. The Climate team produced a valuable data product for the climate change community. **CUMULO**, a legacy of this project consists of a benchmark cloud data set for training and evaluating cloud classifiers at high spatial and temporal resolution.

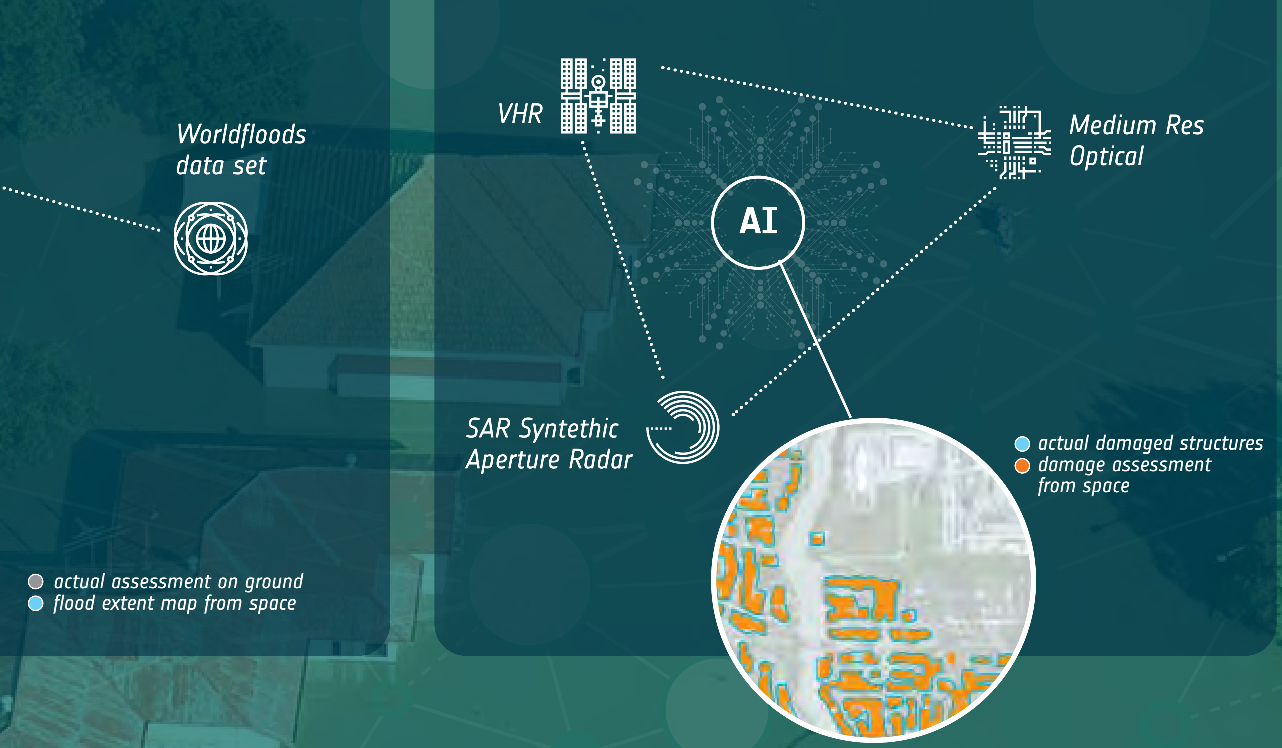
AI4EO CHALLENGE WITH FRONTIER DEVELOPMENT LAB (FDL)

Using roof material to map the world's informal settlements at continental scale and low cost

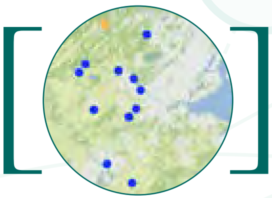


"WorldFloods" dataset trains a global flood detector, running on-board a low-cost satellite

“Multi3Net” ML can quickly and accurately assess damage after a flood or an earthquake



Resilience planning | Disaster Response | Rebuild & Recover

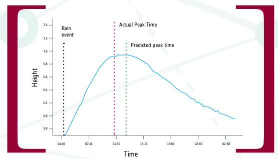


1. Flood Prediction

Data fusion:
- Satellite
- Stream gauge (depth of flood)
- Known rainfall (radar)

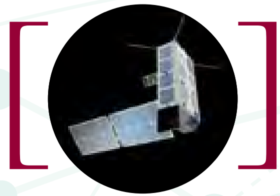
2. Informal Settlement Detection

Data fusion:
- Very High Res (optical)
- Roof types of informal settlements
- Medium res + spectral data



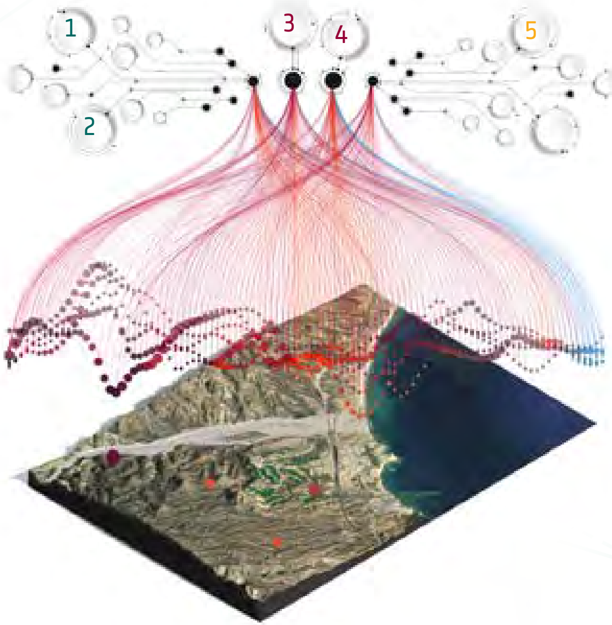
3. Inundation Warning

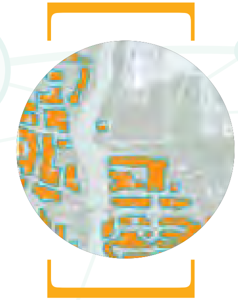
Data fusion:
- Satellite
- Stream gauge (rate)
- Known rainfall (radar)



4. Rapid Extent Mapping

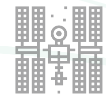
Data fusion:
- ‘Worldfloods’ training data set
- Medium Optical Resolution





4. Damage Assessment

Data fusion:
- VHR (expensive)
- Medium res optical (free)
  - SAR Synthetic Aperture Radar (low temporal)



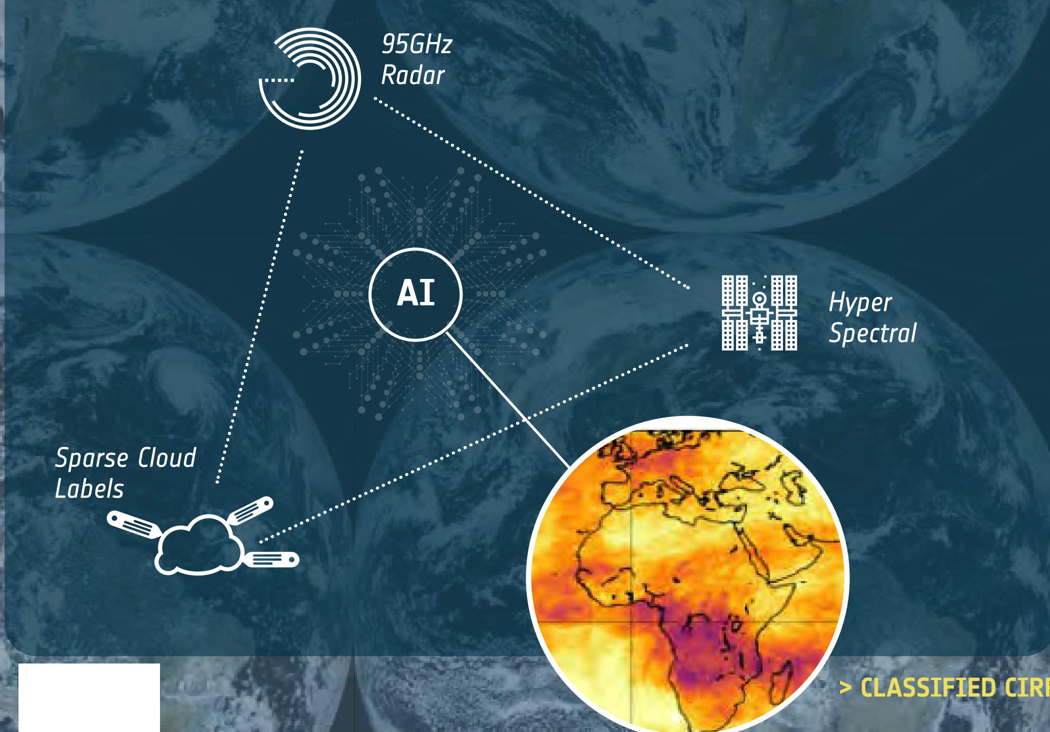
High res visual

Ground Truth (2 weeks)

Data Fusion (with AI) (2 minutes)

AI Capabilities [Flooding]

"CUMULO" dataset trains a global cloud classifier



> CLASSIFIED CIRRUS CLOUDS

Best paper award for the NeurIPS 2019 Workshop on Tackling Climate Change with Machine Learning

AI4EO CHALLENGE WITH Φ-WEEK BOOTCAMP

The Φ-week Bootcamp is hosted by the European Space Agency (ESA) in the framework of Φ-week. It consists of an intensive five-day design 'sprint' process to prototype a solution geared at solving a big industry 'challenge' using Earth observation data. 40 pre-selected participants from all around the world work in teams, contributing with ideas and skills to respond to a specific industry challenge. The solutions are tested with a set of users and then presented to a jury that selects the prototype with the best potential for commercialisation.

Using a workshop-style format, the event is designed to be hands-on and interactive, allowing participants to test the prototypes on real-life users before presenting them on the final day.

The sprint is a 5-day progress for answering critical business questions through design, prototyping, and testing of ideas with customers. It covers business strategy, innovation, behaviour science, design thinking and more – packed into a battle-tested process that any team can use.

In 2018 the winning team, AMIGROW, presented a concept aimed at integrating field information about crop health and management practices with satellite imagery to provide farmers with valuable insights on crop production management in order to maximise crop yield.

Φ-week Bootcamp winning teams (top) PowerPatrol 2019 winning team, (bottom) AMIGROW 2018 winning team. *Credits: PI School*

Presentation by Agnieszka Lukaszczyk of Planet on Space renaissance.

In 2019 the **Φ-week** Bootcamp challenge was set by the energy service producer and provider EDP, through its innovation arm, EDP Inovação (EDPI). The challenge concerned the prevention, detection and mitigation of the risks of wildfires relating to bushes and vegetation, around important energy infrastructures in Portugal. The first prize was won by the team PowerPatrol, which developed an end-to-end demonstrative platform which allows the user to visualise their entire asset network, specify an area of interest, and observe the predicted risk of vegetation-related wildfire and the recommended actions based on this risk. The team won a travel voucher to a space/startup event of their choice and a 3-month licence to Planet's API to further develop their prototype.

Φ-week Bootcamp group pictures. *Credits: PI School*

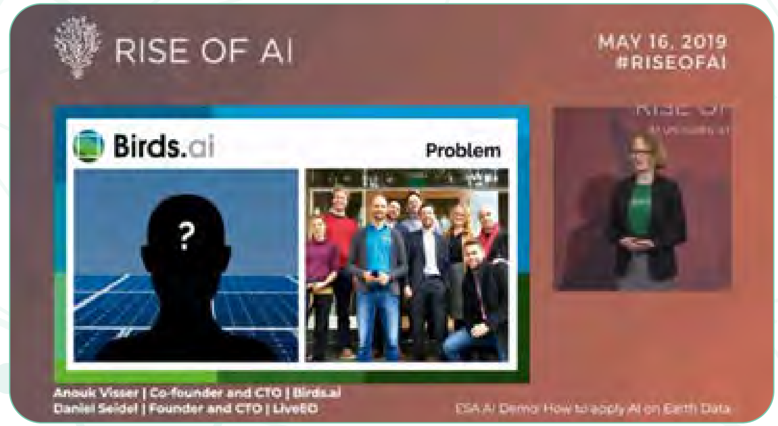
STARTUP PITCH @ RISE OF AI

The **Rise of AI** (riseof.ai) event is one of the leading global conference for AI. In Berlin, 2019 Rise of AI hosted over 800 investors, CEOs, researchers, politicians, entrepreneurs and enterprises.

ESA participated with the **AI demo: how to apply AI on Earth data**. This demo, supported by the **Φ-lab**, focused on how to get insights out of satellite imagery with AI. The demo allowed selected AI start-ups seeking to exploit the opportunities brought by this disruptive domain, to present themselves through a set of talks.

LiveEO mastered the challenge of handling big data and showed how it is disrupting the utility sector. LiveEO uses satellite imagery to provide innovative **infrastructure monitoring** in the verticals railway, electricity and pipelines networks. It provides a solution which uses satellite data to identify risks to operations from vegetation, height changes and third party interaction along networks, enabling operators to work more efficiently.

Another startup, **Bird.ai** showed how they are giving customers insight into the degradation of their assets so they can proactively perform maintenance in order to keep them in top-condition. Their toolset contains artificial intelligence, machine learning, and predictive analytics. With this, Birds.ai creates intelligent analytical software for anomaly checks, functional statistics, and intelligent monitoring.



Start-up pitching at the Rise of AI (top) LiveEO, (bottom) Bird.ai. Credits: Rise of AI

AI4EO CHALLENGE WITH PI SCHOOL

The **Pi School** (picampus-school.com) is a **next-generation** school focused on building a better future through technology and creativity. The Φ-lab brought two EO challenges to Pi school to be developed over the course of eight weeks. The projects exploited free and open data from Europe's Copernicus satellites, whose data is useful for a range of applications in various domains, from climate change to urban planning. Both projects demonstrated new aspects of the use of AI in terms of **climate change and sustainability.**

The first challenge addressed the agriculture sector and aimed at developing a proof of concept that will help monitor the crop fields all over Europe. The proof of concept, **"SmartCrop"**, was the result of work carried out during the eight-week programme. It aims at distinguishing different types of field crops contains (soybean, wheat, corn, etc.) using satellite images and deep learning techniques. Crop classification served as a starting point for several use cases related to improving the current state of the art in the field of agriculture by using AI. Adding space-based observation data to traditional crop-forecasting models enables more accurate harvest predictions and helps farmers to monitor the health of their crops and plan deployment of fertilisers.

Another challenge, also sponsored by Φ-lab was to explore the available solutions that will enable countries to to monitor emissions and improve adaptation and mitigation efforts – through **air pollution forecasting**. This eight-week challenge aimed at predicting levels of air pollution and air quality, through mathematical simulations of how pollutants dispersed in the air. This downscaling air quality forecast used Copernicus Atmosphere Monitoring Service (CAMS), and Sentinel-5P, to create a weather-forecasting model developed with deep-learning techniques.

Jamila Mifdal and Teodora Selea at Φ-lab developing the eight-week challenge and presenting the proof of concept that will help monitor the crop fields all over Europe. *Credits: PI School.*

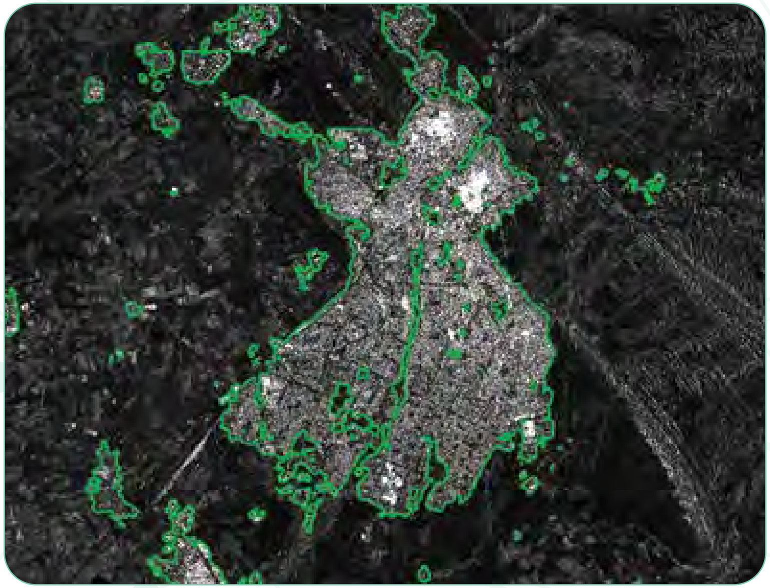
Luka Sachsse and Maximilien Houël teamed up for eight-weeks at Φ-lab and presented the enhancement of weather-forecasting model working at based on EO data and deep-learning techniques. *Credits: PI School.*

AI4EO CHALLENGE WITH UNOSAT

This challenge was organised by Phi-Unet (phi-unet.com) in partnership with UNOSAT (unitar.org/unosat) a technology-intensive programme under the United Nations Institute for Training and Research. The aim of the contest, instigated by UNOSAT in partnership with RUS Copernicus and with the technical support of CERN openlab (home.cern/science/computing/cern-openlab), was to put artificial intelligence and Earth Observation data at the service of a **humanitarian cause**: supporting the Iraqi government in planning reconstruction activities.

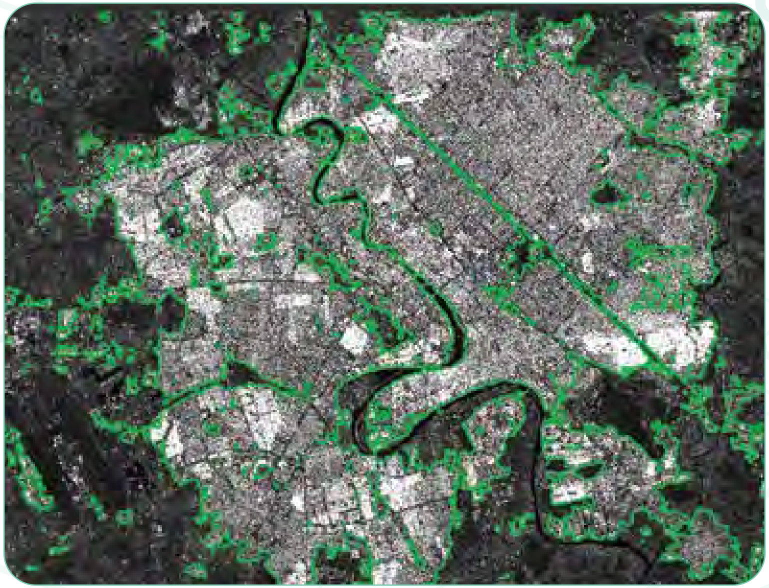
This challenge is focused on the creation and generation of the **building footprints in Iraq**. The building footprint request comes out of a need from the UN Populations Fund (UNFPA). UNFPA is the United Nations sexual and reproductive health agency. They are assisting the Government of Iraq to plan the October 2020 population census, which is crucial for key baseline information in support of reconstruction and development (i.e. fight extreme poverty, inequality and sexual and reproductive health problems, prevention of gender-based violence, climate change resilience). The building footprints are needed to plan the implementation of the on-site survey interviews – and the contest has two phases.

During phase one in late 2019, around 70 AI/EO enthusiasts worked on **urban areas detection** (8 cities in Iraq). Copernicus Sentinel-1 imagery data and Ground Truth data (provided by DLR) were provided as training and evaluation datasets. Their most updated product, the World Settlement Footprint (WSF) 2015 (obtained by combining multi temporal Landsat-8 [30 m] and Sentinel-1 [10 m] imagery, with a final product generated at 10 m resolution), refers to the year 2015. European participants had the possibility to benefit from the services of the Research and User Support (RUS) Service (i.e. Expert Support and Virtual Machines with 4 cores and 16 Go RAM), or use their own computing environment to develop their workflow. The second phase (still ongoing) provides VHR 3-band natural color images for three cities, plus Open Street Map (OSM) data with a total of 722,837 building polygons.



Satellite image of Kirkouk city in VV polarisation with model predicted binary mask, building objects vs non buildings (green contour). Credits: Andrey Malakhov and Alessandro Patruno (Team Zephyros)

In order to process the images provided (with two polarisations - VV-VH) and to perform a semantic segmentation, some participants used a U-Net (fully convolutional network) architecture, with pre-trained VGG11 encoder or Resnet34 backbone and center dilation layer. These results have been evaluated using the F-1 score (a weighted average of the precision and recall). Phase 2 results will be released in early April.



Satellite image of Bagdad city in VV polarisation with model predicted binary mask, building objects vs non buildings (green contour). Credits: Andrey Malakhov and Alessandro Patruno (Team Zephyros)

Predictions overlaid with Bing imagery in QGIS (from top to bottom: Bagdad, Kirkouk, Samawah, Tikrit). Credits: Tomasz Dyczek

AI4EO CHALLENGE WITH BIG DIVE

BIG DIVE is an initiative (bigdive.eu), organised by TOP-IX, ISI Foundation, AXANT offering private and public training on Data Science, Machine and Deep Learning, Data Visualisation, and Data Engineering. The BIG DIVE platform presents itself as an interactive " street-fighting gym" which puts the raw material - in the form of high value datasets - into the hands of "ambitious smart geeks" tutored and mentored by experts in three key areas: Development, Visualization and Data Science. Courses include lectures by experts in the field and the latest resources and technologies.

In 2019, the focus for BIG DIVE 8 (bigdive.eu/bigdive8) was space data and satellite images.

ESA as one of the main data sponsors collaborated from the beginning to decide the datasets for two of the final projects, ensure the challenges were up to date and relevant and prepare the class to tackle the real-world problems in the space sector and machine learning in general. The three-week training covered the necessary programming and data skills, the best approaches to predictive models, and lectures in the space industry by ESA, Target Detection, Starlab, and forefront research centres. The final week the students, mentored by the teachers, worked and presented four final projects on diverse fields such as urbanization, polar regions, agriculture, atomic clocks and space weather.

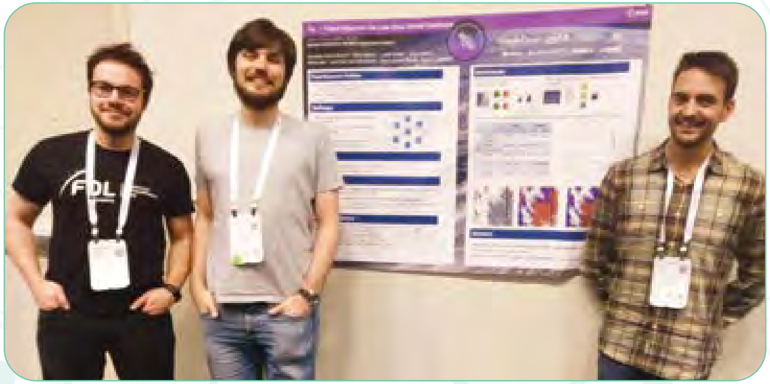
Group picture of BIG DIVE participants. Credits: BIG DIVE.

One of the projects was the set up of an image recognition challenge with sea ice images a dataset of high-resolution Copernicus Sentinel-1 satellite images on a selected polar area. An ML system for the detection of ice presence and concentration was developed in order to automatically produce an ice map. The group explored the concentration of ice using the ASIP sea ice data set of 26 files in NetCDF format. From that data they extracted matrices and transformed them in order to feed the ML model. The trained a U-net model and evaluate the quality of the learning by using the Jaccard coefficient. Normally, the interpretation of the ice charts is done manually, therefore the project was an important attempt in making the process automatically and faster to allow in the future a scalable approach.

BIG DIVE participants working on the challenge. Credits: BIG DIVE.

WORKSHOP ON AI4EO NEURIPS

The **FDL (Frontier Development Lab) Europe** (fdleurope.org) presented the work done at the Neural Information Processing Systems (NeurIPS) annual conference. The purpose of the conference is to foster the exchange of research on neural information processing systems in their biological, technological, mathematical and theoretical aspects. The core focus is peer-reviewed novel research which is presented and discussed in the general session, along with invited talks by leaders in their field. FDL Researcher Valentina Zantedeschi presented Cumulo, a breakthrough dataset and method of fusing radar and image data for improved cloud classification. This was also awarded the best paper in the Climate Change research workshop, which is a fantastic achievement. Josh Veitch-Michaelis, FDL Researcher for the Disaster Response team, presented the Flood Detection on low cost orbital Software at the AI & HADR (Artificial Intelligence for Humanitarian Assistance and Disaster Response) Workshop. The mission support challenge team were also accepted to present their work at the Machine Learning competitions for all workshops.



The FDL (Frontier Development Lab) Europe participants presenting the work done at the NeurIPS annual conference. Credits: FDL Europe.

WORKSHOP ON QUANTUM COMPUTING

This event was conceived during a mini-workshop on **Quantum Computing for Earth Observation** which was held on 15 November 2018 at the ESA Φ-Week in ESA's ESRIN establishment in Frascati, Italy. At this event representatives from the quantum computing community, from both academia and industry, met with Earth Observation practitioners. The objective was to explore possible synergies between the two technologies to stimulate their further development and to accelerate their impact for societal benefit. The focus of the workshop was on the application of quantum computing for downstream data processing and Earth observation data exploitation. The workshop aimed to prepare the ground for the opportunities that will be presented when the quantum community will be able to produce software for quantum-enhanced optimisation problems of direct use in big data management. Together with machine learning, quantum computing has the potential to be a game-changer in data science and applications.

Illustration of a quantum computer.Credits: Gettyimages.

MASSIVE OPEN ONLINE COURSE (MOOC) ON DISRUPTIVE TECH

The **Earth Observation – Disruptive Technology and New Space** (imperativemoocs.com/courses/disruptingeo) is a mini **MOOC from ESA.** It consists of a series of interviews with leading experts across Earth Observation and related technologies. The explosion in EO data from the Sentinel programme, a new generation of commercial satellites, and emerging constellations of small-sats, has created one of the greatest "big data" challenges in the world today. This course explores technologies such as AI, 3D data visualisation, cloud computing technologies and blockchain, and how they are meeting the needs of the ever-growing data analytics and data navigation challenges in EO.

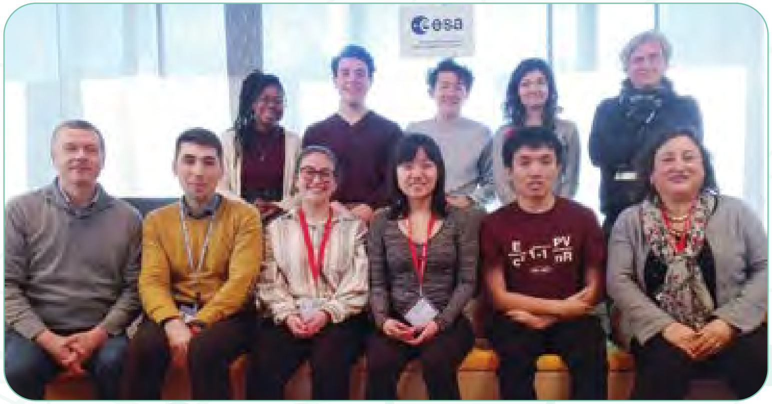
The course is composed of four modules: 1) AI, Big Data Analytics and Data Visualisation, 2) The New "Internet of Data", 3) EO – What comes next… and 4) Responding to Digital Trends.

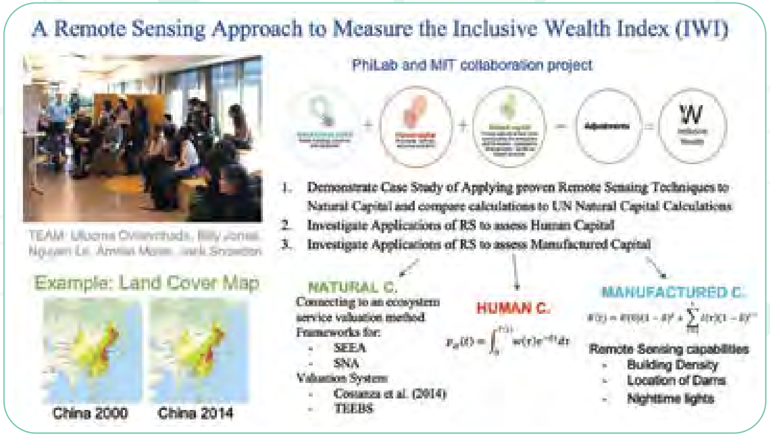
Participants of MOOC: Adrien Muller, Satellite Applications Catapultm*; Silvia Ullo,* University of Sannio*; Alexis Smith,* IMGeospatial*; Philip Briscoe,* Rezatec*; Steffen Fritz,* IIASA*; Adina Gillespie,* GHGSat. *Credits: Imperative Space.*

COLLABORATION WITH MIT

A group of students from **MIT (Massachusetts Institute of Technology)** of Computer Science and Economics and Engineering partnered with the Telecommunications group of the **University of Sannio** and the Φ-lab to work together on research cases. These students carried out specific projects on cutting-edge issues compared to the state of the art at an international level in the field of engineering, economics, law and biology. The **MIT and Unisannio** students collaborated on specific projects with Professors from the University of Sannio, as well as with other students, postdoctoral research fellows and PhD candidates to carry out advanced projects on remote sensing and artificial intelligence. All these projects aimed to improve human health and security, ranging from infrastructures and geohazard monitoring, to countries development measurements.

The idea behind the project was to **link** students of MIT and University of Sannio, through the **commonality of studies that binds humanity across its borders**. Three main projects were conducted over a one month period. One of the research activities was to demonstrate a case study of applying proven remote sensing techniques to Natural Capital and compare calculations to UN Natural Capital Calculations, followed by investigations on applications of remote sensing to assess Human Capital and Manufactured Capital. Another research project aimed at classifying landslides with Copernicus Sentinel-1 and Sentinel-2 data, and a third research project focused on utilising Sentinel-1 images for infrastructure monitoring.





Group photo of MIT student visit to Φ-lab.
Workflow of MIT and Φ-lab collaboration project.
Credits: ESA.

“Working with the Φ-lab team has enabled us to make our ideas applicable to real world situations and to help us mainstream usage of drones in the mapping of agricultural livelihoods.”

Rogerio Bonifacio
Senior Climate and Earth Observation Advisor Research, UN World Food Programme

“R&I activities between SatCen RTDI and ESA Φ-lab teams were fundamental to understand how to apply AI to monitor critical infrastructures in an operational way.”

Michele Lazzarini
RTDI Project Manager, European Union Satellite Centre

“Gravity force does not keep you on the ground at ESA’s VR experience. It was a fantastic way to interact with EO data using VR technology. Great project!”

Rochelle Schneider
Research Fellow in Geospatial Data Science, LSHTM

“VR will be a key technology in the preparation of space missions, the dissemination of their results and to enable and engage the public to “experience space” with their own eyes.”

Matthias Maurer
Astronaut, ESA

Φ-lab fosters novel data endeavours and tries to reach beyond classic remote sensing, a place that allows ideas to grow.

Michael Schultz
Researcher, Heidelberg University

Pi School collaborated with ESA and the Φ-lab team to develop two open-source AI+EO projects this summer. This worked beyond expectations. Collaborating with them is invigorating, smooth, and always very productive for both parts. We're in AI, and we think that Φ-lab is looking at just the right problems on the interface between Earth Observation and AI these days.

Sébastien Bratières
Managing Director, PI School

Innovation at UNICEF is about doing new things to solve problems and improve the lives of children around the world. Working with Phi-lab at ESA — using data and new technologies like Earth Observation and AI — allows us to match today's challenges with tomorrow's solutions. The team at Phi-lab plays a pivotal role — connecting different actors from the international development and research fields; facilitating creative collaborations; and delivering critical insights into the environmental conditions, challenges and opportunities facing children and young people around the world.

Do-Hyung Kim
Space Tech Lead, UNICEF

The ELLIS-ESA collaboration will make top-notch research on Earth sciences a reality.

Gustau Camps-Valls
Researcher & Professor, ELLIS

PUBLICATIONS

This section highlights a selection of publications resulting from of the Φ-lab research and pilot projects with academia and industry.

Nalepa, J., et al., (2019), **Unsupervised segmentation of hyperspectral images using 3-D convolutional autoencoders** – IEEE Geoscience and Remote Sensing Letters, pp. 1-5, DOI: 10.1109/LGRS.2019.2960945.

Nalepa, J., et al., (2019), **Transfer learning for segmenting dimensionally reduced hyperspectral images** – IEEE Geoscience and Remote Sensing Letters, pp. 1-5, DOI:10.1109/LGRS.2019.2942832.

Nalepa, J., et al., (2020), **Training and test-time data augmentation for hyperspectral image segmentation** – IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 2, pp. 292-296.

Nalepa, J., et al., (2019), **Validating hyperspectral image segmentation** – IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 8, pp. 1264-1268.

Ribalta Lorenzo, P., et al., (2020), **Hyperspectral band selection using attention-based convolutional neural networks** – IEEE Access, vol. 8, pp. 42384-42403.

Peressutti, D., et al., (2019), **Crowd-sourcing the open dataset for water detection** – Medium.

Lubej, M., et al., (2018), **Land Cover Classification with eo-learn** – Medium.

Bollinger, D., et al., (2019), **Use eo-learn with AWS SageMaker** – Medium.

Bollinger, D., et al., (2019), **Tracking a Rapidly Changing Planet** – Medium.

Lubej, M., et al., (2019), **QueryPlanet: Machine Learning with EO big data at scale** – Living Planet Symposium.

Zupanc, A., et al., (2019), **Blue Dot Water Observatory** – Living Planet Symposium.

Rudner, T., et al., (2019), **Multi3Net: Segmenting Flooded Buildings via Fusion of Multiresolution, Multisensor, and Multitemporal Satellite Imagery** – AAAI Conference on Artificial Intelligence.

Gram-Hansen, B., et al., (2019), **Mapping Informal Settlements in Developing Countries using Machine Learning and Low-Resolution Multispectral Data** – AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society.

Rudner, T., et al., (2019), **Rapid Computer Vision-aided Disaster Response via Fusion of Multiresolution, Multisensor, and Multitemporal Satellite Imagery** – Conference on Neural Information Processing Systems (NeurIPS).

Gram-Hansen, B., et al., (2018), **Generating Material Maps to Map Informal Settlements** – Conference on Neural Information Processing Systems (NeurIPS).

Fil, J., et al, (2018), **Multi3Net: Segmenting Flooded Buildings via Fusion of Multiresolution, Multisensor and Multitemporal Satellite Imagery, Conference on Neural Information Processing Systems (NeurIPS)** – Modeling and Decision-Making in the Spatiotemporal Domain Workshop.

Helber, P. et al., (2019), **Mapping Informal Settlements in Developing Countries with Multi-resolution, Multi-spectral Data, International Conference on Learning Representations (ICLR)** – AI for Social Good Workshop.


Artificial Intelligence
Chatbot
Earth Observation
Space
Digital Twin Earth
Neural Network
Autonomy
Supervised Learning
Data Science
Deep Learning
Artificial Intelligence
Space
Earth Observation
Neural Network
Space
Chatbot
Segmentation
Machine Learning
Deep Learning
Data Science
Clustering
Autonomy
Segmentation
Data Mining
Digital Twin Earth
Machine Learning
Transfer Learning
Reinforcement Learning
On board processing
Deep Learning
Clustering
Space
Autonomy
Space
Space
Clustering
Space Transfer Learning
Chatbot
Neural Network
Data Mining
Reinforcement Learning
Autonomy
Earth Observation
Chatbot
Machine Learning
Recurrent Neural Network

TERMINOLOGY

Earth Observation (EO) is the gathering of information about our planet's physical, chemical and biological systems. Within this document, EO refers to the measurements specifically made from space via satellites. EO remote sensing data are used to monitor and assess the state and changes in our environment. EO satellites are providing unprecedented volumes of data and Artificial Intelligence (AI) techniques offer the potential to fully exploit this Big Data resource to extract relevant information for science and society.

Artificial Intelligence (AI) is intelligence exhibited by machines that can observe, perceive and act upon their environment to maximise their performance over a variety of tasks. It refers to the capacity of an algorithm for assimilating information to perform tasks that are characteristic of human intelligence, such as recognising objects and sounds, contextualising language, learning from the environment, and problem solving. Within this report, the term "AI" will therefore be used mainly as a generic term to refer to Machine Learning adapted to work with geospatial data. The term "AI4EO" will refer to the mining of EO data with AI techniques.

- ***Machine Learning (ML)*** is a branch of AI relying on algorithms that are capable of learning from both data and through human interactions (e.g. supervision) to enable prediction, and are also used for data mining (i.e. discovery of unknown properties and patterns). ML is a field of statistical research for training computational algorithms that split, sort, transform a set of data in order to maximise the ability to classify, predict, cluster or discover new patterns in target datasets. ML is all about using computers to learn how to deal with problems without programming. In fact, ML generates models by taking some data for training a model, and then makes predictions. ML relies on a wide variety of algorithms (supervised and unsupervised), ranging from simple Symbolic Regression, Neural Network, decision tree, Support Vector Machine, up to genetic programming and ensemble methods such as random forest.
- ***Deep Learning (DL)*** is a type of ML algorithm that aims to solve the same kind of problems by mimicking the biological structure of the brain and construct hierarchical architectures of increasing sophistication. There is a wide variety of network architectures including Convolutional Neural Networks (CNN) (e.g. GoogleNet, Res-Net, YOLO), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, Generative Adversarial Networks (GAN), Deep Belief Networks, and stacked auto-encoders. Today, DL is reaching high-level accuracy going beyond human performance, with the potential to substitute handcrafted feature extraction, thereby enabling totally automatic image recognition of big data (including EO) and opening huge opportunities for new science and business.
- ***Reinforcement Learning (RL)*** is an area of ML concerned with how software agents ought to take actions in an environment in order to maximise some notion of cumulative reward. e.g. learning from "mistakes".

Neural Networks (NN) are a set of algorithms modeling loosely the human brain connections, that are designed to recognise patterns. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. There are a lot of variety of NN architectures including for example:

- ***Convolutional Neural Networks (CNN)*** is a class of Deep Learning networks (and one of the most influential innovations in the field of computer vision), most commonly applied to analysing visual imagery. A CNN takes the input (a tensor with a shape depending on number of images, width and height and number of bands), and passes it through a convolutional layer (instead of the more traditional matrix multiplications) abstracting into a feature map. A CNN is able to successfully capture the spatial and temporal dependencies in an image through the application of relevant filters - and are an ideal tool for processing regularly sampled data (such as is the case of satellite data).
- ***Recurrent Neural Networks (RNN)*** is a class of neural networks where the connections between nodes from a directed graph run along a temporal sequence, allowing to exhibit temporal dynamic behaviour. As such, RNNs can use their internal state (memory) to process variable length sequences of inputs, and is a powerful technique to enhance temporal information.
- ***Generative Adversarial Networks (GANs)*** are algorithm architectures that use two neural networks, that compete with each other (thus the term "adversarial"). Given a training set, this technique learns to generate new data with the same statistics as the training set. GANs enable the generation of new, synthetic instances of data that look like real data. They are used widely in image generation, video generation and voice generation.

Big Data Analytics is a suite of analysis techniques aiming to deliver "value" from big datasets, whose Volume, Velocity, Variety, Veracity is beyond the ability of traditional tools to capture, store, manage and analyse. Within this report, the word Big Data analytics will mainly address ML.

Crowdsourcing (CS) is the practice of public participation and collaboration in a common goal. Within this report, the term will also be used as a synonym for "Citizen Science" when the goal is to do research. Citizen scientists can help in processing/analysing EO data (e.g. visual interpretation of land cover and identification of other features visible from VHR images) but also in generating new observations (e.g. air quality measurements using a mobile phone – ispex.nl – or a variety of new mobile apps for recording observations on the ground) for a myriad of applications, ranging from land cover validation to animal tracking to humanitarian response. There are also many emerging synergies between AI and CS, in both directions, where AI can help analyse the data of citizens, and where citizens can train the AI through generation of data sets (e.g. labelled observations).

Computer Vision (CV) is a field concerned with the automatic extraction, analysis, and understanding of useful information from a single image or sequence of images (e.g. videos). It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding.

Inference is the application of pre-trained Machine Learning algorithms on newly sensed and real-world data. The inference is the result of a trained neural network making predictions based on new data input.

Transfer Learning is a ML method focused on storing knowledge gained while solving one problem and applying it to a different but related problem. For example, knowledge gained while learning to recognise a type of feature that could apply when trying to recognise other similar or related features. Practically in this method a model developed for a task is reused as the starting point for a model on a second task.

Transformative EO Technologies also called radical, deep or disruptive technologies refer in this document to technologies that can make a big impact on the EO sector, by helping to address big societal and environmental challenges in a new way with EO data, by shaping the future of new EO "smart" and "connected" satellites, but also by having the power to create their own markets or disrupt existing industries. In 2019, Gartner (world leader in innovation) identified a suite of top-10 transformative technologies. It is worth noting that these techs cannot be addressed in isolation as they in fact highly interconnected and together shape and accelerate the changes.

Digital Twin is a virtual representation of a physical asset enabled through data and simulations for real-time prediction, monitoring, control and optimisation of the asset for improved decision-making throughout the life-cycle of the asset and beyond [Rasdeed et al., 2019].

Digital Twin Earth (DTE) is an interactive "digital replica" of the entire planet that can facilitate a shared understanding of the multiple relationships between the physical and natural environments and society. DTE enables scientists and users to quantify past, present and future changes on our planet, by integrating data from models and observations and technologies such as AI to advance our understanding of the impact of human activities on our global environment and society.

ACKNOWLEDGEMENTS
We express our sincere gratitude to the lab fellows and team and to EOGB (Federica Mezzo and Giorgia Marino in particular) for helping with amazing graphics and our communication team (Julie Oakley and Gaia Maggi in particular) for comments and proof reading and to Sara Aparicio for supporting the coordination. Thanks for their patience, motivation, enthusiasm, and knowledge.

How to reference this publication:
Mathieu, PP. et al., (2020), Research & Innovation activities and strategy of the philab Explore Office, ESA

The man who has no imagination has no wings.

Muhammad Ali
(1942-2016)

www.esa.int/eo **phi-lab@esa.int**