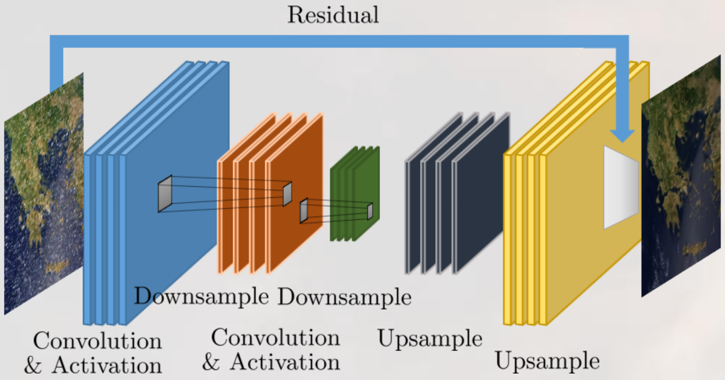
TITAN updates

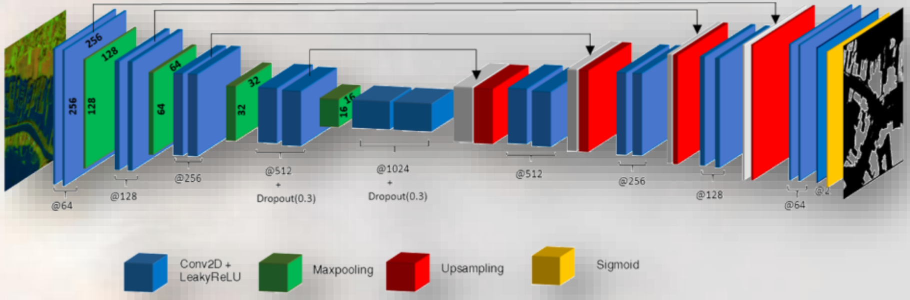
- Tensor-based recovery
- TITAN Datathon
- AOB

Paradigms in ML/DL

Generative models

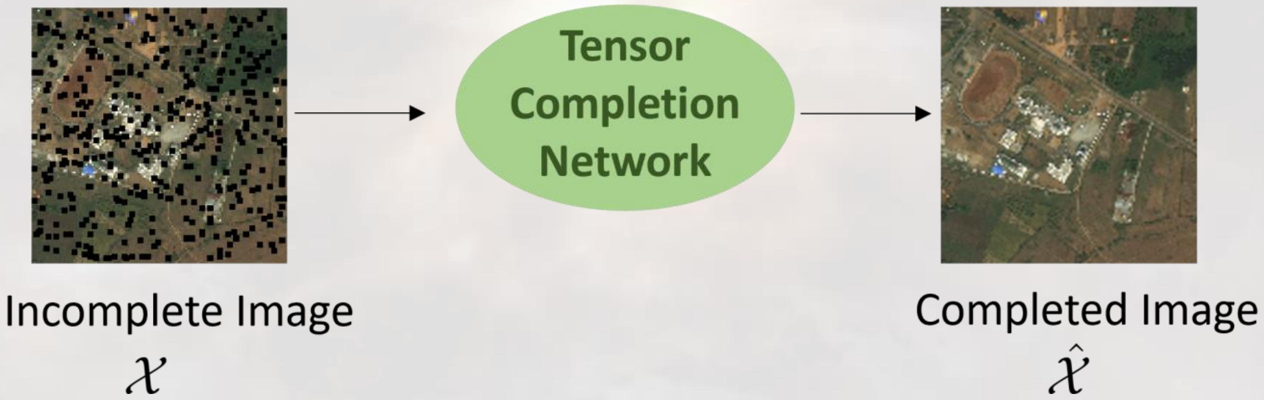


Discriminative models



Tensor Completion Neural Network

We have created a *tensor completion network* to recover the missing entries of an image by converting a model-based method into a neural network using *Tucker decomposition* and the *algorithmic unrolling* technique, without having access to the actual measurements of the missing entries.

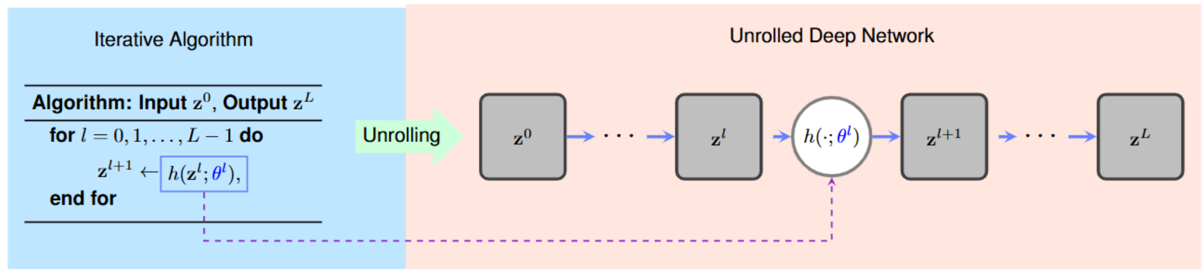


Algorithmic unrolling

Starting with an abstract iterative algorithm,

- map one iteration (described as the function h parametrized by θ into a single network layer,
- stack a finite number of layers together to form a deep network.

Feeding the data forward through an L-layer network ⇔ executing the iteration L times
The parameters θ are learned from real datasets by end-to-end optimization.

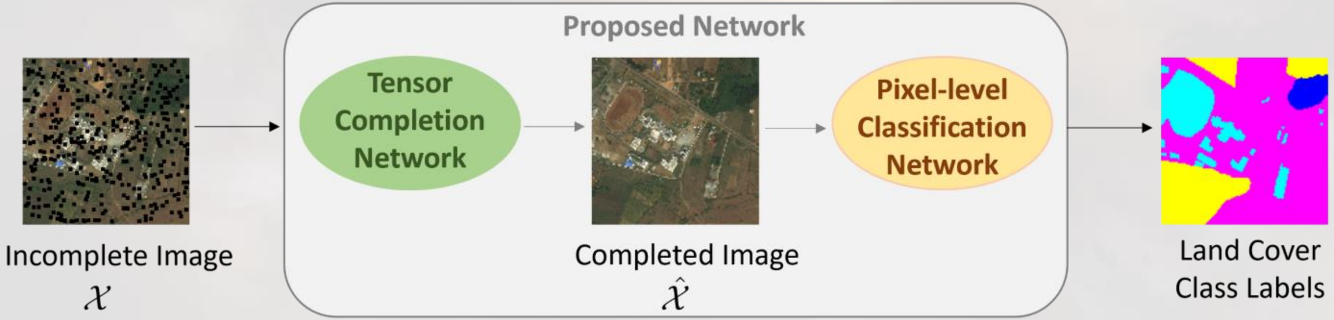


Monga, Vishal, Yuelong Li, and Yonina C. Eldar. "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing." *IEEE Signal Processing Magazine* 38.2 (2021): 18-44.

Tensor Completion Neural Network

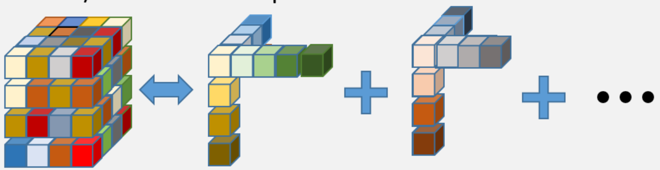
Deep learning formulation of tensor models:

- Exploit the benefits of both tensor analysis and deep learning techniques
- Create a tensor completion network
- Combine the tensor network with other popular networks
- Perform two tasks simultaneously

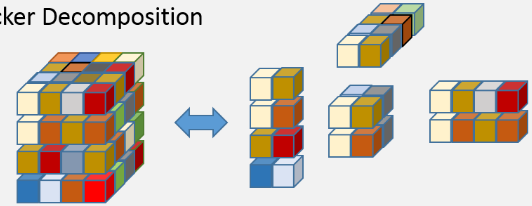


Tensor Decomposition

CANDECOMP/PARAFAC Decomposition



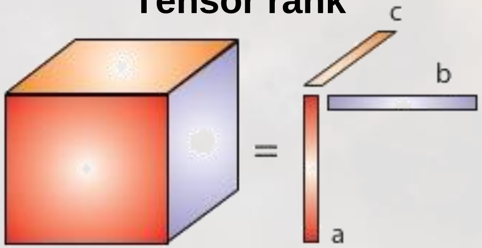
Tucker Decomposition



$$\mathcal{X} = \sum_{r=1}^{J} \lambda_r \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$$

$$\mathcal{X} = \mathcal{G} \times_1 \boldsymbol{A}^{(1)} \times_2 \boldsymbol{A}^{(2)} \times_3 \ldots \times_N \boldsymbol{A}^{(N)}$$

Tensor rank



Kronecker Product

$$\left.\begin{array}{l} \boldsymbol{X} \in \mathbb{R}^{I \times J} \\ \\ \boldsymbol{Y} \in \mathbb{R}^{K \times L} \end{array}\right\} \boxed{\mathbf{Z} = \mathbf{X} \otimes \mathbf{Y} = \begin{bmatrix} x_{11}\mathbf{Y} & x_{12}\mathbf{Y} & \cdots & x_{1J}\mathbf{Y} \\ x_{21}\mathbf{Y} & x_{22}\mathbf{Y} & \cdots & x_{2J}\mathbf{Y} \\ \vdots & \vdots & \ddots & \vdots \\ x_{I1}\mathbf{Y} & x_{I2}\mathbf{Y} & \cdots & x_{IJ}\mathbf{Y} \end{bmatrix} \in \mathbb{R}^{IK \times JL}}$$

Funded by
the European Union

"Low-rank tensor completion via tucker decompositions", Shi, Jiarong, et al., *J. Comput. Inf. Syst* (2015)

$$\min \frac{1}{2}\|\mathcal{G} \times_1 \boldsymbol{D}_1 \times_2 \cdots \times_N \boldsymbol{D}_N - \mathcal{Z}\|_F^2$$

$$\text{s.t. } \mathcal{P}_\Omega(\mathcal{Z}) = \mathcal{P}_\Omega(\mathcal{X}) \quad \text{and} \quad \mathbf{D}_n^T \cdot \boldsymbol{D}_n = \boldsymbol{I}_{R_n}, n = 1, .., N$$

Lagrange function:

$$L(\mathcal{G}, \mathbf{D}_1, .., \mathbf{D}_N, \mathcal{Z}, \mathcal{Y}) = \frac{1}{2}\|\mathcal{G} \times_1 \boldsymbol{D}_1 \times_2 \cdots \times_N \boldsymbol{D}_N - \mathcal{Z}\|_F^2 - < \mathcal{Y}, \mathcal{P}_\Omega(\mathcal{Z}) - \mathcal{P}_\Omega(\mathcal{X}) >$$

At each iteration l, we update:

$$\begin{aligned}
\boldsymbol{D}_n &= \mathrm{QR}(\boldsymbol{Z}_{(n)}^{l-1} \cdot \boldsymbol{C}_{n_{(n)}}^{-1}) \text{ where } \mathcal{C}_n = \mathcal{G} \times_{i=1, i \neq n}^{N} \boldsymbol{D}_i \\
\mathcal{G} &= \mathcal{Z}^{l-1} \times_1 \boldsymbol{D}_1^T \times_2 \cdots \times_N \boldsymbol{D}_N^T \\
\mathcal{Z}^l &= \mathcal{G} \times_1 \boldsymbol{D}_1 \times_2 \cdots \times_N \boldsymbol{D}_N + P_\Omega(\mathcal{X}) - P_\Omega(\mathcal{G} \times_1 \boldsymbol{D}_1 \times_2 \cdots \times_N \boldsymbol{D}_N)
\end{aligned}$$

LRTC-Net

Trainable Parameters:
Factor matrices $\boldsymbol{D}_n, n = 1, .., N$ (the same for all layers),
Weights (from convolutional layers)

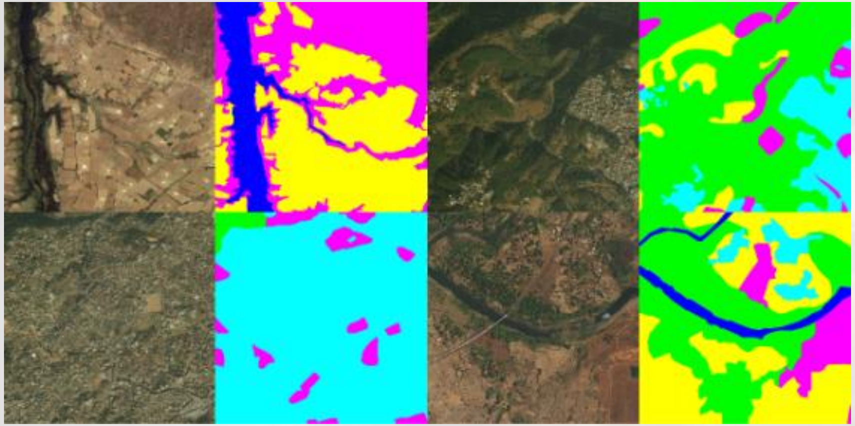
At each **layer** l we update:

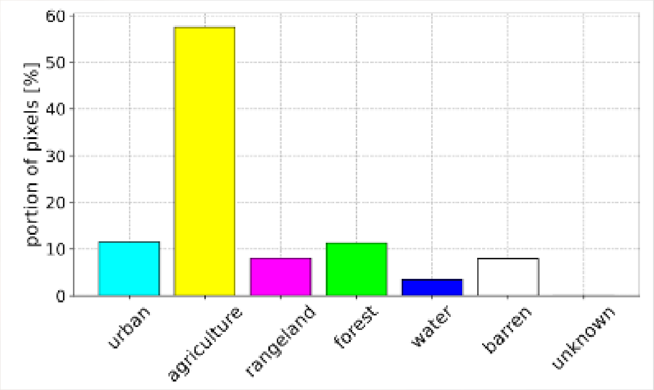
$$\begin{aligned} \mathcal{G} &= \mathcal{Z}^{l-1} \times_1 \boldsymbol{D}_1^T \times_2 \cdots \times_N \boldsymbol{D}_N^T \\ \hat{\mathcal{Z}}^l &= \mathcal{G} \times_1 \boldsymbol{D}_1 \times_2 \cdots \times_N \boldsymbol{D}_N + \mathcal{P}_\Omega(\mathcal{X}) - \mathcal{P}_\Omega(\mathcal{G} \times_1 \boldsymbol{D}_1 \times_2 \cdots \times_N \boldsymbol{D}_N) \\ \mathcal{Z}^l &= \text{CNN}(\hat{\mathcal{Z}}^l) \text{ with 4 layers} \end{aligned}$$

Loss function $= \text{MSE}(\mathcal{P}_\Omega(\mathcal{X}), \mathcal{P}_\Omega(\mathcal{Z}^L))$
where L is the number of layers.

Data

- DeepGlobe Land Cover Classification Challenge (2018)
- 803 RGB satellite images of size 2448 x 2448 x 3 (50cm pixel resolution)
- Mask images for land cover annotation with 7 classes



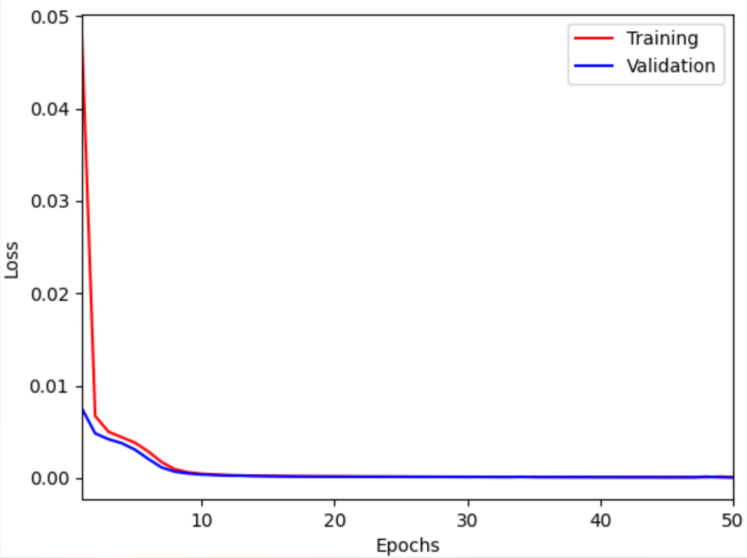


Data pre-processing: Resize to 320 x 320 x 3
Data augmentation: Horizontal and vertical flipping, random brightness and contrast

LRTC-Net Results

Training: 642 images
Validation: 80
Test: 81

50 Epochs
5 layers

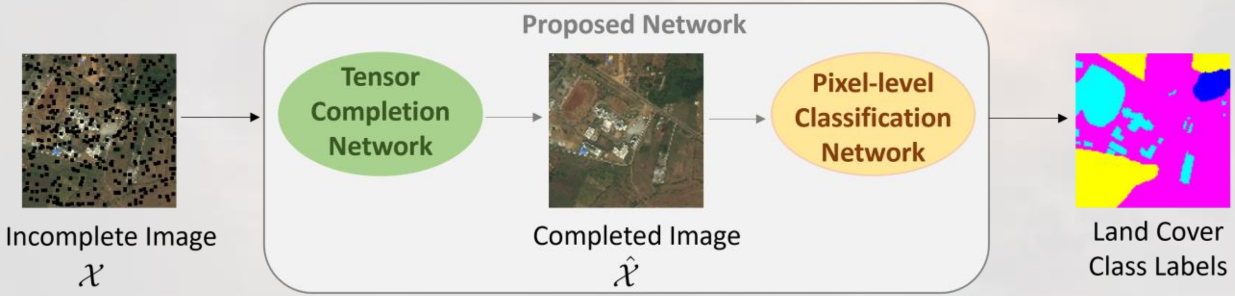


NRMSE	20% Missing values	50% Missing values	80% Missing values
LRTC-Net	0.05147	0.1208	0.1933
LRTC-Tucker	0.1156	0.2425	0.3966
LRSETD	0.06823	0.1148	0.1652

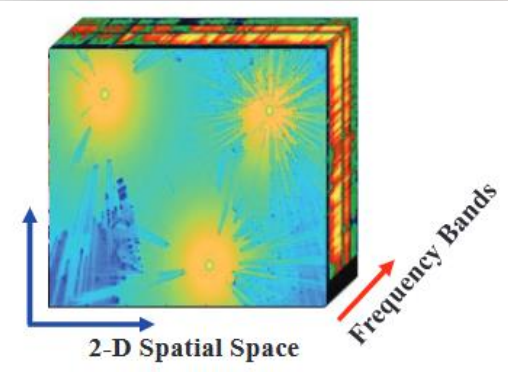
TITAN Artificial Intelligence in Astrophysics
FORTH Institute of Computer Science
Signal Processing Laboratory
CosmoStat
cea
Funded by the European Union

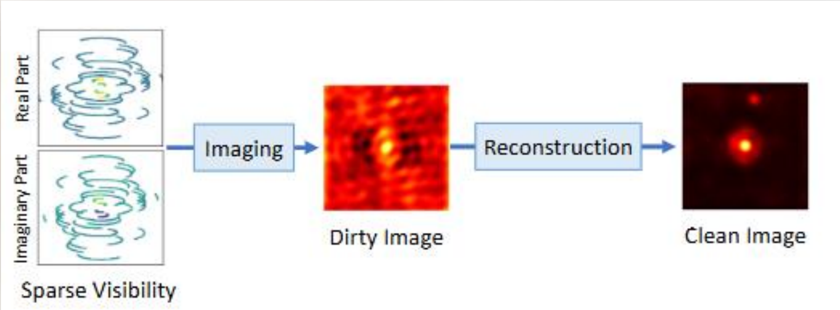
Tensor-based Models in the Deep Learning Framework

- Leverage the benefits of both tensor analysis and deep learning techniques
- Analyze high-dimensional data in all dimensions
- Improve the performance of standard models
- Use prior domain knowledge
- Interpretable networks
- Combination of tensor-based networks with other popular networks to perform two tasks simultaneously
  - ❖ Recovery of missing or corrupted measurements in combination with classification problems in multitemporal data

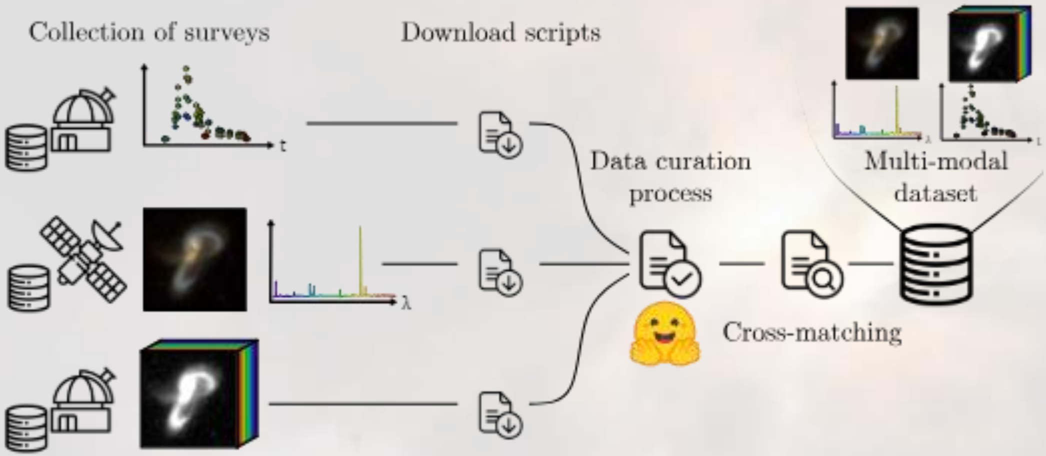


Tensor/matrix completion in radio astronomy





Multi-modal Astronomy



The Multimodal Universe: Enabling Large-Scale Machine Learning with 100 TB of Astronomical Scientific Data

The Multimodal Universe Collaboration

Eirini Angeloudi[1,2], Jeroen Audenaert[3], Micah Bowles[4,5], Benjamin M. Boyd[6], David Chemaly[6], Brian Cherinka[7], Ioana Ciucă[8,9,10], Miles Cranmer[6,5] Aaron Do[6], Matthew Grayling[6], Erin E. Hayes[6], Tom Hehir[6,5] Shirley Ho[11,12,13,5], Marc Huertas-Company[1,2,9], Kartheik G. Iyer[14,11,9], Maja Jablonska[10,9] Francois Lanusse[11,5,15], Henry W. Leung[16], Kaisey Mandel[6], Juan Rafael Martínez-Galarza[17,18] Peter Melchior[13], Lucas Meyer[11,5], Liam H. Parker[11,5,19], Helen Qu[20] Jeff Shen[13], Michael J. Smith[21,9], Connor Stone[22,23,24], Mike Walmsley[16], John F. Wu[7,25]

[1]Instituto de Astrofisica de Canarias [2]Universidad de La Laguna [3]Massachusetts Institute of Technology [4]University of Oxford [5]Polymathic AI [6]University of Cambridge [7]Space Telescope Science Institute [8]Stanford University [9]UniverseTBD [10]Australian National University [11]Flatiron Institute [12]New York University [13]Princeton University [14]Columbia University [15]Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM [16]University of Toronto [17]Center for Astrophysics, Harvard & Smithsonian [18]AstroAI [19]University of California, Berkeley [20]University of Pennsylvania [21]Aspia Space [22]Université de Montréal [23]Ciela Institute [24]Mila [25]Johns Hopkins University

Modality	Source Survey	N_c	Shape	Number of samples	Main science
Images	Legacy Surveys DR10 [43]	4	160×160	124M	Galaxies
	Legacy Surveys North [43, 134]	3	152×152	15M	Galaxies
	HSC [5, 3]	5	160×160	477K	Galaxies
	BTS [56, 114, 120]	3	63×63	400K	Supernovae
	JWST[13, 14, 50]	6-7	96×96	300K	Galaxies
Spectra	Gaia BP/RP [59]	-	110[1]	220M	Stars
	SDSS-II [1]	-	Variable	4M	Galaxies, Stars
	DESI [41]	-	7081	1M	Galaxies
	APOGEE SDSS-III [6]	-	7514	716k	Stars
	GALAH [28]	-	Variable	325k	Stars
	Chandra [51]	-	Variable	129K	Galaxies, Stars
	VIPERS [126]	-	557	91K	Galaxies
Hyperspectral Image	MaNGA SDSS-IV [2]	4563	96×96	12k	Galaxies
Time Series	PLAsTiCC[2][138]	6	Variable	3.5M	Time-varying objects
	TESS [121, 33]	1	Variable	1M	Exoplanets, Stars
	CfA Sample [68, 69, 18, 70]	5-11	Variable	1K	Supernovae
	YSE [7]	6	Variable	2K	Supernovae
	PS1 SNe Ia [127]	4	Variable	369	Supernovae
	DES Y3 SNe Ia [24]	4	Variable	248	Supernovae
	SNLS [63]	4	Variable	239	Supernovae
	Foundation [53, 81]	4	Variable	180	Supernovae
	CSP SNe Ia [36, 135, 86]	9	Variable	134	Supernovae
	Swift SNe Ia[26]	6	Variable	117	Supernovae
Tabular	Gaia [59]	-	-	220M	Stars
	PROVABGS [65]	-	-	221K	Galaxy
	Galaxy10 DECaLS [147, 92]	-	-	15K	Galaxy

TITAN Datathon

Goal: Attract students of UoC

Who: Academics & Industry partners

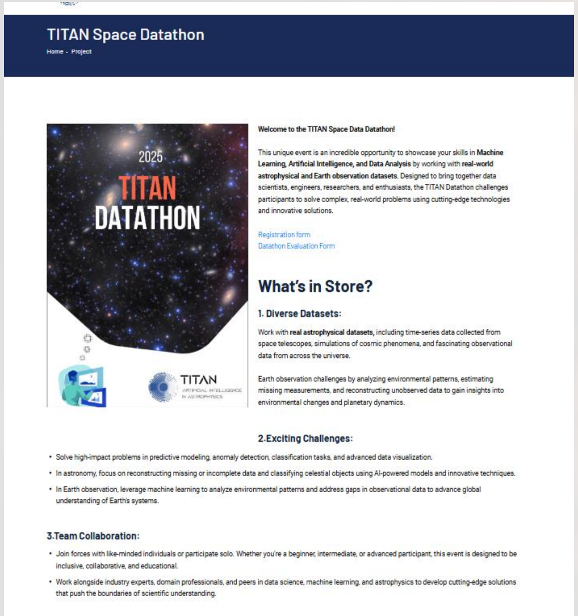
Time frame: Spring 2025

Location: FORTH/Crete

Case studies

- Astronomy (**Victor**)
- Earth Observation (**Anastasia**)

TITAN Datathon



TITAN Space Datathon

Home - Project

2025 TITAN DATATHON

Welcome to the TITAN Space Data Datathon!

This unique event is an incredible opportunity to showcase your skills in **Machine Learning, Artificial Intelligence, and Data Analysis** by working with **real-world astrophysical and Earth observation datasets**. Designed to bring together data scientists, engineers, researchers, and enthusiasts, the TITAN Datathon challenges participants to solve complex, real-world problems using cutting-edge technologies and innovative solutions.

Registration form
Datathon Evaluation Form

What's in Store?

1. Diverse Datasets:

Work with **real astrophysical datasets,** including time-series data collected from space telescopes, simulations of cosmic phenomena, and fascinating observational data from across the universe.

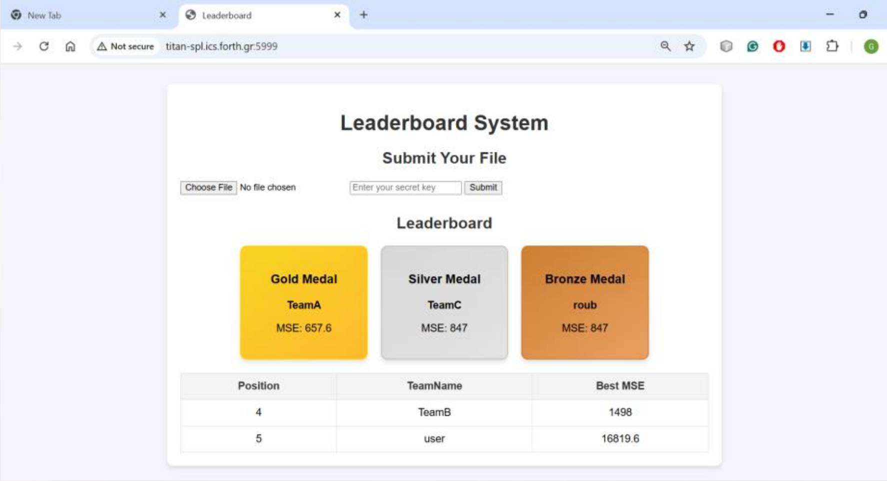
Earth observation challenges by analyzing environmental patterns, estimating missing measurements, and reconstructing unobserved data to gain insights into environmental changes and planetary dynamics.

2.Exciting Challenges:

- Solve high-impact problems in predictive modeling, anomaly detection, classification tasks, and advanced data visualization.
- In astronomy, focus on reconstructing missing or incomplete data and classifying celestial objects using AI-powered models and innovative techniques.
- In Earth observation, leverage machine learning to analyze environmental patterns and address gaps in observational data to advance global understanding of Earth's systems.

3.Team Collaboration:

- Join forces with like-minded individuals or participate solo. Whether you're a beginner, intermediate, or advanced participant, this event is designed to be inclusive, collaborative, and educational.
- Work alongside industry experts, domain professionals, and peers in data science, machine learning, and astrophysics to develop cutting-edge solutions that push the boundaries of scientific understanding.



Position	TeamName	Best MSE
4	TeamB	1498
5	user	16819.6

Funded by the European Union

Datathon

Goal: Detection of the effects of extreme events in multi-temporal multispectral satellite images

Event: Wildfire

Before the event

After the event

Mask

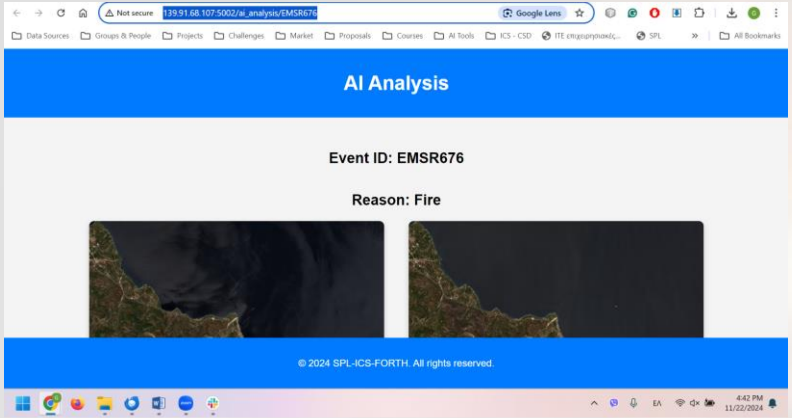
Funded by the European Union

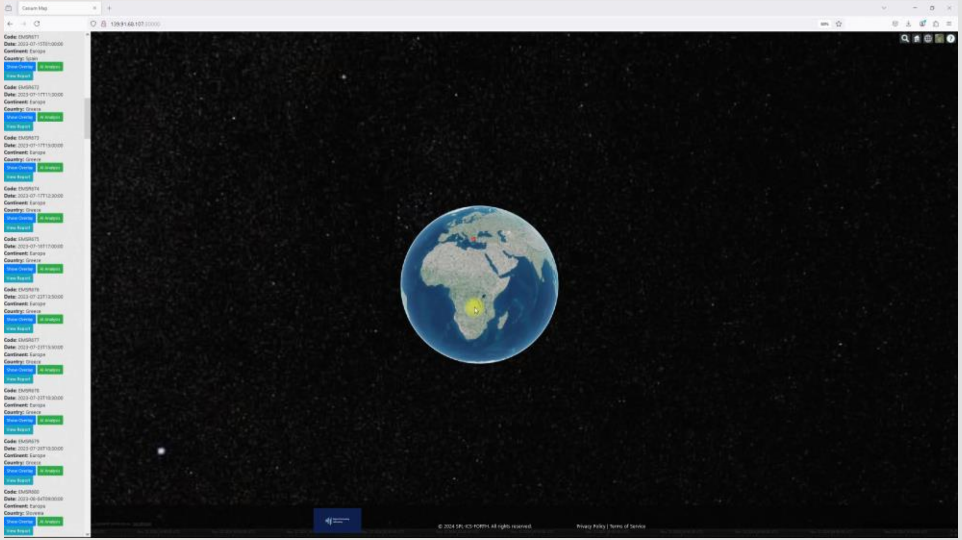
Dataset

- 20 events in Greece (e.g. wildfires, floods)
- Multi-temporal multi-spectral images acquired by Sentinel-2 with cloud cover < 50%
  (12 images before the event, 1 image after the event, 12 spectral bands)
- Synthetic Aperture Radar (SAR) images acquired by Sentinel-1 (12+1 days, 2 bands)
  (not affected by cloud cover or lack of illumination)
- Masks indicating the location of events (1-event, 0-no event)

TITAN ARTIFICIAL INTELLIGENCE IN ASTROPHYSICS FORTH INSTITUTE OF COMPUTER SCIENCE Signal Processing Laboratory COSMOSTAT cea Funded by the European Union

Open Science Platform





TITAN ARTIFICIAL INTELLIGENCE IN ASTROPHYSICS FORTH INSTITUTE OF COMPUTER SCIENCE Signal Processing Laboratory CosmoStat cea Funded by the European Union

ToDo

- Create video(s) for project -> YouTube
- Joint publications between ICS and IA
- Joint publications with SMEs