StellarFold

Challenge: Hunting for Exoplanets using AI

EVANGELOS MITIKAS
Electrical &
Electronics
Engineer

IOANNA FLORAKI
Biomedical
Engineering
Student

LOONatics

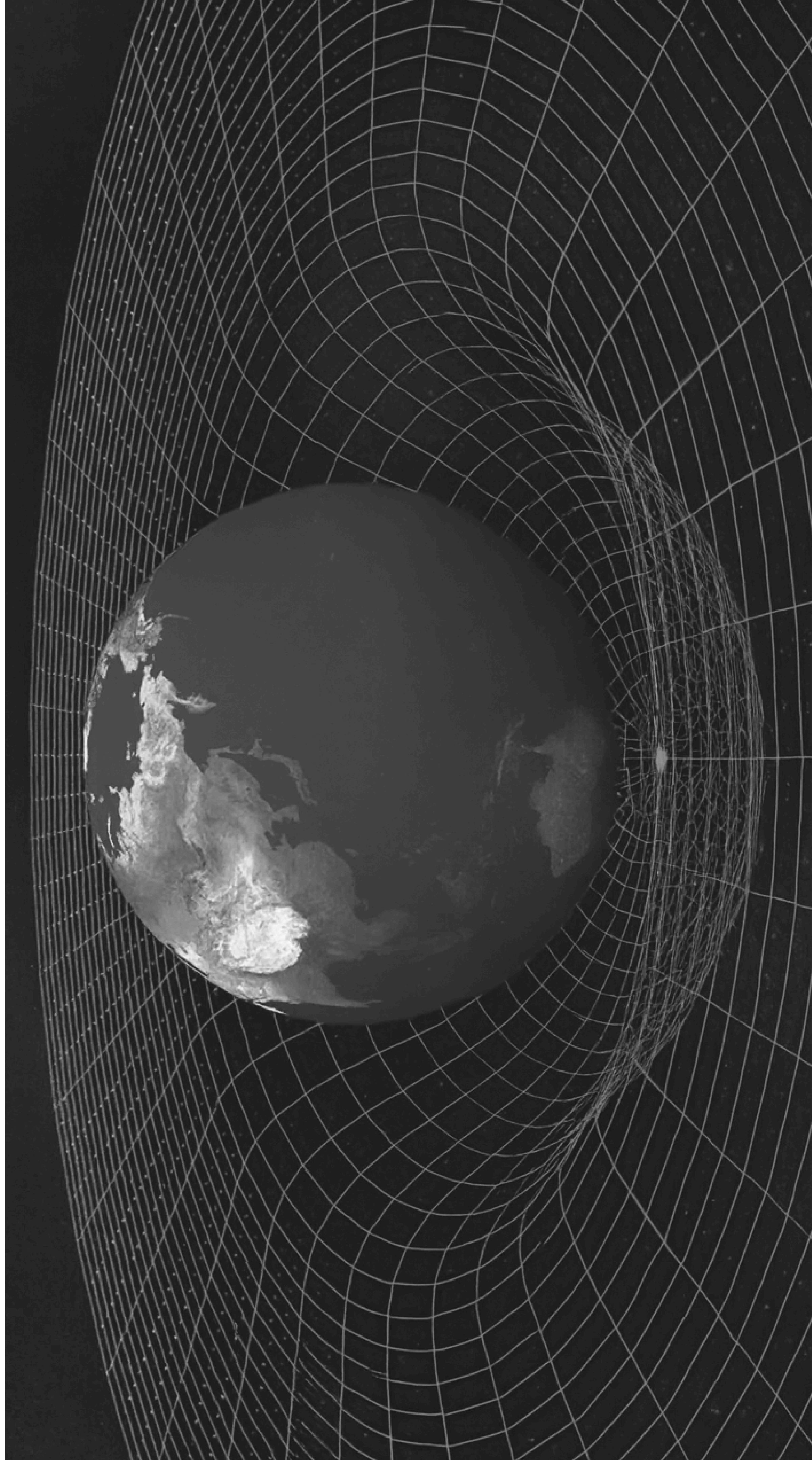
StellarFold

Scientific tool that detects exoplanet existence in orbit of a star. It is utilized on:

1. newly obtained data
2. pre-existing data for re-evaluation

In contrast to popular models (e.g. transformers, neural networks), the detection is done using differential geometry (manifolds), which:

1. can learn the geometrical sctructure of data
2. fit better for an astronomical application



DATA COLLECTION

EXOPLANET DATASETS

- Kepler Objects of Interest (KOI)
- TESS Objects of Interest (TOI)
- K2 Planets and Candidates

https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblSearch/nph-tblSearchInit?app=ExoTbls&config=k2targets

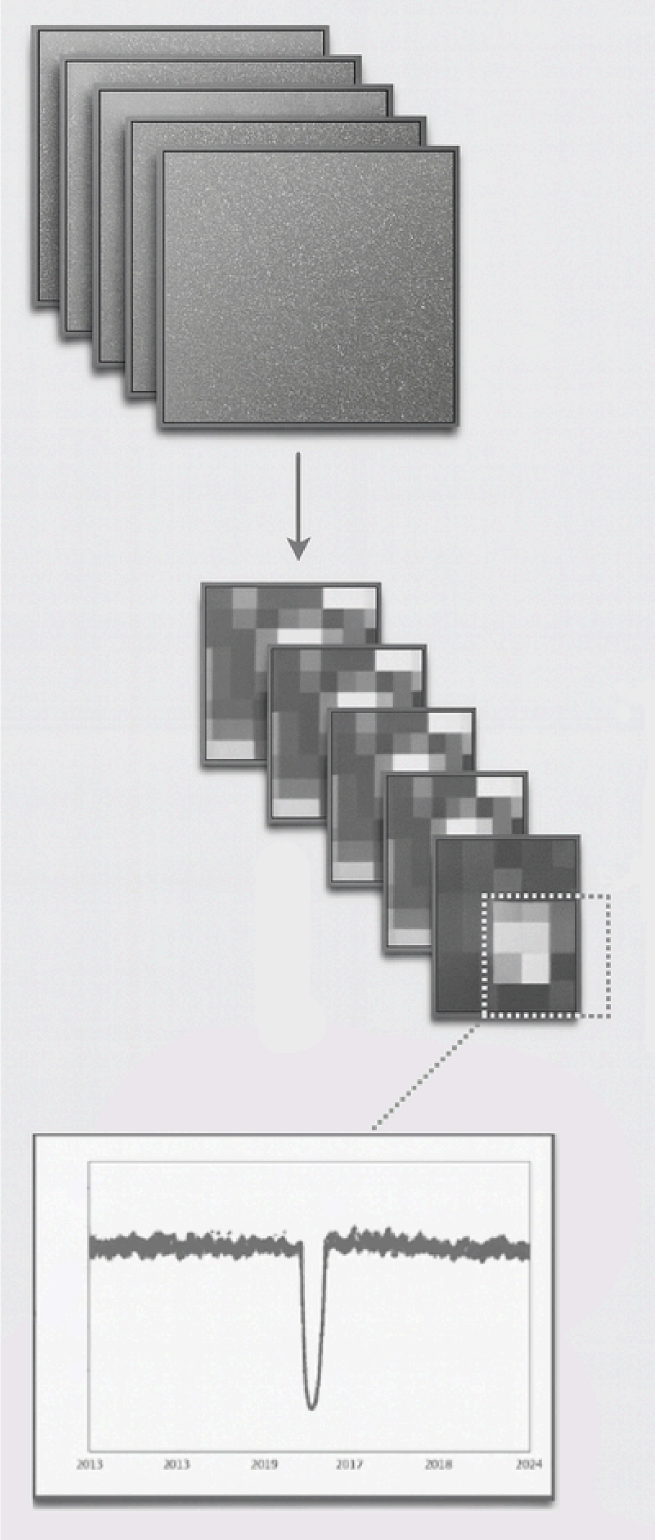
LIGHT CURVE DATA

Each mission captures Full Frame Images (FFIs) every 30" and Targeted Pixel Files (TPFs) every 2" of target stars.

The Light Curve Files (LCFs) are created from the light flux density calculated in a specified areas of a series of TPFs.

We use LCFS to extract features that encapsulate exoplanet information and can be used for our machine learning models.

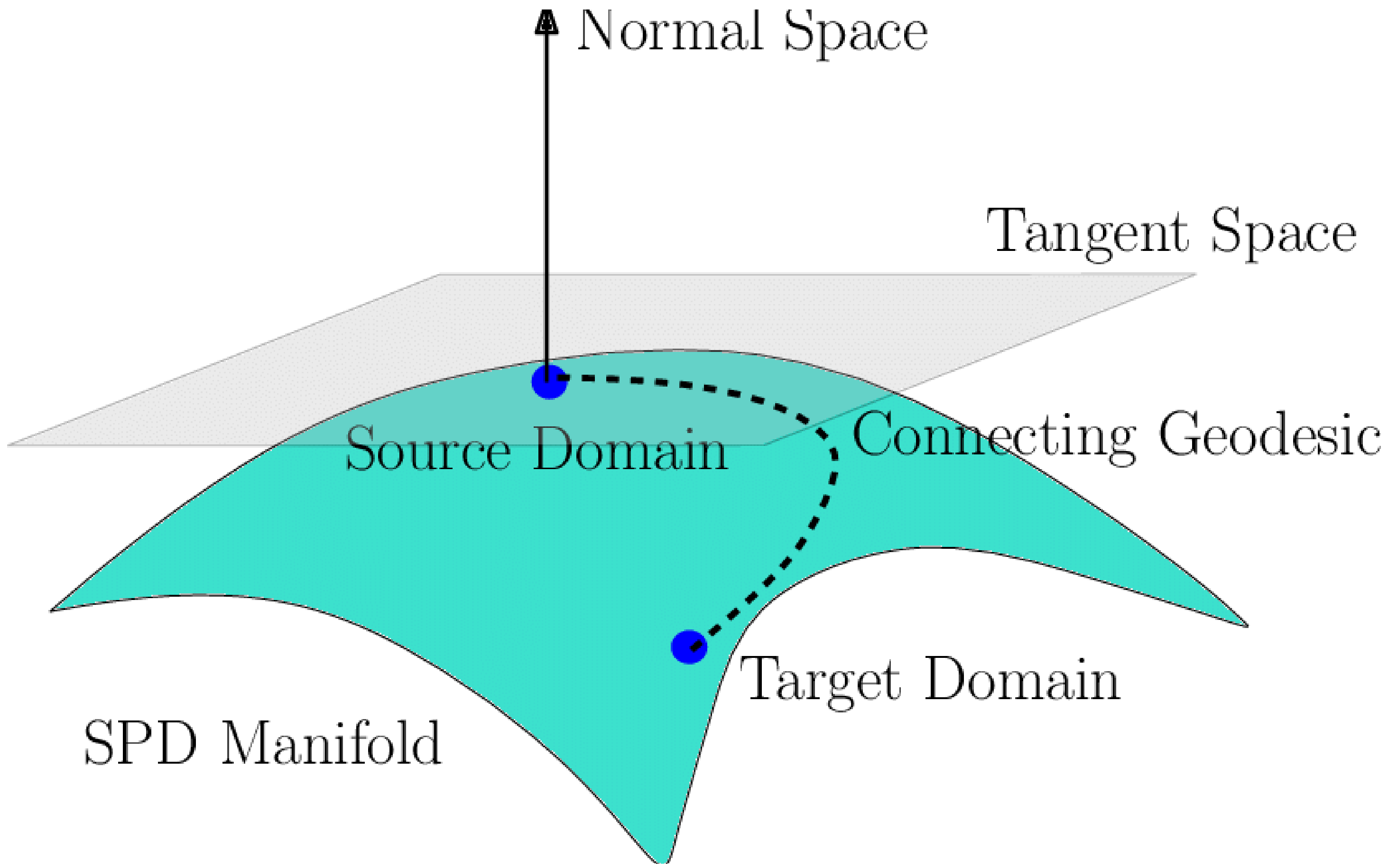
https://heasarc.gsfc.nasa.gov/docs/tess/data-products.html
https://academic.oup.com/mnras/article/513/4/5505/6472249?login=false



SPD FEATURES & TRAINING

FEATURE EXTRACTION

Features are picked for each TPF (mean, std, kurtosis, skewness, etc.) to create vectors. The covariance matrix of these vectors is used as a descriptor of each star for the model.



TRAINING

A set of cluster centroids is learned that represents four classes of stars based their exoplanets' status:

- >=1 confirmed
- 0 confirmed and >=1 candidate
- 0 confirmed, 0 candidate and >=1 refuted
- 0 confirmed, 0 candidate, 0 refuted and >=1 FP

These classes have been created based on the annotations on the available exoplanet data.

After training, the model will calculate an exoplanet existence indicator that can be used for scientists to pick the star most likely to have an exoplanet.

https://www.researchgate.net/figure/The-space-of-symmetric-positive-definite-matrices-forms-a-Riemannian-manifold-as_fig1_324860037

MANIFOLD-BASED CLUSTERING PROCESS

The cluster model is learned by minimizing a classification objective function using the Riemannian Conjugate Gradient algorithm.

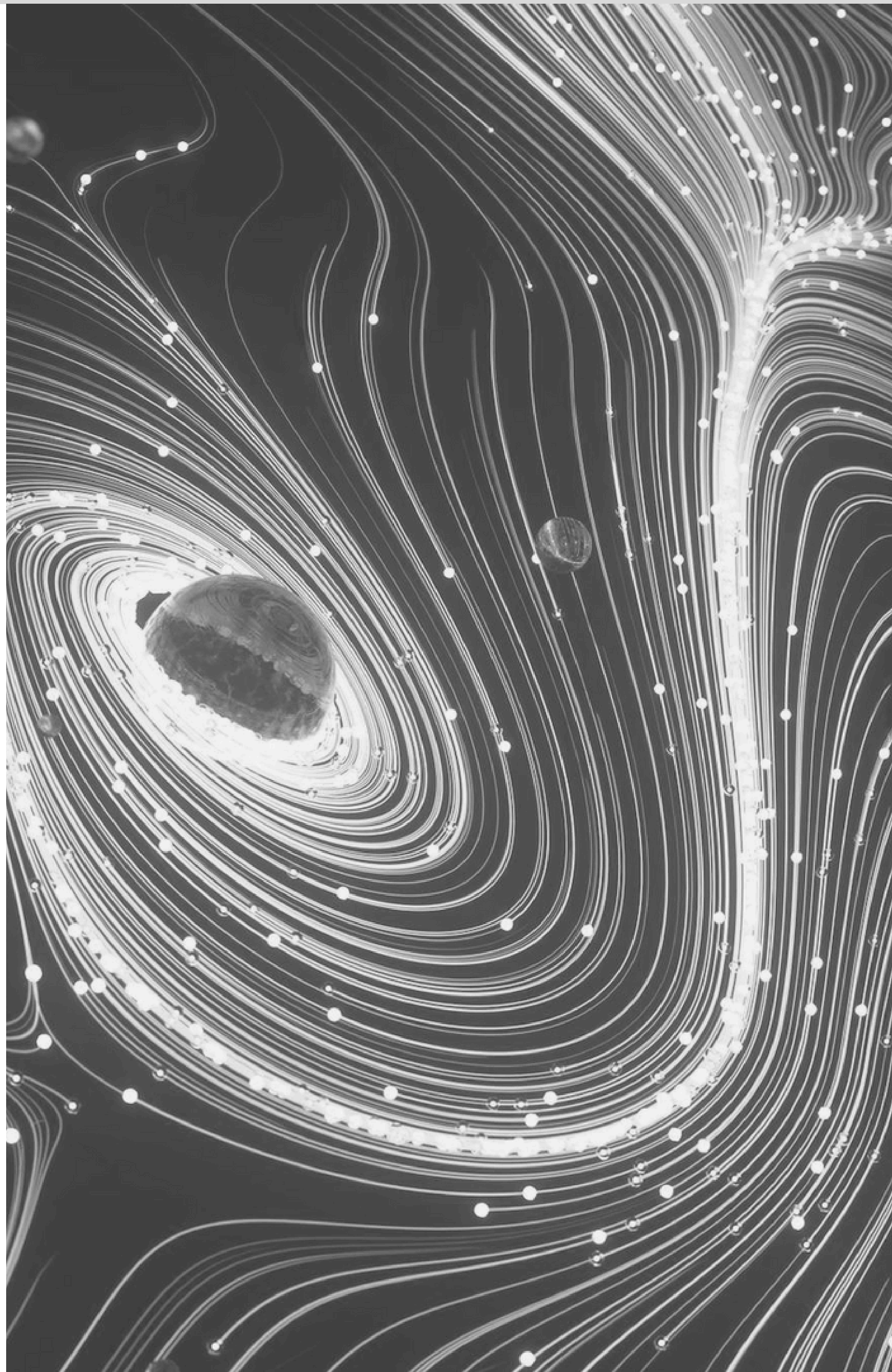
Example function:

$$\min_{\mathbb{C}} f(x, y, z) = \sum_{i=1}^{N} y_i * max(0, D_{ne} - D_{he}) + (1 - y_i) * max(0, D_{he} - D_{ne}),$$

$$\text{where } D_{he} = w_{conf}\,\mathrm{D}(X_i, C_{conf}) + w_{cand}\,\mathrm{D}(X_i, C_{cand}) + w_{ref}\,\mathrm{D}(X_i, C_{ref})$$

$$D_{ne} = w_{fp}\,\mathrm{D}(X_i, C_{fp})$$

https://openaccess.thecvf.com/content_CVPR_2020/papers/Gao_Learning_to_Optimize_on_SPD_Manifolds_CVPR_2020_paper.pdf



Project Demo

GitHub Repository: **vaggm/StellarFold**

Implementation in **MATLAB** using the **Manopt** optimization toolbox:

1. README.pdf
2. main.mlx
3. additional functions

References:

MATLAB

Manopt

THANK YOU
FOR YOUR ATTENTION!