Australia’s National Science Agency

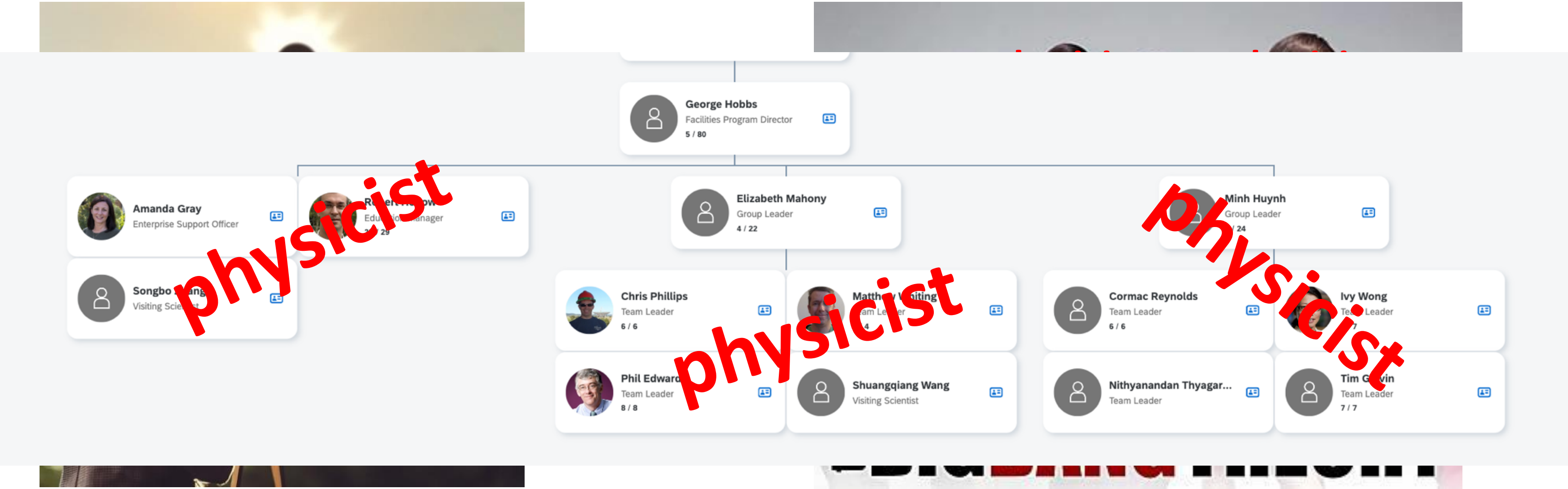
Astronomy through a Different Lens

A non-Astronomer's Perspective

Li Wang | CERC Fellow | 27/11/2024

A Growing Curiosity for the Sky

A Journey of Curiosity: From Childhood Wonder to Exploring the Stars



A child observing the sky with curiosity. *AI Generated

The Big Bang Theory

First Steps into Astronomy

My Job Interview: Stars, Data Cubes, and a Bit of Confusion

Interview problem (Job 90751)

Please propose a machine/deep learning- based solution to the following problem. Note that non-ML solutions may exist for this problem but we are asking you specifically to describe an ML/DL-based solution. You will be given 5 minutes at the interview to pitch us your ML solution. We do not require an end-to-end working demonstration of your solution but we would be interested in having a short discussion based on your proposed solution.

Context: The WALLABY survey using the Australian SKA Pathfinder (ASKAP) telescope will perform a large-area survey of the sky for atomic Hydrogen (HI) via its 21-cm hyperfine transition emission out to a distance or lookback time of ~1 billion years.

Fig 1 provides an example of an HI data cube and the extracted spectra (flux intensity as a function of frequency or velocity) from 2 regions within the cube. As emitting sources increase in distance, the HI emission line will become redshifted to lower frequencies. In Fig 1, the detection of the HI emission line is clearly seen from the region bound by the green box. On the other hand, there is no HI detection in the region bound by the red box.

To further complicate matters, the baseline spectra around the emission line may also have residual noise ripples which can be confused with a true HI emission line. Therefore as the signal-to-noise of the HI emission decreases, it becomes increasingly difficult to determine true detections from false detections.

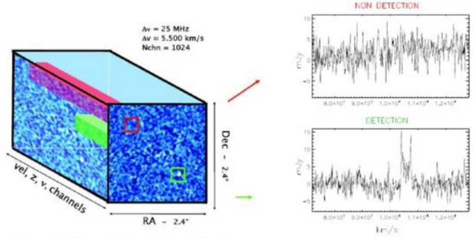
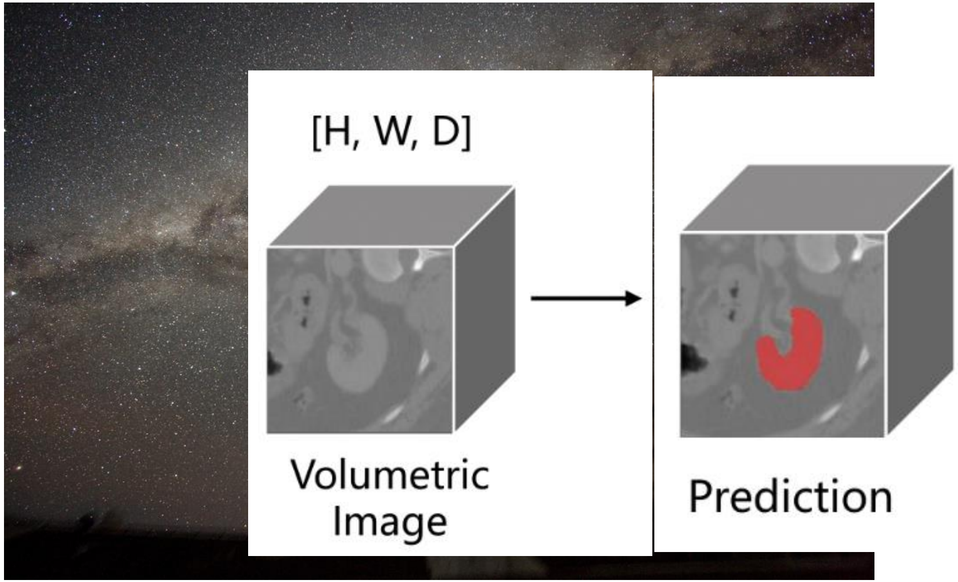


Figure 1: Example HI data cube from Fabello et al 2011.

****

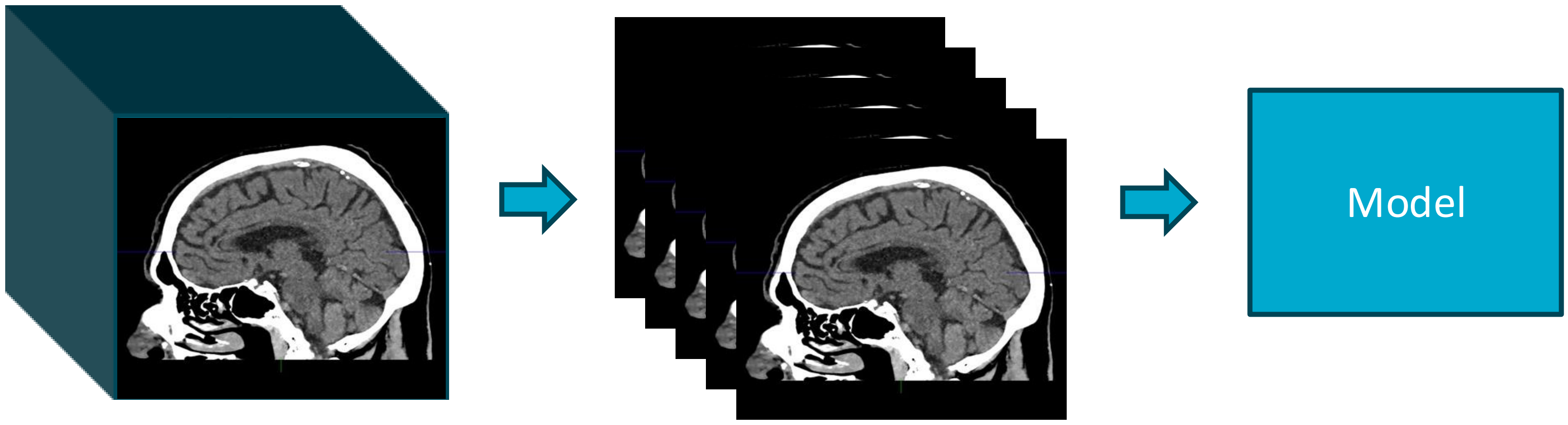
We provide an example ASKAP datafile in the original FITS format (astronomer's format)[1] with this interview problem which contain several true HI sources as well as potential false positives. This file has also been gzipped and so will need to be gunzipped. Please note that this is a smaller dataset than a single ASKAP pointing but large enough to be challenging to traditional methods for automated source recovery. What would you propose as your ML/DL-based solution that could automatically tell apart the true from the false detections, especially at low signal-to-noise levels?



From Computer Vision to Astronomy

Same Tools, Different Challenges

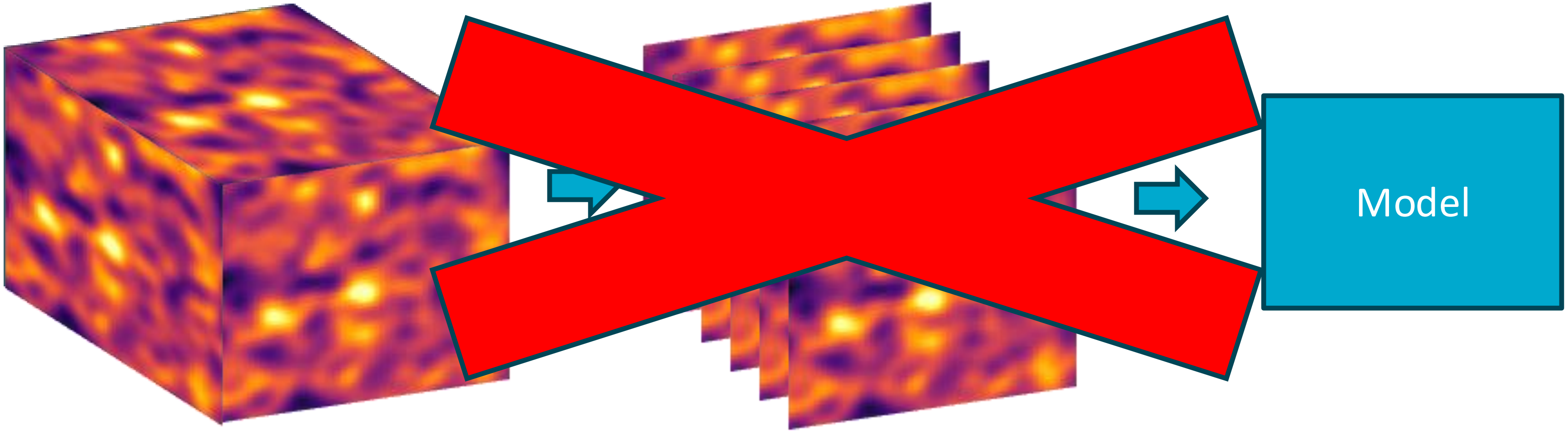
- **Medical Imaging Approach**:
  - In CT/MRI, the 3D data represents physical structures (e.g., tissues or organs) that are inherently spatially continuous.
  - Slicing these volumes into 2D images retains the continuity of features across layers, making it effective for analysis.



From Computer Vision to Astronomy

Same Tools, Different Challenges

- **Radio Image Cube**: A 3D representation of HI (neutral hydrogen) signals capturing the spatial and velocity distribution of gas in galaxies.



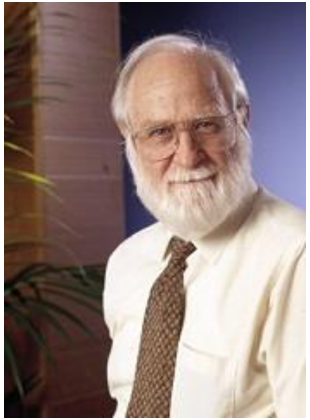
From Computer Vision to Astronomy

Why It Didn’t Work: The Physics Gap

- **Challenges with Radio Images**:
    - **High Noise Levels**: Radio signals are faint and often buried in noise.
    - **Irregular Structures**: The HI data cube is also 3D, but the third dimension isn’t spatial—it represents **frequency**, which correlates with the velocity of neutral hydrogen gas through the Doppler effect.
    - **Sparse Data**: HI signals can be highly scattered across slices, with weak continuity between 2D layers.

“To do the useful science, you must understand the instrument well.”

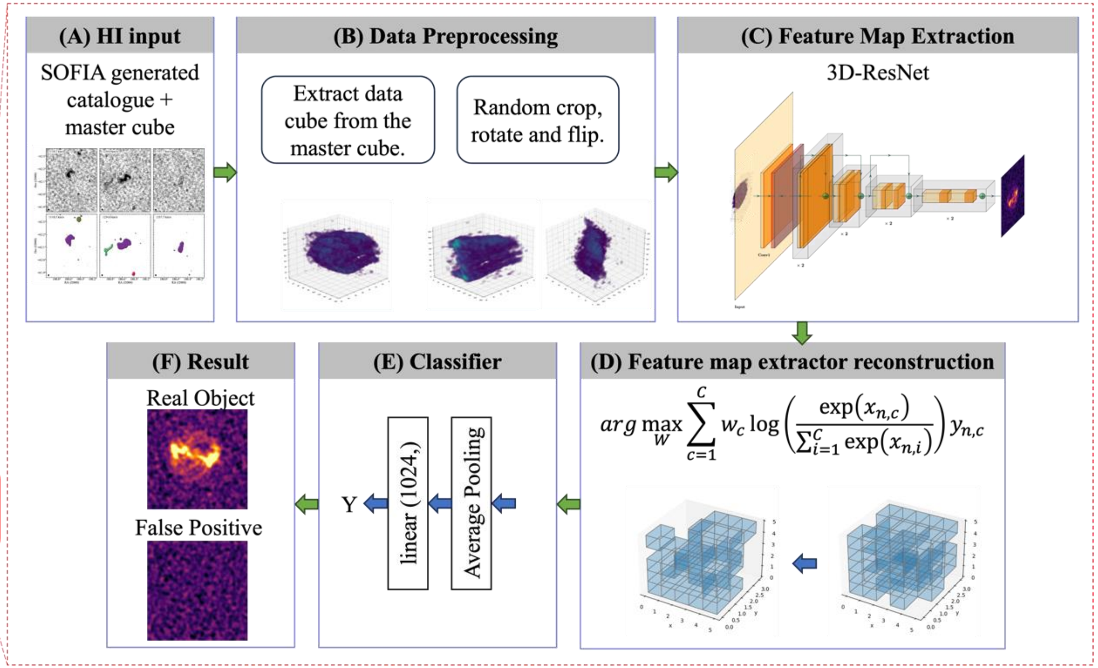
-- Prof. Ron Ekers



From Computer Vision to Astronomy

Building My First Working Model

The ML method rejection (This paper)



L Wang, et al. *PASA.* 2024

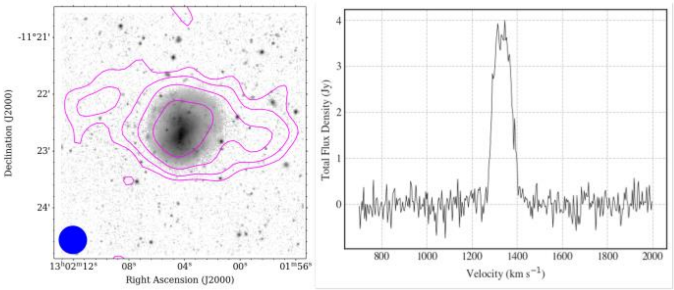
From Computer Vision to Astronomy

What I Learned: Physics Guides AI

The Core Realization: Data is Not Just Numbers

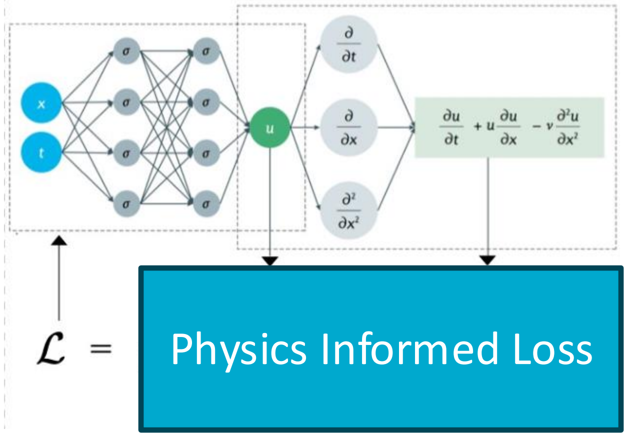
```
array([[0.14213333, 0.82429343, 0.51460849, 0.73955439, 0.66099584],
       [0.26197975, 0.48108974, 0.71687447, 0.39189053, 0.33559641],
       [0.40682764, 0.89710629, 0.8468038 , 0.23186316, 0.34631472],
       [0.39914504, 0.94851763, 0.64665897, 0.34734862, 0.62033593]])
```

vs



Each value represents physical phenomena.

Physics Can Guide AI



Physics provides the structure and context that makes the data meaningful.

Bridging the Gap Between AI and Astro

What I Learned: Aligning AI with Astronomy

1. Physics-Aware AI

- Embed physical principles into AI models to improve performance.
- Examples:
  - Use continuity in frequency space to model relationships in HI cubes.

2. Understanding the Data Pipeline

- Understand how astronomers process **visibility data** into **image data.**
- Cleaning is critical

Bridging the Gap Between AI and Astro

What I Learned: Aligning AI with Astronomy

3. Explainable AI

- Build trust by making AI predictions more explainable:
  - Use techniques like saliency maps to show what the model "sees."
  - Ensure predictions align with physical intuition and scientific reasoning.

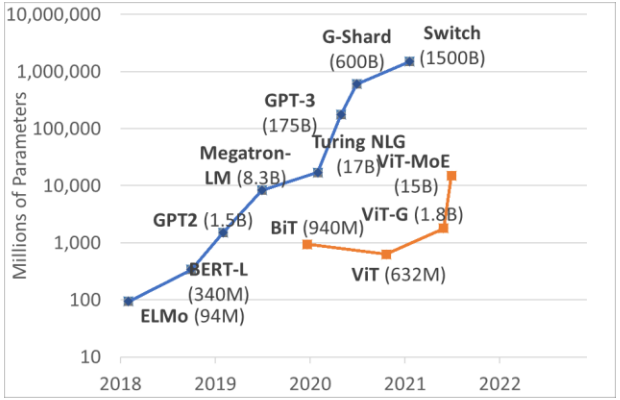
4. Collaboration is Key

- Combine the strengths of AI researchers and astronomers:
  - AI contributes computational tools and scalable solutions.
  - Astronomy provides domain expertise and context.
- Together, create hybrid approaches to tackle complex problems.

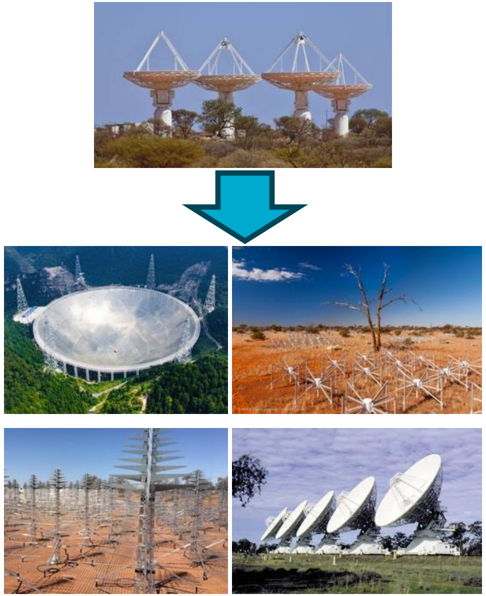
Current work: Scaling Up AI for the Universe

Applying Large Models to Astronomy

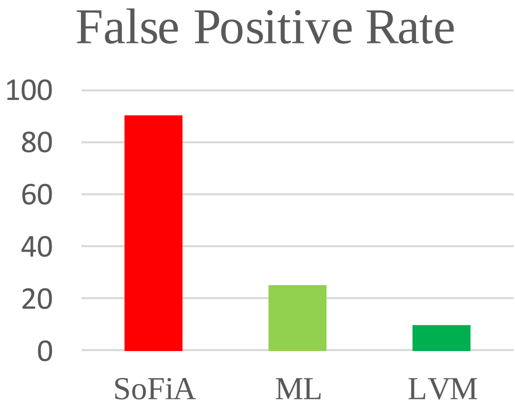
Handling Complex Patterns



Generalization Capability



Enhanced Accuracy



The Future Role of AI in Astronomy

What’s Next for Astronomy?

1. AI as a Core Tool in Astronomy

- **Automating Repetitive Tasks**
    - Source finding、 Denoising、 Morphology.
- **Real-Time Discovery**:
    - Detecting transient events (e.g., supernovae, fast radio bursts) as they happen and triggering immediate follow-up observations.
- **Predictive Modeling**:
    - Forecasting future observations or simulating cosmic events.

2. Telescope Integration with AI

- **Smart Telescopes**:
    - Telescopes equipped with onboard AI for noise reduction, anomaly detection, and dynamic observation adjustments.

Key Takeaways

Looking at Astronomy Through a Different Lens

1. Data is Not Just Numbers

Understanding the physical meaning behind data is crucial for designing effective AI models.

2. AI is a Tool, Not a Solution:

The best results come from combining AI with scientific intuition and interpretation. The Future of Astronomy is AI-Enhanced.

3. Collaboration is the Key to Success

Next-generation telescopes and astronomical datasets demand deeper collaboration between AI researchers and astronomers.

Thank you

Space and Astronomy
Li WANG
CERC Fellow

li.wang1@csiro.au