Applications of Computer Vision to Space and Plasma Physics

Discovery Learning Apprenticeship 2024-2025

Erick White

April 5, 2025

Ann and H.J. Smead Department of Aerospace Engineering Sciences

Smead Aerospace
UNIVERSITY OF COLORADO **BOULDER**

Introduction

Background: Space Physics

- **Magnetosphere:** Region of space surrounding the Earth and permeated by its magnetic field
- **Plasma Wave:** Charged particle event that causes a measurable disturbance in the magnetosphere
  - **EMIC (electromagnetic ion cyclotron) waves** are a type of plasma wave

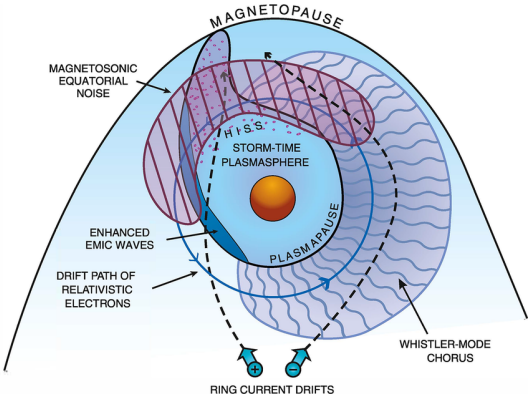


Illustration of magnetosphere and plasma events. Figure from Thorne (2010).

Background: Machine Learning

- **Object Detection:** A branch of machine learning dedicated to identifying certain categories of objects in images
- **YOLO:** You Only Look Once, a near-real-time object detection model used in this project

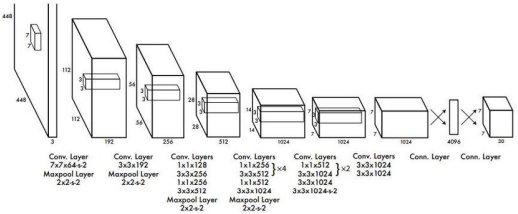


Illustration of YOLO model architecture. Figure from Redmon *et al.* (2016).

Motivation

- Van Allen Probes were launched in 2012 to observe Earth's magnetosphere
  - Recorded data for around seven years
  - Impossible to analyze entirely by hand
- Increased prevalence of machine learning
  - Great potential for scientific data analysis
  - "Proof-of-concept" for similar applications to other missions

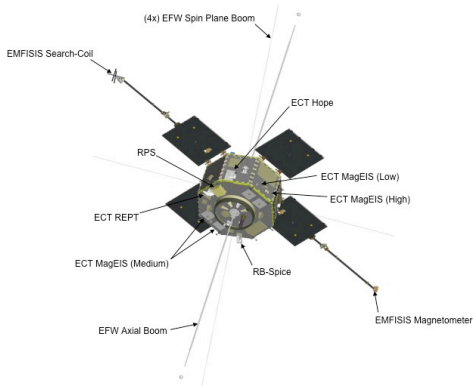
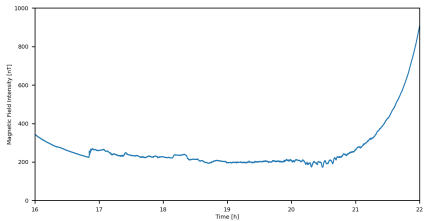


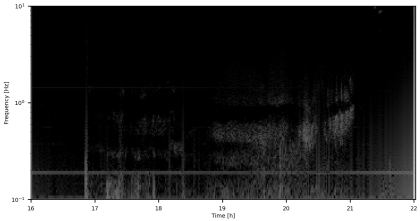
Image of Van Allen Probe including instrument labels. Figure courtesy of JHUAPL.

Methods

Plasma Wave Detection

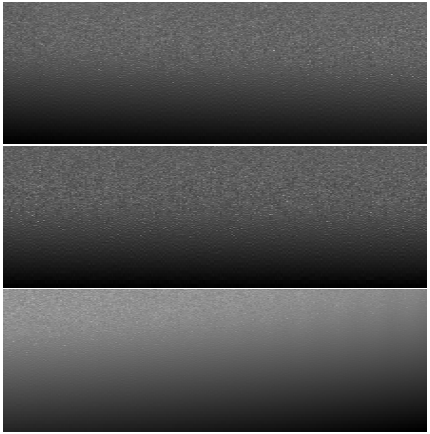
- Plasma waves disturb the magnetosphere in a measurable way
  - Can use short-time Fourier transforms to detect these signatures
- Optimize spectrograms for EMIC wave detection
  - Typically restricted to the 0.1 – 10 Hz range
  - 0.3125 – 8.281 25 Hz range chosen for this project





Spectrogram Generation

- Lots of null data
  - A problem for working at scale
  - Too much background data leads to a poorly-trained model
- How to find meaningful data?
  - Inglis *et al.* database (August 2024)
    - Large, but lots of false positives and negatives
    - Thousands of generated images led to only a few hundred annotations



Examples of null spectrograms

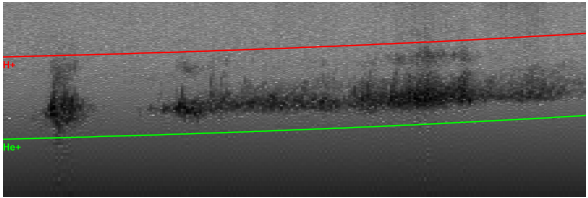
Verification Process

- Gyrofrequency annotation
  - Are events where we expect them to be?
- Visual inspection
  - Do they *look* right?

Governing Equations

$$\omega_c = \frac{qB}{m}$$

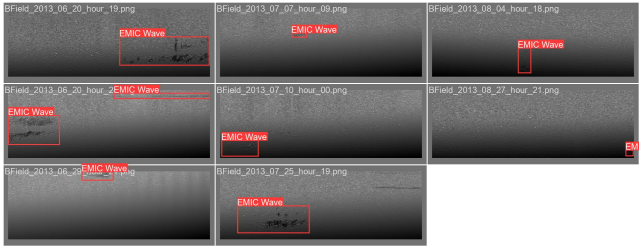
$$|B| = B_0 \left(\frac{R_E}{r}\right)^3 \sqrt{1 + 3\sin^2(\lambda)}$$



A spectrogram with H+ and He+ gyrofrequencies annotated - note how well they line up!

Spectrogram Annotation

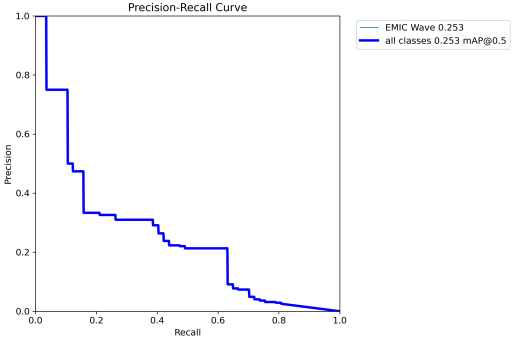
- Annotations make the model work
  - Telling the model where EMIC waves are to train on
- Tool in question: CVAT
  - Can batch upload and annotate directly
- Lots of judgement calls
  - What is an EMIC wave?
  - What isn't?



An example of EMIC wave annotations for model training.

Model Training

- Model is data-ready
  - YOLO can directly process annotated images
- A couple of big pitfalls
  - Null images messed up the model pretty bad
  - Incorrect annotations can lead to bad predictions!
  - Finding enough events can be a challenge

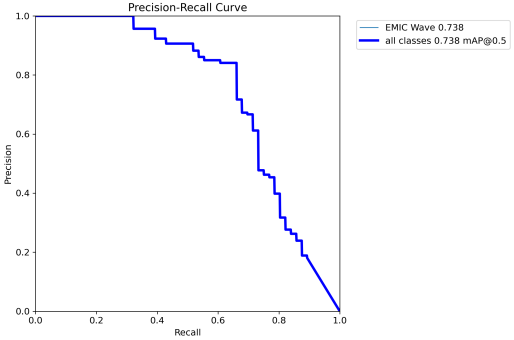


Example of poor model performance when trained with too many null images.

Results

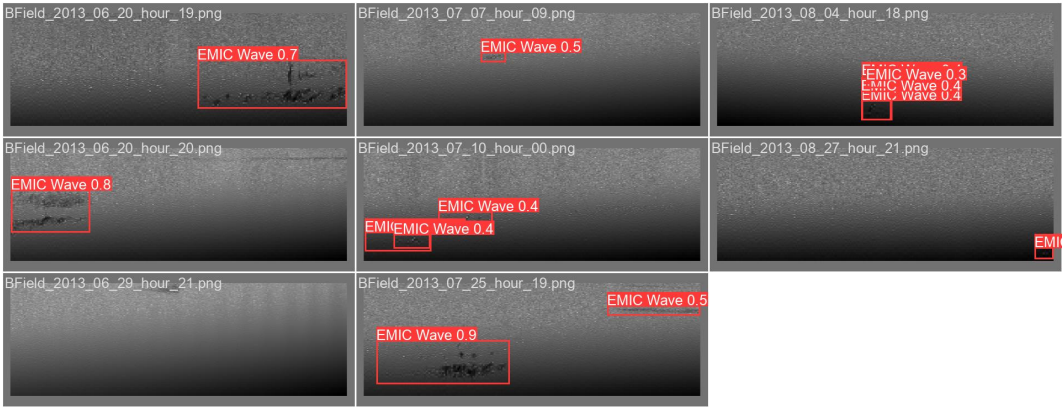
Model Results

- Results seem promising!
  - Even with only a few hundred images, the model can identify EMIC waves relatively well
    - Have only annotated around two and a half years so far
    - Can expect at least around 2.5 times this number of events
    - Annotations are the biggest bottleneck at this point
- Model can automatically quantify confidence of detections
  - Able to filter results by confidence



Precision-recall curve for the most recent version of the model.

Model Results



Model predictions for several spectrograms with likely EMIC wave events.

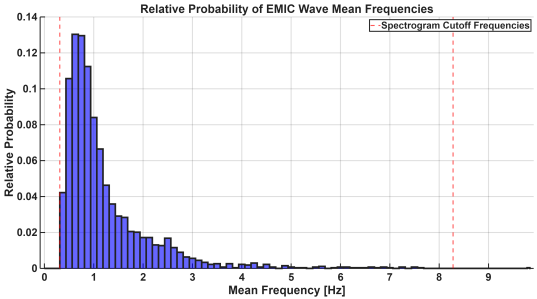
Conclusions

Feasibility and Applicability

- Worthwhile track to pursue?
  - Seems promising!
  - Potentially more versatile than current methods employed
  - Reduces the amount of data scientists need to look through
- Machine learning can be integrated efficiently into a scientific workflow
  - Heavy lifting is in configuration work
  - Annotation is time-consuming, but relatively straightforward
    - Still need someone familiar enough to recognize problems

Analysis of EMIC Waves

- Events have been found with longer duration than previously expected
  - Contradicts Inglis *et al.* results
- Lacking the distinct shape some plasma waves have
  - Potentially harder to identify

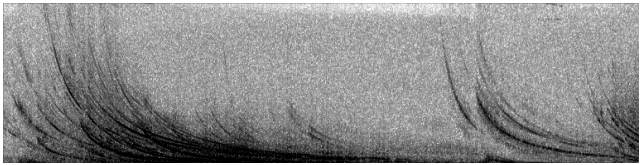


Distribution of EMIC wave mean frequencies according to Inglis *et al.* (2024).

Future Work

Analysis of EMIC Waves

- Quantifying where these waves occur
  - Polar plots
  - Do they happen where we think they do?
  - How do they compare to other plasma waves?



Distinctive appearance of whistler plasma waves in spectrograms. Figure courtesy of Paraksh Vankawala.

Analysis of EMIC Waves

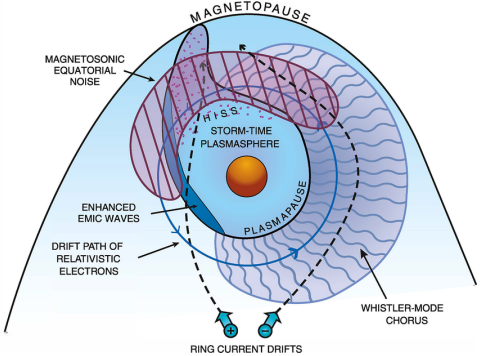
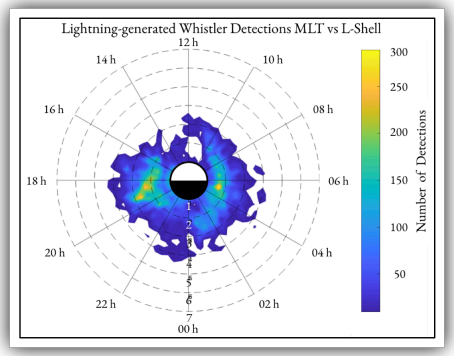


Illustration of magnetosphere and plasma events. Figure from Thorne (2010).



Polar plot of whistler-type wave distribution. Figure courtesy of Paraksh Vankawala.

Acknowledgments

Acknowledgments

A big thank you to the following teams and individuals for their support, guidance, and funding throughout this project:

The LAIR Lab

- Paraksh Vankawala (DLA Mentor)
- Dr. Robert Marshall (Principal Investigator)

Discovery Learning Apprenticeship Team

- Dr. Sharon Anderson (DLA Coordinator)
- Grace Griffith (Coordinator of Undergraduate Research Programs)

Thank You

Questions?

References

References

[1] Ronald L. Allen and Duncan W. Mills. ***Signal Analysis: Time, Frequency, Scale, and Structure.*** John Wiley & Sons, 2004.

[2] Andrew R. Inglis, Kyle R. Murphy, and Alexa Halford. **“A Survey of EMIC Waves in Van Allen Probe Data”.** In: *Journal of Geophysical Research: Space Physics* (2024).

[3] C. A. Kletzing et al. **“The Electric and Magnetic Field Instrument Suite and Integrated Science (EMFISIS) on RBSP”.** In: *Space Science Review* (2012).

[4] Johns Hopkins University Applied Physics Laboratory. ***Van Allen Probes: Spacecraft.*** URL: https://vanallenprobes.jhuapl.edu/Spacecraft/index.php.

[5] W. Li and M. K. Hudson. **"Earth's Van Allen Radiation Belts: From Discovery to the Van Allen Probes Era".** In: *Journal of Geophysical Research: Space Physics* (2018).

[6] Robert A. Marshall. **"Module 2: The Magnetosphere".** Lecture notes. 2021.

[7] Robert A. Marshall and Paraksh Vankawala. ***Computer Vision-Generated Database of Plasma Waves in the Inner Magnetosphere.*** Grant proposal. 2019.

[8] Keiron O'Shea and Ryan Nash. **"An Introduction to Convolutional Neural Networks".** arXiv preprint. 2015.

[9] Joseph Redmon et al. **“You Only Look Once: Unified, Real-Time Object Detection”.** In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 779–788.

[10] Richard Mansergh Thorne. **“Radiation belt dynamics: The importance of wave-particle interactions”.** In: *Geophysical Research Letters* (2010).

[11] Daniel Vech and David M. Malaspina. **“A Novel Machine Learning Technique to Identify and Categorize Plasma Waves in Spacecraft Measurements”.** In: *Journal of Geophysical Research: Space Physics* (2021).

[12] Zheng Xiang et al. **“Understanding the Mechanisms of Radiation Belt Dropouts Observed by Van Allen Probes”.** In: *Journal of Geophysical Research: Space Physics* (2017).

Backup Slides

Detection Challenges

Theorem: Spectrogram Limiting Resolution

A time-frequency atom with time width σ_t and frequency width σ_ω cannot exceed a resolution limited by

$$\sigma_t \sigma_\omega \geq \frac{1}{2}$$

Aside: Drawing Connections

The Uncertainty Principle

$$\sigma_t \sigma_\omega \geq \frac{1}{2}$$

Aside: Drawing Connections

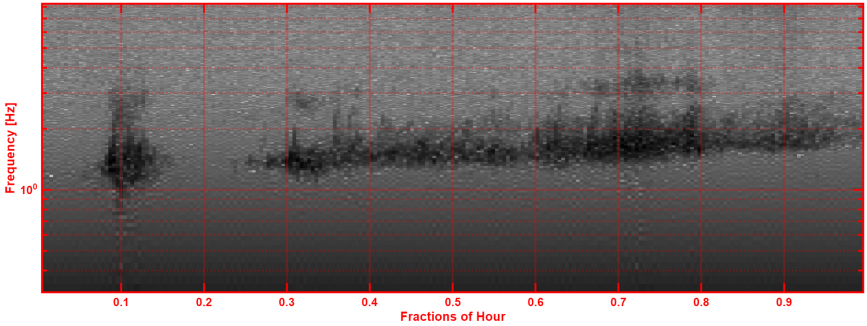
The Uncertainty Principle

$$\sigma_t \sigma_\omega \geq \frac{1}{2}$$

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

$$\Delta E \Delta t \geq \frac{\hbar}{2}$$

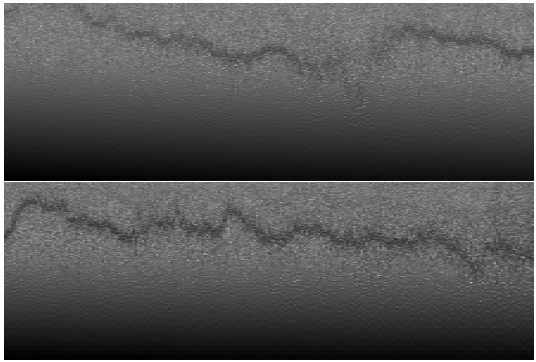
Spectrogram Generation



Spectrogram containing a likely EMIC wave event!

Verification Process

- Problem: How to verify that a signal is an EMIC wave?
  - Lots of unknown signals
  - Not much to verify against
- Lots of weird stuff going on in space
  - Must ensure that model is training on EMIC only



Two hours of an unknown 13-hour signal on 11/22/2013!

Expanding to Other Missions

- Main issue: consistent spectrograms
  - Image size affects model results
    - Frequency range?
    - Time scale?
  - Feasibility of a standardized spectrogram generation method
    - How much can be carried over?