Classification of High-Energy Particle Precipitation Events Using Computer Vision

Caleb Bettcher
University of Colorado Boulder

INTRODUCTION 01

INTRODUCTION

Radiation Belts

- Donut shaped regions around the Earth
- Protons and electrons trapped in orbit

Energetic Particle Precipitation (EPP)

- Particles from the radiation belts falling into Earth's atmosphere
- From disruptions such as a coronal mass ejection or other space weather events

Effects of EPP

- Interrupt telecommunications
- Destroy sensitive electronic components
- Deplete atmospheric ozone

* EPP plays a significant role in the formation of the Aurora Borealis

MOTIVATION

Large-Scale Mission

Determine the atmospheric impact of the varying sources of EPP in order to build better predictive models of space weather and climate.

This Work

Develop tools to automatically create databases of EPP events and their sources from satellite data.

DATA COLLECTION

* The two ELFIN twins prior to launch! Each one is about the size of a loaf of bread

ELFIN Cube Satellites

- Electron Losses and Fields Investigation (ELFIN)
- Launched 9/15/18 by UCLA
- Sensors measure energy, incoming angle, and particle flux (rate of particle measurements)
- Direction of an incoming particle determines if it will precipitate into the atmosphere or be trapped in orbit

METHODOLOGY 02

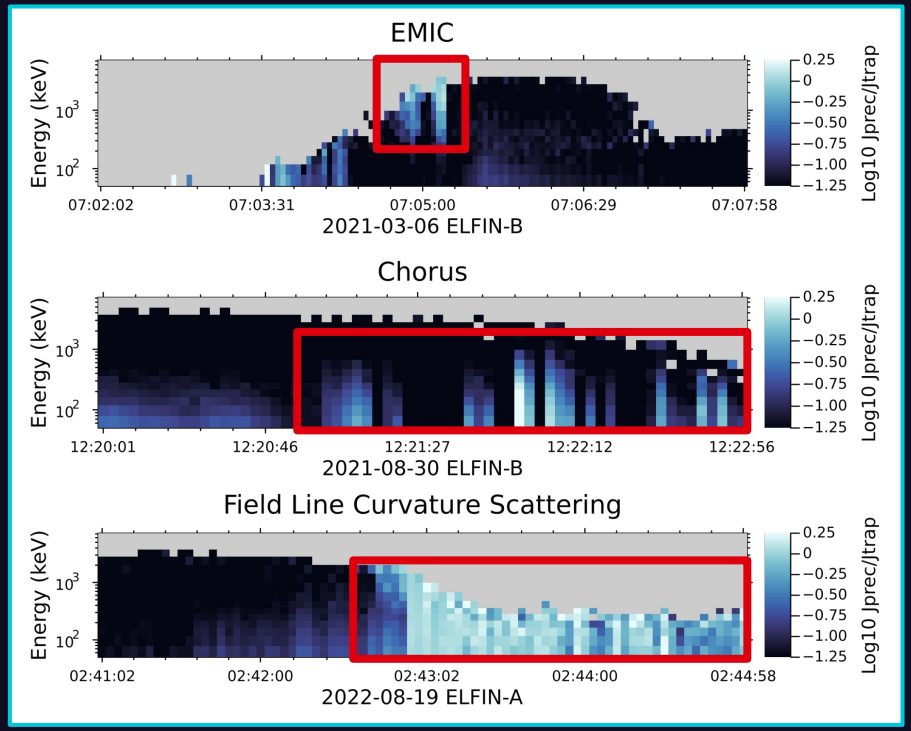
DECIPHERING EPP EVENTS

J_{Prec}/J_{Trap} Ratio

- Strength of precipitation relative to non-precipitating population
- Brighter spots indicate more precipitation

Unique Signature

- Different sources of EPP have unique data signatures
- Used to visually differentiate EPP events



IDENTIFYING EPP EVENTS

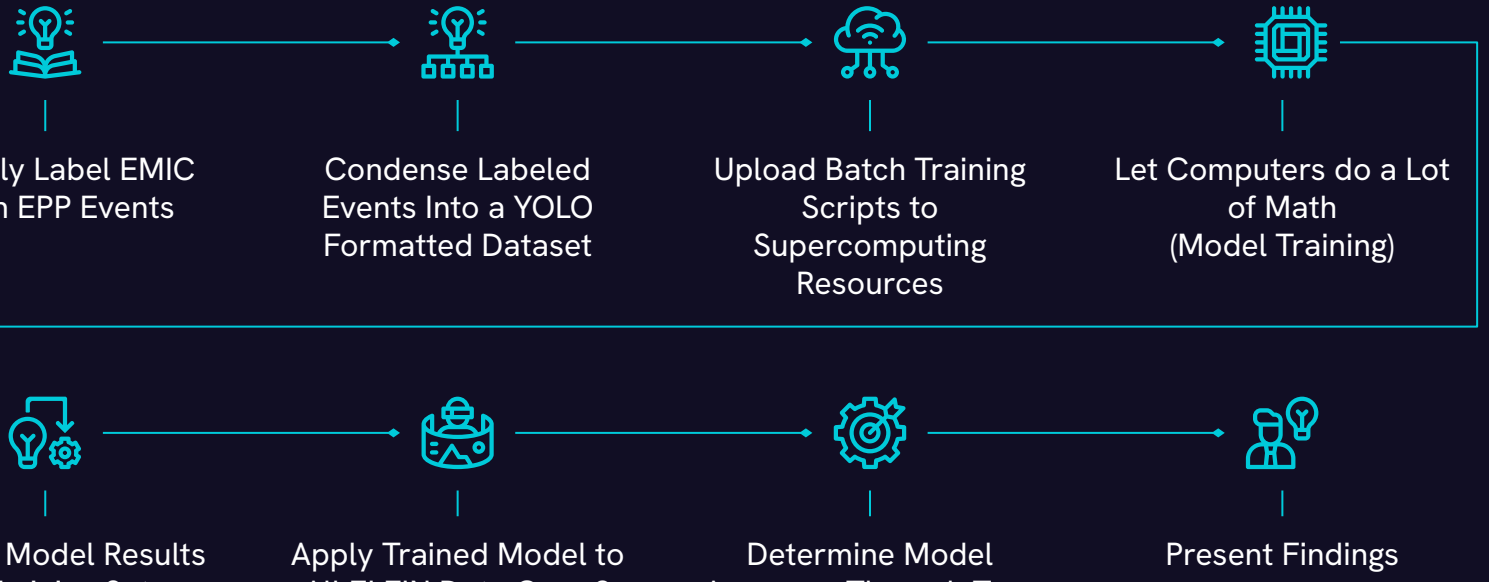
EMIC Waves

- Electromagnetic Ion Cyclotron Waves (EMIC)
- Elevated precipitation rates at higher energies which tapers off (teardrop shape)
- High community interest due to unique atmospheric impacts (more data!)

Computer Vision

- Easy to visually decipher various EPP events by source
- Computer Vision is a subset of artificial intelligence (and not just buzzwords to boost stock price)
- You Only Look Once (YOLO) CV model due to support and documentation

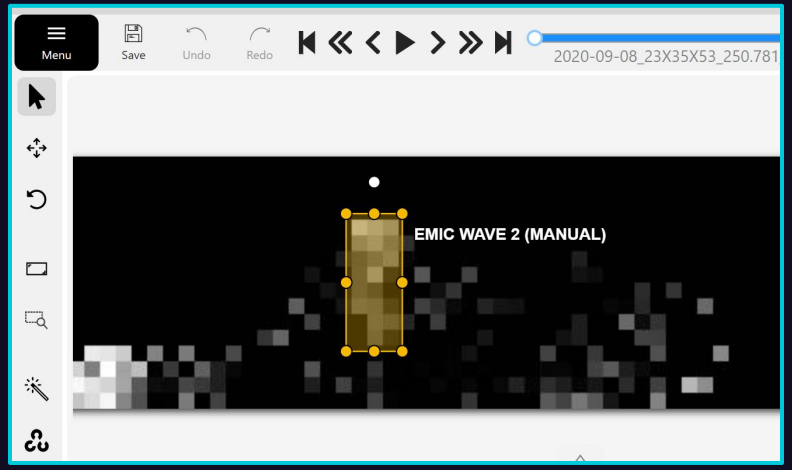
COMPUTER VISION METHODOLOGY



TRAINING SETS

First Dataset

- 142 Labeled Images
- 1 Type of EPP (EMIC Events Only)
- Higher error surrounding labels (less precise labeling)



* A screenshot of the image data labeling software we used

Final Dataset

- 306 Labeled Images
- 4 Types of EPP to Avoid Confounding Events
- Higher label precision

RESULTS

03

OUR MODELS

* Total # Detections found by applying model to lifetime of ELFIN data

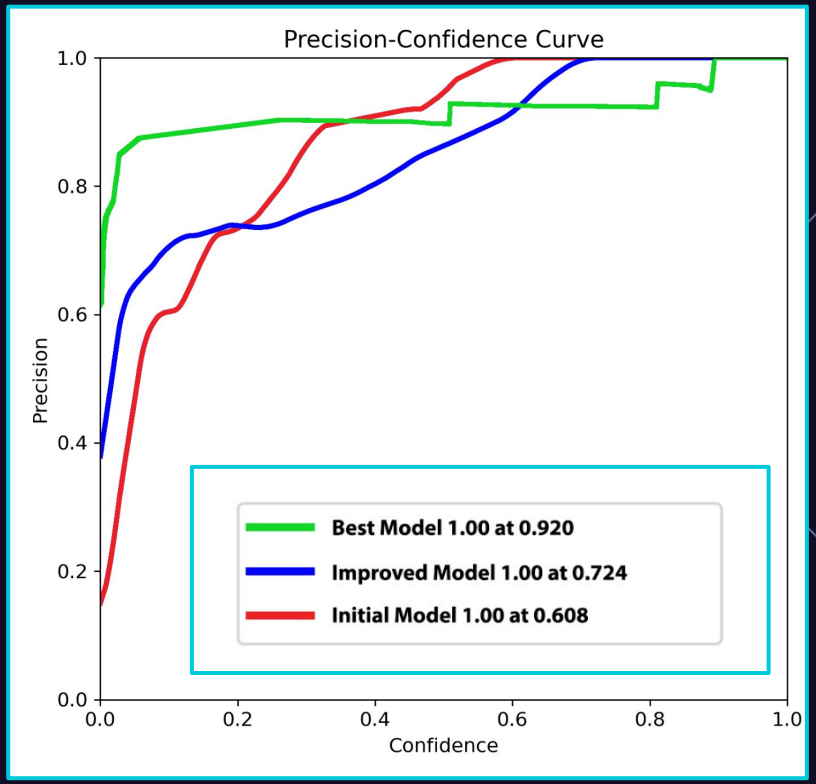
	Total Epochs	Total # Detections
1) Initial Model - Trained overnight on a laptop - YOLOv8s (2nd smallest YOLO Model) - Great initial performance	100	3837
2) Improved Model - Refined Training Techniques - Leveraged CU Supercomputer - YOLOv8x (largest YOLO model) - Still used our initial training set	100	1310
3) Final Model (Best) - Improved training set - Trained over days on supercomputer - YOLOv8x - Significantly improved performance due to handling of confounding EPP events in training set	344	786

MODEL PERFORMANCE

	Training Set Recall	Total # Detections
Best Model	73.6%	786
Improved Model	70.1%	1310
Initial Model	75%	3837

$$\text{Precision} = \frac{\text{Correct Detections}}{\text{Detections}}$$

$$\text{Training Set Recall} = \frac{\text{Correctly Detected Training Events}}{\text{Events in Training Set}}$$



SUPPORTING ANALYSIS

L-Shell:

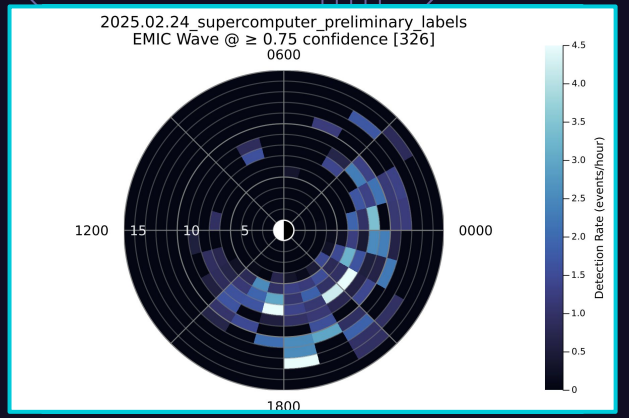
- "# Earth's radius from Earth"
- Indicated by the smaller sub circles inside of the graph

Magnetic Local Time:

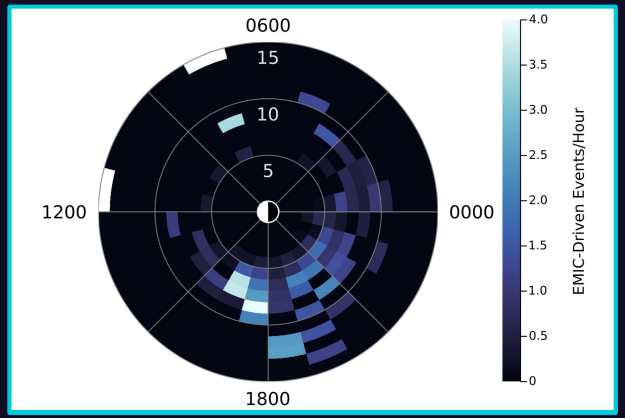
- Satellite's orientation relative to Earth and Sun at recording
- Noon (12:00) always points towards sun (24 hour clock!)

Comparing Location:

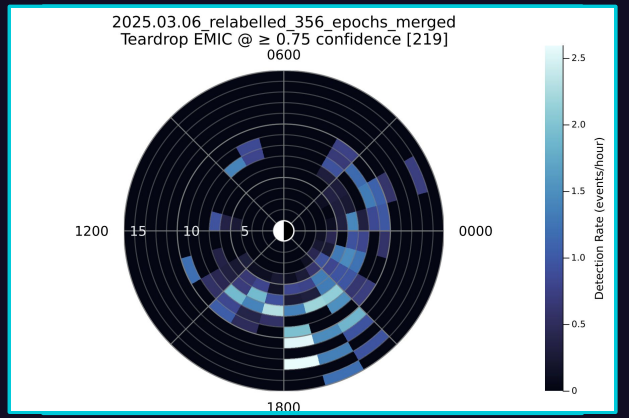
- Events of the same type tend to occur in the same area relative to Earth/Sun



First Model Detections



Manual Labels from Previous Work



Final Model Detections

CONCLUSION

04

THE FUTURE OF OUR RESEARCH

Improved Model Performance

- Our biggest limiting factor was training set size
- Larger training sets and experimenting with different algorithms may yield even more accurate models

Avoid Confounding False Positives

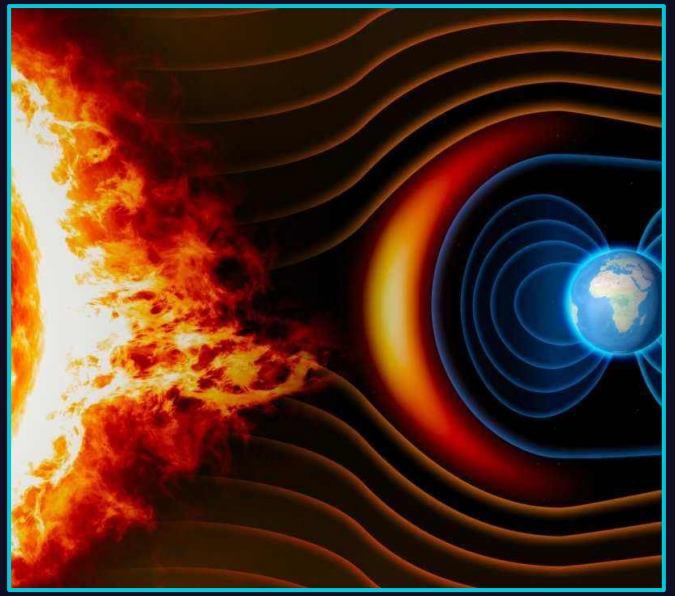
- Our second training set used 4 EPP classes instead of just 1
- We saw a drastic reduction in false positives due to confounding data
- Further dataset improvements may continue this trend

Impact on the Study of EPP

- Create large dataset to help find patterns in data and improve predictive models
- Save time and resources (to be used elsewhere!)

CONCLUSION

Our results indicates a promising future for the use of computer vision (and Artificial Intelligence) to identify EPP events in large datasets

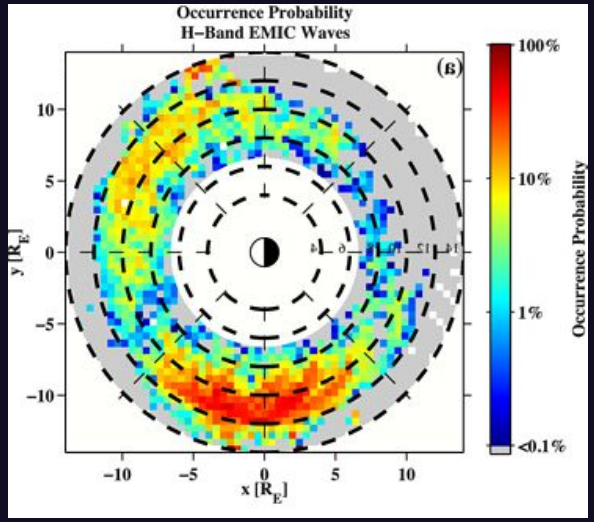


Contact

caleb.bettcher@colorado.edu

Special thank you to Julia Claxton for her guidance, teachings, and **patience!**

(emphasis on patience)

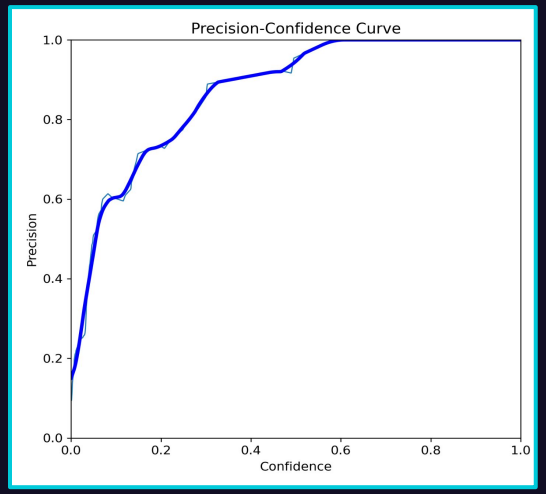

Occurrence Probability
H−Band EMIC Waves
(B)
100%
10%
1%
<0.1%
Occurrence Probability
y [R_E]
x [R_E]
10
5
0
−5
−10

Min+ 2012

Original Model Performance Slides

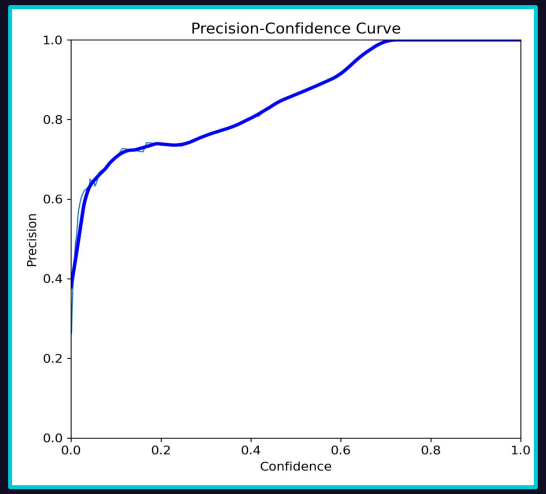
MODEL PERFORMANCE

EMIC Wave
all classes 1.00 at 0.608



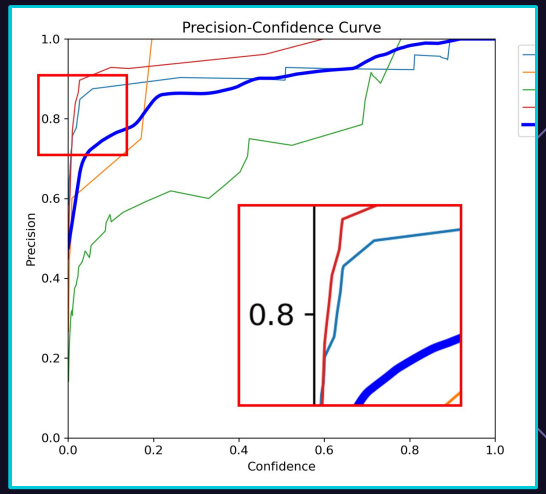
Initial Model

75%	3837
Training Set Recall	Total # Detections



Refined Training Techniques

70.1%	1310
Training Set Recall	Total # Detections



Updated Training Set

73.6%	786
Training Set Recall	Total # Detections

Training Set Recall:
Correctly Detected Training Events
Events in Training Set

Precision:
Correct Detections
Detections