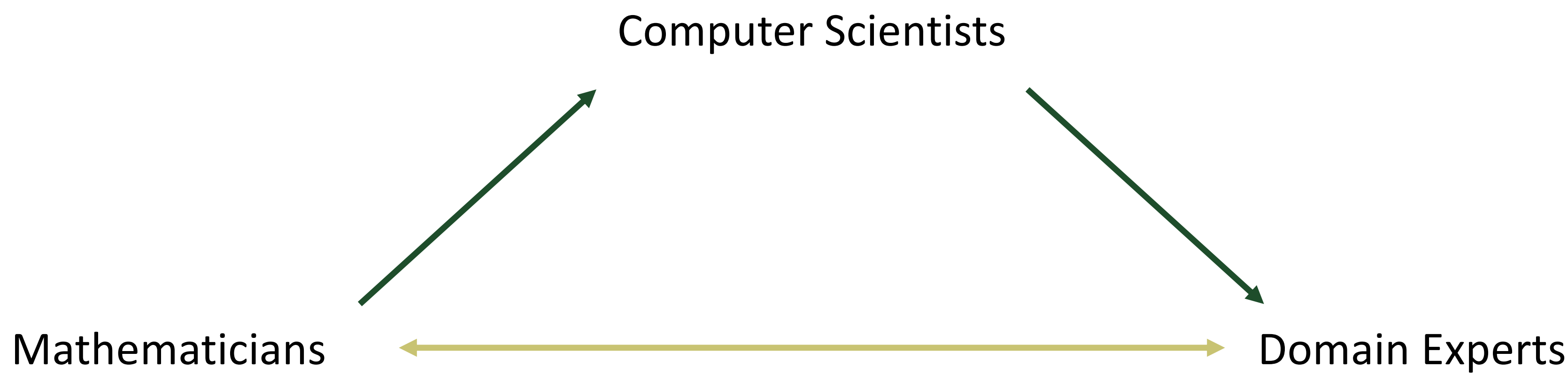
Using Mathematical Techniques to Leverage Domain Knowledge in Image Analysis for Environmental Science

PhD Defense for Lander Ver Hoef

Doctoral Committee:
Co-advisor: Henry Adams
Co-advisor: Emily J. King
Jess Ellis Hagman
Imme Ebert-Uphoff

February 23rd, 2023

Overall Theme



Projects we'll discuss

- Classifying mesoscale organization of clouds using topological data analysis
- Comparing rotationally invariant and conventional convolutional neural networks on storm forecast data
- Enhancing gravity waves in satellite imagery

Acknowledgements

- Many thanks to the following:
  - Imme Ebert-Uphoff
  - CIRA
  - AI2ES
  - DJ Gagne
  - NCAR, and MILES specifically

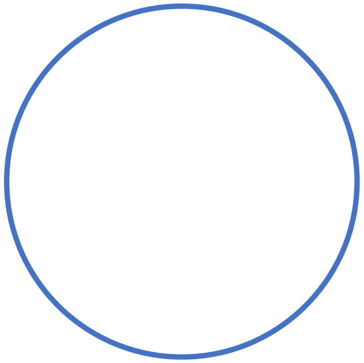
Project 1: A Primer on Topological Data Analysis to Support Image Analysis Tasks in Environmental Science

Mesoscale Cloud organization via TDA

- Paper has been published in Artificial Intelligence for the Earth Sciences, a journal of the American Meteorological Society
- Online early release paper is available at: https://doi.org/10.1175/AIES-D-22-0039.1
- Funding for this work was provided by the Strategic Funds for Machine Learning of the Cooperative Institute for Research in the Atmosphere at CSU, and by the National Science Foundation under Grant No. OAC-1934668 and Grant No. ICER-2019758 which funds the NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography.

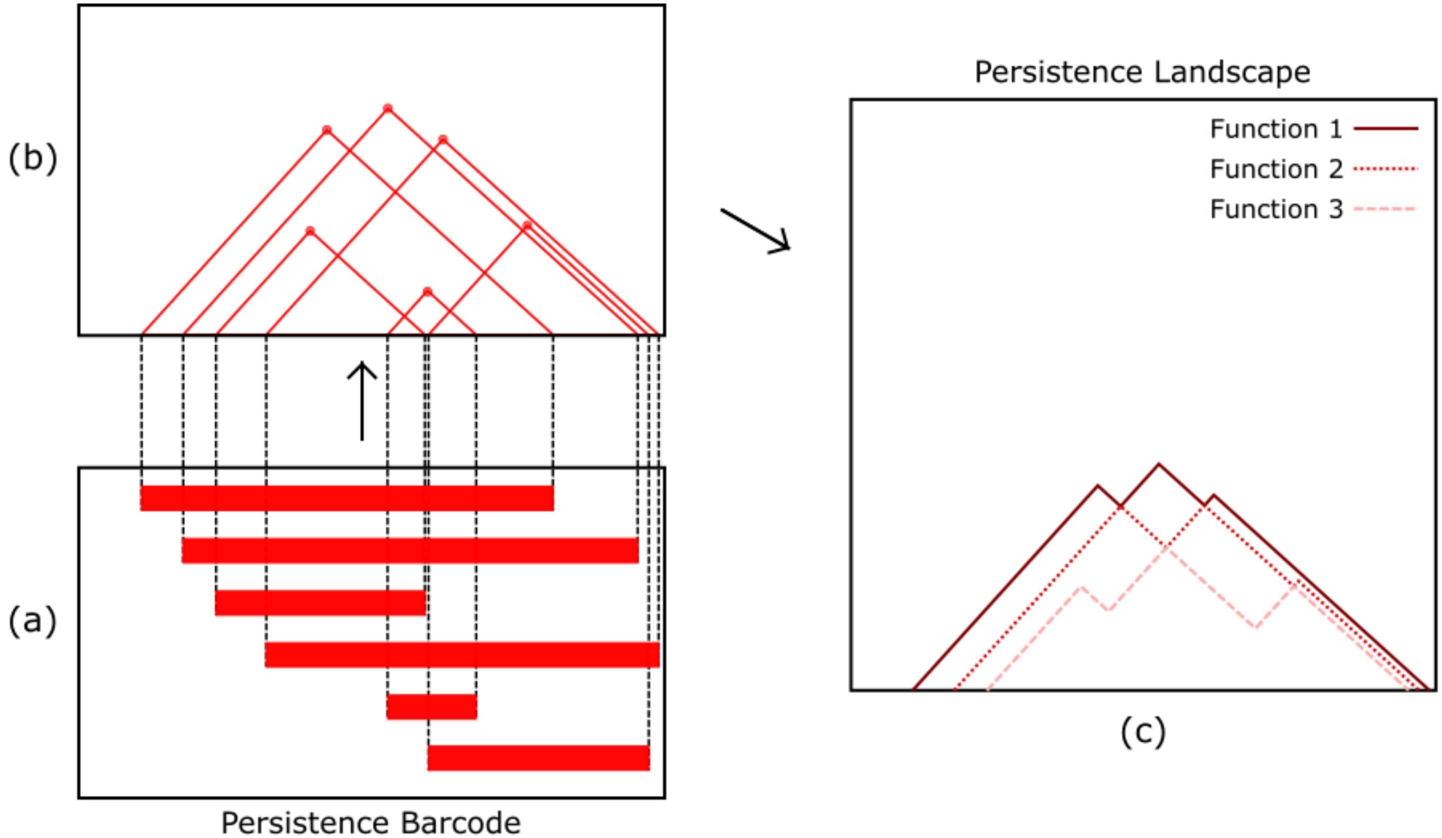
What is *homology*?

- Counts number of holes of each dimension, as well as connected components
- Can be computed efficiently
- Invariant to a broad class of transformations known as *homeomorphisms*

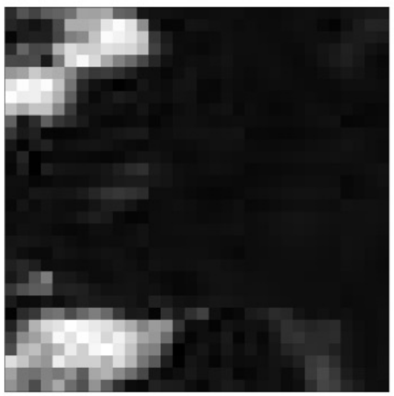


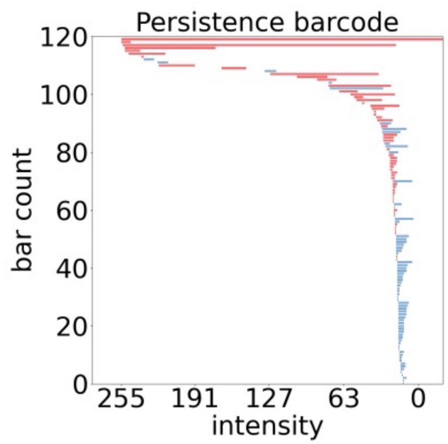
What is *persistent* homology?

What are persistence *landscapes?*

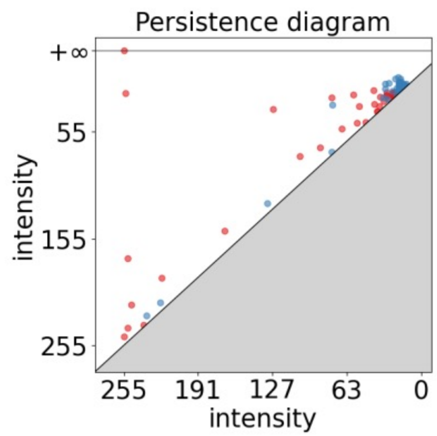


What does this look like on cloud satellite data?

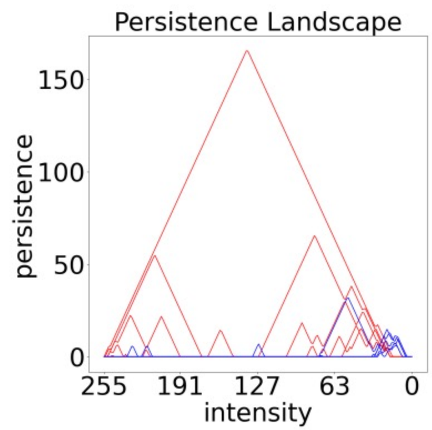

(a)



(b)



(c)



(d)

Mesoscale cloud organizations

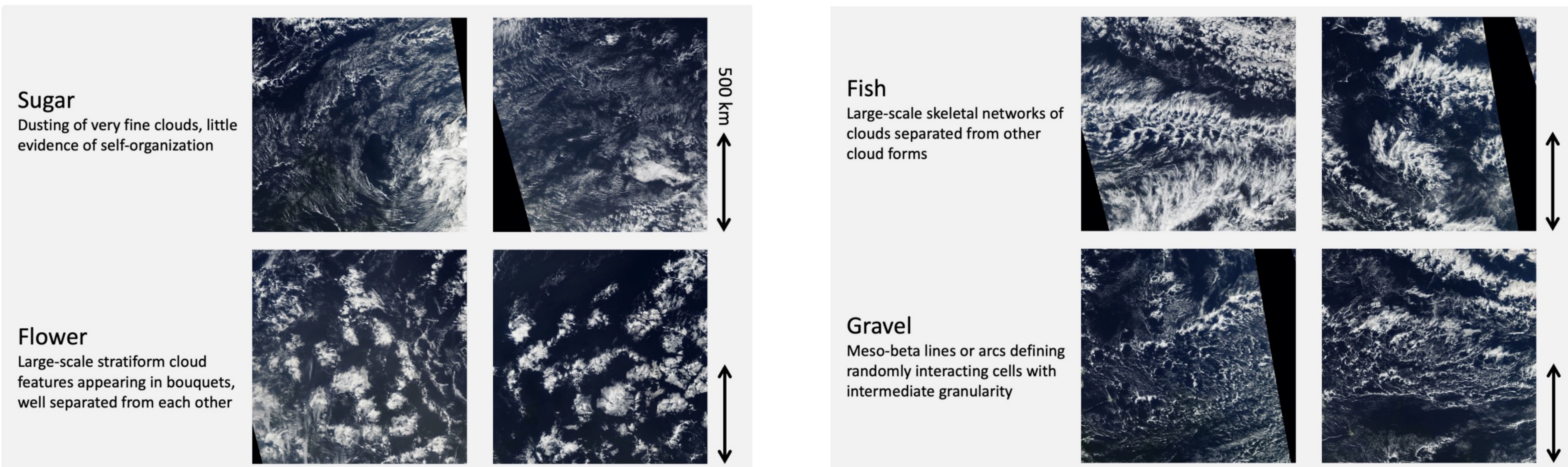
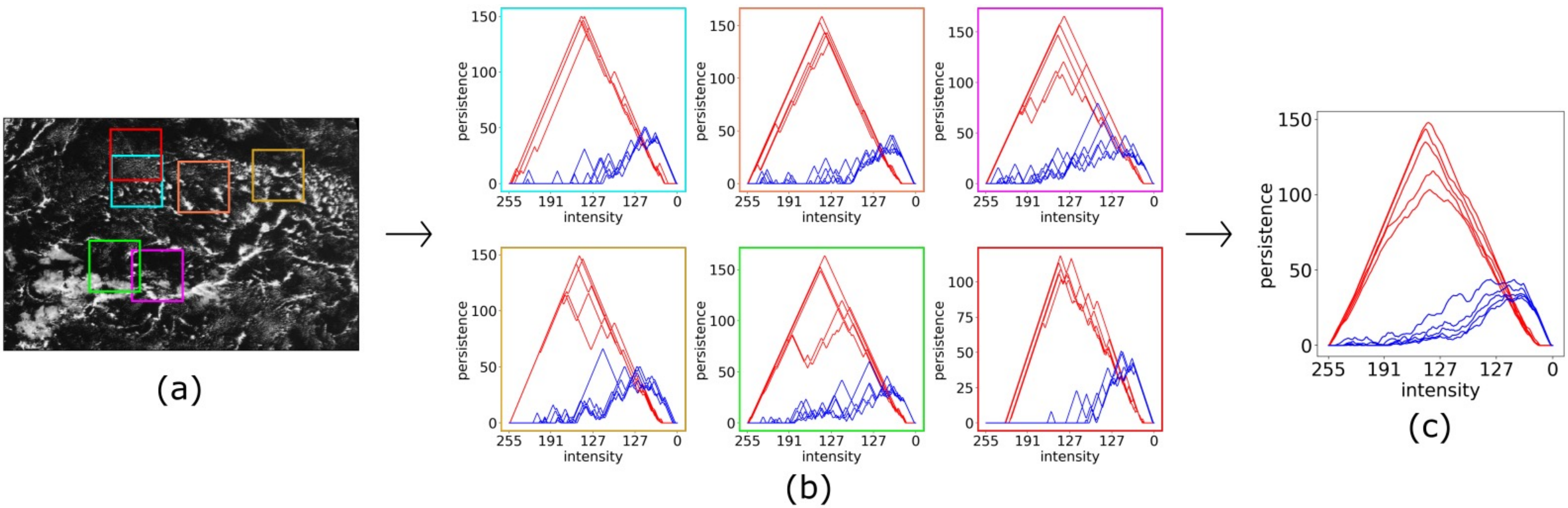
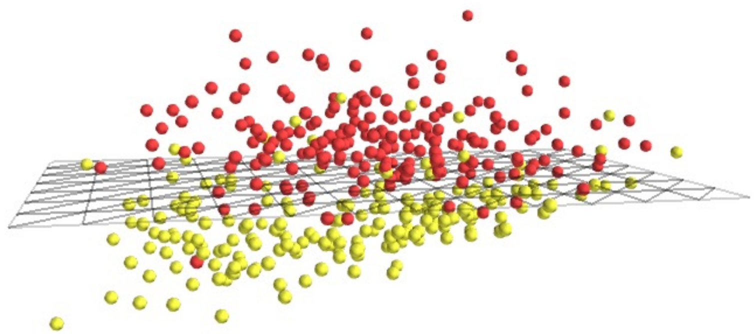


Image credit: Rasp, S., H. Schulz, S. Bony, and B. Stevens, 2020: Combining crowdsourcing and deep learning to explore the mesoscale organization of shallow convection. *Bulletin of the American Meteorological Society,* **101 (11)**, E1980-E1995.

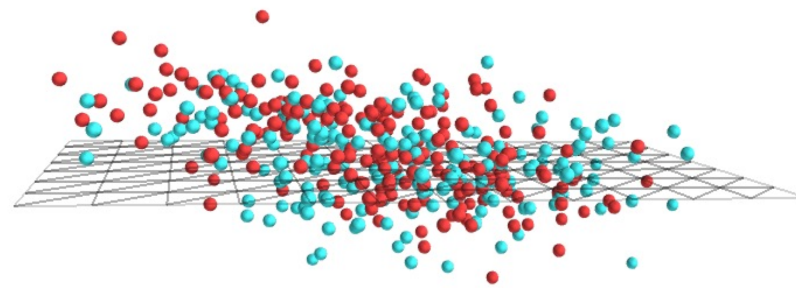
How did we apply TDA?



What sort of separations did we get?



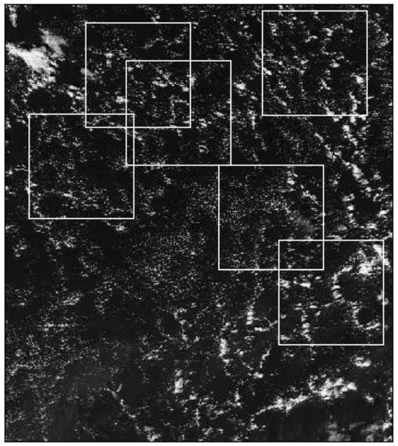
Sugar vs. flowers

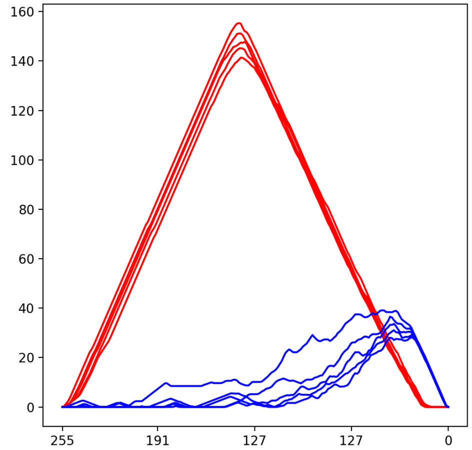


Fish vs. flowers

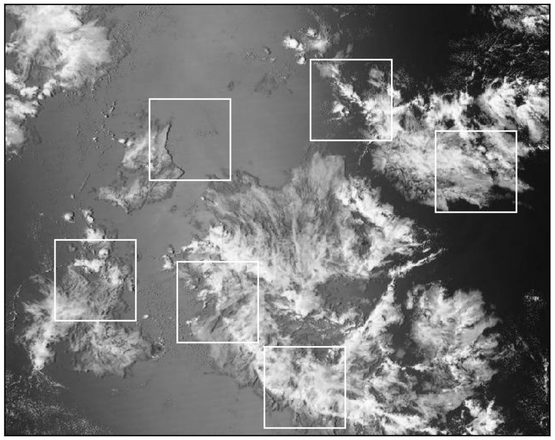
How can we analyze the sugar vs. flowers separation?

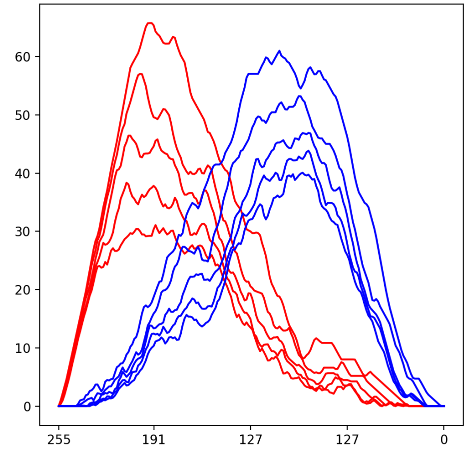
Most extreme
sugar example





Most extreme
flowers example





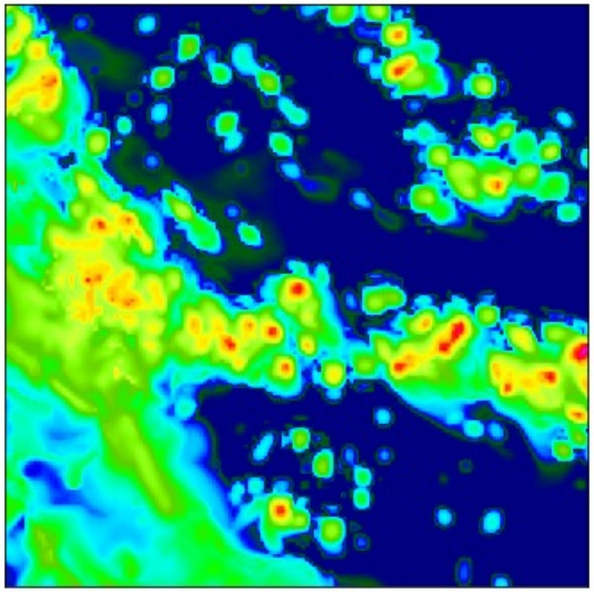
Project 2: Comparing Rotationally Invariant and Conventional CNNs

Storm Mode with Geometric Deep Learning

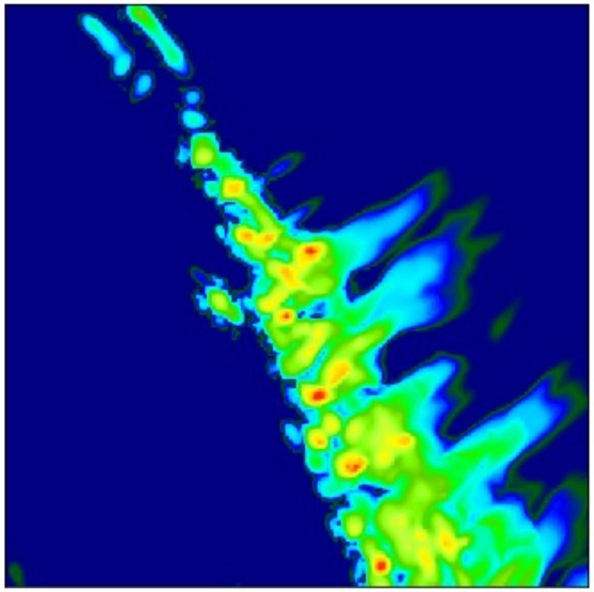
- Work done while a visitor at the National Center for Atmospheric Research, working with the Machine Integration and Learning for Earth Science (MILES) group under the supervision of Dr. David John Gagne.

- Funding for this work was provided the National Science Foundation under Grant No. OAC-1934668.

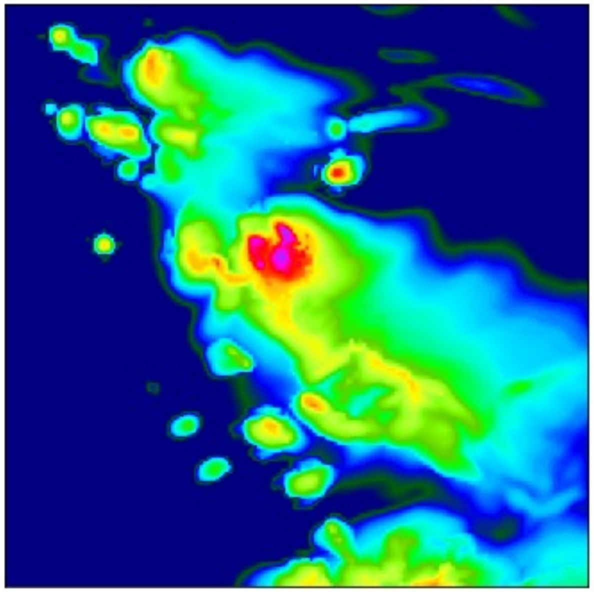
Storm Mode



Disorganized

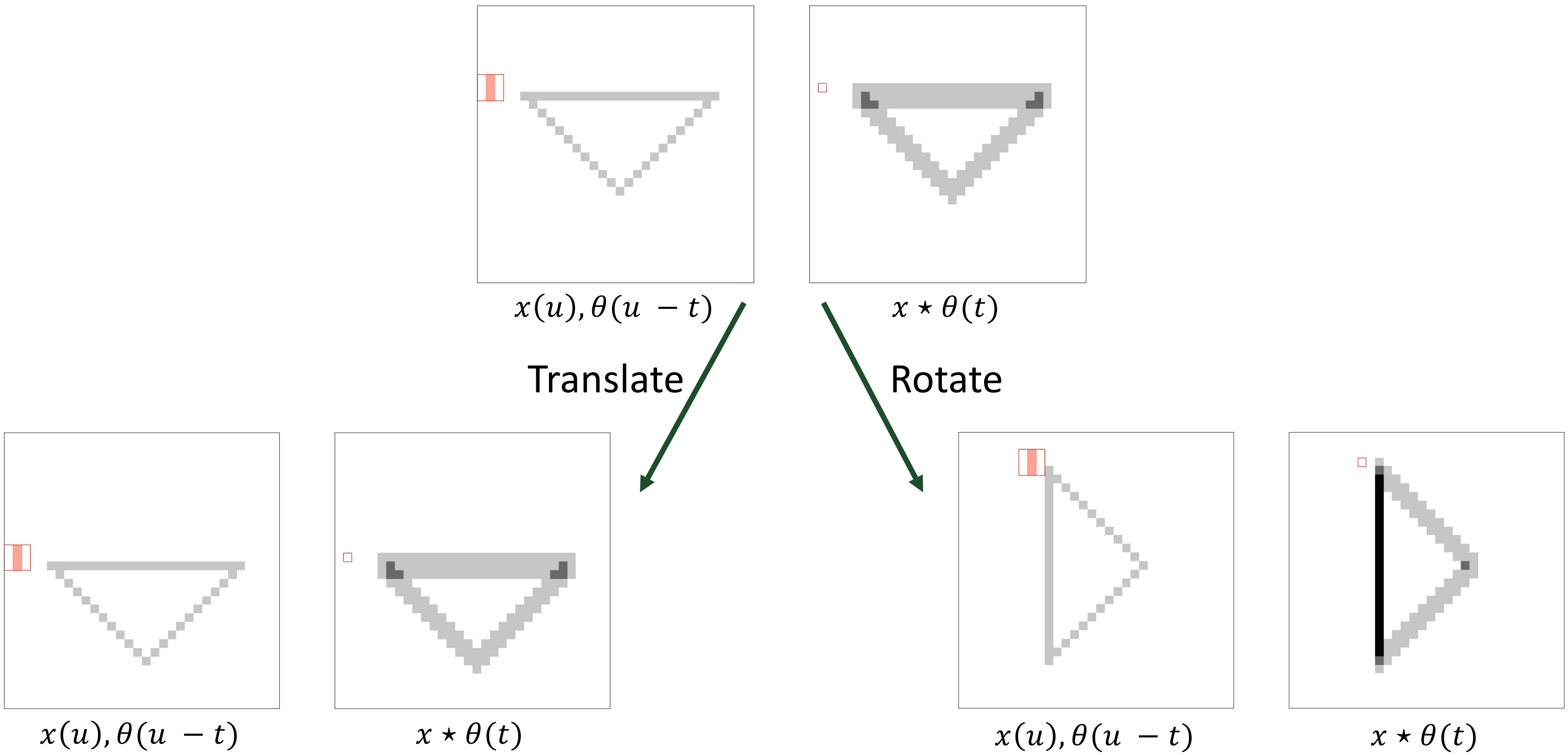


Quasi-linear System

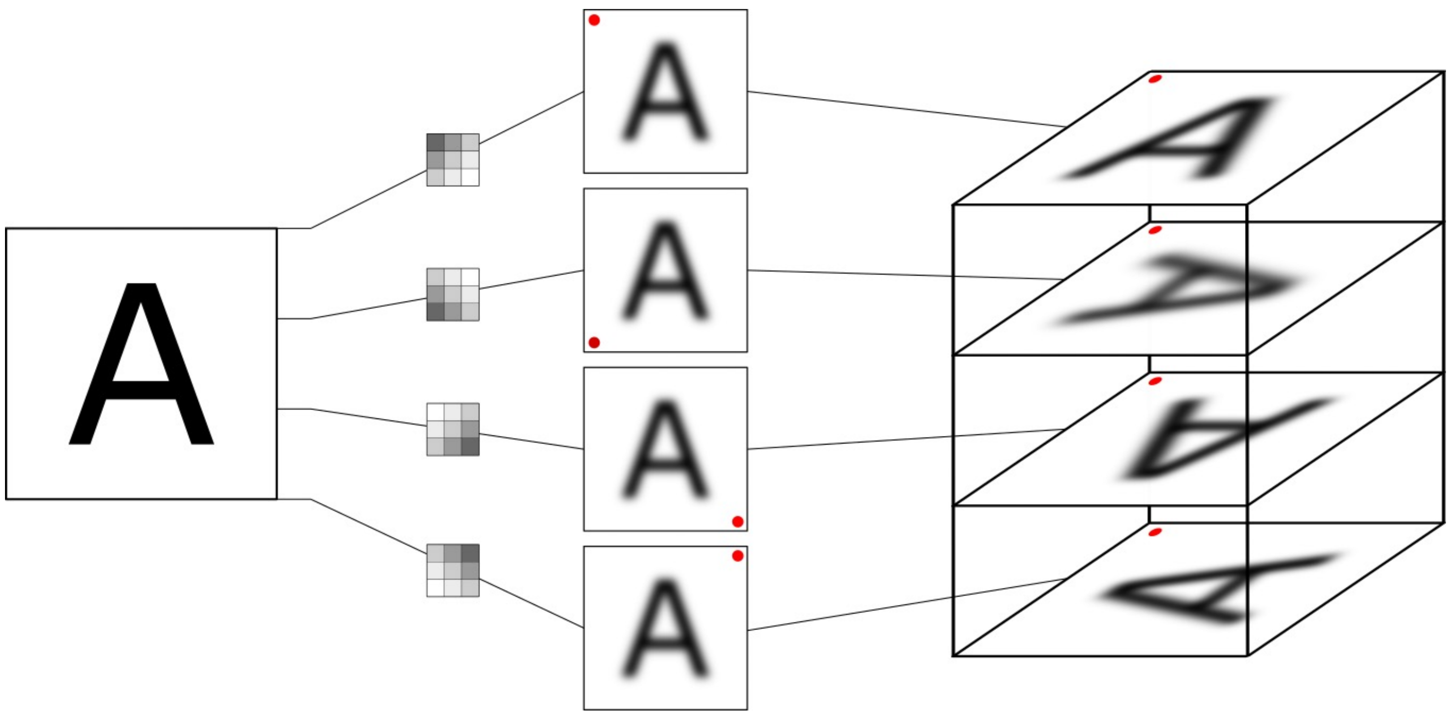


Supercell

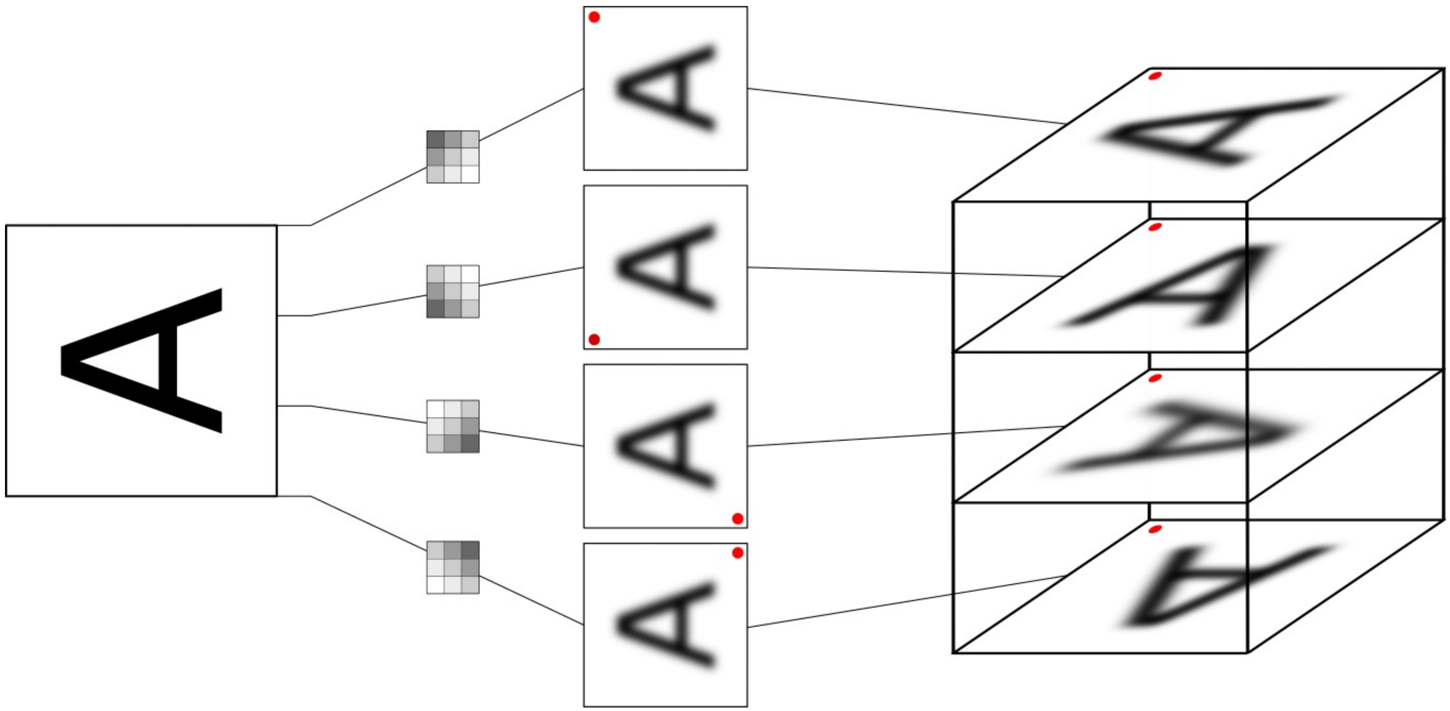
Translational vs Rotational Equivariance



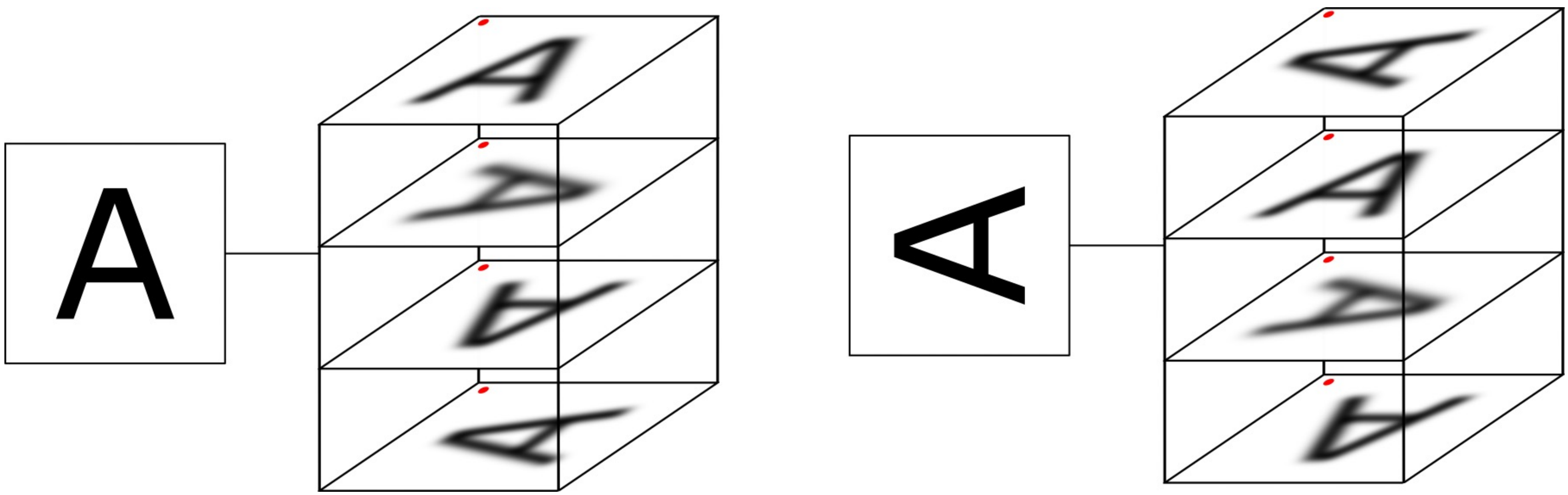
A Rotationally Equivariant Convolution (Initial Layer)



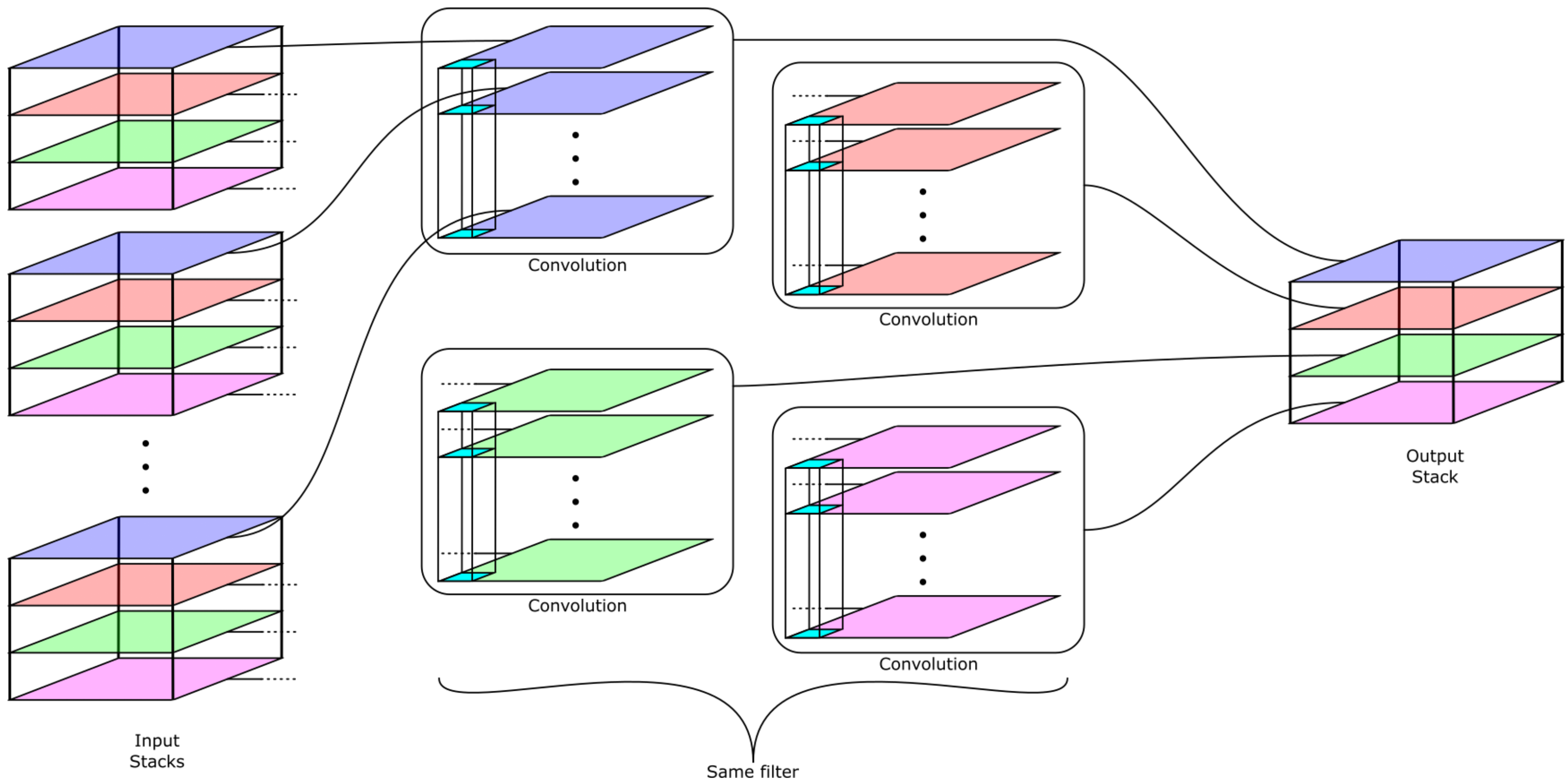
A Rotationally Equivariant Convolution (Initial Layer)



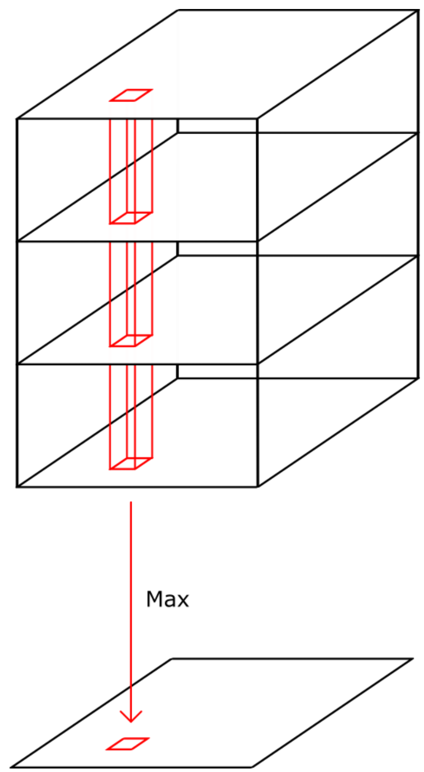
Equivariance Achieved



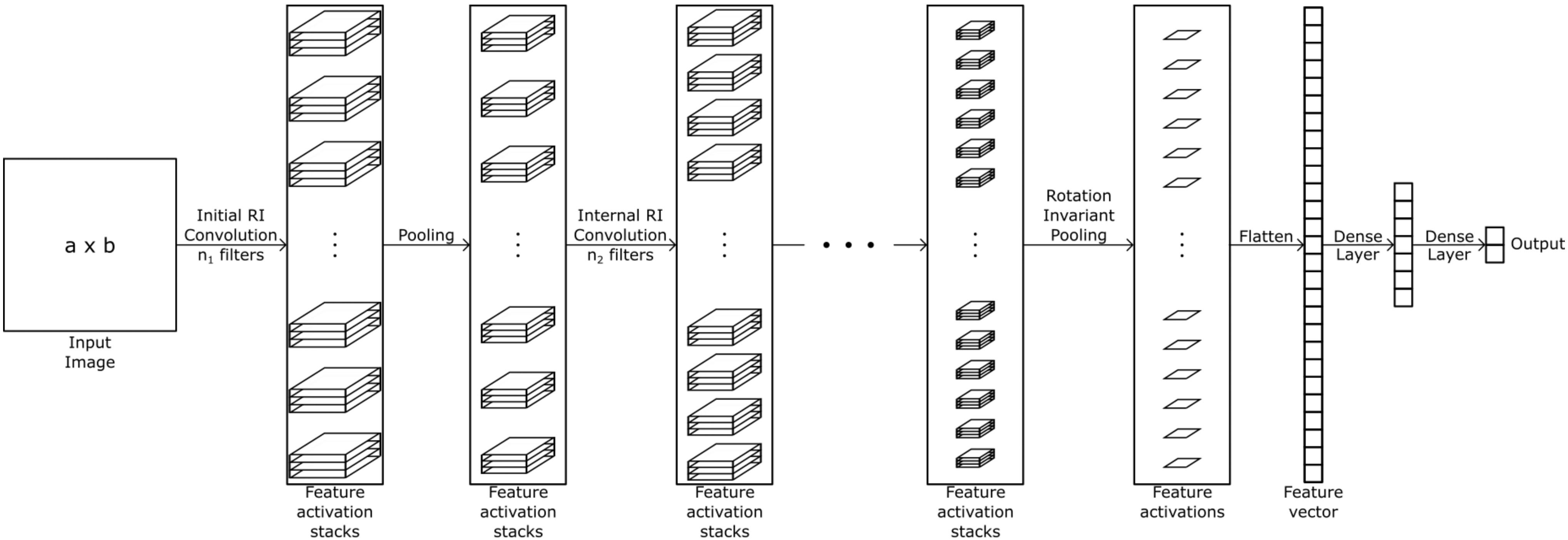
A Rotationally Equivariant Convolution (Internal Layer)



Rotational Invariant Pooling

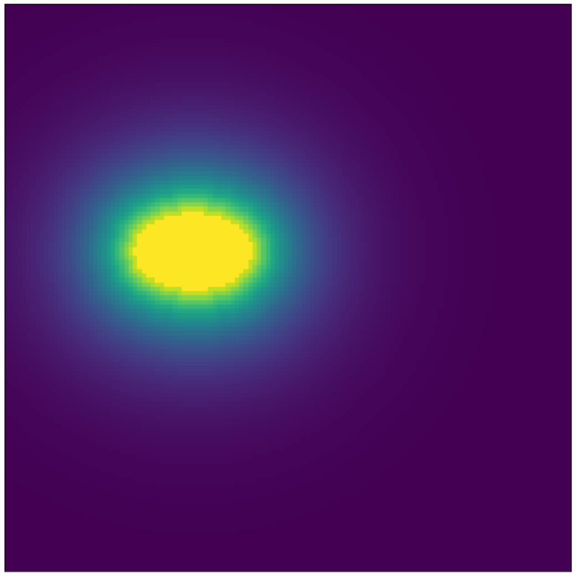


Rotationally Invariant CNN Architecture



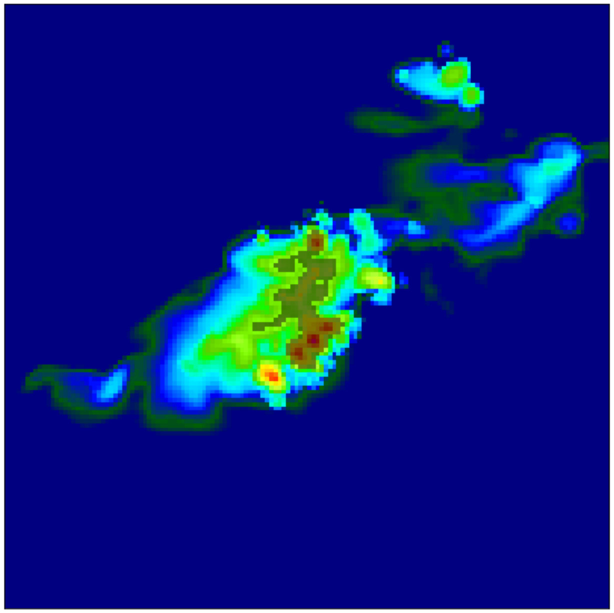
Rotational Invariance on Synthetic Dataset

Training ellipses all had major axes within $\frac{\pi}{6}$ of horizontal. Test ellipses had ellipses of any orientation.



	Final Training Error	Final Validation Error	Testing Error
CNN	0.0162	0.0196	0.7100
Aug. CNN	0.0082	0.0071	0.0075
RICNN	0.0111	0.0110	0.0141

Shifting to Storm Data



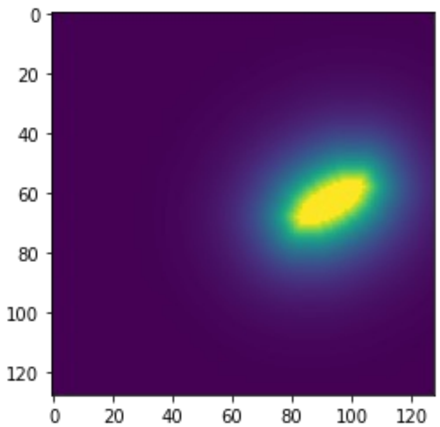
Initial Results

	Final Training Error	Final Validation Error	Testing Error
CNN	0.0365	0.0608	0.0705
Aug. CNN	0.0256	0.0278	0.0364
RICNN	0.0363	0.0409	0.0483

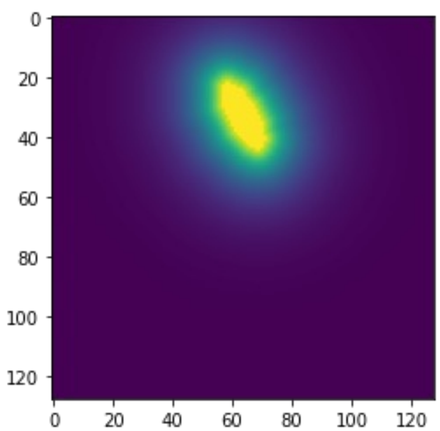
Separated on Storm Motion

	Final Training Error	Final Validation Error	Testing Error
CNN	0.0290	0.0684	0.0281
Aug. CNN	0.0304	0.0743	0.0333
RICNN	0.0317	0.0642	0.0255

Approximate vs True Invariance

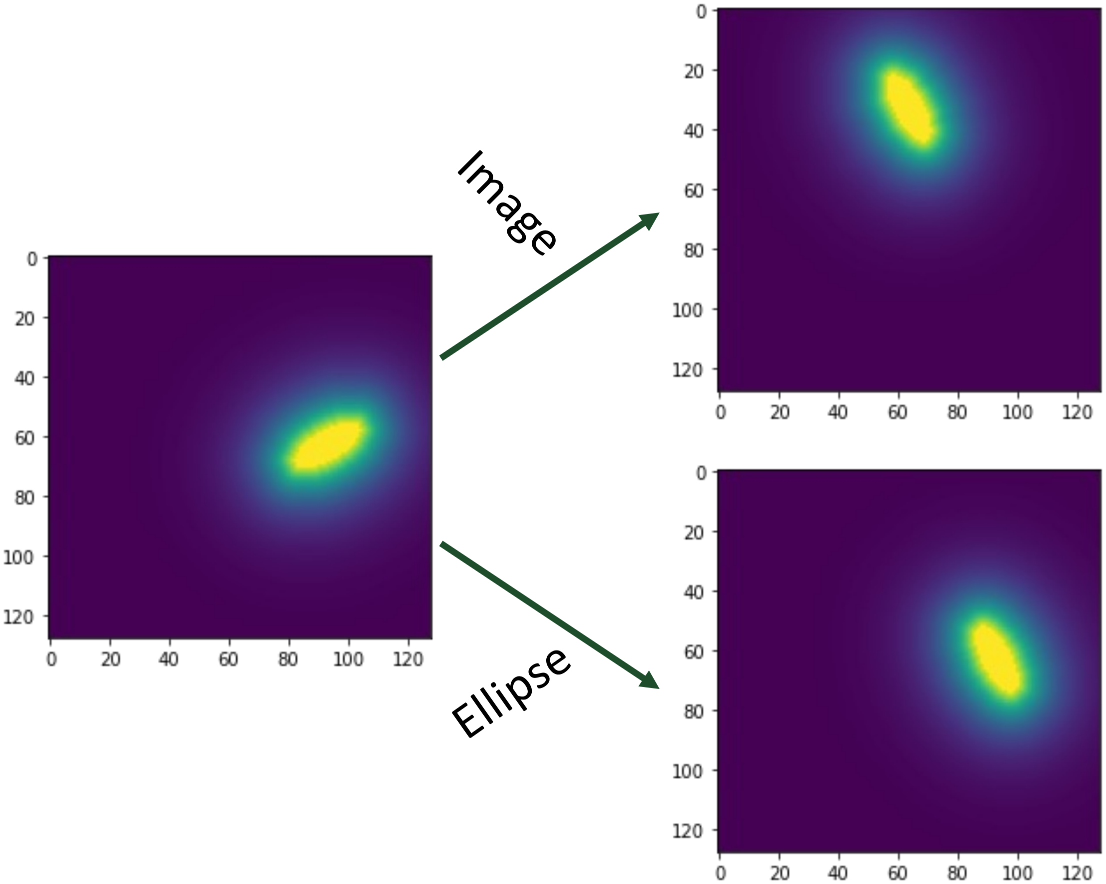


	Major Axis Len.	Minor Axis Len.
True	25	10
Aug. CNN	24.576	10.420
RICNN	23.982	9.855



	Major Axis Len.	Minor Axis Len.
True	25	10
Aug. CNN	25.444	10.005
RICNN	23.982	9.855

Approximate vs True Invariance



	Major Axis Var.	Minor Axis Var.
Aug. CNN	0.00379	0.00231
RICNN	0.00497	0.00228

	Major Axis Var.	Minor Axis Var.
Aug. CNN	0.00277	0.00277
RICNN	0.00407	0.00217

Project 3: Using Harmonic Analysis Techniques to Enhance Gravity Waves in the Day-Night Band

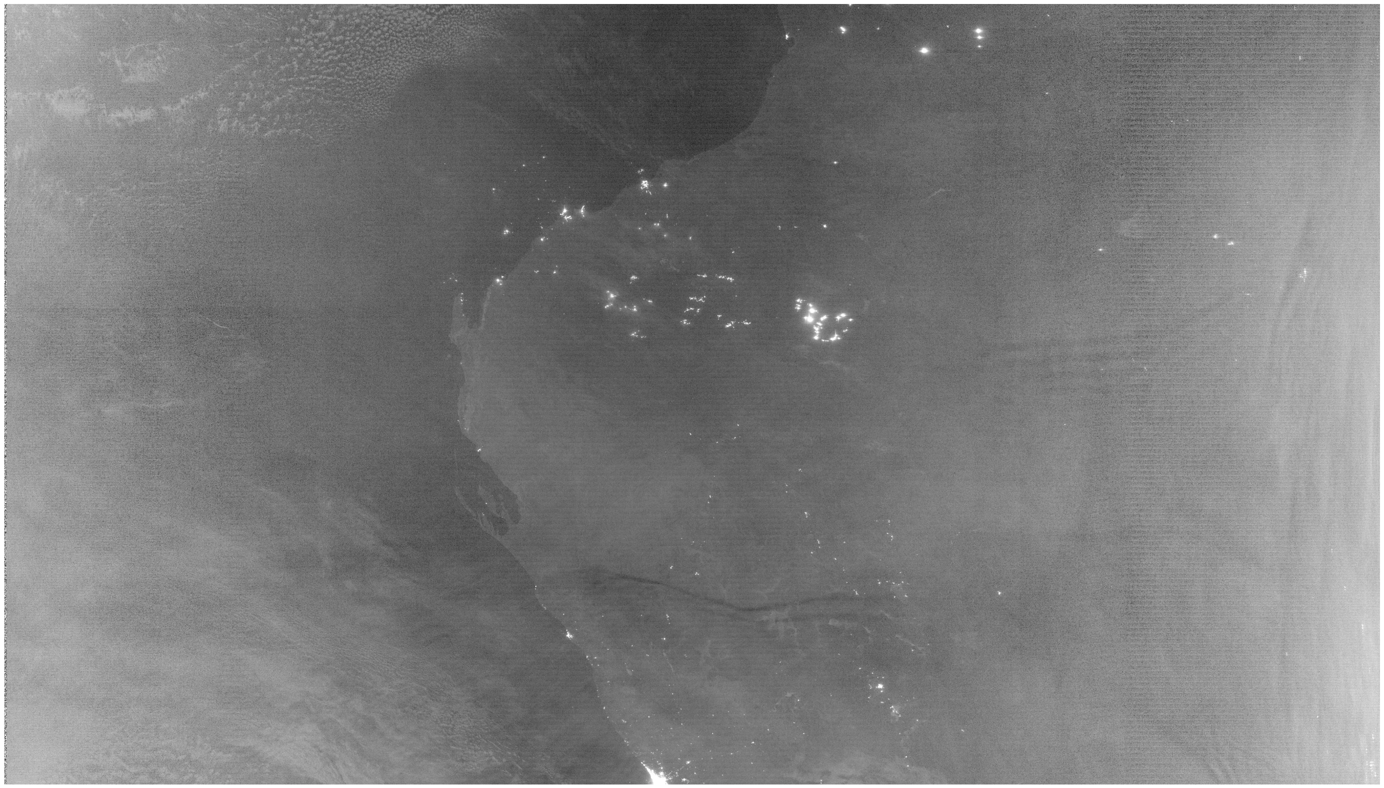
Gravity Waves in the Day-Night Band

- Work done in collaboration with researchers at the Cooperative Institute for Research in the Atmosphere (CIRA)
- Funding for this work was provided the National Science Foundation under Grant No. OAC-1934668.

Gravity Waves in the Day-Night Band

- Gravity waves are an important energy transfer mechanism, not modeled by large-scale climate models
- Gravity waves near the mesopause (~90 km above the surface) appear in nightglow
- This is detectable in data from the Day/Night Band (DNB) of the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor suites on the Suomi National Polar-orbiting Partnership (NPP) and NOAA-20 satellites

Gravity Waves Examples



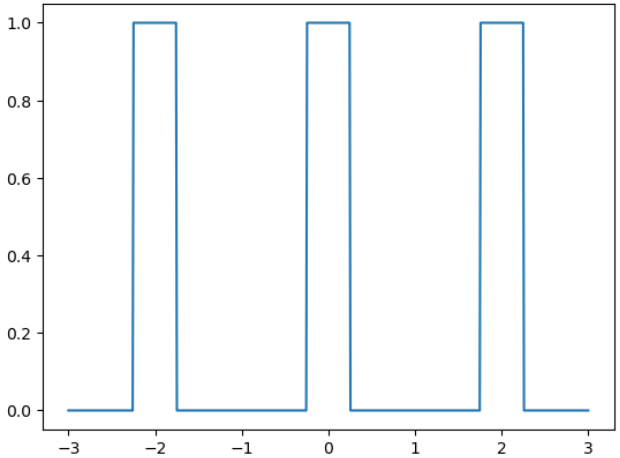
Structured Search

- Focus on two key properties:
  - Periodicity
  - Linearity
- We used three methods to try and enhance the gravity waves:
  - Local autocorrelation
  - Wavelet-based ridge detection
  - Finite Radon Transform

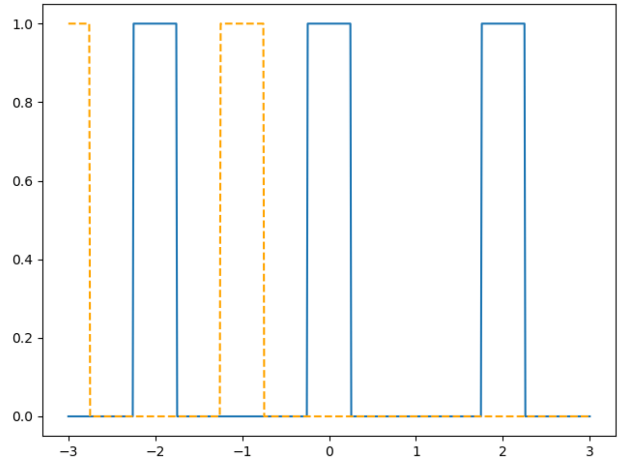
Local autocorrelation

- Based on autocorrelation, a technique for detecting periodicity in a signal
- Done locally on small patches to find local periodicity
- Current implementation is computationally costly, but could be sped up using the Fast Fourier Transform

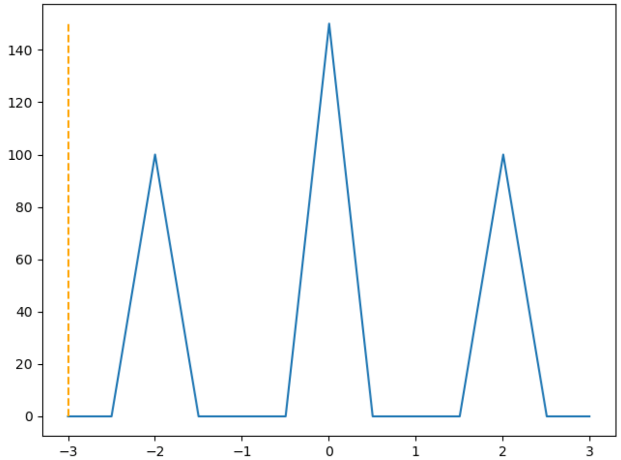
1D Autocorrelation



Original Signal

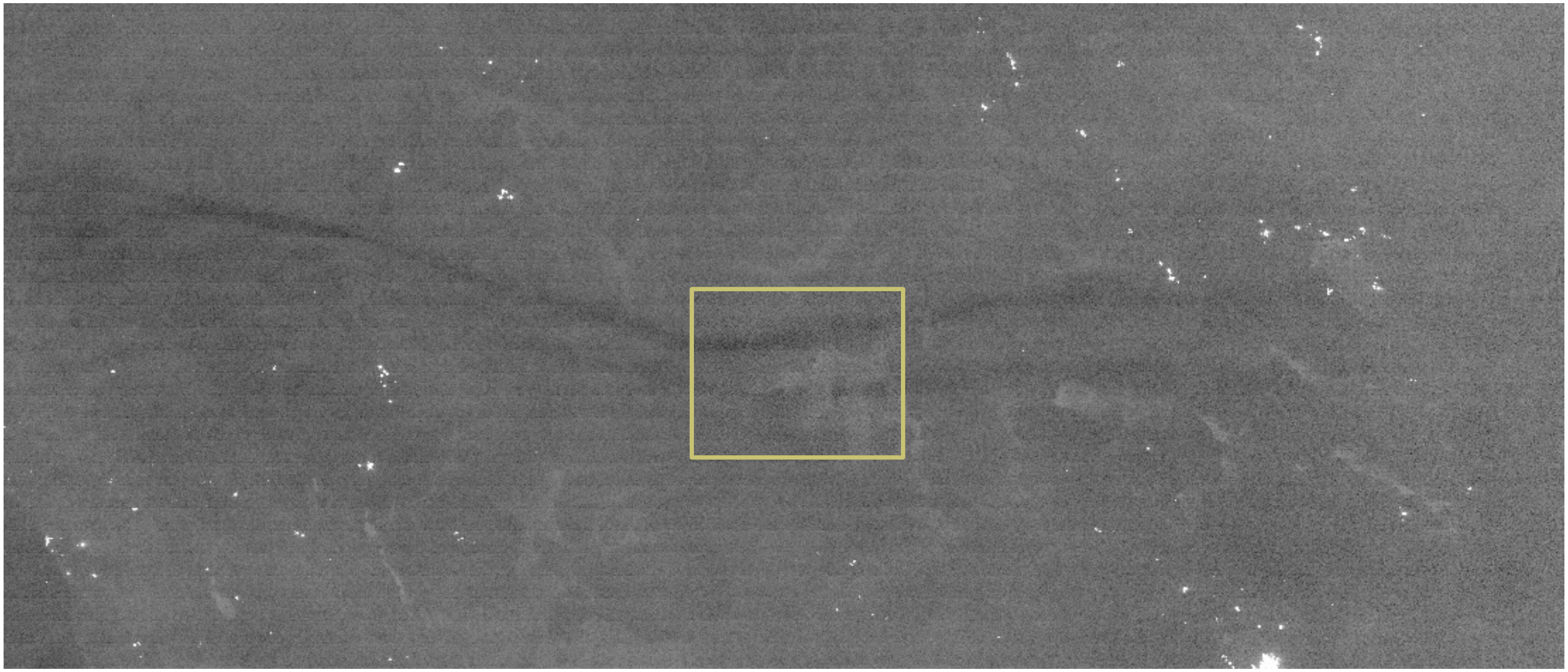


Original and shifted signal

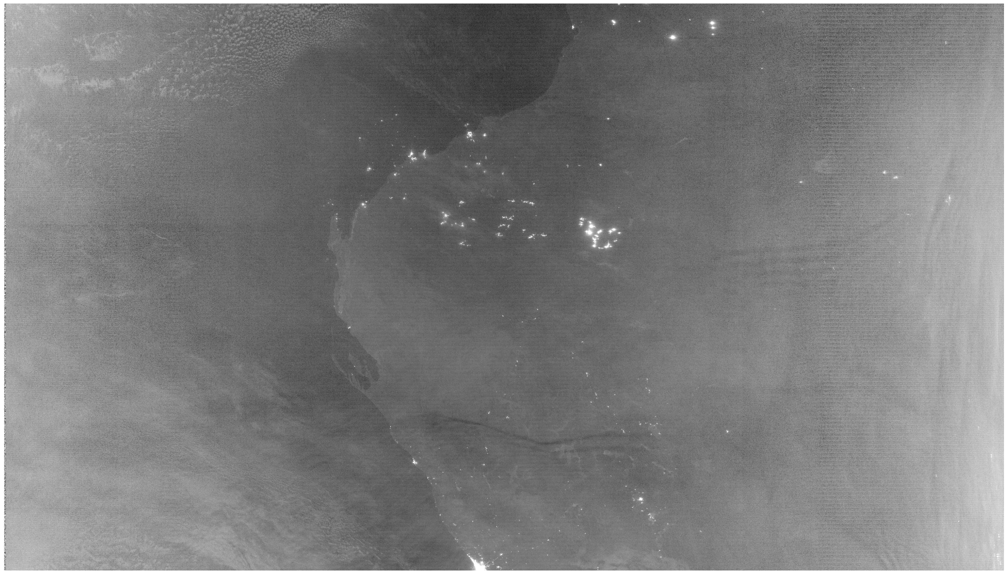


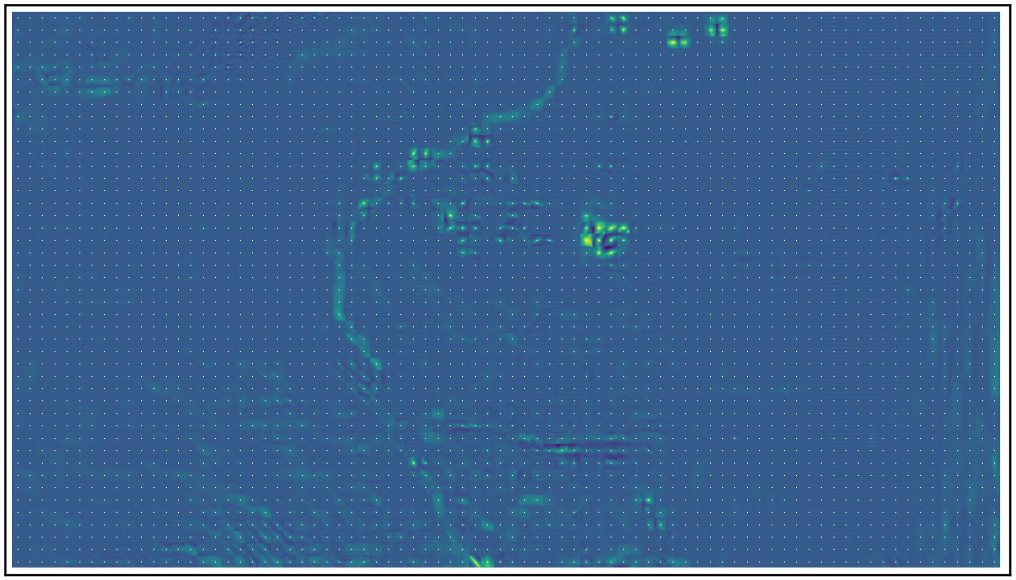
Autocorrelation

Local Autocorrelation



Local Autocorrelation Results



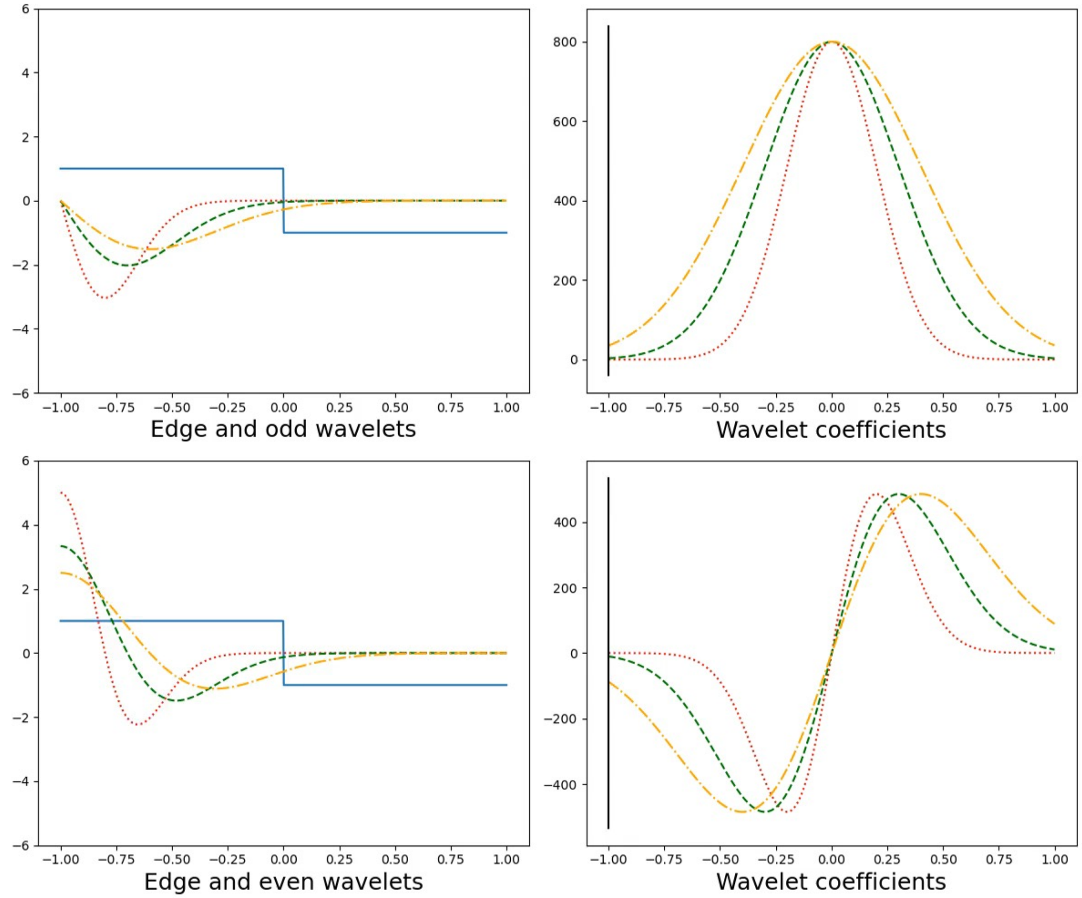


Wavelet-Based Ridge Detection

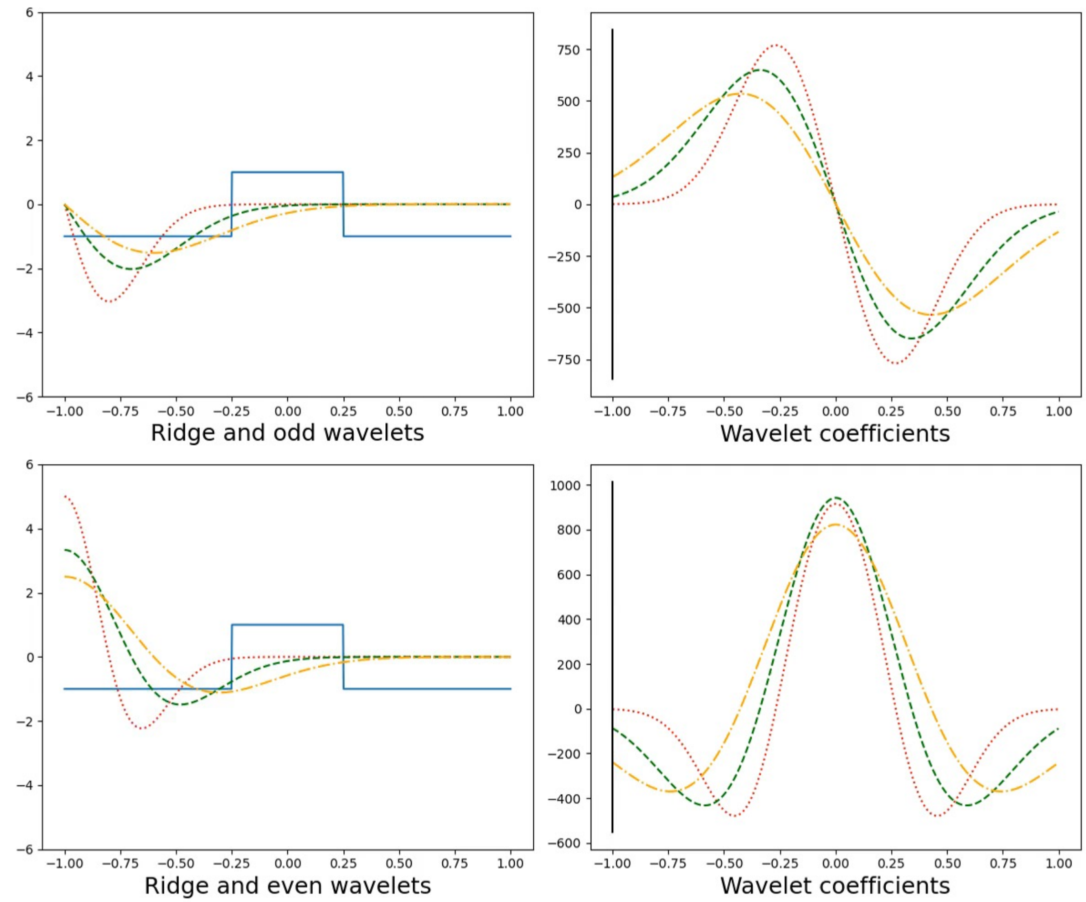
- Uses an algorithm developed by Dr. Raphael Reisenhofer[1]
- Gives at each location a measure of how "ridge-like" the local structure is, and the scale and orientation of that ridge
- Computed using a MATLAB package, SymFD, also developed by Reisenhofer

[1]Rafael Reisenhofer and Emily J. King. Edge, ridge, and blob detection with symmetric molecules. *SIAM Journal on Imaging Sciences*, 12(4):1585–1626, January 2019.

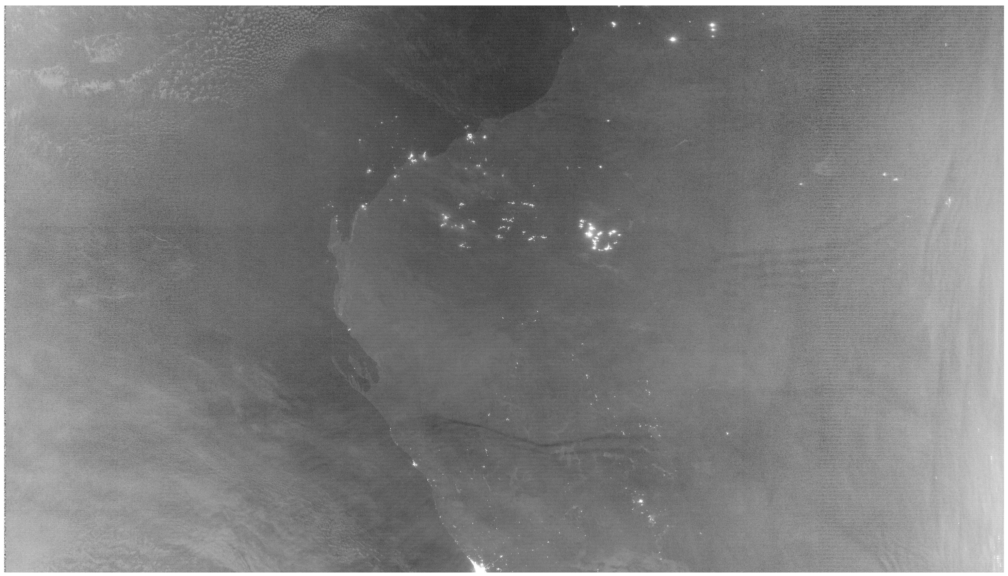
1D Edge Detection

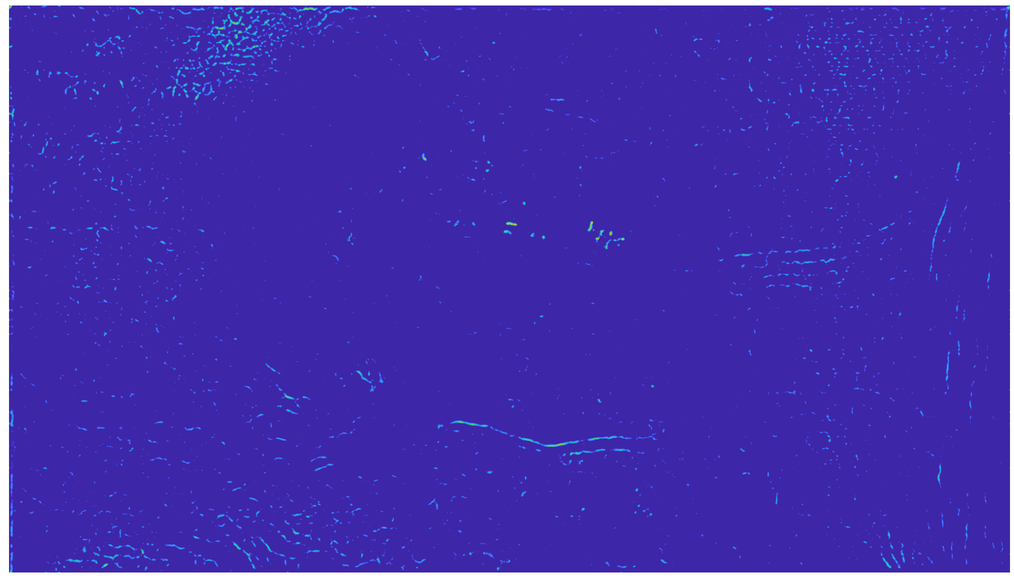


1D Ridge Detection



Ridge Detection Results





Radon Transform

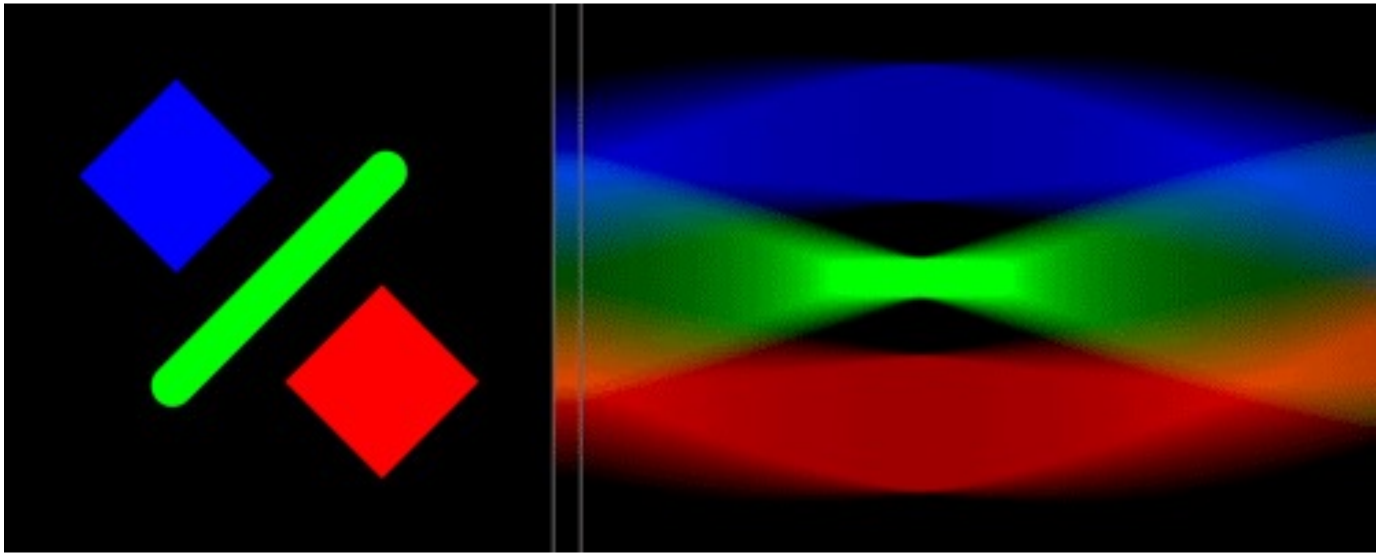
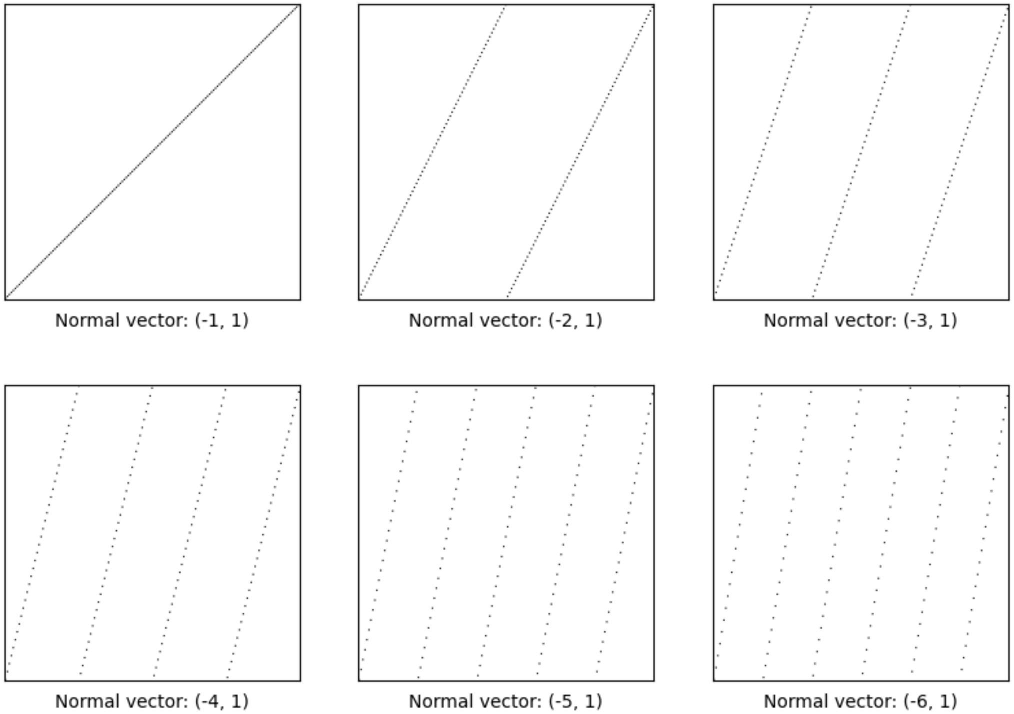


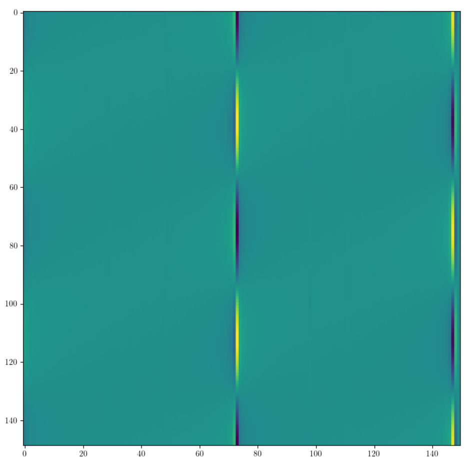
Image credit: LucasVB via Wikimedia Commons, at https://commons.wikimedia.org/wiki/File:Radon_transform_sinogram.gif. Public domain image.

Finite Radon Transform (FRT)

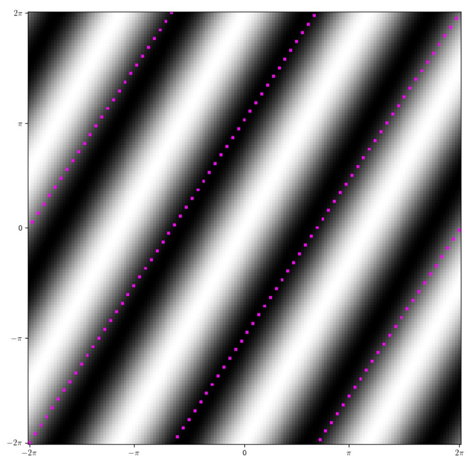
- Generalization of continuous Radon transform
- Relies on finite geometry structure of p x p patch for prime p
- Output is the sum along all finite lines through the space



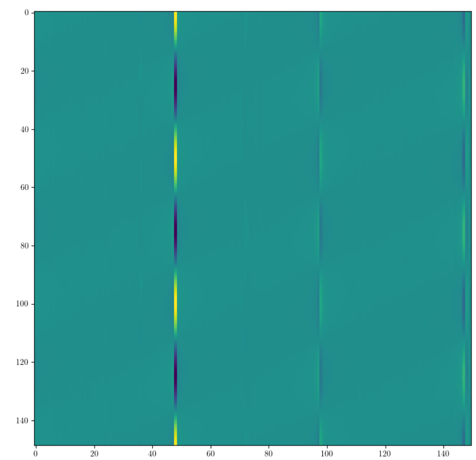
FRT on Synthetic Data



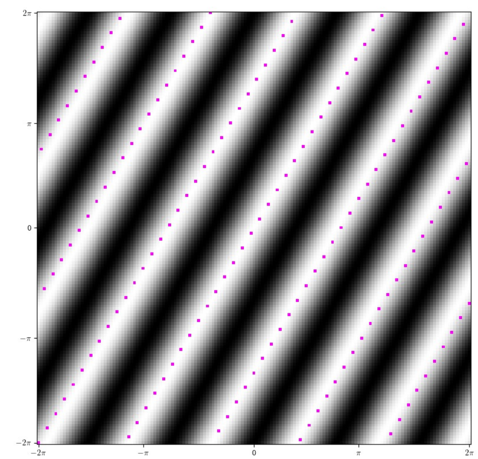
Example 1
Sinogram



Example 1
Image and example line

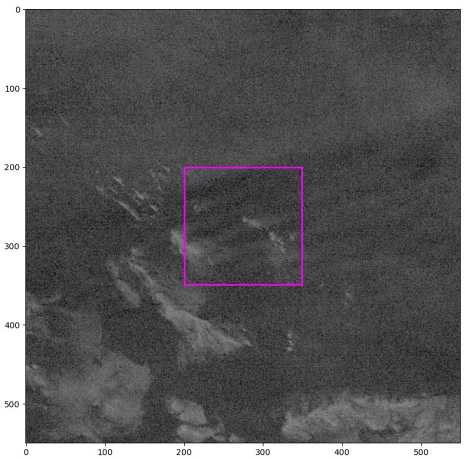


Example 2
Sinogram

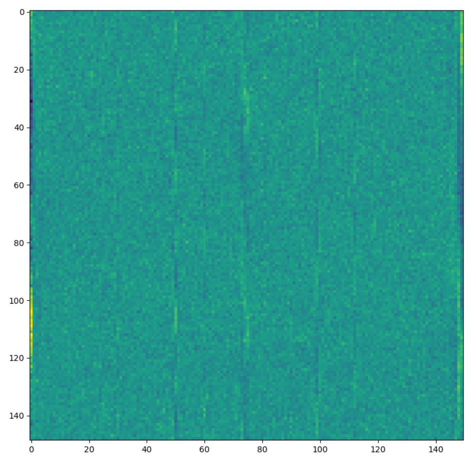


Example 2
Image and example line

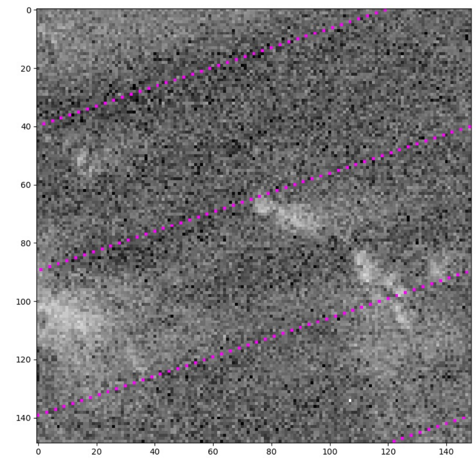
FRT on Real Data



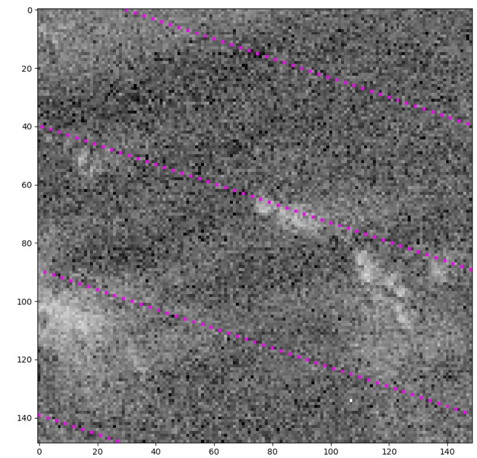
Patch with context



Sinogram



Patch with example line 1



Patch with example line 2

Conclusions and Future Work

- None of these algorithms fully utilize the concepts of linearity and periodicity, although all show some promise
- Two ideas to extend this work:
  - Combine ridge detection and local autocorrelation
  - Mojette transform

Conclusions

Conclusions

- Three mathematically very different topics
- Common application area with similar data types
- Most interesting theme: as a mathematician, working closely with domain experts and the data to inform algorithmic choices is vital

What questions do you have?

Thank you!