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

Mathematicians → Computer Scientists → Domain Experts

Mathematicians ← Computer Scientists ← Domain Experts
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?
Mesoscale cloud organizations

How did we apply TDA?

(a) → (b) → (c)
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
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

Translate

Rotate

\[ x(u, \theta(u - t)) \]

\[ x \star \theta(t) \]
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.

<table>
<thead>
<tr>
<th></th>
<th>Final Training Error</th>
<th>Final Validation Error</th>
<th>Testing Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>0.0162</td>
<td>0.0196</td>
<td>0.7100</td>
</tr>
<tr>
<td>Aug. CNN</td>
<td>0.0082</td>
<td>0.0071</td>
<td>0.0075</td>
</tr>
<tr>
<td>RICNN</td>
<td>0.0111</td>
<td>0.0110</td>
<td>0.0141</td>
</tr>
</tbody>
</table>
### Initial Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Final Training Error</th>
<th>Final Validation Error</th>
<th>Testing Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>0.0365</td>
<td>0.0608</td>
<td>0.0705</td>
</tr>
<tr>
<td>Aug. CNN</td>
<td>0.0256</td>
<td>0.0278</td>
<td>0.0364</td>
</tr>
<tr>
<td>RICNN</td>
<td>0.0363</td>
<td>0.0409</td>
<td>0.0483</td>
</tr>
</tbody>
</table>

### Separated on Storm Motion

<table>
<thead>
<tr>
<th>Model</th>
<th>Final Training Error</th>
<th>Final Validation Error</th>
<th>Testing Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>0.0290</td>
<td>0.0684</td>
<td>0.0281</td>
</tr>
<tr>
<td>Aug. CNN</td>
<td>0.0304</td>
<td>0.0743</td>
<td>0.0333</td>
</tr>
<tr>
<td>RICNN</td>
<td>0.0317</td>
<td>0.0642</td>
<td>0.0255</td>
</tr>
</tbody>
</table>
## Approximate vs True Invariance

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True</strong></td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>Aug. CNN</td>
<td>24.576</td>
<td>10.420</td>
</tr>
<tr>
<td>RICNN</td>
<td>23.982</td>
<td>9.855</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True</strong></td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>Aug. CNN</td>
<td>25.444</td>
<td>10.005</td>
</tr>
<tr>
<td>RICNN</td>
<td>23.982</td>
<td>9.855</td>
</tr>
</tbody>
</table>
## Approximate vs True Invariance

<table>
<thead>
<tr>
<th></th>
<th>Major Axis Var.</th>
<th>Minor Axis Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug. CNN</td>
<td>0.00379</td>
<td>0.00231</td>
</tr>
<tr>
<td>RICNN</td>
<td>0.00497</td>
<td>0.00228</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Major Axis Var.</th>
<th>Minor Axis Var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug. CNN</td>
<td>0.00277</td>
<td>0.00277</td>
</tr>
<tr>
<td>RICNN</td>
<td>0.00407</td>
<td>0.00217</td>
</tr>
</tbody>
</table>

Image: A comparison of an image and an ellipse, showing the invariance properties of different models.
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 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
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

- 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

---

1D Edge Detection
1D Ridge Detection

Ridge and odd wavelets

Ridge and even wavelets

Wavelet coefficients

Wavelet coefficients
Ridge Detection Results
Radon Transform

Finite Radon Transform (FRT)

- Generalization of continuous Radon transform
- Relies on finite geometry structure of $p \times p$ 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!