A Data-driven Cloud Classification Framework
Based on a Rotationally Invariant Autoencoder

Takuya Kurihana¹,⁴, Elisabeth Moyer³,⁴, Rebecca Willett¹,², and Ian Foster¹,⁵

¹Department of Computer Science, University of Chicago; ²Department of Statistics, University of Chicago; ³Department of Geoscience, University of Chicago; ⁴Center for Robust Decision Making on Climate and Energy Policy (RDCEP); ⁵Argonne National Laboratory

Supervised cloud classification restricts models in artificial predefined category e.g. WMO, ISCCP
Issue of artificial cloud category:
- Effective for “classic” examples or deterministic definition
- Non-functional for intermediate / complex cloud types
- Mean cloud properties do not capture relevant physics and spatial information

Unsupervised Cloud Classification (CC) enables novel classification
Supervised cloud classification restricts models in artificial predefined category e.g. WMO, ISCCP

Rotation-Invariant (RI) Autoencoder
- Adapt shift-transform invariant autoencoder (Matuso et.al 2017) to map different orientations of identical inputs into an uniform orientation (Cloud class is independent to its orientation!)
- Rotation dependence problem in autoencoder

\[ L = \lambda_{inv}L_{inv} + \lambda_{res}L_{res} \]

- Transform-Invariant loss
  \[ L_{inv} = \frac{1}{N} \sum_{x \in S} \left\| \tilde{x} - D(E(Th_j(x))) \right\|^2 \]
- Restoration loss
  \[ L_{res} = \min_{x \in S} \left\| T_{\Theta_j}(x) - \hat{x} \right\|^2 \]

Use of MODIS Satellite Data
- MOD02: Use for train and test data
- MOD35: Detec cloud pixels
- MOD06: Evaluate physical association

Validation of RICC
- Design four protocols for evaluating the meteorological utility of the novel clusters produced by RICC system
- Tested RICC system on Phys & Test dataset (see data section) to validate whether our RICC system satisfies our evaluation protocols
- Results show physically reasonable, spatially coherent, cohesive in latent space and rotational invariant

Comparison with CUMULO
- Compared resultant cloud classes obtained with unsupervised and supervised classifier
- Use CUMULO dataset (Zantedeschi et. al 2019)
- Computed percentage of categories of CUMULO in our respective cluster

Take-Away Points
- Address rotation dependence problem to cloud image
- Yield physically reasonable, spatially coherent which learns spatial and rotation invariant features