Studying convective organisation with Deep Learning

Leif Denby, ICAS, University of Leeds
3/7/19
RMetS Atmospheric Science Conference
Barbados

~100km
“flowers”
Aim

• produce software to **automatically segment and classify** a satellite image **into regions with differently organised convection**

Motivation

• Form of organisation affects radiative properties (albedo) and cloud-radiative feedback contributes **majority of climate sensitivity uncertainty** (Bony et al 2015 and many more)

• **Relative importance of local and large-scale** factors driving convection into specific forms of organisation are **unknown**

• **Use tool on satellite images** to identify times where different classes have formed and **correlate with large-scale state** diagnosed from reanalysis data (e.g. ERA-Interim)
“Archetypes” of convective organisation?

“sugar”

“fish”

“gravel”

“flower”

Stevens et al 2019, submitted
“Archetypes” of convective organisation?
Machine learning modelling aim

- Produce for every tile \( (t) \) of a satellite image an *embedding*. A point in \( N \)-dimensional space

\[
\mathbf{f}(t) = [0.12, 0.82, \ldots]
\]

- Enforce that tiles with similar cloud structure are close in this \( N \)-dimensional space

- Previous successful application in Google’s word2vec (Mikolov et al 2013):
  - \( \mathbf{f}(\text{“santa”}) - \mathbf{f}(\text{“christmas”}) \sim \mathbf{f}(\text{“man”}) \)
  - \( \mathbf{f}(\text{“london”}) - \mathbf{f}(\text{“england”}) \sim \mathbf{f}(\text{“copenhagen”}) - \mathbf{f}(\text{“denmark”}) \)

- Using technique of Tile2Vec (Jean et al 2018) which learnt land-use classification
Using convolutional network to produce embedding

- Training done with **fastai** (built on **pytorch**)
- Use pre-trained Resnet34
- Replace last layer by fully connected layer
- Currently using $N_d=100$ embedding length
Using convolutional network to produce embedding

• Training done with **fastai** (built on **pytorch**)
• Use pre-trained Resnet34
• Replace last layer by fully connected layer
• Currently using $N_d=100$ embedding length

• How does this work?
  • See workshop in next session
Model training

• Every training example consists for three tiles (triplet) the anchor \((t_a)\), neighbour \((t_n)\) and distant \((t_d)\) tiles.
Model training

• Every training example consists for three tiles (triplet) the anchor ($t_a$), neighbour ($t_n$) and distant ($t_d$) tiles.

• Use loss function which optimises for anchor and neighbour tiles being close in embedding space and distant tile being far away (measured by Euclidian distance):
Model training

• Every training example consists for three tiles (triplet) the anchor ($t_a$), neighbour ($t_n$) and distant ($t_d$) tiles.

• Use loss function which optimises for anchor and neighbour tiles being close in embedding space and distant tile being far away (measured by Euclidian distance):

$$L(t_a, t_n, t_d) = \max(0, \| f_{\theta}(t_a) - f_{\theta}(t_n) \|_2 - \| f_{\theta}(t_a) - f_{\theta}(t_d) \|_2 + m)$$
What does training data look like?

First 150 "anchor" tiles for a GOES-16 training dataset
What does this embedding look like?

First four dimensions of embedding with a 10 random examples highlighted
So what does this embedding give us?

Can rank tiles by distance (in embedding space) to specific tile of interest (showing half of entire study set)
So what does this embedding give us?

- Can do (hierarchical) clustering to find out how tiles clump in \textit{embedding space}.
- \textit{Nested} clusters share similar features.
- Vertical distance in \textit{dendrogram} measure of persistence of clusters.
Do different cloud structures have different radiative properties?

- Per-cluster mean of channel 1 (visible) and channel 9 (IR), error in the mean as error bars. Nearest tile to mean rendered as example.
- Separation of clusters indicate each has specific radiative properties.
Summary

• a neural network can, without labelled training data:
  • automatically discover different forms of cloud organisation
  • through this learn to group input images containing similar cloud structures together

• different cloud structures have distinct radiative properties

• benefits of unsupervised model:
  1. can be applied to any spatial dataset with limited effort as no hand-labelling is required
  2. automatically discovers the types of structures present in the input
  3. produces a representation of the similarity between these structures.
Thank you!

Questions?