EGU2020-9870
https://doi.org/10.5194/egusphere-egu2020-9870
EGU General Assembly 2020

Detecting Tropical Cyclones using Deep Learning Techniques

Daniel Galea, Bryan Lawrence, and Julian Kunkel
University of Reading, Computer Science, United Kingdom of Great Britain and Northern Ireland

Finding and identifying important phenomena in large volumes of simulation data consumes time and resources. Deep Learning offers a route to improve speeds and costs. In this work we demonstrate the application of Deep Learning in identifying data which contains various classes of tropical cyclone. Our initial application is in re-analysis data, but the eventual goal is to use this system during numerical simulation to identify data of interest before writing it out.

A Deep Learning model has been developed to help identify data containing varying intensities of tropical cyclones. The model uses some convolutional layers to build up a pattern to look for, and a fully-connected classifier to predict whether a tropical cyclone is present in the input. Other techniques such as batch normalization and dropout were tested. The model was trained on a subset of the ERA-Interim dataset from the 1st of January 1979 until the 31st of July 2017, with the relevant labels obtained from the IBTrACS dataset. The model obtained an accuracy of 99.08% on a test set, which was a 20% subset of the original dataset.

An advantage of this model is that it does not rely on thresholds set a priori, such as a minimum of sea level pressure, a maximum of vorticity or a measure of the depth and strength of deep convection, making it more objective than previous detection methods. Also, given that current methods follow non-trivial algorithms, the Deep Learning model is expected to have the advantage of being able to get the required prediction much quicker, making it viable to be implemented into an existing numerical simulation.

Most current methods also apply different thresholds for different basins (planetary regions). In principle, the globally trained model should avoid the necessity for such differences, however, it was found that while differing thresholds were not required, training data for specific regions was required to get similar accuracy when only individual basins were examined.

The existing version, with greater than 99% accuracy globally and around 91% when trained only on cases from the Western Pacific and Western Atlantic basins, has been trained on ERA-Interim data. The next steps with this work will involve assessing the suitability of the pre-trained model for different data, and deploying it within a running numerical simulation.