Segmenting and Tracking Extreme Climate Events using Neural Networks

Mayur Mudigonda[*1], **Sookyung Kim**[*2], **Ankur Mahesh**[*1], **Samira Kahou**[3],
Karthik Kashinath[4], **Dean Williams**[2], **Vincent Michalski**[5], **Travis O'Brien**[4] **and Mr Prabhat**[4]
University of California Berkeley[1], Lawrence Livermore National Lab[2], Microsoft Research[3]
Lawrence Berkeley National Lab[4], Montreal Institute for Learning Algorithms[5]
mudigonda@berkeley.edu, kim79@llnl.gov, ankur.mahesh@berkeley.edu,
samira.ebrahimi.kahou@gmail.com, kkashinath@lbl.gov, williams13@llnl.gov,
michalskivincent@gmail.com, taobrien@lbl.gov, prabhat@lbl.gov

Abstract

Predicting extreme weather events in a warming world is one of the most pressing and challenging problems that humanity faces today. Deep learning and advances in the field of computer vision provide a novel and powerful set of tools to tackle this demanding task. However, unlike images employed in computer vision, climate datasets present unique challenges. The channels (or physical variables) in a climate dataset are manifold, and unlike pixel information in computer vision data, these channels have physical properties. We present preliminary work using a convolutional neural network and a recurrent neural network for tracking cyclonic storms. We also show how state-of-the-art segmentation algorithms can be used to segment atmospheric rivers and tropical cyclones in global climate model simulations. We show how the latest advances in machine learning and computer vision can provide solutions to important problems in weather and climate sciences, and we highlight unique challenges and limitations.

1 Introduction

Analyzing extreme events in large datasets poses a significant challenge in climate science research. Conventional tools to analyze extreme events are built upon human expertise, and they require subjective thresholds of relevant physical variables to define specific events. There is a vast literature on applications of deep learning methods to speech recognition (Hinton et al. [2012], Bahdanau et al. [2016]), computer vision (Krizhevsky et al. [2012], Szegedy et al. [2015]) and natural language processing (Bahdanau et al. [2016], Kalchbrenner et al. [2014]). However, in climate science research the adoption of deep learning techniques is relatively new and limited. In this work we present a preliminary study, showing the benefits of employing such techniques to two applications: (i) tracking trajectories of extra-tropical cyclones (ETCs), and (ii) segmentation of atmospheric rivers (ARs) and tropical cyclones (TCs).

TCs, ETCs, and ARs are important and impactful extreme weather events. Current methods to detect storms rely on sequential processing of the same data to detect each class of storm (TCs, ETCs, ARs, etc.). It would be significantly more efficient to detect all types of extreme weather events based on features/patterns that exist in multivariate climate datasets. Deep learning methods could achieve this goal when they are applied to physical variables such as integrated water vapor, surface pressure, and wind speed. Furthermore, traditional detection methods resort to subjective, arguably arbitrary thresholds, which may change with global warming. Accurate, efficient, and automatic tracking of

[*] These authors equally contributed.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

extreme events can play a critical role in weather prediction if the network can learn precursors to these events. Deep learning may serve as an automated detector and tracker of extreme weather that relies on spatiotemporal patterns, not thresholds, in climate model simulations. With this tool, scientists can better study the environmental drivers that control the frequency, intensity, and location of extreme weather events and how they may change in a warming world. We present preliminary results on segmentation and tracking of extreme weather events using deep learning.

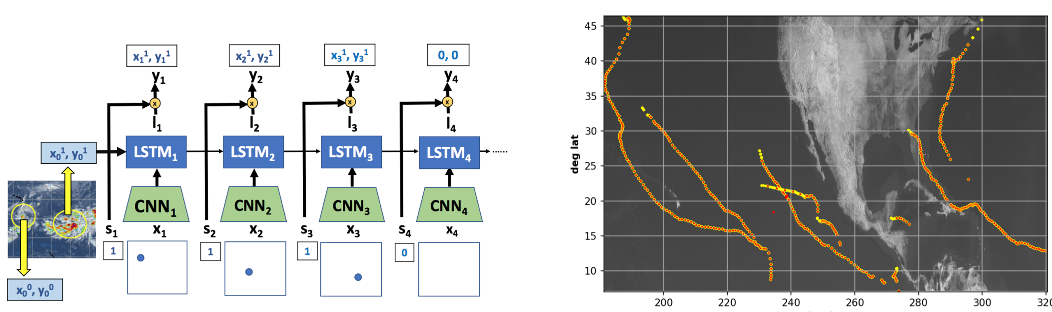
2 Materials and Methods

In this study, we analyze output from a 20-year run (from 1996 to 2015) of the *Community Atmospheric Model v5 (CAM5)* (Neale et al. [2010]). Each snapshot of the global atmospheric state in the CAM5 model output is comprised of multiple physical variables such as integrated water vapor, surface and atmospheric temperature, pressure, wind velocity, etc. For the tracking problem, ground-truth labeling of ETCs is obtained from the Toolkit for Extreme Climate Analysis (TECA) (Prabhat et al. [2012]). Implementing state-of-the-art heuristics and the *MapReduce* paradigm (Dean and Ghemawat [2008]), TECA uses thresholds to detect different types of extreme weather events in global climate model images. While climate models are run on a 3D grid, with the vertical dimension corresponding to 30 pre-determined heights, we only consider surface quantities (i.e. 2D data) in this initial study.

For the semantic segmentation problem, we are interested in both TCs and ARs. We created AR labels by identifying contiguous regions at least 1500 km in length with greater than 95 percentile Integrated Vapor Transport (IVT). To create TC labels, we used TECA to find the TC center and radius (category 0) and create a bounding box. Within this bounding box, we applied the Otsu local method (Otsu [1979]) (histogram-based foreground-background segmentation) of binarization to separate TC regions from the background.

Unlike other machine learning problems, no ground truth or baseline in segmentation and tracking of extreme weather is obviously available. Here, we present preliminary scientific exploration using labels from current threshold-based methods. While we do not have baselines with which to compare our tracking and segmentation models, we show that deep learning is able to (in some cases) rectify errors made by methods used to generate ground truth data. In the near future, we hope to come up with a ground truth dataset with accurate labels employing both expert hand-labelling and unsupervised methods.

3 Tracking Extreme Climate Events



(a) Model architecture

(b) Tracking results

Figure 1: (a) Model: Given the initial position of the target trajectory, the model predicts the remaining positions at each time step. (b) Tracking results: Comparison between ground-truth (yellow) and model prediction (red).

3.1 Model architecture

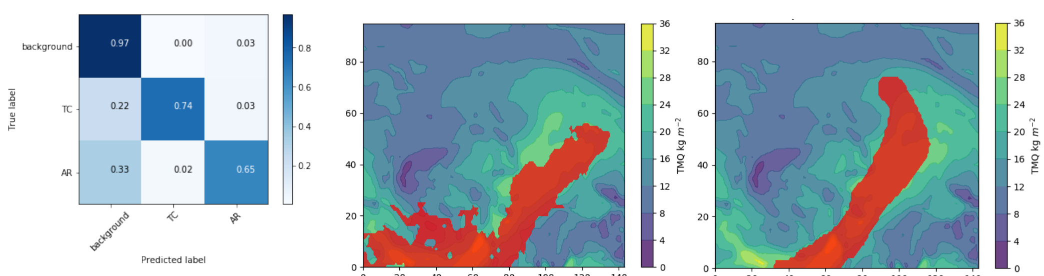
In this work, we track one ETC trajectory at a time. Given the initial position (x_0, y_0), spatiotemporal inputs $(X_1, ..., X_T)$, and state inputs $(s_1, ..., s_T)$, the model determines the positions until the end of the trajectory, $(x_1, y_1), ..., (x_T, y_T)$, using a Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber [1997]). The proposed model is shown in Fig. 1a which consists of two sub-networks: an embedding network using a Convolutional Neural Network (CNN) to represent the input feature maps and a tracking network using an Long Short-Term Memory (LSTM). The embedding network compresses the raw input at each time step, X_i, into a hidden state vector. The tracking network takes the sequence of embedded inputs (hidden state vector) and generates coordinate vectors l. Coordinate values after the trajectory is finished are excluded from loss estimation, and y is represented as element-wise multiplication with coordinate vector l and state vector s. Loss is represented as mean-squared error (MSE) between ground truth ($\hat{y}$) and predicted output (y), and it is minimized using the Adam optimizer (Kingma and Ba [2014]).

3.2 Experiment and Results

We use spatiotemporal CAM5 data with 2 variables (channels) for a fixed time length of 72 hours (24 time steps) in the region of 180 $\deg$ to 340 $\deg$ longitude and 0 $\deg$ to 80 $\deg$ latitude. Label coordinates are collected using TECA and are normalized to be between 0 and 1. We choose precipitation and zonal wind as input channels, given their relevance to tropical cyclone (TC) identification. In order to stay within memory constraints during training, we reduce the resolution of CAM5 data with quarterly max-pooling. After this reduction, the input at a single time step is a 129x86 image with 100-km resolution.

Input sequences are collected from 1480 trajectories with different starting times. The total size of data is 14136, and we used 84.7% (11976) for training and 15.3% (2160) for testing. After training, the model's average mean squared error is $0.68 \deg$ (75.48 km). The ETC tracking results on the test set are shown in Fig. 1b. Ground truth trajectories (denoted as yellow dots) are overlaid together with predicted trajectories (denoted as red dots). The figure illustrates different trajectories of ETCs, randomly selected from the test set. As shown, our model successfully tracks each trajectory by learning the spatiotemporal representation of ETCs in multi-variable climate data.

4 Segmentation of Extreme Climate Events



(a) Confusion Matrix for a Tiramisu semantic segmentation model

(b) Ground truth label of an AR

(c) Model Prediction of an AR

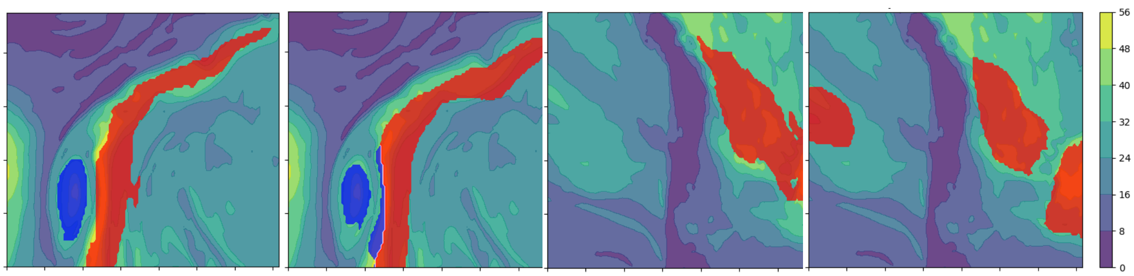
Figure 2: (a) Shows the confusion matrix for a Tiramisu semantic segmentation model. There are three classes in our model - background, ARs and TCs. The model is biased towards predicting classes as the background class, but it still performs greater than chance in all the categories. (b) Current AR-labelling algorithms, based on subjective thresholds, can lead to messy, arbitrary identifications of ARs. (c) Deep learning semantic segmentation models can yield smoother AR identifications.

4.1 Model architecture

Semantic segmentation aims to classify every pixel in an image as a member of a class. We explored semantic segmentation using the Tiramisu model proposed by Jégou et al. [2017]. Briefly, the Tiramisu model applies the DenseNet architecture (Huang et al. [2016]), used for image classification, to semantic segmentation. DenseNets organize fully convolutional layers into a sequence of "blocks." Within a block, the input to each layer is the output of the previous layer concatenated with the inputs of all previous layers. There is a *down-path* (that extracts features) and a symmetrical *up-path* (that reconstructs the image) in the model with skip connections between each corresponding block in the *down-path* and *up-path*. The main advantage of DenseNet is that each block gets direct supervision from its input, resulting in improved performance. Here, with IWV as its only input, DenseNet classifies pixels as tropical cyclone (TC), atmospheric river (AR), or background.

4.2 Experiment and Results

One of the challenges in semantic segmentation of climate data is the class-imbalance problem (2% of the pixels belong to foreground classes). A segmentation model would achieve 98% accuracy by only predicting zeros. Additionally, GPUs do not have enough memory to run segmentation algorithms on high resolution global climate images of size $(768, 1152)$.



(a) Ground Truth (b) Prediction (c) Ground Truth (d) Prediction

Figure 3: (a) Sample cropped ground truth image with AR delineated in red and TC delineated in blue. (b) Corresponding (to Fig.3a) sample cropped predicted image, with AR in red and TC in blue. (c) Sample cropped ground truth image (d) Model prediction for Figure 3c. Here the model fails to isolate the AR on the right side as a contiguous region, and it finds a false positive for an AR on the left side. These errors may have been introduced when cropping global images.

To address the issues above, we crop the global model image into sub-images of size $(96, 144)$. To combat the class imbalance problem, we then choose samples that had at least 10% of the pixels as belonging to non-background classes. This results in about 16,000 images for training and 1,500 images each for validation and testing. We use RMS Prop as our optimizer with an initial learning rate of $1e - 4$ along with weight decay. We explore Tiramisu architectures (modified) with 3 and 5 blocks. Our test accuracy is about 92% (matching both training and validation accuracy). Here, accuracy refers to pixel classification accuracy per image across the whole test set. Figures 3a, 3b, 3c, and 3d show sample ground truths and predictions, and in Figure 3b, the model detects both ARs and TCs. Note that the AR labels were generated from thresholds on IVT, which contains rich information about horizontal transport of atmospheric water but is not commonly output in climate model simulations. However, the model was trained with IWV data, which contains only information about the presence, not the transport, of atmospheric water. Despite IWV's lower information content, the model accurately learned the representation of ARs. Likewise, TECA used surface pressure, surface winds, upper atmospheric temperature, and geopotential height to generate TC labels, but DenseNet learned these labels with IWV alone, showing the benefits of a deep learning approach.

5 Conclusion

We present preliminary results on tracking and semantic segmentation with climate data. CNNs and LSTMs are able to track storms as long as they are trained one at a time. In the future, we aim to apply CNNs and LSTMs to simultaneous storm tracking. Additionally, we show promising preliminary results of semantic segmentation of ARs and TCs. In the future, we hope to train larger models on global, not cropped, climate images, and we plan to explore hyperparameters more rigorously with methods such as Spearmint (Snoek et al. [2012]. It is impressive and promising that the model is able to segment ARs and TCs with only IWV as a feature. Providing more physical variables may improve model performance and allow for detection of other extreme weather events.

References

Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

Dzmitry Bahdanau, Jan Chorowski, Dmitriy Serdyuk, Philemon Brakel, and Yoshua Bengio. End-to-end attention-based large vocabulary speech recognition. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 4945–4949. IEEE, 2016.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188*, 2014.

Richard B Neale, Chih-Chieh Chen, Andrew Gettelman, Peter H Lauritzen, Sungsu Park, David L Williamson, Andrew J Conley, Rolando Garcia, Doug Kinnison, Jean-Francois Lamarque, et al. Description of the ncar community atmosphere model (cam 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, 2010.

Prabhat, Oliver Rübel, Surendra Byna, Kesheng Wu, Fuyu Li, Michael Wehner, Wes Bethel, et al. Teca: A parallel toolkit for extreme climate analysis. *Procedia Computer Science*, 9:866–876, 2012.

Jeffrey Dean and Sanjay Ghemawat. Mapreduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1):107–113, 2008.

Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8): 1735–1780, 1997.

Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Simon Jégou, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*, pages 1175–1183. IEEE, 2017.

Gao Huang, Zhuang Liu, Kilian Q Weinberger, and Laurens van der Maaten. Densely connected convolutional networks. *arXiv preprint arXiv:1608.06993*, 2016.

Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951–2959, 2012.