Cooperative Institute for Climate and Satellites–North Carolina
Inspire. Advance. Engage.

Determining Weather Front Locations Using a Deep Learning Convolutional Neural Network

James Biard, Kenneth Kunkel – Cooperative Institute for Climate and Satellites, Asheville, NC, USA; Evan Racah – University of Montreal, Montreal, Canada

Science Problem

Weather fronts are important to studies of weather and climate because of the variety of significant weather events associated directly with them, including severe thunderstorms and a wide spectrum of precipitation types and amounts. In a recent study, fronts were found to be the direct cause of more than half of observed extreme precipitation events in the contiguous U.S. (Kunkel et al. 2012). They have complex spatial patterns, often represented by complex curved lines.

Fronts are identified visually based on the approximate spatial coincidence of a number of quasi-linear localized features - a trough in air pressure in combination with gradients in air temperature and/or humidity and a shift in wind direction. (Stull 2015) Fronts are categorized as cold, warm, stationary, or occluded, with each type exhibiting somewhat different characteristics.

The locations of weather fronts are currently determined by meteorologists performing visual analysis, and there is no long historical record of the results of those analyses available in a form that can be used for climatological studies. An automated method for determining front locations is needed in order to study fronts over climatological time scales.

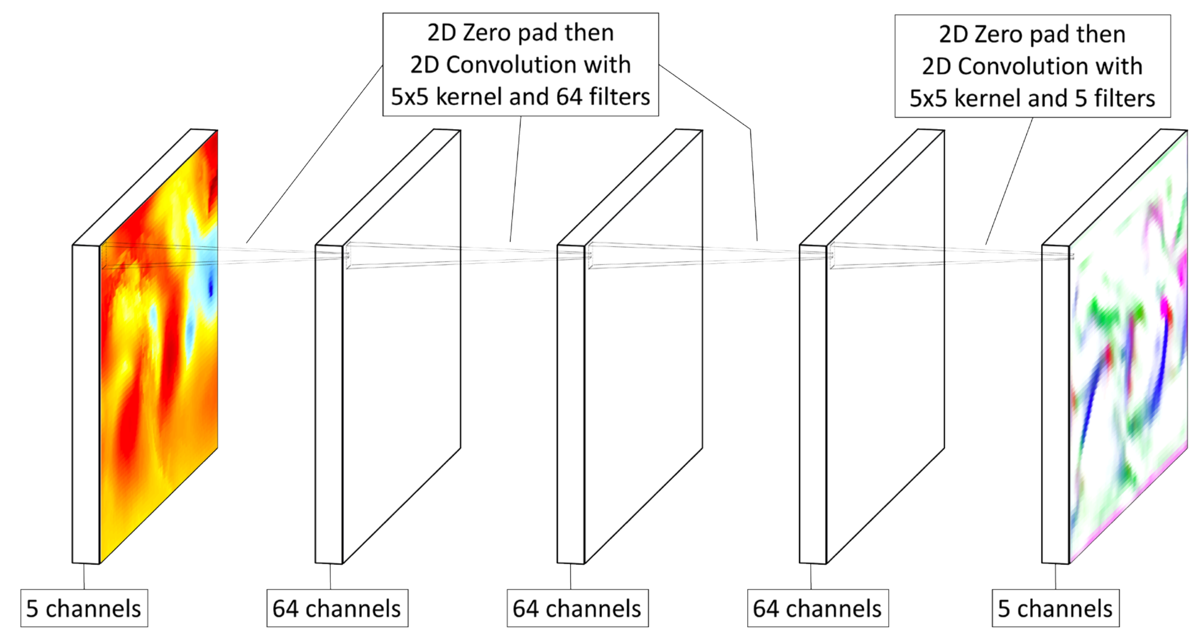
Analytics Problem

A supervised 2-dimensional Convolutional Neural Network (2D CNN) was implemented to investigate whether it could imitate the visual fronts recognition task. The goal for our front classification CNN is to estimate the likelihood that a given pixel in an image composed of grids of weather data lies within a front.

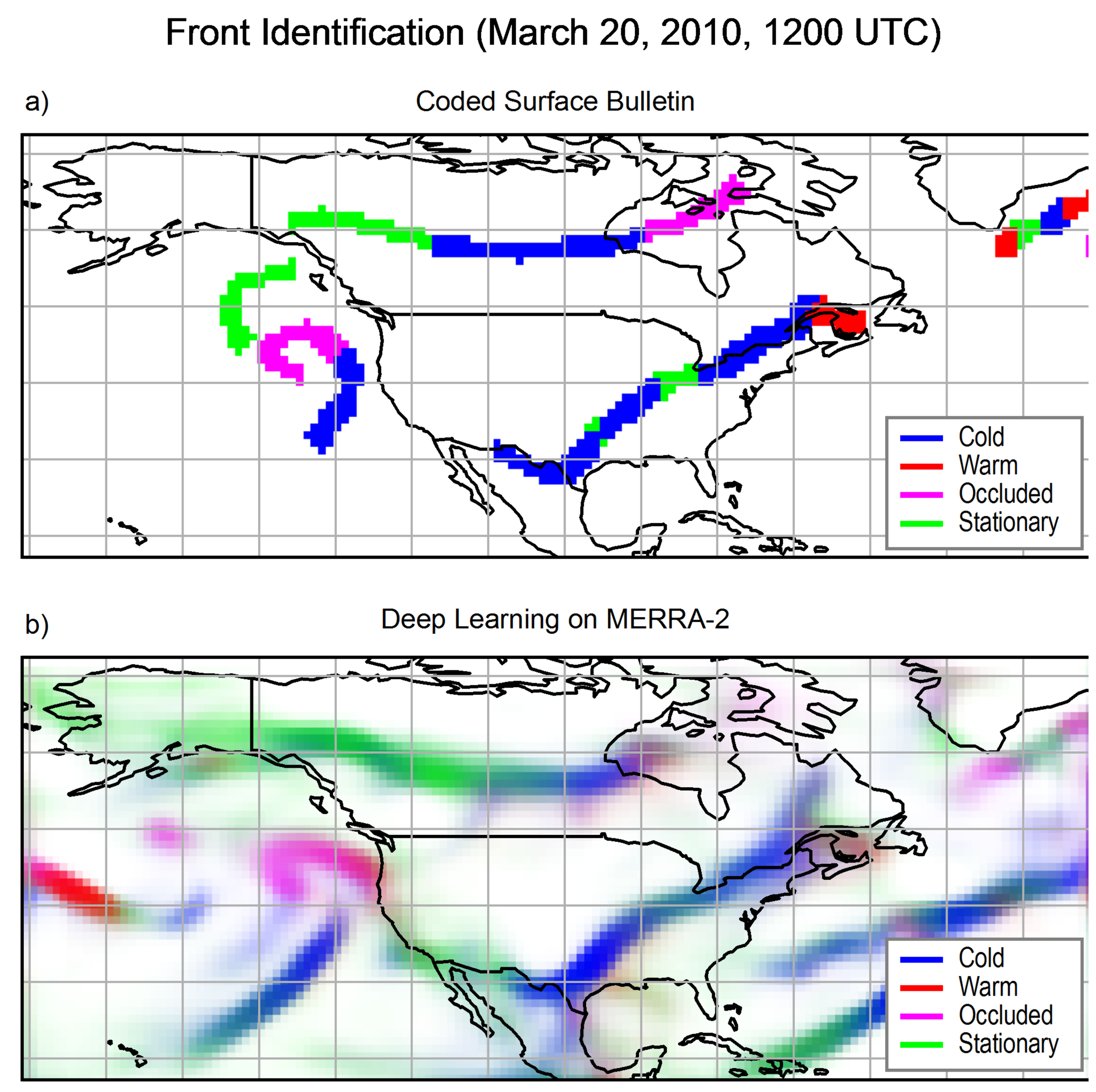
The CNN architecture is trained by optimizing the values of the pixels in the convolution filters to minimize the difference between the truth dataset and the output of the network applied to the input dataset as measured using a cost, or loss, function. We used the categorical cross-entropy loss function, which has the form

$$H(p,t) = -\sum_{i=1}^{I}\sum_{c=1}^{C} w_c \log(p_{ic}) t_{ic}$$

where p is a set of output pixels, t is a set of truth pixels, w is a per-category weight, I is the number of pixels, and C is the number of categories. There are five possible categories including four for different types of fronts and a fifth for the absence of a front (or ``no-front''). Each truth pixel is assigned one and only one category. The lower the likelihood value for the corresponding output category, the larger the contribution to the loss. The per-category weights are used to adjust the relative significance of the contributions from the different categories.



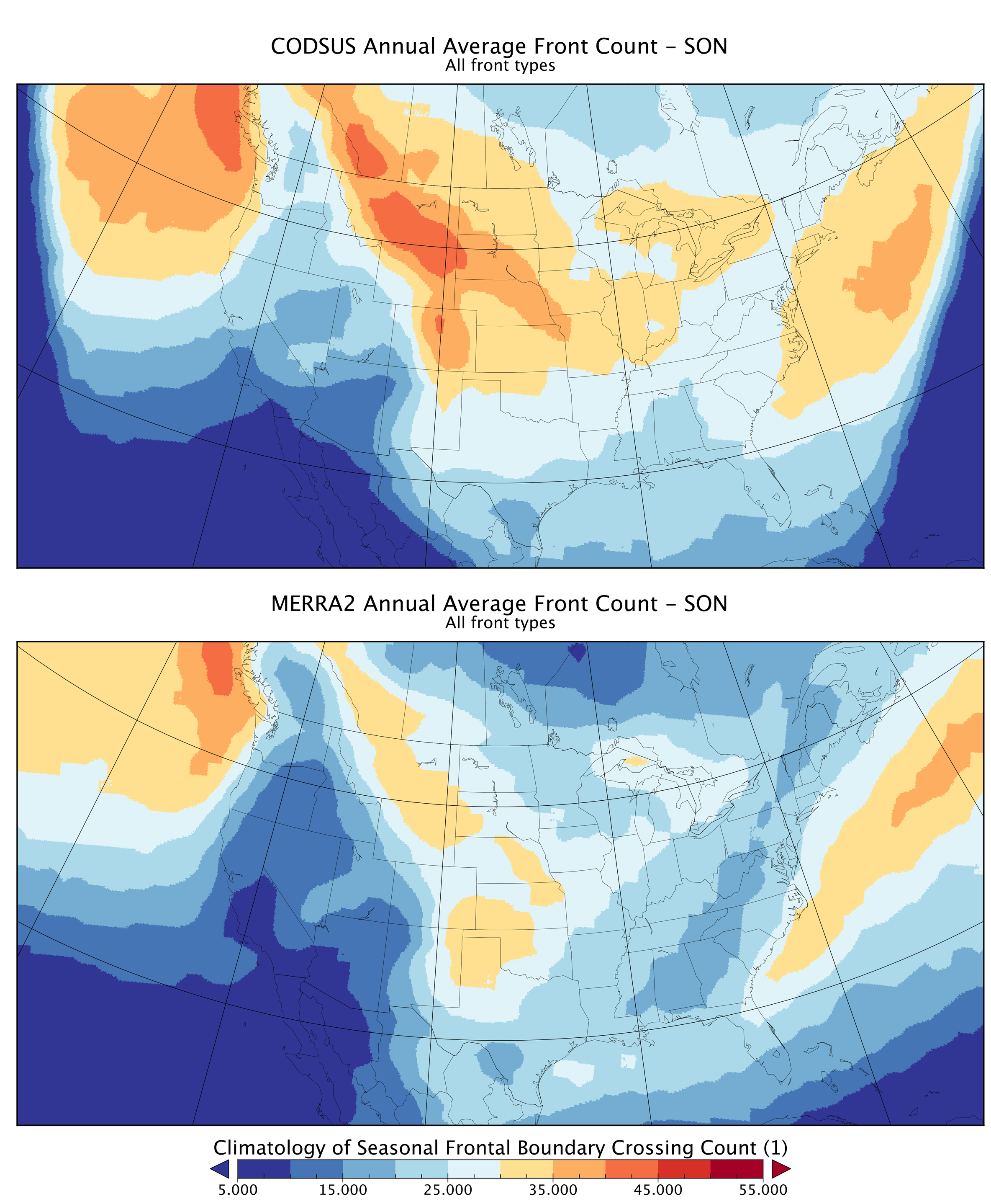
2D CNN Architecture



Sample Truth and CNN Results

Truth\Pred	warm	occluded	cold	stationary
warm	0.44	0.13	0.14	0.29
occluded	0.05	0.75	0.14	0.06
cold	0.01	0.03	0.79	0.17
stationary	0.04	0.03	0.15	0.79

Confusion as Fraction of Truth



Seasonal Climatology of Front Crossing Counts
Sep – Nov, 2003-2015

Datasets

The input dataset consisted of gridded fields of five surface variables taken from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2). The variables were 3-hourly instantaneous values of 2m air temperature, 2m specific humidity, sea level pressure, and the 10m wind velocity.

Our truth dataset was extracted from the Coded Surface Bulletin (CSB). Each text bulletin contains latitudes and longitudes specifying the locations of pressure centers, fronts, and troughs identified visually. Each front and trough is represented by a polyline. We obtained all the bulletins possible for 2003-2016 and produced an image for each time step by drawing the front lines into latitude / longitude grids with one degree cell size. Each front was drawn with a transverse extent of three degrees to account for the fact that a front is not a zero-width line, and to add tolerances for slight lateral differences in position between the CSB and any MERRA-2 front signatures. The quantitative evaluation was restricted to regions where the frequency of fronts was at least 40 per year.

Processing

The network was implemented in Keras, Theano, and Scikit-learn and trained with the data for 2003-2007 using an 80%-20% training-test split and 3-fold cross-validation. We tested the training results by calculating the confusion matrix and per-category Intersection-over-Union (IoU) (ratio of correctly categorized pixels to total pixels in that category in either truth or CNN data, computed over each category) for the entire set of images.

The results were further processed to extract polylines describing the fronts in each time step by tracing out the lines following the maxima for each type of front. These were used to calculate the annual average number of front crossings at each 96 km x 96 km grid cell in a Lambert Conformal Conic map covering North America. We then compared the results with the annual average number of front crossings found using the original polylines from the CSB.

Conclusion

The current front detection CNN is correctly reproducing the overall spatial distribution and frequency of fronts. It consistently under-detects fronts by ~20% according to climatology. The confusion is greatest for warm fronts, followed by occluded fronts. Improvement is likely as the CNN architecture is refined, including adding sensitivity to the time dimension.

References

Kunkel, K. E., D. R. Easterling, D. A. R. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2012: Meteorological causes of the secular variations in observed extreme precipitation events for the conterminous united states. Journal of Hydrometeorology, **13 (3)**, 1131–1141

Stull, R., 2015: Practical Meteorology: An Algebra-based Survey of Atmospheric Science. University of British Columbia.

MERRA-2 (https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/)

CSB (http://www.nws.noaa.gov/noaaport/html/noaaport.shtml)