Automating weather forecasts based on convolutional networks

Pablo Rozas Larraondo [1] **Iñaki Inza** [2] **Jose A. Lozano** [2 3]

Abstract

Numerical weather models generate a vast amount of information which requires human interpretation to generate local weather forecasts. Convolutional Neural Networks (CNN) can extract features from images showing unprecedented results in many different domains. In this work, we propose the use of CNN models to interpret numerical weather model data which, by capturing the spatial and temporal relationships between the input variables, can produce local forecasts. Different architectures are compared and a methodology to introspect the models is presented.

1. Introduction

Weather forecasting is based on Numerical Weather Predictions (NWP) that capture the state of the atmosphere and simulate its evolution based on physical and chemical models. Global NWP models normally provide a large number of parameters representing different physical variables in space and time. Because of the lack of spatial and temporal resolution, these fields need to be interpreted by highly qualified personnel to produce forecasts for any specific region. This is still today a human based process, which relies on specifically trained and experienced professionals to interpret modeled and observed data (Wilson et al., 2017; Gravelle et al., 2016). NWP variables define the state of the atmosphere and its changes through space and time. NWPs define a highly structured dataset in which the relationships between its variables are defined by physics equations, such as conservation of mass, momentum and energy.

Recent advances in neural networks have proven that by increasing the number of general hidden layers, unprecedented results can be achieved in many different domains. More specifically, research around Convolutional Neural Networks (CNN) (Krizhevsky et al., 2012) has proven to be very effective in solving image classification and segmentation problems.

Machine learning has been applied to different areas of weather forecasting, such as downscaling (Tripathi et al., 2006) or nowcasting (Xingjian et al., 2015). The main deficiency of the traditional methodologies is their inability to incorporate both the spatial and temporal components present in the data. Most of the existing research in this field has been based on manually extracting the points in a model representing a certain location, and training models with the resulting data. The problem with this approach is that weather is a dynamic system, and analysing individual points in isolation misses important information contained in the synoptic and meso-scales.

CNNs enable analysis and extraction of the spatial information in images, building from fine grained details into higher level structures. The temporal dimension can be added to these networks by adding a third axis to the convolutional kernels. This work demonstrates how CNNs can be used to interpret the output of Numerical Weather Prediction (NWP) automatically to generate local forecasts. The main outcomes of this work are:

- CNNs are able to provide a model to interpret numerical weather model fields directly and to generate local weather forecasts.
- Class Activation Mapping (CAM) provides a valuable mechanism to assess the spatial and temporal correlations of the different fields visually, helping to introspect and develop new models.
- 3D convolutions can naturally incorporate the temporal component into neural networks, significantly improving the accuracy of the results.

2. Datasets

For this work, we propose the use of NWP and observed precipitation data from different locations, to experiment with different configurations of CNN models. The objective is to train a model which predicts the event of rain for a

[1]National Computational Infrastructure, Australian National University, Australia [2]University of the Basque Country, Spain [3]Basque Center for Applied Mathematics, Spain. Correspondence to: Pablo Rozas Larraondo <pablo.larraondo@anu.edu.au>.

Proceedings of the ICML 17 Workshop on Deep Structured Prediction, Sydney, Australia, PMLR 70, 2017. Copyright 2017 by the author(s).

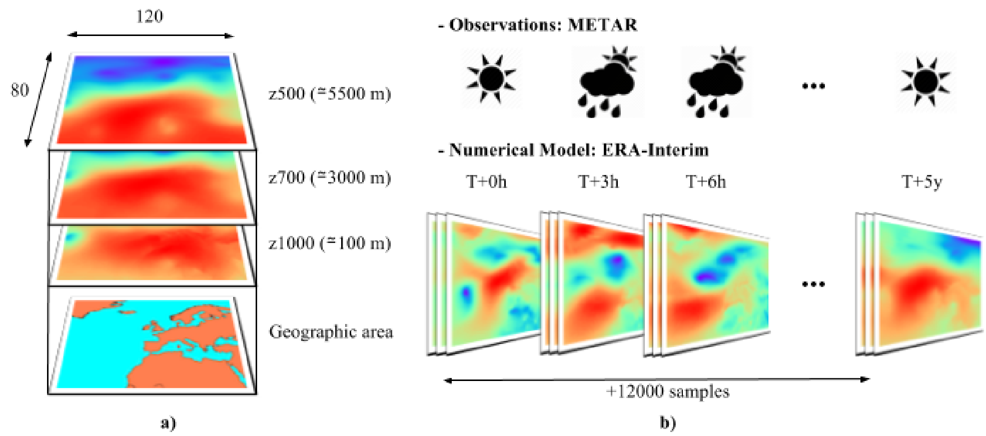

Figure 1. a) represents the 3 geopotential subsets extracted from ERA-Interim, corresponding to different heights of the atmosphere, stacked over a map to represent the spatial extent. b) Represents the whole extracted time series and the alignment of both datasets.

particular location, using numerical weather model data as input. In this section, we describe how these datasets have been generated.

ERA-Interim (Dee et al., 2011) is a publicly available meteorological reanalysis dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF). This dataset is generated using a numerical weather model which simulates the state of the atmosphere for the whole planet, with a spatial resolution of approximately 80 km. There is data available since the year 1979, with a temporal resolution of 3 hours. The output is presented in the form of regular numerical grids and there is a large number of physical parameters available, representing variables such as temperature, wind speed and relative vorticity.

Aviation Routine Weather Reports (METARs) are operational aviation weather text reports that encode observed meteorological variables for every commercial airport in the world. METARs are produced with an hourly or half-hourly frequency and are also made publicly available through the World Meteorological Organisation (WMO) communications system. Each report is uniquely identified by its header, which contains the International Civil Aviation Organization (ICAO) airport code and a UTC time stamp.

We considered 5 main airports located in different cities across Europe and a period of 5 years (2012-2017) to perform our experiments. The airports and their corresponding ICAO codes are: Helsinki-Vanta (EFHK), Amsterdam-Schiphol (EHAM), Dublin (EIDW), Rome-Fiumicino (LIRF) and Vienna (LOWW).

We extracted an extended area over Europe from ERA-Interim, creating a 3 hourly series of images composed by 3 bands, corresponding to the geopotential height **z** at the 1000, 700 and 500 pressure levels of the atmosphere. This parameter represents the height in the atmosphere at which a certain pressure value is reached and the levels correspond typically to 100, 3000 and 5500 metres above the mean sea level respectively.

The reason for selecting these fields is that weather forecasters normally base their predictions on these. They contain information about the shape, location and evolution of the pressure systems in the atmosphere.

Using the METAR data, the precipitation conditions [*rain*, *dry*] were extracted for each airport for the same time period and frequency. The resulting dataset time series contains over 12000 samples. Figure 1 represents a sample of the considered ERA-Interim fields with their size and geographical extent on the left. The right side, shows how the ERA-Interim data aligns with the observed precipitation for a sample location.

3. Experiments and Results

The objective of the proposed experiment is to predict precipitation events for the considered airports using ERA-Interim geopotential data as the input and METAR observations to annotate the samples [*rain*, *dry*]. Two different CNNs are used. The first model performs 2D convolutions and the second incorporates the temporal dimension based on 3D convolutions (Ji et al., 2013). We aim to prove that these models can capture part of the mental and intuitive process that human forecasters follow when interpreting numerical weather data.

3.1. CNN architecture

To perform the experiments, a 2 layer CNN is used. Each convolution layer uses a 3x3 kernel followed by a 2x2 max-pooling layer. After the convolution operations, a fully connected layer is used to connect the output [*rain*, *dry*] using a 'softmax' activation function.

Table 1. Rain forecasting accuracies for the different locations comparing 2D and 3D CNNs with the reference accuracy of climatology.

AIRPORT	RAIN CLIM.	2D CNN	3D CNN
EFHK	60.8	73.6	75.4
EHAM	74.2	77.8	79.3
EIDW	61.2	70.7	72.6
LIRF	83.1	87.3	88.2
LOWW	75.7	77.1	78.8

For the 3D CNN, the configuration is similar to the previous version, but the kernels in both the convolution and max-pooling layers have an extra dimension, with sizes 3x3x3 and 2x2x2 respectively. The 3D CNN, is trained by aggregating the input dataset in groups of 8 consecutive images. This aggregation represents the evolution of the atmosphere over 24 hours. The neural network can then extract information out of the temporal dimension, using the observation corresponding to the last image of the series as output.

The 2D and 3D CNNs were implemented in TensorFlow (Abadi et al., 2016) and trained per airport over 300 epochs using 80% of the data. The remaining 20% was used as validation to test the accuracy of the models.

3.2. Results

Table 1 contains the results produced by the different models using the validation dataset. The accuracy values represent the success rate of the model when predicting either *rain* or *dry* conditions for each location. The climatology for each location, number of rain observations over the total number of observations, is also included in the results as a reference. A model whose output is always '*dry*' would have that success rate.

Figure 2 represents the results using a stacked bar chart. The relative improvement over climatology achieved with the 2D and 3D convolutional models is represented by the green and red fractions of the bar.

3.3. Class Activation Mapping

Class Activation Mapping (CAM) (Zhou et al., 2016) is a technique that localises class-specific image regions in a trained CNN.

This technique uses the last layer of a CNN to create a graphical representation for a particular output based on its weights. The resulting image is a heat-map representing which parts of the image have a higher influence in the output.

For example, Figure 3 depicts two different CAM representations for the 2D CNN models trained using the precipitation data of Helsinki and Rome. Warmer colours in the image represent higher weight values, so the network makes its decisions mostly based on the features located in those areas. The images in Figure 3 corroborate the intuitive idea that local weather patterns have a higher influence than distant ones when forecasting the weather of a particular location. The images have been overlaid with a coast map to serve as a reference for the relative position of the structures in the heat-maps.

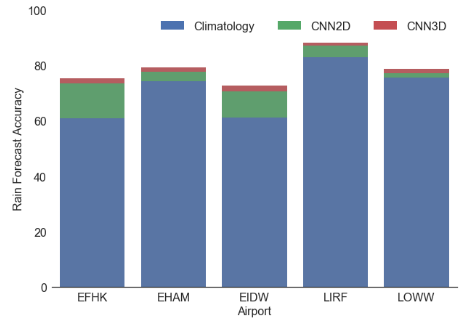

Figure 2. Accuracy results obtained for the different airports and methodologies.

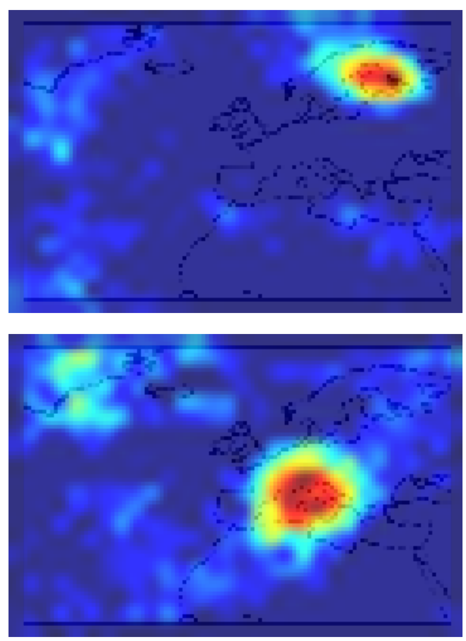

Figure 3. Example of the resulting Class Activation Maps for an ERA-Interim CNN trained using the observed precipitation at Helsinki-Vantaa airport, EFHK, (top) and Rome Fiumicino airport, LIRF, (bottom). Coastlines have been overlaid as a reference for readers.

This technique has been proven very useful for visually assessing the soundness of a CNN model. Another use could be for input variable selection, identifying NWP parameters which show a higher correlation with respect to the class to be predicted.

3.4. Software and Data

The code used to run all the experiments included in this work and instructions on how to access the corresponding datasets are available at the following repository: http://github.com/prl900/DeepWeather

4. Conclusions and future work

This work demonstrates how CNNs can be directly applied to the output of numerical weather models by using observed data to annotate the samples. The design of the CNNs used in our experiments is very simple compared to some of the state-of-the-art architectures (Simonyan & Zisserman, 2014; Szegedy et al., 2015). Despite their simplicity, results show that convolutional layers can be used to interpret the output of weather models.

The NWP parameters used in the experiments are not directly correlated to the precipitation output variable. NWPs have many other variables, such as humidity, vorticity or even total precipitation, that could be used to forecast precipitation patterns with better accuracy. The purpose of this initial experiment was to demonstrate that CNNs can learn certain configurations of the atmospheric pressure systems and associate them with precipitation events (fronts, convection, etc).

Apart from weather model interpretation, these techniques open a new research pathway for the automatic generation of derived products. Some of the variables contained in NWPs are computed based on parameterisations or statistical models instead of physical equations. We think that these variables can be computed using CNN based models, potentially offering better results.

Acknowledgements

The authors wish to acknowledge funding from the Australian Government Department of Education, through the National Collaboration Research Infrastructure Strategy (NCRIS) and the Education Investment Fund (EIF) Super Science Initiatives through the National Computational Infrastructure (NCI), Research Data Storage Infrastructure (RDSI) and Research Data Services Projects.

References

Abadi, Martín, Agarwal, Ashish, Barham, Paul, Brevdo, Eugene, Chen, Zhifeng, Citro, Craig, Corrado, Greg S, Davis, Andy, Dean, Jeffrey, Devin, Matthieu, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.

Dee, DP, Uppala, SM, Simmons, AJ, Berrisford, Paul, Poli, P, Kobayashi, S, Andrae, U, Balmaseda, MA, Balsamo, G, Bauer, P, et al. The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656):553–597, 2011.

Gravelle, Chad M, Runk, Kim J, Crandall, Katie L, and Snyder, Derrick W. Forecaster evaluations of high temporal satellite imagery for the goes-r era at the nws operations proving ground. *Weather and Forecasting*, 31(4): 1157–1177, 2016.

Ji, Shuiwang, Xu, Wei, Yang, Ming, and Yu, Kai. 3d convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):221–231, 2013.

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

Simonyan, Karen and Zisserman, Andrew. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

Szegedy, Christian, Liu, Wei, Jia, Yangqing, Sermanet, Pierre, Reed, Scott, Anguelov, Dragomir, Erhan, Dumitru, Vanhoucke, Vincent, and Rabinovich, Andrew. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

Tripathi, Shivam, Srinivas, VV, and Nanjundiah, Ravi S. Downscaling of precipitation for climate change scenarios: a support vector machine approach. *Journal of Hydrology*, 330(3):621–640, 2006.

Wilson, Katie A, Heinselman, Pamela L, Kuster, Charles M, Kingfield, Darrel M, and Kang, Ziho. Forecaster performance and workload: Does radar update time matter? *Weather and Forecasting*, 32(1):253–274, 2017.

Xingjian, SHI, Chen, Zhourong, Wang, Hao, Yeung, Dit-Yan, Wong, Wai-Kin, and Woo, Wang-chun. Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems*, pp. 802–810, 2015.

Zhou, Bolei, Khosla, Aditya, Lapedriza, Agata, Oliva, Aude, and Torralba, Antonio. Learning deep features for discriminative localization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2921–2929, 2016.