EARTH SCIENCE DEEP LEARNING: APPLICATIONS AND LESSONS LEARNED

Manil Maskey[1], Rahul Ramachandran[1], J.J. Miller[2], Jia Zhang[3], Iksha Gurung[2]

[1]NASA Marshall Space Flight Center
[2]University of Alabama in Huntsville
[3]Carnegie Mellon University

ABSTRACT

Deep learning has revolutionized computer vision and natural language processing with various algorithms scaled using high-performance computing. At the NASA Marshall Space Flight Center (MSFC), the Data Science and Informatics Group (DSIG) has been using deep learning for a variety of Earth science applications. This paper provides examples of the applications and also addresses some of the challenges that were encountered.

***Index Terms*—** Deep learning, neural network, Earth science, classification, large scaled labeled data, training

1. INTRODUCTION

Deep learning is a subfield of machine learning that includes algorithms inspired by the function of the brain. It consists of multilayer neural network of neurons (simple computational units). The lower layers learn low-level features, such as edges, and then the higher layers progressively learn high-level representations, such as complex shapes, followed by object parts [1]. The first layer is composed of the inputs to the neural network, followed by one or more hidden layers, with the last layer containing the outputs of the network. The difference between a traditional neural network and deep learning is that deep learning receives a set of inputs and performs progressively complex calculations to output a solution. In this hierarchical layout, each layer receives input from the output of the previous layer, breaking down complex patterns into a series of simpler patterns. Deep learning algorithms have proven to be a powerful tool for various machine learning problems. Unlike traditional approach of domain experts engineering hand-crafted features, deep learning algorithms learns the features without the need for feature engineering to solve the same problems. Here, we present use of deep learning to address several outstanding Earth science problems. Each application is unique and presents challenges: mainly to construct features from data, the algorithms require large amount of labeled training data.

There are many cases in industry where deep learning has scaled successfully. For example: Facebook translates about 2 billion user posts per day in more than 40 languages. Microsoft products such as Bing and Xbox uses deep learning for speech-recognition. Google uses deep learning for almost all of its services. In medical science, deep learning is used for diagnosis and language translation. Self-driving cars are the latest advancements driven by deep learning. One thing that is in common with all of these applications is that large amount of training data exists. We highlight and address this challenge for Earth science applications.

2. EARTH SCIENCE APPLICATIONS

In this section, we discuss our deep learning-based applications for Earth science.

2.1. Earth Science Phenomena Detection

Researchers typically use event (an instance of a phenomenon) data for case study analysis. Earth science data search systems are currently limited to specifying a query parameter that includes space and time of an event. This is a current limitation that results in researchers spending a considerable amount of time sorting through data to study an event. An alternative search paradigm is to use browse images-based search. Before search based on images can be performed, the images in the Earth science database need to be classified. For most Earth science data, data archives also distribute corresponding browse images, which are much smaller in size compared to actual data files and include rendering of the data values. The DSIG team applied the Convolutional Neural Networks CNNs [2][3] to classify images in Earth science database to improve the search experience for event study [4]. The training dataset for the application was constructed using the NASA's Land, Atmosphere Near real-time Capability for EOS (LANCE) rapid response which supports end users in monitoring and analysis of various phenomena. Domain experts also labeled images to further increase the size of training dataset. Sample labeled images of few phenomena are shown Figure 1.

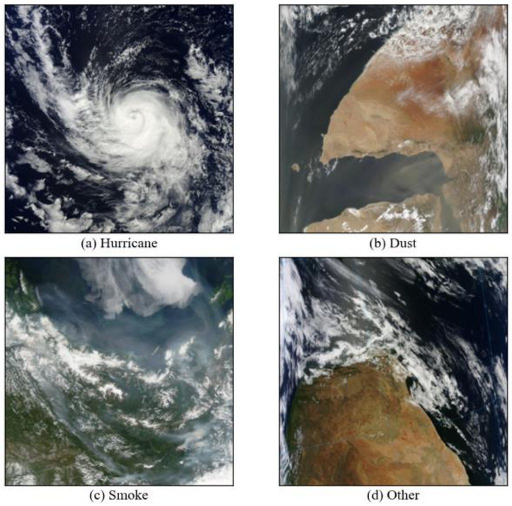

Fig 1 Sample labeled images for various phenomena: (a) hurricane, (b) dust storm, (c) smoke, and (d) none.

2.2. Tropical Cyclone Intensity Estimation

There are inherent issues with current techniques to estimate wind speed of tropical cyclones that rely on the Dvorak technique [5]. Mainly the issues relate to human subjectivity and generalization. The issues were apparent for the most recent hurricane Ophelia where *two human experts at Tropical Analysis and Forecast Branch (TAFB) and NOAA/NESDIS Satellite Analysis Branch (SAB) differed by 20 knots in their Dvorak analyses, and the automated version at the University of Wisconsin was 12 kt lower than either of them*. The DSIG team has adapted the Convolutional Neural Networks (CNNs), a deep learning algorithm that is most suitable for computer vision tasks to address the issue of objectively and accurately estimating the intensity of tropical cyclone using satellite imageries [6]. The training dataset for this application was constructed by using two different datasets: tropical cyclone centric imageries from Naval Research Laboratories (*http://www.nrlmry.navy.mil*) and wind speed information from HURDAT2, the tropical cyclone best track renalysis data (*http://www.nhc.noaa.gov/data/#hurdat*). Figure 2 illustrates a set of feature maps (outputs of convolution filters) for a hurricane.

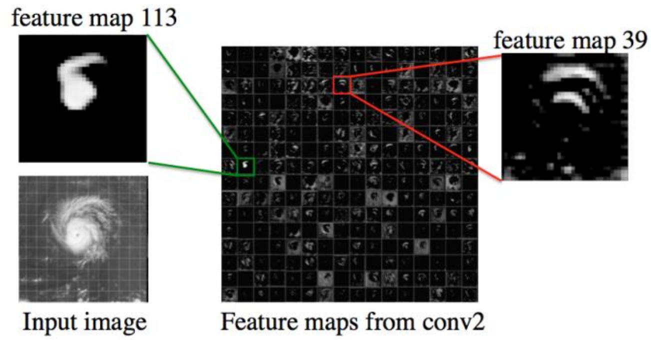

Fig. 2 CNN feature maps for a hurricane image.

2.3. Severe Storm (Hailstorm) Detection

Being able to detect hailstorm from radar imagery has implications to human safety and property protection. Many current hailstorm detection techniques rely on domain knowledge and substantial preprocessing. To avoid this laborious and tedious process, the DSIG team has applied a parameter optimized CNN for hailstorm detection with superior accuracy than existing techniques. The training dataset was constructed by combining known instances of hailstorms from storms reports with corresponding NEXRAD images available from Iowa Environmental Mesonet [7] images. Figure 3 shows correctly classified sample radar images with presence of hail.

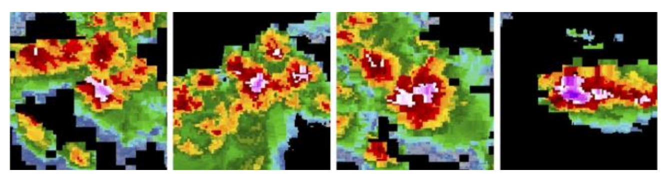

Fig. 3 An example that test images labeled as "Hail" are classified as "Hail" by our trained model.

2.4. Earth Science Knowledge Graph Construction

Published Earth science resources contain enormous amount of knowledge that is not easy to extract. The DSIG team is attempting to accurately extract entities and relations across published Earth science resources and construct the Earth science knowledge graph that can be used to answer more advanced questions and discover new insights. The approach taken includes deep learning methods for natural language processing to extract semantic entities from Earth science literature trained using known vocabularies and limited expert knowledgebase.

2.4. Transverse Bands Detection

Transverse cirrus bands are ice clouds that are irregularly spaced bandlike cirrus clouds and often form in association with other weather phenomena such as mesoscale convective systems, hurricanes, and jet streaks (Knox et al., 2010). These bands are known to be associated with clear air turbulence. Thus, automated detection of transverse cirrus bands in satellite imagery is of utility to aviation weather forecasting. The DSIG team have used CNN to detect the transverse bands in satellite imageries with both spectral and morphological information [8]. Domain expert manually created the training dataset for this application.

2.6. Ephemeral Water Detection

Ephemeral water is temporary water body formed due to direct response to precipitation. Ephemeral water is extremely important for parts of Africa that receives very little precipitation. Detecting such water body from satellite imageries can allow cattle farmers to direct their livestock for grazing. However, detection of such water bodies is a difficult problem since spatially the water body may be represented by only a pixel or two within satellite images of highest resolution. The DSIG team is attempting to solve this problem using stacked auto encoder on Landsat imageries. Training dataset was generated using water index and known water body shapefiles.

3. ADDRESSING IDENTIFIED CHALLENGES

Next, we share our lessons learned after applying deep learning on several outstanding Earth science problems over the past four years. We identified two main challenges: dealing with deep learning black box and creating labeled training datasets.

3.1. Deep Learning Black Box

Even though deep learning performance for above mentioned applications was impressive, there is no clear understanding of why it performs so well, or how it could be improved. From a scientific point of view, it is important to bring insight into the internal operation and behavior of the complex model. Domain scientists are skeptical of the "black box" that is deep learning and want to know what physical conditions or mechanisms contribute to a given result. They prefer to better understand the learned features, the importance of features, and how they relate to their science.

To address this challenge an evaluation component that is geared towards understanding the physical meaning of the model is needed to provide a level of confidence for the scientists. Specifically, in the case of tropical cyclone intensity estimation, we developed visualization techniques to reveal the input pixels that are highly discriminative at any layer in the model. Such visualization allow us to track evolution of features during training. We also applied deconvolution network to project the filter outputs back to input pixel space.

3.2. Labeled Training Data

For each of the applications, constructing training datasets was by far the most tedious and time-consuming step. Deep learning algorithms can be adapted and tuned for most applications, however, the performance of the algorithms depends heavily on the size and quality of the training dataset. Large number of data points are needed to learn large number of parameters in the model that machines have to learn. Generic large-scale labeled datasets such as the ImageNet [8] are the fuel that drives the impressive accuracy of deep learning results. Creating large scaled labeled datasets in the Earth science domain is a big challenge. Manually creating labeled training data is a bottleneck and not scalable. While there are ways to apply deep learning using limited labeled datasets, there is a need in the Earth sciences for creating large-scale labeled datasets for benchmarking and scaling deep learning applications.

From our observations, there is an interesting almost linear relationship in the amount of data required and the size of the model. Basic reasoning is that model should be large enough to capture relations in your data along with specifics of your problem. Initial layers of the model capture high level relations between the different parts of the input. Later layers capture information that helps make the final decision; usually information that can help discriminate between the desired outputs. Since most Earth science problems are rather constrained (For example: satellite image classification into 8 classes), the training dataset can be substantially smaller than a generic image classifier (For example: ImageNet). Next, we present few approaches to address the challenge of creating large scaled labeled dataset for Earth science.

3.2.1. Data Augmentation

Data augmentation is an artificial way of increasing the number of training sample with label preserving transformations. Data augmentation is especially useful for computer vision tasks as there are several image transformation techniques that can be used without affecting the class label. Rotation, cropping, color shifting of images are just a few data augmentation techniques. All of our CNN-based applications use some form of data augmentation.

3.2.2. Transfer Learning

Transfer learning is a method where a model developed for a task is reused as the starting point for a model on a different task. When a model is trained ("pre-trained") the network gains knowledge from training data and compiles weights of the network. The weights can be extracted and then transferred to another network. In this way, instead of training network from scratch, learned features are "transferred"; hence requiring smaller training dataset. There are several ways to fine tune the pre-trained model for specific case: (a) use for feature extraction only by removing the output layer, (b) use the network architecture of the pre-trained model but reinitialize the weights, and (c) use only few layers from the pre-trained model while retraining the other layers. There are empirically validated rules depending on the training data size and data similarity which can help determine how to fine tune pre-trained model or start from scratch. We successfully applied transfer learning to the transverse bands detection application by re-using the network architecture of the pre-trained model.

3.2.3. Permutation Invariance

Permutation invariance occurs when a model produces the same output regardless of the order of elements in the input. It can be used to represent data that does not have spatial relationship. Thus, we can use permutation invariance for constructing a large dataset of related words (entities) for initial training set to build knowledge graph.

3.2.4. Data Programming

Data programming involves programmatic creation of training dataset where experts provide weak supervision strategies and a discriminative model to label the unlabeled data. A few ways to perform weak supervision are include: (i) domain rules and heuristics, (ii) distant supervision: existing ground-truth data that is no exact fit, (iii) weak classifiers, and (iv) non-expert annotations or crowdsourcing. Consider applying data programming to create a labeled data for dust storm study using sample text from literature: “Pronounced changes in the aerosol optical parameters, derived from AERONET, have been observed during dust storms”. If our weak supervision consists of labeling functions as shown in Figure 4, then we will extract the relevant entities to study dust storms. Here, the *labelingFunction1* leverages existing Earth Science knowledge base such as SWEET [9] and the *labelingFunction2* applies domain heuristics.

```
def labelingFunction1(input):
    concept = (input.phenomenon,input.property)
    return 1 if concept in DOMAIN_KB else 0


def labelingFunction2(input):
    found = re.search(r'.*derived.*',input.text.between)
    return 1 if found else 0
```

Fig. 4 Labeling functions used for extracting entities and relations from text.

Recently, we have started exploring a data programming framework, snorkel [10], which seems promising.

6. CONCLUSION

This paper presents applications of deep learning for Earth science. Such applications are not without challenges that persist, including improving scientists' trust in the developed model and creating large-scaled labeled datasets. The DSIG group dealt with both issues in a systematic way. We also present other possible approaches to address the challenges. Our observation suggests that the deep learning algorithms can be adapted and tuned to tackle Earth science problems, however, the value is in the large scaled labeled datasets. We believe that labeled training dataset will be the barrier for using deep learning for Earth science. Thus, we recommend management of existing and future datasets in a catalog for curation, search and discovery, preservation.

11. REFERENCES

[1] Y. Bengio, “Learning deep architectures for AI,” in *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.

[2] Y. LeCun and Y. Bengio, “Convolutional networks for images, speech, and time-series,” in The Handbook of Brain Theory and Neural Net- works, M. A. Arbib, Ed. Cambridge, MA, USA: MIT Press, 1998.

[3] LISA Lab. Convolutional Neural Networks (LeNet)—DeepLearning 0.1. Accessed: Feb. 2016. [Online]. Available: http://deeplearning.net/tutorial/lenet.html

[4] M. Maskey, R. Ramachandran, J. Miller, "Deep learning for phenomena-based classification of Earth science images," in Journal of Applied Remote Sensing, vol. 11, no. 4, 2017.

[5] V. Dvorak, "Tropical Cyclone Intensity Analysis and Forecasting from Satellite Imagery," in Monthly Weather Review, vol. 103, no. 5, pp. 420-430, 1975.

[6] R. Pradhan, R. S. Aygun, M. Maskey, R. Ramachandran and D. J. Cecil, "Tropical Cyclone Intensity Estimation Using a Deep Convolutional Neural Network," in IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 692-702, Feb. 2018.

[7] D. Herzmann, R. Arritt, and D. Todey, Iowa environmental mesonet. Available at http://mesonet.agron.iastate.edu/request/coop/fe.phtml (Accessed August 2017). Iowa State Univ., Dep. of Agron., Ames, IA, 2004.

[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. IEEE Computer Vision and Pattern Recognition (CVPR), 2009.

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

[9] R. Raskin, M. Pan, "Knowledge representation in the semantic web for Earth and environmental terminology (SWEET)," In Computers & Geosciences, vol. 31, no. 9, pp. 1119-1125, 2005.

[10] A. Ratner, S. Bach, H. Ehrenberg, and C. Ré, "Snorkel: Fast Training Set Generation for Information Extraction," in Proceedings of the 2017 ACM International Conference on Management of Data, SIGMOD '17, pp. 1683--1686, 2017.