Convolutional Generative Adversarial Networks for Estimating Precipitation

Ambiorix Cruz Angeles, Sami Abed, Yuan Zhou, Airton Prado, Satyam Sharma
Prof. Michael D. Grossberg
The City College of New York
Department of Computer Science

Abstract.

With the availability of satellites orbiting the earth and the increase in quality and quantity of the data they generate, deep learning models are an ideal tool for leveraging this data. We aim to apply Convolutional Generative Adversarial Networks to the task of predicting precipitation using multiple spectral bands obtained from weather satellites. This can be used to estimate precipitation over the oceans where it is hard to obtain but is essential to weather forecasting and meteorological studies.

Introduction.

In meteorology and climate studies, precipitation in its various forms (rain, snow, drizzle, etc.) is essential to weather prediction such as nowcasting. Ideally, we want to have daily estimates for precipitation over the entire globe. Sources of precipitation data include gauge measurements, Radar (Weather Radar), and Satellite imagery:

Gauge measurements are accurate and give the actual amount of rain that reaches the ground, but it’s sparse and is mostly only available on land.

Radar is a high-resolution precipitation data source but it is hard to obtain because it has limited coverage especially over remote areas and oceans or where radar beams are blocked, and because it is susceptible to anomalies. For example, non-weather objects can influence the radar and be mis-identified as rain. Also, not all precipitation is reliably detectable by radar. Targets have to be 10 times smaller than radar waves to register, so tiny objects like cloud droplets are un-detectable [1].

Weather Satellites can give images of multiple bands in the visible, near IR and IR (thermal) spectrums. Satellites are classified based on their orbit as Geostationary Satellites, such as GEOS-12 and GEOS-13, or Polar-orbiting Satellites such as NOAA-15, which are closer to the earth's surface but have lower sampling frequency. Satellites have a good sampling frequency (every 30 minutes or less), even Polar-orbiting Satellites can provide high sampling frequency by making use of multiple satellites.

Unlike radar, satellites can cover both land and oceans, but, we do not have long-term satellite data. Being able to obtain high quality precipitation data over the oceans is essential to climate simulation and other weather studies, and this can give us a lot of data since oceans cover the majority of earth’s surface (over 70%.)

We will be applying deep learning to train a model to predict a radar precipitation image given satellite images covering the same area. This can then be used to predict precipitation over oceans using satellite imagery. For this, we will be using a model known as Pix2Pix [3], a variation on a generative machine learning model [2] that has recently demonstrated unprecedented results on a multitude of image processing and machine learning tasks.

Background.

There has been numerous attempts at trying to learn and detect precipitation over oceans and other bodies of water dating as far back as back as 1979 [5] where measurements and observation by ships had been used to study precipitation, and satellite microwave measurements were used to further verify these studies.

With the increase of effectiveness and popularity of deep and neural network based machine learning algorithms, several papers have recently been published. A similar work [6] to ours has just been published in Nov, 2018 which studies techniques to detect precipitation over oceans at low winds using GPS satellites.

PERSIANN has been continuously developed [7], moving from pixel-wise predictors to patches of clouds.

Also recently, a related paper [8] studies the impact of incorporating multiple features including IR channels as well as cloud properties in rainfall estimation.

Convolutional neural networks have also been used [9] for doing image segmentation for detecting clouds instead of the more traditional pixel-based approaches.

Method.

Data Collection and Preprocessing:

We are working with two main datasets. One is multi-band satellite images from the GEOS-13 satellite which is publicly available, and the other is Stage IV radar precipitation that is corrected using local measurements, made available to us by our domain expert.

Radar dataset covers the 41N-31N 102W-82W area (mostly land) and is of hourly frequency. Satellite dataset comes from the GEOS-13 satellite and covers a larger area than radar, but was cropped to match the same coordinates. The satellite dataset consists of 5 spectral bands [10]:

GOES Imager Band	Name	Central Wavelength (μm)	Objective
1	Visible	0.63	Cloud cover and surface features during the day, smoke, etc.
2	Shortwave window	3.9	Low cloud/fog, fire detection, winds, etc.
3	Water vapor	6.48	Upper-level water vapor, winds, etc.
4	Longwave window	10.7	Surface or cloud-top temperature, precipitation, etc.
5	N/A	N/A	N/A
6	CO_2 band	13.3	CO_2 band: Cloud detection, etc.

Satellite dataset is also of hourly frequency, but it is 15 minutes behind the radar dataset.

Since naturally there is little change from one hour to the next, radar dataset has been reduced by filtering out completely blank or almost completely blank images.

The satellite images have also been cropped and resized to 256x256 pixels from their original size of about 190x560 pixels to make sure satellite and radar images match in size and are square which is a required for the U-Net implementation of the Generator network (Discussed later). This size was best for minimizing the amount of wasted area in each image while also minimizing image-resizing and the distortion that comes with it. This also gives us more samples out of our original dataset (at the cost of reduced image sizes that is)

Due to the small sample size of the datasets, (which gets even smaller after pairing satellite images with their corresponding radar image,) the datasets have also been augmented by doing a 90-degree rotation and a horizontal left-right flip. This resulted in the dataset size getting increased by a factor of 3 going from 1300 hourly samples (5 satellite images and a corresponding radar image per sample) to 3900 hourly samples.

Finally, we split the data into test and train sets with the test set size having a ratio of 20% of the 3900 samples. This leaves 3120 samples for training the model.

Model:

Our problem is supervised and generative in nature. A Generative Adversarial Network (GAN) [2] is a game-theoretic unsupervised model that was invented in 2014. It consists of 2 networks, a Generator (G) and a Discriminator (D), that are trained in lockstep. The generator generates images (y) starting from a random noise vector (z). The Discriminator's job is to tell whether a given image is from the true data distribution or was generated by the Generator.

Conditional Generative Adversarial Networks [4] is a variation on GANs which adds restrictions on the generated image by using the provided labels (x), which are fed to both the Discriminator and the Generator:

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{x,y}[\log D(x,y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x,z))]$$

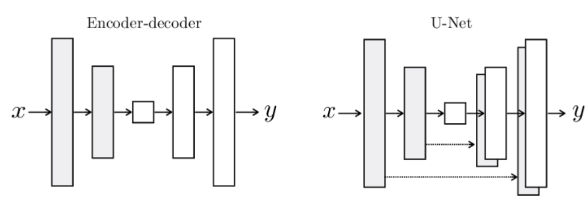
Pix2pix [4] is a model based on conditional GANs which was proposed by Philip et al. It adds an extra loss term (*L1* difference between the generated image and the ground truth) to the images generated by the Generator. And with that, the final objective function is:

$$\arg\min_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{L1}(G).$$

where the Generator is trying to minimize this function and the Discriminator is trying to maximize it.

In our case, x is an observed satellite image, and y is the corresponding ground truth radar image

Pix2pix uses a convolutional architecture and ReLU as activation function for both its generator and discriminator as well as skip connections in the generator network (U-Net: encoder-decoder + skip connections)

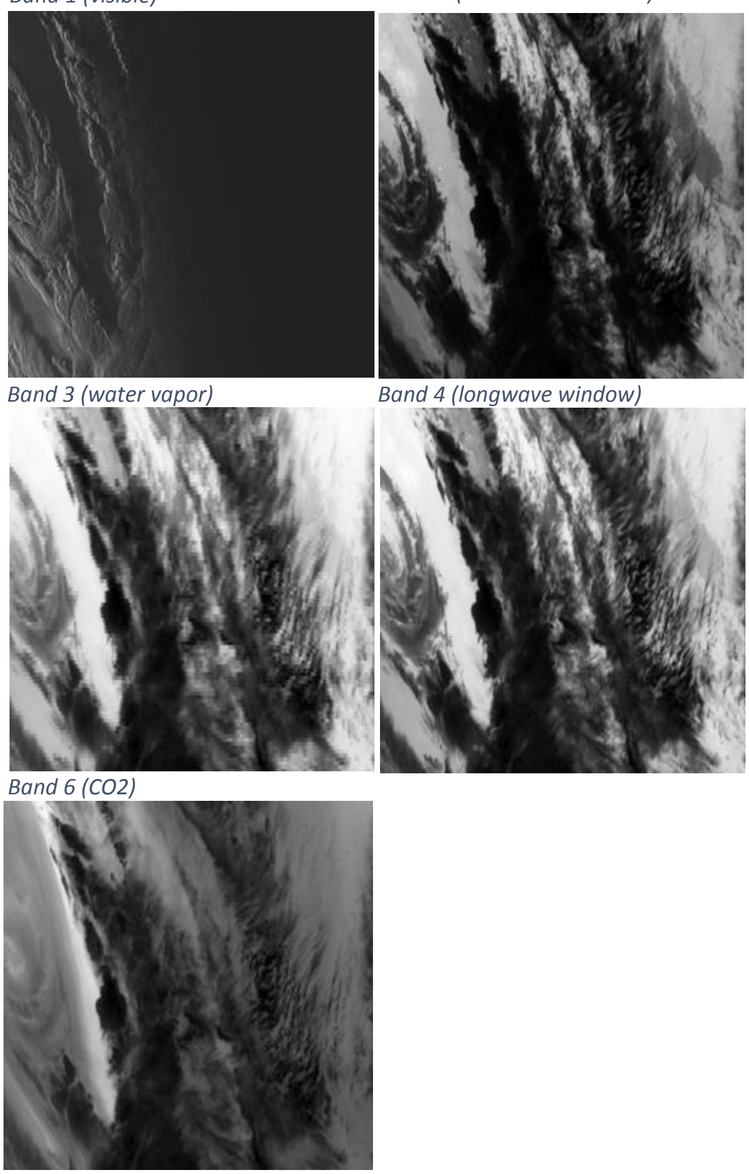


Philip et al [3]. Architecture of Encoder-decoder vs. U-Net

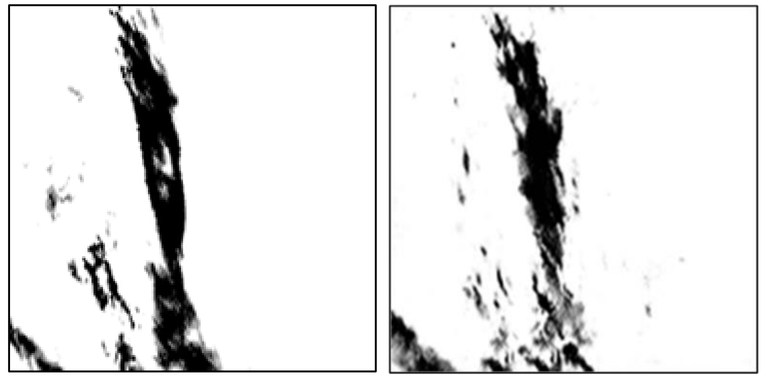
Training and Hyper-parameters:

For optimization, ADAM is used with an initial learning rate of 0.0002, which stays constant for the first 150 epochs, and decays to 0 over the next 100 epochs, making for a total of 250 training epochs. Training for longer than this did not yield improvement in the quality of the generated images. Here's an training sample showing the input and output:

Input Channels (Satellite, 5 Bands)



Band 1 (visible) *Band 2 (shortwave window)* *Band 3 (water vapor)* *Band 4 (longwave window)* *Band 6 (CO2)*

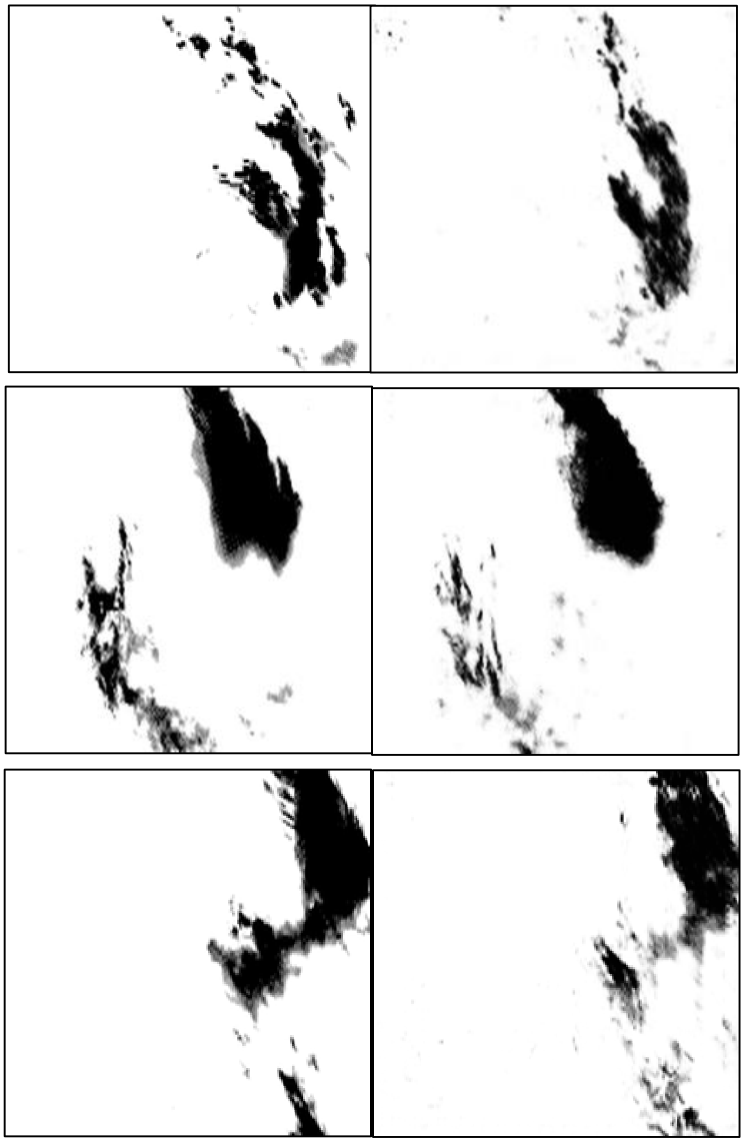


Ground Truth (Radar) *Generated (Radar)*

An obvious, yet important, advantage of this model over the simpler pixel-based models is that it is able to leverage the spatial information in the images and not just the pixel values. Another important advantage of using this model is that it can improve and make use of more data, more so than traditional machine learning models whether that data is entire new features or just more samples of the same data. Perhaps the biggest dis-advantage is the time it takes to train the model as well as the massive compute resources it needs. We trained our model on server-grade machines with an NVIDIA K80 GPU and that still took over 20 hours on a rather small dataset and 250 epochs.

Results and Evaluation.

Here we can see a few examples of the predicted radar images (right) and their corresponding ground truth (left):



We run a pixel-wise comparison on each (ground truth, predicted) pair of the test set images to determine the correct classification of each pixel, that is, whether it's a cloud or non-cloud pixel. This gives us the following confusion matrix:

	Predicted non-cloud	Predicted cloud
Actual non-cloud	8.1%	4.1%
Actual cloud	6.8%	81%

We can see that our model was able to get a total of 89.1% accuracy (81% true positive + 8.1% true negative.)

A simple linear regression baseline that does pixel-based fitting got a 50% accuracy (true positive + true negative.) This can probably be attributed to the fact that there is a slight shift between the input satellite bands and the radar images because of the 15 minute gap. Our GAN can mitigate this and make better use of the spatial relationship between the pixels in the image. An 89.1 percent accuracy is a very big improvement over the pixel-based baseline.

Conclusion.

Being able to accurately estimate precipitation over large bodies of water is an important problem in meteorology. Here we have shown that Generative Adversarial Networks are very promising, even though we have not tested our model on satellite imagery directly above the ocean. It is worth highlighting the fact that we worked with a small dataset, and these results could be improved by incorporating more samples and additional features that would help in predicting precipitation, such as weather model data. Deep learning can yield impressive results on huge datasets.

References.

[1] Doviak, R. J.; Zrnic, D. S. (1993). Doppler Radar and Weather Observations (2nd ed.). San Diego CA: Academic Press. ISBN 0-12-221420-X.

[2] Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." In *Advances in neural information processing systems*, pp. 2672-2680. 2014.

[3] Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." *arXiv preprint* (2017).

[4] Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." *arXiv preprint arXiv:1411.1784* (2014).

[5] Dorman, Clive E., and Robert H. Bourke. "Precipitation over the Pacific Ocean, 30 S to 60 N." *Monthly Weather Review* 107, no. 7 (1979): 896-910.

[6] Asgarimehr, M., Zavorotny, V., Wickert, J., & Reich, S. (2018). Can GNSS Reflectometry detect precipitation over oceans? *Geophysical Research Letters*, 45, 12,585–12,592. https://doi.org/10.1029/2018GL079708

[7] Hong, Y.; Hsu, K.; Sorooshian, S.; Gao, X. Precipitation Estimation from Remotely Sensed Imagery
Using an Artificial Neural Network Cloud Classification System. J. Appl. Meteorol. **2004**, 43, 1834–1853, doi:10.1175/JAM2173.1.

[8] Beusch, L.; Foresti, L.; Gabella, M.; Hamann, U. Satellite-based rainfall retrieval: From generalized linear models to artificial neural networks. Remote Sens. **2018**, 10, 939.

[9] Drönner, Johannes, Nikolaus Korfhage, Sebastian Egli, Markus Mühling, Boris Thies, Jörg Bendix, Bernd Freisleben, and Bernhard Seeger. "Fast Cloud Segmentation Using Convolutional Neural Networks." *Remote Sensing* 10, no. 11 (2018): 1782.

[10] Schmit, Tim. *The GOES Imager Spectral Bands*. PDF. NOAA NESDIS, 2014.