Abstract

Greenhouse gas emissions and detection have become very important as climate change becomes more severe. Methane is one of these prominent greenhouse gases. In our project, we tried to expand on the work of past researchers who attempted to accurately detect the location and extent of methane plumes. Our approach differs in our solution to the problem of past researchers having limited high-quality methane plume image data. We propose to use a deep-learning-based algorithm called generative adversarial networks(GAN) to generate a more comprehensive set of plume images using only limited but high-quality images. We then used a convolutional neural network as well as the discriminator itself as a classifier to evaluate our success in using the generated images to detect methane plumes and compare them with the paper we attempted to build upon, "Detecting Methane Plumes using PRISMA: Deep Learning Model and Data Augmentation" by Alexis Groshenry et al.

Background

Methane plumes are an important issue to be aware of and a contributor to the large amounts of methane that get released from man-made products, which is very detrimental to the environment. Many of these come from "super-emitters," a term used by NASA to describe facilities or equipment that emit methane at high rates. Methane has one of the highest global warming potentials of any greenhouse gas, almost 80 times that of carbon dioxide, and is the second-highest emitted GHG[3]. It is a crucial problem to be able to detect big methane emitters and deal with them either through policies or other measures. Part of our inspiration to work on this project came from our study of the paper, "Detecting Methane Plumes using PRISMA: Deep Learning Model and Data Augmentation," by Alexis Groshenry et al[1]. Researchers of the paper used high-resolution images of methane plumes taken from the PRISMA satellite to train models on identifying and measuring methane plumes. They encountered a problem when data obtained from PRISMA was not large enough to effectively train a model for accurate detection. To increase the amount of input data, they attempted to use simulated data from Large Eddy Simulations (LES). However, it was very computationally expensive. They also tried to photoshop plume images from another satellite of lower resolution where the data was more readily available onto PRISMA images without plumes to generate more data. One final solution was to use transfer learning to train a model on images from another satellite in hopes of being able to better train it on PRISMA images as a result.[1] Our idea was to use a GAN (generative adversarial network) to generate images that would be as similar to the methane plume images taken from PRISMA as possible. We felt that a GAN could be a more effective approach to the problem of not having enough data, and would require less loss of information in comparison to simply photoshopping lower-quality methane plume data. We plan on using how successful models trained with our generated synthetic images are in detecting methane plumes.

Methods

Our original plan was to obtain the same PRISMA and Sentinel-2 data because we wanted to tackle the same data problem as the researchers[1] in a different way and be able to compare our results directly by replicating the rest of the image segmentation methods. We also tried accessing the PRISMA data directly through a request on their website, but their website was down, and the request form we found elsewhere asked for some personal information(i.e. address). We weren't sure if it was safe to request given that it seems no one is managing it, and even if we did it seems unlikely to be approved within the timeframe of this project. We decided to use the simulated data graciously provided by Professor Lindsay from another research project[2].

The simulated data was generated using the Weather Research and Forecasting model (WRF) in large eddy simulation (LES) mode with data of atmospheric columns and added noise(Varon). There are a total of 354 simulated images of size 120x120 pixels, all of which contain methane plumes. This means we have no image data without plumes in them. As an alternative, we used random noises produced from a normal distribution to replicate images without plumes.

After checking the data, we reshaped the arrays into (354, 120, 120, 1) shape and normalized the data to between -1 and 1. We also tried not normalizing the data because a histogram of the data points show it is not evenly distributed through a range. We also tried feature selection using PCA. After data was split into train and test sets, data augmentation techniques RandomCrop and RandomRotation were used on the training data only to increase the training data size tenfold. Crop and rotate were chosen because they result in images that still resemble methane plumes. Both training and test data were resized to 80x80 pixels. Data was then put into a batched and shuffled state in preparation for training, with a standard batch size of 128 and a buffer size of 3500 to accommodate for the new data size.

We decided a DCGAN would be suitable to generate images for our purposes because it is a simple GAN friendly to those new to using GAN and provides stable results. Each training of the GAN took a long time due to performance limitations. If we decrease the amount of epoch too much it decreases the quality of resulting images, therefore we decided to manually test a few different hyperparameters. We tried increasing the number of transposed convolutional layers in both the discriminator and generator(both at the same time and separately) but gained worse results. This was surprising because even though we risk overfitting, we presumed it would provide better results. We also adjusted the kernel size, number of filters, training rate, and different numbers of epochs to find the values of these hyperparameters that generate images that are most visually similar to the original data. A binary cross entropy loss function was used because of the GAN's nature of the two networks, discriminator and generator, essentially competing against each other.

A simple K-means image segmentation technique was used on the original plume images, and we found that perhaps due to the simplicity of the simulated images, even a basic K-means technique was successful in identifying the shape, size, and location of the plumes in the images[Figure 7]. We concluded that comparing image segmentation results with the original research paper[1] would be pointless, so we focused on comparing the plume detection results using a classification model. A CNN was implemented to classify images with plumes and images without plumes using image data we generated from our GAN(Figure 1) and random noise(Figure 2). The discriminator of the DCGAN was also tested as a classifier against the original data, with an interest in measuring how well the GAN did. We compared the result of the CNN with both the discriminator and the U-net image segmentation model of the research paper[1]. We will be using precision, recall, and f1-score as evaluation metrics so to compare with the result of the research paper we are expanding upon[1] since they also used these same detection evaluation metrics.

Results

The DCGAN training on normalized data and without PCA showed some success in generating images visually similar to those in the original data(Figure 3). The quality of the images increased by a lot up to approximately 400-500 epochs then the increasing trend of quality became pretty noticeable, but randomly generated images after about 1000 epochs still seem to perform better. Training on data that are not normalized and with PCA provided unsuccessful results even after 500 epochs(Figure 4). After PCA the data images were still very visually similar to the original, and so we presumed training using data after applying PCA could increase the performance of the algorithm through compressing the images to more important parts. Somehow it caused the opposite result.

The results of using the discriminator and the CNN as a classifier are shown in Figure 5, in comparison to the result in Figure 6 from the original research paper[1].

Discussion

The successful image generation process of the DCGAN on the simulated data images shows the potential in generating reliable hyperspectral images of methane plumes, and therefore the use of GAN-generated images in training a model to detect methane plumes better than the photoshopped data used to resolve the problem of lack of data in the research paper. Due to performance limitations, we weren't able to use techniques like cross-validation to tune different hyperparameters of the GAN more to see if its result became as good as we wanted it to. If we were to continue working on this project, we would like to potentially see if there are any improvements we can make to either output decent results within fewer epochs or run each epoch quicker. The high performance of the discriminator was expected as it was trained on the training data and tested with the test data, but the high performance of CNN was unexpected. We believe it might have to do with the images we generated with random noise and labeled as images without a plume. The model was likely overfitting a lot on the data, but we do not have other images without plumes to train the model on. One future consideration, apart from finding more data, is to find a way to generate more realistic images with no plumes in them. Overall, even though the result of our classification was abnormally high, we likely cannot compare our results with the results from the research[1], but GAN does show potential in being used to generate more reliable images in the event of a lack of data.

If we were to continue working on this we might want to try implementing a CycleGAN. We could use real plume image data to generate false positive images. These false methane plume detections described in the paper[1] as a result of other things in the atmosphere having a high response from the radiance used by satellites like PRISMA, if we can identify them better would increase our detection precision more. A Semi-Supervised GAN (SGAN) could also be useful given the small amount of high-quality plume data available but large amounts of lower-quality ones and false positives.

Figure 1.

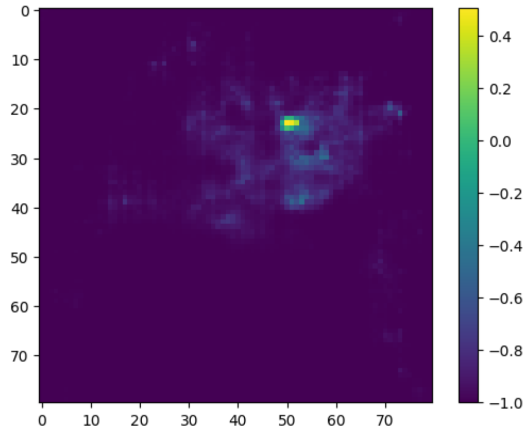


Figure 2

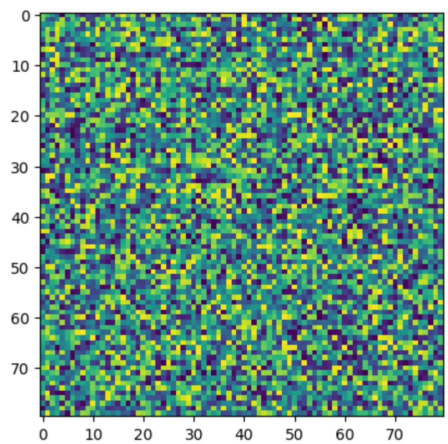


Figure 3(left: image from original data, right: GAN generated image)

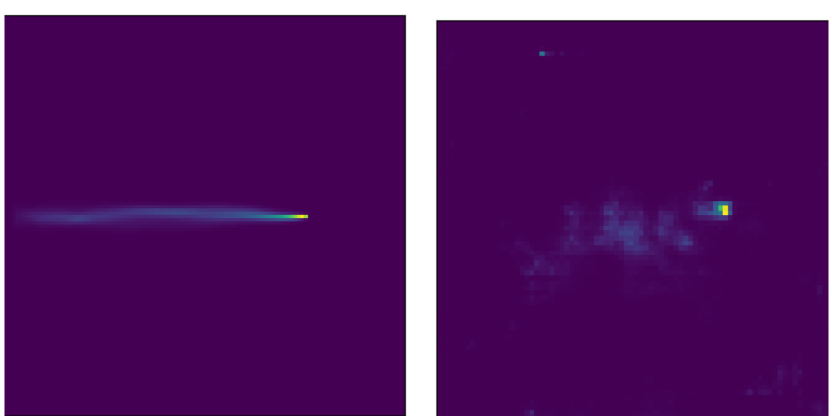


Figure 4

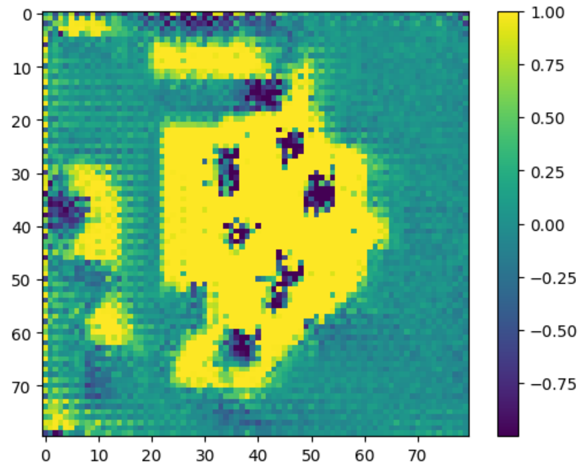


Figure 5(Top: DCGAN Discriminator, Bottom: CNN)

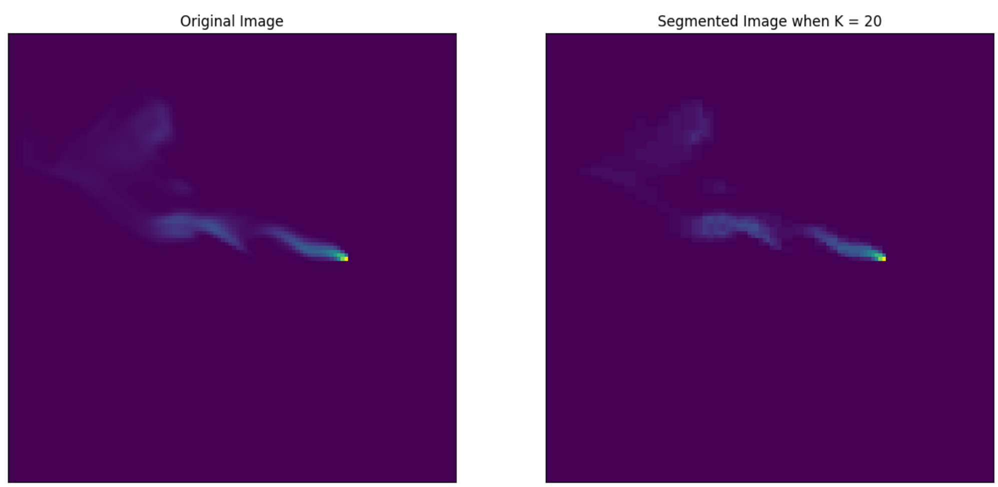
```
-------------  ------------------  ------------------
precision_dis  recall_dis          f1_score_dis
1.0            0.7464788732394366  0.8548387096774194
-------------  ------------------  ------------------

-------------  ----------  ------------
precision_cnn  recall_cnn  f1_score_cnn
1.0            1.0         1.0
-------------  ----------  ------------
```

Figure 6

	detection metrics		
	precision	recall	f1-score
Transfer Learning	0.28	0.53	0.37
Plumes Transfer	0.88	0.42	0.57

Figure 7



References

1. Groshenry, Alexis, et al. "Detecting Methane Plumes using PRISMA: Deep Learning Model and Data Augmentation." arXiv preprint arXiv:2211.15429 (2022).
2. Varon, D. J., Jacob, D. J., McKeever, J., Jervis, D., Durak, B. O. A., Xia, Y., and Huang, Y.: Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes, Atmos. Meas. Tech., 11, 5673–5686, https://doi.org/10.5194/amt-11-5673-2018, 2018.
3. Ge, Mengpin. “World Greenhouse Gas Emissions: 2016”, https://www.wri.org/data/world-greenhouse-gas-emissions-2016