I tried to expand on the work of previous researchers who attempted to accurately detect the location and extent of methane plumes using spectrograph data from the high resolution PRISMA satellite. My approach differed in how I handled the problem of having a very limited amount of high-quality methane plume image data. I decided to use a deep-learning-based algorithm called generative adversarial networks(GAN) to generate a more comprehensive set of plume images from the PRISMA data. I 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.

After data was split into train and test sets, data augmentation techniques such as Random Crop and Rotation were used on the training data to increase the size tenfold. Crop and rotate were chosen because they resulted in images that still resemble methane plumes. Both the training and test data were resized to 80x80 pixels because that allowed for training to be more efficient while not compromising the quality of the images. Data was then batched and shuffled in preparation for training, with a standard batch size of 128 and a buffer size of 3500 to accommodate for the new data size.

I tried altering the number of transposed convolutional layers in both the discriminator and generator(both at the same time and separately) to observe the results. I 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 having two networks, the discriminator and generator, essentially working opposite each other.

A simple K-means image segmentation technique was used on the original plume images, and I found that perhaps due to the relative simplicity of the simulated images, even a basic K-means technique such as this was successful in identifying the shape, size, and location of the plumes in the images[Figure 7]. 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 was not normalized and with PCA provided unsuccessful results even after 500 epochs(Figure 4).