Proceedings of the ASME 2025
International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC-CIE2025
August 17–20, 2025, Anaheim, CA

IDETC-CIE2025-168671

AVIRIS-NG-LIKE SMART VIRTUAL REMOTE SENSING VIA SPECTRA-AWARE PHYSICS INFORMED GANS

Sachin Giri[1,2], Rafal Krzysiak[1], Derek Hollenbeck[1], YangQuan Chen[1,2]

[1]Mechatronics Embedded Systems and Automation Lab, Dept. of Mechanical Engineering, University of California, Merced, CA 95343, USA
[2]EECS Graduate Program, University of California, Merced, CA 95343, USA

ABSTRACT

This paper aims to create a physics informed virtual replica of the hyperspectral image captured by NASA's Airborne Visible InfraRed Imaging Spectrometer - Next Generation (AVIRIS-NG) sensor equipped on manned aircraft. Few image samples are selected from study site around New Mexico, USA from flight mission ran in 2019. Out of 425 bands, 8 bands are utilized. For each band, reflectance spectra are chosen from United States Geological Survey (USGS) based on site specific geographical features. These spectras are infused during the image generation process with the correction check using matched filters. Moreover, we include the methane plumes along with other closely related hydrocarbons during the image generation process. Generative Adversarial Networks (GANs) architecture is employed with physics informed loss function for generating realistic and physically plausible images. Additionally, we also propose a new light weight dataset for creating the virtual replica of the AVIRIS-NG sensor on selected 8 bands in the visible light spectrum and the short-wave infrared region.

Keywords: AVIRIS-NG, GANS, Physics Informed virtual replicas, Matched Filters

1. INTRODUCTION

AVIRIS-NG [1] was developed by National Aeronautics and Space Administration - Jet Propulsion Laboratory (NASA-JPL) to provide continued access to high signal-to-noise ratio imaging spectroscopy measurements in the solar reflected spectral range. Numerous flight missions were ran in various parts of the world including North America, Europe, South America, India. While the dataset from the hyperspectral sensor has been widely adopted in landmass classification, crop detection, green house gas mapping, tar leak detection and other remote sensing applications, missions using such technology are not only costly but also demand considerable skilled manpower for meticulous post-processing and analysis. Moreover, dynamic changes in geography, environmental conditions, and other critical factors necessitate repeated monitoring. In this research, we propose to create synthetic hyperspectral replicas that accurately mimic the AVIRIS-NG sensor response characteristics, leveraging datasets from previous missions to simulate and analyze sensor performance efficiently without the need for frequent, costly field deployments.

Different methodologies have been adopted to synthesize the Hyperspectral images (HSI). Jakubowski et al. [2] generated Hyperspectral image using the Digital Imaging and Remote Sensing Image Generation Model (DIRSIG) based on numerous spectral and spatial ground truth measurements. Chang [3] designed a set of standardized synthetic hyperspectral images with two types of scenarios, Target Implantation(TI) and Target Embeddedness (TE), to simulate how a target can be inserted into the image background. Badola et al. [4] synthesize AVIRIS-NG data from widely available Sentinel-2 multispectral data using the Universal Pattern Decomposition Method (UPDM) spectral unmixing technique.

Recently, artificial algorithms (AI) algorithms are used for HSI generation. Generative Adversarial Networks [5] are a type of artificial intelligence algorithm designed to solve the generative modeling problem. This algorithm has been widely used for generation of hyperspectral images [6] with popular Pavia University [7] and Indian Pines datasets [8]. Another work is HSIGene [9], HSI generation foundation model which is based on latent diffusion and supports multi condition control, where Rectangular Guided Attention Network (RGAN) is trained for guided HSI super resolution. Liu et al. [10] proposed Physics-informed

Corresponding author: YangQuan Chen. ychen53@ucmerced.edu and +1-209-2284672. The authors are all supported by the Center for Methane Emission Research and Innovation (CMERI methane.ucmerced.edu) through the Climate Action Seed Funds grant (2023-2026) at the University of California, Merced. mechatronics.ucmerced.edu

Deep Adversarial Spectral Synthesis (PDASS) for hyperspectral generation using GANs. None of these AI based methods employ the generation of AVIRIS-NG hyperspectral data.

Methane, which traps 80 times more heat than carbon dioxide,is a significant green house gas that contributes to climate change and global warming. We also explore different methane retrieval algorithms applied to AVIRIS-NG hyperspectral images. Kuo [11] proposed a model for generating synthetic images of plumes as viewed from a hyperspectral sensor using DIRSIG. However, this method does not account for all the physical interactions that occur between the plume and the environment. Thorpe [12] applied the iterative maximum a posteriori differential optical absorption spectroscopy (IMAP-DOAS) method to AVIRIS-NG data and generated gas concentration maps for methane, carbon dioxide, and water vapor plumes. This method is computationally expensive. Foote [13] presented a computationally efficient algorithm that applies sparsity and an albedo correction to matched filter retrieval of trace gas concentration path length. HyperSTARCOP [14] improvised on methane retrieval using the features from classical methane enhancement products and outperforms strong matched filter baseline [13].

We selected the HyperSTARCOP study site where strong methane plume emissions are captured and propose a novel framework to generate a virtual replica of AVIRIS-NG sensor using a GAN architecture that incorporate the methane spectra from known methane emission sites, along with 74 other spectras for 7 different classes of materials using a USGS spectral library. Moreover, we propose a lightweight dataset for AVIRIS-NG image generation using Generative Adversarial Network models.

2. METHODS

2.1 AVIRIS-NG

Airborne Visible Infrared Imaging Spectrometer - Next Generation (AVIRIS-NG), using a push broom spectral mapping system, measures the wavelength range from 380 nanometer (nm) to 2510 nm with 5 nm spectral resolution. The spatial resolution or approximate ground sampling is 3 to 8 meter. AVIRIS-NG L2 data [15] is a product that provides orthocorrected and atmospherically corrected surface reflectance data. This dataset, free from distortions due to sensor geometry and accounts for solar irradiance variations and atmospheric absorption effects, has a total of 425 bands. We use the total of 8 bands, 3 RGB and 5 from Short wave infrared(SWIR).

2.2 Spectral Library and Hyperspectral Image model

United States Geological Survey (USGS) spectral library version 7b [16] provides the spectral signatures for seven different classes: Artificial Materials, Coatings, Liquids, Minerals, Organic Compounds, Soil and Mixtures, and Vegetation, with a total of 2457 types of spectra collected from these materials. Based on the geography, vegetation, landscape, roads and buildings and other site features, we select spectral signatures for 75 types of materials. In addition to methane, we also select spectral signatures for nine closely related hydrocarbons, including ethane, ethylene, pentane, and others, to enhance the specificity and robustness of our analysis. Since the study location is mostly desert, 10 different species of vegetations are chosen. The exact number of materials chosen for each class is shown in Table 1. The selected reflectance signatures are linearly interpolated across eight bands of the AVIRIS-NG images, providing each material type with a distinct spectral signature value across all eight selected bands.

TABLE 1: SELECTION OF SPECTRAL SIGNATURES

Class	USGS	Ours
Artificial Materials	290	10
Coatings	12	11
Liquids	24	14
Minerals	1276	0
Organic Compounds	360	10
Soils and Mixtures	209	20
Vegetations	286	10
Total	**2457**	**75**

With the linear mixing model, a single hyperspectral pixel P can be represented as:

$$P = \sum_{i=1}^{N} a_i r_i + n, \quad \text{subject to} \quad a_i \geq 0, \quad \sum_{i=1}^{N} a_i = 1, \tag{1}$$

where N represents the number of endmembers or materials. $P \in \mathbb{R}^L$ is the observed pixel spectrum of L dimensions, with L denoting the number of spectral bands. a_i is the abundance of the i-th endmember in a pixel, signifying the fractional contribution of a specific material. For physical validity, abundance values must be non negative and sum to one. r_i is the spectral signature of the i-th endmember, and n represents perturbations including sensor noise and modeling errors.

According to the AVIRIS-NG sensor specifications in NASA's COMEX Final Report [17], the instrument achieves a signal-to-noise ratio (SNR) exceeding 2000 at 600 nm and 1000 at 2200 nm, corresponding to over 66 dB and 60 dB respectively. It also maintains a radiometric accuracy within 5%, implying a maximum relative uncertainty of 5%. In our formulation, the additive noise n is modeled as zero mean Gaussian noise:

$$n \sim \mathcal{N}(0, \Sigma),$$

where $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_L^2)$ is a diagonal covariance matrix with band dependent variances defined by the signal-to-noise ratio:

$$\sigma_\ell^2 = \frac{\mathbb{E}[P_\ell^2]}{10^{\text{SNR}_\ell/10}}, \quad \ell = 1, \ldots, L. \tag{2}$$

To simulate realistic field conditions with surface variability, reflectance darkening, and atmospheric distortion, we use conservative empirical SNR values for each of the $L = 8$ spectral bands. These include three visible bands (460 nm, 550 nm, 640 nm) with SNRs of 45 dB each, and five shortwave infrared (SWIR) bands (2004 nm, 2109 nm, 2310 nm, 2350 nm, 2360 nm) with SNRs of 38, 35, 35, 33, and 33 dB respectively. These values were used to construct the per band noise model in training.

Liu et al. [10] employed a generator network that accepts an RGB image as input to produce an abundance map. Similarly, our generator function $G(I;\theta_g)$ uses an RGB image I, extracted from the original 8 channel AVIRIS-NG hyperspectral image, as input, where θ_g denotes the trainable parameters of the generator. Hence our final image model becomes:

$$P = \sum_{i=1}^{N} a_i r_i + n, \quad \text{where} \quad a_i = G_i(I;\theta_g), \quad n \sim \mathcal{N}(0,\Sigma). \quad (3)$$

2.3 Matched Filters

A matched filter (MF) [18] is an optimal linear filter designed to maximize the signal-to-noise ratio (SNR) at its output when detecting the presence of a known signal against background noise. Additionally, the matched filter can be intuitively viewed as performing a correlation between the observed spectral signal and the known target spectrum, effectively identifying pixels with spectral characteristics closely resembling the target material. It has been widely applied in hyperspectral imaging due to its robustness in differentiating target signatures from complex backgrounds, especially in scenarios with low SNR or subtle spectral differences.

We compute the matched filter output, $y_{MF}(P)$, for each pixel by comparing the pixel spectral data vector, P, to the known target spectrum, s_t, adjusted by the background statistics (mean μ and covariance Σ):

$$y_{MF}(P) = \frac{(P-\mu)^T \Sigma^{-1}(s_t-\mu)}{(s_t-\mu)^T \Sigma^{-1}(s_t-\mu)}. \quad (4)$$

The known target spectrum s_t is derived from a USGS spectral library as shown in Table 1.

2.4 Network Architecture

Generative Adversarial Networks (GANs) [5] have an ideal architecture that contain both a generator and discriminator to be applied for synthetic hyperspectral image generation. The detailed diagram for the network architecture is shown in (Figure 1). The RGB portion of the hyperspectral image was extracted and fed into the generator network whose output is the abundance map. The abundance map for each object type in each pixel location is multiplied with the corresponding spectral signatures to generate the hyperspectral image, shown mathematically in (3). A U-net architecture [19] is employed as a generator function. Spatial and spectral discriminators are used, where the spatial discriminator evaluates the spatial authenticity of generated hyperspectral images, and the spectral discriminator assesses the spectral fidelity by distinguishing generated spectral signatures from the real ones, thereby enhancing the overall quality of synthesized hyperspectral images.

Multiple loss terms were included to train the GAN effectively to ensure accurate hyperspectral image reconstruction. The generator loss combines three distinct components: the pixel-wise reconstruction loss (L1 loss), the cosine similarity loss for spectral shape alignment, and the matched filter (MF) loss to enhance detection accuracy of target spectral signatures.

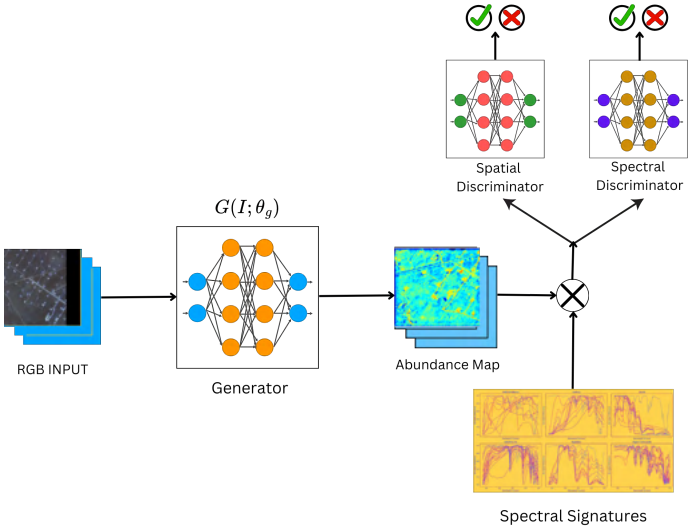


FIGURE 1: NETWORK ARCHITECTURE

Let $x_{\text{gen}}^{(i)}$ and $x_{\text{real}}^{(i)}$ denote the generated and real hyperspectral images respectively, each having dimensions (C,H,W), where C is the number of spectral channels, and H,W represent spatial dimensions. N is the batch size.

L1 Loss The L1 loss evaluates pixel wise absolute differences between generated and real hyperspectral images, promoting precise reconstruction of spectral intensities:

$$\mathcal{L}_{L1} = \frac{1}{N}\sum_{i=1}^{N} \|x_{gen}^{(i)} - x_{real}^{(i)}\|_1. \quad (5)$$

Cosine Similarity Loss The cosine similarity loss measures the alignment between spectral vectors at each pixel location, ensuring generated spectra closely match the shape and direction of real spectra:

$$\mathcal{L}_{COS} = 1 - \frac{1}{N}\sum_{i=1}^{N} \frac{x_{gen}^{(i)} \cdot x_{real}^{(i)}}{\|x_{gen}^{(i)}\| \|x_{real}^{(i)}\|}. \quad (6)$$

Matched Filter Loss The matched filter loss was calculated by comparing the matched filter responses of the generated and real images; ensuring the accurate spectral representation of specific target materials,

$$\mathcal{L}_{MF} = \lambda_{MF} \cdot \frac{1}{N}\sum_{i=1}^{N} \|y_{MF}(x_{gen}^{(i)}) - y_{MF}(x_{real}^{(i)})\|_2^2, \quad (7)$$

where the matched filter response for a pixel spectrum P is computed in (4). Out of 75 different material types, we selected 11 target spectrum s_t, which are highly abundant in the desert type geography for computing the matched filter loss. Namely, the selected target spectrum object types are sand, desert varnish, concrete road, asphalt road, sheet metal, painted aluminum, roof, water, black brush, juniper, and methane. These material types can be seen in the study site discussed in next section.

Spectral Angle Loss The Spectral Angle Mapper (SAM) loss measures the spectral similarity between generated and real hyperspectral images by computing the angle between their spectral vectors at each pixel. This loss ensures that the generated spectra align closely with the real spectra in terms of their angular orientation, which is critical for maintaining spectral fidelity in hyperspectral imaging:

$$\mathcal{L}_{SAM} = \frac{1}{NHW} \sum_{i=1}^{N} \sum_{h,w} \cos^{-1}\left(\frac{x_{gen}^{(i,h,w)} \cdot x_{real}^{(i,h,w)}}{\|x_{gen}^{(i,h,w)}\| \|x_{real}^{(i,h,w)}\|} \right), \quad (8)$$

where $x_{gen}^{(i,h,w)}$ and $x_{real}^{(i,h,w)}$ are the spectral vectors at pixel (h, w) for the i-th generated and real images, respectively, and N, H, W denote the batch size, height, and width of the images.

Spectral Correlation Loss The spectral correlation loss evaluates the correlation between spectral bands of the generated and real hyperspectral images, promoting consistency in inter band relationships. It is defined as:

$$\mathcal{L}_{SpecCorr} = 1 - \frac{1}{NC} \sum_{i=1}^{N} \sum_{c=1}^{C} \frac{\sum_{h=1}^{H} \sum_{w=1}^{W} (x_{gen}^{(i,c,h,w)} - \bar{x}_{gen}^{(i,c)})(x_{real}^{(i,c,h,w)} - \bar{x}_{real}^{(i,c)})}{\sqrt{\sum_{h=1}^{H} \sum_{w=1}^{W} (x_{gen}^{(i,c,h,w)} - \bar{x}_{gen}^{(i,c)})^2} \sqrt{\sum_{h=1}^{H} \sum_{w=1}^{W} (x_{real}^{(i,c,h,w)} - \bar{x}_{real}^{(i,c)})^2}}, \quad (9)$$

where $\bar{x}_{gen}^{(i,c)}$ and $\bar{x}_{real}^{(i,c)}$ are the mean intensities of the c-th channel for the i-th generated and real images, respectively, computed as $\bar{x}_{gen}^{(i,c)} = \frac{1}{HW} \sum_{h=1}^{H} \sum_{w=1}^{W} x_{gen}^{(i,c,h,w)}$ and $\bar{x}_{real}^{(i,c)} = \frac{1}{HW} \sum_{h=1}^{H} \sum_{w=1}^{W} x_{real}^{(i,c,h,w)}$, and C is the number of spectral channels. This loss enhances the preservation of spectral correlations across bands.

Combined Generator Loss The individual losses are combined to define the final generator loss function as follows:

$$\mathcal{L}_G = \lambda_{L1}\mathcal{L}_{L1} + \lambda_{COS}\mathcal{L}_{COS} + \lambda_{MF}\mathcal{L}_{MF} + \lambda_{SAM}\mathcal{L}_{SAM} + \lambda_{SpecCorr}\mathcal{L}_{SpecCorr}. \quad (10)$$

The hyperparameters λ_{L1}, λ_{COS}, λ_{MF}, λ_{SAM}, and $\lambda_{SpecCorr}$ control the contribution of each loss term to the overall optimization objective.

The adversarial losses from the spatial and spectral discriminators were optionally included during training to further enhance spatial realism and spectral fidelity. The spatial and spectral discriminators' detailed loss equations are standard binary cross-entropy (BCE) formulations, as described in the literature for GANs [5]. This combination ensures robust synthesis of hyperspectral imagery that closely resembles authentic spectral distributions and spatial textures.

Spatial Discriminator Loss The spatial discriminator loss ensures spatial realism of generated hyperspectral images by distinguishing between real and generated image pairs:

$$\mathcal{L}_{D_{spa}} = \frac{1}{2} \left[\mathcal{L}_{BCE}(D_{spa}(x_{real}), 1) + \mathcal{L}_{BCE}(D_{spa}(x_{gen}), 0) \right]. \quad (11)$$

Spectral Discriminator Loss The spectral discriminator loss assesses spectral fidelity by distinguishing real spectral signatures from generated ones:

$$\mathcal{L}_{D_{spe}} = \frac{1}{2} \left[\mathcal{L}_{BCE}(D_{spe}(x_{real}), 1) + \mathcal{L}_{BCE}(D_{spe}(x_{gen}), 0) \right] \quad (12)$$

3. EXPERIMENT

In this section, we discuss about the study location for AVIRIS-NG replica generation, data preprocessing, and experiment setup used to train our model.

3.1 Study Site

The dataset [1] we choose is around Carlsbad region in southeastern New Mexico, USA. This region is known for its proximity to oil and gas extraction sites, which aligns with our context of capturing the methane plumes in generated Hyperspectral replica. The reason for selecting this site is mainly because of known methane emission sources as predicted by HyperSTARCOP [14].

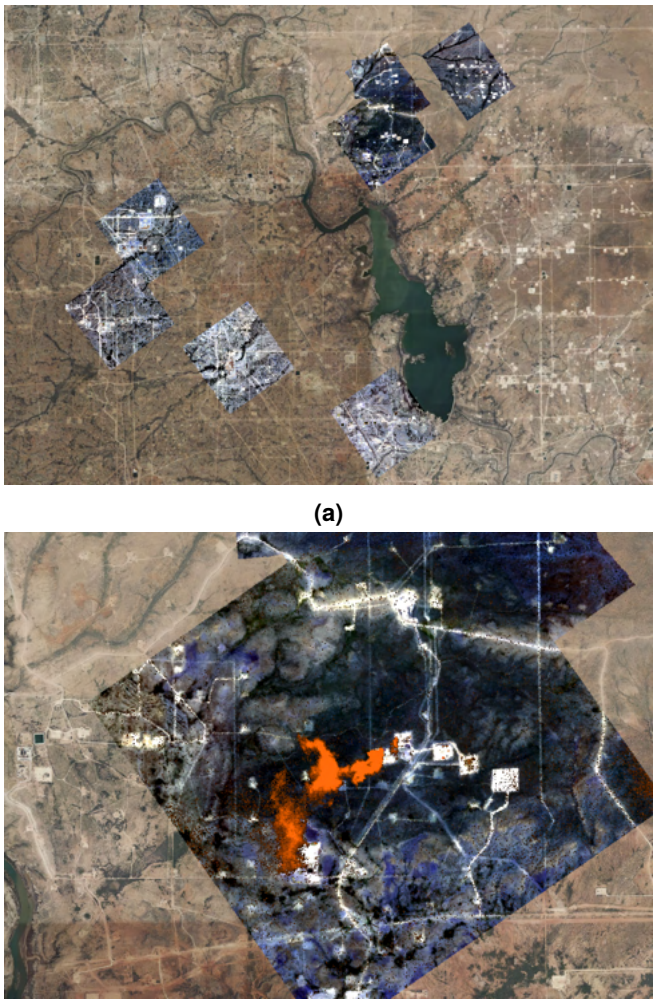
Shown in Figure 2 (a) are superimposed data samples from the AVIRIS-NG flight locations highlighting points of interest within a desert setting. While Figure 2 (b) displays a zoomed in superimposed data sample view with enhanced perspective of material types such as roads, desert, sand, vegetation, buildings and other land features as well as a methane plume in orange.

3.2 Data Preprocessing

The original AVIRIS-NG hypercube is of dimension $\approx 25000 \times 1500 \times 425$, spectral bands ranging from 380 nm to 2510 nm with 5 nm interval. We choose the flight paths from 2019 where the flight paths show abundance of methane emission from HyperSTARCOP [14] and extract 8 bands out of these 425 bands. The reason for choosing only 8 bands is that these selected channels are in RGB and SWIR spectrum and the spectral signature of the material types in the selected flight paths fall in the same spectrum. Additionally, reducing the number of bands significantly reduces the dataset complexity and computation time. The chosen band wavelengths are in the visible spectrum (460 nm, 550 nm, 640 nm) and the remaining in Shortwave infrared (SWIR) spectrum (2004 nm, 2109 nm, 2310 nm, 2350 nm, 2360 nm). We also resize the spatial dimension to 512×512, resulting in total of 56 tiles of shape $512 \times 512 \times 8$. Out of these data tiles, 50 are used for training and rest are used for testing. RGB bands are extracted and fed into the generator network as input. The total size of this dataset is 0.5 GB.

3.3 Experiment Setup

The value for hyperparamter λ_{L1} is set to 100, λ_{COS} to 1000, λ_{SAM} to 100, $\lambda_{SpecCorr}$ to 0.4, and λ_{MF} to 1 based on few random experimentation. The model was trained on an Intel Core i9-14900K processor with 64 GB of system memory. GPU acceleration was provided by two NVIDIA GeForce RTX 4090 GPUs, each equipped with 24 GB of dedicated video memory, totaling 48 GB of GPU memory utilized during training. The model is trained for a total of 1500 epochs. To address the risk of overfitting, given the small dataset size of 0.5 GB with 56 tiles,



(a)

(b)

FIGURE 2: STUDY SITES SHOWING (a) THE LOCATION OF OIL AND GAS EXTRACTION AREA, AND (b) HYPERSPECTRAL DATA HIGHLIGHTING METHANE PLUME [14] ALONG WITH OTHER GEOGRAPHICAL FEATURES.

several stabilization techniques were employed. The learning rate for the generator starts at 10^{-4}, while the discriminators use 10^{-5}, both following a cosine annealing schedule with a decay period of 800 epochs, starting after 400 epochs, and reaching a minimum of 10^{-7}. This gradual decay ensures smooth convergence and prevents large updates late in training that could lead to overfitting. The generator's multi-objective loss, combining L1, cosine, matched filter, spectral angle, and spectral correlation terms with their respective weights, acts as a regularizer to balance optimization objectives. Band wise Gaussian noise, calibrated to AVIRIS-NG's signal-to-noise ratio specifications (45 dB for RGB bands and 38 to 33 dB for SWIR bands), was added to the generated images as a form of data augmentation to enhance generalization. Additionally, the Adam optimizer uses $\beta_1 = 0.9$ and $\beta_2 = 0.999$, adjusted to improve training stability, and a conditional training strategy allows delaying discriminator updates, although in this setup, they were applied from the start to further stabilize GAN training.

RESULTS AND DISCUSSION

After successfully training the model, we ran the inference on the test images using the model trained for 1500 epochs. Six of the original images and the corresponding generated images are presented in Figure 3 (top row). Moreover, to evaluate the performance of the model, various test metrics are used to compute the reconstruction accuracy of the generated hyperspectral cube. These values are shown in Table 2.

TABLE 2: RECONSTRUCTION METRICS FOR ALL 8 SYNTHESIZED BANDS

Indicators	Value
RMSE	5333.23
MRAE	69.64
SAM	0.466
MSSIM	0.569
MPSNR	37.40
spec corr	0.625
FSIM	0.955

The reconstruction accuracy of the hyperspectral images, evaluated over six test samples, is quantified by several metrics that together highlight the performance of the proposed approach. The RMSE (Root Mean Square Error) value of 5333.23 and MRAE (Mean Relative Absolute Error) of 69.64 reflect the challenges in achieving pixelwise accuracy, attributed to the high dimensionality, dynamic range of the data, noise added and the diversity of material types. The Spectral Angle Mapper (SAM) value of 0.466 indicates good spectral fidelity, as the angular difference between reconstructed and reference spectra remains relatively small. The Mean Structural Similarity Index Measure (MSSIM) of 0.569 suggests moderate structural similarity, while the Mean Peak Signal to Noise Ratio (MPSNR) of 37.40 dB indicates an acceptable signal-to-noise ratio(SNR). Notably, the spectral correlation of 0.625 shows improved preservation of inter band relationships, and the high Feature Similarity Index Measure (FSIM) of 0.955 demonstrates excellent retention of key perceptual features.

Overall, these results reflect the effectiveness of the enhanced loss function, incorporating SAM and spectral correlation losses, in improving spectral fidelity and inter band consistency. While the higher RMSE and MRAE suggest areas for further optimization, the current method establishes a robust benchmark for synthesizing AVIRIS-NG hyperspectral images using a lightweight dataset with 8 bands in the RGB and SWIR spectrum.

4. CONCLUSION

A novel physics-informed GAN framework is presented for generating virtual replicas of AVIRIS-NG hyperspectral images using a lightweight dataset with selected RGB and SWIR bands. The integration of spectral signatures from a USGS spectral library v7b and a matched filter loss has enabled the model to capture essential spectral characteristics, as demonstrated by promising evaluation metrics such as a low SAM and high FSIM and spectral correlation. Despite moderate RMSE and MRAE,which

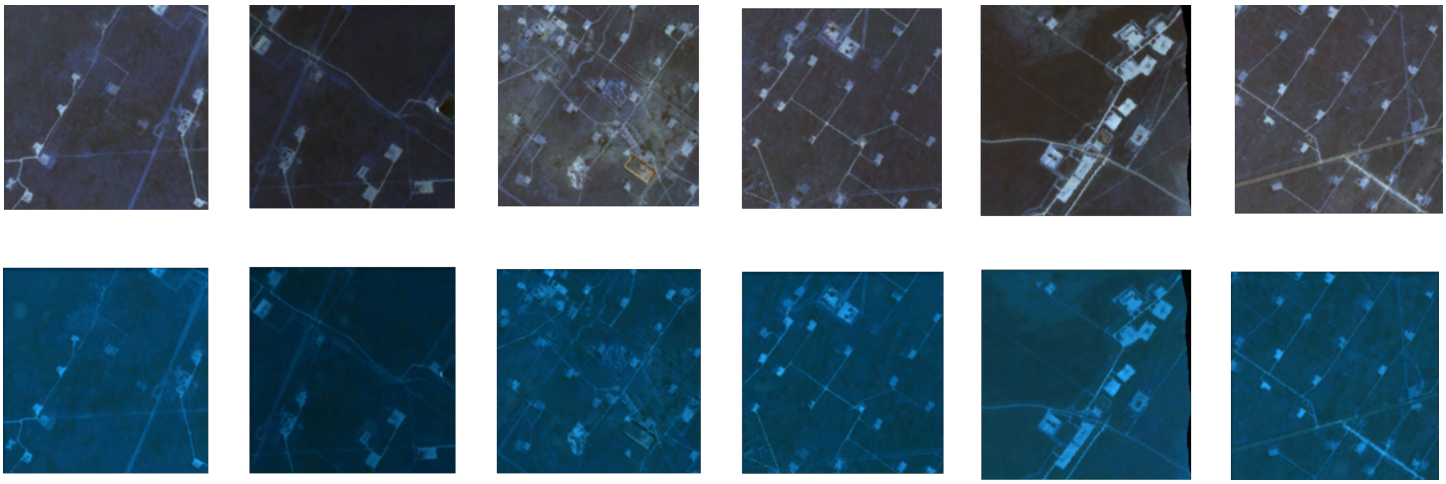


FIGURE 3: REAL IMAGES ON TOP ROW(RGB BANDS) AND CORRESPONDING SYNTHESIZED IMAGES ON BOTTOM ROW(RGB BANDS)

are also effects of added noise, these results underscore the potential of the proposed approach to deliver realistic and physically plausible hyperspectral reconstructions.

As part of the future work, we plan to implement this architecture on the full 425 bands and use advanced filtering techniques, which capture the spectral signatures of the materials with improved overall reconstruction accuracy. Moreover, we also plan to incorporate material specific filtering techniques, such as greenhouse gas filters, water level indicators, vegetation indicators, etc., in order to capture and detect such material types.

ACKNOWLEDGMENTS

We are thankful to CMERI and Climate Action seed fund grants, and all the lab members at MESA Lab, UC Merced.

REFERENCES

[1] Laboratory, Jet Propulsion. "AVIRIS-NG." National Aeronautics and Space Administration (NASA). URL https://avirisng.jpl.nasa.gov/index.html.

[2] Jakubowski, Marek K, Pogorzala, David, Hattenberger, Timothy J, Brown, Scott D and Schott, John R. "Synthetic data generation of high-resolution hyperspectral data using DIRSIG." *Imaging Spectrometry XII*, Vol. 6661: pp. 153–163. 2007. SPIE.

[3] Chang, Yu-Cherng Channing, Ren, Hsuan, Chang, Chein-I and Rand, Robert S. "How to design synthetic images to validate and evaluate hyperspectral imaging algorithms." *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV*, Vol. 6966: pp. 489–499. 2008. SPIE.

[4] Badola, Anushree, Panda, Santosh K, Roberts, Dar A, Waigl, Christine F, Jandt, Randi R and Bhatt, Uma S. "A novel method to simulate AVIRIS-NG hyperspectral image from Sentinel-2 image for improved vegetation/wildfire fuel mapping, boreal Alaska." *International Journal of Applied Earth Observation and Geoinformation* Vol. 112 (2022): p. 102891.

[5] Goodfellow, Ian, Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair, Sherjil, Courville, Aaron and Bengio, Yoshua. "Generative adversarial networks." *Communications of the ACM* Vol. 63 No. 11 (2020): pp. 139–144.

[6] Audebert, Nicolas, Le Saux, Bertrand and Lefèvre, Sébastien. "Generative adversarial networks for realistic synthesis of hyperspectral samples." *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*: pp. 4359–4362. 2018. IEEE.

[7] Computational Intelligence Group, University of the Basque Country. "Hyperspectral Remote Sensing Scenes: Pavia Centre and University." (2024). URL https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Pavia_Centre_and_University.

[8] Baumgardner, Marion, Biehl, Larry and Landgrebe, David. "220 band AVIRIS hyperspectral image data set: June 12, 1992 indian pine test site 3." .

[9] Pang, Li, Cao, Xiangyong, Tang, Datao, Xu, Shuang, Bai, Xueru, Zhou, Feng and Meng, Deyu. "Hsigene: A foundation model for hyperspectral image generation." *arXiv preprint arXiv:2409.12470* (2024).

[10] Liu, Liqin, Li, Wenyuan, Shi, Zhenwei and Zou, Zhengxia. "Physics-informed hyperspectral remote sensing image synthesis with deep conditional generative adversarial networks." *IEEE Transactions on Geoscience and Remote Sensing* Vol. 60 (2022): pp. 1–15.

[11] Kuo, Shiao Didi, Schott, John R and Chang, Chia Y. "Synthetic image generation of chemical plumes for hyperspectral applications." *Optical Engineering* Vol. 39 No. 4 (2000): pp. 1047–1056.

[12] Thorpe, Andrew K, Frankenberg, Christian, Thompson, David R, Duren, Riley M, Aubrey, Andrew D, Bue, Brian D, Green, Robert O, Gerilowski, Konstantin, Krings, Thomas, Borchardt, Jakob et al. "Airborne DOAS retrievals of methane, carbon dioxide, and water vapor concentrations at high spatial resolution: application to AVIRIS-NG." *At-*

mospheric Measurement Techniques Vol. 10 No. 10 (2017): pp. 3833–3850.

[13] Foote, Markus D, Dennison, Philip E, Thorpe, Andrew K, Thompson, David R, Jongaramrungruang, Siraput, Frankenberg, Christian and Joshi, Sarang C. "Fast and accurate retrieval of methane concentration from imaging spectrometer data using sparsity prior." *IEEE Transactions on Geoscience and Remote Sensing* Vol. 58 No. 9 (2020): pp. 6480–6492.

[14] Růžička, Vít, Mateo-Garcia, Gonzalo, Gómez-Chova, Luis, Vaughan, Anna, Guanter, Luis and Markham, Andrew. "Semantic segmentation of methane plumes with hyperspectral machine learning models." *Scientific Reports* Vol. 13 No. 1 (2023): p. 19999.

[15] ORNL DAAC. "AVIRIS-NG L2 Surface Reflectance Data Product." URL https://daac.ornl.gov/AVIRIS/guides/AVIRIS-NG_L2_Reflectance.html.

[16] Kokaly, Raymond F., Clark, Roger N., Swayze, Gregg A., Livo, K. Eric, Hoefen, Todd M., Pearson, Neil C., Wise, Richard A., Benzel, William M., Lowers, Heather A., Driscoll, Rhonda L. and Klein, Andrew J. "USGS Spectral Library Version 7." U.S. Geological Survey (2017). URL https://doi.org/10.3133/ds1035.

[17] NASA Airborne Science Program. "COMEX Final Report: CO2 and Methane Experiment (COMEX)." (2019). URL https://ntrs.nasa.gov/api/citations/20190025386/downloads/20190025386.pdf. NASA Technical Report, Document ID: 20190025386.

[18] Turin, George. "An introduction to matched filters." *IRE Transactions on Information Theory* Vol. 6 No. 3 (1960): pp. 311–329.

[19] Ronneberger, Olaf, Fischer, Philipp and Brox, Thomas. "U-net: Convolutional networks for biomedical image segmentation." *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*: pp. 234–241. 2015. Springer.

APPENDIX A. DATASET AND MODEL

The dataset can be accessed at Kaggle: https://www.kaggle.com/datasets/sachingiri348/aviris-ng-hyperspectral-data

The trained model can be accessed at: https://www.kaggle.com/models/sachingiri348/physics-informed-gan-for-aviris-ng