LD-GAN: Low-Dimensional Generative Adversarial Network for Spectral Image Generation with Variance Regularization

Emmanuel Martinez, Roman Jacome, Alejandra Hernandez-Rojas and Henry Arguello

Universidad Industrial de Santander

{emmanuel2162134@correo,roman2162474@correo,maria.hernandez26@correo,henarfu@}.uis.edu.co *

Abstract

Deep learning methods are state-of-the-art for spectral image (SI) computational tasks. However, these methods are constrained in their performance since available datasets are limited due to the highly expensive and long acquisition time. Usually, data augmentation techniques are employed to mitigate the lack of data. Surpassing classical augmentation methods, such as geometric transformations, GANs enable diverse augmentation by learning and sampling from the data distribution. Nevertheless, GAN-based SI generation is challenging since the high-dimensionality nature of this kind of data hinders the convergence of the GAN training yielding to suboptimal generation. To surmount this limitation, we propose low-dimensional GAN (LD-GAN), where we train the GAN employing a low-dimensional representation of the dataset with the latent space of a pretrained autoencoder network. Thus, we generate new low-dimensional samples which are then mapped to the SI dimension with the pretrained decoder network. Besides, we propose a statistical regularization to control the low-dimensional representation variance for the autoencoder training and to achieve high diversity of samples generated with the GAN. We validate our method LD-GAN as data augmentation strategy for compressive spectral imaging, SI super-resolution, and RBG to spectral tasks with improvements varying from 0.5 to 1 [dB] in each task respectively. We perform comparisons against the non-data augmentation training, traditional DA, and with the same GAN adjusted and trained to generate the full-sized SIs. The code of this paper can be found in https://github.com/romanjacome99/LD_GAN.git

1. Introduction

Spectral imaging involves acquiring a collection of 2D images which contain specific electromagnetic radiation from light, known as a spectral image (SI). The spectral information allows the estimation of unique characteristics and distribution of the different materials. SIs have been used for computer vision tasks related to object classification [24], image segmentation [9], salient object detection [47], and object tracking [17], which have been widely applied in fields such as medical applications [27], earth observation [30], food quality [35], and surveillance [10], among others.

The recent advances in data-driven deep learning (DL) methods have opened new frontiers for SI processing, acquisition, and its applications [32]. Some examples are in hyperspectral, and multispectral image fusion [22, 23, 43], classification [6, 24], recovery methods for snapshot compressive spectral imaging (CSI) recovery [7, 20] or mapping RGB images to SI [44]. While one of the main reasons for the great success of DL in a wide range of applications is that the models can extract the intrinsic structure of large datasets [26] improving its generalization performance, in SI applications the datasets are limited in the number of available samples due to the expensive and long acquisition times [14]. Thus the performance of the DL methods is still restricted to the constrained available data.

To address this issue, data augmentation (DA) strategies are employed [39]. Traditional DA performs geometrical transformations, such as flipping, rotation, and resizing the original dataset to generate new training samples. However, these approaches generate few invariances on the original dataset, which are easily learned by the network. Recent approaches employ generative adversarial networks (GANs) [13] to synthesize new samples based on learning the probability distribution of the original dataset and generating synthetic samples from the learned distribution [2]. In the context of SI generation, many works have been developed as a data augmentation strategy for SI classification [15], where conditional GANs are employed to generate new labeled data. A more related work is SHS-GAN [16], which is a conditional GAN that creates new SI samples by mapping SIs from RGB images through a generator and employs a critic network to inject spectral information in the

---

mapped SIs. Also, methods based on variational autoencoders (VAE) have been proposed to generate only spectral signatures in [33]. Unlike current SI generation methods that are conditioned on labels or RGB images, we look to generate new SI samples from unconditional GANs leveraging the available spectral image datasets without additional information.

However, training GAN for SI generation is challenging due to the high dimensionality and the limited training of SI samples. Theoretical studies describe these issues by deriving the convergence rate of the GAN estimator, yielding that for high dimensional data and a low number of samples, the convergence rate decreases, producing a sub-optimal performance of the sample generation [5, 19, 28]. Experimentally, projected GAN [38] has addressed this issue by training GANs using LD image datasets obtained from several feature or codification networks. Unlike the projected GAN, we look to develop a straightforward strategy for generating a unique LD representation of the SIs for training a GAN. This approach differs from existing generative models based on dimensionality reduction such as VAE [25]. These methods aim to learn the dataset distribution via latent variables that represent the mean and variance of the distribution and with a Kullback-Leibler divergence regularization approximate this distribution to Gaussian. Here, we propose to learn the latent distribution with a GAN and decode the image with a pre-trained decoder network. Additionally LD representations of SIs are very suitable since the high dimensionality of SIs produces high information redundancy in this type of data. Other SI applications have harnessed this in tasks such as the recovery of SIs from CSI measurements [29], classification [1] or unmixing [45].

The proposed LD-GAN for generating SIs trains an autoencoder (AE) network using a SI dataset to obtain an LD image dataset. Then, a GAN is trained adversarially using the LD image dataset to generate new LD image samples. Finally, the generated LD image samples are decoded through the AE network to obtain generated SIs. We propose a statistical regularization function over the LD representation images in the AE training. This regularization function promotes a variance minimization on LD space along the batch of the training data, thus having a more compact representation of the SI, yielding optimal decoding. This criterion follows from contraction AE that improves the LD representation by reducing the variability of the representation concerning the decoder's input [37]. Additionally, since we are reducing the dimensionality of the generated data, in order to produce more diverse LD samples that results in more variability of the SI, the generated LD images are regularized for the GAN training by maximizing the variance of the LD image.

Since we are proposing an LD adversarial training method, we also develop a high dimensional adversarial training with the same GAN model in order to perform comparisons but adjusting the subnetworks to generate the full-sized SIs, which we will call from now on S-GAN. Experimental results show that the proposed LD-GAN improves upon the convergence of the GAN compared with S-GAN. This improvement is further supported by the visual quality of the generated SIs, which show that the proposed method produces more realistic spatial and spectral information compared to the S-GAN. We validate the proposed LD-GAN as a DA technique for CSI recovery, SI spatial-spectral super-resolution, and RGB to spectral tasks. We present improvements varying from 0.5 to 1 [dB] with the proposed method compared with the non-DA training, traditional DA, and with a GAN trained to generate the full spectral image.

The rest of the paper is organized as follows. Section 2 is formulated as the traditional GAN. Thereafter, in section 3, we formulate the proposed LD-GAN. Then, section 4 describes the statistical regularization over the AE latent space and the GAN-generated LD samples. Section 5 presents the mathematical formulation of the SI applications employed to validate the proposed methods. The experimental results are disseminated in section 6. Finally, conclusions are presented in 7.

2. Generative Adversarial Network

Given a dataset $\left\{\boldsymbol{x}^{(i)}\right\}_{i=1}^{K}$, with K images and $\boldsymbol{x}^{(i)} \in \mathbb{R}^{MNL}$, where M, N are the spatial dimensions and L the number of spectral bands and assuming that the dataset follows a distribution $p_{\text{data}}(\boldsymbol{x})$, which is unknown. Then, a generative network $\mathcal{G}$ will learn the distribution p_g from data $\boldsymbol{x}$. For the generative network, a prior is assumed on input noise variables $p_{\boldsymbol{z}}(\boldsymbol{z})$ where $\boldsymbol{z} \in \mathbb{R}^m$ and m is the random variable dimension, usually a Gaussian distribution, which are mapped to the desired generated images. Then, a discriminative network $\mathcal{D}$ is also defined. This network will receive the generated samples or dataset samples. The loss function of a GAN [13] is

$$\mathcal{L}_{gan}(\mathcal{D}, \mathcal{G}, \mathcal{I}) = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})} \left[\log\left(\mathcal{D}(\mathcal{I}(\boldsymbol{x}))\right) \right] + \\ \mathbb{E}_{\boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z})} \left[\log\left(1 - \mathcal{D}\left((\mathcal{G}(\boldsymbol{z}))\right)\right) \right], \quad (1)$$

where $\mathcal{G}$ and $\mathcal{D}$ are the generator and distriminator networks, respectively. Additionally, $\mathcal{I}$ is an identity operator. In the adversarial training, each network $\{\mathcal{G}, \mathcal{D}\}$ competes to achieve its goals: $\mathcal{G}$ will generate fake samples, and $\mathcal{D}$ will predict if the received samples are real or fake. This is called adversarial training and can be represented as

$$\{\hat{\mathcal{G}}, \hat{\mathcal{D}}\} = \underset{\mathcal{D}}{\arg\min} \, \underset{\mathcal{G}}{\arg\max} \, \mathcal{L}_{gan}(\mathcal{D}, \mathcal{G}, \mathcal{I}). \quad (2)$$

Despite GANs having achieved state-of-the-art results in the generation of images, achieving proper convergence of
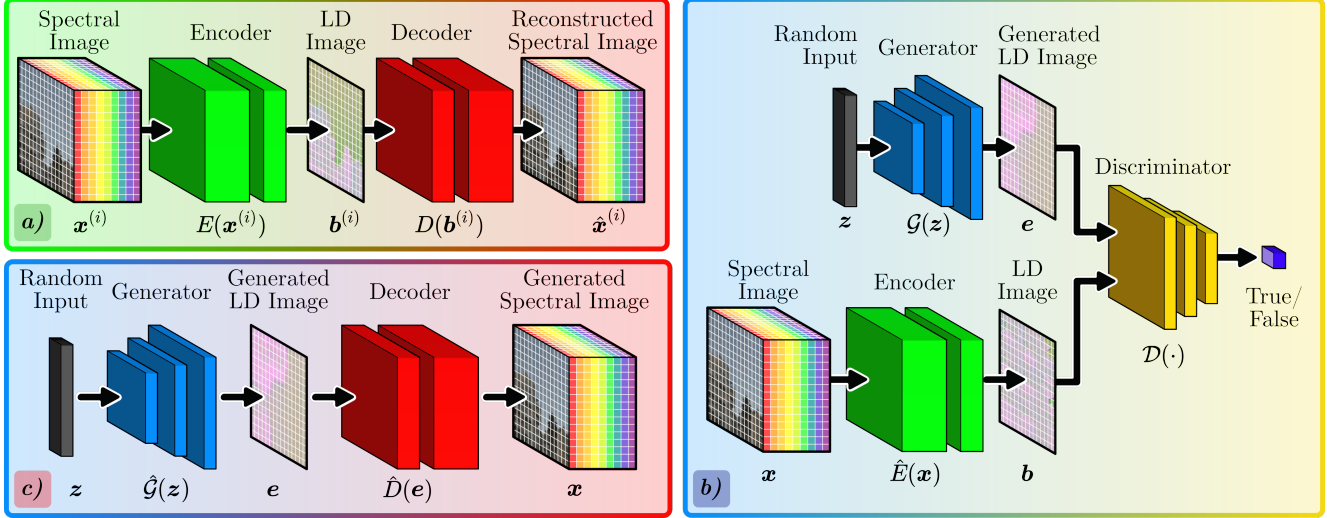
Figure 1. Autoencoder and LD-GAN networks. a) The AE network allows obtaining an LD representation of the SI through an encoder subnetwork and then reconstructs the original SI with a decoder subnetwork. b) The LD representation obtained by the AE is employed to train a GAN to generate new LD representation images. c) Once the GAN is trained, the new LD images can be decoded to recover SIs.

GAN through adversarial training is challenging when the employed dataset samples have high dimensionality and it is limited in the number of samples, such as the generation of SIs. In the following section, we will analyze the convergence of the GANs.

2.1. Analysis on the GAN Convergence Rate

Theoretical studies of the convergence properties of GAN have been carried out concerning the dimensionality of the data and the number of training samples. For instance, in [28] derives an asymptotic bound of the total-variation metric between an estimated distribution and the real source data distribution, where the sample complexity relies on the squared dimension of the data (in our case MNL), and a logarithmic factor of the number of samples of the training dataset [28, Theorem 19]. Also, in [19] provides an error analysis of GANs through convergence rates of the integral probability metric, yielding sample complexity that increases in terms of the number of training samples and the dimension of the data [19, Theorem 5]. Moreover, the convergence rate is improved when the source data can be represented in an LD manifold [19, Theorem 19]. Consequently, based on these theoretical insights on the GANs convergence, we propose generating an LD representation of the SIs to improve convergence on GANs and generate high-spectral dimensional SIs by decoding them through the pretrained AE network, described in the following section.

3. Generative Adversarial Networks with Low-Dimensional Representation

This section will describe the proposed LD-GAN where LD samples are generated by a GAN and the synthetic SIs

are obtained by decoding the LD via a decoder network.

3.1. Spectral Image Autoencoder

An AE network is employed to obtain an LD representation of a SI dataset $\{\boldsymbol{x}^{(i)}\}_{i=1}^{K}$ with K samples. Two subnetworks give the structure of an AE, as shown in Fig. 1(a). An encoder network E, which will compress the spectral information of SIs, and a decoder network D, which will decode the underlying SI. The optimization of these subnetworks is given by

$$\{\hat{E}, \hat{D}\} = \sum_{i=1}^{K} ||\boldsymbol{x}^{(i)} - D(E(\boldsymbol{x}^{(i)}))||_2^2, \qquad (3)$$

where $\hat{E}$ and $\hat{D}$ are the encoder and decoder, respectively. Thus, an LD representation of a SI can be obtained as $\boldsymbol{b}^{(i)} = E(\boldsymbol{x}^{(i)})$ where $\boldsymbol{b}^{(i)} \in \mathbb{R}^{MNc}$ and $c < L$. Finally, SI can be recovered by decoding this LD image as $\hat{\boldsymbol{x}}^{(i)} = D(\boldsymbol{b}^{(i)})$.

3.2. Low-Dimensional Generative Adversarial Network

We employ a GAN that will be optimized with respect to the LD representation of a SI dataset. Then, the adversarial training from the equation 2 can be rewritten as

$$\{\hat{\mathcal{G}}, \hat{\mathcal{D}}\} = \arg\min_{\mathcal{D}} \arg\max_{\mathcal{G}} \mathcal{L}_{gan}(\mathcal{D}, \mathcal{G}, \hat{E}), \qquad (4)$$

where the generative network can generate LD images as $\boldsymbol{e} = \hat{\mathcal{G}}(\boldsymbol{z})$ and new SI samples can be obtained using the decoder as $\hat{\boldsymbol{f}} = \hat{D}(\boldsymbol{e})$.

4. Statistical Regularization for the AE and GAN

Towards improving the LD representations of the AE and the diversity of generated LD images by the GAN, we propose a variance minimization regularizer in the AE training that allows a compact representation of the SI dataset in the LD space, which improves AE recovery performance. Then, we employed a variance maximization on the generated LD space for the GAN training to produce diverse data and more quality on the generated SI dataset. This variance-based regularization criterion has also been used for self-supervised learning [8] and sparse-coding [12]. First, define set $\mathcal{A} = \{e^{(i)}\}_{i=1}^{K}$ that contains the LD representation of the AE and define the set of generated LD images by GAN as $\mathcal{W} = \{d^{(i)}\}_{i=1}^{K}$. The variance of $\mathcal{A}$ is denoted by $\sigma_{\mathcal{A}}^2 \in \mathbb{R}_+^d$, and of the GAN generated representations $\sigma_{\mathcal{W}}^2 \in \mathbb{R}_+^d$ pixel-wise across the training batch dimension. The proposed statistical regularization function is given by

$$\mathcal{R}(\cdot) = \|\sigma_{(\cdot)}^2\|_2. \qquad (5)$$

Then, the regularized AE training is performed as follows

$$\{\hat{E}, \hat{D}\} = \arg\min_{E,D} ||x^{(i)} - D(E(x^{(i)}))||_2^2 + \mu_{ae}\mathcal{R}(\mathcal{A}). \qquad (6)$$

and the final GAN optimization problem is given by

$$\{\hat{\mathcal{G}}, \hat{\mathcal{D}}\} = \arg\min_{\mathcal{D}} \arg\max_{\mathcal{G}} \mathcal{L}_{gan}(\mathcal{D}, \mathcal{G}, \hat{E}) - \mu_{gan}\mathcal{R}(\mathcal{W}). \qquad (7)$$

Note that the negative sign in the variance regularization function is because we want to maximize the variance of the generated LD representation of the SI. In both optimization problems, the AE and the GAN, the parameters μ_{ae} and μ_{gan} are regularization hyperparameters that control how much we concentrate the AE LD space or how much we increase the variability of the generated images by the GAN.

5. Spectral Image Applications

To validate the performance of our proposed method, we address the following DL-based SI tasks:

CSI Recovery [7]: CSI is a technique that allows sensing the whole 3D information of the SI with a 2D sensor in a single shot [46] through coded projections of the high-dimensional data and a computational algorithm to recover the underlying SI. Several optical architectures have been proposed for CSI. We employed the coded aperture snapshot spectral imager (CASSI) [40], which is the flagship architecture in CSI [4]. This sensing phenomenon is expressed in a matrix-vector multiplication as $y = Hx + n$, where $y \in \mathbb{R}^{M(N+L-1)}$ are the compressed measurements, $H \in \mathbb{R}^{M(N+L-1) \times MNL}$ is the CASSI sensing matrix, $x \in$ $\mathbb{R}^{MNL}$ is the vectorization of the SI, and $n \in \mathbb{R}^{M(N+L-1)}$ is additive noise from the acquisition process.

Single Image Spatial-Spectral Image Super-Resolution: A well-known task for SI is to recover a high spatio-spectral resolution SI from a low spatio-spectral resolution SI [44]. In this task, the SI can be downsampled spatially and spectrally by a decimation matrix $D \in \mathbb{R}^{\frac{MN}{k_s}\left(\frac{L}{k_l}\right) \times MNL}$, where k_s and k_l represent the decimation factor for the spatial and spectral resolution of the SI, respectively. The low spatio-spectral resolution image is represented as $y = Dx + n$.

SI Recovery from RGB Images: Another task that has gathered significant attention from the research community is the mapping from RGB image to SI [3]. This task consists on recover a SI from an RGB image with a count with less spectral information considering a known spectral response function $R \in \mathbb{R}^{MN3 \times MNL}$. Then, the RGB image can be represented as $y = Rx + n$.

Since the mentioned tasks are ill-posed problems, the objective is to recover the SI x from the observed data $y = Ax$, where A could represent the sensing matrix H, the decimation matrix D or the spectral response function R, according to the selected problem. Then, we can solve the DL-based SI computational tasks through the optimization problem

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{K}\sum_{k=1}^{K} \|x^{(k)} - \mathcal{M}_{\theta}(y^{(k)})\|_2^2, \qquad (8)$$

These are ill-posed inverse problems that are challenging. Therefore, we aim to increase the performance of the network $\mathcal{M}_\theta$ for each case by adding synthetic samples generated by the LD-GAN.

6. Numerical Experiments

In this section, we perform several experiments with the AE network with different channels for the LD representation. For the GAN architecture, the deep convolutional generative adversarial network [36] was adapted to the spatial size of the employed dataset, and the experiments are performed with respect to the LD image dataset obtained from the AE network against the entire-sized SI dataset. All experiments were performed on GPU with an NVIDIA RTX 3090 graphic card. The dataset employed for all the experiments is the ARAD 1K [3], which was preprocessed, reducing the spatial resolution to 256×256 and keeping the 31 spectral bands. This dataset contains 900 samples for training and 50 samples for testing. We extract unique patches with a spatial resolution of 128×128 obtaining a total of 3600 patches for training and 200 patches for testing. Recovery performance is measured with the peak-noise-to-signal ratio (PSNR) and the structural similarity index measure (SSIM) [18].
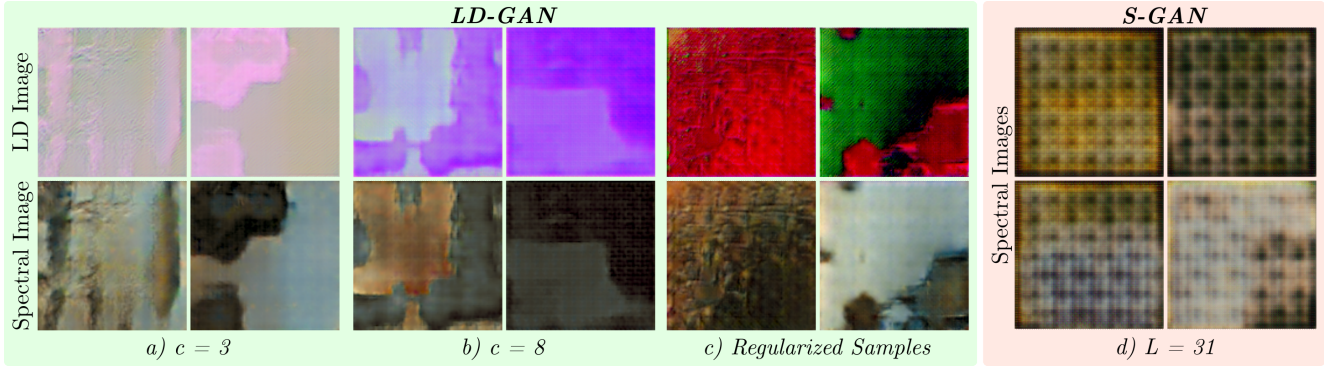
Figure 2. Generated images for an LD representation with a) 3 bands and b) 8 bands. c) Generated SI samples from LD representation employing the statistical regularizer with $\mu_{gan} = \mu_{ae} = 0.001$. For a-c) the top images refer to the generated LD images and the bottom images are the obtained SI by decoding the LD representation. d) Generated SIs without an LD representation. The bottom row represents a false RGB mapping of generated SIs from their LD representations, except for S-GAN, which SIs is directly generated by the GAN.

6.1. AE Experiments

The compression of the SIs is only performed in the spectral dimension, where several values for the number of output channels $c = \{1, 3, 8, 16\}$ of the encoder subnetwork are evaluated. The encoder and subnetwork consist of 7 convolutional layers with ReLU activations, which decrease the number of feature channels from $16c$ to c and the decoder network has also 7 layers increasing the number of features from c to $16c$ and a final convolutional layer with $L = 31$ features to match the SI spectral dimension. Table 1 shows that for the evaluated number of channels, $c = 3$ has the best performance. Thus, this level of compression will be highly considered for the performance of the SI applications when the proposed method is applied.

Channels	PSNR [dB] ↑	SSIM ↑
1	33.72	0.9513
3	39.32	0.9739
8	38.37	0.9722
16	38.31	0.9679

Table 1. AE recovery performance for different latent channel sizes. The highlighted green values are the best performance and the blue ones are the second best.

6.2. GAN Experiments

We employ the LD images from the SI dataset using the AE above mentioned to optimize the GAN. All GAN experiments were trained for 50 epochs with a learning rate of $2e^{-4}$ for both the discriminator and the generator. The generator and the discriminator networks employ 2D transpose convolution and 2D convolution layers, respectively. Each subnetwork counts with batch normalization, and Leaky ReLU activations, after the convolutional layer, following the implementation given by [36]. All the experiments with GANs employed a random input $z \in \mathbb{R}^{100} \sim \mathcal{N}(\mathbf{0}, \mathbf{1})$. In Fig. 2, some of the generated LD samples and their respective SIs are shown for the LD-GAN with $c = \{3, 8\}$,

LD-GAN with the statistical regularization, and S-GAN. To compare the effectiveness of the proposed method, we adopt the employed GAN for directly generating SIs without an LD representation, denoted as **S-GAN**. In Fig. 3, the discriminator loss function is shown for different channels c. The performance shows this behavior properly when the GAN is trained on the LD image dataset. This performance is far from the theoretical optimum when training is done with S-GAN.

We analyzed the generated spectral information with the proposed LD-GAN compared with the S-GAN. To this end, we employ a widely used SI analysis technique which is the linear mixture model [11]. This model states that a spectral signature of a given pixel of the SI is represented by the linear combination of the q most representative signatures of the SI. These signatures are denoted as endmembers and the coefficients of the linear model are named abundance. To extract the endmembers and the abundances of a SI, we employ the hyperspectral vertex component analysis (HyperVCA) [31] algorithm. Specifically, we employ HyperVCA to extract the $q = 4$ endmembers from the original ARAD dataset, from the generated dataset by the LD-GAN
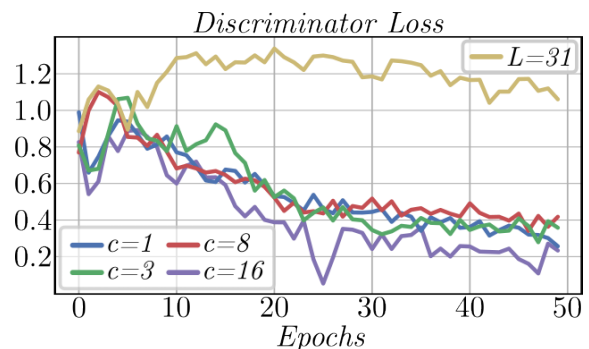

Figure 3. Discriminator loss function convergence for different channels. $L = 31$ represents the performance by S-GAN.
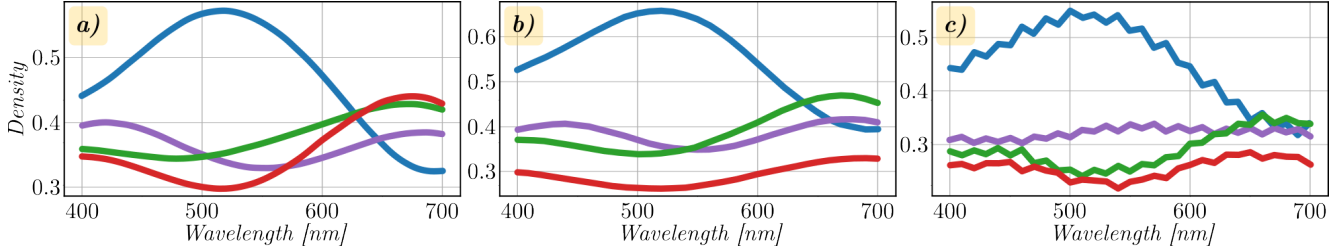
Figure 4. Linear mixture model analysis of the generated images. Extracted 4 endmembers from 100 SIs of the original SI dataset (a) the generated SI dataset with the proposed LD-GAN (b) and with the S-GAN (c).

a) CSI Recovery
PSNR [dB] | Baseline: 30.71

	1	3	8	16	31 S-GAN
100%	31.52	31.74	31.47	31.45	30.15
50%	31.52	30.79	31.21	31.25	30.45
20%	31.09	31.18	30.95	31.18	30.68

b) RGB to Spectral
PSNR [dB] | Baseline: 36.89

	1	3	8	16	31 S-GAN
100%	37.07	37.49	37.28	37.3	27.0
50%	36.95	32.52	37.15	32.42	36.62
20%	37.43	28.2	37.11	31.49	23.67

c) Spectral Super Resolution
PSNR [dB] | Baseline: 35.75

	1	3	8	16	31 S-GAN
100%	36.29	36.38	35.89	36.04	35.43
50%	35.57	34.77	35.24	35.46	35.51
20%	36.15	36.1	35.95	36.13	35.7

SSIM | Baseline: 0.86 (CSI Recovery)

	1	3	8	16	31 S-GAN
100%	0.873	0.88	0.874	0.872	0.843
50%	0.87	0.878	0.871	0.871	0.85
20%	0.868	0.873	0.869	0.87	0.855

SSIM | Baseline: 0.9816 (RGB to Spectral)

	1	3	8	16	31 S-GAN
100%	0.9839	0.9844	0.9832	0.9832	0.9251
50%	0.983	0.9615	0.9816	0.9589	0.9806
20%	0.9839	0.953	0.9831	0.9563	0.8721

SSIM | Baseline: 0.9512 (Spectral Super Resolution)

	1	3	8	16	31 S-GAN
100%	0.9581	0.9589	0.9559	0.9561	0.9501
50%	0.9541	0.9477	0.9529	0.9522	0.9506
20%	0.9569	0.9588	0.9553	0.9554	0.9514

Figure 5. SI recovery performance in terms of PSNR and SSIM employing the proposed LD-GAN considering the number of channels and compared with the S-GAN for a different number of generated SIs based on the percentage of samples from the original SI dataset as a DA strategy for SI applications: (a) CSI recovery, (b) RGB to spectral recovery, and (c) SI super-resolution.

and by the S-GAN, as shown in Fig. 4 with the mean of 100 images. The obtained results illustrate that generated mean endmembers of the proposed LD-GAN are very similar with respect to the original dataset. This result allows concluding that the generated spectral data follow similar behavior to the original dataset. While the S-GAN has many artifacts along the spectrum, which could affect negatively the training of networks a DA strategy for DL-based SI applications.

6.3. DA Applications

We validate the proposed method by generating a synthetic SI dataset and employing it as a DA for DL-based, CSI, SSSR, and RGB to spectral tasks. All of the following networks were trained for 100 epochs with a learning rate of $1e^{-3}$. From the obtained results, we report the best performance from testing images for each experiment.

- *CSI recovery setting:* The unrolled network deep spatial-spectral prior [41] was used for CSI recovery. The measurements were simulated through the CASSI system using a random CA with 50% of transmittance.

- *SI recovery from RGB images Setting:* A variation of the UNET architecture described for the SSSR case is employed. UNET-based models have been previously proposed for RGB to spectral task [44]. To obtain the RGB input image for the network training and validation, the spectral response provided by the ARAD dataset [3] was used. For the training of this network, the learning rate decays exponentially to achieve a more stable convergence.

- *SI super-resolution setting:* We employed a UNET network with 4 downsampling and 4 upsampling convolutional blocks to reconstruct the high-resolution SI from low-spatial low-spectral measurements. The selection of the type of architecture is based on recent SSSR works [42]. The spatial and spectral decimator factors are set as $k_s = 2, k_l = 4$.

6.4. Quantitative Results

For each SI application, we perform the recovery task considering each of the generated SI datasets by LD-GAN and S-GAN with 20%, 50%, and 100% new samples, (per-
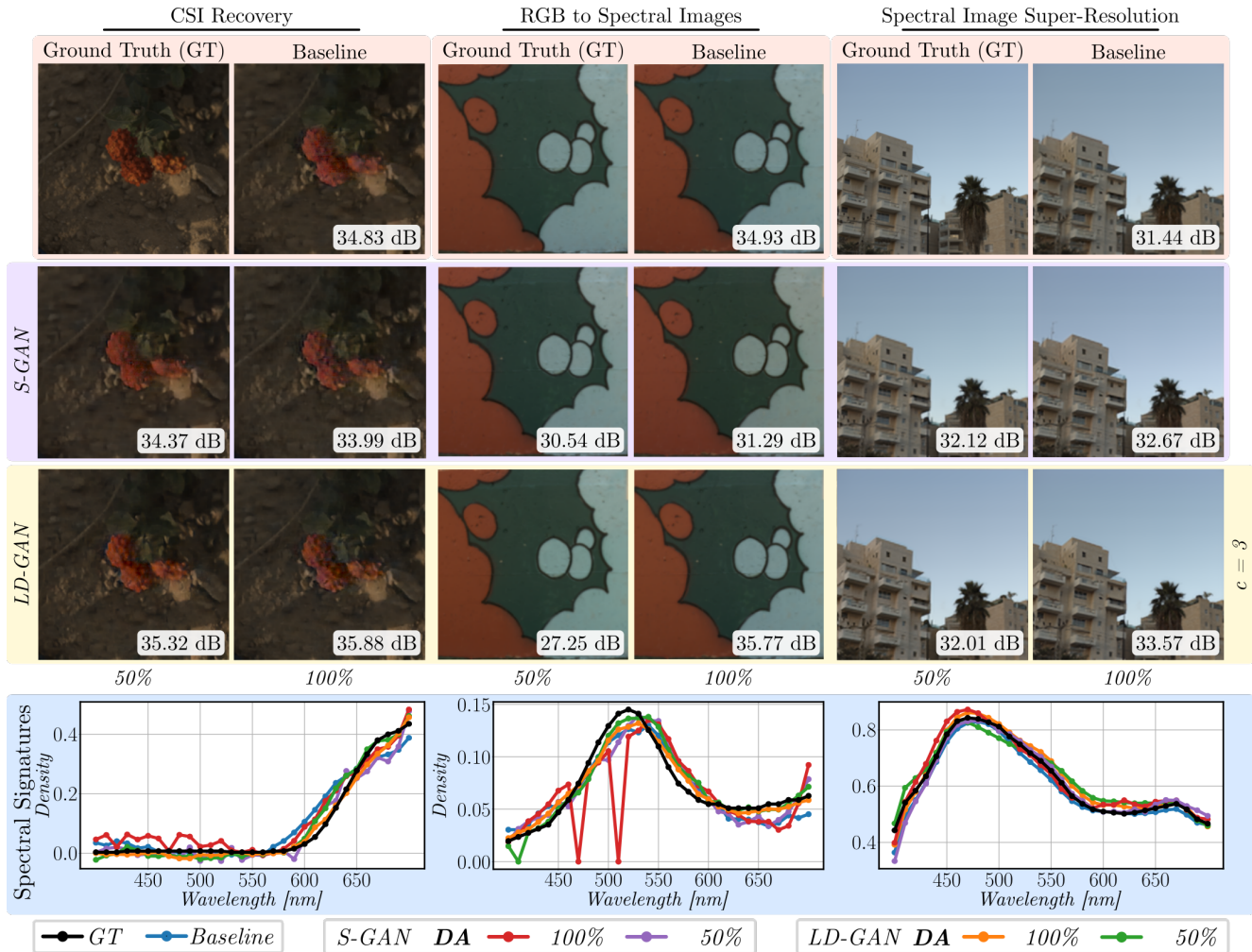
Figure 6. DL-based SI applications. Top to bottom: Ground Truth and baseline, S-GAN, and LD-GAN from 3 bands LD image generation. From left to right: SI application for baseline (no DA based on GANs), 50%, and 100% of DA based on GANs. Last row: Spectral signature from the middle coordinate for each application.

Figure 7. SI recovery performance in terms of PSNR and SSIM employing the proposed LD-GAN with different values of the regularization parameters μ_{gan} and μ_{ae} for each SI application.

centages based on the number of samples from the original dataset). As shown in Fig. 5, the higher performances in terms of PSNR are achieved by LD-GAN with $c = 3$. Furthermore, when a comparison between S-GAN and LD-

GAN is performed, we can determine that not only LD-GAN is better, but S-GAN has a lower performance than the baseline in several cases. We also observe that the greater the number of data generated employed as DA the higher improvement of the performance for the SI applications is obtained with the proposed approach.

6.5. Qualitative Results

Visual results of a reconstructed SI from the test dataset are shown in Fig. 6, where a high number of new samples allows a higher performance in terms of PSNR. The spectral signature extracted from the ground truth SI and its reconstructions shows that the training with the LD SI images achieves a density more similar to the ground truth data. The results suggest that the models trained with the synthetic samples from the proposed dataset have higher reconstruction quality in each task.
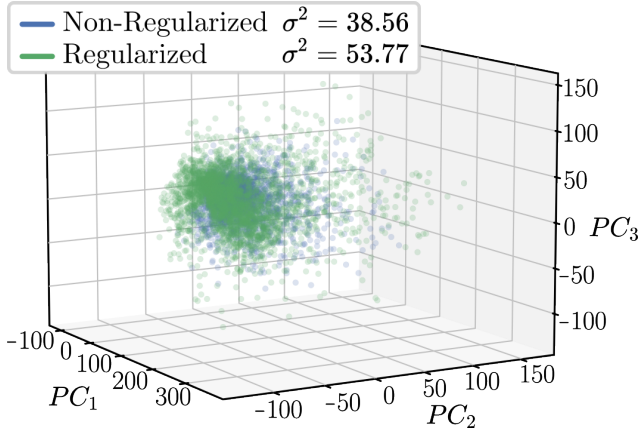
Figure 8. Three principal components of the generated dataset and its respective variance with the regularized training and the non-regularized. PC_1, PC_2, and PC_3 refer to the first, second, and third principal components of the dataset respectively.

6.6. Statistical Regularization Experiments

To validate the effectiveness of the proposed regularized training in the AE and GAN, we perform a hyperparameter study of the regularization parameters μ_{gan} and μ_{ae}. The parameters $\mu_{ae} = \mu_{gan} = \{0, 1e^{-5}, 1e^{-3}\}$ were changed for each task. The augmented SI dataset was 100% of the original dataset. Fig. 7 shows the performance of the mentioned experiment. The highest performance in each task is obtained at higher regularization parameters showing the effectiveness of this training method.

To visualize the effect of maximizing the variance of the generated LD representation in the GAN training, in Fig. 8, the three first principal components of 3000 synthesized SIs were computed for the non-regularized GAN and the proposed LD-GAN with $\mu_{ae} = \mu_{gan} = 0.001$ showing that the last one has more variability than the first one.

6.7. Comparison with traditional DA techniques

We compared the performance of the proposed method against the traditional DA consisting of rotations (rot), and vertical and horizontal flip (VFlip and HFlip), respectively, obtaining the same number of samples for each experiment reported in Table 2. The results obtained show the efficacy of the proposed method compared with the S-GAN and only geometric DA approaches. Additionally, the performance with the geometric DA achieves similar performance to the same experiments without this DA strategy, indicating that is not always helpful to use geometric DA.

7. Conclusion and Future Work

We proposed LD-GAN, a method that allows the generation of SIs through a GAN. To deal with the challenging issue of synthesizing the high-dimensional SI, we propose to train the GAN with an LD representation of a SI dataset

Method	Geometric DA	CSI PSNR [dB]	SSIM	RGB to spectral PSNR [dB]	SSIM	SISR PSNR [dB]	SSIM
Baseline	X	29.75	0.827	37.01	0.982	36.14	0.957
	✓	30.96	0.858	36.89	0.981	35.13	0.951
S-GAN	X	30.46	0.851	36.98	0.981	35.72	0.954
	✓	30.15	0.843	27.00	0.925	35.43	0.950
LD-GAN	X	31.51	0.874	37.38	0.978	36.10	0.957
	✓	31.74	0.880	37.49	0.984	36.38	0.958

Table 2. Comparison of performance for the SI computational tasks of the models trained with geometric DA and the GAN-based methods. The highlighted green values are the best performance and the blue ones are the second best

obtained via the latent space of an AE. Then the generative model synthesizes new LD samples mapped to the high-dimensional spectral image space with a decoder network. This approach significantly improves the convergence of the GAN. Moreover, a regularization based on data variability was proposed to optimally train the AE and GAN. We validate the proposed SI generation model as a DA strategy for some DL-based SI applications, such as CSI recovery, SISR, and SI recovery from RGB images, where an improvement is achieved from 0.5 to 1 dB. As future studies, improved performance of the LD-GAN can be obtained via a conditional generation [21], using, for instance, RGB images to guide the generation process. Training GAN with LD representation of the data can also be applied to other types of high-dimensional data such as video, polarized images, and 3D images for medical applications.

Additionally, VAE can also be included instead of the AE, where the LD dataset employed to train the generative network is obtained from sampling the latent distribution of the VAE. Also, beyond GAN-based models, our method can be extended to state-of-the-art approaches based on diffusion probabilistic models (DPM). Some DPM generative approach employ LD representation from a VAE to condition the generation towards meaningful latent representation [34]. Different from this approach, the extension of our method to DPM aims to perform the corrupting process and the denoising of the inverse process over the LD representations of the dataset, and then, decode the generated image to obtain the desired synthetic data.

References

[1] Bilal Alhayani and Haci Ilhan. Hyper spectral image classification using dimensionality reduction techniques. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, 5(4):71–74, 2017. 2

[2] Antreas Antoniou, Amos Storkey, and Harrison Edwards. Data augmentation generative adversarial networks. *arXiv preprint arXiv:1711.04340*, 2017. 1

[3] Boaz Arad, Radu Timofte, Rony Yahel, Nimrod Morag, Amir Bernat, Yuanhao Cai, Jing Lin, Zudi Lin, Haoqian Wang, Yulun Zhang, et al. Ntire 2022 spec-

tral recovery challenge and data set. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 863–881, 2022. 4, 6

[4] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2014. 4

[5] Sanjeev Arora, Rong Ge, Yingyu Liang, Tengyu Ma, and Yi Zhang. Generalization and equilibrium in generative adversarial nets (gans). In *International Conference on Machine Learning*, pages 224–232. PMLR, 2017. 2

[6] Jorge Bacca, Alejandra Hernandez-Rojas, and Henry Arguello. Deep coding patterns design for compressive near-infrared spectral classification. In *2022 30th European Signal Processing Conference (EUSIPCO)*, pages 548–552, 2022. 1

[7] Jorge Bacca, Emmanuel Martinez, and Henry Arguello. Computational spectral imaging: a contemporary overview. *JOSA A*, 40(4):C115–C125, 2023. 1, 4

[8] Adrien Bardes, Jean Ponce, and Yann Lecun. Vicreg: Variance-invariance-covariance regularization for self-supervised learning. In *ICLR 2022-10th International Conference on Learning Representations*, 2022. 4

[9] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson. Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. *IEEE Signal Processing Magazine*, 31(1):45–54, 2014. 1

[10] Simon Denman, Todd Lamb, Clinton Fookes, Vinod Chandran, and Sridha Sridharan. Multi-spectral fusion for surveillance systems. *Computers & Electrical Engineering*, 36(4):643–663, 2010. 1

[11] Olivier Eches, Nicolas Dobigeon, Corinne Mailhes, and Jean-Yves Tourneret. Bayesian estimation of linear mixtures using the normal compositional model. application to hyperspectral imagery. *IEEE Transactions on Image Processing*, 19(6):1403–1413, 2010. 5

[12] Katrina Evtimova and Yann LeCun. Sparse coding with multi-layer decoders using variance regularization. *arXiv preprint arXiv:2112.09214*, 2021. 4

[13] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020. 1, 2

[14] Nathan A. Hagen and Michael W. Kudenov. Review of snapshot spectral imaging technologies. *Optical Engineering*, 52(9):1 – 23, 2013. 1

[15] Renlong Hang, Feng Zhou, Qingshan Liu, and Pedram Ghamisi. Classification of hyperspectral images via multitask generative adversarial networks. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2):1424–1436, 2020. 1

[16] Jonathan Hauser, Gal Shtendel, Amit Zeligman, Amir Averbuch, and Menachem Nathan. Shs-gan: Synthetic enhancement of a natural hyperspectral database. *IEEE Transactions on Computational Imaging*, 7:505–517, 2021. 1

[17] Hien Van Nguyen, A. Banerjee, and R. Chellappa. Tracking via object reflectance using a hyperspectral video camera. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pages 44–51, 2010. 1

[18] Alain Hore and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th international conference on pattern recognition*, pages 2366–2369. IEEE, 2010. 4

[19] Jian Huang, Yuling Jiao, Zhen Li, Shiao Liu, Yang Wang, and Yunfei Yang. An error analysis of generative adversarial networks for learning distributions. *Journal of Machine Learning Research*, 23(116):1–43, 2022. 2, 3

[20] Longqian Huang, Ruichen Luo, Xu Liu, and Xiang Hao. Spectral imaging with deep learning. *Light: Science & Applications*, 11(1):1–19, 2022. 1

[21] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017. 8

[22] Roman Jacome, Jorge Bacca, and Henry Arguello. Deep-fusion: An end-to-end approach for compressive spectral image fusion. In *2021 IEEE International Conference on Image Processing (ICIP)*, pages 2903–2907. IEEE, 2021. 1

[23] Roman Jacome, Jorge Bacca, and Henry Arguello. D 2 uf: Deep coded aperture design and unrolling algorithm for compressive spectral image fusion. *IEEE Journal of Selected Topics in Signal Processing*, 2022. 1

[24] Román Jácome, Carlos López, Hans Garcia, and Henry Arguello. Deep learning-based object classification for spectral images. In Alvaro David Orjuela-Cañón, Jesus Lopez, Julián David Arias-Londoño, and Juan Carlos Figueroa-García, editors, *Applications of Computational Intelligence*, pages 147–159, Cham, 2021. Springer International Publishing. 1

[25] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 2

[26] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015. 1

[27] Chao Li, Souleymane Balla-Arabé, and Fan Yang. Embedded multi-spectral image processing for real-time medical application. *Journal of Systems Architecture*, 64:26 – 36, 2016. Real-Time Signal Processing in Embedded Systems. 1

[28] Tengyuan Liang. How well generative adversarial networks learn distributions. *The Journal of Machine Learning Research*, 22(1):10366–10406, 2021. 2, 3

[29] Brayan Monroy, Jorge Bacca, and Henry Arguello. Jr2net: a joint non-linear representation and recovery network for compressive spectral imaging. *Applied Optics*, 61(26):7757–7766, 2022. 2

[30] David Morales-Norato, Sergio Urrea, Hans Garcia, Julian Rodriguez-Ferreira, Elizabeth Martinez, Henry Arguello, Alberto Silva-Lora, Rafael Torres, Ignacio F Acero, Francisco L Hernández, et al. Hyperspectral camera as a compact payload architecture for remote sensing applications. *Applied Optics*, 62(8):C88, 2023. 1

[31] José MP Nascimento and José MB Dias. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE transactions on Geoscience and Remote Sensing*, 43(4):898–910, 2005. 5

[32] Akin Ozdemir and Kemal Polat. Deep learning applications for hyperspectral imaging: a systematic review. *Journal of the Institute of Electronics and Computer*, 2(1):39–56, 2020. 1

[33] Tessa Phillips and Waleed Abdulla. Variational autoencoders for generating hyperspectral imaging honey adulteration data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 214–221, 2022. 2

[34] Konpat Preechakul, Nattanat Chatthee, Suttisak Wizadwongsa, and Supasorn Suwajanakorn. Diffusion autoencoders: Toward a meaningful and decodable representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10619–10629, June 2022. 8

[35] Jianwei Qin, Kuanglin Chao, Moon S Kim, Renfu Lu, and Thomas F Burks. Hyperspectral and multispectral imaging for evaluating food safety and quality. *Journal of Food Engineering*, 118(2):157–171, 2013. 1

[36] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015. 4, 5

[37] Salah Rifai, Grégoire Mesnil, Pascal Vincent, Xavier Muller, Yoshua Bengio, Yann Dauphin, and Xavier Glorot. Higher order contractive auto-encoder. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2011, Athens, Greece, September 5-9, 2011, Proceedings, Part II 22*, pages 645–660. Springer, 2011. 2

[38] Axel Sauer, Kashyap Chitta, Jens Müller, and Andreas Geiger. Projected gans converge faster. *Advances in Neural Information Processing Systems*, 34:17480–17492, 2021. 2

[39] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019. 1

[40] Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Opt.*, 47(10):B44–B51, Apr 2008. 4

[41] Lizhi Wang, Chen Sun, Ying Fu, Min H Kim, and Hua Huang. Hyperspectral image reconstruction using a deep spatial-spectral prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8032–8041, 2019. 6

[42] Xianghai Wang, Xinying Wang, Keyun Zhao, Xiaoyang Zhao, and Chuanming Song. Fsl-unet: Full-scale linked unet with spatial–spectral joint perceptual attention for hyperspectral and multispectral image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2022. 6

[43] Q. Xie, M. Zhou, Q. Zhao, D. Meng, W. Zuo, and Z. Xu. Multispectral and hyperspectral image fusion by ms/hs fusion net. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1585–1594, 2019. 1

[44] Yiqi Yan, Lei Zhang, Jun Li, Wei Wei, and Yanning Zhang. Accurate spectral super-resolution from single rgb image using multi-scale cnn. In *Pattern Recognition and Computer Vision: First Chinese Conference, PRCV 2018, Guangzhou, China, November 23-26, 2018, Proceedings, Part II 1*, pages 206–217. Springer, 2018. 1, 4, 6

[45] Jiarui Yi and Miguel Velez-Reyes. Dimensionality reduction using superpixel segmentation for hyperspectral unmixing using the cnmf. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIII*, volume 10198, pages 444–451. SPIE, 2017. 2

[46] Xin Yuan, David J Brady, and Aggelos K Katsaggelos. Snapshot compressive imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 38(2):65–88, 2021. 4

[47] Lei Zhang, Yanning Zhang, Hangqi Yan, Yifan Gao, and Wei Wei. Salient object detection in hyperspectral imagery using multi-scale spectral-spatial gradient. *Neurocomputing*, 291:215 – 225, 2018. 1