This CVPR Workshop paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Generative adversarial networks for spectral super-resolution and bidirectional RGB-to-multispectral mapping

Kin Gwn Lore, Kishore K. Reddy, Michael Giering
United Technologies Research Center
411 Silver Lane, E. Hartford, CT.
{lorek,reddykk,gierinmj}@utrc.utc.com

Edgar A. Bernal
University of Rochester
260 E. Main St., Rochester NY 14604
edgar.bernal@rochester.edu

Abstract

Acquisition of multi- and hyperspectral imagery imposes significant requirements on the hardware capabilities of the sensors involved. In order to keep costs manageable, and due to limitations in the sensing technology, tradeoffs between the spectral and the spatial resolution of hyperspectral images are usually made. Such tradeoffs are usually not necessary when considering acquisition of traditional RGB imagery. We investigate the use of statistical learning, and in particular, of conditional generative adversarial networks (cGANs) to estimate mappings from three-channel RGB to 31-band multispectral imagery. We demonstrate the application of the proposed approach to (i) RGB-to-multispectral image mapping, (ii) spectral super-resolution of image data, and (iii) recovery of RGB imagery from multispectral data.

1. Introduction

Hyperspectral image acquisition involves the use of specialized sensors to collect information across the electromagnetic spectrum. Hyperspectral imaging is widely used for various applications ranging from precision agriculture and biotechnology [15], environmental monitoring and remote sensing [8, 21], to manufacturing of pharmaceutical products and production of thin-films [24, 9]. Hyper- and multispectral image analysis techniques have demonstrated significant success in identifying signatures of interest to aid vision-based decision-making processes.

However, due to the complexity and cost of hyperspectral sensors, tradeoffs between spectral and spatial resolution capabilities of the sensors are usually made. Acquisition of hyperspectral data, in addition to involving long reconstruction time, also requires significant storage space resources and high-speed data transmission infrastructures which can make the technology prohibitively costly or infeasible in certain applications. The full potential of hyperspectral imaging has yet to be realized due to the aforementioned technical challenges.

Significant effort has been devoted to addressing these challenges, which are not as prevalent in acquistion of traditional RGB imagery. In this paper, we focus on leveraging inexpensive and readily available RGB images to estimate the corresponding hyperspectral imagery of a scene via statistically learned mappings. Some authors have applied machine learning models to learn direct mappings from RGB images to hyperspectral data based on statistical learning approaches; related work includes [17, 2, 12] where the authors proposed to fuse low-dimensional hyperspectral images with high-resolution RGB images to reconstruct high-resolution hyperspectral images.

In this paper, we explore the use of conditional generative adversarial networks (cGAN) to learn a mapping from three-channel RGB images to 31-band multispectral images. cGANs are a popular architecture within the broader category of deep learning, which encompasses a set of machine learning algorithms that are able to learn hierarchies of abstract representation of data. Currently, cGANs are capable of producing state-of-the-art results in a range of conditional image generation tasks such as image spatial super-resolution, edge mapping and automatic grayscale colorization. While RGB data is virtually inexhaustible due to the ubiquity and low cost of the sensors involved, hyperspectral data is relatively limited. We propose to use cGANs to tackle an image translation task. Specifically, we demonstrate the application of cGANs for (i) RGB-to-multispectral image mapping, (ii) spectral super-resolution of image data, and (iii) recovery of RGB imagery from multispectral data. A high-level overview of the first two tasks is provided in Fig. 1.

The paper is organized as follows. In Sec. 2, related work on mapping from RGB to multispectral imagery is discussed. Specific details on the proposed approach, in-

cluding the the description of the algorithms, the problem formulation, and the data preparation procedure are presented in Sec. 3. Evaluation metrics are outlined in Sec. 4 along with discussions on the experimental results. Lastly, we conclude the paper in Sec. 5.

2. Related work

As mentioned, acquisition, processing and storage of hyperspectral imagery imposes significant demands on the involved hardware. A common workaround to address these limitations is to make a trade off between spectral and spatial resolution, inevitably incurring a loss in the amount of available information. Methods to infer missing spectral and spatial information from low-resolution imagery are being actively researched.

A number of papers in the literature address the task of infering spectral information from low-fidelity imagery. Spectral super-resolution can be thought of as a spectral upsampling problem. Consider the task of automatic colorization of grayscale images [20, 30, 6, 13] where a single-channel grayscale image is used to infer the red, green, and blue components of the image. In a more general setting, different color spaces or spectral bands can be reconstructed from a given set of available data.

Early methods of spectral super-resolution investigated and formulated upsampling functions to convert low spectral-resolution images to high spectral-resolution spectral images by leveraging image statistics [5, 1]. This trend gradually shifted in favor of dictionary learning of image patches leveraging sparsity priors [26, 27, 29]. Some authors have introduced model-based methods that rely on characterization of the spectral response of an RGB camera to estimate muti- and hyperspectral information [2, 3, 17].

Leveraging ever-increasing hardware capabilities, more recent research has exploited high-performance GPU computing to implement statistical learning frameworks that perform spectral super-resolution. Galliani *et al.* [10] employed deep CNNs to map between RGB and multispectral images. More recently, generative adversarial networks (GANs), a deep architecture largely based on CNNs, have been shown to generate high quality, high spatial resolution images from low spatial resolution counterparts [22, 18, 23, 16]. While GANs are noted for their success in *spatial* super-resolution task, there has not been much focus on *spectral* super-resolution.

With regards to RGB estimation from multi- and hyperspectral data, previous work has leveraged visualization techniques of imagery with high spectral resolution on RGB displays such as minimum estimated abundance covariance (MEAC) [25]. Other approaches have attempted to preserve the maximal amount of information within the reduced spectral space but ultimately the image had to be displayed in false color. To overcome the limitations, manifold learning and alignment has been proposed [19]. In this work, we demonstrate a direct mapping from multispectral imagery to RGB images using conditional generative adversarial networks.

3. Methods

This section outlines the proposed data-driven approach for estimation of missing spectral data; training methods and data preprocessing steps are described in detail.

3.1. Conditional GANs

Tne introduction of Generative Adversarial Networks (GAN) [11] represented a significant breakthrough in the field of unsupervised learning. A GAN consists of two modules, a generator and a discriminator, which are usually implemented in the form of neural networks. The generative module captures the distribution of the data, while the discriminative module performs probability estimation on whether a sample to which it is exposed comes from the training data or is synthetically generated. Specifically, generator $G(z;\theta_g)$ builds a mapping function from a noise distribution $p_z(z)$ to the data space, while discriminator $D(x;\theta_d)$ produces a single scalar output representing the probability that observed sample x comes from the actual training data distribution rather than from the learned p_g.

G and D are trained simultaneously. The parameters for G are learned through gradient descent and error backpropagation to minimize loss function $\log(1-D(G(z)))$. At the same time, the parameters for D are learned by minimizing $\log D(X)$. Optimizing both networks can be posed as a two-player minimax game with value function $V(G,D)$:

$$\begin{aligned}&\min_G \max_D V(D,G)\\&= E_{x\sim p_{data}(x)}\left[\log D(x)\right] + E_{z\sim p_{x(z)}}\left[1-D(G(z))\right]\end{aligned} \tag{1}$$

The GAN framework can be naturally extended to learn conditional distributions by conditioning the generator and discriminator on additional information y, rather than on the noise vector z; the networks that result by following this formulation are known as *conditional* generative adversarial networks (cGANs). Conditioning input y can be any auxiliary information such as class labels, actual images, or any data from other modalities. In cGANs, the prior input noise $p_z(z)$ is combined with y to form joint hidden layer representations, which results in a generative model that is capable of transforming samples from one domain into another domain. Applications of such domain transformations include image-to-image translation [16].

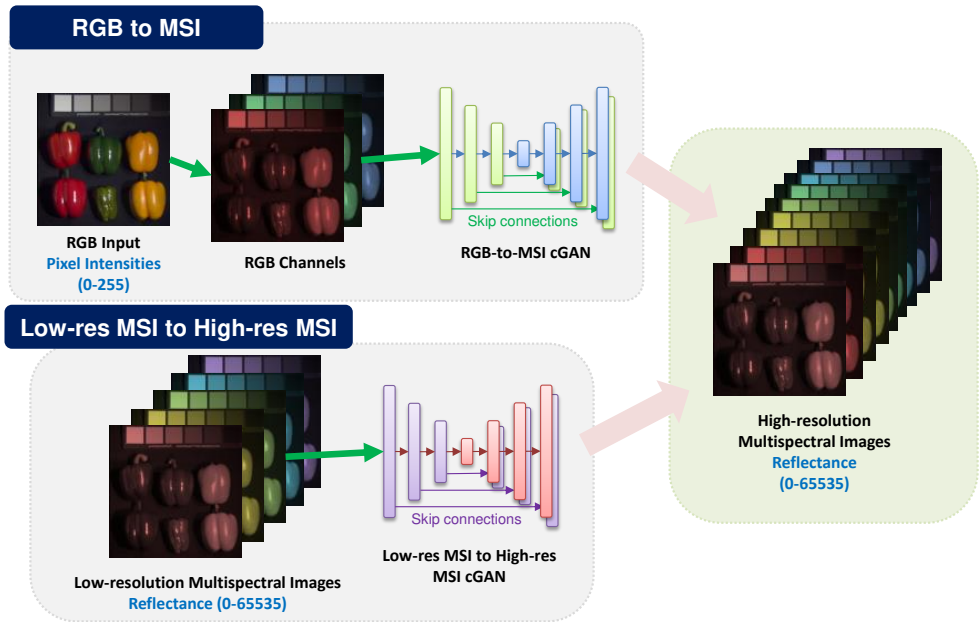

Figure 1. An illustration of the problem formulation. The first approach involves learning a mapping directly from the RGB to the high *spectral*-resolution image domain. The second approach learns a mapping between the downsampled multispectral images (MSI) and the full multispectral images. The former operates on pixel intensities while the latter operates on spectral reflectance values.

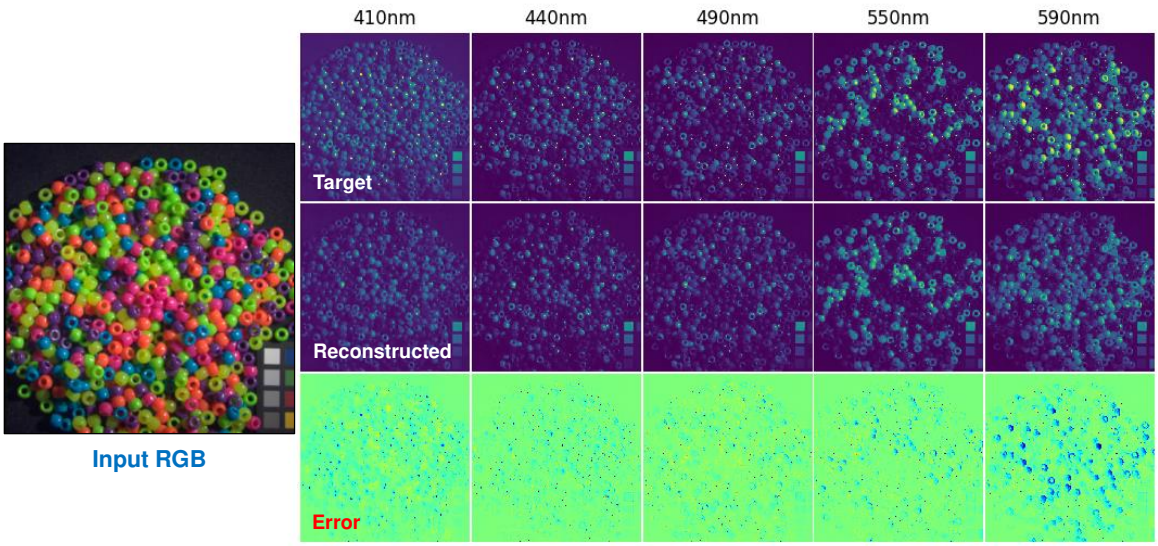

Figure 2. Reconstruction of five selected spectral bands using an RGB input. The Error maps are pseudo colored with the `jet` colormap, with red, green and blue indicating negative, zero, and positive errors respectively.

The objective function of the conditional two-player minimax game is:

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x|y)}\left[\log D(x)\right] + E_{z\sim p_x(z)}\left[1 - D(G(z|y))\right] \quad (2)$$

3.2. Problem formulation

In this paper, we propose a framework to perform learning of a spectral mapping function from an input image with C_i channels to an output image with C_o channels using cGANs. Our framework supports an arbitrary number of input and output channels, in practice limited only by the

available computational resources.

Spectral super-resolution from RGB data. This task involves mapping a three-channel ($C_i = 3$) RGB pixel image to a 31-band multispectral image ($C_o = 31$). RGB intensity values are the post-processed result from what has been captured by the camera sensor. On the other hand, multispectral image data usually corresponds to the spectral reflectance measured by the sensor. Therefore, the relationship between the two domains is expected to be highly non-linear.

Spectral super-resolution from downsampled bands. In this task, the 31 bands are downsampled to $C_i < C_o$ bands for training; we consider scenarios where $C_i \in \{6, 7, 11, 16, 31\}$. We keep the first and last channels fixed across the different scenarios, while the intermediate channels are uniformly sampled in order to yield the desired number of bands.

Inverse mapping from multispectral to RGB images. Traditionally, rendering RGB images from multispectral images is not a trivial task. Oftentimes, such operations have high complexity and may involve complicated non-linear transformations that leverage domain knowledge. In this task, a function that maps multispectral data to RGB imagery is learned.

3.3. Data description

In this experiment, we used the CAVE dataset [28], which consists of 32 scenes divided into 5 sections. Imagery for each scene includes full spectral resolution reflectance data from 400nm to 700nm at 10nm steps resulting in 31 total bands. Each band is stored as a 16-bit grayscale PNG image. For training and testing, we preserve the order of the filenames listed in the official website and split the data into training and testing sets in alternating order. In other words, channels 1, 3, 5, ..., 31 are used for training whereas images 2, 4, 6, ..., 32 are used for testing.

Data preprocessing. The available reflectance data has a depth of 16 bits (i.e., ranging from 0 to 65535) while the RGB pixel intensities have a depth of 8 bits (i.e., ranging between 0 and 255). We find that by linearly stretching the RGB pixel intensities to the full 16-bit range, the learning process converged faster. In other words, the conversion from 8-bit to 16-bit images was achieved by computing:

$$I_{\text{16-bit}} = 256(I_{\text{8-bit}} + 1) - 1$$

Due to the relatively small number of training images, we performed on-the-fly data augmentation where the training sample was scaled randomly by a factor of $\pm 5\%$. Nearest-neighbor interpolation was used as needed during the resizing process.

3.4. Training the model

In our experiments, we observed that when C_i is much smaller than C_o (e.g., when attempting to upsample 3 channels into 31 channels), training took longer to converge. Hence, the number of training epochs was selected as a function of input channel size; specifically, we set the number of epochs to $E_1 - E_0 C_i$, where $E_1 = 5700, E_0 = 300$ are constant values chosen. The model was trained using the ADAM optimizer with a learning rate of 0.0002 and momentum of 0.5.

4. Results and Discussion

We first provide a brief description of the choices of evaluation metrics, and then report the performance of the model based on these metrics.

4.1. Evaluation metrics

We employed multiple evaluation metrics to evaluate the quality of the reconstructed data. All metrics are computed by treating 8-bit integers as a floating point value.

Performance metrics. We measure the *root mean-squared error* (RMSE) between the reconstructed spectral map $\hat{y}$ and the ground truth y according to:

$$\text{RMSE}(y, \hat{y}) = \left(\frac{1}{whc} \sum_{i=0}^{whc} (\hat{y}_i - y_i)^2 \right)^{1/2}$$

We also employed the relative root mean squared error (rRMSE) as outlined in [10]. To that end, the RMSE is divided by the mean value of the ground truth according to:

$$\text{rRMSE}(y, \hat{y}) = \frac{\text{RMSE}(y, \hat{y})}{whc} \sum_{i=0}^{whc} y_i$$

Some authors [10] choose to clip values beyond the 8-bit 0-255 range. However, they fail to mention how they implement the 16-bit to 8-bit conversion. In this paper, the conversion takes place by computing:

$$I_{\text{8-bit}} = \left[\frac{1}{256} (I_{\text{16-bit}} + 1) - 1 \right]$$

where $[\cdot]$ is the nearest-integer function.

4.2. Model performance

Overall performance. Fig. 3 includes RMSE computed between the ground truth imagery and the reconstructed imagery for different reconstruction tasks. Generally speaking, our framework was able to reconstruct pixel values in missing spectral bands within approximately 3% of their

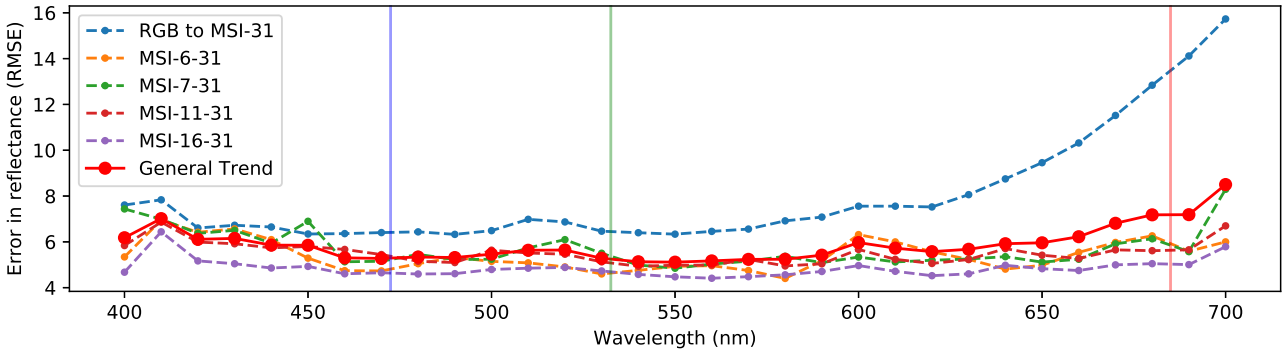

Figure 3. Reconstruction RMSE with respect to the ground truth. Labels are formatted as MSI-C_i-C_o, where C_i and C_o are the number of input and output channels, respectively. The vertical colored lines indicate the center of the red, green, and blue channels.

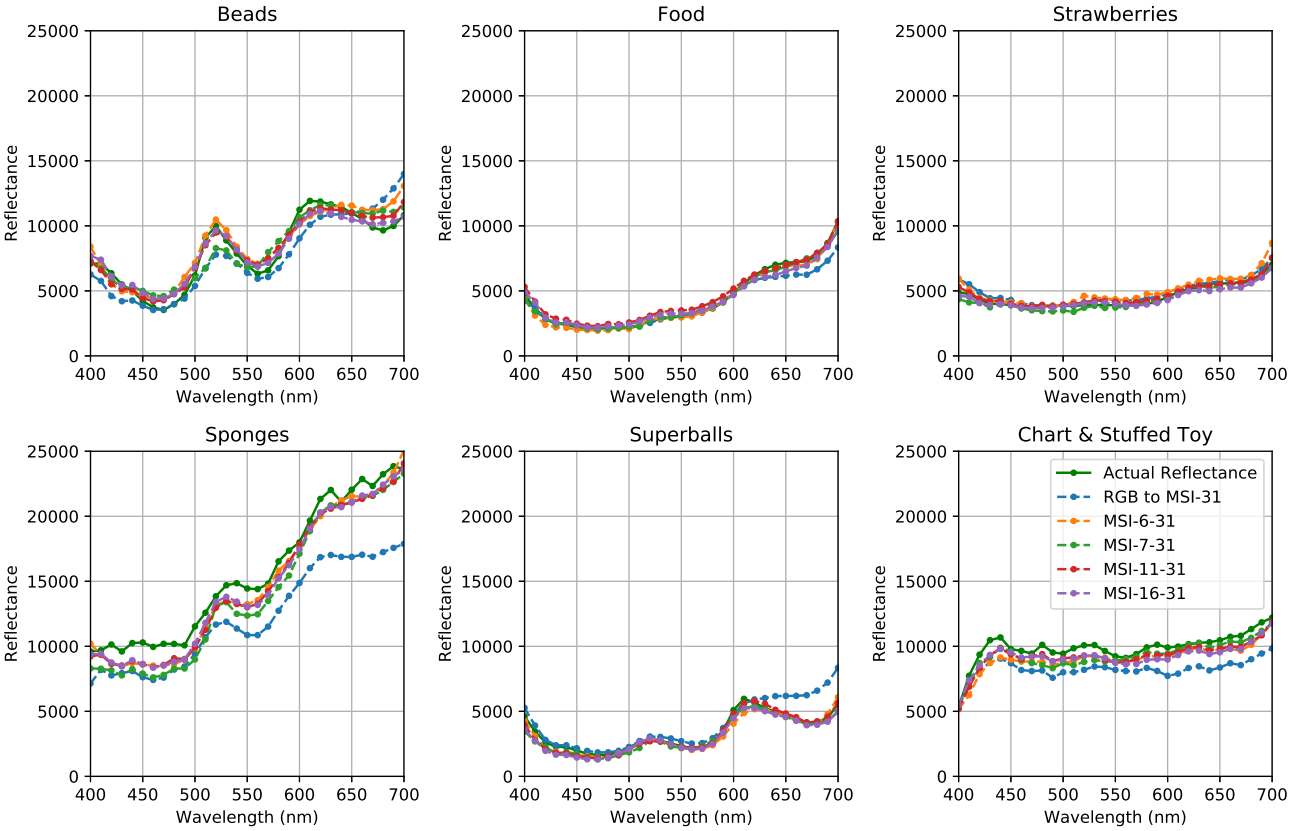

Figure 4. Spectral signatures of six examples from the test set computed by taking the mean of reflectance values for each channel. RGB to MSI approach (cross domain) struggled more compared to direct spectral super-resolution (same domain).

original value. As illustrated in the plots, the reconstruction error has a tendency to increase as the wavelength of the reconstructed band increases. This trend is observed across all cases investigated. Inspection of the data set revealed that pixel values tend to have larger variance in bands with longer wavelengths, which is likely an effect of noise. Furthermore, the effect of the increasing reconstruction error is more pronounced in RGB to MSI reconstruction, and is less obvious in MSI to MSI spectral super-resolution. We hypothesize that this is due to the fact that RGB pixel intensities have different physical meaning than the target values which effectively represent spectral reflectance. In contrast, input-output values for the spectral super-resolution approach lie within the same domain and there is less dis-

Figure 5. Visual evaluation of the MSI-to-RGB approach.

parity between the physical meaning of the values. Consequently, the reconstruction performance is better in the latter case.

Table 1. Resulting RMSE and rRMSE for the three-channel RGB to 31-band Multispectral reconstruction task.

CAVE RGB - MSI 31	Arad [4]	Galliani [10]	Ours
RMSE	5.40	**4.76**	8.06
rRMSE	-	**0.2804**	0.3210

RGB to multispectral reconstruction. The reconstruction of spectral bands using RGB inputs is visualized in Fig. 2. The proposed approach is compared with the algorithms introduced by Arad et al. [4] and Galliani et al. [10] in terms of RMSE and rRMSE (see Table 1). Our method did not fare too well with the RGB-to-Multispectral super-resolution formulation and is largely attributed to the challenge of GANs "hallucinating" the spectral bands instead of learning a good mapping function between the two domains, as mentioned in prior work [7, 16, 14]. In addition, [10] performed 8-bit clipping before computation of the metric; while we performed computation on our 16-bit model output (0-65536) rescaled to 8-bit (0-255) without clipping.

For visualization purposes, spectral signatures of six examples from the test set are shown in Fig. 4. For *Food*, *Strawberries*, and *Beads*, the reconstructed reflectance (blue dotted lines) accurately reflects the ground truth reflectance (green solid lines). As before, reconstruction errors are relatively small in shorter wavelengths of *Superballs* and tend to increase at higher wavelengths. An almost constant bias of mapped values is observed in *Sponges* and *Chart & Stuffed Toy*, potentially due to the scene having visual clutter in the foreground, in addition to the 2D chart in the background.

Spectral super-resolution. We compare our approach with the spectral interpolation (SpctInt) approach where missing bands are filled with interpolated values between the two neighboring available channels. Formally, the matrix for the i-th channel between two channels D_{start} and D_{end} is described by:

$$D_i = D_{start} + \frac{(D_{end} - D_{start})(i-1)}{c+1}$$

where c is the number of missing channels in between. For example, $c = 5$ when a 6-channel input is upsampled to 31 channels (i.e., $31 = 1+5+1+5+1+5+1+5+1+5+1$ where 1s are channels that are retained, and 5s in between indicate the number of missing channels.) Results for spectral super-resolution (i.e., from low spectral resolution to high spectral resolution) are shown in Table 2. As hypothesized, the general observation is that as the number of discarded bands, the performance of the model improves. It is also observed that the proposed approach has lower RMSE compared to naive spectral interpolation.

Table 2. Resulting RMSE and rRMSE for competing spectral super-resolution formulations. Labels are formatted as MSI-C_i-C_o, where C_i and C_o are the number of input and output channels respectively.

CAVE Dataset	SpctInt RMSE	SpctInt rRMSE	Ours RMSE	Ours rRMSE
RGB to MSI-31	-	-	8.0622	0.2062
MSI-6-31	7.2942	0.4236	**5.9907**	**0.1664**
MSI-7-31	5.9417	0.3514	**5.3824**	**0.1420**
MSI-11-31	4.8440	0.2881	**4.7001**	**0.1293**
MSI-16-31	5.0178	0.3074	**4.8740**	**0.1326**

Multispectral to RGB. Quantitative results for this task are shown in Fig. 3. The outputs are visualized in Fig. 5. From inspection of the examples provided, we conclude the model can accurately reconstruct RGB color values. However, the model appears to have introduced some blur to the generated outputs. Since the framework accepts the full dimension of the image as the input, it relies on the global image statistics while reconstructing the output. With this in mind, the framework could be extended to learn finer local features to produce crisper and sharper images. This can be achieved by training a local model that processes overlapping subimages from the original input and averaged at the output.

Table 3. Resulting RMSE and rRMSE for the MSI to RGB mapping task.

CAVE Dataset	RMSE	rRMSE
MSI 31 - RGB	5.6490	0.1389

5. Conclusion

We proposed a framework to super-resolve images in the spectral domain using conditional generative adversarial networks. We demonstrated via experimental validation that such framework is not only able to learn the mapping between two domains having distinct physical meaning, but that it can also learn to effectively map between low- and high-dimensional features in the same domain. We compared our forward-mapping approach with competing state-of-art methods, in addition to proposing the same framework to solve the *inverse mapping* problem which traditionally has been solved using high complexity band-selection methods. Our framework was able to learn global image statistics from a limited number of training data and achieve good performance. We surmise that as mapping methods such as the ones proposed increase in capabilities, many applications where performance is affected by the expense of hyperspectral sensing will benefit. The downside of relying on purely data-driven generative model such as GANs is their susceptibilty to hallucination as shown in our RGB-to-Spectral translation and also in prior work. Future research directions are delineated below:

1. Extend the framework to operate on a patch basis so that high-frequency local image structures can be better reconstructed and appear sharper.

2. Investigate multi-scale framework to enhance performance by leveraging statistics extracted from hierarchical multi-resolution image structures.

3. Devise methods to reduce susceptibility to hallucinations.

References

[1] T. Akgun, Y. Altunbasak, and R. M. Mersereau. Super-resolution reconstruction of hyperspectral images. *IEEE Transactions on Image Processing*, 14(11):1860–1875, 2005. 2

[2] N. Akhtar, F. Shafait, and A. S. Mian. Sparse spatio-spectral representation for hyperspectral image super-resolution. In *ECCV (7)*, pages 63–78, 2014. 1, 2

[3] N. Akhtar, F. Shafait, and A. S. Mian. Hierarchical beta process with gaussian process prior for hyperspectral image super resolution. In *ECCV (3)*, pages 103–120, 2016. 2

[4] B. Arad and O. Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In *European Conference on Computer Vision*, pages 19–34. Springer International Publishing, 2016. 6

[5] A. S. Charles and C. J. Rozell. Spectral superresolution of hyperspectral imagery using reweighted spatial filtering. *IEEE Geoscience and Remote Sensing Letters*, 11(3):602–606, 2014. 2

[6] G. Charpiat, M. Hofmann, and B. Schölkopf. Automatic image colorization via multimodal predictions. *Computer Vision–ECCV 2008*, pages 126–139, 2008. 2

[7] J. P. Cohen, M. Luck, and S. Honari. Distribution matching losses can hallucinate features in medical image translation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 529–536. Springer, 2018. 6

[8] M. T. Eismann. Hyperspectral remote sensing. SPIE Bellingham, 2012. 1

[9] W. Fountain, K. Dumstorf, A. E. Lowell, R. A. Lodder, and R. J. Mumper. Near-infrared spectroscopy for the determination of testosterone in thin-film composites. *Journal of pharmaceutical and biomedical analysis*, 33(2):181–189, 2003. 1

[10] S. Galliani, C. Lanaras, D. Marmanis, E. Baltsavias, and K. Schindler. Learned spectral super-resolution. *arXiv preprint arXiv:1703.09470*, 2017. 2, 4, 6

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014. 2

[12] H. Grahn and P. Geladi. *Techniques and applications of hyperspectral image analysis*. John Wiley & Sons, 2007. 1

[13] R. K. Gupta, A. Y.-S. Chia, D. Rajan, E. S. Ng, and H. Zhiyong. Image colorization using similar images. In *Proceedings of the 20th ACM international conference on Multimedia*, pages 369–378. ACM, 2012. 2

[14] K. Gwn Lore, K. Reddy, M. Giering, and E. A. Bernal. Generative adversarial networks for depth map estimation from RGB video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1177–1185, 2018. 6

[15] W. Huang, D. W. Lamb, Z. Niu, Y. Zhang, L. Liu, and J. Wang. Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precision Agriculture*, 8(4):187–197, 2007. 1

[16] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:1611.07004*, 2016. 2, 6

[17] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi. High-resolution hyperspectral imaging via matrix factorization. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2329–2336. IEEE, 2011. 1, 2

[18] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. *arXiv preprint arXiv:1609.04802*, 2016. 2

[19] D. Liao, Y. Qian, J. Zhou, and Y. Y. Tang. A manifold alignment approach for hyperspectral image visualization with natural color. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6):3151–3162, 2016. 2

[20] Q. Luan, F. Wen, D. Cohen-Or, L. Liang, Y.-Q. Xu, and H.-Y. Shum. Natural image colorization. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pages 309–320. Eurographics Association, 2007. 2

[21] F. Melgani and L. Bruzzone. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on geoscience and remote sensing*, 42(8):1778–1790, 2004. 1

[22] M. Mirza and S. Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014. 2

[23] A. Odena, C. Olah, and J. Shlens. Conditional image synthesis with auxiliary classifier gans. *arXiv preprint arXiv:1610.09585*, 2016. 2

[24] Y. Roggo, A. Edmond, P. Chalus, and M. Ulmschneider. Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms. *Analytica Chimica Acta*, 535(1):79–87, 2005. 1

[25] H. Su, Q. Du, and P. Du. Hyperspectral image visualization using band selection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):2647–2658, 2014. 2

[26] R. Timofte, V. De Smet, and L. Van Gool. Anchored neighborhood regression for fast example-based super-resolution. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1920–1927, 2013. 2

[27] J. Yang, J. Wright, T. S. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE transactions on image processing*, 19(11):2861–2873, 2010. 2

[28] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE transactions on image processing*, 19(9):2241–2253, 2010. 4

[29] R. Zeyde, M. Elad, and M. Protter. On single image scale-up using sparse-representations. In *International conference on curves and surfaces*, pages 711–730. Springer, 2010. 2

[30] R. Zhang, P. Isola, and A. A. Efros. Colorful image colorization. *arXiv preprint arXiv:1603.08511*, 2016. 2