

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,000

Open access books available

148,000

International authors and editors

185M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Unsupervised Deep Hyperspectral Image Super-Resolution

Zhe Liu and Xian-Hua Han

Abstract

This chapter presents the recent advanced deep unsupervised hyperspectral (HS) image super-resolution framework for automatically generating a high-resolution (HR) HS image from its low-resolution (LR) HS and high-resolution RGB observations without any external sample. We incorporate the deep learned priors of the underlying structure in the latent HR-HS image with the mathematical model for formulating the degradation procedures of the observed LR-HS and HR-RGB observations and introduce an unsupervised end-to-end deep prior learning network for robust HR-HS image recovery. Experiments on two benchmark datasets validated that the proposed method manifest very impressive performance, and is even better than most state-of-the-art supervised learning approaches.

Keywords: deep learning, unsupervised learning, hyperspectral image, super-resolution, generative network

1. Introduction

Hyperspectral images (HSI) feature hundreds of bands with extensive spectral qualities that are helpful for a range of visual tasks, such as computer vision [1], mineral exploration [2], medical diagnosis [3], remote sensing [4], and so on. Due to technology restrictions, it is harder to capture high-quality HSI, and the acquired HSI has substantially lower resolution. As a result, super-resolution (SR) has been applied to obtain a HR-HSI, but it is a challenge because of texture blurring and spectral distortion problems at high magnifications. Thus, researchers frequently combine high-resolution PAN and low-resolution HSI [5] to achieve SR tasks. In recent years, it is a trend to fuse a high-resolution multispectral/RGB (HR-MS/RGB) image and a low-resolution hyperspectral (LR-HS) image for generating a high-resolution hyperspectral (HR-HS) image, which is called hyperspectral image super-resolution (HSI-SR). The HSI-SR methods are classified into two primary categories based on reconstruction principles: conventional mathematical model-based methods and deep learning-based approaches in a supervised/unsupervised manner. The following sections go into further information about each of these categories.

1.1 Mathematical model-based methods

Since HSI-SR is typically an inverse problem, a mathematical model-based approach yields a solution space that is far bigger than the actual result needed. In order to tackle this issue, mathematical model-based HSI-SR constrains the solution space using hand-crafted prior knowledge, regularizes the mathematical model, and then optimizes the model by minimizing the reconstruction errors. This method aims at establishing a mathematical formulation that simulates the transformation of HR-HS images into LR-HS and HR-RGB images. This process is extremely difficult, and direct optimization of the formed mathematical model might result in very unreliable solutions, as the known variables in the LR-HS/HR-RGB images under consideration are significantly less than the unknown variables to be estimated in the latent HR-HS images. In order to narrow the set of possible solutions, existing approaches often utilize a variety of priors to modify the mathematical model.

Based on prior knowledge of various structures, three categories of mathematical model-based HSI-SR methods can currently be distinguished: spectral unmixing-based methods [6], sparse representation-based methods [7], and tensor factorization-based methods [8]. For spectrum unmixing-based methods, Yokoya et al. [9] proposed a coupled non-negative matrix decomposition approach (CNMF), which alternatively unmixes LR-HS images and HR-RGB images to estimate HR-HS images. Later, Lanaras et al. [6] proposed a similar framework to jointly unmix the two input images by decoupling the initial optimization problem into two constrained least square problems. Dong et al. [7] incorporated alternating multiplication method (ADMM) techniques for solving the spectra unmixing model. Additionally, the sparse representation is frequently used as an alternative mathematical model for HSI-SR. In this model, the underlying HR-HS image is recovered by first learning the spectral dictionary from the LR-HS image under consideration, and then calculating the sparse coefficient of the HR-RGB image. Inspired by the existed spectral similarity of the neighboring pixels in the latent HS image, Akhtar et al. [10] proposed to perform group sparse and non-negativity representation within a small patch, while Kawakami et al. [11] applied a sparse regularizer for the decomposition of spectral dictionaries. Moreover, the tensor factorization-based method demonstrated that it could be used to resolve the HSI-SR problem. He et al. [8] factorized the HR-HS image into two low-rankness constraint matrices and achieved great super-resolution performances, which were motivated by the intrinsic low dimensionality of the spectrum space and the three-dimensional structure of the HR-HS image.

Despite some advancements in handcrafted prior, HSI-SR performance tends to be inconsistent and can cause severe spectral distortion due to the under-representation of handcrafted prior, depending on the content of the image under investigation.

1.2 Deep learning-based methods

Hyperspectral super-resolution is a hot field of research in hyperspectral imaging, as it can improve low-resolution images in both the spatial and spectral domains, turning them into high-resolution hyperspectral images. HSI-SR is a classic inverse problem, and deep learning has a lot of promise for resolving it. Depending on whether a training dataset is provided, supervised and unsupervised learning are the two approaches used in deep learning-based HSI-SR. A labeled training dataset is necessary for supervised learning in order to create a function or model from which

subsequent data is fed in order to generate accurate predictions. But a labeled training dataset is not necessary for unsupervised learning.

1.2.1 Deep supervised learning-based methods

Different vision tasks have been successfully resolved by DCNNs. As a result, DCNN-based methods have been suggested for HSI-SR tasks, which eliminate the requirement to investigate various manually handcrafted priors. With the LR-HS observation only, Li et al. [12] presented an HSI-SR model by combining a spatial constraint (SCT) strategy with a deep spectral difference convolutional neural network (SDCNN). Han et al. [13] utilized three straightforward convolutional layers in the groundbreaking HS/RGB fusion work, whereas later work utilized more advanced CNN architectures, such as ResNet [14] and DenseNet [15], in an effort to attain more robust learning capabilities. By resolving the Sylvester equation using a fusion framework, Dian et al. [16] first provided an optimization technique, and then they investigated a DCNN-based strategy to enhance the initialization results. Further, Han et al. [17] proposed a multi-layer, multi-level spatial, and spectral fusion network that successfully fused existing LR-HS and HR-RGB images. In order to investigate an MS/HS fusion network and optimize the suggested MS/HS fusion system, Xie et al. [18] employed a low-resolution imaging model and spectral low-level knowledge of HR-HS images. In order to solve HS image reconstruction difficulties effectively and accurately, Zhu et al. [19] researched the progressive zero-centric residual network (-PZRes-Net), a lightweight deep neural network-based system. All the DCNN-based methods mentioned above take training with a large number of pre-prepared training instances that contain not only LR-HS and HR-RGB images but also the corresponding HR-HS images as labels, that is, the set of training triples, despite the fact that the reconstruction performance was significantly improved.

1.2.2 Deep unsupervised learning-based methods

Although HS images are difficult to obtain in the real world, deep learning networks for HSI-SR require a lot of hyperspectral images as training data. It is rather challenging to collect good quality HSIs due to hardware restrictions, and the resolution of the acquired HSIs is relatively low. For supervised learning, which needs big training datasets to succeed, this is an unsolvable problem. As a result, unsupervised learning is one of the key research areas. Unlike supervised learning, unsupervised learning does not require any HR-HS image as a ground-truth image and uses only easily accessible HR-MS/RGB images and LR-HS images to generate HR-HS images.

It is well known that the corresponding training triplets, especially the HR-HS images, are extremely hard to be collected in real applications. Thus, the quality and amount of the collected training triplets generally become the bottleneck of the DCNN-based methods. Most recently, Qu et al. [20] attempted to solve the HSI super-resolution problem in an unsupervised way and designed an encoder-decoder architecture for exploiting the approximate low-rank prior structure of the spectral model in the latent HR-HS image. This unsupervised framework did not require any training samples in an HSI dataset and could restore the HR-HS image using a CNN-based end-to-end network. However, this method needed to be carefully optimized step-by-step in an alternating way, and the HS image recovery performance was still not enough. Liu *et al.* [21] proposed an unsupervised multispectral and hyperspectral image fusion (UnMHF) network using the observations of the under-studying scene

only, which estimates the latent HR-HS image with the learned encoder-decoder-based generative network from a noise input and can only be adopted to the observed LR-HS and HR-RGB image with the known spatial downsampling operation and camera spectral function (CSF). Later, Uezato et al. [22] exploited a similar method for unsupervised image pair fusion, dubbed a guided deep decoder (GDD) network for the known spatial and spectral degradation operation only. Thus, the UnMHF [21] and GDD [22] can be categorized into the non-blind paradigm, and lack of generalization in a real scenario. Zhang et al. [23] proposed two steps of learning methods via modeling the common priors of the HR-HS image in a supervised way and then adapting to the under-studying scene for modeling its specific prior in an unsupervised manner. In addition, the unsupervised adaptation is capable of learning the spatial degradation operation of the observed LR-HS image but can only deal with the observed HR-HS image with known CSF, and thus it would be categorized as a semi-blind paradigm for possibly learning the spatial degradation operations only in the observed LR-HS image. Moreover, Fu et al. [24] exploited an unsupervised hyperspectral image super-resolution method using the designed loss function formulated by the observed LR-HS and HR-RGB images only and integrated a CSR optimization layer after the HSI super-resolution network to automatically select or learn the optimal CSR for adapting to the target RGB image possibly captured by various color cameras, which is also divided into the semi-blind paradigm for possibly learning the spectral degradation operation: CSF only. Further, the unsupervised adaptation subnet in ref. [23] and the method [24] utilize the under-studying observed images only instead of the requirement of additional training samples for guiding the network training, which achieved impressive performance as an unsupervised learning strategy. However, these learning methods based on the under-studying observed images only are easy to drop into a local solution, and the final prediction heavily depends on the initial input of the network. Our method is also formulated in this unsupervised learning paradigm, and we are going to clarify the distinctiveness of our method in the next sub-section.

2. The proposed unsupervised learning-based methods

In this section, we first describe the problem formulation in the HSI-SR task and then present the proposed deep unsupervised learning-based method.

2.1 Problem formulation

Let us consider image pairs: a LR-HS image $X \in R^{w \times h \times L}$, where w and h are the width and height, and a HR-RGB image $Y \in R^{W \times H \times 3}$, where w and h are the width and height of Y and Z , respectively. A HR-HS image $Z \in R^{W \times H \times L}$, where L is the number of spectral channels in the HR-HS image, is what we are trying to reconstruct for HSI-SR. The following formula can be used to represent the degradation between the HR-HS target image and the observed images: X and Y .

$$X = k^{(spa)} \otimes Z^{(spa)} \downarrow + n_x, Y = Z * C^{(Spec)} + n_y, \quad (1)$$

where \otimes stands for the convolution operator, $(Spa) \downarrow$ for the spatial domain downsampling operator, and $k^{(Spa)}$ for the two-dimensional blur kernel in the spatial

domain. Three one-dimensional spectral filters $C^{(Spec)}$ constitute the spectral sensitivity function of RGB cameras, which translates L spectral bands to RGB bands. The additive white Gaussian noise (AWGN) with noise level σ is represented by n_x and n_y . We rephrase the degenerate model as a matrix formulation to quantify the problem, that is,

$$X = DBZ + n_x, Y = ZC + n_y, \quad (2)$$

where B is the spatial blur matrix, D is the downsampling matrix, and C is the transformation matrix representing the spectral sensitivity function (CSF). According to Eqs. (1) and (2), a general HSI-SR task should evaluate $k^{(Spa)}$ (or B), $(Spa) \downarrow$ (or D), and $C^{(Spec)}$ (or C) from observed image pairs X and Y , which makes it very complicated to obtain the latent Z . It is a challenging problem that has rarely been studied in the HSI-SR task. Therefore, the general solution is to assume that the blur kernel type and spectral sensitivity function (CSF) of the RGB camera are known and to approximate them by some mathematical operations in the application. The current study followed to the previous setting in principle, but we also investigated whether it was possible to reconstruct HR-HS images without knowing the kind of CSF or the blur kernel beforehand as a generic solution for a specific scenario.

Let us begin by defining the generic formula of the HSI-SR task generally. The maximum a posterior (MAP) framework is the foundation formula of the majority of classical approaches.

$$Z^* = \underset{Z}{\operatorname{argmax}} \Pr(Z|X, Y, B, C) = \underset{Z}{\operatorname{argmax}} \Pr_{(B, C)}(X, Y|Z) \Pr(Z), \quad (3)$$

where $\Pr(Z)$ performs prior modeling of latent HR-HS images and $\Pr_{(B, C)}(X, Y|Z)$ is the likelihood of the fidelity term corresponding to the known kernel type and CSF matrix. With regard to the latent HR-HS image Z , which we define as $-\log(\Pr_{(B, C)}(X, Y|Z))$, it is specifically assumed that the reconstruction errors of the fidelity terms X and Y are independent Gaussian distribution in general. The prior modeling of HR-HS images is subjected to the regularization requirement $-\log(\Pr(Z)) = \phi(Z)$. The reconstruction model of the MAP-based HSI SR in Eq. (3) can be redefined using the following formula.

$$Z^* = \underset{Z}{\operatorname{argmin}} \alpha \beta_1 \|X - DBZ\|_F^2 + (1 - \alpha) \beta_2 \|Y - ZC\|_F^2 + \lambda \phi(Z), \quad (4)$$

where $\|\cdot\|_F$ represents the Frobenius norm. It is generally necessary to introduce normalization weights, such as $\beta_1 = 1/N_1$ and $\beta_2 = 1/N_2$, where N_1 and N_2 are multiples of the number of pixels and spectral bands in LR-HS and HR-HS images, respectively. This is because HR-RGB and LR-HS images have different numbers of elements. In addition, we further modify the contribution of these two reconstruction errors using the hyperparameter α ($0 \leq \alpha \leq 1$). On the other hand, the trade-off adjustment parameter is λ . We have experimentally developed appropriate prior parameters as regularization term $\phi(Z)$ in order to obtain a robust solution. Numerous prior restrictions have been present. The employed priors, however, are often manually determined and fall short of adequately describing the intricate structure of HR-HS images. Furthermore, the established priors should vary depending on the details of the situation being studied, and choosing the suitable priors for a specific scenario is still an art.

The DCNN method is one of the most recent deep learning-based HSI-SR techniques. It effectively captures prospective HS image features (common prior) in a fully supervised learning manner utilizing previously trained training samples (external datasets). Particularly supervised deep learning methods seek to learn joint CNN models by minimizing such loss functions given N trainable triples.

$$(X_i, Y_i, Z_i)(i = 1, 2, \dots, N).$$

$$\theta^* = \operatorname{argmin}_{\theta} \sum_i^N \|Z_i - F_{CNN}(X_i, Y_i)\|_F^2, \quad (5)$$

where F_{CNN} stands for a DCNN network transform with θ learning parameters. In contrast to directly searching in the ground-truth image space Z , these approaches are trained to extract the optimal parameters θ^* of the network, and they are able to identify common prior variables concealed in the training samples utilizing the powerful and effective DCNN modeling capabilities. The underlying HR-HS images for each observation (X_t, Y_t) in the research can be simply rebuilt as: $\hat{Z}_t = F_{CNN}^{\theta^*}(X_t, Y_t)$ after learning the network parameters θ^* . Although these supervised deep learning methods have shown encouraging results, it is necessary to provide a substantial training dataset that includes LR-HS, HR-RGB, and HR-HS images—all of which are particularly challenging to gather in HSI-SR tasks—in order to learn a good model.

2.2 The overview motivation

Recent deep learning-based HSI-SR techniques have demonstrated that DCNNs perform well and are capable of accurately capturing the underlying spatial and spectral structure (joint prior information) of potential HS images. The training labels (HR-HS images) for these algorithms, which are typically performed in a fully supervised way and need large-scale training datasets containing LR-HS, HR-RGB, and HR-HS images, are challenging to gather. Numerous studies on natural image generation (DCGAN [25]) and its variations have demonstrated that high-resolution, high-quality images with specific features and attributes can be produced from noisy random input data without the supervision of high-quality ground-truth data. This indicates that originating from a random initial image and scanning the parameter space of a neural network can capture the inherent structure (a prior) of possible images with certain features. DIPs [26] have also been utilized to properly perform a number of natural image restoration tasks, including image separation, blurring, and super-resolution extraction, using just the degraded version of a scene to guide them. This unsupervised paradigm is used in the current study, which tries to learn the precise spatial and spectral structure (a prior) of HR-HS latent images from degraded data (LR-HS and HR-RGB images).

The spatial and spectral structure of the underlying HR-HS image Z was specifically modeled using the generative neural network G_{θ} (θ is a network parameter that must be learnt). The fusion-based HSI-SR model can be rebuilt as follows by substituting Z with G_{θ} in Eq. (4) and deleting the regularization term $\phi(Z)$ connected with the prior acquired automatically by the generative network.

$$\theta^* = \operatorname{argmin}_{\theta} \alpha \beta_1 \|X - DBG_{\theta}(Z_{in})\|_F^2 + (1 - \alpha) \beta_2 \|Y - G_{\theta}(Z_{in})C\|_F^2, \quad (6)$$

where $G_\theta(Z_{in})_i$ is the i -th component of the HR-HS estimation and Z_{in} is the input to the generative neural network. Eq. (6) tries to explore the parameter space of the generative neural network G_θ by leveraging the powerful modeling capability to generate a more reliable HR-HS image, instead of directly searching the exceedingly vast, non-uniquely determined raw HR-HS space.

To solve the above unsupervised HSI-SR task, there are still several issues to be needed to elaborately address: (i) How to design the generative network's architecture so that both spectral correlations and low-level spatial statistics can be effectively modeled during training. (ii) What kind of input to the generative network should be employed so that the local minimization point can be avoided. (iii) How to implement an end-to-end learning framework for incorporating different degradation operations following the generative network. In the next sections, we embody the solutions to the aforementioned issues.

2.3 Architecture of the generative neural network

Generative neural networks G_θ can be implemented using arbitrary DCNN architectures. A generative neural network G_θ is required to offer acceptable modeling skills due to the diversity of information, including potentially significant structures, rich textures, and complicated spectra in HR-HS images. It has been demonstrated that various generative neural networks have a great deal of promise for producing high-quality natural images [Pix2pix and others], for example, in adversarial learning settings [27]. In this study, a multi-level feature learning architecture is employed, along with simplified encoder-decoder features and an encoder-decoder architecture that allows for feature reuse via skip connections between the encoder and the decoder. **Figure 1** shows a thorough representation of a generative neural network.

Five blocks compensate the encoder and decoder, and they both learn representative features at various scales. To reuse the extracted detailed features, the output of each of the 5 encoder-side blocks is straight-through forwarded to the corresponding decoder. A maximum clustering layer with a 2×2 kernel is used to reduce the size of the feature map between encoder blocks, and an upconversion layer is used to double

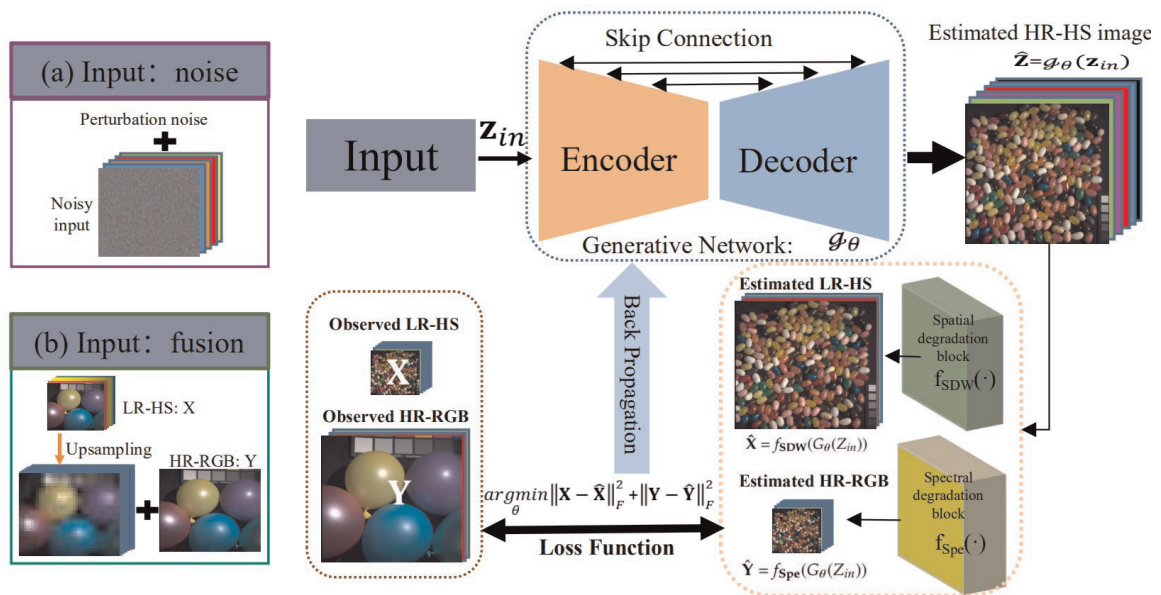


Figure 1. Conceptual diagram of the proposed unsupervised deep HSI-SR.

the size of the feature map between decoder blocks for recovery. Each block is comprised of three convolutional layers that each follow the RELU activation function. Finally, the HR-HS images are estimated using the convolutional output layer. The training state of the generative neural network cannot be estimated or guided in an unsupervised learning environment as there is no ground-truth HR-HS image. The assessment criteria listed in Eq. (6) are then generated using the observed HR-RGB and LR-HS images.

2.4 Input data to the generative neural network

We classify the input data into two types. The first is a noisy input with a random perturbation added to check the robustness, corresponding to the deep unsupervised fusion learning (DUFL) model; in particular, to contrast with the addition of random perturbation, we also perform experiments without random perturbation, that is, the DUFL+ model. The second input data is the fusion context of fused observations HR-RGB and LR-HS, which corresponds to the deep self-supervised HS image reconstruction (DSSH) framework.

2.4.1 The noise input

The deep image prior network (DIP) [26] was developed to get low spatial statistics using inputs of uniformly distributed noise vectors generated at random. Nevertheless, because the noise vectors are chosen at random, DIP has a limited ability to discover spectral and spatial correlations and is more challenging to tune. Motivated by the DIP, we proposed a deep unsupervised fusion learning (DUFL) model, in which a common generative neural network is trained to generate target images with predetermined features; typically, a randomly selected noise vector based on a distribution function (for example, Gaussian or uniform distribution) is used as input to ensure that the generated images have enough diversity and variability. The observed degradation (LR-HS and HR-RGB images) of the corresponding HR-HS images is required for our HSI-SR task. Therefore, it makes sense to determine the best network parameter space for searching a given HR-HS image as the previously sampled noise vector Z_{in}^0 . However, a constant noise input could lead to a local minimum in the generative neural network. As a result, the HR-HS image's estimate is inaccurate. Therefore, it is suggested to disturb the fixed initial input with a small randomly generated noise vector in each training step to avoid the local minimum condition. For a training step, the input vector i -th can be represented as follows:

$$Z_{in}^i + Z_{in}^0 + \Delta n_i, \quad (7)$$

where Δ stands for the interference level (small scale value) and n_i is the noise vector randomly sampled in the i th training. The final estimated HR-HS image utilized for prediction is the fixed noise vector $Z^* = G_\theta(Z_{in}^0)$, which is created by feeding perturbed inputs into a neural network with coefficient G_θ .

This deep unsupervised fusion learning model employs noise vectors produced at random and sampled from a uniform distribution as input to provide low-level spatial statistics. But this research is less effective at identifying spectral and spatial correlations and is more challenging to optimize due to random noise vectors. We propose a solution to this issue in the next section. In the next part, we substitute observed LR-

HS and HR-RGB images for entirely artificial noise. Additionally, we approximate the degradation operation using two distinctive convolutional layers that can be applied as learning or fixed degradation models for a variety of real-world scenarios.

2.4.2 The fusion context

To deal with the mentioned problems, we improved the DUFL model above. The underlying prior structure of HR-HS images is reflected by an internally designed network structure in the deep self-supervised HS image reconstruction (DSSH) framework, which also learns the network parameters exclusively using observed LR-HS and HR-RGB images. In the proposed DSSH framework, we use the observed fusion context in network learning to gain insight into specific spatial and spectral priorities given the observed images: X reflecting hyperspectral properties of the underlying HR-HS image although with lower spatial-resolution, and Y showing the high-resolution spatial structure although with fewer spectral channels. To be more specific, we utilize an upconversion layer to first transform the LR-HS image to the same spatial dimension as the HR-RGB image before merging them, as seen below.

$$Z_{\text{in}}^0 = \text{Stack}(UP(X), Y). \quad (8)$$

A simple fused context can be used as input, but this generally results in local minimum convergence. To train a more reliable model in this section that takes into account specific spatial and spectral priors, we add additional perturbations. The model is then represented as follows:

$$Z_{\text{in}}^i = Z_{\text{in}}^0 + \lambda\mu, \quad (9)$$

where λ is a small number indicating the intensity of the perturbation and μ is a sample of a 3D tensor generated at random from a uniform distribution equal to the connection context Z_{in}^0 . In this section, λ is set at 0.01 and reduced by half every 1000 steps throughout the training phase. The perturbation is applied to the generative network G_θ during each training phase.

Our suggested approach is capable of using any DCNN architecture for the G_θ generative network construction. Potential HR-HS images frequently have complicated spectra, expressive patterns, and rich textures, all of which demand the full modeling power of the generative network G_θ . Significant advancements have been achieved in generating higher natural images [28], and several generative architectures have been presented, for instance in adversarial learning situations [29].

2.5 Degradation modules

2.5.1 Non-blind degradation module

We apply degradation operations to get approximations of the LR-HS and HR-RGB images from the HR-HS images predicted by the generative network in order to provide evaluation criteria for training the network. However, this part of the network is removed and cannot be included in an integrated training system if only mathematical operations are utilized to approximate the degraded model. In this work, after constructing the backbone, we approximate the degradation model as a conventional learning system utilizing two parallel blocks. To specifically accept

blurred and downsampling transformations, we modified the conventional deep convolutional layers. We apply the same kernel to various spectral bands in the depth-wise convolution layer and set the step space expansion coefficients and bias terms to “false” since the identical blurring and downsampling operations are applied to each spectral band in a real scene. The blurring and downsampling transformations’ equations are written as follows:

$$\hat{X} = f_{SDW}(G_{\theta}(Z_{in})) = k_{SDW} \otimes G_{\theta}(Z_{in})^{(Spa)} \downarrow, \quad (10)$$

where the convolution layer’s specific depth performs the role of $f_{SDW}(\bullet)$. To be more precise, we refer to the same kernel that was used in the depth-wise convolution layer to convolve $G_{\theta}(Z_{in})$ in the HR-HS images generated with each channel independently as $k_{SDW} \in R^{1 \times 1 \times s \times s}$. False bias is accomplished by using conventional two-dimensional mathematical convolution and nearest downsampling operations to transform the spatial expansion factor of $f_{SDW}(G_{\theta}(Z_{in}))$. If the spatially degraded blur kernel is known, we simply set the values to be trained as false values and initialize the weights of each layer based on the known kernel. Similar to this, we simply automatically learned kernel weights of 1×1 during the network training phase or assigned kernel weights of $f_{Spe}(\bullet)$ based on the known RGB camera CSF. Additionally, we employ a conventional convolution kernel with output channels of 3 and a kernel size of 1×1 to implement the spectral transform. We similarly set the stride to 1 and the bias term to false, as shown in the following expression.

$$\hat{Y} = f_{Spe}(G_{\theta}(Z_{in})) = k_{Spe} \otimes G_{\theta}(Z_{in})^{(Spe)} \downarrow, \quad (11)$$

where the activity of the spectral convolution layer is indicated by $f_{Spe}(\bullet)$. The detailed spectra of the obtained HR-HS images are transformed into degenerate RGB images using the convolution kernel $k_{Spe} \in R^{L \times 3 \times 1 \times 1}$. Additionally, the kernel of k_{Spe} minimization that needs to be trained has the same dimension as the $C^{(Spec)}$ that represents the spectrum sensitivity function of an RGB sensor, allowing us to approximate it in our joint network. These two modules can be used concurrently in our integrated learning model by employing the mentioned framework.

2.5.2 (semi-) blind degradation module

This section focuses on automatically learning the transform parameters of the convolutional blocks embedded in the unknown decomposition. For spatially semi-blind, the weight parameter of k_{SDW} in Eq. (10) can either be automatically learned when the blur process is unknown while the weight parameter of k_{Spe} can be predetermined by changing to the parameter of a known CSF kernel. Thus, we can easily extract the approximation LR-HS image from the generated HR-HS image G_{θ} using a specified deep convolutional layer f_{SDW} with a fixed k_{Spe} convolutional kernel. Similarly, it is adaptable to implement the opposite operation to achieve a spectrally semi-blind process. Hence, these two modules can be learned concurrently in our integrated learning framework as a blind degradation module. As a result, the investigated learning model is extremely adaptable and simple to fit into many real-world scenarios. The loss function that was used to train our deep self-supervised network can be rebuilt as follows by substituting the decomposition operation with an improved convolutional block.

$$\begin{aligned} (\theta^*, \theta_{SDW}^*, \theta_{Spe}^*) = \underset{\theta}{\operatorname{argmin}} & \alpha \beta_1 \|X - f_{SDW}(G_\theta)(Z_{in})\|_F^2 \\ & + (1 - \alpha) \beta_2 \|Y - f_{Spe}(G_\theta)(Z_{in})\|_F^2 \quad \text{s.t. } 0 \leq G_\theta(Z_{in}) \leq 1 \forall i. \end{aligned} \quad (12)$$

As can be observed from Eq. (12), in order to rebuild the target well, we learn the generative network parameters rather than directly optimizing the underlying HR-HS image. In our network optimization procedure, the generative network G_θ is trained using only test image pairs (i.e., observed LR-HS and HR-RGB images), and no HR-HS images are provided. This can be seen as a “zero-shot” self-supervised learning method [30]. As a result, we refer to our model as a self-supervised learning model for HSI-SR.

3. Experiment results

3.1 Experimental settings

3.1.1 Datasets

The efficiency of the suggested method was evaluated using two benchmark HSI datasets, namely, CAVE [31] and Harvard [32]. 32 HS images with a spatial resolution of 512×512 are included in the CAVE dataset, which includes various real-world materials. The Harvard dataset includes 50 images of various natural settings, each with a resolution of 1392/1040 pixels and 31 bands of spectral-resolution between 420 and 720 nm. In the experiments, a part of the 1024×1024 sub-image in the top left corner of the Harvard dataset’s original HS image was cropped, resulting in a 512×512 -pixel image that served as the HS image’s main basis. Using different spatial extraction factors (8 and 16) for the bicubic degradation, the observed LR-HS images were generated from the actual HS images of the two datasets, yielding sizes of $64 \times 64 \times 31$ and $32 \times 32 \times 31$. The observed HR-RGB images were also generated by multiplying the HR-HS image by the spectral Nikon D700 camera response function [9].

3.1.2 Evaluation metrics

The proposed method is evaluated against various state-of-the-art methods using five widely used metrics, including root-mean-square error (RMSE), signal-to-noise ratio (PSNR), structural similarity index (SSIM), spectral angle mapping (SAM), and relative dimensional global error (ERGAS). The generated HR-HS image and the ground-truth image were both acquired from the same spatial position. The recovered HR is measured by RMSE, PSNR, and ERGAS which are quantitatively distinct from the reference image to assess the spatial accuracy. Then, SAM offers the average spectral angle of the two spectral vectors to show the spectral accuracy. Additionally, SSIM was employed to evaluate how much the spatial organization of the two images resembled one another. A greater PSNR or SSIM and a lower RMSE, ERGAR, or SAM often indicate superior performance. Bold values mean promising results.

3.1.3 Details of the network implementation

Pytorch has adopted the suggested approach. The input noise was first set to the same size as the HR-HS image that would be generated. Utilizing the Adams optimizer

and a loss function based on the L_2 criteria, the generated network was trained. Initial settings for the learning rate included $1e-3$ with a decrease of 0.7 per 1000 steps. Additionally, the perturbation was reduced by 0.5 every 1000 steps after being initially set at 0.05. After 12,000 iterations, the optimization process was terminated, and all ground-truth HR-HS images from various datasets with various upscale factors were used. Using a Tesla K80 GPU in a training environment, all experiments were carried out. According to our experiments, it takes around 20 minutes to learn an image with a 512×512 size. Across all experiments, we first adjusted the hyperparameter α in the loss function of Eq. (12) to 0.5.

3.2 Performance evaluation

In the study of HS image super-resolution, there are three main paradigms: 1) traditional optimization methods that form image priors based on practical knowledge or physical properties, 2) fully supervised deep learning methods that learn external image priors (training algorithms), and 3) unsupervised methods that learn image priors automatically.

3.2.1 Comparison with traditional non-blind optimization-based methods

The generalization of simultaneous orthogonal matching pursuit (G-SOMP+) method [33], sparse non-negative matrix factorization (SNNMF) method [34], couple spectral unmixing (CSU) method [9], non-negative structured sparse representation (NSSR) method [7], Bayesian sparse representation (BSR) method [35], and other optimization-based HSI-SR methods have all recently been presented. To rebuild stable HS images, conventional optimization-based approaches often employ a variety of hand-crafted priors. The degradation processes (spatial blurring/downsampling and spectral transformations) are a requirement for all approaches. To automatically learn specific priors for latent HR-HS images, we propose a deep unsupervised learning network. In cases when the degradation pattern is unknown, this can yield results for reconstruction. First, we approximated the bicubic decomposition using the Lanczos kernel to initialize the weights of the spatial decomposition blocks, and then we initialized the spectral transform blocks using the CSF of the Nikon D700 camera without learning these blocks in order to make a fair comparison. We evaluate the efficacy of 8 and 16 spatial expansion factors, and compared results on the CAVE and Harvard datasets are shown in **Table 1**. And the visualization results are shown in **Figure 2**.

3.2.2 Comparison with deep non-blind learning-based methods

Deep learning-based methods have recently been thoroughly investigated in the HSI-SR tasks, the majority of them in both fully supervised and unsupervised ways. The unsupervised sparse Dirichlet-net (uSDN) [20], deep hyperspectral image prior (DHP) [36], and GDD method [22] are just a few examples of works that have attempted to use unsupervised strategies in HSI-SR tasks. Our approach comes within the unsupervised branch of HSI-SR methods. In this part, we compare supervised and unsupervised deep learning algorithms, such as SSF-Net [33], ResNet [14], DHSIS [16], uSDN [20], and DHP [36]. Only 12 test images from the CAVE dataset and 10 test images from the Harvard dataset were compared because supervised deep learning methods need training examples to learn the model. The results of the

Up-scale factor = 8										
Dataset	CAVE					Harvard				
Method	RMSE ↓	PSNR ↑	SSIM ↓	SAM ↓	ERGAS ↓	RMSE ↓	PSNR ↑	SSIM ↓	SAM ↓	ERGAS ↓
GOMP [33]	5.69	33.64	—	11.86	2.99	3.79	38.89	—	4.00	1.65
SNNMF [34]	1.89	43.53	—	3.42	1.03	1.79	43.86	—	2.63	0.85
BSR [35]	1.75	44.15	—	3.31	0.97	1.71	44.51	—	2.51	0.84
CSU [9]	2.56	40.74	0.985	5.44	1.45	1.40	46.86	0.993	1.77	0.77
NSSR [7]	1.45	45.72	0.992	2.98	0.80	1.56	45.03	0.993	2.48	0.84
DUFL (Our)	2.08	42.50	0.975	5.35	1.15	2.38	42.16	0.965	2.35	1.09
DUFL+ (Our)	1.96	42.98	0.977	5.22	1.10	2.12	43.23	0.971	2.30	1.01
DSSH (Our)	1.44	45.61	0.992	3.27	0.79	1.17	48.27	0.993	1.75	0.77
Up-scale factor = 16										
GOMP [33]	6.08	32.96	—	12.60	1.43	3.85	38.56	—	4.16	0.77
SNNMF [34]	2.45	42.21	—	4.61	0.66	1.93	43.31	—	2.85	0.45
BSR [35]	2.36	41.57	—	4.57	0.58	1.93	43.56	—	2.74	0.42
CSU [9]	2.87	39.83	0.983	5.65	0.79	1.60	45.50	0.992	1.95	0.44
NSSR [7]	1.78	44.01	0.990	3.59	0.49	1.65	44.51	0.993	2.48	0.41
DUFL (Our)	2.61	40.71	0.967	6.62	0.70	2.81	40.77	0.953	3.01	0.75
DUFL+ (Our)	2.50	41.03	0.969	6.43	0.67	2.56	41.66	0.959	2.95	0.72
DSSH (Our)	1.76	43.84	0.999	3.76	0.49	1.32	47.16	0.992	1.99	0.47

Table 1. Compared results of the conventional non-blind optimization methods with DUFL and DSSH methods in the CAVE and Harvard datasets for up-scale factors: 8 and 16.

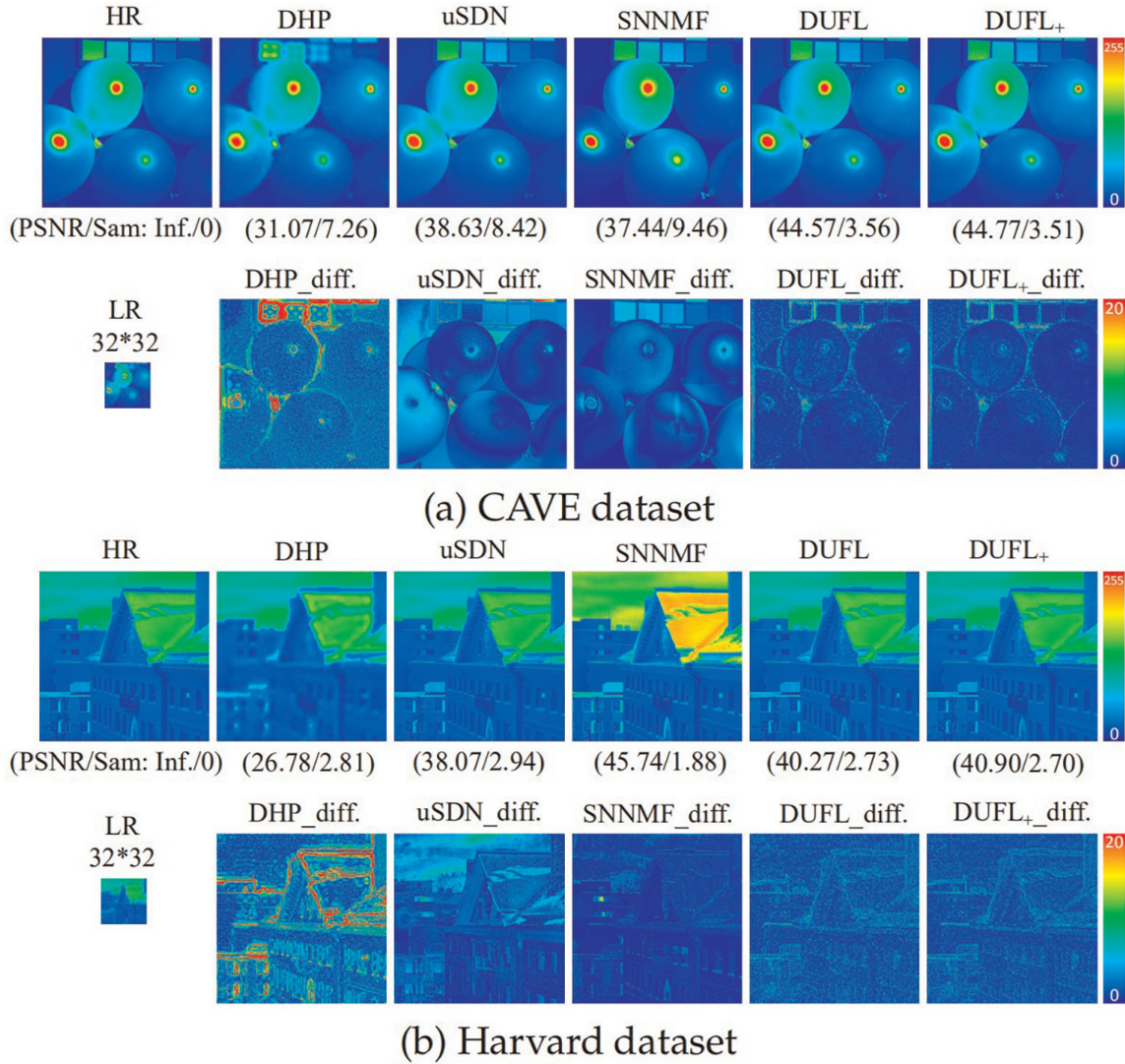


Figure 2. Visualization of the DHP [36], uSDN [20], SNNMF [37], and difference images between the proposed DUFL+ method and the ground-truth/reconstructed images in CAVE and Harvard datasets for an up-scale factor 16.

comparison between the CAVE and Harvard datasets are shown in **Table 2**, with two spatial expansion factors: 8 and 16. It is clear from **Table 2** that our proposed method can perform noticeably better than unsupervised methods based on deep learning, as well as better than supervised methods. And the visualization results are shown in **Figure 3**.

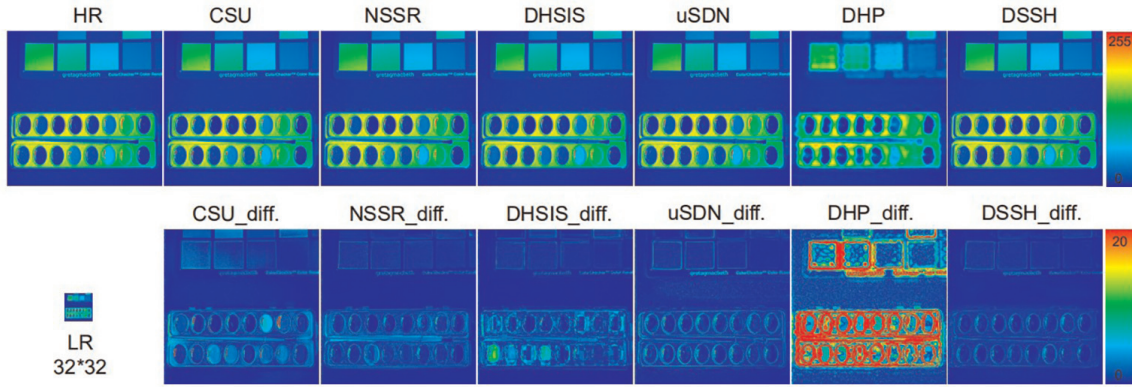
3.2.3 Comparison with (semi-)blind methods

Our proposed method is exploited in a unified framework, which is capable of reconstructing the HR-HS image from the observations not only with the known spatial and spectral degradation operations but also with the unknown spatial or spectral degradation operations or both unknown. Thus, our proposed method can be implemented in a semi-blind setting (the unknown spatial downsampling kernel for LR-HS image or the unknown CSF for HR-RGB image). Consequently, our suggested solution can also be used in total blind mode (unknown spatial degradation operations for LR-HS images and unknown CSF for HR-RGB images). The compared results using our proposed method with semi-blind and complete-blind settings, the state-

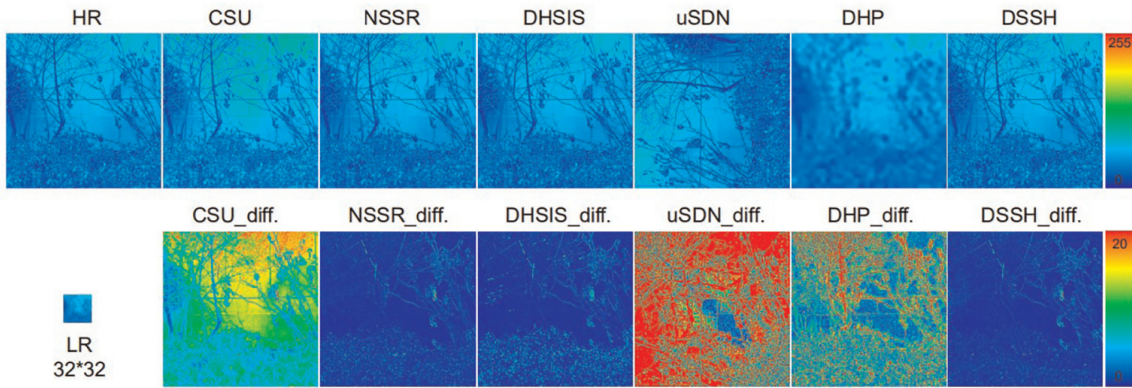
Up-scale factor = 8											
Data		CAVE					Harvard				
	Method	RMSE ↓	PSNR ↑	SSIM ↓	SAM ↓	ERGAS ↓	RMSE ↓	PSNR ↑	SSIM ↓	SAM ↓	ERGAS ↓
Supervised	SSFNet [13]	1.89	44.41	0.991	3.31	0.89	2.18	41.93	0.991	4.38	0.98
	ResNet [14]	1.47	45.90	0.993	2.82	0.79	1.65	44.71	0.984	2.21	1.09
	DHSIS [16]	1.46	45.59	0.990	3.91	0.73	1.37	46.02	0.981	3.54	1.17
Unsupervised	uSDN [20]	4.37	35.99	0.914	5.39	0.66	2.42	42.11	0.987	3.88	1.08
	DHP [36]	7.60	31.40	0.871	8.25	4.20	7.94	30.86	0.803	3.53	3.15
	GDD [22]	1.68	44.22	0.987	3.81	0.96	1.30	47.02	0.990	1.94	0.90
	DUFL (Our)	2.10	42.53	0.978	5.30	1.12	2.15	42.63	0.975	2.32	1.01
	DUFL+ (Our)	2.09	42.39	0.977	4.54	0.91	2.75	40.41	0.965	0.03	0.58
	DSSH (Our)	1.44	45.61	0.992	3.27	0.79	1.17	48.27	0.993	1.75	0.77
Up-scale factor = 16											
Supervised	SSFNet [13]	2.18	41.93	0.991	4.38	0.98	1.94	43.56	0.980	3.14	0.98
	ResNet [14]	1.93	43.57	0.991	3.58	0.51	1.83	44.05	0.984	2.37	0.59
	DHSIS [16]	2.36	41.63	0.987	4.30	0.49	1.87	43.49	0.983	2.88	0.54
Unsupervised	uSDN [20]	3.60	37.08	0.969	6.19	0.41	9.31	39.89	0.931	4.65	1.72
	DHP [36]	11.31	27.76	0.805	10.66	3.09	10.38	38.44	0.754	4.57	2.08
	GDD [22]	2.12	42.24	0.983	4.41	0.61	1.66	44.64	0.986	2.50	0.64
	DUFL (Our)	2.60	40.75	0.970	6.42	0.70	9.46	38.14	0.876	8.52	7.71
	DUFL+ (Our)	2.95	40.56	0.948	2.25	1.15	3.12	39.79	0.945	2.76	0.66
	DSSH (Our)	1.76	43.84	0.999	3.76	0.49	1.32	47.16	0.992	1.99	0.47

Table 2.

Compared results of the deep non-blind learning-based methods with DUFL and DSSH methods in the CAVE and Harvard datasets for up-scale factors: 8 and 16.



(a) CAFE dataset



(b) Harvard dataset

Figure 3.

Visualization of the traditional optimization-based method: CSU [9] and NSSR [7], the supervised deep learning-based methods: DHSIS [16], and the unsupervised deep learning-based methods: uSDN [20], DHP [36], and the proposed DSSH method in the CAFE and Harvard datasets for an up-scale factor 16.

of-the-art unsupervised semi-blind methods: UAL method [23] for spatial blind only, and the spatial blind implementation of NSSR [7] via setting the incorrect spatial kernel, have been given in **Table 3**.

3.2.4 Ablation study

We adjusted the hyperparameters α to 0.3, 0.5, and 0.7 in order to assess the impact of various data circumstances on the loss function of the DUFL method. The comparative results are shown in **Table 4**. The quantitative measurements of our DUFL+ method, PSNR, SAM, and ERGAS, are also shown in **Table 4**, and they demonstrate that the performance of overfitting is not significantly affected by the specific assignment of the hyperparameter α . Similarly, the performance of the DSSH reconstruction method in the ablation study was then evaluated by adjusting α between 0 and 1.0 with an interval of 0.2, and the compared results are shown in **Table 5**.

4. Conclusions

In order to address the super-resolution issue for hyperspectral images, we provide an unsupervised deep hyperspectral image super-resolution framework. A deep

Method	Real downsampling kernel	CAVE					Harvard				
		RMSE ↓	PSNR ↑	SSIM ↓	SAM ↓	ERGAS ↓	RMSE ↓	PSNR ↑	SSIM ↓	SAM ↓	ERGAS ↓
NSSR (Bic) [7]	Bicubic	3.41	38.03	0.968	5.35	1.52	2.76	39.77	0.981	2.00	1.30
NSSR (Ave) [7]	Average	2.76	39.77	0.981	2.00	1.30	3.27	38.55	0.972	5.17	1.78
UAL [23]	K1	1.85	43.23	0.986	6.72	—	2.08	42.38	0.982	2.67	—
	K2	2.01	42.72	0.986	6.78	—	—	—	—	—	—
DSSH (Our) (Spatial blind)	K1	1.47	45.14	0.990	3.54	0.66	1.15	47.59	0.994	1.70	0.78
	K2	1.56	44.71	0.989	3.64	0.69	1.12	47.75	0.994	1.70	0.79
	Bicubic	1.70	44.05	0.988	3.70	0.75	1.33	46.28	0.992	1.95	0.93
DSSH (Our) (Spectral blind)	Bicubic	1.64	44.36	0.989	3.66	0.72	1.28	46.67	0.992	1.86	0.89
DSSH (Our) (Complete blind)	Bicubic	1.68	44.10	0.988	3.72	0.74	1.32	46.44	0.992	1.91	0.91

Table 3. Compared results of the (semi-)blind methods with DUFL and DSSH methods in the CAVE and Harvard datasets for an up-scale factor 8.

Up-scale Factor	α	CAVE			Harvard		
		PSNR \uparrow	SAM \downarrow	ERGAS \downarrow	PSNR \uparrow	SAM \downarrow	ERGAS \downarrow
8	0.3	42.19	5.09	0.95	43.07	2.16	0.93
	0.5	42.91	4.40	0.86	41.68	2.19	1.06
	0.7	42.16	4.75	0.92	41.85	2.18	1.09
16	0.3	40.74	5.71	0.55	40.95	2.90	0.66
	0.5	40.75	5.87	0.54	40.79	2.70	0.62
	0.7	40.42	5.64	0.58	41.90	2.48	0.52

Table 4.

Ablation results of the DUFL+ method with different weights α values of 0.3, 0.5, and 0.7 in the CAVE and Harvard datasets for up-scale factors: 8 and 16.

Dataset		CAVE			
α	RMSE \downarrow	PSNR \uparrow	SSIM \downarrow	SAM \downarrow	ERGAS \downarrow
0.0	25.98	19.97	0.631	40.02	12.50
0.2	1.52	44.99	0.990	3.24	0.67
0.4	1.45	45.45	0.991	3.16	0.63
0.5	1.46	45.35	0.991	3.13	0.64
0.6	1.49	42.26	0.991	3.15	0.66
0.8	1.47	45.20	0.991	3.13	0.66
1.0	3.33	38.36	0.961	4.73	1.51

Table 5.

Ablation results of the DSSH method with different weights α values of 0.0 to 1.0 in the CAVE and Harvard datasets for an up-scale factor 8.

convolutional neural network is used to automatically learn the spatial and spectral features of latent HR-HS images from perturbed noisy input data and the fusion context that naturally collects a significant quantity of low-level image statistics. A special depth-wise convolution layer is designed to achieve degenerate transformations between observations and desired targets, and this generates a universally learnable module that only uses low-quality observations. Without requiring training samples, the proposed unsupervised deep learning framework can efficiently take advantage of the HR spatial structure of HR-RGB images and the detailed spectral characteristics of LR-HS images to deliver more accurate HS image reconstruction. We simply train the network parameters using the observed LR-HS and HR-RGB images and a generative network structure to reconstruct the underlying HR-HS images. Extensive research using the CAVE and Harvard datasets demonstrate promising results in the quantitative evaluation.

IntechOpen

IntechOpen


Author details

Zhe Liu and Xian-Hua Han*

Graduate School of Sciences and Technology for Innovation, Yamaguchi University,
Yamaguchi, Japan

*Address all correspondence to: hanxhua@yamaguchi-u.ac.jp

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Xu JL, Riccioli C, Sun DW. Comparison of hyperspectral imaging and computer vision for automatic differentiation of organically and conventionally farmed salmon. *Journal of Food Engineering*. 2017;**196**:170-182
- [2] Bishop CA, Liu JG, Mason PJ. Hyperspectral remote sensing for mineral exploration in Pulang, Yunnan Province, China. *International Journal of Remote Sensing*. 2011;**32**(9): 2409-2426
- [3] Barnes M, Pan Z, Zhang S. Systems and methods for hyperspectral medical imaging using real-time projection of spectral information. Google Patents; 2018. US Patent 9,883,833
- [4] Bioucas-Dias JM, Plaza A, Camps-Valls G, Scheunders P, Nasrabadi N, Chanussot J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*. 2013;**1**(2):6-36
- [5] Laben CA, Brower BV. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. Google Patents; 2000. US Patent 6,011,875.
- [6] Lanaras C, Baltasavias E, Schindler K. Hyperspectral super-resolution by coupled spectral unmixing. In: *Proceedings of the IEEE International Conference on Computer Vision*. Santiago, Chile: ICCV; 2015. pp. 3586-3594
- [7] Dong W, Fu F, Shi G, Cao X, Wu J, Li G, et al. Hyperspectral image super-resolution via non-negative structured sparse representation. *IEEE Transactions on Image Processing*. 2016;**25**(5): 2337-2352
- [8] He W, Zhang H, Zhang L, Shen H. Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration. *IEEE Transactions on Geoscience and Remote Sensing*. 2015; **54**(1):178-188
- [9] Yokoya N, Zhu XX, Plaza A. Multisensor coupled spectral unmixing for time-series analysis. *IEEE Transactions on Geoscience and Remote Sensing*. 2017;**55**(5):2842-2857
- [10] Akhtar N, Shafait F, Mian A. Sparse spatio-spectral representation for hyperspectral image super-resolution. In: *European Conference on Computer Vision*. Zurich, Switzerland: Springer; 2014. pp. 63-78
- [11] Kawakami R, Matsushita Y, Wright J, Ben-Ezra M, Tai YW, Ikeuchi K. High-resolution hyperspectral imaging via matrix factorization. In: *CVPR 2011*. Colorado Springs, CO, USA: IEEE; 2011. pp. 2329-2336
- [12] Li Y, Hu J, Zhao X, Xie W, Li J. Hyperspectral image super-resolution using deep convolutional neural network. *Neurocomputing*. 2017;**266**: 29-41
- [13] Han XH, Shi B, Zheng Y. Ssf-cnn: Spatial and spectral fusion with cnn for hyperspectral image super-resolution. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*. Athens, Greece: IEEE; 2018. pp. 2506-2510
- [14] Han XH, Sun Y, Chen YW. Residual component estimating CNN for image super-resolution. In: *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*. Singapore: IEEE; 2019. pp. 443-447

- [15] Han XH, Chen YW. Deep residual network of spectral and spatial fusion for hyperspectral image super-resolution. In: 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM). Singapore: IEEE; 2019. pp. 266-270
- [16] Dian R, Li S, Guo A, Fang L. Deep hyperspectral image sharpening. *IEEE Transactions on Neural Networks and Learning Systems*. 2018;**29**(11): 5345-5355
- [17] Han XH, Zheng Y, Chen YW. Multi-level and multi-scale spatial and spectral fusion CNN for hyperspectral image super-resolution. In: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop. Seoul, Korea: ICCVW; 2019
- [18] Xie Q, Zhou M, Zhao Q, Meng D, Zuo W, Xu Z. Multispectral and hyperspectral image fusion by MS/HS fusion net. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, California, USA: CVPR; 2019. pp. 1585-1594
- [19] Zhu Z, Hou J, Chen J, Zeng H, Zhou J. Residual component estimating CNN for image super-resolution. *Hyperspectral Image Super-resolution via Deep Progressive Zero-centric Residual Learning*. 2020;**30**:1423-1428
- [20] Qu Y, Qi H, Kwan C. Unsupervised sparse dirichlet-net for hyperspectral image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah, USA: CVPR; 2018. pp. 2511-2520
- [21] Liu Z, Zheng Y, Han XH. Unsupervised multispectral and hyperspectral image fusion with deep spatial and spectral priors. In: Proceedings of the Asian Conference on Computer Vision Workshops. Kyoto, Japan: ACCV; 2020
- [22] Uezato T, Hong D, Yokoya N, He W. Guided deep decoder: Unsupervised image pair fusion. In: European Conference on Computer Vision. Glasgow, United Kingdom: Springer; 2020. p. 87-102
- [23] Zhang L, Nie J, Wei W, Zhang Y, Liao S, Shao L. Unsupervised adaptation learning for hyperspectral imagery super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: CVPR; 2020. pp. 3073-3082
- [24] Fu Y, Zhang T, Zheng Y, Zhang D, Huang H. Hyperspectral image super-resolution with optimized rgb guidance. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, California, USA: CVPR; 2019. pp. 11661-11670
- [25] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:151106434. 2015
- [26] Ulyanov D, Vedaldi A, Lempitsky V. Deep image prior. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah, USA: CVPR; 2018. pp. 9446-9454
- [27] Seeliger K et al. Generative adversarial networks for reconstructing natural images from brain activity. *NeuroImage*. 2018;**181**:775-785
- [28] Zou C, Huang X. Hyperspectral image super-resolution combining with

- deep learning and spectral unmixing. *Signal Processing: Image Communication*. 2020;**2020**:115833
- [29] He Z, Liu H, Wang Y, Hu J. Generative adversarial networks-based semi-supervised learning for hyperspectral image classification. *Remote Sensing*. 2017;**9**(10):1042
- [30] Imamura R, Itasaka T, Okuda M. Zero-shot hyperspectral image denoising with separable image prior. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. Seoul, Korea: ICCV; 2019
- [31] Yasuma F, Mitsunaga T, Iso D, Nayar SK. Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing*. 2010;**19**(9):2241-2253
- [32] Chakrabarti A, Zickler T. Statistics of real-world hyperspectral images. In: *CVPR 2011*. Colorado Springs, CO, USA: IEEE; 2011. pp. 193-200
- [33] Sims K et al. The effect of dictionary learning algorithms on super-resolution hyperspectral reconstruction. In: *2015 XXV International Conference on Information, Communication and Automation Technologies (ICAT)*. Kyoto, Japan: IEEE; 2015. pp. 1-5
- [34] Kim H, Park H. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics*. 2007;**23**(12):1495-1502
- [35] Akhtar N, Shafait F, Mian A. Bayesian sparse representation for hyperspectral image super resolution. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, Massachusetts, USA: CVPR; 2015. pp. 3631-3640
- [36] Sidorov O, Yngve HJ. Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. Seoul, Korea: ICCVW; 2019
- [37] Wycoff E, Chan TH, Jia K, Ma WK, Ma Y. A non-negative sparse promoting algorithm for high resolution hyperspectral imaging. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE; 2013. pp. 1409-1413