Spectral–Spatial Generative Adversarial Network for Super-Resolution Land Cover Mapping With Multispectral Remotely Sensed Imagery

Cheng Shang, Shan Jiang, Feng Ling, Xiaodong Li, Yadong Zhou, and Yun Du

***Abstract*—Super-resolution mapping (SRM) can effectively predict the spatial distribution of land cover classes within mixed pixels at a higher spatial resolution than the original remotely sensed imagery. The uncertainty of land cover fraction errors within mixed pixels is one of the most important factors affecting SRM accuracy. Studies have shown that SRM methods using deep learning techniques have significantly improved land cover mapping accuracy but have not coped well with spectral–spatial errors. This study proposes an end-to-end SRM model using a spectral–spatial generative adversarial network (SGS) with the direct input of multispectral remotely sensed imagery, which deals with spectral–spatial error. The proposed SGS comprises the following three parts: first, cube-based convolution for spectral unmixing is adopted to generate land cover fraction images. Second, a residual-in-residual dense block fully and jointly considers spectral and spatial information and reduces spectral errors. Third, a relativistic average GAN is designed as a backbone to further improve the super-resolution performance and reduce spectral–spatial errors. SGS was tested in one synthetic and two realistic experiments with multi/hyperspectral remotely sensed imagery as the input, comparing the results with those of hard classification and several classic SRM methods. The results showed that SGS performed well at reducing land cover fraction errors, reconstructing spatial details, removing unpleasant and unrealistic land cover artifacts, and eliminating false recognition.**

***Index Terms*—Deep learning (DL), generative adversarial network (GAN), land cover fractions, spectral–spatial errors, super-resolution mapping (SRM).**

NOMENCLAUTURE

HC	Hard classification.
DL	Deep learning.
GAN	Generative adversarial network.
EIDS	End-to-end, Image-based, and DL-based SRM.
RaGAN	Relativistic average GAN.
RRDB	Residual-in-residual dense block.
MRF	Markov Random Field.
G_{SGS}	Geneative network of SGS.
StdGAN	Standard GAN.
RMSE	Root-mean-square error.
$\mathrm{SGS_U}$	SGS removes CCUL and performs spectral unmixing individually.
$\mathrm{SGS_S}$	SGS uses the StdGAN discriminative network as the backbone.
SRM	Super-resolution mapping.
CNN	Convolutional neural network.
TFDS	Two-step, Fraction-based, and DL-based SRM.
SGS	Spectral-spatial GAN for SRM.
CCUL	Cube-based convolution unmixing layout.
SRCM	Super-Resolution and Conversion Module.
SSI	Spectral and spatial Integration-based SRM.
D_{SGS}	Discriminative network of SGS.
FID	Fréchet inception distances.
OA	Overall accuracy.
$\mathrm{SGS_R}$	SGS does not consider spectral errors without RRDB.
FCLS	Fully constrained least squares.

Manuscript received 24 September 2022; revised 8 November 2022 and 22 November 2022; accepted 3 December 2022. Date of publication 12 December 2022; date of current version 19 December 2022. This work was supported in part by the Natural Science Foundation of Hubei Province under Grant 2022CFB689, in part by the Joint Funds of the National Natural Science Foundation of China under Grant U22A20567, in part by the National Natural Science Foundation of China under Grant 62071457, in part by the Key Scientific Research Projects of Water Conservancy in Hubei Province, China, under Grant HBSLKY202103, and in part by the Application Foundation Frontier Project of Wuhan under Grant 2020020601012283. *(Shan Jiang and Feng Ling contributed equally to this work.) (Corresponding authors: Shan Jiang; Feng Ling.)*

Cheng Shang and Shan Jiang are with the School of Geosciences, Yangtze University, Wuhan 430100, China (e-mail: shangcheng@yangtzeu.edu.cn; jiangshan0712@126.com).

Feng Ling, Xiaodong Li, Yadong Zhou, and Yun Du are with the Key laboratory of Monitoring and Estimate for Environment and Disaster of Hubei province, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430077, China (e-mail: lingf@whigg.ac.cn; lixiaodong@whigg.ac.cn; zhouyadong@asch.whigg.ac.cn; duyun@whigg.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3228741

I. INTRODUCTION

LAND cover is a key component of global environmental information [1], [2], [3]. Satellite sensors are increasingly available as sources for land cover mapping at varying resolutions [4]. Mixed pixels, which may be composed of multiple land cover classes, are essential in land cover mapping with multispectral remotely sensed imagery [5], [6]. Hard classification (HC), by which each pixel is assigned exclusively to one single land cover class, is unsatisfactory, especially when the image has a relatively low spatial resolution [7], [8]. Spectral unmixing, which is designed to predict the fraction values of all land cover classes within each mixed pixel, is viewed as an alternative and equivalent solution to the problem of mixed pixels [9], [10]. However, the spatial distribution of land cover classes within each mixed pixel cannot be correctly predicted [11], [12], [13].

Super-resolution mapping (SRM), which can be viewed as the postprocessing of spectral unmixing, aims to generate a fine-resolution land cover map with the observed remotely sensed

imagery [14]. SRM has been successfully used in mapping water [15], [16], [17], urban buildings [18], [19], and forests [20], [21], [22]. Although SRM has been proven effective in indicating spatial distributions of land cover at subpixel scale, the accuracy of SRM is often limited by spectral errors and spatial errors, which are combined and called spectral–spatial errors. Spectral–spatial errors are often caused by the problem of mixed pixels. Within one mixed pixel, many different spectral signatures and spatial patterns may respond to fine-resolution land cover classes [23], and this many-to-one relationship may reduce SRM accuracy. Spectral errors are often caused by spectral unmixing that provides land cover fraction information, and spatial errors are often caused by spatial dependence models used to describe the spatial patterns of land cover.

The two-step SRM is commonly used to reduce spectral–spatial errors, in which land cover fraction images are first produced, and then input to super-resolution analysis to generate a fine-resolution land cover map [24], [25], [26]. This kind of SRM can be interpreted as fraction-based, as its input is the fraction image generated by spectral unmixing [27]. Although fraction-based SRM is simple and intuitive, it is constrained by spectral unmixing. Because accurate land cover fraction images by spectral unmixing are always unavailable, which is an open problem [28], [29]. Various classical fraction-based SRM methods have been proposed to represent spatial dependence, such as Hopfield neural networks [30], pixel swapping [27], and spatial dependence [31].

Owing to the excellent learning and generalization performance of artificial intelligence, deep learning (DL), such as convolutional neural networks (CNNs) [32], generative adversarial networks (GANs) [33], and graph convolutional networks [34], have gained increased attention in the field of SRM. These DL-based models applied a typical two-step strategy, which first unmixes the remote sensing image to fraction images and then inputs them into a super-resolution DL model to map subpixel land covers. The results showed that these two-step, fraction-based, and DL-based SRM (TFDS) models performed better than conventional non-DL-based fraction-based SRMs.

However, limitations of spectral–spatial errors from TFDS models still exist. In general, these models focus on translating the SRM to the task of super-resolution CNN for land cover fraction images, which are not end-to-end. Therefore, only low- and high-level spatial features for land cover fractions are concerned with a nonlinear relationship, and the uncertainty of land cover fraction images determined by spectral signatures of endmembers and mixed pixels is also not considered. Thus, the accuracy of TFDS models is still strongly influenced by spectral errors from land cover fractions [33], [34].

An alternative, image-based SRM method is directly applied to the observed coarse-resolution multi/hyperspectral remotely sensed imagery. Unlike SRM models that use land cover fraction images as input, image-based SRM integrates spectral unmixing, and super-resolution in a unified model, which may somewhat reduce the impact of land cover fraction errors on the final SRM map. In essence, image-based SRM aims to find the global optimal solution, which comprises spatial and spectral terms [35], [36]. However, this regularization optimization is an ill-posed problem. To convert the ill-posed image-based SRM to well-posed, examples include the Markov random field [37], fuzzy c-means [11], [35], and spectral and spatial integration [36].

Considering the advantage of DL in dealing with spectral information and spatial land cover features, a possible alternative to address spectral–spatial errors lies in end-to-end, image-based, and DL-based SRM (EIDS) models. Attempts have included CNNs [38], [39], [40], [41], [42], GANs [43], [44], and deep residual networks [45]. Consequently, EIDSs avoid spectral errors associated with the production by spectral unmixing. However, these models also have two shortcomings.

First, most end-to-end, image-based, and DL-based SRM (EIDS) models, including CNNs [39], [40], [41], GAN [44], and DRN [45], directly learn the nonlinear relationship between land cover classes and the reflectance of the remotely sensed imagery, and view land cover mapping as an image-segmentation task. In this way, the effect of spatial variability in spectral properties of land covers in remotely sensed imagery has not been fully considered. Thus, there is always uncertainty in EIDSs due to the lack of spatial information about subpixel land cover patterns.

Second, some EIDSs, including CNNs [38], [42] and GAN [43], separately design spectral unmixing and super-resolution layouts and simply incorporate the two layouts into one framework. The spectral errors of spectral unmixing layout will be directly transferred to the super-resolution layout. In turn, the latter super-resolution layout only considers spatial errors during the spatial reconstruction process and may not deal with spectral errors that were passed down along the spectral unmixing layout. Furthermore, the objective functions of these models only calculate the divergence of final land cover classes, and not that of land cover fractions from the spectral unmixing procedure. In practice, these EIDSs are formally considered end-to-end, but are actually two-step, which does not solve spectral errors from land cover fraction images.

Thus, it is important to propose an ideal DL-based SRM model, which deals with spectral unmixing and super-resolution in parallel. This model allows for the consideration of spectral–spatial errors by transferring them to the extraction layout, jointly considering spectral and spatial information, and reducing spectral–spatial errors.

To address the limitations of current methods, we propose a spectral–spatial GAN for super-resolution land cover mapping (SS-GAN-SRM, or SGS), which takes the relativistic average GAN (RaGAN) as an essential framework, comprising a cube-based convolution unmixing layout (CCUL), residual-in-residual dense block (RRDB), and super-resolution and conversion module (SRCM). The motivation of this work is to improve the accuracy of resultant subpixel land cover maps by end-to-end, image-based, and DL-based SRM through the following.

1) CCUL, which plays a role in spectral unmixing and reduces spectral errors from land cover fraction images.
2) RRDB, which jointly considers the complementary information of spectral signatures and spatial textures by transferring the errors to the next architecture.

3) SRCM, which reconstructs more high-frequency spatial details during the super-resolution procedure for land cover fractions and further reduces spectral–spatial errors.

The contributions of this work are as follows.

1) A novel end-to-end, image-based, and deep-learning-based (EIDS) model is proposed to generate subpixel land cover maps from observed multispectral remotely sensed imagery. To the best of authors' knowledge, this is the first GAN framework of DL to simultaneously deal with spectral–spatial errors in SRM.
2) Spectral errors for spectral unmixing from land cover fraction images are reduced by CCUL. Spatial errors for spatial distributions from land cover patterns are reduced by SRCM and RaGAN. By reducing these two errors, the resultant fine resolution land cover maps achieve high performances than conventional SRM models.
3) Unlike traditional SRMs that consider spectral errors and spatial errors independently, these two errors are jointly taken into account under the guidance of RRDB thorough information interaction. Thus, spectral–spatial errors can be effectively reduced.

The rest of this article is organized as follows. Section II introduces the overall network architecture, methodology, detailed structure, loss function, and implementation of the proposed model. Section III compares the performance of the proposed SGS against HC and three state-of-art SRM methods using one synthetic and two realistic experiments with multi/hyperspectral remotely sensed imagery, including spectral and spatial integration-based (SSI), Markov random field-based (MRF), and CNN of DL. Section IV describes the comparative analysis, ablation studies, computational efficiency, and discusses several influencing factors. Section V concludes this article. For convenience, the list of some important acronyms and abbreviations throughout this article is presented in the Nomenclature.

II. Methodology

A. Related Works

1) Super-Resolution Mapping: Assume coarse-resolution multispectral remotely sensed imagery **I** with B bands, and corresponding coarse-resolution land cover fraction image **F** with C land cover classes, where **I** and **F** contain $i \times j$ pixels with spatial resolution R. The objective of SRM is to predict a fine-resolution land cover map **M** from **I** or **F**. By setting the scale factor s, the output high-spatial resolution land cover map contains $(i\times s)\times(j\times s)$ pixels with spatial resolution $r = R/s$, which divides each mixed pixel into s^2 subpixels.

In general, SRM is proving to be a promising method for predicting the spatial distribution of each land cover class at subpixel scale. It takes the land cover fraction values yielded by spectral unmixing and uses these as intermediate input to retrieve an appropriated spatial location for specific land cover fractions. The resultant super-resolution land cover maps often have uncertainty as no information about subpixel patterns within coarse-resolution pixels is used in the model. Thus, SRM can generally be formulated as an inverse, underdetermined, and ill-posed problem. The key issue is to regularize spectral–spatial errors of the mixed pixel problem by characterizing the spatial distributions of the land cover within a mixed pixel. EIDS incorporates the conversion from spectral information to land cover that performs spectral unmixing (**I** to **F**) and coarse-to-fine enhancement that performs super-resolution (**F** to **M**) in one end-to-end network (**I** to **M**). All these operations are conducted by convolutional layers.

2) GAN for SRM: Among DL techniques, GAN is widely used to estimate generative models via an adversarial process [46]. GAN is proposed to realize functional nonlinear mapping in an end-to-end manner, achieving superior performance against previous works. A high correlation of local features and invariance to location shift of input and output images are two fundamental properties of GAN. A GAN consists of a generative network G, and a discriminative network D, which are trained simultaneously. G is trained to generate data from input, and D to distinguish references from the outputs of G [46], [47], [48], [49], [50]. The two networks are trained by playing a two-player minimax game until a Nash equilibrium [46] is reached

$$\min_G \max_D V(G,D) = \mathbb{E}_{x\sim P_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim P_G(z)}[\log(1-D(G(z)))] \quad (1)$$

where x is the real data from true distribution p_{data}, and z is an input sampled from distribution p_z.

GAN has been found capable of solving the spatial dependence of the ill-posed SRM problem [33]. Therefore, we adopt a GAN as an essential framework for EIDS.

B. Overview of the Proposed Model

The proposed SGS model includes a generative network G_{SGS} and discriminative network D_{SGS}, which are simultaneously implemented by a CNN (see Fig. 1). To implement SGS, it is assumed that enough pairs of a training dataset can be collected, each containing a fine-resolution land cover map and corresponding coarse-resolution image. Each pair is used as input to train G_{SGS}, and the role of G_{SGS} is to learn the relationship from the coarse spatial resolution remotely sensed imagery to the fine. D_{SGS} is adversarially trained to output a scalar probability, and the role of D_{SGS} is to further predict the similarity between the generated fine spatial resolution generated by G_{SGS} and real from the training dataset through this scalar probability. After training, G_{SGS} can be used to generate **M** from **I**.

First, spectral unmixing affects the accuracy of land cover fraction images, because spectral unmixing can be simulated as a combination of subpixel spectral signatures within one mixed pixel. From this perspective, it is highly difficult to minimize the difference between the simulated and actual spectral signatures. Therefore, CCUL in G_{SGS} has been proposed to meet the many-to-one nonlinear relationship and reduce spectral errors.

Second, the spatial dependence model is essential to make SRM results spatially smooth, because the super-resolution task is inherently underdetermined inverse, of which the solution is not unique. From this viewpoint, it is a straightforward and effective way to maximize spatial correlations and extract rich and high-level features. Therefore, SRCM in G_{SGS} has been

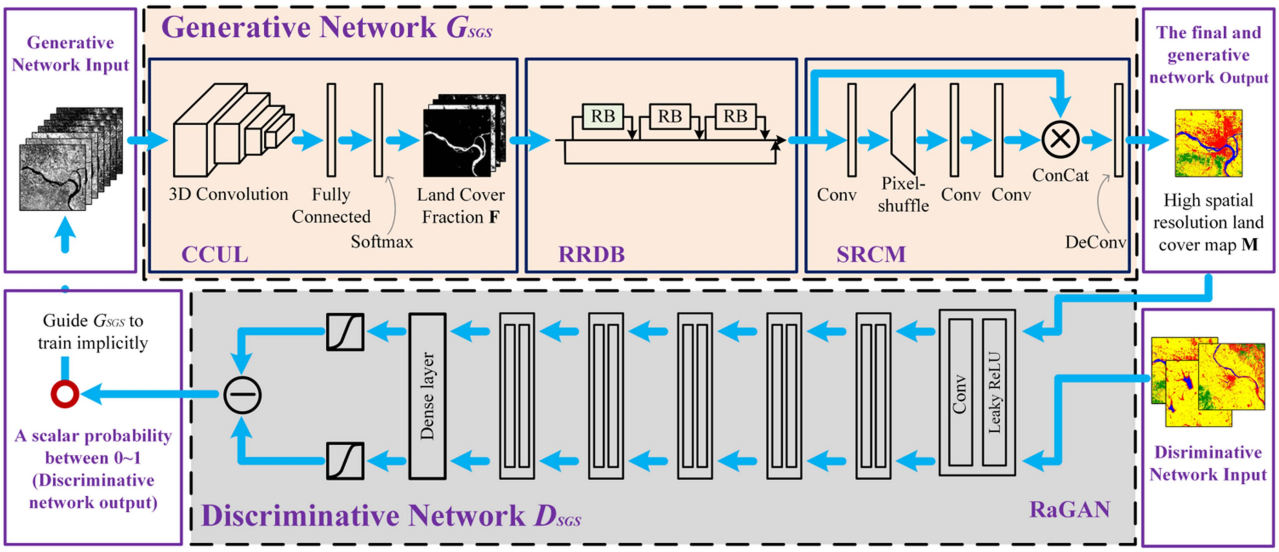


Fig. 1. Network architecture of the proposed SGS.

proposed to mitigate the ill-posed situation and reduce spatial errors.

Third, spectral errors of land cover fraction images are fed directly into super-resolution, meaning that spectral information cannot be fully exploited when dealing with spatial information. Therefore, integrating spectral errors and spatial errors into RRDB and jointly dealing with them can take full advantage of spectral–spatial information.

Finally, D_{SGS} is proposed to further reduce uncertainty through adversarial training mechanisms, contributing to preserving the most valuable spatial details.

In brief, the functions of SGS can be formulated as extracting spectral information, reconstructing spatial details, and reducing spectral–spatial errors; these three functions are embedded in an integrated network, further reducing spectral–spatial errors.

C. Generative Network G_{SGS}

The input of G_{SGS} is the observed coarse-resolution remotely sensed imagery **I** of size $i\times j\times B$, and the output is the coarse-resolution land cover map **M** of size $(i\times s)\times(j\times s)\times 1$. G_{SGS} is a 39-layer CNN that consists of CCUL, RRDB, and SRCM, which realize the interaction of spectral and spatial information and reduce spectral–spatial errors.

CCUL performs spectral unmixing, which fuzzily classifies the input multi/hyperspectral bands from remotely sensed imagery and outputs land cover fraction images. It includes a 3-D convolution, two fully connected layers, and the output layer (a total of seven layers). The 3-D convolution contains four cube-based convolutional layers as higher level feature extractors with a kernel size of 16×16 and 128 feature maps at the first layer and halved at each layer. CCUL can automatically utilize the 3-D data cube so as to avoid the curse of dimensionality from multispectral bands and obtain more representative features without losing helpful information, as does traditional spectral unmixing [51]. Compared with the conventional endmember selection process in spectral unmixing, which only processes spectral information, the cube-based convolutional layer simultaneously handles spectral and spatial information. Hence, it reduces spectral errors during spectral unmixing and realizes spectral and spatial information interaction at the beginning, which makes full use of complementary spectral–spatial feature information, and thus improves the performance and generality of the proposed model.

RRDB reduces the spectral errors, and transfers the reduced errors from spectral unmixing, augmenting network capacity without increasing complexity. The input is the land cover fraction image with large spectral errors from CCUL, and the output is still the land cover fraction image but with reduced spectral errors. RRDB [see Fig. 2(a)] has a multilevel residual-in-residual structure with dense blocks in the main path, and a residual is added every two layers in each dense block. Each dense block contains nine layers, including five 2-D convolutional layers, the first four of which are followed by the LeakyReLU activation function. The feature map of each layer in the dense block is the same as that of the land cover fraction image with a size of $i\times j\times C$. First, the core of RRDB is an autoencoder by which the hidden layers can capture high-level representations and eliminate small changes in the input. Thus, it shows a superior performance in reducing spectral errors. Second, the foundation of RRDB is the residual, which enables the transfer of spectral errors from CCUL to SRCM, yielding spectral and spatial information interaction by reusing spectral features; hence, the two kinds of complementary information can be jointly considered in the proposed model. Third, it is empirically observed that complex and deep network layers are more likely to limit generalization ability. Hence, the proposed model is easier to train and has a high capacity.

SRCM [33] upsamples the land cover fraction feature maps to size $(i\times s)\times(j\times s)$ and converts them to a land cover categorical map. This involves five layers, including three convolutional layers, a pixel-shuffle layer, and one deconvolutional layer. The

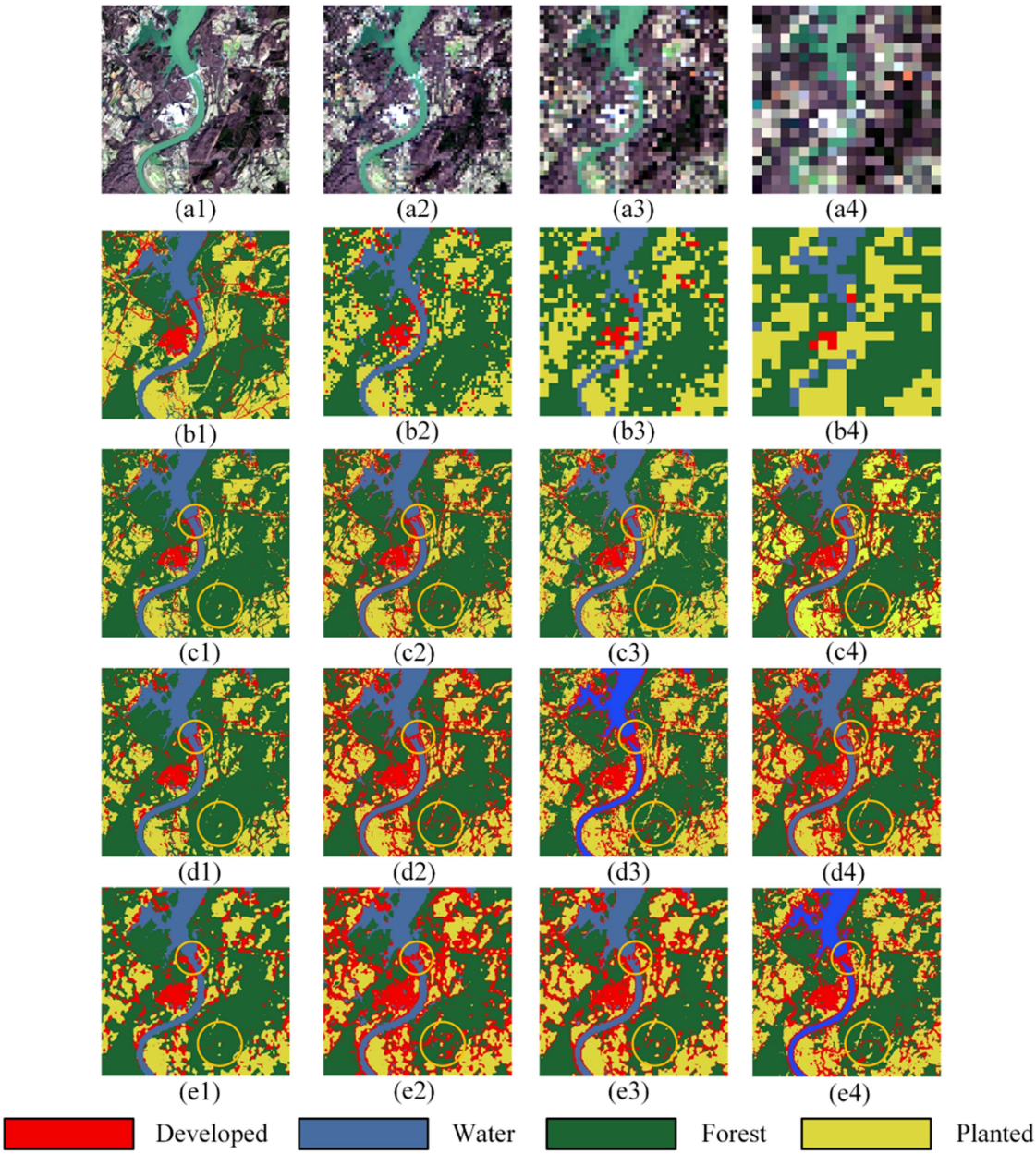


Fig. 2. Input, resultant land cover maps, and references in synthetic experiment. (a1)–(a4) Original and degraded Landsat-8 OLI images (bands 4-3-2, 30m, 120 m, 180 m, 240 m). (b1)–(b4) Land cover maps generated by SVM classifier from (a1) to (a4). Final subpixel land cover maps generated by (c1)–(e1) MRF, (c2)–(e2) SSI, (c3)–(e3) CNN, (c4)–(e4) proposed SGS. Resultant subpixel land cover maps at scale factor: (c1)–(c4) s=4, (d1)–(d4) s=6, (e1)–(e4) s=8.

pixel-shuffle layer is designed for downscaling [46], enabling the high-frequency preservation of spatial details. The sizes of convolutional and deconvolutional layers are the same as that of the output $((i\times s)\times(j\times s)\times 1)$. Combined with three convolutional layers and a deconvolutional layer, SGS can lead to a more reasonable conversion from land cover fraction values to a single land cover class by simultaneously dealing with spectral errors and spatial uncertainties [33].

The generative network G_{SGS} creates a more robust EIDS model, as it should be able to jointly extract complementary spectral and spatial information and simultaneously reduce spectral–spatial errors.

D. *Discriminative Network* D_{SGS}

The discriminative network D_{SGS} trains to correctly predict whether the land cover map to D_{SGS} is synthetic or realistic. The input is a pair of fine-resolution land cover maps of size $(i\times s)\times(j\times s)\times 1$ from the training dataset and G_{SGS}. The output is a scalar between 0 and 1, which can guide its training. The first part of its network architecture is a series of convolutional mapping layers, where the dimension of feature maps is halved, and the number of feature map channels is doubled from the previous layer. The convolutional mapping network consists of 12 convolutional layers and LeakyReLU layers-two for each

feature ($16 \times 16 \times 1$ to $1024 \times 1024 \times 1$). The remaining parts consist of a dense layer, a sigmoid activation function, and a minus operation.

The training process of standard GAN (StdGAN) is unstable because the architecture of the discriminative network is nontransformed and saturated [46], [52]. RaGAN is an extension of StdGAN where the discriminative network takes a relativistic and nonsaturating form [52], [53], [54]. Therefore, rather than measuring the probability that the input land cover map $\mathbf{M}$ is realistic and generated in StdGAN, RaGAN predicts whether a land cover map $\mathbf{M}_r$ from the training dataset is more realistic than a fake one, $\mathbf{M}_f$ ($\mathbf{M}_f = G(\mathbf{I})$). Hence, its output can be translated from $D(\mathbf{M}_r) = \sigma(D_{\mathrm{GS}}(G_{\mathrm{GS}}(\mathbf{F})))$ to $D_{\mathrm{EG}}(\boldsymbol{M}_r, \boldsymbol{M}_f) = \sigma(D_{\mathrm{EG}}(\mathbf{M}_r) - \mathbb{E}_{\mathbf{M}f}[D_{\mathrm{GS}}(G_{\mathrm{GS}}(\mathbf{F}))])$, where σ is the sigmoid function, and $\mathbb{E}_{\mathbf{M}f}[\cdot]$ represents the operation of taking the average for all fake land cover maps in the mini-batch. This modification allows RaGAN to benefit from the gradients from generated and realistic land cover maps in adversarial training, while in StdGAN only the generated part takes effect.

For an existing EIDS, the sources, properties, and influences of spectral information and spatial features are utterly different from those of spectral–spatial errors. Hence, it is difficult to improve super-resolution accuracy and reduce errors under a unified framework. Compared with a conventional EIDS, which only uses a unified G_{SGS} for land cover mapping, D_{SGS} further reconstructs more spatial details and minimizes spectral–spatial errors through a dual-level hierarchical analysis. In other words, the functions of D_{SGS} are reinforcement and strengthening.

E. Loss Function

Based on the idea of SRM and the GAN framework in (1), the loss function is rewritten as

$$L^{\mathrm{SGS}} = L^{\mathrm{Spectral}} + \lambda L^{G} + \eta L^{\mathrm{RaD}} \tag{2}$$

where L^{Spectral} is spectral loss, and ($\lambda L^{G} + \eta L^{\mathrm{RaD}}$) is spatial loss. L^{Spectral} and L^{G} are the generative loss trained from G_{SGS}, L^{RaD} is the relativistic discriminative loss trained from D_{SGS}, and λ and η are tradeoff coefficients to balance different loss terms.

The spectral loss

$$L^{\mathrm{spectral}} = \frac{1}{Cs^2ij} \sum_{i=1}^{si} \sum_{j=1}^{sj} \sum_{c=1}^{C} \left\| \boldsymbol{F} - \frac{\boldsymbol{M}}{s^2ij} \right\|_1 \tag{3}$$

is the $L1$ norm spectral constraint of land class fractions from matching the spectral information between the final subpixel land cover map $\mathbf{M}$ and original remotely sensed imagery $\mathbf{I}$, which is trained by CCUL of G_{SGS}.

The generative loss

$$L^{G} = \frac{1}{s^2ij} \sum_{i=1}^{si} \sum_{j=1}^{sj} (\boldsymbol{M} - G_{\mathrm{SGS}}(\boldsymbol{I})) - \sum_{n=1}^{N} \log D_{\mathrm{SGS}}(G_{\mathrm{SGS}}(\boldsymbol{I})) \tag{4}$$

provides a rough training direction from observed coarse-resolution remotely sensed imagery $\mathbf{I}$ in the fine-resolution land cover map $\mathbf{M}$, which is trained by G_{SGS}.

The relativistic discriminative loss

$$\begin{aligned} L^{\mathrm{RaD}} = & -\log(1 - D_{\mathrm{SGS}}(\boldsymbol{M} - G_{\mathrm{SGS}}(\boldsymbol{I}))) \\ & - \log(D_{\mathrm{SGS}}(G_{\mathrm{SGS}}(\boldsymbol{I})) - \boldsymbol{M}) + \gamma \|\boldsymbol{M} - G_{\mathrm{SGS}}(\boldsymbol{I})\|_1 \end{aligned} \tag{5}$$

is trained by D_{SGS}, where γ is a tradeoff coefficient to balance the discriminative and the $L1$ norm terms. It minimizes the difference between the resultant and reference fine-spatial resolution land cover maps. Thus, through RaGAN and the $L1$ norm, unpleasant or unrealistic land cover artifacts can be removed, and spatially smooth and sharp edges produced.

F. Model Implementation

Backpropagation [55] is employed to minimize the objective loss function in (2)–(5). Hyperparameters are set manually. There are 16 mini-batches and 2000 iterations. The learning rate is initialized as 2×10^{-3} and halved at [200, 400, 800, 1600] iterations. The loss function of SGS has three terms L^{spectral}, L^{G}, and L^{Rad}; Hence, the settings of λ, η, and γ play a vital role. It is found that the optimal ranges of parameters λ, η, and γ were from 4×10^{-2} to 4×10^{-4}, from 0.1 to 0.001, and from 2×10^{-2} to 2×10^{-4}, respectively. Finally, the three tradeoff parameters were measured through empirical and continuous tests, and remained unchanged, with $\lambda = 4 \times 10^{-3}$, $\eta = 1 \times 10^{-2}$, and $\gamma = 2 \times 10^{-2}$. The model uses the Adam optimizer [56], the Python programming language, and the PyTorch (version 1.6.0) deep-learning library. The implementation is done with an Nvidia RTX 3080Ti GPU (Nvidia Corp., Santa Clara, California, USA).

During each mini-batch iteration, black-box preprocessing is required to provide the land cover fraction or spectral information. First, CCUL in G_{SGS} is pretrained during preprocessing. The input is the fine-resolution land cover map, and the output is the corresponding coarse-resolution land cover fraction images. After pretraining, the spectral loss L^{Spectral} [see (3)] is calculated, and the weights and biases of CCUL are obtained. Thus, the coarse-resolution land cover fraction image with large spectral errors is generated as an intermediate output. Second, G_{SGS} is trained, and the fine-resolution land cover map with reduced spectral errors $G_{\mathrm{SGS}}(\mathbf{I})$ is produced by G_{SGS}. The generative loss L^{G} from training is calculated [see (4)], the weights and biases of CCUL are updated, and the weights and biases of the network are obtained. Third, the realistic and fake/generated fine-resolution land cover maps ($\mathbf{M}$/$G_{\mathrm{SGS}}(\mathbf{I})$) are simultaneously fed to D_{SGS} to estimate the probability that the real land cover map from the training dataset is more realistic than the fake one. In this dual-level hierarchical analysis, the relativistic discriminative loss [see (5)] is calculated. This relativistic adversarial convergence process will emerge as a gradient ∇, guiding G_{SGS} to generate more realistic spatial details. The weights and biases of the network are updated using the Adam optimizer by ascending ∇. The optimization process stops if the maximum number of iterations (2000) is reached, or if no less than 0.01% of the loss function is changed between two iterations. Once trained, G_{SGS} of SGS can be used to generate a fine resolution land cover map with the learned parameters of networks.

III. Experiments and Results

A. Benchmarks and Indicators

The performance of the proposed SGS is compared with that of HC and several state-of-art SRM algorithms, including the MRF [37], SSI [36], and CNN-based (CNN) [38]. MRF and SSI are implemented in MATLAB (version r2019b) (MathWorks. Inc, Natick, MA, USA) using an Intel i5-9400 CPU. CNN is implemented in the Keras framework with TensorFlow as the back end, using the same hardware as the proposed SGS.

Overall accuracy (OA) at the subpixel scale is adopted to validate the land cover maps produced by HC and different SRM methods. The Fréchet inception distances (FID) and training error (1-accuracy) of the training dataset and the root-mean-square error (RMSE) of the land cover fraction images are employed to measure the performance.

B. Synthetic Landsat-8 and National Land Cover Dataset (NLCD) Pair Dataset

1) Dataset Preparation: Synthetic Landsat-8 operational land imager (OLI) images and land cover maps from the NLCD were used to assess the performance of SGS. A subset from NLCD (version 2016), which has been applied consistently across the United States at a spatial resolution of 30 m, was used as the test scenario [see Fig. 2(d2)]. This raster-based 16-class land cover map was clustered into four primary land cover classes: developed, forest, water, and planted [33], [34], [57]. The land cover map contained 320 × 320 pixels near the Watts Bar Nuclear Plant and Spring City (35°37'N and 84°50'W), TN, USA. The multispectral image was obtained from Landsat-8 OLI remotely sensed imagery (Path 020, Row 035) on February 28, 2016 [see Fig. 2(a1)]. Synthetic coarse-resolution input images were generated by downsampling the original Landsat image with scale factors of $s = 4$, 8, and 16 [see Fig. 2(a1)–(a3)].

For the training dataset, a subset land cover map of 80 000 × 40 000 pixels (34°13'N–46°6'N and 115°28'W–87°39'W) was selected and divided into 20 000 subimages of 400 × 400 pixels. Corresponding coarse-resolution land cover fraction images in the training dataset were generated by the NLCD land cover maps by degradation for pretraining. Coarse-resolution Landsat-8 images from spring and summer (March–August) 2016 were downloaded from USGS as the input for the training dataset.

For benchmarks, HC was conducted by an support vector machine (SVM) classifier, with samples chosen manually and randomly based on pixel numbers and spatial distributions from NLCD, containing a total of 10 000 pixels. The parameters, settings, and selection of endmembers of MRF, SSI, and CNN were in accordance with previous studies [36], [37], [38]. MRF and SSI do not need training images; the CNN training dataset does not include land cover images and has only fine-resolution subpixel land cover maps and corresponding coarse-resolution remotely sensed imagery.

2) Results: Fig. 2 shows synthetic coarse-resolution remotely sensed imagery as input and resultant fine-resolution subpixel land cover maps produced by HC, MRF, SSI, CNN, and the proposed SGS. A visual comparison demonstrates that SRM methods provided more detailed information than traditional per-pixel HC. In addition, with the increase of scale factors, land cover spatial patterns were more difficult to reproduce, because HC can only generate a land cover map at the pixel scale, and much land cover information within pixels is eliminated. SRM methods could reconstruct the subpixel land cover spatial information to a great extent.

TABLE I
OA of Land Cover Maps Generated by HC and Different SRM Methods Compared With NLCD at Different Scale Factors in the Synthetic Landsat Experiment

	Scale factor *s=4*	Scale factor *s=8*	Scale factor *s=16*
HC	69.32	61.34	47.77
MRF	73.59	66.09	54.11
SSI	76.53	69.53	59.60
CNN	81.38	74.60	65.45
SGS	**85.91**	**79.64**	**70.60**

The best performance is denoted in bold.

For SRM approaches, SGS recovered rich spatial details and prevailed over MRF, SSI, and CNN, as shown in Fig. 4(c4)–(e4). It is noticed that MRF and SSI results had many isolated patches, particularly when $s = 16$ [see Fig. 2(c4) and (d4)], because both models adopt the maximal spatial dependence principle to describe land cover spatial features. Furthermore, linear objects, especially urban roads in the orange circles, were not correctly mapped in the MRF results [see Fig. 2(c2)–(e2)]; they were continuous in the SSI results, but with many fractures [see Fig. 2(c3)–(e3)]. By contrast, these linear objects as generated by EIDS (CNN and SGS) were more abundant and smooth [see Fig. 2(c3)–(e3) and Fig. 2(c4)–(e4)].

For the two EIDSs, the presence of littered plants in the background [forest, also in the orange circle, Fig. 2(c3)–(e3)] indicates that CNN may produce some pepper and salt noise. However, these plaques and noises disappeared in the resultant land cover map by SGS [see Fig. 2(c4)–(e4)]. This is because the primary function of RRDB is denoising. The land cover fraction images, which are generated by CCUL, are approximately the same as those generated by fine-resolution land cover maps, with trivial differences, perhaps due to spectral errors. RRDB, which is a type of autoencoder, is suitable for decreasing spectral divergence within each pixel and denoising.

This demonstrates that SGS significantly improves the SRM performance by removing unpleasant and unrealistic land cover artifacts with fewer noises.

The OA is shown in Table I. OA values decreased with increasing scale factors for all methods because the uncertainty increased with the scale factor. Compared with the result generated from HC, SGS has the highest OA, followed by CNN, SSI, and MRF. The decrease rates of OA in SGS are only about 7.30% and 11.35%, respectively, when the scale factor changes from 4 to 8 and 8 to 16, which are lower than those in the other SRM methods (10.19% and 22.12% in MRF, 9.14% and 14.28% in SSI, 8.33% and 12.26% in CNN). This indicates that SGS was less affected than conventional methods by scale factors.

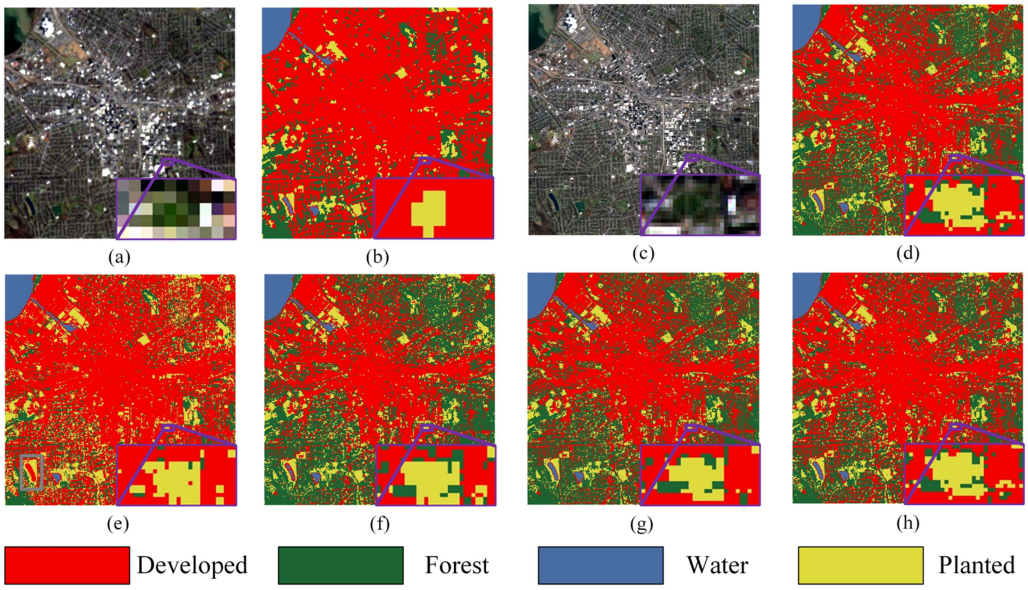


Fig. 3. Input, resultant land cover maps, and references in realistic Landsat-Sentinel experiment. (a) Landsat-9 image (bands 4-3-2, 30 m, 200 × 200 pixels). (b) Land cover maps generated by SVM classifier from (a). (c) Sentinel-2 image (bands 4-3-2, 10 m, 600 × 600 pixels). (d) Reference generated by SVM classifier from (c). Final subpixel land cover maps generated by: (e) MRF, (f) SSI, (g) CNN, (h) proposed SGS.

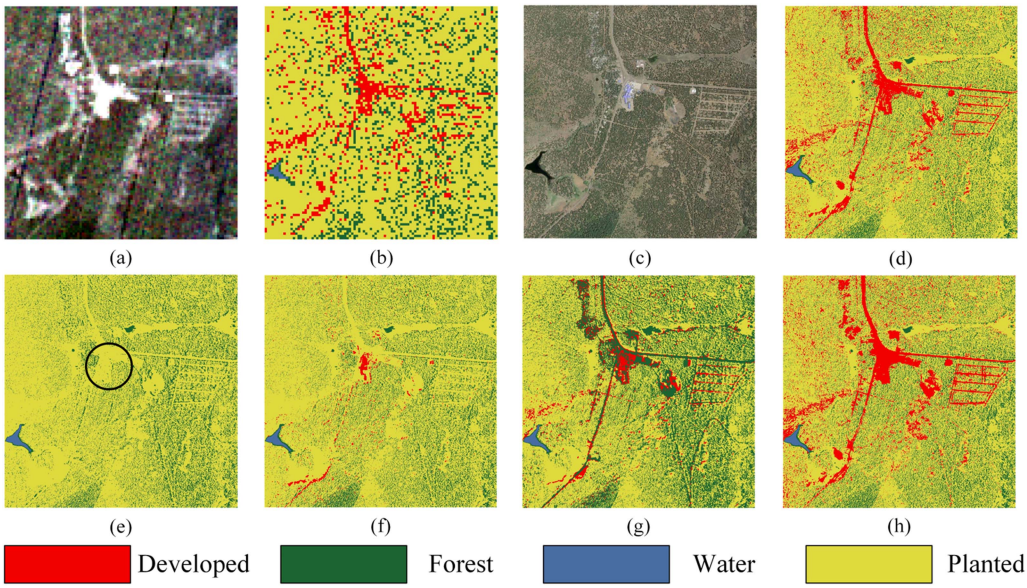


Fig. 4. Input, resultant land cover maps, and references in realistic EO-1 Hyperion experiment. (a) EO-1 image (bands 8-9-10 as red-green-blue, 30 m, 100 × 100 pixels). (b) Land cover map by SVM classifier from (a). (c) Google Digitized image (5 m, 600 × 600 pixels). (d) Refrence by SVM classifier from (c). Final subpixel land cover maps generated by: (e) MRF, (f) SSI, (g) CNN, (h) proposed SGS.

C. Realistic Landsat-9 and Sentinel-2 Pair Multispectral Dataset

1) Dataset Preparation: To further validate the proposed model in real circumstances, we applied a multispectral Landsat-9 image, acquired on March 21, 2022, with a spatial resolution of 30 m [37], consisting of 200 × 200 pixels of seven multispectral bands, covering a portion of the Syracuse University campus (43°02'N and 76°08'W, New York, USA), characterized mainly by artificial infrastructures with a clear structural context. A Sentinel-2 image acquired from the same day and area with a spatial resolution of 10 m was used to produce the subpixel land cover map. This reference was generated by the SVM classifier. The endmembers used in SVM were chosen randomly and manually from the digitized image from Google Earth and contained a total of 5000 pixels. The SVM accuracy was 97.17% when compared with the chosen endmembers from a 5-m digitized image. The scale factor was set to 3.

The training dataset was a set of 20 chips from the southern study area (central area 53°35'N and 112°57'W) of the Boreal Ecosystem-Atmosphere Study of fine-resolution land cover maps in 2021, called the Esri Sentinel-2 10-meter Land

Use/Land Cover (Esri Sentinel-2). The Esri Sentinel-2 training dataset is a nine-class land cover map with a spatial resolution of 10 m over 2 000 000 Sentinel-2 Earth observations (EOs) from six spectral bands from 2017 to 2021, and the land cover was clustered into developed, forest, water, and planted classes. The coarse resolution land cover fraction images were generated from Esri Sentinel-2 by degradation. Corresponding coarse resolution remotely sensed imagery was obtained by Landsat-9 in November and December 2021 from USGS. All land cover fraction images and corresponding remotely sensed imagery required geometry and radiographic correction using the land cover maps from Esri Sentinel-2 as a reference.

The model setups of MRF, SSI, and CNN mimicked the synthetic experiment and previous studies [36], [37], [38]. CNN shared the training dataset of SGS except for land cover fraction images, which CNN did not need.

2) Results: The realistic coarse-resolution Landsat-9 remotely sensed imagery was input and the resultant fine-resolution subpixel land cover maps produced by HC, MRF, SSI, CNN, and the proposed SGS are shown in Fig. 3. A similar trend to the Landsat-NLCD experiment could also be observed in this realistic experiment. As shown in Fig. 3(b), jagged boundaries appeared in the resultant pixel-level land cover map by HC, and an amount of detailed spatial information was lost. For the result of MRF, many land cover patterns were wrongly mapped or mapped as discontinuous. For example, the lake of southeast Syracuse in the scenario [see gray box, Fig. 3(c)], which should be classified as water, was wrongly mapped into developed. For the SSI result in Fig. 3(d), although the continuity of the resultant fine-resolution land cover maps could be basically maintained, numerous indented patches and linear artifacts appeared around the patch boundaries. Compared with MRF and SSI, many isolated small-sized patches caused by spatial patterns were eliminated by CNN and SGS, as shown in Fig. 3(e) and (f).

Visual comparison of the subresults in the purple boxes in Fig. 3 also shows that SGS was more effective than HC and other SRMs. Although there were still some minor differences between the resultant land cover map produced by SGS and the reference, the improvement was noticeable. For example, the central land cover (mainly planted) map by SGS in the purple boxes was more similar to the reference maps than those of MRF, SSI, and CNN. Furthermore, for two EIDSs, the right part of the planted was not preserved by CNN in Fig. 3(e), while it was presented more precisely by SGS in Fig. 3(f). This is because CCUL utilizes deep CNN and 3-D convolutional layer as a higher level feature extractor and a data cube across many data planes comprising multispectral bands. Thus, it can tolerate shift, scale, and distortion invariance, reducing spectral error. For conventional CNNs, the loss function can be roughly interpreted as minimizing the approximate spatial divergence in L^G [see (4)]. The spectral divergence L^{Spectral} [see (3)] was also calculated in SGS, meaning that the joint spectral–spatial errors, rather than spatial errors alone, were fully considered. As a result, the proposed model could avoid high dimensionality and extract more representative spectral–spatial features.

The accuracies of overall and each land cover by different methods are presented in Table II. The OA and accuracy of each land cover of the proposed SGS had the highest values, followed by CNN, SSI, MRF, and HC. The qualitative and quantitative results show that the proposed SGS could improve the accuracy compared with the traditional HC and SRM method directly applied to the coarse-resolution multispectral remotely sensed imagery.

TABLE II
ACCURACY STATISTICS OF LAND COVER MAPS GENERATED BY HC AND DIFFERENT SRM METHODS COMPARED WITH REFERENCE IN THE REALISTIC LANDSAT-SENTINEL EXPERIMENT

	Developed	Forest	Water	Planted	Overall
HC	65.80	68.52	66.72	61.76	65.70
MRF	78.55	73.59	78.53	72.41	75.77
SSI	81.90	78.63	84.38	74.45	79.84
CNN	82.24	78.89	85.80	78.14	81.27
SGS	**87.24**	**82.09**	**91.79**	**86.22**	**86.84**

The best performance is denoted in bold.

D. Realistic EO-1 Hyperion Hyperspectral Dataset

1) Dataset Preparation: In this experiment, the effectiveness of the proposed SGS was validated by a subset of EO-1 Hyperion hyperspectral images with a spatial resolution of 30 m on July 2, 2014, located at Indian Pine, AZ, USA (34°4'N and 109°54'W). A test scenario of 600 × 600 pixels [58] was considered. A total of 44 channels without radiometric calibration (bands 1–7, 58–76, and 225–242) and 22 noisy or water absorption channels (bands 77, 78, 121–127, 167–178, and 224) were removed, and the remaining 186 hyperspectral channels with ample spectral and spatial information were used. The digitized image from Google Earth with a spatial resolution of 5 m, which was assumed to be pure pixels, was used as the ground-truth reference. The subpixel land cover map as reference was generated by SVM, where the endmembers were selected randomly and manually and contained a total of 40 000 pixels. The SVM accuracy was 95.22% compared with the chosen endmembers from a 5-m digitized image.

For the training dataset, the fine-resolution land cover maps were generated from the GF-2 satellite image by SVM. All Gaofen (GF)-2 images were distributed in the 2019 Remote Sensing Image Sparse Representation and Intelligence Analysis Contest (Information Science Department, National Natural Science Foundation of China), which were taken from 2015 to 2016 over developed, forest, water, and planted. The classification samples were selected randomly from Google Earth. Each image has a size of 600 × 600 pixels with a spatial resolution of 4 m, which was then resampled to 5 m (480 × 480 pixels). The EO-1 Hyperion remotely sensed imagery as the same scene and closing time was downloaded from USGS as the training dataset. The scale factor was set to 6. Corresponding coarse resolution land cover fraction images were generated by degradation from fine land cover maps based on GF-2, which were only used in SGS. The model setup was presented in a manner similar to the Landsat-NLCD, Landsat-Sentinel experiment, and previous studies [36], [37], [38].

TABLE III
ACCURACY STATISTICS OF LAND COVER MAPS GENERATED BY HC AND DIFFERENT SRM METHODS COMPARED WITH GOOGLE DIGITIZED IMAGE IN THE REALISTIC EO 1 HYPERION EXPERIMENT

	Developed	Forest	Water	Planted	Overall
HC	60.94	77.13	66.15	78.44	70.67
MRF	65.01	81.29	74.50	77.29	74.52
SSI	75.57	84.31	79.50	79.69	79.77
CNN	78.55	86.10	86.58	86.61	84.46
SGS	**85.30**	**89.38**	**87.89**	**89.79**	**88.09**

The best performance is denoted in bold.

2) Results: Fig. 4 shows the realistic coarse-resolution EO-1 remotely sensed imagery as input, and the resultant fine-resolution land cover maps produced by MRF, SSI, CNN, and SGS, from which it is seen that SRMs [see Fig. 4(c)–(f)] provided more detailed spatial information than HC [see Fig. 4(b)]. HC could only generate land cover maps at the pixel scale, losing a considerable amount of land cover information within pixels, especially the difference between forest and planted. For MRF, many developed and forest pixels were not correctly mapped. It is noticed that in the Indian Pine Restaurant [middle of test scene, black circle, Fig. 4(c)], developed and planted were easily misclassified. For SSI, it can be seen clearly that the main roads and slender linear artificial surface objects with a single-pixel width almost disappeared. By contrast, the linear spatial distributions were approximately mapped by two EIDSs. In addition, for the resultant land cover map by CNN, the developed-planted boundaries were not sufficiently regular and smooth, and some planted patches adjacent to the linear road were wrongly classified as forest, while the SGS results could effectively reduce these.

The accuracy measures of these approaches are shown in Table III. Similar to the previous two experiments, the statistical results indicate the potential of the proposed SGS method in SRM analysis.

IV. DISCUSSION

A. Comparative Analysis

1) Effectiveness of SRM Beyond HC: In all three experiments, the OA of all used SRM methods yielded at least 5%–10% improvements over HC. Theoretical studies [7], [9], [29], [30], [59] indicate that SRM considers the subpixel spatial distribution of land cover fraction information within and between pixels. However, pixel-level HC fails to account for the land cover fraction spatial distribution within the mixed pixel. Hence, SRM revealed a considerable increase in land cover mapping accuracy over HC.

2) Effectiveness of Deep-Learning Over Nondeep-Learning: We analyze additional measures of all three experiments to demonstrate the effectiveness of two EIDSs in coping with mixed pixels. Considering that the accuracy of SRM may be affected by the proportions of pure and mixed pixels in coarse-resolution remotely sensed imagery, the adjusted OA (AOA) after removing pure pixels is summarized in Table IV. It can be observed that the AOA values of all SRM results on all experiments had a certain degree of reduction after removing pure pixels; the Landsat-NLCD experiment had the least reduction, followed by Landsat-Sentinel and EO-1 experiments. These reductions may be related to the spatial distribution and geographic characteristics of the land cover classes in different datasets. For example, the EO-1 experiment is the most fragmented between planted and forest, making the SRM process more difficult because more complicated land cover must be distributed. By contrast, the spatial distribution of land cover in the Landsat-NLCD experiment is the most polymerized, which can bring a more accurate and straightforward SRM process. The quantitative gaps between various SRM methods, especially the DL-based models and conventional nondeep-learning-based methods, further increased.

When dealing with spectral and spatial information, for MRF, it is assumed that the super-resolution land cover map has MRF properties, i.e., adjacent pixels are more likely to belong to the same land cover class than different classes. Thus, the statistical correlation of intensity levels among neighboring pixels can be exploited by maximum likelihood estimation. Moreover, for each subpixel land cover configuration within a mixed pixel in the SSI, the spectral signature of the pixel can be simulated as a combination of the subpixel spectral signatures of its components. From this perspective, the SSI task can be formulated as a multiobject optimization aiming to maximize the spatial dependence and minimize the spectral signature difference by the maximum spatial dependence (MSD) model. Consequently, when no prior information about the spatial distribution of subpixels within neighbor pixels can be provided, these are considered to be randomly distributed. Thus, the distance between the considered subpixel and neighbor pixels, i.e., the spatial dependence, can be calculated by some artificially designed spatial pattern. In practice, however, these subpixels are indeed not distributed randomly, and thus, these artificially designed spatial patterns (MRF or MSD) are inaccurate.

Instead of conventionally extracting handcrafted features, EIDSs automatically learn high-level latent features from the additional training dataset without manual intervention. Furthermore, because the information about the spatial pattern from the training dataset is prior, the spatial pattern description of DL-based models is often closer to reality than conventional SRM methods.

In short, these adjusted experimental results and theoretical analyzes suggest that two EIDSs (CNN and SGS) are more capable than the conventional nondeep-learning-based MRF and SSI of alleviating the mixed pixel problem and accurately characterizing spatial patterns.

3) Effectiveness of GAN Beyond CNN: Previous studies [33], [42], [43], [44] have demonstrated that a GAN has advantages over CNN in reconstructing spatial details. Simple experiments here confirm this conclusion. The FIDs between the generated and realistic land cover maps from the training dataset of all three experiments were calculated during model training, and are shown in Table V. It can be observed that SGS significantly improves FID over CNN, by 10%–20%. This corroborates that the FID values of SGS were lower than those of CNN; thus, SGS

TABLE IV
AOA OF SUBPIXEL LAND COVER MAPS GENERATED BY DIFFERENT SRM METHODS BY REMOVING PURE PIXELSIN ALL EXPERIMENTS

	Landsat-NLCD (*s=4*)	Landsat-NLCD (*s=8*)	Landsat-NLCD (*s=16*)	Landsat-Sentinel (*s=3*)	EO-Google (*s=6*)
MRF	60.53	56.32	51.60	60.14	66.73
SSI	65.59	59.32	55.06	65.60	67.73
CNN	69.53	65.60	59.45	71.13	71.99
SGS	**74.38**	**71.64**	**67.60**	**77.94**	**78.67**

The best performance is denoted in bold.

TABLE V
FIDS FOR COMPARISON BETWEEN SGS AND CNN (LOWER IS BETTER), CALCULATED USING ALL LAND COVER MAPS FROM TRAINING DATASET

	SGS	CNNSRM
Landsat-NLCD (*s=4*)	**2.47**	3.37
Landsat-NLCD (*s=6*)	**3.26**	4.52
Landsat-NLCD (*s=8*)	**5.71**	6.95
Landsat-Sentinel	**7.27**	8.75
EO-Google	**10.42**	11.87

The best performance is denoted in bold.

TABLE VI
RMSE VALUES OF INTERMEDIATE RESULTANT LAND COVER FRACTION IMAGES GENERATED BY SGS_U AND SGS FROM INPUT EO-1 HYPERSPECTRAL REMOTELY SENSED IMAGERY

	Developed	Forest	Water	Planted	Average
SGS_U	6.51	11.14	0.59	13.03	7.82
SGS	**5.61**	**10.73**	**0.47**	**12.63**	**7.56**

The best performance is denoted in bold.

could generate more similar land cover maps than CNN when training. Hence, GAN exhibited a considerable improvement over CNN, confirming previous observations.

B. Ablation Analysis

To study the effectiveness of each component in the proposed SGS, an ablation study was conducted by gradually modifying the baseline model and comparing the differences. Comparisons of ablation models are provided as follows.

SGS_U (SGS removes CCUL and performs spectral unmixing individually): spectral unmixing is conducted by fully constrained least squares (FCLS), and the ablation network uses land cover fraction images estimated by FCLS as input.

SRS_R (SGS does not consider spectral errors without RRDB): This model does not use RRDB; alternatively, its main path does not contain residuals and remains RB. Hence, errors from spectral unmixing are not transferred to super-resolution.

SGS_S (SGS uses the StdGAN discriminative network as the backbone): the discriminative network is trained to directly estimate the probability, indicating whether a land cover map is real or fake.

We conducted three ablation studies on the EO-Google pair hyperspectral experiment ($s = 3$) to demonstrate the effectiveness of CCUL, RRDB, and RaGAN.

1) Effectiveness of CCUL (SGS Versus SGS_U): CCUL is first employed to generate land cover fraction images from the original remotely sensed imagery.

Fig. 5 shows the resultant land cover maps by SGS and SGS_U. Although their spatial distributions are roughly similar to those of the reference map, it can be seen from the result by SGS_U [see Fig. 5(b5)] that the integrity of realistic small-patch and linear spatial features could not be well preserved. SGS retained more spatial details, especially the narrow linear distributions [two black circles, Fig. 5(b5) and (c5)]. Fig. 5(a)–(c) compares spectral unmixing performance by showing the extracted land cover fraction images when coding for the two models. The land cover fraction images by SGS_U-FCLS [see Fig. 5(b1)–(b4)] are very intense looking, spatially sparse and visually binary. Moreover, there are some obvious errors in the water fraction [yellow circle, Fig. 5(b3)], yet SGS avoided this error [see Fig. 5(c3)]. As the tree shadow spectrum of the forest is close to the water spectrum in magnitude, FCLS cannot effectively distinguish between tree, tree shadow, and water, so tree shadows were misclassified as water. This indicates that the proposed CCUL can correctly extract spectral information and obtain more accurate land cover fraction images than conventional spectral unmixing algorithms.

Table VI reports the RMSE values of each land cover fraction image. Again, consistent with visual perception, the average and each RMSE value of SGS scored better, nearly 5%–10% lower than SGS_U-FCLS. These experimental results and theoretical analyzes demonstrate that CCUL effectively utilized spectral–spatial information and achieved superior spectral unmixing performance.

2) Effectiveness of RRDB (SGS Versus SGS_R): RRDB takes full advantage of complementary spectral–spatial information by transferring spectral errors to the super-resolution procedure and reducing them. RRDB has a residual-in-residual structure with dense blocks in the main path. The foundation of RRDB is to realize complementary information interaction. After transferring, the next module of the network architecture benefits from the reuse, recalculation, exploitation, and exploration of spectral–spatial errors, so as to extract superior spectral and spatial features. Additional residual learning inside dense blocks was added to augment network capacity without increasing complexity. Accordingly, the total loss [see (5)] incorporates spectral loss L^{Spectral} [see (3)] that calculates spectral divergences, and generative loss [see (4)] that calculates spatial divergences, so as to combine spectral–spatial features in one model.

Experimentally, the proposed SGS and ablation model SGS_R were trained on the training dataset of the EO-Google hyperspectral experiment without manually setting iteration times until the

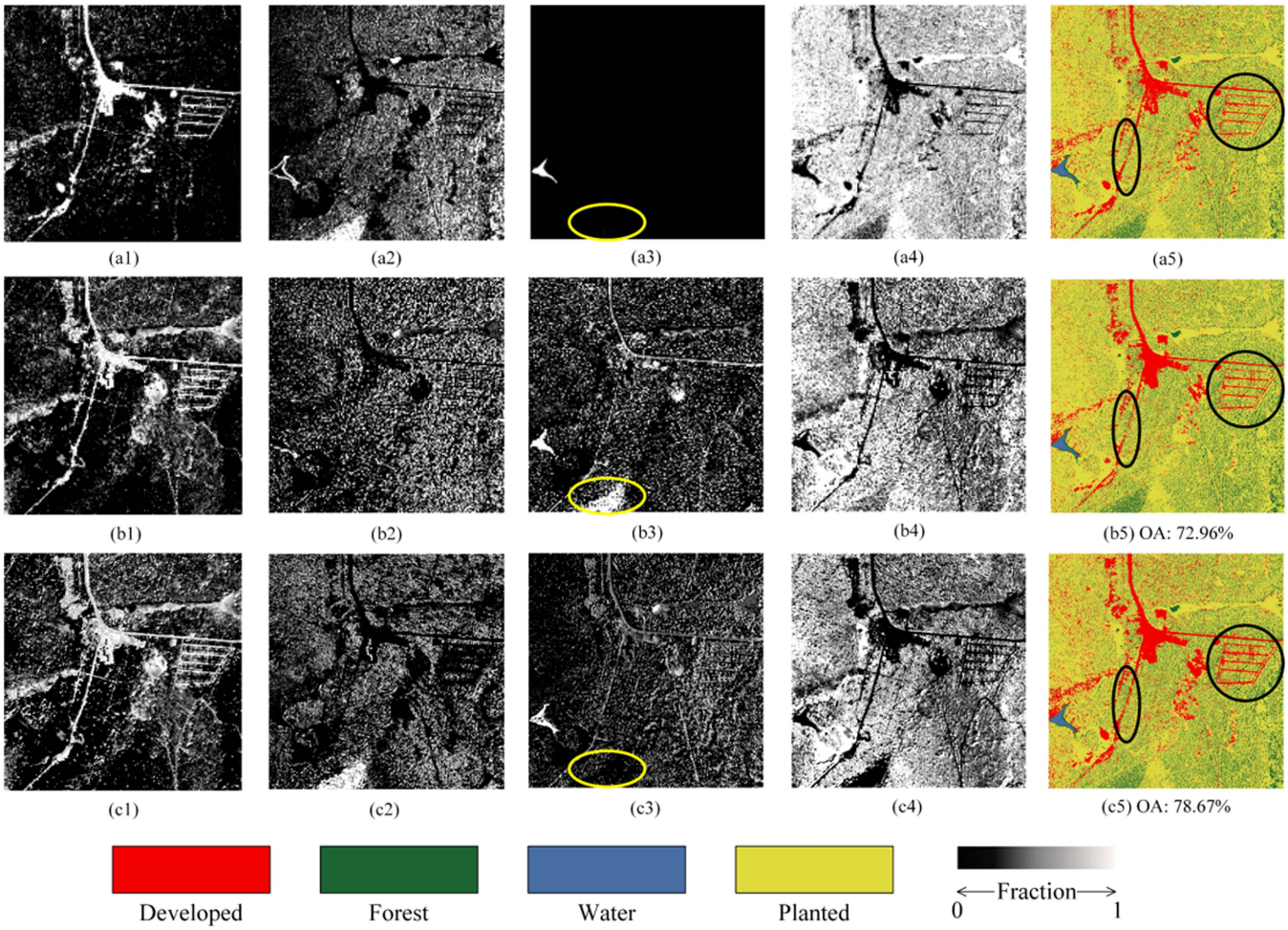


Fig. 5. Ablation experiment 1 conducted on realistic EO-1 (s=6) experiment. Land cover fraction images produced by: (a1)–(a4) degradation with subpixel land cover map, (b1)–(b4) SGS_U-FCLS, (c1)–(c4) SGS with input EO-1 hyperspectral remotely sensed imagery. Subpixel land cover maps: (a5) reference, (b5) SGS_U, (c5) SGS.

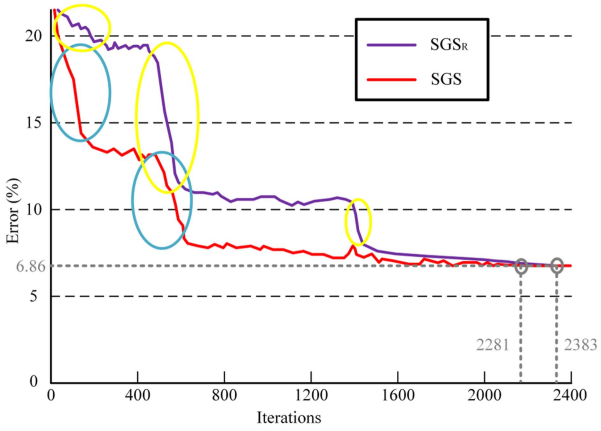


Fig. 6. Ablation experiment 2 (Effectiveness of RRDB) conducted on realistic EO-1 (s=6) experiment. Purple and red curves donote training error (%, 1-accuracy) of SGS_R and SGS, respectively.

training error remained constant. Fig. 6 compares the training errors, and we make four observations.

First, after training, the training error of SGS was lower than that of CNN (i.e., 5.72% versus 6.86%, respectively), which means that SGS performs better in reducing errors compared to CNN. This demonstrates that RDDB enhances the SRM performance by essentially reducing spectral errors, because SGS_R only calculates spatial divergence in the generative loss [see (4)], while SGS calculates both the spectral divergence in the spectral loss [see (3)] and the spatial divergence in the generative loss [see (4)].

Second, SGS has considerably lower error throughout training, indicating that the problem of vanishing/exploding gradients is well addressed.

Third, it is noted that SGS reached the optimal error value at iteration 2281, while SGS_R was still able to find solutions to reach the optimum after about 100 more iterations (2281 to 2383). This means that RRDB helps the network to converge.

Finally, SGS_R attained the optimum after gradient descent (which is indicated by three yellow circles in Fig. 6). In contrast, SGS reached the optimum after rapidly gradients descent (which is indicated by two blue circles in Fig. 6). These studies demonstrate that SGS reached optimization through faster convergence at the early stage of training.

These experimental comparisons verify the effectiveness of RRDB at reducing spectral errors, transferring spectral errors

TABLE VII
STATISTICS OF RESULTANT LAND COVER MAPS OF EO-1 HYPERION EXPERIMENT AFTER CANNY EDGE DETECTION; N_t = TOTAL NUMBER OF EDGE PIXELS; N_f = TOTAL NUMBER OF 4-CONNECTED PIXELS; N_e = TOTAL NUMBER OF 8-CONNECTED PIXELS

	SGS_S	SGS
N_t	31240	**32137**
N_f	16025	**13582**
N_e	1854	**895**
N_e/N_t ($\times 10^{-4}$)	593.47	**278.50**
N_e/N_f ($\times 10^{-4}$)	1156.94	**658.96**

The best performance is denoted in bold. Small Ne/Nt and Ne/Nf mean strong continuity with few interruptions.

TABLE VIII
COMPUTATIONAL EFFICIENCY OF BENCHMARKS, THE PROPOSED MODEL AND ITS ABLATION MODELS IN ALL EXPERIMENTS

	Synthetic Landsat Experiment	Realistic Landsat 9-Sentinel 2 Experiment	Realistic EO-1 Hyperion Experiment
MRF	1236	1537	1966
SSI	1387	1654	2068
CNN	655	826	1204
SGS	**1013**	**1197**	**1570**
SGS_U	828	1108	1429
SGS_R	1183	1217	1550
SGS_S	770	1090	1400

The computational effeciency was recored by the training time (s), the processing time of the proposed model is denoted in bold.

to the super-resolution procedure to jointly consider spectral–spatial information, and making training easier.

3) Effectiveness of RaGAN Beyond StdGAN (SGS Versus SGS_S): The standard discriminative network of StdGAN may produce annoying sharp-shape spatial textures with some unpleasant and unrealistic land cover artifacts [33], [46], [60]. Theoretically, this is because if the discriminative network of StdGAN reaches optimality, the gradient will stop learning what it means for land cover maps to be real, and the training will focus entirely on fake land cover maps, completely ignoring real data [52]. By contrast, for RaGAN, both real and fake land cover maps from the training dataset equally contribute to the gradient of loss functions L^G and L_1. This means that when training, the discriminative network D_{SGS} from SGS acts more globally than StdGAN. As a result, spectral–spatial errors are minimized even more through an RaGAN, which uses both real and fake land cover maps (i.e., dual-level hierarchical analysis), unlike StdGAN, which uses only fake land cover maps.

Canny edge detection, which can reflect the spatial pattern and structural connectivity of land cover, further illustrates why RaGAN is better than StdGAN. From the experimental results in Table VII, the ratios of the total numbers of 8-connected/edge pixels (N_e/N_t) and 8-connected/4-connected pixels (N_e/N_f) of RaGAN were smaller than those of StdGAN, meaning that the improvement of SGS over SGS_S is mainly embodied in the connection degree of the edge lines of land covers. The edge line was the boundary between different land cover classes, and the boundary was a mixed pixel. If the values of N_e/N_t and N_e/N_f were low, then the performance in dealing with boundary/mixed pixels was good.

Thus, it can be concluded that RaGAN of SGS performs better at reducing spectral–spatial errors within mixed pixels, with more robust land cover continuity and fewer interruptions.

C. Computational Efficiency

Table VIII reports the processing time by the proposed method, benchmarks, and the ablation models in all three experiments to evaluate computational efficiency.

Compared to nondeep-learning-based SRM models, deep-learning-based SRMs were much faster than on testing. This advantage is because deep-learning-based SRMs mainly conduct models on GPU, which may be more appropriate for image processing, yet nondeep-learning-based SRMs conduct models on CPU with less computational capability than GPU. The processing time of SGS is rather time-consuming than CNN and ablation models (except SGS_R) because its network is more complex. Indeed, we found this sacrifice of 5% to 10% higher computational costs of SGS to be acceptable, because the proposed models created a significant performance. In addition, the processing time of SGS is faster than SGS_R, proving that RRDB in the proposed model can reduce computational complexity, as previously reported in Section IV-B-2.

D. Future Work

While the abovementioned experiments highlight the accuracy and overall performance of the proposed SGS compared with SRM and EIDSs, several works should be addressed in the future.

1) Scale Factor: The scale factor affects the analysis of remotely sensed imagery and is particularly important for SRM. For example, based on the synthetic Landsat experiment with scale factors $s = 4$, 8, and 16, the mapping accuracy gradually decreased with an increasing scale factor for every SRM method, as can be observed in Table I. This may be because the undetermined spatial location of a subpixel land cover may have multiple solutions with mixed pixels, making the SRM

process more complicated [61]. However, SGS can perform better than the others under all scale factors, demonstrating the generalization and tolerance of the proposed model under various scales.

2) Spatial Extent and Temporal Coverage: Regarding spatial extent and except for synthetic Landsat-8 images, the remotely sensed imagery as input and land cover maps as reference cover substantially the same ground, but not exactly the same—sensors from the Landsat-9/Sentinel-2 and the EO-1 Hyperion/Google Earth have different spatial extents. Regarding temporal coverage, there is a slight difference in acquisition time between the test scene and the reference. Thus, image registration may affect accuracy. The misregistration of spatial extent and temporal coverage may decrease the detection accuracy of fractional unmixing. Subpixel and fully automatic registration approaches should be further studied.

V. Conclusion

We proposed a spectral–spatial RaGAN to solve the ill-posed super-resolution land cover mapping problem. This end-to-end DL-based SRM method with multispectral remotely sensed imagery includes a CCUL, RRDB, and SRCM. In the proposed model, spectral information and spatial information are incorporated in a one-objective loss function, and are jointly considered rather than calculated separately.

The performance of SGS was validated using three datasets with varying spatial resolutions and different multi/hyperspectral remotely sensed imagery. Compared with conventional HC, state-of-art MRF, SSI, CNN of DL, and ablation models, experimental results demonstrated that SGS could yield a better SRM performance, both qualitatively and quantitatively, especially in reconstructing contiguous linear features, eliminating land cover artifacts, and generating geographically realistic spatial details. Future research will focus on the image registration algorithm, parameter value selection, and improvement of the model.

References

[1] J. Townshend, C. Justice, W. Li, C. Gurney, and J. McManus, "Global land cover classification by remote sensing: Present capabilities and future possibilities," *Remote Sens. Environ.*, vol. 35, no. 2/3, pp. 243–255, 1991.

[2] P. M. Atkinson, "Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison study," *Int. J. Remote Sens.*, vol. 30, no. 20, pp. 5293–5308, 2009.

[3] M. A. Friedl et al., "Global land cover mapping from MODIS: Algorithms and early results," *Remote Sens. Environ.*, vol. 83, no. 1/2, pp. 287–302, 2002.

[4] W. B. Cohen and S. N. Goward, "Landsat's role in ecological applications of remote sensing," *Bioscience*, vol. 54, no. 6, pp. 535–545, 2004.

[5] Y. Zhang, F. Ling, X. Li, and Y. Du, "Super-resolution land cover mapping using multiscale self-similarity redundancy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 11, pp. 5130–5145, Nov. 2015.

[6] X. Yang et al., "Spatio-temporal super-resolution land cover mapping based on fuzzy C-means clustering," *Remote Sens.*, vol. 10, no. 8, 2018, Art. no. 1212.

[7] P. M. Atkinson, "Super-resolution mapping using the two-point histogram and multi-source imagery," *GeoENV VI–Geostatistics Environ. Appl.*, vol. 15, pp. 307–321, 2008.

[8] X. Li et al., "Locally adaptive linear mixture model-based super-resolution land-cover mapping based on a structure tensor," *Int. J. Remote Sens.*, vol. 37, no. 23/24, pp. 5802–5825, 2016.

[9] G. M. Foody, "The role of soft classification techniques in the refinement of estimates of ground control point location," *Photogrammetric Eng. Remote Sens.*, vol. 68, no. 9, pp. 897–903, 2002.

[10] P. M. Atkinson, "Mapping sub-pixel boundaries from remotely sensed images," *Innov. GIS*, vol. 4, pp. 184–202, 1997.

[11] X. Li, F. Ling, and Y. Du, "Super-resolution mapping based on the supervised fuzzy c-means approach," *Remote Sens. Lett.*, vol. 3, no. 6, pp. 501–510, 2012.

[12] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[13] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[14] X. Li, Y. Du, and F. Ling, "Sub-pixel-scale land cover map updating by integrating change detection and sub-pixel mapping," *Photogrammetric Eng. Remote Sens.*, vol. 81, no. 1, pp. 59–67, 2015.

[15] G. M. Foody, A. M. Muslim, and P. M. Atkinson, "Super-resolution mapping of the waterline from remotely sensed data," *Int. J. Remote Sens.*, vol. 26, no. 24, pp. 5381–5392, 2005.

[16] F. Ling et al., "Waterline mapping at the subpixel scale from remote sensing imagery with high-resolution digital elevation models," *Int. J. Remote Sens.*, vol. 29, no. 6, pp. 1809–1815, 2008.

[17] W. Li et al., "Locally adaptive super-resolution waterline mapping with MODIS imagery," *Remote Sens. Lett.*, vol. 7, no. 12, pp. 1121–1130, 2016.

[18] F. Ling and B. Fu, "Super-resolution mapping of urban buildings with remotely sensed imagery based on prior shape information," *Joint Urban Remote Sens. Event*, pp. 1–5, 2009, doi: 10.1109/URS.2009.5137529.

[19] M. W. Thornton, P. M. Atkinson, and D. A. Holland, "Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping," *Int. J. Remote Sens.*, vol. 27, no. 3, pp. 473–491, 2006.

[20] Y. Zhang, P. M. Atkinson, X. Li, F. Ling, Q. Wang, and Y. Du, "Learning-based spatial–temporal superresolution mapping of forest cover with MODIS images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 1, pp. 600–614, Jan. 2017.

[21] Y. Zhang et al., "Mapping annual forest cover by fusing PALSAR/PALSAR-2 and MODIS NDVI during 2007–2016," *Remote Sens. Environ.*, vol. 224, pp. 74–91, 2019.

[22] Y. Zhang et al., "Updating landsat-based forest cover maps with MODIS images using multiscale spectral-spatial-temporal superresolution mapping," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 63, pp. 129–142, 2017.

[23] L. Wang and Q. Wang, "Subpixel mapping using markov random field with multiple spectral constraints from subpixel shifted remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 598–602, May 2013.

[24] D. He et al., "Generating continuous fine-scale land cover mapping by edge-guided maximum a posteriori based spatiotemporal sub-pixel mapping," *Sci. Remote Sens.*, vol. 5, 2022, Art. no. 100041.

[25] Y. Chen et al., "Bayesian subpixel mapping of hyperspectral imagery via discrete endmember variability mixture model and markov random field," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 6506–6517, Jul. 2022, doi: 10.1109/JSTARS.2022.3194065.

[26] M. Dastjani, M. J. Valadan Zoej, and M. Jannati, "Soft-then-hard super resolution mapping based on spatial-spectral attraction model and the new class allocation approach based on units of class technique," *Eng. J. Geospatial Inf. Technol.*, vol. 10, no. 1, pp. 11–20, 2022.

[27] P. M. Atkinson, "Sub-pixel target mapping from soft-classified, remotely sensed imagery," *Photogrammetric Eng. Remote Sens.*, vol. 71, no. 7, pp. 839–846, 2005.

[28] X. Li et al., "A spatial-temporal hopfield neural network approach for super-resolution land cover mapping with multi-temporal different resolution remotely sensed images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 93, pp. 76–87, 2014.

[29] K. C. Mertens et al., "A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models," *Int. J. Remote Sens.*, vol. 27, no. 15, pp. 3293–3310, 2006.

[30] A. J. Tatem et al., "Super-resolution land cover pattern prediction using a Hopfield neural network," *Remote Sens. Environ.*, vol. 79, no. 1, pp. 1–14, 2002.

[31] F. Ling et al., "Sub-pixel mapping of remotely sensed imagery with hybrid intra-and inter-pixel dependence," *Int. J. Remote Sens.*, vol. 34, no. 1/2, pp. 341–357, 2013.

[32] F. Ling and G. M. Foody, "Super-resolution land cover mapping by deep learning," *Remote Sens. Lett.*, vol. 10, no. 6, pp. 598–606, 2019.

[33] C. Shang, X. Li, G. M. Foody, Y. Du, and F. Ling, "Superresolution land cover mapping using a generative adversarial network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Sep. 2022, Art. no. 6000105, doi: 10.1109/JSTARS.2022.3194065.
[34] X. Zhang, Y. Ge, F. Ling, J. Chen, Y. Chen, and Y. Jia, "Graph convolutional networks-based super-resolution land cover mapping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 7667–7681, Jul. 2021, doi: 10.1109/JSTARS.2022.3194065.
[35] Y. Zhang, Y. Du, X. Li, S. Fang, and F. Ling, "Unsupervised subpixel mapping of remotely sensed imagery based on fuzzy c-means clustering approach," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 1024–1028, May 2014.
[36] F. Ling, Y. Du, F. Xiao, and X. Li, "Subpixel land cover mapping by integrating spectral and spatial information of remotely sensed imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 408–412, May 2012.
[37] T. Kasetkasem, M. K. Arora, and P. K. Varshney, "Super-resolution land cover mapping using a Markov random field based approach," *Remote Sens. Environ.*, vol. 96, no. 3/4, pp. 302–314, 2005.
[38] Y. Jia et al., "Super-resolution land cover mapping based on the convolutional neural network," *Remote Sens.*, vol. 11, no. 15, 2019, Art. no. 1815.
[39] V. Alhassan et al., "A deep learning framework for land-use/land-cover mapping and analysis using multispectral satellite imagery," *Neural Comput. Appl.*, vol. 32, no. 12, pp. 8529–8544, 2019.
[40] C. D. Storie and C. J. Henry, "Deep learning neural networks for land use land cover mapping," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2018, pp. 3445–3448.
[41] W. Bupphawat et al., "Super-resolution land cover mapping based on deep learning and level set method," in *Proc. Int. Conf. Elect. Eng./Electron., Comput., Telecommun. Inf. Technol.*, 2017, pp. 557–560.
[42] D. He, Y. Zhong, X. Wang, and L. Zhang, "Deep convolutional neural network framework for subpixel mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9518–9539, Nov. 2021.
[43] J. Xie, L. Fang, B. Zhang, J. Chanussot, and S. Li, "Super resolution guided deep network for land cover classification from remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Oct. 2022, Art. no. 5611812, doi: 10.1109/JSTARS.2022.3194065.
[44] M. Liu, Q. Shi, A. Marinoni, D. He, X. Liu, and L. Zhang, "Super-resolution-based change detection network with stacked attention module for images with different resolutions," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Feb. 2022, Art. no. 4403718, doi: 10.1109/JSTARS.2022.3194065.
[45] C. Wang et al., "Deep residual convolutional neural network for hyperspectral image super-resolution," in *Proc. Int. Conf. Image Graph.*, 2017, pp. 370–380.
[46] I. Goodfellow et al., "Generative adversarial nets," *Adv. Neural Inf. Process. Syst.*, vol. 27, no. 3, pp. 2672–2680, 2014.
[47] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proc. Int. Conf. Learn. Representations*, 2015, doi: 10.48550/arXiv.1511.06434.
[48] S. Reed et al., "Generative adversarial text to image synthesis," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1060–1069.
[49] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4217–4228, Dec. 2021.
[50] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," *Comput. Ind.*, vol. 106, pp. 85–93, 2019.
[51] J. Leng, T. Li, G. Bai, Q. Dong, and H. Dong, "Cube-CNN-SVM: A novel hyperspectral image classification method," in *Proc. IEEE 28th Int. Conf. Tools Artif. Intell.*, 2016, pp. 1027–1034.
[52] A. Jolicoeur-Martineau, "The relativistic discriminator: A key element missing from standard GAN," in *Proc. Int. Conf. Learn. Representations*, 2018, doi: 10.48550/arXiv.1807.00734.
[53] S. Zhang, D. Cheng, D. Jiang, and Q. Kou, "Least squares relativistic generative adversarial network for perceptual super-resolution imaging," *IEEE Access*, vol. 8, pp. 185198–185208, 2020.
[54] H. Yu et al., "A super-resolution generative adversarial network with simplified gradient penalty and relativistic discriminator," in *Proc. Int. Joint Conf. Neural Netw.*, 2019, pp. 1–8.
[55] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Proc. Int. Joint Conf. Neural Netw.*, vol. 1, pp. 593–605, 1989, doi: 10.1109/IJCNN.1989.118638.
[56] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, 2017, doi: 10.48550/arXiv.1412.6980.
[57] F. Ling et al., "Learning-based superresolution land cover mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3794–3810, Jul. 2016.
[58] F. Ling et al., "Superresolution land cover mapping with multiscale information by fusing local smoothness prior and downscaled coarse fractions," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5677–5692, Sep. 2014.
[59] F. Ling, X. Li, Y. Du, and F. Xiao, "Super-resolution land cover mapping with spatial-temporal dependence by integrating a former fine resolution map," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 5, pp. 1816–1825, May 2014.
[60] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *Comput. Sci.*, pp. 2672–2680, 2014, doi: 10.48550/arXiv.1411.1784.
[61] F. Ling, X. Li, F. Xiao, and Y. Du, "Superresolution land cover mapping using spatial regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4424–4439, Jul. 2014.

Cheng Shang received the B.S. degree in land resource management from the China University of Geosciences, Wuhan, China, in 2014 and the Ph.D. degree in physical geography from the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, China, and also from the University of Chinese Academy of Sciences, Beijing, China, in 2021.

He is currently a Lecturer with the School of Geosciences, Yangtze University, Wuhan, China. He has been supported by the Hubei Province Natural Science Fund for Youth Project. His research interest is geo-information science and technology, including remote sensing and its application in geosciences.

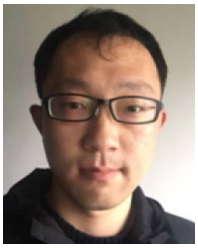
Shan Jiang received the B.S. degree in geological exploration of oil and gas from Jianghan Petroleum College, Jingzhou, China, in 2001, the M.S. degree in mineral survey and exploration from Yangtze University, Jingzhou, China, in 2004, and the Ph.D. degree in mineral survey and exploration from the China University of Geosciences, Beijing, China, in 2011.

He is currently an Associate Professor and the Director of the Department of Geoscience Information Science and Technology, School of Geosciences, Yangtze University, Wuhan, China. His research interest is oil-gas pool numerical simulation and geoscience informatization.

Feng Ling received the B.S. degree in geology and the M.S. degree in geophysical prospecting and information technology from the China University of Geosciences, Wuhan, China, in 1999 and 2002, respectively, and the Ph.D. degree in system analysis and integration from the Huazhong University of Science and Technology, Wuhan, China, in 2006.

He is currently a Professor with the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences. He has also been a Visiting Scholar with the University of Nottingham, Nottingham, U.K., in 2016. He has authored or coauthored more than 100 refereed journal articles. His research interests include super-resolution land cover mapping, spatial-temporal fusion of remotely sensed imagery, and water resources monitoring.

Dr. Ling is currently an Associate Editor for Remote Sensing in Ecology and Conservation, and Remote Sensing Letters. He was the recipient of the 2016 Boeing Award for Best Scientific Paper in Image Analysis and Interpretation from the American Society for Photogrammetry and Remote Sensing and the 2020 Taylor and Francis Remote Sensing Letters Award.



Xiaodong Li received the B.S. degree in geographic information system from the China University of Geosciences, Wuhan, China, in 2006 and the M.S. and Ph.D. degrees in physical geography from the Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, China, in 2009 and 2012, respectively.

He is currently a Professor with the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences. He has been supported by the Youth Innovation Promotion Association, CAS, and the Hubei Province Natural Science Fund for Distinguished Young Scholars. His research interest is super-resolution mapping of remotely sensed imagery.

Yadong Zhou received the Ph.D. degree in physical geography from the University of Chinese Academy of Sciences, Beijing, China, in 2021.

She has been a Visiting Scholar with the Institute for Electromagnetic Sensing of the Environment, National Research Council, Milan, Italy, from 2019 to 2021. She is currently a Postdoctoral Researcher with the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences. Her research interests include remote sensing and Earth observation, water monitoring in inland and coastal environments, geographic information systems, and machine learning.

Yun Du received the B.S. degree in geomorphology and quaternary geology from Nanjing University, Nanjing, China, in 1989 and the M.S. degree in physical geography from the Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, China, in 1992, and the Ph.D. degree in historical geography from Wuhan University, Wuhan, China, in 1999.

He is currently a Professor with the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences. His research interests are remote sensing applications and hydrology.