Region-Kernel-Based Support Vector Machines for Hyperspectral Image Classification

Jiangtao Peng, Yicong Zhou, *Senior Member, IEEE*, and C. L. Philip Chen, *Fellow, IEEE*

***Abstract*—This paper proposes a region kernel to measure the region-to-region distance similarity for hyperspectral image (HSI) classification. The region kernel is designed to be a linear combination of multiscale box kernels, which can handle the HSI regions with arbitrary shape and size. Integrating labeled pixels and labeled regions, we further propose a region-kernel-based support vector machine (RKSVM) classification framework. In RKSVM, three different composite kernels are constructed to describe the joint spatial–spectral similarity. Particularly, we design a desirable stack composite kernel that consists of the point-based kernel, the region-based kernel, and the cross point-to-region kernel. The effectiveness of the proposed RKSVM is validated on three benchmark hyperspectral data sets. Experimental results show the superiority of our region kernel method over the classical point kernel methods.**

***Index Terms*—Composite kernel, hyperspectral image (HSI) classification, region kernel, support vector machine (SVM).**

I. Introduction

HYPERSPECTRAL data contain a set of images with the same geographic scene. These images correspond to different spectral bands of electromagnetic radiation. Fixed a band, the hyperspectral data reduce to a single image containing the scene structure information of different materials. Fixed an image coordinate, it obtains a spectral curve vector, which is called a pixel [here, "pixel" refers to "sample" in a hyperspectral image (HSI)]. Different materials have different absorptions or reflections at a certain spectral band. Thus, it can identify and classify the materials based on their spectral curves. Traditional classifiers, such as the Bayesian classifier, the k-nearest neighbor classifier, and neural networks, use the spectral signatures in the HSI classification.

However, these traditional methods usually encounter difficulty because of the measured noise, interruption, nonlinear spectral responses [1], and the Hughes phenomenon caused by the high-dimensional small samples [2]. These make the HSI classification an extremely challenging problem [1], [3], [4]. Support vector machine (SVM) and other kernel-based methods [5]–[8] overcome these limitations at a certain extent because of their good abilities for handling high-dimensional small-sample-sized nonlinear and noisy data [1]. In spite of these desirable properties, SVM is rapidly turning out to be insufficient to exploit the rich HSI information [1]. In the SVM classification, each labeled HSI pixel is considered as a sample point and processed independently. It ignores the correlations between the spatial neighboring pixels. For an HSI, spatial neighboring pixels have similar spectral characteristics and usually belong to the same class [9], [10]. The relevance between neighboring pixels can be exploited to increase the consistency and accuracy of the SVM classification [4], [11]–[13]. In fact, for a given pixel, we can extract the neighborhood size, shape, and structure distribution information [14]. For the pixels belonging to different materials, the corresponding spatial geometrical structure information will be different. This spatial information can be used to improve the classification of pixels that are hard to discriminate using spectral features alone. Therefore, to achieve an excellent HSI classification performance, it would be better to combine the spectral and spatial information.

Recently, many spatial–spectral classification methods have been proposed to improve the HSI classification performance [4], [15]. Morphological-transformation-based spatial–spectral classifier extracts the morphological profile (MP) features from the first several principal components of hyperspectral data and fuses the extracted feature and the original spectral feature into one stacked vector for the SVM classification [16], [17]. There are also other spatial–spectral feature extraction methods, such as multiple spatial–spectral features [18] and tensor features [19], [20]. Markov random fields (MRFs) are usually combined with SVM to perform a joint spatial–spectral classification in an integrated Bayesian framework, where SVM is used to estimate the Bayesian-like probability outputs, and MRF regularization refines the SVM classification results using the spatial contextual information [13], [21], [22]. In SVM with composite kernels (SVM-CK), a composite kernel was designed using both the spectral and spatial features, where the spatial feature is represented as the mean or standard deviation of pixels within a spatial pixel neighborhood [11]. Based on SVM-CK, a new composite spatial–spectral kernel was proposed in [14] and [23], where the spatial feature for each pixel is represented as the vector median of pixels in a morphological neighborhood obtained by the area filtering.

Manuscript received June 2, 2014; revised August 17, 2014, November 20, 2014, and January 12, 2015; accepted March 4, 2015. This work was supported in part by the Macau Science and Technology Development Fund under Grant FDCT/017/2012/A1; by the Research Committee at University of Macau under Grant MYRG113(Y1-L3)-FST12-ZYC, Grant MRG001/ZYC/2013/FST, and Grant MYRG2014-00003-FST; and by the National Natural Science Foundation of China under Grant 11371007. *(Corresponding author: Yicong Zhou.)*

J. Peng is with the Faculty of Mathematics and Statistics and the Hubei Provincial Key Laboratory of Applied Mathematics, Hubei University, Wuhan 430062, China, and also with the Department of Computer and Information Science, University of Macau, Macau 999078, China (e-mail: pengjt1982@126.com).

Y. Zhou and C. L. P. Chen are with the Department of Computer and Information Science, University of Macau, Macau 999078, China (e-mail: yicongzhou@umac.mo; Philip.Chen@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2015.2410991

0196-2892 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

All of the aforementioned spatial–spectral methods combine the useful spatial information with the powerful kernel-based SVM classifier. Whether the spatial information is used at the feature level or the decision level, the final SVM classification is performed on feature vectors, such as MP features used in the morphological-based classification methods and mean or standard deviation vectors used in the SVM-CK methods. That is, although the kernel method (SVM) is used for similarity metric and classification, the processing features are individual pixels or sample points. Thus, these spatial–spectral classifiers can be regarded as the sample-point-based methods, in which a sample vector cannot fully capture the spatial local neighborhood variability of the spectral signature. Because an HSI contains many homogeneous regions, the pixels in a homogeneous region generally belong to the same class [9]–[11], [22], [24], [25]. It can classify these local homogeneous regions directly rather than classifying the feature vectors (i.e., mean or standard deviation vectors) extracted from the regions. One of the existing region classification methods is the object-based classification methods. They group the spatially adjacent pixels into homogeneous objects and then classify these homogeneous objects [9], [10], [24]. Although these methods use regions or homogeneous objects as features, they have never used the kernel metric to measure the similarity between different regions.

In this paper, based on the HSI spatial homogeneous regions and the kernel similarity metric, we propose a region-kernel-based SVM (RKSVM) to classify the local homogeneous region of each HSI pixel. In the region classification, a key point is to measure the distance or the similarity between different regions, which is achieved by a novel region kernel in RKSVM. In detail, for each pixel $\boldsymbol{x}$ (spectral curve vector), a corresponding local region $\boldsymbol{R}$ containing $\boldsymbol{x}$ and its spatial neighbors is generated from either a distance similarity neighborhood or an area-filtering-based morphological neighborhood. As the local homogeneous distribution of an HSI, pixel $\boldsymbol{x}$ and its spatial neighbors show similar spectral features. That is, the spectral values of pixels in the local region $\boldsymbol{R}$ change in a small range. Based on the range of spectral values in each band, it can determine an interval in each band and hence a regular box to be the Cartesian product of multiple intervals, which reflects the spectral variations of local homogeneous pixels in the region. To describe spectral variations more accurately and stably, multiscale boxes with different sizes are generated. For different regions, we can construct multiscale boxes for each region separately and compute the corresponding pairwise box-to-box similarities of different regions by means of box kernels [26]. Then, the similarity between different regions is defined as a weighted average of pairwise box-to-box kernels, which is called the region kernel in this paper. Based on the region similarity metric, we can classify HSI regions.

Different from the sample-point-based methods that operate on the individual samples (spectral curves), compute the distance similarity metric between different samples, and classify a sample based on the computed sample-point-based similarity metric, the region kernel method operates on the regions (spectral curve sets) and computes the region-to-region kernel similarity metric to classify each region. In the region kernel method, each sample has a corresponding region. Therefore, by classifying the regions, it can get the classification results of each sample.

Different from the object-based classification methods, such as the Extraction and Classification of Homogeneous Objects (ECHO) [9], which segments the image into homogeneous objects and then classifies these homogeneous objects based on the minimum distance strategy or the maximum-likelihood strategy, our RKSVM classifies local regions using the region kernel. Instead of using homogeneous objects, RKSVM uses individual spectral pixels and their corresponding local regions or local homogeneous objects as features. In RKSVM, each pixel corresponds to a local region. The classification samples consist of labeled pixels and labeled local regions, where the number of labeled local regions is equal to the number of labeled pixels, whereas in ECHO, the classification samples are homogeneous objects. When the number of homogeneous objects is small, the classification performance may be affected [24]. Because the maximum-likelihood strategy for region classification in ECHO needs to estimate the probability density function, it is usually a difficult problem, particularly in the case of small sample sizes. However, the region distance metric or the classification method is the region kernel in RKSVM. This is because kernel-based classification is widely used in the fields of machine learning and HSI classification [1], [5], [6], [11]. By means of the region kernel, it is easy to perform the linear and nonlinear classifications of the regions.

The proposed RKSVM performs the region classification by integrating both the spectral pixel points and the spatial local regions. It has at least the following characteristics.

1) It is the first to use a kernel to measure the region-to-region distance similarity for the HSI classification. By means of a region kernel, it is easy to describe the linear or nonlinear relations between different regions.
2) It designs a novel stack composite kernel, which integrates the point-to-point kernel, the region-to-region kernel, and the cross point-to-region kernel.
3) Based on the distance similarity or the morphological connectivity, the region can be an arbitrary shape. Using the percentile-based method, the extracted interval or box features from the region can resist noisy pixels and abrupt points.
4) It generalizes the state-of-the-art SVM-CK methods from point to region and is more effective in the case with limited training samples.

The rest of this paper is organized as follows. In Section II, the proposed RKSVM is described in details. The experimental results and analysis are provided in Section III. Finally, Section IV gives a summary of our work.

II. Proposed Method

In the SVM and other kernel-based methods, one can use any kernel function $K(\cdot,\cdot)$ that fulfills Mercer's condition (symmetric and positive semidefinite) to measure the similarity between sample points. The commonly used kernel functions are linear, polynomial, and Gaussian radial basis functions (RBFs) [5], [6]. Because the general kernel function describes the

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

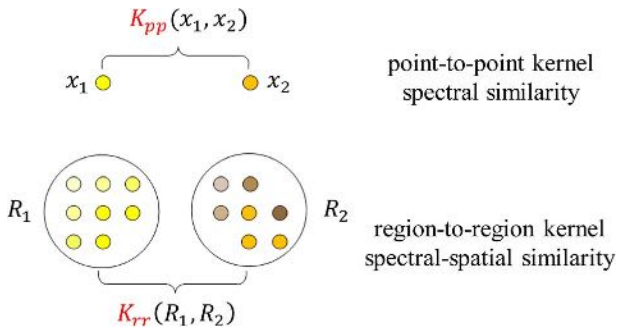


Fig. 1. Point-to-point kernel and region-to-region kernel.

point-to-point similarity, we call it the point-to-point kernel (point kernel) in this paper.

The point kernel is a symmetric and positive semidefinite Mercer's kernel and is widely used in the HSI classification [5], [6], [11]. Given a set of HSI training sample points $\mathcal{L} = \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_\ell, y_\ell)\}$, with $\boldsymbol{x}_i \in R^d$, based on the Gaussian RBF, the point kernel between two points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ can be expressed as

$$K_{pp}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2}\right) \tag{1}$$

where σ is the width of the RBF kernel.

For HSIs, each sample point is a spectral pixel or a spectral curve vector. The point-kernel-based methods use a point kernel as defined in (1) to measure the spectral similarity between sample points for classification. However, neglecting the information related to the spatial arrangement of the pixels in the scene [17], the point-kernel-based spectral similarity metric is insufficient to discriminate complex HSI pixels, such as spectrally similar pixels, i.e., pixels from several subclasses of a material (three subclasses, namely, "Soybean-notill," "Soybean-mitill," and "Soybean-clean," of "Soybean" material appeared in the Indian Pines data set). Although these pixels have similar spectral characteristics, their spatial neighborhood structures are usually different [14]. These spatial neighborhood information combined with the spectral information is able to identify subtle critical differences between different pixels.

As shown in Fig. 1, each circled point represents a sample (spectral vector); two spectrally similar target pixels $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are quite difficult to be discriminated by the point-kernel-based spectral similarity metric $K_{\rm pp}(\boldsymbol{x}_1, \boldsymbol{x}_2)$. Taking into account spatial neighboring information, we can construct a pixel set or a local region $\boldsymbol{R}_i$ for each pixel $\boldsymbol{x}_i$. The region $\boldsymbol{R}_i$ contains the pixel point $\boldsymbol{x}_i$ and its spatial similar or connected neighbors. If we can measure the spectral–spatial kernel similarity metric between two regions $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$, that is, construct a region-to-region kernel (region kernel) $K_{\rm rr}(\boldsymbol{R}_1, \boldsymbol{R}_2)$, then the spectrally similar pixels $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are more likely to be correctly classified. The region kernel is embedded with the spatially structural similarity of spectrally similar target pixels $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and is usually more accurate than the point kernel when the regions are homogeneous.

In order to measure the region-to-region kernel similarity, we propose a new RKSVM classification system, as shown in Fig. 2. It aims to perform kernel-based classification on the labeled pixel points and the labeled regions of HSIs. Each HSI sample is a pixel point and can be extended to a pixel region (or spatial neighborhood), which contains the sample itself and its spatial neighbors. The point-to-point kernel (point kernel) is applied to measure the spectral similarity between different pixel points. The region-to-region kernel (region kernel) and the cross point-to-region kernel (cross kernel) are designed to measure the similarity between different pixel regions and between pixel points and regions, respectively. Then, different combinations of the point kernel, the cross kernel, and the region kernel are used in the SVM classification.

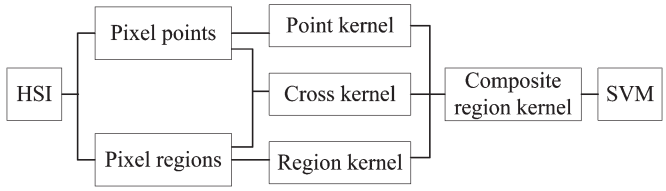


Fig. 2. Flowchart of the proposed RKSVM classification system.

From the flowchart in Fig. 2, we can see that the key issues of the RKSVM classification are how to construct the pixel regions, how to define the region kernels, and how to design an appropriate composite region kernel framework.

A. Spatial Pixel Region

In order to discriminate from the pixel point, we use the name "pixel region." The pixel region is actually a spatial neighborhood. The commonly used spatial neighborhoods are squared neighborhood [11] and adaptive morphological neighborhood [14], [23]. In the fixed-shape squared neighborhood, there usually are pixels from different structures (e.g., noise and background) or/and from different materials. In order to eliminate those interrupt pixels and make the local region more spectrally consistent, we delete the pixels in the squared neighborhood that have large spectral distances to the central pixel, resulting in the distance similarity region. For the construction of adaptive morphological neighborhood, an area filtering method proposed in [14] and [23] is exploited to extract the adaptively connected flat zones.

1) Distance Similarity Region: For sample $\boldsymbol{x}_i$, the squared neighborhood centered at $\boldsymbol{x}_i$ is

$$N(\boldsymbol{x}_i) = \left\{\boldsymbol{x} | \boldsymbol{x} \triangleq (p, q) \in [p_i - a, p_i + a] \times [q_i - a, q_i + a]\right\} \tag{2}$$

where (p_i, q_i) is the pixel coordinate of sample $\boldsymbol{x}_i$, $a = (\omega - 1)/2$, and the odd number ω is the width of the squared window. The size of the squared neighborhood for each pixel is the same, i.e., $\omega \times \omega$. The pixels in the squared neighborhood $N(\boldsymbol{x}_i)$ are denoted by $\boldsymbol{x}_i, \boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{is}$, where $s = \omega^2 - 1$ is the number of neighbors of $\boldsymbol{x}_i$.

The fixed-shape squared neighborhoods may overlap and may contain background pixels or pixels from different structures, particularly for pixels on the boundary of a material. In this case, the neighborhood cannot faithfully reflect the

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

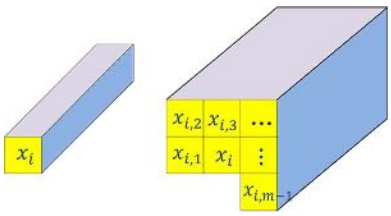


Fig. 3. (Left) Pixel point and (right) corresponding region.

interpixel correlations. In order to eliminate the effect of noisy or inhomogeneous pixels, we design a distance similarity region by choosing the most similar pixels in the squared neighborhood. In detail, in the squared neighborhood $N(\boldsymbol{x}_i)$, we compute the distance between the central pixel $\boldsymbol{x}_i$ and its neighboring pixels $\boldsymbol{x}_{ik}$, i.e., $d_k = \|\boldsymbol{x}_i - \boldsymbol{x}_{ik}\|^2$, and select $H-1$ neighboring pixels with the smallest distances. The $H-1$ selected neighboring pixels and the central pixel $\boldsymbol{x}_i$ form the distance similarity region of $\boldsymbol{x}_i$, i.e.,

$$S(\boldsymbol{x}_i) = \{\boldsymbol{x}_i, \hat{\boldsymbol{x}}_{i,1}, \ldots, \hat{\boldsymbol{x}}_{i,H-1}\}. \tag{3}$$

2) Adaptive Morphological Region: The adaptive morphological region is obtained using the self-complementary flat zone area filters proposed in [14], [23], and [27]. A flat zone is a connected region with constant gray level [28]. The flat zone filter removes all flat zones whose areas are smaller than a given parameter λ. The area is defined as the number of pixels in the flat zone. After area filtering, an image is partitioned into many marked flat zones. For a given pixel $\boldsymbol{x}_i$, the corresponding flat zone is called the morphological region $M(\boldsymbol{x}_i)$ [14], i.e.,

$$M(\boldsymbol{x}_i) = \left\{\boldsymbol{x} | \boldsymbol{x} \in \text{Zone}_{m(\boldsymbol{x}_i)}\right\} \tag{4}$$

where $m(\boldsymbol{x}_i)$ is the mark of the flat zone containing $\boldsymbol{x}_i$, and $\text{Zone}_{m(\boldsymbol{x}_i)}$ represents the corresponding flat zone.

As proposed in [14] and [23], the area filtering is performed on the first principal component image of HSIs, and then, the morphological neighborhood mask extracted from the first principal component image is used to cover HSIs on each band to obtain the final adaptive morphological region.

From the aforementioned two kinds of regions, we can find the following: 1) by deleting dissimilar pixels from a regular squared neighborhood, the distance similarity region can be an irregular shape, as shown in Fig. 3; and 2) by extracting connected neighbors, the morphological region also can be an arbitrary shape.

B. Box

Now, we construct and extract a box feature from a region.

A box can be represented by multidimensional intervals in d-dimensional input space and can be represented as d-ary Cartesian product as follows:

$$\begin{aligned}\boldsymbol{B} &= \left\{\boldsymbol{x} \in R^d : x^z \in [a^z, b^z], z = 1, \ldots, d\right\} \\ &\triangleq [a^1, b^1] \times [a^2, b^2] \times \cdots \times [a^d, b^d]\end{aligned}$$

where d is the number of dimensions or spectral bands, and a^z, b^z are the lower and upper bounds of the interval in the zth spectral band. If $d = 2$, the box $\boldsymbol{B}$ is a rectangle in the plane, which contains all the points located in the rectangle. If $d = 3$, the box $\boldsymbol{B}$ is a cube in the 3-D space. For $d > 3$, the box $\boldsymbol{B}$ is a hyperrectangle.

To construct a box, it needs to determine only the lower and upper bounds of the interval in each band. We first give the following notations. Each pixel $\boldsymbol{x}_i$ is a d-band spectral vector, i.e., $\boldsymbol{x}_i = [x_i^1, \ldots, x_i^d]^{\mathrm{T}}$. For a target pixel $\boldsymbol{x}_i$, a corresponding region $\boldsymbol{R}_i$ has been constructed in Section II-A. Assume that $\boldsymbol{R}_i$ contains m pixels, i.e., $\boldsymbol{R}_i = \{\boldsymbol{x}_i, \boldsymbol{x}_{i,1}, \ldots, \boldsymbol{x}_{i,m-1}\}$. Fixed a band z, the pixels in the region $\boldsymbol{R}_i$ reduce to a vector, which is denoted by $\boldsymbol{v}_i = \boldsymbol{R}_i^z = [x_i^z, x_{i,1}^z, \ldots, x_{i,m-1}^z]^{\mathrm{T}}$. A simple way to define an interval for the band z is to set the lower and upper bounds of the interval as the minimum and maximum numbers in $\boldsymbol{v}_i$. However, this extreme interval fails to reflect the variations of most samples and may enlarge the region.

In order to eliminate the abrupt points and to obtain a mild interval, we set the lower bound a_i^z and the upper bound b_i^z of the interval in band z as the ℓth $(0 < \ell < 50)$ and uth $(50 < u < 100)$ percentiles of the values in the vector $\boldsymbol{v}_i$, respectively, i.e., $a_i^z = \text{prctile}(\boldsymbol{v}_i, \ell)$ and $b_i^z = \text{prctile}(\boldsymbol{v}_i, u)$. For example, if $\boldsymbol{v}_i = [1, 4, 4, 5, 5, 6, 7, 7, 10]^{\mathrm{T}}$ with the target pixel in the fifth position, let $\ell = 25$ and $u = 75$, it obtains the lower bound $a_i^z = \text{prctile}(\boldsymbol{v}_i, 25) = 4$, the upper bound $b_i^z = \text{prctile}(\boldsymbol{v}_i, 75) = 7$, and the interval $[a_i^z, b_i^z] = [4, 7]$. From this example, we can see the following:

1) The interval can reflect the variation range of spectral values of the homogeneous pixels in the region, i.e., [4, 7] reflects the variations of seven values [4, 4, 5, 5, 6, 7, 7]. Compared with SVM-CK that uses a mean to represent these values, i.e., $\text{mean}(\boldsymbol{v}_i) = 5.44$, the interval contains more information.
2) The percentile-based interval selection method can eliminate abrupt points, i.e., 1 and 10 in $\boldsymbol{v}_i$, avoiding extreme interval [1, 10]. In the construction of region, we have used two alternative methods to eliminate the interrupts, i.e., deleting dissimilar pixels from a fixed-shape squared neighborhood or choosing adaptively connected neighboring pixels in a morphological neighborhood. The percentile-based method can be considered as a complementary to further eliminate the interrupts or potential abrupt points.
3) The interval based on the percentiles is around the target point, i.e., [4, 7] containing the target point 5. Due to the local homogeneous distribution of an HSI, the target pixel and its neighboring pixels show similar spectral characteristics. Our interval feature is always around the target point and can reflect this neighborhood similarity. That is, the interval will not be deviated from the target point far away. In the extreme case that the lower and upper bounds are set as the mean, the interval is reduced to a mean point, and our method is reduced to SVM-CK, whose reasonability is validated in [11].

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

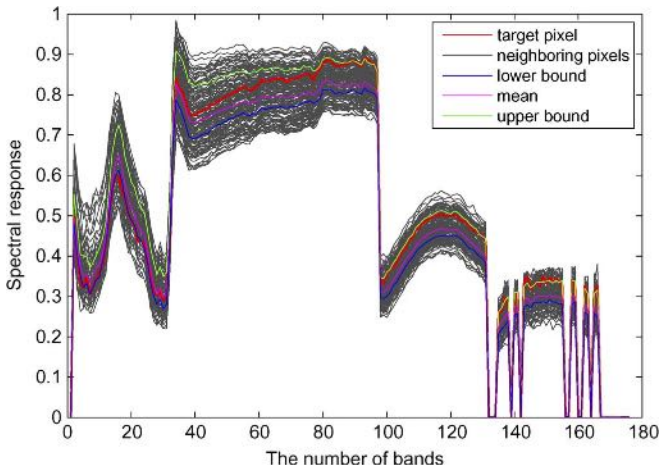


Fig. 4. Pixels in a spatial region: target pixel (red), neighboring pixels of the target pixel (gray), lower bound vector (blue), upper bound vector (green), and mean vector (magenta).

The above conclusion can also be seen from a real example based on the KSC data (see Section III-A) in Fig. 4. It shows the pixels in a distance similarity region and the extracted interval features. For a target pixel in the red color, we have constructed a local region containing the target pixels and its spatial neighbors in the gray color. Based on the percentiles of values in each band, we extract a lower bound $(l = 25)$ and an upper bound vector $(u = 75)$ shown in the blue and green colors, respectively. The mean vector of pixels in the region is shown in the magenta color. In general, the extracted interval bounds most of the central pixels, including the target pixel and mean pixel.

Considering all bands $1 \le z \le d$, we obtain a lower bound vector $\boldsymbol{a}_i = [a_i^1, \ldots, a_i^d]^{\mathrm{T}}$ and an upper bound vector $\boldsymbol{b}_i = [b_i^1, \ldots, b_i^d]^{\mathrm{T}}$ of the region $\boldsymbol{R}_i$. Based on these multidimensional lower and upper bounds, the box is expressed as

$$\begin{aligned}\boldsymbol{B}_i &= \left\{\boldsymbol{x} = (x^1, x^2, \ldots, x^d)^{\mathrm{T}}, x^z \in [a_i^z, b_i^z]\right\} \\ &\triangleq \left[a_i^1, b_i^1\right] \times \left[a_i^2, b_i^2\right] \times \cdots \times \left[a_i^d, b_i^d\right]. \end{aligned} \tag{5}$$

C. Box Kernel

For each region, we have constructed a box feature in Section II-B. Now, we use a box kernel to measure the distance similarity between different boxes. Recall that the general point kernel, such as the Gaussian RBF kernel defined in (1), measures the similarity between sample points; the box kernel extends the point kernel from points to boxes. It can be computed by convolving the point kernel with the characteristic function of the box [26].

The box-to-point kernel K_{bp} and box-to-box kernel K_{bb} are computed as follows [26]:

$$K_{\mathrm{bp}}(\boldsymbol{B}_j, \boldsymbol{x}) = \frac{1}{\nu_j} \prod_{i=1}^{d} \frac{\sqrt{2\pi}\sigma}{2} \left(h\left(\frac{x^i - b_j^i}{\sqrt{2}\sigma}\right) - h\left(\frac{x^i - a_j^i}{\sqrt{2}\sigma}\right)\right) \tag{6}$$

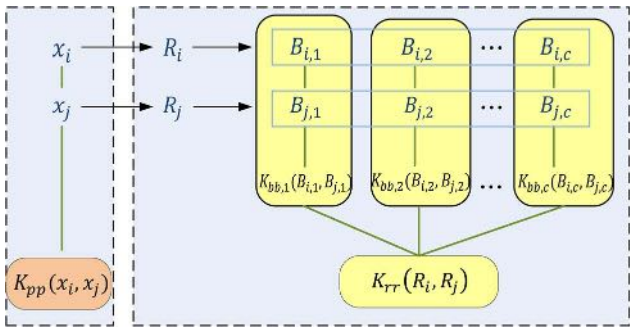


Fig. 5. Flowchart of the (left) point kernel and the (right) region kernel.

$$\begin{aligned}K_{\mathrm{bb}}(\boldsymbol{B}_p, \boldsymbol{B}_q) = \frac{(\sqrt{\pi}\sigma^2)^d}{\nu_p \nu_q} \prod_{i=1}^{d} \big(&\phi\left(b_p^i, b_q^i\right) - \phi\left(a_p^i, b_q^i\right) \\ &- \phi\left(b_p^i, a_q^i\right) + \phi\left(a_p^i, a_q^i\right)\big) \end{aligned} \tag{7}$$

where ν_j is the volume of the box $\boldsymbol{B}_j$, and

$$\phi(a, b) := t \cdot h(t) - \frac{1}{\sqrt{\pi}} e^{-t^2}, \qquad t = \frac{a - b}{\sqrt{2}\sigma}$$

and h is the complementary error function [26].

D. Region Kernel

For each pixel point $\boldsymbol{x}_i$, we have constructed a pixel region $\boldsymbol{R}_i$ (distance similarity region $S(\boldsymbol{x}_i)$ or morphological region $M(\boldsymbol{x}_i)$) in Section II-A and extracted a box feature $\boldsymbol{B}_i$ from the local region $\boldsymbol{R}_i$ and provided a box kernel to measure the similarity between different boxes in Sections II-B and C, respectively.

Now, we need to measure the distance similarity between different pixel regions. For this purpose, we first extract a set of multiscale box features from each region, then compute the box kernels between pairwise boxes from different regions, and finally design a region kernel to measure the similarity between two regions as a weighted average of pairwise box-to-box kernels of two regions. The flowchart of the region kernel is shown in Fig. 5, where the left part is the point kernel and the right part is the region kernel.

As shown in Section II-B, for a pixel region, by setting different lower and upper thresholds ℓ_k and u_k, $k = 1, \ldots, c$, we can obtain different scales of boxes $\boldsymbol{B}_{i,k}$ of a region. Assume that each of two regions $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ can be represented as c boxes, namely, $\boldsymbol{B}_{i,1}, \ldots, \boldsymbol{B}_{i,c}$ and $\boldsymbol{B}_{j,1}, \ldots, \boldsymbol{B}_{j,c}$, respectively. The boxes in two regions with the same scale are compared and combined as a box kernel. The region kernel between regions $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ is defined as the linear combination of the pairwise box-to-box kernels as follows:

$$K_{\mathrm{rr}}(\boldsymbol{R}_i, \boldsymbol{R}_j) = \sum_{k=1}^{c} \beta_k K_{\mathrm{bb},k}(\boldsymbol{B}_{i,k}, \boldsymbol{B}_{j,k}) \tag{8}$$

where $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_c]^{\mathrm{T}}$ is a linear combination coefficient vector, which measures the contribution of different box kernels

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

$K_{\mathrm{bb},k},\ k = 1, \ldots, c$. Because the pairwise box-to-box kernels in the same scale are considered, no matter in the sequential processing or the random processing of the given c scales, the region kernel is the same.

Next, we determine the coefficient vector $\boldsymbol{\beta}$ based on the kernel alignment [29]–[31]. The kernel alignment measures the similarity between two kernels K_1 and K_2 [29], i.e.,

$$A(K_1, K_2) = \frac{\langle K_1, K_2 \rangle_F}{\sqrt{\langle K_1, K_1 \rangle_F \langle K_2, K_2 \rangle_F}} \tag{9}$$

where $\langle \cdot, \cdot \rangle_F$ is the Frobenius norm, and it can be used to choose the optimal kernel in a set of candidate kernels. Given a benchmark kernel, the importance of each candidate kernel is evaluated by comparing its kernel alignment value with the benchmark kernel. Here, the benchmark kernel is chosen as the ideal kernel [29], [32], which is defined as

$$K_0(\boldsymbol{x}_i, \boldsymbol{x}_j) = \begin{cases} 1, & y_i = y_j \\ 0, & y_i \neq y_j. \end{cases} \tag{10}$$

The ideal kernel leads to a perfect classification inspired from an "oracle," two samples $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ should be considered as "similar" (with kernel value 1) if and only if they belong to the same class $(y_i = y_j)$ [29], [32].

Based on the ideal kernel K_0, the importance of each candidate box kernel $K_{\mathrm{bb},k}$ is assessed by the kernel alignment value $A(K_0, K_{\mathrm{bb},k})$. The corresponding coefficient β_k of box kernel $K_{\mathrm{bb},k}$ in (8) is now computed as

$$\beta_k = \frac{A(K_0, K_{\mathrm{bb},k})}{\sum_{k=1}^{c} A(K_0, K_{\mathrm{bb},k})}. \tag{11}$$

Similarly, when a region is degenerated to a point in Fig. 5, the region-to-point cross kernel between a region $\boldsymbol{R}_j$ and a point $\boldsymbol{x}_i$ is defined as

$$K_{\mathrm{rp}}(\boldsymbol{R}_j, \boldsymbol{x}_i) = \sum_{k=1}^{c} \beta'_k K_{\mathrm{bp},k}(\boldsymbol{B}_{j,k}, \boldsymbol{x}_i) \tag{12}$$

where β'_k is the normalized kernel alignment value between $K_{\mathrm{bp},k}$ and K_0.

Now, we have obtained the point kernel, the region kernel, and the cross kernel in (1), (8), and (12), respectively. It can be seen that the point kernel operates on pixel points and can be computed directly according to the general kernel function, such as the commonly used Gaussian function in (1), whereas the region kernel operates on pixel regions and is relatively difficult to compute, where complexity convolution calculations between the point kernel and characteristic functions of multiscale boxes are needed, as shown in (6) and (7). In spite of different operation objects and computing methods, the region kernel and the point kernel have the same form and meaning if we consider the region as a point. In particular, when the regions are reduced to points, the region kernel is the point kernel.

Now, we give some prosperities of the region kernel.

Proposition 1: The region kernels K_{rr} and K_{rp} are symmetric and positive definite kernels.

Proof: According to [26, Proposition 4.1], the box kernels $K_{\mathrm{bb},k}$ and $K_{\mathrm{bp},k}$ are symmetric and positive definite kernels. Thus, the region kernel as a linear combination of finite box kernels is also a symmetric and positive definite kernel.

Proposition 2: If K_{pp} is a point kernel and K_{rr} is a region kernel, then $K_{\mathrm{pp}} + K_{\mathrm{rr}}$ is a valid kernel.

Proof: Denote $K = K_{\mathrm{pp}} + K_{\mathrm{rr}}$; the elements of kernel K can be represented as

$$K\left((\boldsymbol{x}_i, \boldsymbol{R}_i), (\boldsymbol{x}_j, \boldsymbol{R}_j)\right) = K_{\mathrm{pp}}(\boldsymbol{x}_i, \boldsymbol{x}_j) + K_{\mathrm{rr}}(\boldsymbol{R}_i, \boldsymbol{R}_j). \tag{13}$$

Based on the symmetry of K_{pp} and K_{rr}, it is easy to validate the symmetry of the new kernel K as follows:

$$K\left((\boldsymbol{x}_i, \boldsymbol{R}_i), (\boldsymbol{x}_j, \boldsymbol{R}_j)\right) = K\left((\boldsymbol{x}_j, \boldsymbol{R}_j), (\boldsymbol{x}_i, \boldsymbol{R}_i)\right). \tag{14}$$

Because K_{pp} and K_{rr} are positive definite kernels, the kernel K is positive definite. Therefore, K is a valid kernel.

Proposition 3: Given a point kernel K_{pp} and region kernels K_{rr} and K_{rp}. If $K_{\mathrm{rr}} - K_{\mathrm{rp}} K_{\mathrm{pp}}^{-1} K_{\mathrm{pr}}$ or $K_{\mathrm{pp}} - K_{\mathrm{pr}} K_{\mathrm{rr}}^{-1} K_{\mathrm{rp}}$ is positive semidefinite, then the following function is a valid kernel:

$$K = \begin{bmatrix} K_{\mathrm{pp}} & K_{\mathrm{pr}} \\ K_{\mathrm{rp}} & K_{\mathrm{rr}} \end{bmatrix}. \tag{15}$$

Proof: Because K_{pp}, K_{rp}, and K_{rr} are symmetric matrices, the stack matrix K is symmetric. Because K_{pp} and K_{rr} are positive definite kernels, based on the Schur complement condition for positive definiteness [33], if $K_{\mathrm{rr}} - K_{\mathrm{rp}} K_{\mathrm{pp}}^{-1} K_{\mathrm{pr}}$ or $K_{\mathrm{pp}} - K_{\mathrm{pr}} K_{\mathrm{rr}}^{-1} K_{\mathrm{rp}}$ is positive semidefinite, then the stack kernel K is positive semidefinite.

E. Composite Region Kernel Framework

The composite region kernel method deals with both the labeled points and the labeled regions. Each labeled pixel $\boldsymbol{x}_i$ can be extended to a region $\boldsymbol{R}_i$. The region $\boldsymbol{R}_i$ contains the pixel $\boldsymbol{x}_i$ and its spatial neighbors and can represent pixel $\boldsymbol{x}_i$ at a certain extent. The pixel $\boldsymbol{x}_i$ and its label y_i consist of a supervised sample $(\boldsymbol{x}_i, y_i)$. Similarly, the region $\boldsymbol{R}_i$ and label y_i also consist of a supervised sample $(\boldsymbol{R}_i, y_i)$. Given a collection of labeled points and corresponding labeled regions

$$\begin{aligned} \mathcal{L} &= \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_\ell, y_\ell)\} \\ \mathcal{R} &= \{(\boldsymbol{R}_1, y_1), \ldots, (\boldsymbol{R}_\ell, y_\ell)\} \end{aligned} \tag{16}$$

we can compute the point-to-point kernel (point kernel) $K_{\mathrm{pp}}(\mathcal{L}, \mathcal{L})$, the point-to-region cross kernel $K_{\mathrm{pr}}(\mathcal{L}, \mathcal{R})$, and the region-to-region kernel (region kernel) $K_{\mathrm{rr}}(\mathcal{R}, \mathcal{R})$. When the point- and region-based kernels are computed, we can treat the points and regions as the same and implement the composite region kernel classification.

In the following, we propose three composite region kernels.

1) Single Region Kernel: The single-region-kernel-based RKSVM classification framework uses the region kernel alone, i.e.,

$$K = K_{\mathrm{rr}}(\mathcal{R}, \mathcal{R}). \tag{17}$$

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

2) Weighted Summation Region Kernel: Motivated by the general composite point kernel classification framework in SVM-CK [11], we propose a weighted summation region kernel for the RKSVM classification, i.e.,

$$K = \mu K_{\mathrm{rr}}(\mathcal{R}, \mathcal{R}) + (1-\mu) K_{\mathrm{pp}}(\mathcal{L}, \mathcal{L}) \tag{18}$$

where μ is a combination coefficient balancing the pixel-point-based point kernel and the pixel-region-based region kernel.

3) Stack Region Kernel: Considering the cross-information between the region kernel and the point kernel, we further propose a new stack kernel on the 2ℓ labeled pairs in (16). In the training phase, we compute the stack kernel K that contains the point kernel, the cross kernel, and the region kernel as follows:

$$K = \begin{bmatrix} K_{\mathrm{pp}}(\mathcal{L}, \mathcal{L}) & K_{\mathrm{pr}}(\mathcal{L}, \mathcal{R}) \\ K_{\mathrm{rp}}(\mathcal{R}, \mathcal{L}) & K_{\mathrm{rr}}(\mathcal{R}, \mathcal{R}) \end{bmatrix} \tag{19}$$

and then train the SVM model to obtain the coefficient vector $\boldsymbol{\alpha}$ satisfying $\boldsymbol{f} = K\boldsymbol{\alpha}$. In the prediction phase, for each testing sample $\hat{\boldsymbol{x}}$, we compute the corresponding pixel region $\hat{\boldsymbol{R}}$ and stack kernel $\hat{K}$, i.e.,

$$\hat{K} = \begin{bmatrix} K_{\mathrm{rp}}(\hat{\boldsymbol{R}}, \mathcal{L}) & K_{\mathrm{rr}}(\hat{\boldsymbol{R}}, \mathcal{R}) \end{bmatrix} \tag{20}$$

and predict the label using the decision value $\hat{f} = \hat{K}\boldsymbol{\alpha}$. Note that, different from the single region kernel and the weighted summation region kernel, the stack kernel has different forms in the training and testing phases.

Incorporating the spectral training pixel points and the spatial labeled regions, the proposed RKSVM is summarized in Algorithm 1.

Algorithm 1 RKSVM

Input: Training samples $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^{\ell}$, similarity neighborhood scale ω and number H, or morphological area parameter λ

1 Define the spatial pixel region $\boldsymbol{R}_i$ for each sample $\boldsymbol{x}_i$
  1.1 Find the similarity neighborhood region $S(\boldsymbol{x}_i)$ in (3), or
  1.2 Determine the morphological region $M(\boldsymbol{x}_i)$ in (4)
2 Compute the point and region kernels
  2.1 Construct multiscale boxes from the region
  2.2 Compute box kernels using (6) and (7)
  2.3 Solve the coefficient $\boldsymbol{\beta}$ by (11)
  2.4 Obtain the region kernels (8) and (12)
3 Perform SVM classification using:
  3.1 the single region kernel in (17), or
  3.2 the weighted summation region kernel in (18), or
  3.3 the stack region kernel in (19)

Output: The prediction label for each sample

III. Experimental Results

The proposed RKSVM is compared with the spectral, spatial, and spectral–spatial SVMs on three hyperspectral data sets. In the experiments, K^{ω} represents the spectral SVM using the spectral features alone. K^{s} denotes the spatial SVM using only the spatial features represented as the mean of neighboring pixels. $\mu K^{s+\omega}$ refers to the spectral–spatial SVM with weighted summation composite kernel, i.e., $\mu K^{s+\omega} = \mu K^{s} + (1-\mu)K^{\omega}$. Different neighborhood systems can be used in RKSVM. Using the distance similarity neighborhood in (3) and the morphological neighborhood in (4), the proposed RKSVM is called RKSVM-SN and RKSVM-MN, respectively. In the following, K_S^r and K_M^r refer to RKSVM-SN and RKSVM-MN with the single region kernel defined in (17), respectively. $\mu K_S^{r+\omega}$ and $\mu K_M^{r+\omega}$ refer to RKSVM-SN and RKSVM-MN with the weighted summation region kernel defined in (18), respectively, where the combination coefficient μ is set as 0.8. $K_M^{[r,\omega]}$ and $K_S^{[r,\omega]}$ refer to RKSVM-SN and RKSVM-MN with the stack region kernel defined in (19), respectively. In RKSVM-SN and RKSVM-MN, the lower and upper thresholds $\ell \in [25, 30, 35]$ and $u \in [65, 70, 75]$ are used to generate multiscale boxes. For all SVM-based algorithms, the Gaussian kernel is used, and LIBSVM software is used to implement SVM [34]. The classification performance is assessed on the testing set by the overall accuracy (OA), which is the number of correctly classified samples divided by the number of total testing samples, and by the kappa coefficient (κ), which measures the degree of classification agreement.

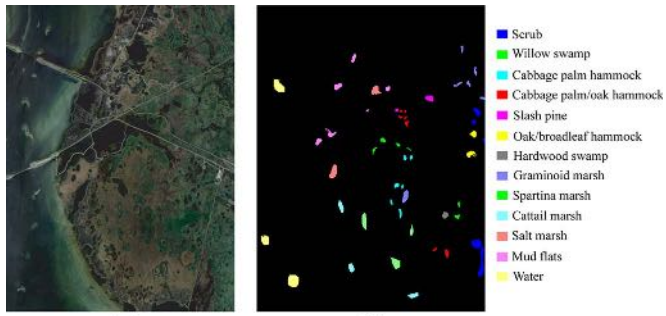


Fig. 6. KSC data set. (a) RGB composite image of bands 31, 21, and 11. (b) Ground-truth map.

A. Hyperspectral Data Sets

Three public HSI data sets are used.[1]

1) Kennedy Space Center: The data set was acquired by the NASA AVIRIS instrument over KSC, Florida, on March 23, 1996 [35]. The image scene has the size of 512 × 614 pixels and 224 spectral channels. After discarding water absorption and noisy bands, 176 bands remain. It contains 13 ground-truth classes. The total number of samples is 5211, ranging from 105 to 927 in each class. The false color composition of bands 31, 21, and 11 and the ground-truth map are shown in Fig. 6.
2) University of Pavia: The data set was acquired in 2001 by the ROSIS instrument over the city of Pavia, Italy [36]. This image scene corresponds to the University of Pavia

[1]Available online: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

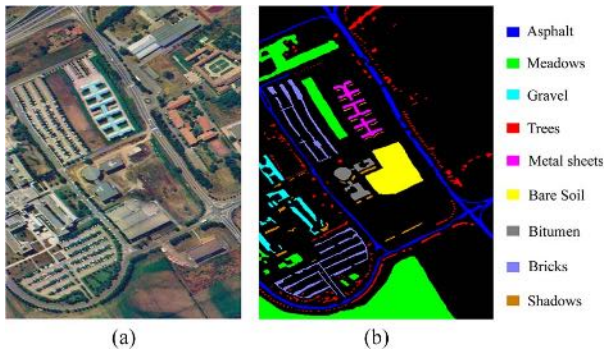


Fig. 7. University of Pavia data set. (a) RGB composite image of bands 60, 30, and 2. (b) Ground-truth map.

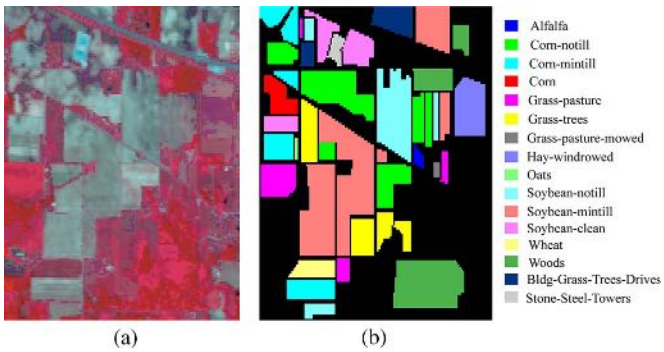


Fig. 8. Indian Pines data set. (a) RGB composite image of bands 50, 27, and 17. (b) Ground-truth map.

and has the size of 610 × 340 pixels and 115 spectral bands. After discarding noisy and water absorption bands, 103 bands are retained. The data contain nine ground-truth classes. The false color composition of bands 60, 30, and 2 and the ground-truth map are shown in Fig. 7.

3) Indian Pines: The data set was acquired by the AVIRIS sensor in 1992. The image scene contains 145 × 145 pixels and 220 spectral bands, where 20 channels were discarded because of the atmospheric affection. There are 16 classes in the data set. The total number of samples is 10 249, ranging from 20 to 2455 in each class. The false color composition of bands 50, 27, and 17 and the ground-truth map are shown in Fig. 8.

B. Parameter Analysis

In RKSVM, there are SVM parameters (regularization parameter C and kernel parameter σ) and spatial region related parameters (neighborhood window ω and the number of similar pixels H in RKSVM-SN or area size λ in RKSVM-MN). We take the KSC data set as an example to investigate the parameters separately. We first investigate the sensitivity of SVM parameters C and σ on the stack-kernel-based RKSVM methods in (19), where C is chosen from the set {1, 10, 100, 1000, 10 000}, and σ varies in the range $\{2^{-4}, 2^{-3}, \ldots, 2^{4}\}$. Fifteen samples in each class are selected as the training set for the parameter selection. We present the classification OAs with different parameter values and choose the optimal parameter values such that OA is maximized. Fig. 9(a) and (b) shows OA as a function of C and σ for the stack-kernel-based RKSVM-SN and RKSVM-MN, respectively. It can be clearly seen that OAs follow a similar trend for both algorithms. OAs are stable when C varies from 100 to 10 000, and σ changes from 2^{-1} to 2^{2}. The proposed RKSVM is not very sensitive to the SVM parameters. In the following experiments for all three data sets, we set $C = 1000$ and $\sigma = 2$. It should also be noted that when σ tends to zero ($\sigma = 2^{-4}$), the SVM classifier is severe underfitting and the entire data space is assigned to the majority classes, which leads to extremely bad results, with OAs only about 20%.

Next, we investigate RKSVM-SN parameters: the neighborhood window ω and the number of similar pixels H. Seven different neighborhood windows, i.e., $3 \times 3, 5 \times 5,$ and 15×15, are considered. The number H in the distance similarity neighborhood $S(\boldsymbol{x}_i)$ in (3) is equal to the number of similar pixels in the square neighborhood $N(\boldsymbol{x}_i)$ in (2). After deleting several dissimilar pixels in $N(\boldsymbol{x}_i)$, the remaining pixels are called similar pixels. Thus, for determining H, it needs to determine only the ratio of deleted pixels to total pixels in $N(\boldsymbol{x}_i)$. Denote the ratio of deleted pixels as η, then $H = (1-\eta) \times \omega^2$. The ratio η is chosen from the set {0, 5%, 10%, 15%, 20%, 30%, 40%, 50%}. Fixed η, we first investigate the neighborhood window ω. Based on obtained window ω, we further investigate the ratio of deleted pixels η. Fig. 10(a) shows the changes of OAs as a function of ω, where the window of 13 × 13 provides the best OA. The OA versus η is shown in Fig. 10(b). By deleting several noisy or inhomogeneous interrupted pixels in the fixed-shape squared neighborhood, the OA increases at a certain extent, as stated in Section II-A1. However, when the ratio of deleted pixels exceeds a certain limit, many useful homogeneous pixels are deleted, and the OA decreases rapidly. The ratio of deleted pixels η corresponding to the peak OA value ($\eta = 10\%$) is used to determine an optimal number of pixels in the distance similarity region.

Now, we show the OA of RKSVM-MN on the area size parameter λ in Fig. 11, where λ is chosen from the set $\{5, 10, 20, \ldots, 100\}$. It can be seen that the OA is insensitive to the area size. In the experiments, the parameter λ is chosen as 20 for the KSC data set.

Similarly, we can find the optimal parameters for the other two data sets. That is, for the University of Pavia data set, $\omega = 7$, $\eta = 15\%$, and $\lambda = 50$; for the Indian Pines data set, $\omega = 11$, $\eta = 40\%$, and $\lambda = 100$.

C. Comparison Results

The effectiveness of the RKSVM methods under different numbers of training samples per class is investigated. We randomly choose $N = 5$, 10, 15, 20, 25, and 30 samples from each class to form the training sets, respectively (For the class less than N samples, half of the total samples are chosen.). The remaining samples consist of the testing sets. In each case, the experiment is repeated ten times with randomly chosen training samples. The results of ten runs are averaged. It should be noted

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

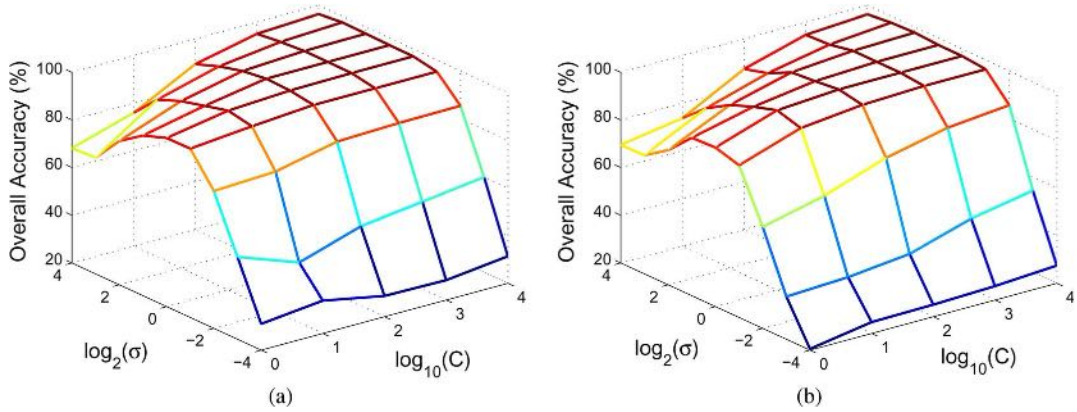


Fig. 9. OA versus SVM parameters C and σ. (a) RKSVM-SN. (b) RKSVM-MN.

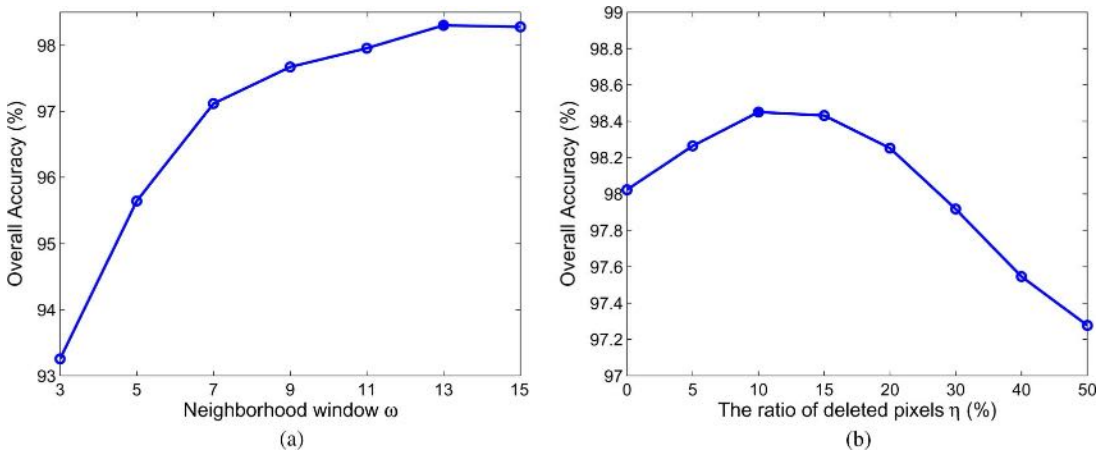


Fig. 10. RKSVM-SN parameters. (a) OA versus the neighborhood window ω. (b) OA versus the ratio of deleted pixels η.

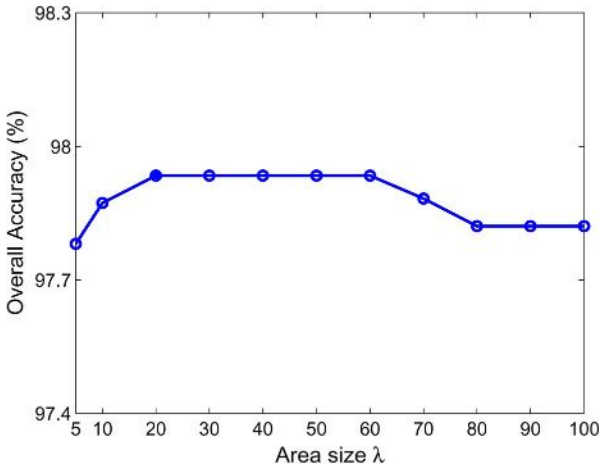


Fig. 11. RKSVM-MN parameter. OA versus the area size λ.

that the samples in different sample sizes are independent and randomly selected.

The classification OAs and kappa coefficients under different numbers of training samples per class are shown in Tables I–III. It should be noted that, for the University of Pavia data set, the morphological region in RKSVM-MN is obtained as the union of morphological neighborhoods corresponding to the first and second principal component images because the second principal component accounts for a considerable contribution. From the tables, we can conclude the following.

1) With the increase of training samples, OAs and kappa coefficients for all algorithms are greatly improved. The proposed composite region kernel methods (RKSVM-SN and RKSVM-MN) show a significant improvement over the point-kernel-based SVM-CK with the single spectral kernel, the single spatial kernel, and the weighted summation spatial–spectral kernel.
2) Compared with SVM-CK, the proposed RKSVM is quite effective in the case with limited training samples, seeing the results in the case of $N = 5$.
3) Comparing the single region kernel K_S^τ in the fourth column (take RKSVM-SN as an example) with the single spectral point kernel K^ω and the spatial point kernel K^s in the first and second columns in the three tables, investigating all the cases ranging from 5 to 30 labeled samples per class for training, the single region kernel K_s^τ improves the single spectral point kernel K^ω by the OA

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

TABLE I
CLASSIFICATION ACCURACIES (%) ON THE KSC DATA SET

N	Point-kernel-based SVM-CK			Region-kernel-based RKSVM-SN			Region-kernel-based RKSVM-MN		
	K^{ω}	K^{s}	$\mu K^{s+\omega}$	K_S^r	$\mu K_S^{r+\omega}$	$K_S^{[r,\omega]}$	K_M^r	$\mu K_M^{r+\omega}$	$K_M^{[r,\omega]}$
5	78.94±2.11	87.00±1.74	87.39±1.54	89.92±1.98	89.96±1.87	**91.79±2.70**	89.59±1.34	89.64±1.34	91.29±2.02
	76.95±2.34	85.53±1.95	85.97±1.72	88.78±2.21	88.82±2.08	**90.86±3.01**	88.42±1.50	88.47±1.50	90.30±2.25
10	84.51±3.16	93.81±1.24	93.88±1.48	96.34±1.11	96.39±1.11	**97.03±1.11**	95.92±1.39	95.91±1.31	96.44±1.05
	82.77±3.50	93.10±1.38	93.19±1.64	95.92±1.23	95.98±1.24	**96.69±1.24**	95.46±1.55	95.44±1.46	96.03±1.17
15	88.25±1.17	96.13±1.12	96.59±0.86	97.99±0.79	98.02±0.74	**98.43±0.66**	97.30±0.77	97.46±0.65	97.92±0.83
	86.91±1.29	95.68±1.25	96.19±0.96	97.76±0.89	97.79±0.82	**98.25±0.73**	96.99±0.86	97.17±0.72	97.68±0.92
20	89.44±1.51	97.35±0.61	97.36±0.71	98.60±0.66	98.55±0.73	**98.74±0.67**	98.02±0.63	98.08±0.69	98.66±0.57
	88.24±1.67	97.04±0.68	97.06±0.79	98.43±0.74	98.39±0.82	**98.60±0.75**	97.79±0.70	97.86±0.77	98.51±0.64
25	90.91±1.77	98.41±0.62	98.36±0.43	**99.38±0.40**	99.33±0.40	99.35±0.44	99.07±0.37	99.06±0.37	99.23±0.34
	89.85±1.97	98.23±0.69	98.17±0.48	**99.31±0.44**	99.25±0.45	99.27±0.50	98.97±0.42	98.95±0.41	99.14±0.38
30	91.45±1.65	98.61±0.58	98.70±0.46	99.60±0.20	99.55±0.20	**99.65±0.15**	99.21±0.41	99.18±0.24	99.46±0.28
	90.45±1.83	98.44±0.64	98.55±0.52	99.55±0.22	99.50±0.22	**99.61±0.17**	99.12±0.28	99.09±0.27	99.40±0.31

of nearly 8%–12% for the KSC data set, 14%–17% for the University of Pavia data set, and 22%–25% for the Indian Pines data set and improves the single spatial point kernel K^s by the OA of 1%–3%, 5%–7%, and 5%–10% for the three data sets, respectively. In a word, the region kernel dramatically improves the spectral and spatial point kernels. This demonstrates that the region in RKSVM can capture local pixel variations more effectively than the spectral vector or the mean vector used in the traditional pixel-point-based SVM-CK methods, and the region-to-region similarity measured by the region kernel is more accurate than the point-to-point similarity.

4) Comparing the weighted summation region kernel $\mu K_S^{r+\omega}$ in the fifth column with the single region kernel K_S^r in the fourth column, $\mu K_S^{r+\omega}$ almost provides same results with K_S^r on the three data sets. It demonstrates that a simple linear combination of the region kernel and the point kernel is ineffective.
5) Comparing the stack region kernel $K_S^{[r,\omega]}$ in the sixth column with the single region kernel K_S^r in the fourth column, $K_S^{[r,\omega]}$ improves K_S^r at a certain extent. In particular, in the case of five labeled samples, stack kernel $K_S^{[r,\omega]}$ improves single region kernel K_S^r by 1.8%, 2%, and 1.7% on the three data sets, respectively. It shows that the stack composite way is more effective than the weighted summation composite way, and the cross point-to-region kernel used in the stack kernel plays a role in increasing the accuracy.
6) Comparing the region kernel in the second three columns and the point kernel in the first three columns, it can be seen that the improvements of classification performance are achieved by using the region kernel and the stack composite classification framework, where the substantial improvements rely on the use of the region kernel.
7) Among the three different kinds of region kernels, the proposed stack region kernel ($K_M^{[r,\omega]}$ or $K_S^{[r,\omega]}$) using additional cross-information (point-to-region kernel) provides the best overall results. For the KSC and University of Pavia data sets, the similarity-neighborhood-based RKSVM-SN methods show slightly better results than the morphological-neighborhood-based RKSVM-MN methods. For the Indian Pines data set, the classification results of RKSVM-MN are much better than those of RKSVM-SN.

Figs. 12 and 13 show the classification maps for the University of Pavia and Indian Pines data sets in the case of 30 labeled samples per class, respectively. The maps correspond to the classification results using different pixel-point-based composite kernels and pixel-region-based composite kernels. It can be seen that the RKSVM methods show relatively better results than the SVM-CK methods in terms of consistent classification results with little noise.

D. Discussion

In the proposed RKSVM, we should note the following.

1) In RKSVM-SN, the distance similarity region is obtained from the fixed-window squared neighborhood by discarding outliers. The selected region can reflect spectral variations of local homogeneous pixels at a certain extent. This can be further improved by using multiwindow (different ω) neighborhood systems and fusing multiple complementary information in different windows for handling complex HSI structures (different sizes and shapes). In RKSVM-MN, the area filtering based morphological region can adaptively catch the connected neighboring pixels belonging to the same structure (flat zone). This neighborhood is local homogeneous and spectrally consistent [14].
2) Using a region and a region kernel to model the local homogeneous pixel relations is feasible. The region is used to capture spatial local homogeneous pixels. The only requirement is that the pixels in the region should have the same label. Due to the homogeneous distribution of an HSI, the local pixel patch or region usually contains the same material. This special characteristic ensures the effectiveness of the region model. Moreover, the region is obtained from the similarity neighborhood or the adaptive morphological neighborhood, in which the homogeneous pixels are kept whereas interrupt pixels are removed at a certain extent. Thus, using a region to capture the local homogeneous pixels is feasible.

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

TABLE II
CLASSIFICATION ACCURACIES (%) ON THE UNIVERSITY OF PAVIA DATA SET

N	Point-kernel-based SVM-CK			Region-kernel-based RKSVM-SN			Region-kernel-based RKSVM-MN		
	K^{ω}	K^{s}	$\mu K^{s+\omega}$	K_S^r	$\mu K_S^{r+\omega}$	$K_S^{[r,\omega]}$	K_M^r	$\mu K_M^{r+\omega}$	$K_M^{[r,\omega]}$
5	58.33±9.56	65.44±4.71	69.10±4.63	72.66±4.10	72.95±4.07	**74.62±6.76**	71.42±5.80	71.52±5.78	73.37±6.31
	49.94±9.04	56.92±5.58	61.45±5.12	65.90±4.70	66.23±4.68	**68.37±7.99**	64.30±6.44	64.42±6.44	66.69±7.07
10	67.04±4.04	76.23±5.06	77.99±5.32	81.54±4.84	81.93±4.87	**82.75±5.05**	80.85±5.64	80.92±5.64	80.65±5.07
	59.26±4.35	70.12±5.70	72.27±6.00	76.71±5.66	77.18±5.69	**78.19±6.00**	75.77±6.59	75.87±6.58	75.55±5.91
15	72.86±4.69	83.69±3.15	85.47±3.50	89.58±2.63	**89.89±2.88**	89.46±2.84	88.74±2.64	89.07±2.69	88.95±2.74
	65.58±5.57	78.93±3.83	81.19±4.33	86.48±3.24	**86.87±3.57**	86.37±3.50	85.37±3.23	85.79±3.30	85.69±3.35
20	76.57±4.16	85.74±3.67	89.08±1.47	92.67±1.70	92.94±1.48	**93.09±1.55**	91.81±2.43	91.86±2.46	92.86±2.10
	70.30±4.67	81.56±4.50	85.72±1.84	90.39±2.14	90.73±1.87	**90.94±1.94**	89.27±3.03	89.34±3.06	90.64±2.62
25	78.62±2.50	88.72±2.66	90.68±2.15	93.40±1.88	93.44±1.81	**93.49±1.54**	91.97±1.02	92.06±1.06	92.25±0.96
	72.64±2.98	85.31±3.34	87.80±2.67	91.33±2.39	91.39±2.30	**91.46±1.96**	89.45±1.29	89.57±1.34	89.85±1.21
30	78.88±3.69	90.42±1.76	91.94±1.33	95.20±1.28	**95.40±1.19**	94.63±1.70	93.56±1.96	93.50±1.96	93.60±2.03
	73.06±4.43	87.43±2.21	89.40±1.71	93.66±1.67	**93.92±1.54**	92.95±2.19	91.52±2.54	91.44±2.55	91.61±2.61

TABLE III
CLASSIFICATION ACCURACIES (%) ON THE INDIAN PINES DATA SET

N	Point-kernel-based SVM-CK			Region-kernel-based RKSVM-SN			Region-kernel-based RKSVM-MN		
	K^{ω}	K^{s}	$\mu K^{s+\omega}$	K_S^r	$\mu K_S^{r+\omega}$	$K_S^{[r,\omega]}$	K_M^r	$\mu K_M^{r+\omega}$	$K_M^{[r,\omega]}$
5	47.19±2.98	59.81±4.12	61.67±2.37	69.30±2.05	69.37±1.92	70.98±1.92	72.04±3.68	72.45±3.24	**73.67±3.41**
	41.14±3.14	55.23±4.60	57.28±2.64	65.64±2.27	65.70±2.11	67.52±2.11	68.71±3.97	69.11±3.54	**70.49±3.71**
10	54.16±3.29	70.74±2.74	72.44±2.96	78.92±2.51	78.46±2.77	80.00±2.83	79.98±2.63	80.23±2.82	**81.88±2.92**
	48.82±3.75	67.19±3.04	69.10±3.22	76.24±2.80	75.74±3.06	77.49±3.16	77.47±2.90	77.72±3.10	**79.60±3.24**
15	60.22±1.61	77.26±2.47	79.68±2.39	85.24±3.11	85.18±3.02	86.04±3.21	87.60±2.65	87.52±2.55	**88.73±2.38**
	55.28±1.64	74.38±2.71	77.08±2.62	83.28±3.42	83.22±3.31	84.19±3.54	85.94±2.87	85.85±2.86	**87.21±2.68**
20	63.69±1.85	82.05±2.33	83.73±2.38	87.50±2.71	87.52±2.55	88.15±2.27	88.67±2.25	88.79±2.10	**89.83±1.94**
	59.19±2.02	79.74±2.58	81.62±2.63	85.84±3.02	85.87±2.84	86.58±2.53	87.15±2.51	87.28±2.34	**88.45±2.18**
25	65.59±1.69	83.60±1.75	85.83±1.55	89.54±1.27	89.56±1.30	90.57±1.22	91.35±2.16	91.56±2.10	**92.17±1.73**
	61.28±1.86	81.44±1.97	83.96±1.73	88.11±1.43	88.14±1.46	89.27±1.38	90.16±2.44	90.39±2.37	**91.08±1.95**
30	67.75±1.61	85.97±1.53	87.59±1.22	90.64±1.32	90.61±1.31	91.13±1.33	91.89±1.60	92.11±1.55	**92.55±1.69**
	63.64±1.79	84.12±1.70	85.93±1.36	89.35±1.48	89.32±1.47	89.91±1.50	90.77±1.80	91.02±1.75	**91.51±1.91**

TABLE IV
WITHIN-CLASS AND BETWEEN-CLASS SIMILARITIES BETWEEN DIFFERENT SAMPLES FROM THE KSC DATA SET

		Within-class similarity					Between-class similarity				
Case 1	PPSM	0.633	**0.880**	0.550	0.655	0.709	0.484	0.407	0.023	0.086	0.110
	RRSM	0.936	0.961	0.935	**0.970**	0.942	0.294	0.237	0.015	0.059	0.128
Case 2	PPSM	0.633	0.880	0.550	0.655	0.709	**0.906**	0.663	0.879	0.869	**0.906**
	RRSM	0.936	0.961	0.935	**0.970**	0.942	0.572	0.260	0.481	0.723	0.478

TABLE V
CLASSIFICATION OAS OF RKSVM IN SINGLE-SCALE BOXES AND MULTISCALE BOX IN THE CASE OF 15 LABELED SAMPLES PER CLASS FOR TRAINING (THE FIRST AND SECOND ROWS FOR EACH DATA SET CORRESPOND TO THE OAS OF THE SINGLE-REGION-KERNEL-BASED RKSVM-SN AND RKSVM-MN, RESPECTIVELY.)

	Single-scale									Multi-scale
KSC	97.46	97.69	97.88	97.68	97.89	**98.12**	97.52	97.68	97.98	97.99
	96.82	97.05	97.42	96.84	97.03	**97.43**	96.80	97.07	**97.43**	97.30
PaviaU	87.34	87.07	88.21	87.66	87.42	87.60	87.90	87.60	87.47	**89.58**
	85.76	85.82	85.88	85.94	85.85	85.85	85.92	85.83	86.20	**87.21**
Indian Pines	84.30	83.67	83.30	84.88	84.12	83.78	85.10	84.70	84.13	**85.24**
	86.80	86.73	86.71	86.82	86.82	87.13	87.23	87.14	87.07	**87.60**

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

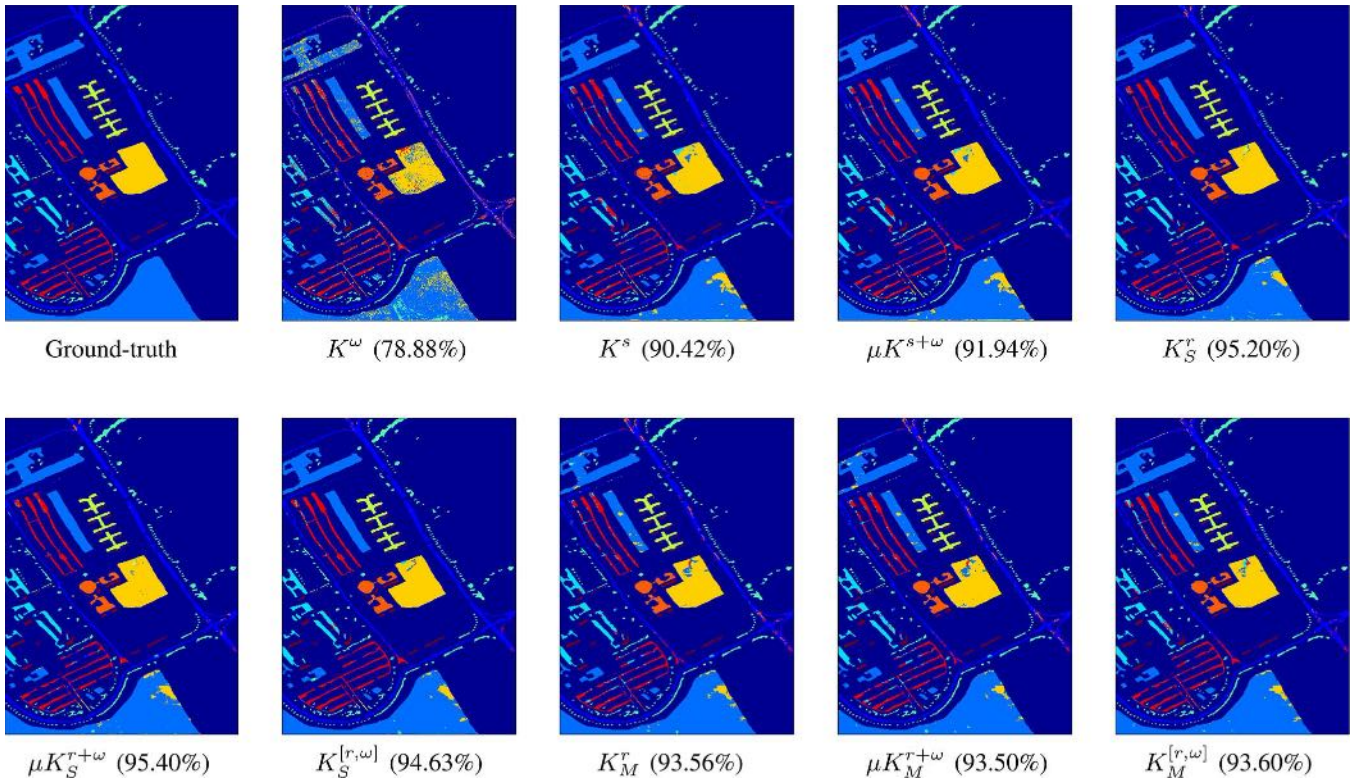


Fig. 12. Classification maps on the University of Pavia data set (OAs are included in the parentheses).

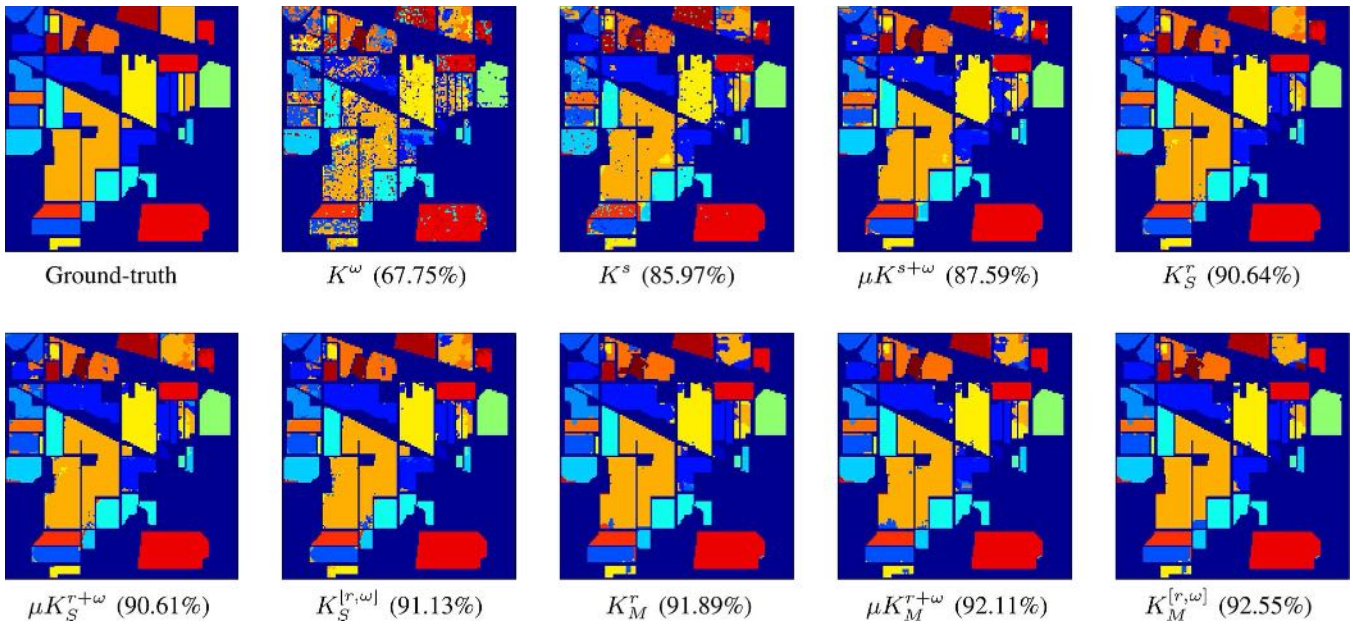


Fig. 13. Classification maps on the Indian Pines data set (OAs are included in the parentheses).

The proposed region kernel can accurately reflect region similarity. We provide a simulation example on the KSC data set to show that the region-kernel-based region-to-region similarity metric (RRSM) in (8) is more accurate than the traditional Gaussian-kernel-based point-to-point similarity metric (PPSM) in (1). We choose a sample from the Scrub class as the benchmark sample and find five other samples from the same class and a different class, respectively. Fig. 14(a) shows the spectral curves of the benchmark sample in red, its five homogeneous samples in blue, and five inhomogeneous samples from the Graminoid marsh class in green. In Fig. 14(b), five inhomogeneous samples in green are from the Willow swamp class.

From the first case in Fig. 14(a), we can see that the benchmark sample is more similar to the five homogeneous samples than to the five inhomogeneous samples, whereas in the second case in Fig. 14(b), the spectral curve

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

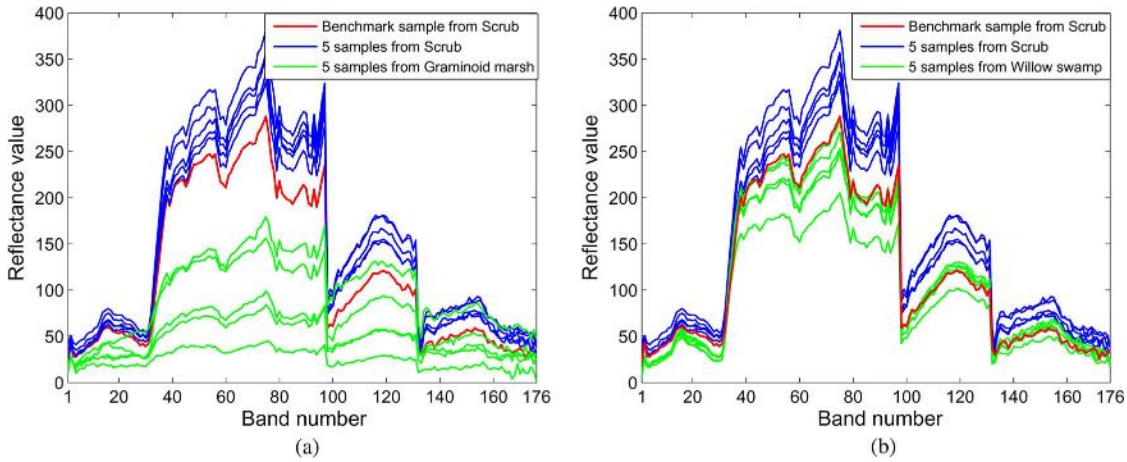


Fig. 14. Spectral curves of selected samples from the KSC data set: a benchmark sample (red) and five homogeneous samples (blue) from the Scrub class, (a) five inhomogeneous samples (green) from the Graminoid marsh class, and (b) five inhomogeneous samples (green) from the Willow swamp class.

of the benchmark sample is closer to that of the inhomogeneous samples. Table IV shows the within-class and between-class similarities between the benchmark sample and its five homogeneous and inhomogeneous samples, in terms of PPSM and RRSM. From Table IV, we can clearly see that the PPSM-based within-class similarities are larger than the between-class similarities in Case 1, but smaller than the between-class similarities in Case 2. Based on the point-to-point spectral similarity metric, traditional spectral-based classifiers will make a correct classification for the benchmark sample in Case 1 but a wrong classification in Case 2. The spectral PPSM is insufficient to reflect the intrinsic similarity between the HSI samples. According to the spatial region distribution characteristics of an HSI, the spatial region structures of different materials will be different and can be used to improve the spectral similarity. By exploiting the spatial region size, shape, and structure information, the RRSM-based within-class similarities are larger than the between-class similarities in both cases. Moreover, compared with PPSM, RRSM increases the within-class similarities, whereas it reduces the between-class similarities. That is, it enlarges the difference between the homogeneous and inhomogeneous samples and increases the class separability. The benchmark sample that is wrongly classified by the spectral PPSM in Case 2 is now correctly classified by using the RRSM.

3) In RKSVM, we construct the region-to-region kernel by means of multiscale box kernels. Considering that one single box kernel is insufficient to describe the similarity between different complex pixel regions, we use a linear combination of multiscale box kernels to approximate the region-to-region distance similarity. In our experiments, a box is generated by choosing a lower threshold $\ell \in [25, 30, 35]$ and an upper threshold $u \in [65, 70, 75]$; thus, there are nine single-scale boxes. We show the classification OAs of the single-region-kernel-based RKSVM-SN and RKSVM-MN in the cases of nine single-scale boxes and multiscale box (the combination of nine boxes) in Table V, where 15 labeled samples per class are used for training. It can be clearly seen that the classification OAs in the multiscale box are better than those in the single-scale boxes for the University of Pavia (PaviaU) and Indian Pines data sets. In particular, for University of Pavia, the multiscale box improves the single-scale boxes by OA of almost 1% and 2% for RKSVM-SN and RKSVM-MN, respectively. For the KSC data set, the OA in the multiscale box is extremely close to the best OA in single-scale boxes. From these results, we find that the combination of multiscale boxes in our RKSVM can approximate or improve the optimal results in the single-scale boxes. Using multiscale boxes, RKSVM is free of the scale selection and can fuse the complementary information in different single-scale boxes to improve the classification performance.
4) It should be noted that the region based on distance similarity measure or morphological operation can be an irregular shape in the spatial domain, whereas the box has a regular shape, in the form of d-ary Cartesian product. The box is constructed based on the spectral values of pixels in the region and consists of multidimensional intervals to bound spectral values in each spectral band. The region captures spatial neighboring pixels, whereas the box reflects the spectral variations of these pixels.
5) Effectiveness of classification is desirable to the case with limited training samples. In RKSVM, all pixels in a region are intuitively assigned as the same class label. The labeled regions can be considered as additional training samples, particularly in the framework (19). Therefore, compared with classic SVM methods, RKSVM is more effective to the case with limited training samples.
6) RKSVM can be considered as a generalization of SVM-CK. Different from SVM-CK that uses the mean of the pixels in a spatial neighborhood, RKSVM uses different lower and upper bounds of the pixels in the neighborhood region to form boxes. That is, SVM-CK operates on pixel points, whereas RKSVM operates on multiscale pixel sets. In the extreme case, when the lower and upper bounds are chosen as the mean, the region is reduced to the mean sample, and RKSVM can revert back to SVM-CK.

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

VI. Conclusion

In this paper, we have proposed a new region-kernel-based SVM framework for classification of hyperspectral remote sensing images. Different from the traditional composite kernel methods operating on pixel points, the proposed RKSVM focuses on the pixel regions and classifies the local regions of an HSI by means of region kernel. It uses a region to capture the local homogeneous pixels and uses a composite region kernel to achieve simultaneous classification of labeled pixels and labeled regions. Experimental results have shown that, by integrating the labeled pixels and the prior knowledge on the spatial homogeneous regions, the proposed pixel-region-based RKSVM outperforms pixel-point-based SVM-CK consistently.

Acknowledgment

The authors would like to thank the Editor-in-Chief, the anonymous Associate Editor, and three reviewers for their insightful comments and suggestions, which have greatly improved this paper. They would also like to thank Prof. M. Crawford for providing the KSC data set, Prof. P. Gamba for providing the University of Pavia data set, Prof. D. Landgrebe for providing the Indian Pines data set, Prof. M. Fauvel for sharing area filtering source codes, and Prof. C. Lin for providing the LIBSVM toolbox.

References

[1] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 45–54, Jan. 2014.

[2] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 55–63, Jan. 1968.

[3] J. M. Bioucas-Dias *et al.*, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[4] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, Sep. 2009.

[5] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[6] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.

[7] D. Tao, X. Tang, X. Li, and X. Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1088–1099, Jul. 2006.

[8] C. Gong, D. Tao, K. Fu, and J. Yang, "Fick's law assisted propagation for semi-supervised learning," *IEEE Trans. Neural Netw. Learn. Syst.*, DOI: 10.1109/TNNLS.2014.2376963, to be published.

[9] R. L. Kettig and D. A. Landgrebe, "Classification of multispectral image data by extraction and classification of homogeneous objects," *IEEE Trans. Geosci. Electron.*, vol. GE-14, no. 1, pp. 19–26, Jan. 1976.

[10] D. A. Landgrebe, "The development of a spectral–spatial classifier for earth observational data," *Pattern Recognit.*, vol. 12, no. 3, pp. 165–175, 1980.

[11] G. Camps-Valls, L. Gomez-Chova, J. Muñoz Maré, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.

[12] G. Camps-Valls, N. Shervashidze, and K. M. Borgwardt, "Spatio-spectral remote sensing image classification with graph kernels," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 741–745, Oct. 2010.

[13] G. Moser and S. B. Serpico, "Combining support vector machines and Markov random fields in an integrated framework for contextual image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2734–2752, May 2013.

[14] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial–spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recognit.*, vol. 45, pp. 381–392, Jan. 2012.

[15] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral–spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[16] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

[17] A. Plaza, P. Martínez, J. Plaza, and R. Pérez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 466–479, Mar. 2005.

[18] L. Zhang, L. Zhang, D. Tao, and X. Huang, "On combining multiple features for hyperspectral remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 879–893, Mar. 2012.

[19] D. Tao, X. Li, X. Wu, and S. Maybank, "General tensor discriminant analysis and gabor features for gait recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1700–1715, Oct. 2007.

[20] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image spectral–spatial feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 242–256, Jan. 2013.

[21] Q. Jackson and D. Landgrebe, "Adaptive Bayesian contextual classification based on Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2454–2463, Nov. 2002.

[22] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.

[23] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Adaptive pixel neighborhood definition for the classification of hyperspectral images with support vector machines and composite kernel," in *Proc. IEEE ICIP*, 2008, pp. 1884–1887.

[24] Y. Gong, "Homogeneous-region analysis of hyperspectral image based on HDA and MRF," Ph.D. dissertation, Photogramm. Remote Sens., Wuhan Univ., Wuhan, China, 2007.

[25] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral–spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.

[26] S. Melacci and M. Gori, "Learning with box kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2680–2692, Nov. 2013.

[27] P. Soille, "Beyond self-duality in morphological image analysis," *Image Vis. Comput.*, vol. 23, no. 2, pp. 249–257, Feb. 2005.

[28] P. Salembier, J. Serra, and A. Plaza, "Flat zones filtering, connected operators, and filters by reconstruction," *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1153–1160, Aug. 1995.

[29] N. Cristianini, J. Shawe-Taylor, A. Elisseeff, and J. Kandola, "On kernel target alignment," in *Advances in Neural Information Processing Systems*, vol. 14. Cambridge, MA, USA: MIT Press, 2002, pp. 367–373.

[30] D. Tuia, G. Camps-Valls, G. Matasci, and M. Kanevski, "Learning relevant image features with multiple-kernel classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3780–3791, Oct. 2010.

[31] Y. Gu *et al.*, "Representative multiple kernel learning for classification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2852–2865, Jul. 2012.

[32] J. T. Kwok and I. W. Tsang, "Learning with idealized kernels," in *Proc. Int. Conf. Mach. Learn.*, 2003, pp. 400–407.

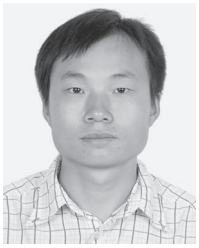
[33] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[34] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 27, Apr. 2011.

[35] J. Ham, Y. Chen, M. M. Crawford, and J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 492–501, Mar. 2005.

[36] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Mar. 2012.

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.

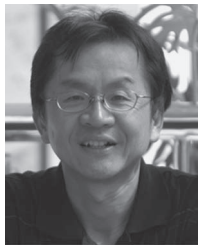


Jiangtao Peng received the B.S. and M.S. degrees from Hubei University, Wuhan, China, in 2005 and 2008, respectively, and the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2011.

He is currently a Lecturer with the Faculty of Mathematics and Statistics, Hubei University. He was visiting the Department of Computer and Information Science, University of Macau. His research interests include machine learning and hyperspectral image processing.

Yicong Zhou (M'07–SM'14) received the B.S. degree from Hunan University, Changsha, China, in 1992 and the M.S. and Ph.D. degrees from Tufts University, Medford, MA, USA, in 2008 and 2010, respectively, all in electrical engineering.

He is currently an Assistant Professor with the Department of Computer and Information Science, University of Macau, Macau, China. His research interests focus on multimedia security, image/signal processing, pattern recognition, and medical imaging.



C. L. Philip Chen (S'88–M'88–SM'94–F'07) received the M.S. degree in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 1985 and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1988.

After having worked in the U.S. for 23 years as a Tenured Professor, a Department Head, and an Associate Dean in two different universities, he is currently the Dean of the Faculty of Science and Technology and a Chair Professor of the Department of Computer and Information Science with the University of Macau, Macau, China.

Dr. Chen is a Fellow of the American Association for the Advancement of Science and Hong Kong Institution of Engineers. He has been a Junior President of the IEEE Systems, Man, and Cybernetics Society and is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS.

Authorized licensed use limited to: Universidade de Macau. Downloaded on April 12,2026 at 16:18:47 UTC from IEEE Xplore. Restrictions apply.