Article

Hybrid Collaborative Representation for Remote-Sensing Image Scene Classification

Bao-Di Liu [1,†,*] , **Wen-Yang Xie [1]** , **Jie Meng [1]** , **Ye Li [2]** and **Yanjiang Wang [1]**

[1] College of Information and Control Engineering, China University of Petroleum (Huadong), Qingdao 266580, China; xiewenyang_upc@163.com (W.-Y.X.); 15564855530@163.com (J.M.); yjwang@upc.edu.cn (Y.W.)
[2] Shandong Provincial Key Laboratory of Computer, Networks Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan 250101, China; liye@sdas.org
* Correspondence: thu.liubaodi@gmail.com; Tel.: +86-157-6421-7948
† Current address: No.66 Changjiang Road West, Huangdao District, Qingdao 266580, China.

check for
updates

Abstract: In recent years, the collaborative representation-based classification (CRC) method has achieved great success in visual recognition by directly utilizing training images as dictionary bases. However, it describes a test sample with all training samples to extract shared attributes and does not consider the representation of the test sample with the training samples in a specific class to extract the class-specific attributes. For remote-sensing images, both the shared attributes and class-specific attributes are important for classification. In this paper, we propose a hybrid collaborative representation-based classification approach. The proposed method is capable of improving the performance of classifying remote-sensing images by embedding the class-specific collaborative representation to conventional collaborative representation-based classification. Moreover, we extend the proposed method to arbitrary kernel space to explore the nonlinear characteristics hidden in remote-sensing image features to further enhance classification performance. Extensive experiments on several benchmark remote-sensing image datasets were conducted and clearly demonstrate the superior performance of our proposed algorithm to state-of-the-art approaches.

Keywords: collaborative representation; hybrid collaborative representation; kernel space; remote-sensing images

1. Introduction

Remote-sensing (RS) images are widely used for land-cover classification, target identification, and thematic mapping from local to global scales owing to its technical advantages such as multiresolution, wide coverage, repeatable observation, and multi/hyperspectral–spectral records [1]. During the past few decades, we have witnessed the rapid development of remote-sensing technology. Nowadays, a volume of heterogeneous RS images, with different spatial and spectral resolutions, provide a vivid service for a myriad Earth observation (EO) applications, from road detection and traffic monitoring to visual tracking and weather reports, respectively. To offer a better performance of such intelligent EO applications, RS image scene recognition has attracted widespread attention. Weifeng Liu et al. [2,3] applied p-Laplacian regularization to scene recognition. Generally, an integrated RS image scene-recognition system includes two components, i.e., a feature-learning approach [4] and a corresponding classifier, and both have a vital effect on the classification result.

As a core problem in image-related applications, image feature representation exhibits the trend of transference from handcrafted to learning-based methods. Specifically, most of the early literature

is based on handcrafted features, such as bag-of-visual-words (BoVW) [5], part-based models [6], and models fusing global and local descriptors [7]. However, handcrafted features are limited in their ability to extract robust and transferable feature representation for image scene classification and ignore many effective cues hiding in the image. Then, an unsupervised feature-learning framework [8] was proposed by Zhang et al. for scene classification. After that, Fu et al. [9] proposed an unsupervised feature-learning method for high-resolution-image scene classification. In this method, a set of filter banks are learned from the unlabeled image patches, and then each image is encoded by these filter banks, generating a final feature vector to excellently represent the image scene. An effective geospatial object-detection framework [10] was proposed by Han et al., combining the weakly supervised learning (WSL) and high-level feature learning. Li et al. [11] developed a new framework combining multiple features including both linear and nonlinear features for the classification of hyperspectral scenes. In 2006, Hinton [12] pointed out that deep neural networks could learn more profound and essential features of objects of interest, which led to tremendous performance enhancement. After that, many attempts have been made to utilize deep-learning methods to feature learning in remote-sensing images. Li et al. [13] presented an unsupervised deep-learning method to learn a data-driven feature for Urban Villages. Then, as one of the most popular deep-learning models in image processing, convolutional neural networks (CNNs) currently dominate computer vision literature, achieving state-of-the-art performance in almost every topic to which they are applied. However, for some classification tasks of remote-sensing images, there are no large amounts of labeled training data. Some efforts have been made to combine CNN with unlabeled data to tackle the problem by performing unsupervised learning. Yu et al. [14] proposed balanced data-driven sparsity to help train CNN in an unsupervised way. Chaib et al. [15] proposed a new feature-learning method, which combined deep features extracted from different pretrained CNN models of the Very High Resolution image scene.

Another core problem is to construct a visual classifier. Visual-classifier design is one of the fundamental issues in computer vision areas. Representation-residual-based classifiers have attracted more attention in recent years due to the emerging paradigm of compressed sensing (CS) [16–18]. The first representation-based classifier, the sparse representation-based classification (SRC) method was proposed within the framework of the nearest subspace. The SRC [19] method firstly obtained the sparse representation of the test sample, and then measured the residual error from the training samples of each class. After that, Zhang et al. [20] proposed the collaborative representation-based classification (CRC) algorithm by using collaborative representation (ℓ_2 norm regularizer) instead of sparse representation (ℓ_1 norm regularizer). Many researchers from the field of remote sensing are attracted by the superior performance of SRC and CRC. Wu et al. [21] introduced an improved sparse representation-based classification method, which represented the test samples with a feature dictionary. Then, a novel sparse representation classification method [22] was proposed by Tang et al., which added the Local Binary Pattern(LBP) feature into the SRC model to extract the local texture of the remote-sensing image. For CRC, Li et al. [16] reviewed several representation-based classification approaches for hyperspectral remote-sensing imagery, including representation-based classification with weighted regularization by measuring similarity between each atom and a testing sample, or with dictionary partition using class-specific labeled samples and other methods. Li et al. [23] proposed a joint collaborative representation (CR) classification method, which uses several complementary features to represent an image, including spectral value and spectral gradient features, Gabor texture features, and DMP features. Then, Jiang et al. [24] proposed a spatial-aware collaborative representation (CR) for hyperspectral images classification, which utilized both the spatial and spectral features to represent images. From the perspective of representation of a testing sample with training samples, both shared attributes and class-specific attributes exist in remote-sensing images, and both of them are vital in the representation of an image. However, these CRC methods used on RS images before did not consider these two image representations at the same time. Generally, existing nearest subspace approaches are mainly categorized into two types: shared representation-based classification (SRC and CRC), and class-specific representation-based classification (LRC). For shared

representation-based classification, Deng et al. [25] proposed a superposed linear representation classifier (SLRC) by representing the test image in terms of superposition of the class centroids and the shared intraclass differences. For class-specific representation-based classification, Liu et al. [26,27] proposed a class-specific representation algorithm that can find the intrinsic relationship between base vectors and the original-image features. Wang et al. [28] proposed a label-constrained specific representation approach to preserve the structural information in the feature space.

In this paper, we propose a hybrid collaborative representation-based classification approach. The proposed method is capable of improving the performance of classifying remote-sensing images by embedding the class-specific collaborative representation to CRC. Moreover, we extend the proposed method to arbitrary kernel space to explore the nonlinear characteristics hidden in remote-sensing image features to further enhance the performance of classification. The scheme of our proposed method is listed in Figure 1. Our work foci are threefold:

- We propose a novel hybrid collaborative representation-based classification method that considers both conventional collaborative representation and class-specific collaborative representation.
- We extend our proposed hybrid collaborative representation-based classification method to arbitrary kernel space to find the nonlinear structures hidden in the image features.
- The proposed hybrid collaborative representation-based classification method is evaluated on four benchmark remote-sensing image datasets and achieves state-of-the-art performance.
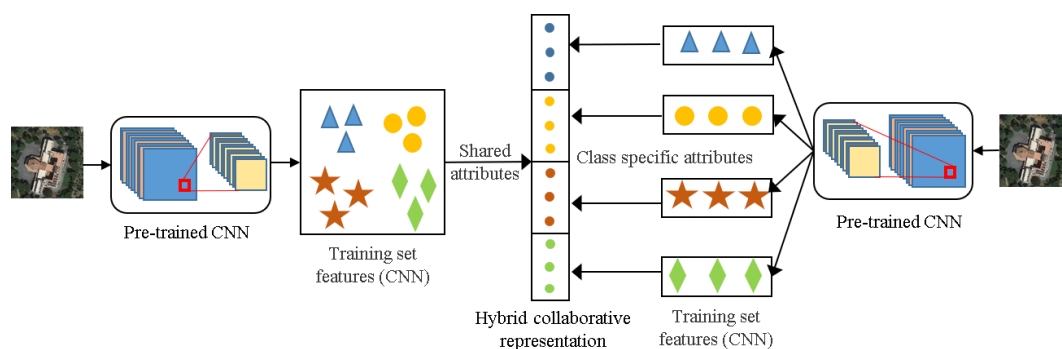
Figure 1. Scheme of our proposed hybrid collaborative representation algorithm. The left part is the collaborative representation to extract the shared attributes while the right part is the class-specific collaborative representation to extract class-specific attributes. Their combination forms our hybrid collaborative representation algorithm.

The rest of the paper is organized as follows. Section 2 overviews several classical visual-recognition algorithms and proposes our hybrid collaborative representation-based classification with kernels. Then, experimental results and analysis are shown in Section 3. Discussion about the experimental results and the proposed method is presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Proposed Method

In this section, we review related work about CRC. Then, we introduce some work about class-specific CRC (CS-CRC). Finally, we focus on introducing our proposed approach.

2.1. Overview of CRC

Zhang et al. [20] proposed the CRC. For CRC, all training samples are concatenated together as the base vectors to form a subspace, and the test sample is described in the subspace. To be specific, given the training samples $X = [X^1, X^2, \cdots X^C] \in \mathbb{R}^{D \times N}$, $X^c \in \mathbb{R}^{D \times N_c}$ represents the training samples from the c_{th} class, C represents the number of classes, N_c represents the number of training samples in

the c_{th} class ($N = \sum\limits_{c=1}^{C} N_c$), and D represents the dimensions of the samples. Supposing that $y \in \mathbb{R}^{D \times 1}$ is a test sample, the objective function of CRC is as follows:

$$\hat{s} = \arg \min_s \left\{ \|y - Xs\|_2^2 + \lambda \|s\|_2^2 \right\}. \tag{1}$$

Here, λ is the regularization parameter to control the tradeoff between fitting goodness and collaborative term (i.e., multiple entries in X participating into representing the test samples). The role of the regularization term is twofold. Firstly, the ℓ_2 norm makes the least-square solution stable. Secondly, it introduces a certain amount of "sparsity" to collaborative representation $\hat{s}$ and indicates that it is the collaborative representation but not the ℓ_1-norm sparsity that makes sparsity powerful for classification. Collaborative representation-based classification effectively utilizes all training samples for visual recognition and the objective function of CRC has analytic solutions.

2.2. Class-Specific Collaborative Representation

For class-specific collaborative representation, the training samples in each category are considered a subspace. A test sample is represented with the samples in the specific class. The objective function of class-specific collaborative representation is as follows:

$$\hat{s} = \arg \min_s \sum_{c=1}^{C} \left\{ \|y - X^c s^c\|_2^2 + \gamma \|s^c\|_2^2 \right\} \tag{2}$$

Here, γ is the regularization parameter to control the tradeoff between fitting goodness and collaborative term. The CS-CRC is capable of describing the sample y in each category.

2.3. Hybrid Collaborative Representation

Theorem 1. *The reconstruction error of collaborative representation is bounded by the reconstruction error of class-specific collaborative representation.*

$$\|y - Xs\|_2^2 \leq \sum_{c=1}^{C} \|y - X^c s^c\|_2^2 \tag{3}$$

Proof of Theorem 1. According to Cauchy inequality,

$$
\begin{aligned}
\|y - Xs\|_2^2 &= \left\| \tfrac{1}{C} (Cy - CXs) \right\|_2^2 \\
&= \left\| \tfrac{1}{C} \left\{ (y - X^1 s^1) + \cdots + (y - X^c s^c) + \cdots + (y - X^C s^C) \right\} \right\|_2^2 \\
&\leq \tfrac{\|y - X^1 s^1\|_2^2 + \cdots + \|y - X^c s^c\|_2^2 + \cdots + \|y - X^C s^C\|_2^2}{C^2} \left(1^2 + \cdots + 1^2 \right) \\
&= \tfrac{1}{C} \sum_{c=1}^{C} \|y - X^c s^c\|_2^2 \\
&\leq \sum_{c=1}^{C} \|y - X^c s^c\|_2^2
\end{aligned}
\tag{4}
$$

$\square$

Shared collaborative representation is advantageous to reduce the reconstruction error (as shown in Theorem 1), while class-specific collaborative representation is conducive to capturing discriminant information. We propose a hybrid collaborative representation algorithm to combine these two approaches. The objective function is as follows:

$$\hat{s} = \arg \min_s \left\{ \|y - Xs\|_2^2 + \lambda \|s\|_2^2 + \tau \sum_{c=1}^{C} \left\{ \|y - X^c s^c\|_2^2 + \gamma \|s^c\|_2^2 \right\} \right\} \tag{5}$$

In Equation (5), the first two terms are the conventional collaborative representation and the latter two terms are the class-specific collaborative representation. Conventional collaborative representation guarantees residual error and robustness, while the class-specific collaborative representation obtains distinctiveness via different classes. Equation (5) can be further arranged as follows:

$$\hat{s} = \arg\ min_s \left\{ \|y - Xs\|_2^2 + \beta \|s\|_2^2 + \tau \sum_{c=1}^{C} \left\{ \|y - X^c s^c\|_2^2 \right\} \right\} \tag{6}$$

Here, $\beta = \lambda + \tau \times \gamma$. The latter form of Equation (6) arose in the estimation of latent-variable graphical models [29].

2.4. Hybrid Collaborative Representation with Kernels

Superior performance of visual recognition is often achieved in Reproducing Kernel Hilbert Space because a nonlinear structure often exists in image features. Our proposed hybrid collaborative representation algorithm is easily extended to arbitrary kernel space. Suppose there exists a kernel function $\kappa(x, y) = \phi^T(x)\phi(y)$. Here, $\phi : R^D \to R^K(D < K)$ maps the image features to the high-dimensional feature space, $\phi(X) = [\phi(X^1), \phi(X^2), \cdots, \phi(X^C)] \in R^{K \times N}$. The objective function of our proposed hybrid collaborative representation algorithm with kernels is as follows:

$$\hat{s} = \arg\ min_s \left\{ \|\phi(y) - \phi(X)s\|_2^2 + \beta \|s\|_2^2 + \tau \sum_{c=1}^{C} \left\{ \|\phi(y) - \phi(X^c)s^c\|_2^2 \right\} \right\} \tag{7}$$

2.5. Optimization of the Objective Function

The objective function in Equation (7) can be rewritten as follows:

$$f(s) = \|\phi(y) - \phi(X)s\|_2^2 + \beta \|s\|_2^2 + \tau \sum_{c=1}^{C} \left\{ \|\phi(y) - [0, \cdots, \phi(X^c), \cdots, 0]s\|_2^2 \right\} \tag{8}$$

Here, let $[0, \cdots, \phi(X^c), \cdots, 0]$ be $\phi(\hat{X})$. Equation (8) can be simplified as follows:

$$
\begin{aligned}
f(s) &= \|\phi(y) - \phi(X)s\|_2^2 + \beta \|s\|_2^2 + \tau \sum_{c=1}^{C} \left\{ \|\phi(y) - \phi(\hat{X})s\|_2^2 \right\} \\
&= trace \left\{ \phi(y)^T \phi(y) - 2\phi(y)^T \phi(X) s + s^T \phi(X)^T \phi(X) s \right\} + \beta trace \left\{ s^T s \right\} \\
&\quad + \tau \sum_{c=1}^{C} \left\{ trace \left\{ \phi(y)^T \phi(y) - 2\phi(y)^T \phi(\hat{X}) s + s^T \phi(\hat{X})^T \phi(\hat{X}) s \right\} \right\} \\
&= (1 + \tau C)\, trace \left\{ \kappa(y, y) \right\} - 2(1 + \tau)\, trace \left\{ \kappa(y, X) s \right\} \\
&\quad + trace \left\{ s^T \left(\kappa(X, X) + \beta I + \tau \sum_{c=1}^{C} \kappa(\hat{X}, \hat{X}) \right) s \right\}
\end{aligned} \tag{9}
$$

It is easy to obtain the optimum $\hat{s}$ for Equation (9), as follows:

$$\hat{s} = \left(\kappa(X, X) + \beta I + \tau \sum_{c=1}^{C} \kappa(\hat{X}, \hat{X}) \right)^{-1} \left\{ (1 + \tau)\, \kappa(y, X) \right\} \tag{10}$$

2.6. Hybrid Collaborative Representation-Based Classification with Kernels

After obtaining collaborative code $\hat{s}$, hybrid collaborative representation-based classification aims to find the minimum value of the residual error for each class:

$$id(y) = \arg\ min_c \|\phi(y) - \phi(X^c)\hat{s}^c\|_2^2. \tag{11}$$

where, $id(y)$ is the label of the testing sample, and y belongs to the class that has the minimum residual error. The procedure of the hybrid collaborative representation-based classification with kernels is shown in Algorithm 1.

Algorithm 1 Algorithm for hybrid collaborative representation-based classification with kernels

---

Require: Training samples $X \in \mathbb{R}^{D \times N}$, β, and test sample y
  1: **for** $c = 1; c \leq C; c++$ **do**
  2:　　Code y with the hybrid collaborative representation algorithm with kernels.
  3:　　Compute the residuals $e^c(y) = \|\phi(y) - \phi(X^c)s^c\|_2^2$
  4: **end for**
  5: $id(y) = \arg\ min_c \{e^c\}$
  6: **return** $id(y)$

---

3. Experimental Results

In this section, we show our experimental results on four remote-sensing image datasets. To illustrate the significance of our approach, we compared our method with several state-of-the-art methods. In the following section, we firstly introduce the experimental settings. Then, we illustrate the experimental results on each aerial-image dataset.

3.1. Experimental Settings

We tested our method on four datasets, the RSSCN7 dataset [30], UC Merced Land Use dataset [5], WHU-RS19 dataset [31], and AID dataset [32]. For all the datasets, we implemented traditional CNN feature representation, where the image is directly fed into the pretrained VGG model [33] and layer fc6 is utilized to extract a 4096-dimensional vector for each image. The final feature of each image is ℓ_2-normalized for better performance [32]. To eliminate the randomness, we randomly (repeatable) split the dataset into the train set and test set for 10 times, respectively. Average accuracy was recorded.

We also expanded our method into kernel space and used four different kernels: the linear kernel ($\kappa(x, y) = x^T y$), the polynomial kernel (POLY, $\kappa(x, y) = (p + x^T y)^q$), the Hellinger kernel ($\kappa(x, y) = \sum_{d=1}^{D} \sqrt{x_d y_d}$), and the radial basis function kernel (RBF, $\kappa(x, y) = \exp\left(-\gamma \|x - y\|_2^2\right)$). Here, we set $p = 4$, $q = 3$ and $\gamma = 2^{-2}$. For convenience, we denote our proposed hybrid collaborative representation algorithm with kernels as Hybrid-KCRC. The proposed Hybrid-KCRC algorithm is compared with other classification algorithms, including nearest-neighbor classification (NN), LIBLINEAR [34], SOFTMAX, CRC [20], CS-CRC, and SLRC-L2 [25].

3.2. Experiment on UC Merced Land Use Dataset

The UC Merced Land Use Dataset [5] was manually extracted from large images from the USGS National Map Urban-Area Imagery collection for various urban areas around the country. The pixel resolution of this public-domain imagery is 1 foot. The UC Merced Land Use Dataset involves 21 categories of 2100 land-use images in total. Each image measures 256×256 pixels. There are 100 images for each of the following classes: agricultural, airplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, freeway, golfcourse, harbor, intersection, medium residential, mobile homepark, overpass, parking lot, river, runway, sparse residential, storage tanks, and tennis court. In Figure 2, we list several samples from this dataset.

Figure 2. Samples of UC Merced Land Use dataset.

3.2.1. Parameter tuning on UC Merced Land Use Dataset

There are two parameters in the objective function of the Hybrid-KCRC algorithm that need to be specified. β is an important parameter in the Hybrid-KCRC algorithm, which is used to adjust the tradeoff between the reconstruction error and the collaborative representation. τ is another important factor in the algorithm, which is used to control the tradeoff between the shared collaborative representation and the class-specific collaborative representation. β and τ are tuned to achieve the best accuracy. β are tuned in the range of 2^{-9}, and 2^2 and τ are tuned in the range of 2^{-10} and 2^{-4}. We randomly chose 20 images as the training samples and testing samples from each category, respectively, in this section. Figure 3 shows the classification rate with different β and τ on four kernels. For the four kernels, linear kernel, POLY kernel, RBF kernel, and Hellinger kernel, the optimal parameter β and τ are $(2^{-4}, 2^{-6})$, $(2^0, 2^{-8})$, $(2^{-7}, 2^{-7})$, $(2^0, 2^{-8})$, respectively.
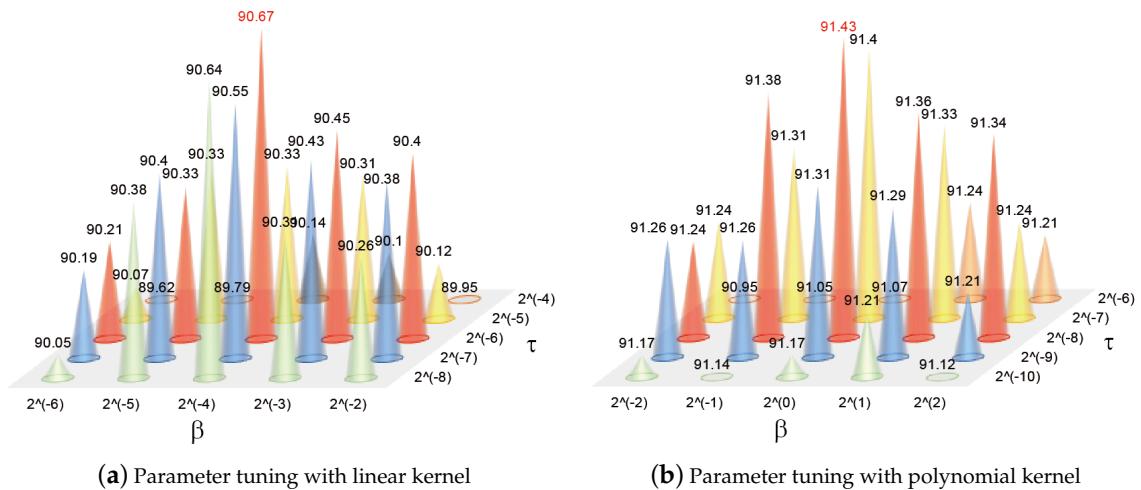

(**a**) Parameter tuning with linear kernel

(**b**) Parameter tuning with polynomial kernel

Figure 3. *Cont.*

(**c**) Parameter tuning with rbf kernel

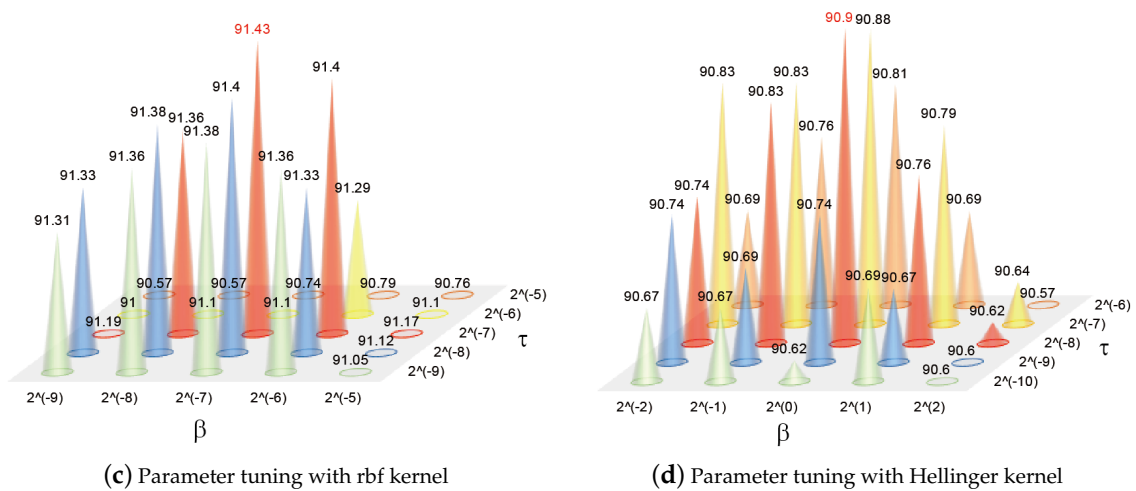(**d**) Parameter tuning with Hellinger kernel

Figure 3. Parameter tuning on UC Merced Land Use Dataset.

### 3.2.2. Comparison with Several Classical Classifier Methods on UC Merced Land Use Dataset

First, we randomly chose 20 images as the training samples and testing samples from each category, respectively, in this section. Table 1 illustrates the effectiveness of Hybrid-CRC for classifying images. For the four kernels, Hybrid-KCRC algorithm achieves the highest accuracy of 91.43% with the POLY kernel and RBF kernel. This is 1.03% higher than the CRC method and 2.33% higher than the CS-CRC method.

Table 1. Comparison with several classical classification methods on the UC Merced Land Use Dataset (%).

Methods\Datasets	UC-Merced
NN	81.88
LIBLINEAR	89.57
SOFTMAX	88.00
SLRC-L2	89.79
CRC	90.40
CS-CRC	89.10
Hybrid-KCRC (linear)	90.67
Hybrid-KCRC (POLY)	91.43
Hybrid-KCRC (RBF)	91.43
Hybrid-KCRC (Hellinger)	90.90

Second, we increased the number of training samples in each category to evaluate the performance of our proposed Hybrid-KCRC method. Figure 4 shows the classification rate on the UC-Merced dataset with 20, 40, 60, 80 training samples in each category. From Figure 4, we can draw the conclusion that our proposed Hybrid-KCRC method achieves superior performance to the liblinear, CRC, and CS-CRC methods.
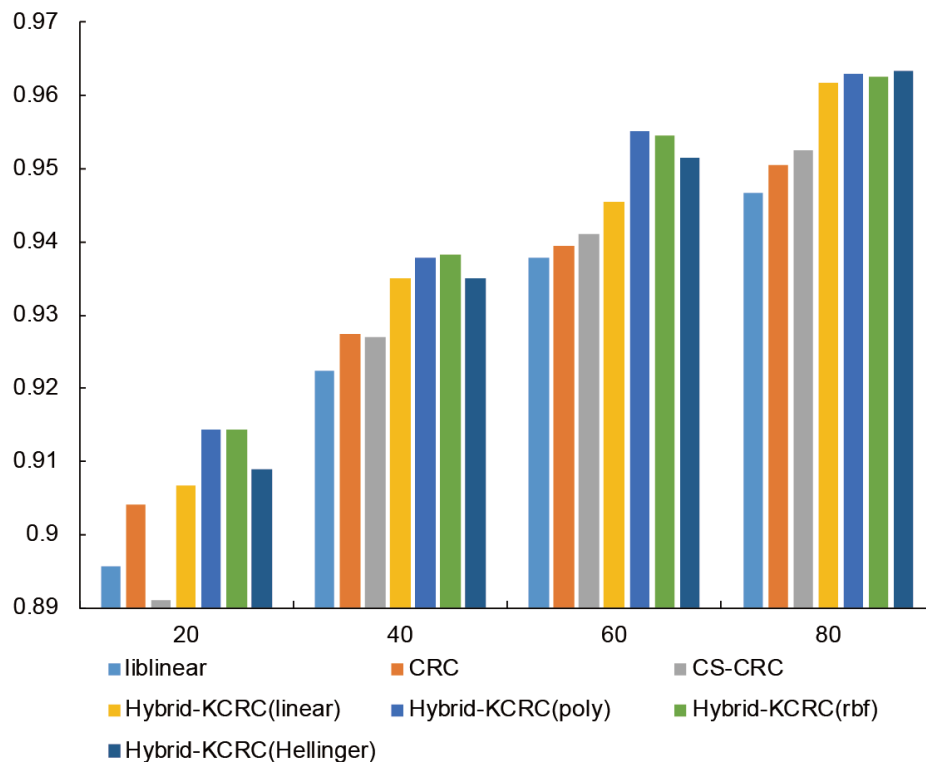
Figure 4. Classification rate on UC-Merced dataset with different number of training samples in each category.

3.2.3. Confusion Matrix on UC Merced Land Use Dataset

To further illustrate the superior performance of our proposed Hybrid-KCRC method, we evaluated the classification rate per class of our method on UC-Merced dataset using a confusion matrix. In this section, we randomly chose 80 images per class as the training samples and 20 images per class as the testing samples. To eliminate randomness, we also randomly (repeatable) split the dataset into a train set and test set for 10 times, respectively. The confusion matrices are shown in Figure 5. From Figure 5, for the Hybrid-KCRC methods, 13, 13, 12, 12 classes achieved classification accuracy greater than or equal to 0.99 for the linear kernel, polynomial kernel, RBF kernel, and Hellinger kernel, respectively. However, 9 and 6 classes achieved classification accuracy greater than or equal to 0.99 for CRC and CS-CRC, respectively. Compared with the CRC method, the Hybrid-KCRC methods achieved a significant performance boost on the denseresidential class. Compared with CS-CRC method, the Hybrid-KCRC methods achieved great performance improvement on the storagetanks and tenniscourt classes. It is worth noting that all methods did not perform well on denseresidental and medimuresidential classes.
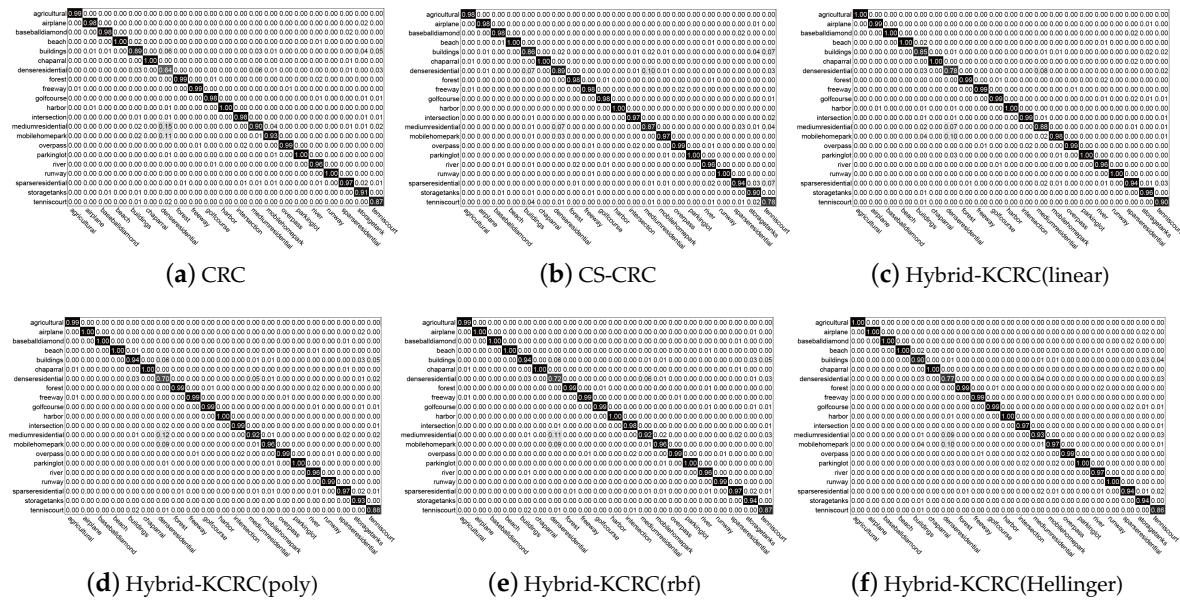
(**a**) CRC

(**b**) CS-CRC

(**c**) Hybrid-KCRC(linear)

(**d**) Hybrid-KCRC(poly)

(**e**) Hybrid-KCRC(rbf)

(**f**) Hybrid-KCRC(Hellinger)

Figure 5. Confusion matrices on UC-Merced dataset.

3.2.4. Comparison with State-of-the-Art Approaches

For comparison, we referred to previous works in the literature [35,36] and randomly selected 80% images from each class as the training set, and the remaining 20% images as the testing set. Several baseline methods (e.g., liblinear, CRC, CS-CRC), as well as state-of-the-art remote-sensing image-classification methods were utilized as benchmarks.

Table 2 shows the overall accuracy of the classification rate with various remote-sensing image-classification methods. First, we compared our proposed Hybrid-KCRC method with the liblinear, CRC, and CS-CRC methods. By comparing our proposed Hybrid-KCRC with the three baseline methods mentioned above, we found that Hybrid-KCRC achieved superior performance to the three baseline methods. It is remarkable that our proposed Hybrid-KCRC is the improvement of the CRC and CS-CRC methods. Second, we compared our Hybrid-KCRC with state-of-the-art remote-sensing image-classification results. It is clear that our proposed Hybrid-KCRC achieved the top performance. It should be noted that the feature utilized by CNN-W + VLAD with SVM, CNN-R + VLAD with SVM, CaffeNet + VLAD is more effective than the feature extracted directly from CNN (e.g., CaffeNet method with 93.42% versus CaffeNet + VLAD method with 95.39%).

Table 2. Experiment on UC-Merced dataset (%).

Methods	Year	Accuracy
SPMK [37]	2006	74%
LDA-SVM [38]	2013	80.33%
SIFT + SC [39]	2013	81.67%
Saliency + SC [8]	2014	82.72%
CaffeNet [40] (without fine-tuning)	2015	93.42%
CaffeNet [41] + VLAD	2015	95.39%
DCGANs [42] (without augmentation)	2017	85.36%
MAGANs [42] (without augmentation)	2017	87.69%
WDM [36]	2017	95.71%
UCFFN [35]	2018	87.83%
CNN-W + VLAD with SVM [43]	2018	95.61%
CNN-R + VLAD with SVM [43]	2018	95.85%

Table 2. *Cont.*

Methods	Year	Accuracy
VGG19 + liblinear		95.05%
VGG19 + CRC		94.67%
VGG19 + CS-CRC		95.26%
VGG19 + Hybrid-KCRC (linear)		96.17%
VGG19 + Hybrid-KCRC (POLY)		96.29%
VGG19 + Hybrid-KCRC (RBF)		96.26%
VGG19 + Hybrid-KCRC (Hellinger)		**96.33**%

3.3. Experiment on the RSSCN7 Dataset

The RSSCN7 dataset was collected from Google Earth 3, which contains 2800 aerial-scene images divided into seven classes, i.e., grassland, forest, farmland, industry, parking lot, residential, and river and lake region. There are 400 images in each class. All images had the same size of 400×400 pixels. Figure 6 shows several sample images from the dataset.
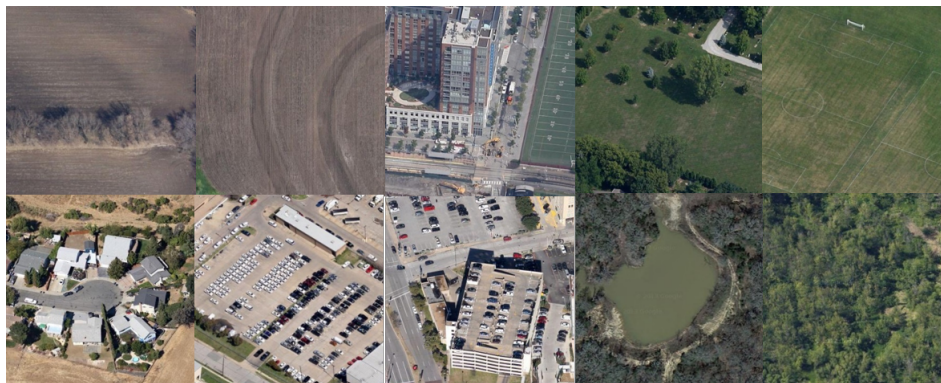

Figure 6. Samples of RSSCN7 dataset.

### 3.3.1. Comparison with Several Classical Classifier Methods on the RSSCN7 Dataset

First, we randomly selected 100 images as the training samples and 100 images as the testing samples from each category. For the four kernels, linear kernel, POLY kernel, RBF kernel, and Hellinger kernel, optimal parameters β and τ were $(2^{-4}, 2^{-7})$, $(2^1, 2^{-6})$, $(2^{-6}, 2^{-6})$, $(2^1, 2^{-6})$, respectively. Recognition accuracy is shown in Table 3.

Table 3. Comparison with several classical classification methods on the RSSCN7 dataset (%).

Methods\Datasets	RSSCN7
NN	76.44
LIBLINEAR	84.84
SOFTMAX	82.14
SLRC-L2	81.99
CRC	85.77
CS-CRC	84.23
Hybrid-KCRC (linear)	86.39
Hybrid-KCRC (POLY)	87.34
Hybrid-KCRC (RBF)	87.29
Hybrid-KCRC (Hellinger)	86.71

From Table 3, we can see that the Hybrid-KCRC algorithm outperformed other conventional methods, achieving accuracy of 86.39%, 87.34%, 86.71%, 87.29% for the linear kernel, POLY kernel, RBF kernel, and Hellinger kernel, respectively. For the four kernels, the Hybrid-KCRC algorithm achieved the highest accuracy, 87.34%, with the POLY kernel. This is 1.57% higher than the CRC method, and 3.11% higher than the CS-CRC method.

Second, we increased the number of training samples in each category to evaluate the performance of our proposed Hybrid-KCRC method. Figure 7 shows the classification rate on the RSSCN7 dataset with 100, 200, and 300 training samples in each category. From Figure 7, we find that our proposed Hybrid-KCRC algorithm achieved superior performance to the baseline methods. The experiment on the CS-CRC method with 200 training samples also achieved a lower classification rate than with 300 training samples. The reason might be that too many training samples make the CRC overfitted. Moreover, both the hybrid-KCRC (poly) and hybrid-KCRC (rbf) achieved top accuracy.
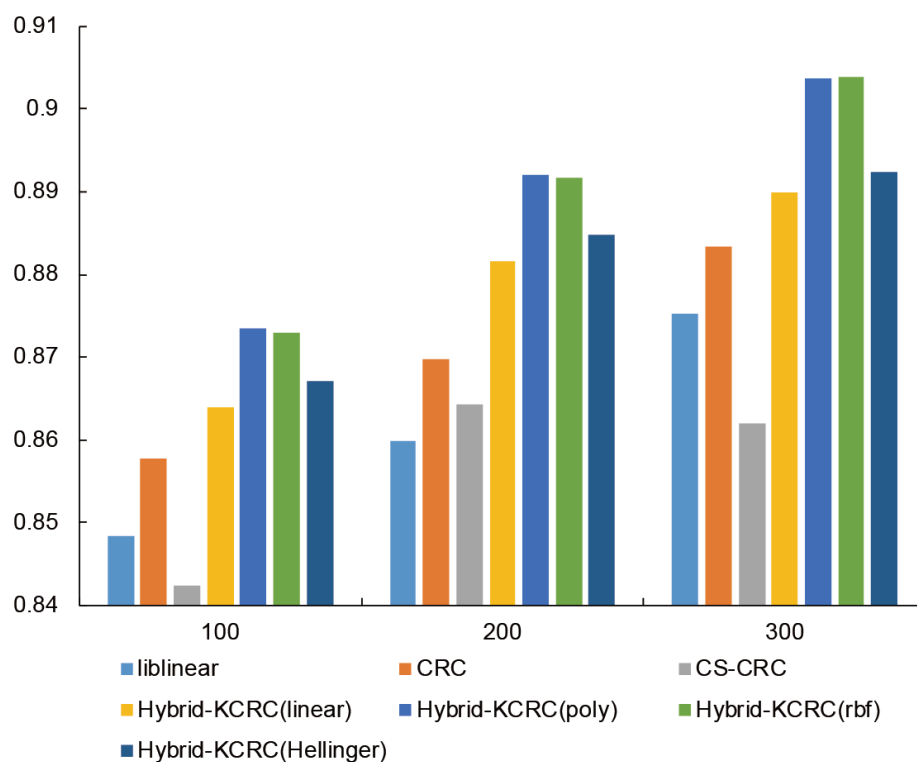
Figure 7. Classification rate on the RSSCN7 dataset with a different number of training samples in each category.

3.3.2. Confusion Matrix on the RSSCN7 Dataset

To further illustrate the superior performance of our proposed Hybrid-KCRC method, we evaluated the classification rate per class of our method on the RSSCN7 dataset using a confusion matrix. In this section, we randomly chose 200 images per class as the training samples and 100 images per class as the testing samples. To eliminate randomness, we also randomly (repeatable) split the dataset into a train set and test set for 10 times, respectively. The confusion matrices are shown in Figure 8. From Figure 8, compared with the CS-CRC method and CRC method, the Hybrid-KCRC methods achieved better performance in most categories.
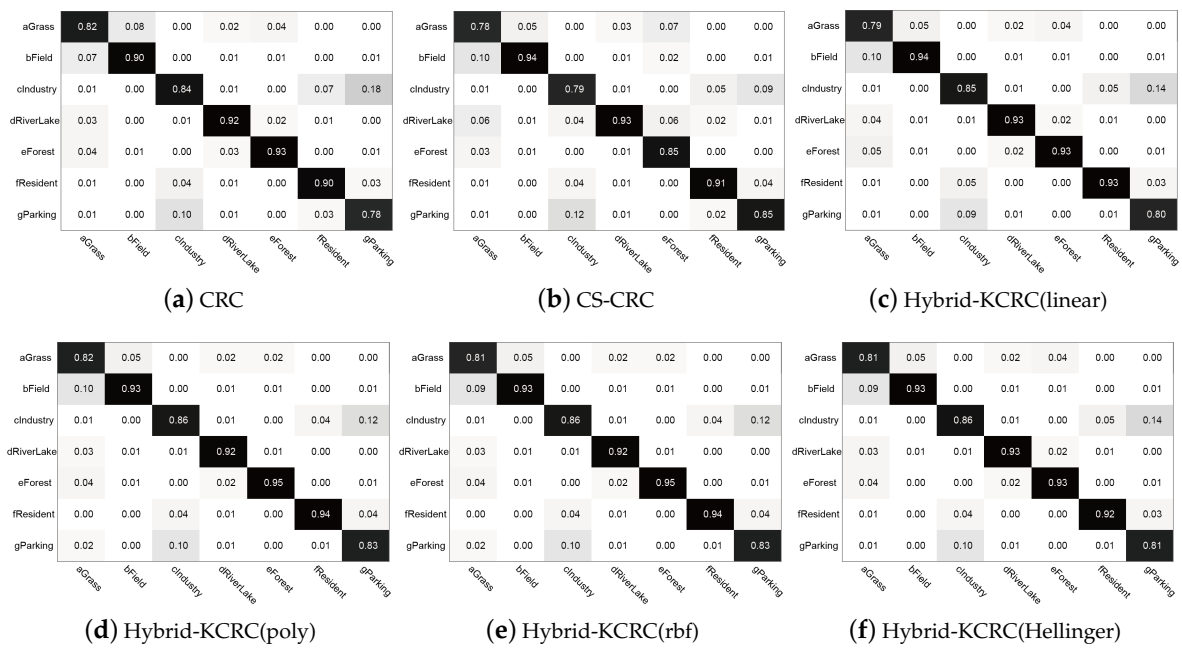
(a) CRC　　　　　　　**(b)** CS-CRC　　　　　　　**(c)** Hybrid-KCRC(linear)

(d) Hybrid-KCRC(poly)　　　　**(e)** Hybrid-KCRC(rbf)　　　　**(f)** Hybrid-KCRC(Hellinger)

Figure 8. Confusion matrices on the RSSCN7 dataset.

3.4. Experiment on the WHU-RS19 Dataset

The WHU-RS19 dataset was collected from Google Earth imagery, which is composed of a total number of 1005 aerial images that belong to 19 classes. Figure 9 shows several samples belonging to this dataset.

Figure 9. Samples of the WHU-RS19 dataset.

We randomly selected 20 and 20 samples per class for training and testing, respectively. For the four kernels, linear kernel, POLY kernel, RBF kernel, and Hellinger kernel, optimal parameters β and τ were $(2^{-2}, 2^{-6})$, $(2^2, 2^{-6})$, $(2^{-6}, 2^{-8})$, $(2^1, 2^{-4})$, respectively. Recognition accuracy is shown in Table 4.

From Table 4, we can see that the Hybrid-KCRC algorithm outperformed other conventional methods, achieving accuracy of 94.76%, 95.34%, 95.34%, 95.39% for the linear kernel, POLY kernel, RBF kernel, and Hellinger kernel, respectively. For the four kernels, the Hybrid-KCRC algorithm achieved the highest accuracy of 95.39% with the Hellinger kernel. This is 0.81% higher than the CRC method and 1.44% higher than the CS-CRC method.

Table 4. Comparison with several classical classification methods on the WHU-RS19 dataset (%).

Methods\Datasets	WHU-RS19
NN	87.74
LIBLINEAR	94.42
SOFTMAX	93.29
SLRC-L2	94.18
CRC	94.58
CS-CRC	93.95
Hybrid-KCRC (linear)	94.76
Hybrid-KCRC (POLY)	95.34
Hybrid-KCRC (RBF)	95.34
Hybrid-KCRC (Hellinger)	**95.39**

Then, we changed the number of training samples in each category to illustrate the performance of our proposed method. Figure 10 shows the classification accuracy on the WHU-RS19 dataset with 10, 20, and 30 training samples in each category. We can clearly see from Figure 10 that our proposed Hybrid-KCRC algorithm achieved superior performance to classical methods.
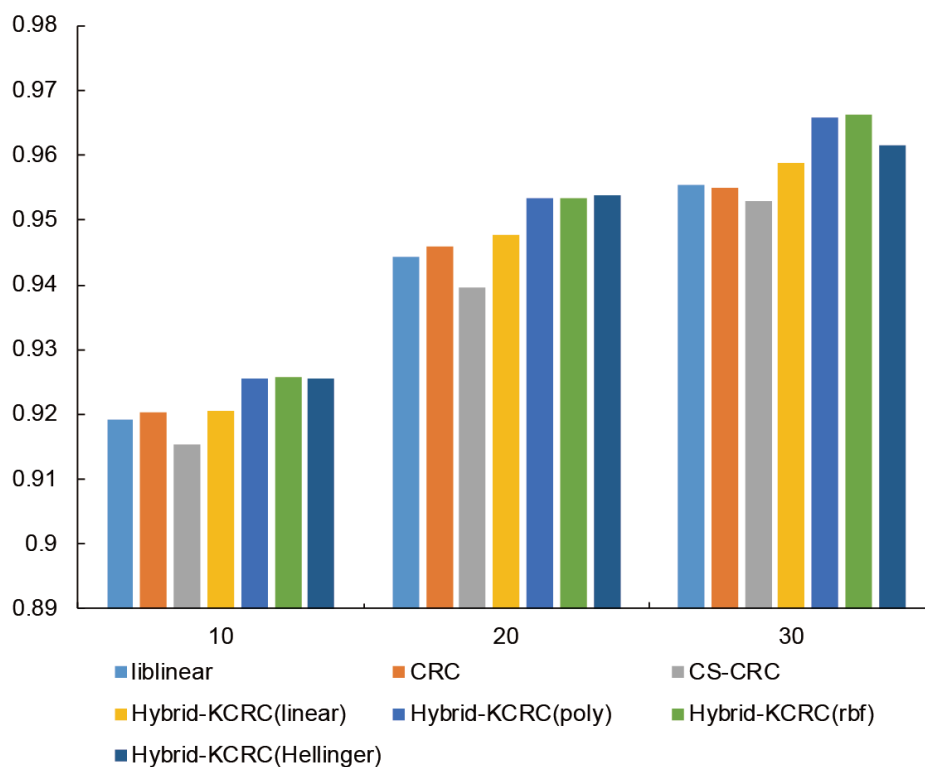
Figure 10. Classification rate on the WHU-RS19 dataset with a different number of training samples in each category.

3.5. Experiment on the AID Dataset

The AID dataset is a new large-scale aerial-image dataset that collects sample images from Google Earth imagery. The AID dataset is the most challenging dataset for the scene classification of aerial images. The dataset is made up of the following 30 aerial-scene types: airport, bare land, baseball field, beach, bridge, center, church, commercial, dense residential, desert, farmland, forest, industrial, meadow, medium residential, mountain, park, parking, playground, pond, port, railway station, resort,

river, school, sparse residential, square, stadium, storage tanks, and viaduct. The size of each aerial image is fixed to 600×600 pixels to cover a scene with various resolutions. There are 10,000 images labeled into 30 categories. In Figure 11, we show several images of this dataset.
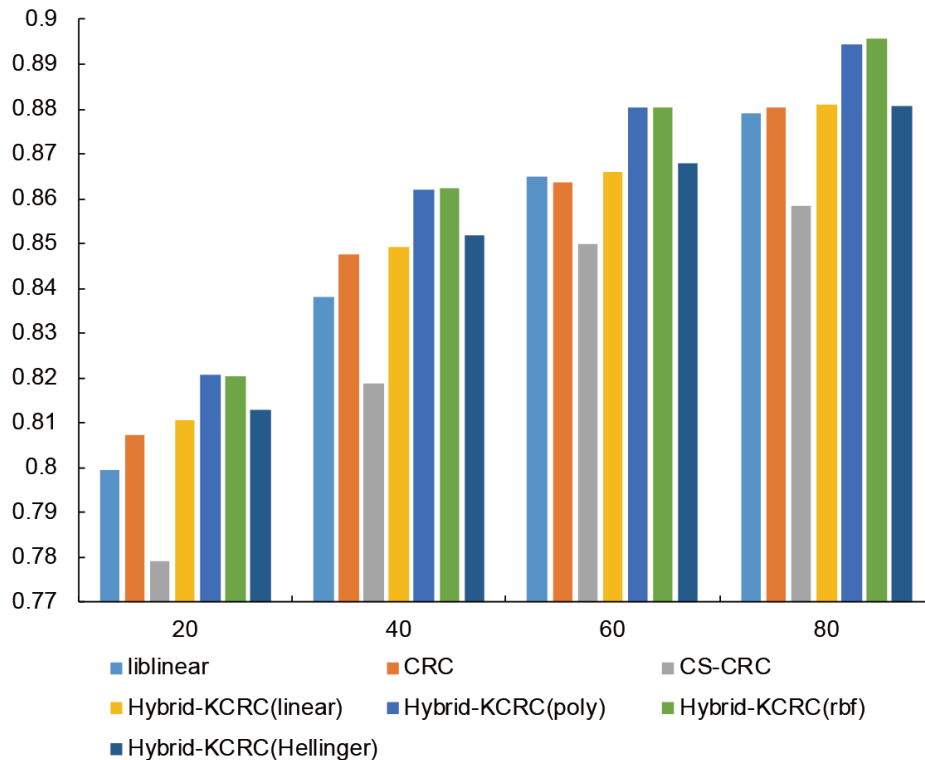

Figure 11. Samples of the AID dataset.

We randomly selected 20 and 20 samples per class for training and testing, respectively. For the four kernels, linear kernel, POLY kernel, RBF kernel, and Hellinger kernel, optimal parameters β and τ were $(2^{-4}, 2^{-7})$, $(2^0, 2^{-8})$, $(2^{-7}, 2^{-9})$, $(2^3, 2^{-6})$, respectively. Recognition accuracy is shown in Table 5.

Table 5. Comparison with several classical classification methods on the AID dataset (%).

Methods\Datasets	AID
NN	65.32
LIBLINEAR	79.93
SOFTMAX	76.13
SLRC-L2	79.27
CRC	80.73
CS-CRC	77.92
Hybrid-KCRC (linear)	81.07
Hybrid-KCRC (POLY)	**82.07**
Hybrid-KCRC (RBF)	82.05
Hybrid-KCRC (Hellinger)	81.28

From Table 5, we can see that the Hybrid-KCRC algorithm outperformed other conventional methods, achieving an accuracy of 81.07%, 82.07%, 82.05%, 81.28% for the linear kernel, POLY kernel, RBF kernel, and Hellinger kernel, respectively. For the four kernels, the Hybrid-KCRC algorithm achieved the highest accuracy of 82.07% with the POLY kernel. This is 1.34% higher than the CRC method and 4.15% higher than the CS-CRC method. We also used different numbers of training samples in each category to evaluate the performance of the Hybrid-KCRC method. The classification rate on the AID dataset using 20, 40, 60, 80 training samples in each category is shown in Figure 12. From Figure 12, we find that our proposed Hybrid-KCRC algorithm was better than several classical classification methods.

Figure 12. Classification rate on the AID dataset with a different number of training samples in each category.

4. Discussion

- For RS image classification, both shared attributes and class-specific attributes are vital to the representation of testing samples with training samples. So, based on CRC, we propose a hybrid collaborative representation-based classification method that can decrease the reconstruction error and improve the classification rate. Through comparison with several state-of-the-art methods for RS image classification, we can see that our proposed method is capable of efficiently promoting the performance of classifying remote-sensing images.

- Because of the existence of a nonlinear structure in image features, we extended our method into Reproducing Kernel Hilbert Space to further improve the performance of our method with kernel functions. From the experimental results of comparing with several classical classification methods, we can see that the classification rates of the Hybrid-KCRC method on these four datasets are all higher than that with the NN, LIBLINEAR, SOFTMAX, SLRC-L2 , CRC and CS-CRC methods. Obviously, our proposed Hybrid-KCRC method achieved superior performance to these methods.

- We took the UC-Merced dataset as an example and evaluated the performance of our proposed Hybrid-KCRC method per class with a confusion matrix. From the confusion matrix, we can see that the Hybrid-KCRC method is better than other methods in most categories.

- It is true that there are several pretrained models to extract features and the performance of the resnet model outperforms the performance of the vgg model. In our paper, however, we paid more attention to the design of the classifier and not feature extraction. We only extracted the features of remote-sensing images to complete the classification task. The vgg is also very popular candidate models for extracting CNN activations of images. That are the reasons why we chose vgg. As a matter of fact, our method could be further improved by using other, better feature-extraction pretrained models. To demonstrate this, we also extracted CNN features with the pretrained Resnet model [44], and layer pool5 was utilized. The feature was a 2048-dimensional vector for

each image. The final feature of each image was ℓ_2-normalized. The experimental results are shown in Table 6. For fair comparison on each dataset, we fixed the ratio of the number of training sets of the UC-Merced dataset, the WHU-RS19 dataset, the RSSCN7 dataset, and the AID dataset to 80%, 50%, 60% and 50%, respectively. From Table 6, we can see that the features extracted via the pretrained resnet model performed better than the features extracted via the pretrained vgg model. Our proposed Hybrid-KCRC methods also achieved superior performance to classical methods, such as SVM and CRC.

- Figures about classification rates with a different number of training samples clearly illustrate that RBF and polynomial kernel functions suit RS images better. Notably, classification rate increased by 1% with the Hybrid-KCRC method from linear kernel to polynomial kernel for the AID dataset. For the RBF kernel function, the metric was different from the linear kernel. The distance between two points x and y from the linear kernel space would be closer in the RBF kernel space if x and y were close, with the contrary conclusion if x and y were far away. This makes representation more discriminative and achieves higher classification accuracy. For the polynomial kernel, the linear kernel can be a special case of the polynomial kernel ($p = 0, q = 1$). Note that kernel function $\kappa(x, y)$ can be approximated by $\phi^T(x)\phi(y)$ [45] to save time for the learning algorithm. We will adopt the approximation of kernel $\kappa(x, y)$ to save time in future works.

- In the literature [23], the author proposed a joint collaborative representation (CR) classification method that uses several complementary features to represent images, including spectral values and spectral gradient features, Gabor texture features, and DMP features. This multifeature fusion can also be implemented via our proposed method. In the literature [24], the author proposed a spatial-aware CR for hyperspectral image classification that utilized both spatial and spectral features to represent an image. The penalty term can also be added into the objective function of our proposed method.

Table 6. Comparison with different CNN pretrained models (%).

Models\Datasets	UC-Merced (0.8)	RSSCN7 (0.5)	WHU-RS19 (0.6)	AID (0.5)
CaffeNet + SVM [32]	95.02	96.24	88.25	89.53
VGG16 + SVM [32]	95.21	96.05	87.18	89.64
GoogleNet + SVM [32]	94.31	94.71	85.84	86.39
VGG19 + SVM	94.67	95.42	85.99	90.35
VGG19 + CRC	95.05	95.63	86.97	89.58
VGG19 + Hybrid-KCRC (linear)	96.17	95.68	88.16	89.93
VGG19 + Hybrid-KCRC (POLY)	96.29	96.42	89.21	91.75
VGG19 + Hybrid-KCRC (RBF)	96.26	96.5	89.17	91.82
VGG19 + Hybrid-KCRC (Hellinger)	96.33	95.82	88.47	90.35
Resnet + SVM	96.90	97.74	91.5	92.97
Resnet + CRC	97.00	98.03	92.47	92.85
Resnet + Hybrid-KCRC (linear)	97.29	98.05	92.89	92.87
Resnet + Hybrid-KCRC (POLY)	97.40	98.16	**93.11**	93.98
Resnet + Hybrid-KCRC (RBF)	**97.43**	98.13	93.07	**94.00**
Resnet + Hybrid-KCRC (Hellinger)	97.36	**98.37**	92.87	93.15

5. Conclusions

In this paper, we proposed a hybrid collaborative representation-based algorithm via embedding class-specific collaborative representation into conventional collaborative representation-based classification to improve the performance of classifying remote-sensing images. The proposed method is capable of balancing class-specific collaborative representation and shared collaborative representation. Moreover, we extended the proposed method to arbitrary kernel space to explore the nonlinear characteristics hidden in remote-sensing image features to further enhance classification

performance. Extensive experiments on four benchmark remote-sensing image datasets have demonstrated the superiority of our proposed hybrid collaborative representation algorithm.

Author Contributions: B.-D.L., W.X., J.M., and Y.L. conceived and designed the experiments; B.-D.L. and W.X. performed the experiments; Y.-J.W. analyzed the data; W.X. and B.-D.L. wrote the paper; all authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CRC	Collaborative Representation-Based Classification
CS-CRC	Class-Specific Collaborative Representation-Based Classification
RS	Remote Sensing
EO	Earth Observation
NS	Nearest Subspace
BovW	Bag-of-Visual-Words
CNNs	Convolutional Neural Networks
LRC	Linear Regression-Based Classification
POLY	Polynomial Kernel
RBF	Radial Basis Function Kernel
NN	Nearest Neighbor
SLRC	Superposed Linear Representation Classifier

References

1. Navalgund, R.R.; Jayaraman, V.; Roy, P.S. Remote sensing applications: An overview. *Curr. Sci.* **2007**, *93*, 1747–1766.
2. Liu, W.; Ma, X.; Zhou, Y.; Tao, D.; Cheng, J. p-Laplacian Regularization for Scene Recognition. *IEEE Trans. Cybern.* **2018**, 1–14. [CrossRef] [PubMed]
3. Ma, X.; Liu, W.; Li, S.; Tao, D.; Zhou, Y. Hypergraph p-Laplacian Regularization for Remotely Sensed Image Recognition. *IEEE Trans. Geosci. Remote Sens.* **2018**, *PP*, 1–11. [CrossRef]
4. Yang, X.; Liu, W.; Tao, D.; Cheng, J.; Li, S. Multiview Canonical Correlation Analysis Networks for Remote Sensing Image Recognition. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1855–1859. [CrossRef]
5. Yang, Y.; Newsam, S. Bag-of-visual-words and spatial extensions for land-use classification. In Proceedings of the Sigspatial International Conference on Advances in Geographic Information Systems, San Jose, CA, USA, 2–5 November 2010; ACM: New York, NY, USA, 2010; pp. 270–279.
6. Cheng, G.; Han, J.; Guo, L.; Liu, Z.; Bu, S.; Ren, J. Effective and Efficient Midlevel Visual Elements-Oriented Land-Use Classification Using VHR Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4238–4249. [CrossRef]
7. Zou, J.; Li, W.; Chen, C.; Du, Q. Scene Classification Using Local and Global Features with Collaborative Representation Fusion. *Inf. Sci.* **2016**, *348*, 209–226. [CrossRef]
8. Zhang, F.; Du, B.; Zhang, L. Saliency-guided unsupervised feature learning for scene classification. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2175–2184. [CrossRef]
9. Fu, M.; Yuan, Y.; Lu, X. Unsupervised feature learning for scene classification of high resolution remote sensing image. In Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing, Chengdu, China, 12–15 July 2015; pp. 206–210.

10. Han, J.; Zhang, D.; Cheng, G.; Guo, L.; Ren, J. Object Detection in Optical Remote Sensing Images Based on Weakly Supervised Learning and High-Level Feature Learning. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3325–3337. [CrossRef]

11. Li, J.; Huang, X.; Gamba, P.; Bioucas-Dias, J.M.; Zhang, L.; Benediktsson, J.A.; Plaza, A. Multiple Feature Learning for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1592–1606. [CrossRef]

12. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef]

13. Li, Y.; Huang, X.; Liu, H. Unsupervised Deep Feature Learning for Urban Village Detection from High-Resolution Remote Sensing Images. *Photogramm. Eng. Remote Sens.* **2017**, *83*, 567–579. [CrossRef]

14. Yu, Y.; Zhong, P.; Gong, Z. Balanced data driven sparsity for unsupervised deep feature learning in remote sensing images classification. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Fort Worth, TX, USA, 23–28 July 2017; pp. 668–671.

15. Chaib, S.; Yao, H.; Gu, Y.; Amrani, M. Deep feature extraction and combination for remote sensing image classification based on pre-trained CNN models. In Proceedings of the International Conference on Digital Image Processing, Hong Kong, China, 19–22 May 2017; International Society for Optics and Photonics: Bellingham, WA, USA, 2017; p. 104203D.

16. Li, W.; Du, Q. A survey on representation-based classification and detection in hyperspectral remote sensing imagery. *Pattern Recognit. Lett.* **2016**, *83*, 115–123. [CrossRef]

17. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [CrossRef]

18. Baraniuk, R.G. Compressive sensing [lecture notes]. *IEEE Signal Process. Mag.* **2007**, *24*, 118–121. [CrossRef]

19. John, W.; Allen, Y.Y.; Arvind, G.; Shankar, S.S.; Yi, M. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 210–227.

20. Zhang, L.; Yang, M.; Feng, X. Sparse representation or collaborative representation: Which helps face recognition? In Proceedings of the IEEE International Conference on Computer vision (ICCV), Barcelona, Spain, 6–13 November 2011; pp. 471–478.

21. Wu, S.; Chen, H.; Bai, Y.; Zhu, G. A remote sensing image classification method based on sparse representation. *Multimed. Tools Appl.* **2016**, *75*, 12137–12154. [CrossRef]

22. Tang, X.; Liu, Y.; Chen, J. Improvement of Remote Sensing Image Classification Method Based on Sparse Representation. *Comput. Eng.* **2016**, *42*, 254–258, 265. [CrossRef]

23. Li, J.; Zhang, H.; Zhang, L.; Huang, X.; Zhang, L. Joint Collaborative Representation With Multitask Learning for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5923–5936. [CrossRef]

24. Jiang, J.; Chen, C.; Yu, Y.; Jiang, X.; Ma, J. Spatial-Aware Collaborative Representation for Hyperspectral Remote Sensing Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 404–408. [CrossRef]

25. Deng, W.; Hu, J.; Guo, J. Face recognition via collaborative representation: Its discriminant nature and superposed representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 2513–2521. [CrossRef]

26. Liu, B.D.; Shen, B.; Wang, Y.X. Class specific dictionary learning for face recognition. In Proceedings of the IEEE International Conference on Security, Pattern Analysis, and Cybernetics (ICSPAC), Wuhan, China, 18–19 October 2014; pp. 229–234.

27. Liu, B.D.; Shen, B.; Gui, L.; Wang, Y.X.; Li, X.; Yan, F.; Wang, Y.J. Face recognition using class specific dictionary learning for sparse representation and collaborative representation. *Neurocomputing* **2016**, *204*, 198–210. [CrossRef]

28. Wang, W.; Yan, Y.; Winkler, S.; Sebe, N. Category specific dictionary learning for attribute specific feature selection. *IEEE Trans. Image Process.* **2016**, *25*, 1465–1478. [CrossRef] [PubMed]

29. Zorzi, M.; Sepulchre, R. AR Identification of Latent-Variable Graphical Models. *IEEE Trans. Autom. Control* **2016**, *61*, 2327–2340. [CrossRef]

30. Zou, Q.; Ni, L.; Zhang, T.; Wang, Q. Deep Learning Based Feature Selection for Remote Sensing Scene Classification. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2321–2325. [CrossRef]

31. Sheng, G.; Yang, W.; Xu, T.; Sun, H. High-resolution satellite scene classification using a sparse coding based multiple feature combination. *Int. J. Remote Sens.* **2012**, *33*, 2395–2412. [CrossRef]

32. Xia, G.S.; Hu, J.; Hu, F.; Shi, B.; Bai, X.; Zhong, Y.; Zhang, L.; Lu, X. AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3965–3981. [CrossRef]

33. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.

34. Fan, R.E.; Chang, K.W.; Hsieh, C.J.; Wang, X.R.; Lin, C.J. LIBLINEAR: A library for large linear classification. *J. Mach. Learn. Res.* **2008**, *9*, 1871–1874.

35. Yu, Y.; Gong, Z.; Wang, C.; Zhong, P. An Unsupervised Convolutional Feature Fusion Network for Deep Representation of Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 23–27. [CrossRef]

36. Lu, X.; Zheng, X.; Yuan, Y. Remote sensing scene classification by unsupervised representation learning. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5148–5157. [CrossRef]

37. Lazebnik, S.; Schmid, C.; Ponce, J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), New York, NY, USA, 17–22 June 2006; pp. 2169–2178.

38. Vaduva, C.; Gavat, I.; Datcu, M. Latent Dirichlet allocation for spatial analysis of satellite images. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2770–2786. [CrossRef]

39. Cheriyadat, A.M. Unsupervised feature learning for aerial scene classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 439–451. [CrossRef]

40. Penatti, O.A.; Nogueira, K.; dos Santos, J.A. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains? In Proceedings of the IEEE International Conference on Computer vision (CVPR) Workshops, Boston, MA, USA, 7–12 June 2015; pp. 44–51.

41. Hu, F.; Xia, G.S.; Hu, J.; Zhang, L. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens.* **2015**, *7*, 14680–14707. [CrossRef]

42. Lin, D.; Fu, K.; Wang, Y.; Xu, G.; Sun, X. MARTA GANs: Unsupervised representation learning for remote sensing image classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2092–2096. [CrossRef]

43. Li, P.; Ren, P.; Zhang, X.; Wang, Q.; Zhu, X.; Wang, L. Region-Wise Deep Feature Representation for Remote Sensing Images. *Remote Sens.* **2018**, *10*, 871. [CrossRef]

44. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

45. Zorzi, M.; Chiuso, A. The Harmonic Analysis of Kernel Functions. *Automatica* **2018**, *94*, 125–137. [CrossRef]