RESEARCH

Two-dimensional bidirectional principal component collaborative projection feature for SAR vehicle target recognition

Tao Tang*[iD], Chudi Zhang and Xiaoyan Zhou

*Correspondence:
tangtao@nudt.edu.cn

College of Electronic Science
and Technology, National
University of Defense
Technology, Changsha 410073,
China

Abstract

With the continuous improvement in the resolution of synthetic aperture radar (SAR), there are many problems in the interpretation of high-resolution SAR images, such as a large amount of data and low efficiency of target recognition. In this paper, a novel SAR target recognition method based on a two-dimensional bidirectional principal component cooperative representation projection feature ($(2D)^2$PCA-CRP) is proposed. First, $(2D)^2$PCA is used to project the image into the low-dimensional feature space, and the redundant information in the high-resolution SAR image is filtered while considering the spatial structure. Then, the spatial global separability feature and local structure feature of the target in the high-resolution SAR image are extracted by CRP to form the $(2D)^2$PCA-CRP feature. Finally, based on this feature, the nearest neighbour classifier is used to complete the target recognition experiments on MSTAR data. The experiments of this study are divided into three parts using standard operation condition (SOC) samples, type change samples and radar incidence angle change data. The experimental results show that the proposed feature achieves better target recognition performance in high-resolution SAR images.

Keywords: Synthetic aperture radar, Target recognition, Feature extraction, Principal component analysis, Collaborative projection

1 Introduction

Target recognition in SAR images is one of the most important applications in SAR image interpretation. SAR target recognition extracts features from target slices and classifies them with a classifier to realize the recognition of target categories, models, attributes, etc. [1, 2]. SAR target recognition technology can primarily be divided into two stages: feature extraction and classification [3, 4]. First, effective target features are extracted from training samples, and then, an appropriate classification model is selected to achieve target recognition for many test samples. Classical SAR target recognition algorithms can be roughly divided into two categories: traditional classification methods and deep learning methods. Traditional classification methods include template matching recognition, K-nearest neighbour, Bayesian [5] and support vector machine [6]. Most of these methods use prior knowledge, such as the probability

Tang *et al. EURASIP Journal on Advances in Signal Processing*   (2022) 2022:91

Page 2 of 15

distribution of SAR images, and their performances are strongly affected by imaging quality, such as SAR data noise and the settings of the classifier parameters. To apply traditional methods to SAR image target recognition more effectively, many scholars have improved feature extraction methods and classification models or introduced new methods in similar fields, such as the sparse representation-based classification (SRC) method. SRC assumes that there are sufficient training samples, which uses the training samples to form an overcomplete dictionary to linearly represent the test samples and then minimizes the reconstruction error of the test samples according to the sparsity of the coding coefficients to achieve target recognition. In recent years, SRC has shown good performance in object recognition [7], face recognition [8], and text detection. SRC was first used for SAR image target recognition by Thiagarajan et al. [7]. Yang et al. [9] proposed a kernel sparse coding classifier using pulse contour transformation to derive the compressed features of the target and clustering pairs in the kernel space and verified its recognition performance on MSTAR data. Hongliang et al. [10] separated the target region from the background and proposed a recognition method based on the sparse representation of the target local dictionary in SAR images. To take advantage of the correlation between SAR multiview targets, Zhang et al. [11] proposed an automatic target recognition method based on multiview joint sparse representation. However, it is typically difficult to obtain multiview images of objects in the same scene in practical situations. Lv et al. [12] achieved the multitask joint classification of multilevel deep features extracted by convolutional neural networks based on a multitask joint sparse representation classifier. Zhang et al. [13] used information decoupling to build a multiresolution dictionary and proposed a joint sparse representation model based on the multiresolution dictionary.

However, SRC must solve the norm minimization problem with constraints, and the computational complexity of this process is too high. Zhang et al. [14] used the L2 norm to replace the L1 norm in the SRC model to make the model have a closed solution and proposed a collaborative representation-based classification (CRC) method. The recognition accuracy of CRC and SRC is similar, but the number of computations is markedly lower. Concurrently, CRC can find the optimal training sample representation coefficient of the reconstructed test sample, which enhances the correlation between the test sample and similar training samples. Zhang et al. [14] confirmed that good classification can be achieved under experimental conditions with few samples. The robustness of its classification has been verified in the fields of face recognition [15, 16], hyperspectral image classification [17], etc. In the field of SAR target recognition, Geng et al. [18] proposed a joint collaborative representation method based on the Wishart distance for polarimetric SAR image classification. Zhang et al. [19] input the 3 types of features of SAR target slices into SRC and CRC to obtain 6 category labels and obtained the final recognition result according to the Yess decision fusion. Based on the model of Zhang et al., Wang et al. [20] added multiple adjacent multiview samples of the current sample as the input of the model and obtained the label of the current test sample under this feature. Finally, the tags are fused by the voting method to obtain the final recognition result.

The deep learning method [21] primarily uses the convolutional neural network (CNN) to complete the automatic extraction of target features in the image and then uses the fully connected layer or other classification models to achieve automatic target

recognition. CNN models represented by AlexNet [22], VGGNet [23], GoogLeNet [24], ResNet [25], DenseNet [26] and SENet [27] have appeared in deep learning methods. CNNs have achieved excellent results in the field of optical image target recognition, and many scholars have also introduced them into SAR image target recognition. Chen S et al. introduced CNN into the SAR field and achieved recognition accuracy far exceeding traditional methods on the MSTAR SOC dataset [28]. Zhang et al. used a ResNet network to extract fixed spatial scatter features from each azimuth image and trained the network using joint supervision of softmax loss and centre loss [29]. Treating SAR images as sequences, Ruihang Xu et al. proposed a target classification method for SAR images based on a spatiotemporal holistic convolutional network (STEC-Net), which achieved better results than basic convolutional neural networks [30]. Currently, a large number of studies of CNN-based SAR image target recognition [31–37, 39] have reported better performances than traditional learning algorithms, but deep learning methods have stricter hardware requirements and are difficult to implement on edge devices. In addition, deep learning methods rely heavily on large-scale and sample-balanced datasets and face problems such as fewer training samples, optimal design of deep models and longer training time.

To develop traditional learning processes into deep learning processes, the primary required change lies in the automatic extraction of effective features. Feature extraction determines the upper limit of the performance of the target recognition algorithm, and the classification model only makes the algorithm approach the upper limit continuously. Therefore, the study of effective feature extraction technology is important for target recognition.

In the field of image processing, feature extraction primarily reduces the dimensionality of image data and obtains a series of features representing the image. In addition to the sparse dictionary method mentioned in the traditional method, it also includes methods such as spatial transformation. Spatial transformation uses mathematical methods such as matrix decomposition and optimization to obtain global or local structural features in the corresponding low-dimensional space [5], which can effectively transform a sample of high-dimensional data into low-dimensional features. This transformation has two advantages for target recognition: one is to remove the correlation between samples and increase the feature difference between samples of different categories; the other is to reduce the dimension of the input classifier by reducing the dimension of the sample features of the input classifier, which can reduce the computational load of the classification process.

The two-dimensional bidirectional principal component analysis feature named $(2D)^2PCA$[6] and collaborative representation-based projections (CRP) feature [7] are two effective methods of data dimension reduction in spatial transformation. $(2D)^2PCA$-CRP is a two-dimensional PCA that can analyse the principal component features of a sample set that is composed of two-dimensional samples. Compared with PCA, $(2D)^2PCA$ reduces the feature dimension and extracts more effective feature information. However, $(2D)^2PCA$ only extracts global structural features of an image and ignores local structural information in the target that may improve recognition. CRP is an unsupervised spatial transformation feature. Similar to Fisher's criterion, CRP must find an optimal projection matrix to maximize the global separability of the sample based on the graph and minimize the local

compactness. It is necessary to convert the SAR image containing redundant information into a one-dimensional vector before extracting the CRP feature, which destroys the spatial structure of the SAR image. Therefore, the extracted CRP features may not be able to characterize the real target characteristics, resulting in a decrease in the performance of SAR target recognition.

To solve the problem of inefficient interpretation of massive high-resolution SAR data, this study combines $(2D)^2$PCA and CRP to develop a spatial transformation feature called the $(2D)^2$PCA-CRP feature, which effectively reduces the dimension of target features and improves the recognition accuracy of targets using the intensity information of SAR images. First, $(2D)^2$PCA is used to project the image to the low-dimensional feature space and filter out the interference redundant information in the high-resolution SAR image without destroying the spatial structure. Then, CRP is used to extract the global structural features and local structural features of the target in the high-resolution SAR image.

2 Two-dimensional bidirectional principal component collaborative projection feature

2.1 Two-dimensional bidirectional principal component analysis

Principal component analysis (PCA) is a commonly used data analysis method that is often used for dimensionality reduction of high-dimensional data and can be used to extract the primary feature components of the data. PCA is often used as a pre-processing method in computer vision and signal processing to reduce the dimension of big data. $(2D)^2$PCA is developed based on PCA. Compared with PCA, $(2D)^2$PCA does not need to vectorize the image matrix but directly processes it. Concurrently, $(2D)^2$PCA performs dimensionality reduction from the row and column directions to eliminate the correlation between rows and columns. Given a training sample set $\mathbf{X} \in \mathbb{R}^{m \times n \times N}$, where $\mathbf{X}$ is a three-dimensional sample tensor composed of N images of size $m \times n$, the specific steps of $(2D)^2$PCA feature extraction are as follows.

To perform row mapping of $\mathbf{Y_1} = \mathbf{XA}$, where $\mathbf{A}$ is a projection matrix composed of a set of orthogonal bases, $\mathbf{Y_1}$ is the result of the original sample $\mathbf{X}$ mapped by $\mathbf{A}$, and the overall scatter matrix $\mathbf{G}_{t,r}$ corresponding to the row transformation is in (1):

$$\mathbf{G}_{t,r} = \frac{1}{N} \sum_{r=1}^{N} \sum_{i=1}^{m} \left(\mathbf{X}_r^i - \overline{\mathbf{X}^i}\right)^{\mathrm{T}} \left(\mathbf{X}_r^i - \overline{\mathbf{X}^i}\right) \tag{1}$$

To perform column mapping of $\mathbf{Y_2} = \mathbf{BX}$, where $\mathbf{B}$ is a projection matrix composed of a set of orthogonal bases, $\mathbf{Y_2}$ is the result of the original sample $\mathbf{X}$ mapped by $\mathbf{B}$, and the overall scatter matrix $\mathbf{G}_{t,l}$ corresponding to the column transformation is shown in (2):

$$\mathbf{G}_{t,l} = \frac{1}{N} \sum_{r=1}^{N} \sum_{j=1}^{n} \left(\mathbf{X}_r^j - \overline{\mathbf{X}^j}\right) \left(\mathbf{X}_r^j - \overline{\mathbf{X}^j}\right)^{\mathrm{T}} \tag{2}$$

where $\mathbf{X}_r^i$ and $\mathbf{X}_r^j$ represent the ith row vector and the jth column vector of the rth sample, respectively; $\overline{\mathbf{X}^i}$ and $\overline{\mathbf{X}^j}$ represent the ith row vector and the jth column vector of the sample mean $\overline{\mathbf{X}}$, respectively. To obtain the optimal projection of the training sample $\mathbf{X} \in \mathbb{R}^{m \times n \times N}$, the eigenvalues and eigenvectors of the scatter matrices $\mathbf{G}_{t,r}$ and $\mathbf{G}_{t,l}$ are

calculated, and the eigenvectors corresponding to the largest d_t and d_r eigenvalues of each are taken to form the projection matrices $\mathbf{A}$ and $\mathbf{B}$. Therefore, the original sample joint mapping form is shown in (3):

$$\mathbf{Y} = \mathbf{B}^{\mathrm{T}}(\mathbf{X} - \overline{\mathbf{X}})\mathbf{A} \tag{3}$$

where $\mathbf{Y_2}$ is the $(2D)^2$PCA feature of the original sample set after spatial projection dimensionality reduction; thus, the feature dimension of the sample project is $d \in \mathbb{R}^{t \times r}$. This process retains the original sample data and spatial structure information as much as possible but also markedly reduces the dimension of the feature space.

2.2 Cooperative representation-based projection

CRP is an unsupervised discriminant projection method based on ℓ_2 norm regularized least squares. The core idea of CRP is to use all samples to represent the specified sample to calculate the edge weight of the specified sample and other samples through collaborative representation theory. In addition, when obtaining the projection matrix, the local information and global information of the sample are considered concurrently, and the local information is minimized. The optimal projection matrix is found by minimizing local compactness and maximizing total separability.

We assume that the training sample set is $\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}] \in \mathbb{R}^{m \times n}$, where $\mathbf{x_i}$ is a column vector composed of the m-dimensional features of the ith sample. First, we construct an ℓ_2 graph according to collaborative representation theory, and the weight $w_i \in \mathbb{R}^{n \times 1}$ of the sample $\mathbf{x_i}$ reconstructed from the ℓ_2 graph is:

$$w_i = \arg \min_{w_i} \left\{ \|\mathbf{x}_i - \mathbf{X}w_i\|_2^2 + \lambda \|w_i\|_2^2 \right\} \tag{4}$$

where λ is the regularization parameter, and $w_{ij}(i \neq j)$ represents the contribution made by sample $\mathbf{x}_j$ when reconstructing sample $\mathbf{x}_i$. We let the derivative of Eq. (4) with respect to w_{ij} be 0, and it is easy to find its closed solution as follows:

$$w_i = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \lambda \mathbf{I}\right)^{-1} \mathbf{X}^{\mathrm{T}} \cdot \mathbf{x}_i \tag{5}$$

According to the ℓ_2 diagram, CRP defines local compactness $\mathbf{J}_L$ and global separability $\mathbf{J}_T$ as:

$$\mathbf{J}_L = \sum_{i=1}^{n} \left\| \mathbf{P}^{\mathrm{T}}\mathbf{x}_i - \sum_{j-1}^{n} w_{ij}\mathbf{P}^{\mathrm{T}}\mathbf{x}_j \right\|_2^2 = \mathbf{P}^{\mathrm{T}}\mathbf{S}_L\mathbf{P} \tag{6}$$

$$\mathbf{J}_T = \sum_{i=1}^{n} \left\| \mathbf{P}^{\mathrm{T}}\mathbf{x}_i - \mathbf{P}^{\mathrm{T}}\bar{\mathbf{x}} \right\|_2^2 = \mathbf{P}^{\mathrm{T}}\mathbf{S}_T\mathbf{P} \tag{7}$$

where the local scatter matrix S_L and the global scatter matrix S_T S are, respectively, expressed as:

$$\mathbf{S}_L = \mathbf{X}(\mathbf{I} - \mathbf{W} - \mathbf{W}^{\mathrm{T}} + \mathbf{W} \cdot \mathbf{W}^{\mathrm{T}})\mathbf{X}^{\mathrm{T}} \tag{8}$$

$$\mathbf{S}_T = \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\mathrm{T}} \tag{9}$$

where $\mathbf{I}$ is identity matrix, $\mathbf{W}$ is the matrix of w_{ij}.

To obtain the best projection matrix $\mathbf{P}$, it is necessary to solve two optimization problems of minimizing local compactness and maximizing total separability concurrently. Therefore, the final optimization function of CRP can be expressed as:

$$\mathbf{J}(\mathbf{P}) = \arg \min_{\mathbf{P}} \frac{\mathbf{P}^{\mathrm{T}} \mathbf{S}_L \mathbf{P}}{\mathbf{P}^{\mathrm{T}} \mathbf{S}_T \mathbf{P}} = \arg \max_{\mathbf{P}} \frac{\mathbf{P}^{\mathrm{T}} \mathbf{S}_T \mathbf{P}}{\mathbf{P}^{\mathrm{T}} \mathbf{S}_L \mathbf{P}} \tag{10}$$

Thus, the projection matrix $\mathbf{P}$ is the first d largest nonzero eigenvalue corresponding to $(\mathbf{S}_L)^{-1}\mathbf{S}_T\mathbf{P} = \lambda\mathbf{P}$. The low-dimensional CRP feature of the final sample is:

$$\mathbf{X}' = \mathbf{P}^{\mathrm{T}} \cdot \mathbf{X} \tag{11}$$

2.3 (2D)²PCA-CRP features

As mentioned in Sect. 2.1, $(2D)^2$PCA is a low-dimensional feature extraction method for two-dimensional data. When it is used in high-resolution SAR images, by compressing the two-dimensional image row and column simultaneously, the redundant information in the SAR image can be removed without destroying the two-dimensional matrix structure of the image, effectively retaining the image. The global scattering information of the target slice is typically removed from the information redundancy caused by noise or background, and the characteristic information related to the target area is retained. However, the disadvantage of $(2D)^2$PCA is that it only focuses on global features and ignores the local structure information in the target that is beneficial for recognition.

CRP is an unsupervised spatial transformation feature. The advantage of CRP is that it considers global and local features concurrently. Similar to Fisher's criterion, the process of CRP to obtain the best projection matrix is the process of simultaneously optimizing the global separability and local compactness based on the ℓ_2 graph. Therefore, the extracted CRP features can effectively represent the global and local features of the SAR image. However, when extracting CRP features, the two-dimensional SAR image matrix containing redundant information must be stretched into a one-dimensional vector, which destroys the spatial structure of the SAR image. After stretching, the redundant information is interlaced and mixed with the effective information, and the extracted CRP features can no longer characterize the effective features of the original SAR image target, which leads to the degradation of the SAR target recognition performance.

Therefore, we combine $(2D)^2$PCA and CRP to design $(2D)^2$PCA-CRP features to effectively reduce the target feature dimension and improve the SAR target recognition rate. First, we use $(2D)^2$PCA to project the high-resolution SAR image into the low-dimensional feature space, remove the interference redundant information in the high-resolution SAR image without destroying the spatial structure, and then use CRP to extract the redundant SAR image. With regard to the characteristic information of the target, the recognition of the target is finally achieved.

3 Methods

In this paper, the SAR vehicle target recognition algorithm flow based on the (2D)2PCA-CRP feature mentioned in Sect. 2.3 is shown in Fig. 1, which effectively reduces the dimension of target features and improves the recognition accuracy of targets in SAR images.

First, $(2D)^2PCA$ is used to project the image to the low-dimensional feature space and filter out the interference redundant information in the high-resolution SAR image without destroying the spatial structure. Then, CRP is used to extract the global structural features and local structural features of the target in the high-resolution SAR image.
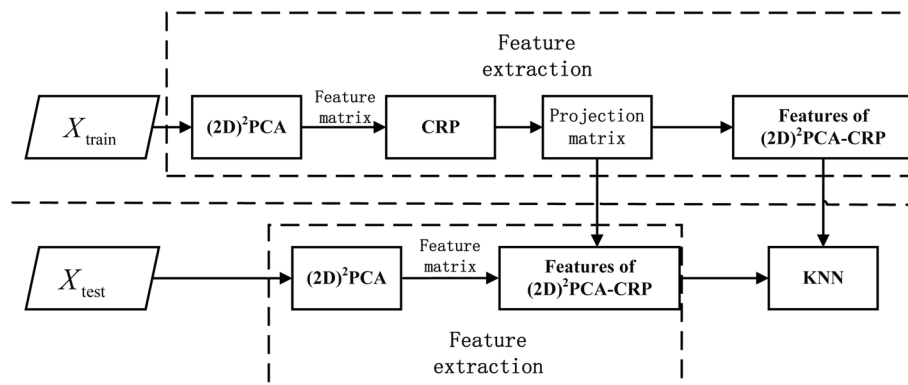
The steps of this method are as follows:

Step 1: (2D) ^2PCA feature matrix extraction. Considering that the original SAR image contains redundant information that interferes with target recognition, the original SAR image sample $\mathbf{X}$ is converted to a lower data dimension using the $(2D)^2PCA$ (2D) 2PCA subspace before CRP feature extraction and using Eqs. (1)-(3). We then denote the sample feature matrix after dimensionality reduction as $\mathbf{X}_{(2D)^2PCA}$.

Step 2: CRP feature projection matrix calculation. First, we establish the ℓ_2 graph about the sample space $\mathbf{X}_{(2D)^2PCA}$ according to Eqs. (4)-(5). Then define the local compactness $\mathbf{J}_L$ and global separability $\mathbf{J}_T$ of the ℓ_2 graph according to Eqs. (6)-(7). Finally the optimization problem of Eq. (10) obtains the characteristic projection matrix $\mathbf{P}$ of CRP by solving Eq. (10).

Step 3: (2D) 2PCA-CRP feature extraction. The CRP characteristics of training samples $\mathbf{X}_{\text{train}}$ and test samples $\mathbf{X}_{\text{test}}$ are expressed as $\mathbf{X}'_{\text{train}} = \mathbf{P}^{\mathrm{T}} \cdot \mathbf{X}_{\text{train},(2D)^2PCA}$ and $\mathbf{X}'_{\text{test}} = \mathbf{P}^{\mathrm{T}} \cdot \mathbf{X}_{\text{test},(2D)^2PCA}$.

Step 4: Nearest neighbour classification. The (2D) 2PCA-CRP features of samples $\mathbf{X}'_{\text{train}}$ and $\mathbf{X}'_{\text{test}}$ are sent to the nearest neighbour classifier, and the test sample is assigned a category label based on the Euclidean distance.

Finally, the complexity of $(2D)^2PCA$-CRP is analysed as following. Suppose there are M m by n training images, the number of projection vectors in PCA, 2DPCA and alternative 2DPCA is p, d and q. Then the compression ratios of PCA, 2DPCA, alternative 2DPCA and $(2D)^2PCA$ are computed as $Mmn/(Mp + mnp)$, $Mmn/(Mmd + nd)$, $Mmn/(Mmq + nq)$ and $Mmn/(Mdq + nd + mq)$, respectively [38].

Fig. 1 Flowchart of the SAR target recognition algorithm based on (2D)^{2}PCA-CRP features

Tang *et al. EURASIP Journal on Advances in Signal Processing*   (2022) 2022:91

Page 8 of 15

4 Results and discussion

Some experiments are performed to demonstrate the effectiveness of the proposed method in SAR vehicle target recognition under several condition and the results are discussed in this section.

The experimental data adopt the moving and stationary target acquisition and recognition named MSTAR published by Sandia Laboratory supported by DARPA. Whether at home or abroad, the research on SAR image target recognition is basically based on this dataset. All images in this dataset are high-resolution spotlight sensor starlos imaging in the X-band that were created in the HH polarization working mode and have an image resolution of 0.3 m $\times$ 0.3 m. The coverage of the imaging azimuth is 0°–360°. A target consists of one type of rocket launcher 2S1; four types of armoured vehicles, including BMP2, BRMD2, BTR60 and BTR70; one type of bulldozer D7; two types of tanks T72 and T62; and one type of air defence unit ZSU23/4. In addition, there are several types of targets with submodels and differences in specific model configurations. In addition, different types of target samples have different sizes, including 128 $\times$ 128 pixels, 158 $\times$ 158 pixels, 178 $\times$ 178 pixels and 192 $\times$ 192 pixels. Therefore, in the experiments, these samples must be pre-processed. For all types of SAR vehicle target samples, the centre interception or zero filling operation is performed, and the sample slice size is uniformly adjusted to 128 $\times$ 128 pixels.

4.1 SOC samples

In this study, the training set and test set under the SOC conditions contain the MSTAR data of 10 types of targets with elevation angles of 17° and 15°. The comparison algorithms are CRP, PCA, (2D)^{2}PCA, (2D)^{2}PCA -CRP and CA-MCNN [39]. CA-MCNN is a multiscale convolutional neural network (CNN) based on component analysis (CA-MCNN) for synthetic aperture radar (SAR) automatic target recognition (ATR). The component information of a target is robust to the local variations of the target, which is not made the best of by traditional CNN-based methods. Specific experimental data settings are shown in Table 1. The training samples and test samples of each type include all submodels, which increases the difficulty of recognition compared to using only a single model. The confusion matrix obtained by the algorithm in this paper is shown in Table 2. The average recognition rate of the 10 types of targets is 95.63%. The average recognition rates obtained based on the other three spatial transformation features are shown in Table 3. The average recognition rates based on the PCA feature and CRP feature are 72.37% and 73.58%, respectively, which are similar. The (2D)^{2}PCA feature recognition rate is marginally higher but is far lower than the recognition accuracy based on (2D)^{2}PCA-CRP features because PCA and CRP both extract features based on the one-dimensional form of the image and lose the spatial structure information that can characterize the detailed features of the target. However, with multitype target recognition, it is particularly important to accurately describe the detailed structural characteristics of the target. Therefore, the (2D)^{2}PCA-CRP feature has marked advantages when used for multiclass MSTAR vehicle target recognition under SOC conditions. It is worth noting that this method uses the multiscale CNN network structure and integrates the target scattering

Table 1 SOC sample division

Type	Train set	Test set
2S1	299	274
BMP2	233 (sn_9563)	195 (sn_9563)
	232 (sn_9566)	196 (sn_9566)
	233 (sn_c21)	196 (sn_c21)
BRDM2	298	274
BTR60	256	195
BTR70	233	196
D7	299	274
T62	299	273
T72	232 (sn_132)	196 (sn_132)
	231 (sn_812)	195 (sn_812)
	228 (sn_s7)	191 (sn_s7)
ZIL131	299	274
ZSU23/4	299	274

Table 2 Confusion matrix of the presented method

Type	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU23/4
2S1	0.8650	0	0	0	0.1314	0	0.0036	0	0	0
BMP2	0	0.9983	0	0	0	0	0	0.0017	0	0
BRDM2	0	0	0.8175	0	0.1825	0	0	0	0	0
BTR60	0	0	0	0.7946	0.2051	0	0	0	0	0
BTR70	0	0	0	0	1	0	0	0	0	0
D7	0	0	0	0	0.0036	0.9891	0.0036	0	0	0.0036
T62	0	0	0	0	0.0110	0.0037	0.9744	0	0	0.0110
T72	0	0	0	0	0	0	0	1	0	0
ZIL131	0.0036	0	0	0	0.0036	0	0	0	0.9927	0
ZSU23/4	0	0	0	0	0	0	0	0	0	1

Table 3 Comparison of recognition rates of various algorithms

Algorithm	PCA	$(2D)^2$PCA	CRP	$(2D)^2$PCA-CRP	CA-MCNN
Average recognition rate	0.7237	0.8820	0.7358	0.9563	0.9861

structure information, and its recognition accuracy is slightly higher than that of this method.

4.2 Type change samples

In specific classification tasks, similar targets typically appear in different models of variants or have different mounts installed. The primary structure of these variants is similar, but there are differences in specific details, such as whether a tank is equipped with fuel tanks, whether an antenna is deployed, whether an armoured vehicle target is equipped with artillery, etc. Therefore, the model recognition of the target brand of the vehicle is important. In this section, two sets of experiments are described. Experiment one

considers four types of targets and a total of 8 specific models, and experiment two considers five submodels of T72 targets to complete two sets of SAR vehicle target model classification experiments.

4.3 Recognition of 4 types of targets

In this section, the training samples and test samples are 14 types of vehicles at elevation angles of 17° and 15°. As shown in Table 4, the BMP2 and T72 targets each contain three different variants. Pairwise methods have similar scattering performance on SAR images [4]. The average recognition rate based on the four spatial transformation features is shown in Table 6. The average recognition rate of the 8 submodels based on the four features from high to low is $CRP > (2D)^2PCA\text{-}CRP > (2D)^2PCA > PCA$. The average recognition accuracy based on $(2D)^2PCA\text{-}CRP$ features is 93.88%, and the performance is marginally lower than when SOC conditions exist. Specifically, from the confusion matrix obtained by the target recognition algorithm based on the $(2D)^2PCA\text{-}CRP$ feature in Table 5, the primary reason for the decline in recognition rate is that the three submodels of the BMP2 target produce serious misclassifications. However, the BMP2 target will not be mistakenly classified into the other three types of targets and thus has almost no impact on the performance of the SOC's large-type target recognition experiment. In general, under the current experimental conditions, the $(2D)^2PCA\text{-}CRP$ feature achieves good performance when recognizing the 14 types of SAR targets (Table 6).

4.4 T72 target type classification

To examine how the proposed method performs with specific model changes in the T72 target, the experimental conditions are set as shown in Table 7. The training samples and test samples are the A32, A62, A63, A64 and SN_S7 submodels of the T72 target at elevation angles of 17° and 15°. The confusion matrix of the algorithm in this paper for the recognition results of the five submodels of the T72 target is shown in Table 8. The recognition rates of each type of submodel are 99.27%, 97.81%, 94.89%, 95.26% and

Table 4 Sample division of 14 types of target model identification

Type	Training set	Test set
BMP2 (sn_9563)	233	195
BMP2 (sn_9566)	232	196
BMP2 (sn_c21)	233	196
BTR60	256	195
T62	299	273
T72 (sn_132)	232	196
T72 (sn_812)	231	195
T72 (sn_s7)	228	191
2S1	299	276
BRDM2	298	274
BTR70	233	196
D7	299	274
ZIL131	299	274
ZSU23/4	299	274

Table 5 Confusion matrix of the presented method for 14 types of targets

Type	BMP2 (sn_9563)	BMP2 (sn_9566)	BMP2 (sn_c21)	BTR60	T62	T72 (sn_132)	T72 (sn_812)	T72 (sn_s7)	2S1	BRDM2	BTR70	D7	ZIL131	ZSU23/4
BMP2 (sn_9563)	0.8923	0.0359	0.0256	0.0308	0.0103	0	0.0051	0	0	0	0	0	0	0
BMP2 (sn_9566)	0.0051	0.9592	0.0153	0.0051	0.0102	0	0	0.0051	0	0	0	0	0	0
BMP2 (sn_c21)	0.0255	0.0255	0.9133	0.0051	0.0051	0.0102	0.0102	0	0	0	0	0	0	0
BTR60	0	0	0	1	0	0	0	0	0	0	0	0	0	0
T62	0	0	0	0	1	0	0	0	0	0	0	0	0	0
T72 (sn_132)	0	0	0	0	0.0051	0.9592	0.0153	0.0204	0	0	0	0	0	0
T72 (sn_812)	0	0	0	0	0.0051	0.0051	0.9846	0.0051	0	0	0	0	0	0
T72 (sn_s7)	0	0	0	0	0.0419	0.0157	0.0628	0.8796	0	0	0	0	0	0
2S1	0	0	0	0	0	0.0013	0.0012	0.0011	0.9891	0	0	0	0.0073	0
BRDM2	0	0	0	0	0	0.0054	0.0032	0.0023	0	0.9854	0	0	0	0.0234
BTR70	0.0051	0.0051	0.0051	0.0459	0	0	0	0	0	0	0.9286	0	0.0102	0
D7	0	0	0	0	0	0.0031	0.0021	0.0021	0	0	0	0.9927	0	0
ZIL131	0	0	0	0	0	0.0031	0.0021	0.0021	0	0	0	0	0.9891	0.0036
ZSU23/4	0	0	0	0	0	0.0036	0	0	0	0	0	0	0	0.9964

Tang *et al. EURASIP Journal on Advances in Signal Processing*   (2022) 2022:91

Page 12 of 15

Table 6 Comparison of recognition rates of various algorithms for 14 types of targets

Algorithm	PCA	(2D)^{2}PCA	CRP	(2D)^{2}PCA-CRP
Average recognition rate	0.8905	0.9143	0.9435	0.9388

Table 7 Data division of 5 specific types of T72 targets

Type	Training set	Test set
A32	298	274
A62	299	274
A63	299	274
A64	299	274
SN_S7	228	191

Table 8 Confusion matrix of the classification algorithm for 5 specific types of T72 targets

Type	A32	A62	A63	A64	SN_S7
A32	0.9927	0.0036	0	0.0036	0
A62	0.0036	0.9781	0.0146	0.0036	0
A63	0.0109	0.0292	0.9489	0.0073	0.0036
A64	0.0073	0.0255	0.0073	0.9526	0.0073
SN_S7	0	0	0	0.0105	0.9895

Table 9 Comparison of recognition rates of various algorithms for 5 specific types of T72 targets

Algorithm	PCA	(2D)^{2}PCA	CRP	(2D)^{2}PCA-CRP
Average recognition rate	0.8042	0.8539	0.7428	0.9713

Table 10 Data division under the change in elevation angle

Type	Training set (17°)	Test set (30°)
2S1	299	288
BRDM2	298	287
ZSU23/4	299	288

98.95%, and the average recognition rate is 97.13%. These results are markedly higher than the recognition results of the other three features in Table 9, particularly the two features based on one-dimensional vectors, which is consistent with the previous theoretical analysis.

4.5 Change in elevation angle

Affected by SAR imaging characteristics, there are marked differences in the imaging of samples at different elevation angles, and the greater the elevation angle difference is, the smaller the similarity between samples. To verify the robustness of the proposed algorithm under the change of the elevation angle, this section sets up the target recognition

Table 11 Confusion matrix of the classification algorithm under the change in elevation angle

Type	2S1	BRDM2	ZSU23/4
2S1	0.9931	0.0069	0
BRDM2	0.0035	0.9965	0
ZSU23/4	0.0069	0.0069	0.9861

Table 12 Comparison of recognition rates of various algorithms under the change of elevation angle

Algorithm	PCA	$(2D)^2$PCA	CRP	$(2D)^2$PCA-CRP
Average recognition rate	0.9779	0.9623	0.4276	0.9919

experiment under the large elevation angle. The training set still uses the samples at 17°, and the test set changes from 15° to the samples at the larger 30° pitch angle. Target types include three types, 2S1, BRDM2 and ZSU23/4, and the specific sample size of each category is shown in Table 10. The confusion matrix of $(2D)^2$PCA-CRP features at large pitch angles is shown in Table 11. The recognition rates of the three types of targets are 99.31%, 99.65% and 98.61%. All have achieved good recognition results. Compared with the recognition rate of the other three features shown in Table 12, the $(2D)^2$PCA-CRP feature achieves a recognition rate of 99.19%, which is higher than the two features of PCA and $(2D)^2$PCA. However, the recognition rate of the CRP feature for the 2S1 and BRDM2 targets is only 16.67% and 11.5%, respectively, which is strongly affected by the change in the pitch angle. The experimental results thus show that compared with the original CRP feature, the improved $(2D)^2$PCA-CRP feature is insensitive to changes in the pitch angle.

5 Conclusion

To solve the problem of a large amount of high-resolution image data and low interpretation efficiency, this study uses three types of PCA, $(2D)^2$PCA and CRP to effectively reduce the spatial transformation characteristics of the SAR image feature space. In addition, we analysed the proposed method's performance in a high-resolution SAR target recognition task. Then, $(2D)^2$PCA and CRP were combined to develop a spatial transformation feature that effectively reduces the feature space dimension and improves the SAR target recognition rate. First, $(2D)^2$PCA is used to project the image into a low-dimensional feature space. While considering the spatial structure, the interference redundant information in the high-resolution SAR image was filtered out. Then, CRP was used to extract the spatial global separability feature and local structure feature of the target in the high-resolution SAR image. Finally, based on the $(2D)^2$PCA-CRP feature, the nearest neighbour classifier was used to complete the target recognition task. Experiments were performed under four experimental conditions in the MSTAR dataset, including SOC, 8 types of target changes, T72 model changes and pitch angle changes. (2D) 2PCA-CRP features achieved average recognition rates of 95.63%, 93.89%, 97.13%, and 99.19%, which were markedly higher than those of the three features of PCA, $(2D)^2$PCA and CRP.

Tang *et al. EURASIP Journal on Advances in Signal Processing*   (2022) 2022:91

Page 14 of 15

When extracting spatial transformation features, this method only removes the correlation and reduces the feature dimension of the current sample, and does not use the category information of the sample. Similar targets usually have similar target characteristics, while different targets have obvious differences in target characteristics. Therefore, the next research will use the sample category information to extract more effective spatial transformation features.

Abbreviations

SAR   Synthetic aperture radar
PCA   Principal component analysis
CRP   Collaborative representation projection
SOC   Standard operating condition
EOC   Extended operating condition
CRC   Collaborative representation-based classification
CNN   Convolutional neural network

Availability of data and materials
Please contact author for data requests.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
The picture materials quoted in this article have no copyright requirements, and the source has been indicated.

Competing interests
The authors declare no competing interests.

References
1. F. Biondi, C. Clemente, D. Orlando, An eigenvalue-based approach for structure classification in polarimetric SAR images. IEEE Geosci. Remote Sens. Lett. **17**(6), 1003–1007 (2019)
2. L. Pallotta, A. De Maio, D. Orlando, A robust framework for covariance classification in heterogeneous polarimetric SAR images and its application to L-band data. IEEE Trans. Geosci. Remote Sens. **57**(1), 104–119 (2018)
3. L.M. Novak, G.J. Owirka, A.L. Weaver, Automatic target recognition using enhanced resolution SAR data. IEEE Trans. Aerosp. Electron. Syst. **35**(1), 157–175 (1999)
4. P. Tait, *Introduction to Radar Target Recognition* (IET, London, 2005)
5. C. Li, G. Liu, Block sparse Bayesian learning over local dictionary for robust SAR target recognition. Int. J. Opt. **2020** (2020)
6. C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST). **2**, 1–27 (2011)
7. J.J. Thiagarajan, K.N. Ramamurthy, P. Knee, A. Spanias, V. Berisha. Sparse representations for automatic target classification in SAR images. In: *2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)* (2010) p. 1–4. https://doi.org/10.1109/ISCCSP.2010.5463416
8. Y. Peng, L. Li, S. Liu, J. Li, X. Wang, Extended sparse representation-based classification method for face recognition. Mach. Vis. Appl. **29**, 991–1007 (2018)
9. S. Yang, Y. Ma, M. Wang, D. Xie, Y. Wu, L. Jiao, Compressive feature and kernel sparse coding-based radar target recognition. IET Radar Sonar Navig. **7**, 755–763 (2013)

10. H. Hongliang, B. Yonglei, L. Wei, F. Fan, W. Jianhua, WITHDRAWN: SAR image target recognition method based on sparse representation of local dictionary (2021)
11. H. Zhang, N.M. Nasrabadi, T.S. Huang, Y. Zhang, Joint sparse representation based automatic target recognition in SAR images. In: *Algorithms for Synthetic Aperture Radar Imagery XVIII* Vol. 8051 (International Society for Optics and Photonics, 2011), p. 805112
12. J. Lv, Exploiting multi-level deep features via joint sparse representation with application to SAR target recognition. Int. J. Remote Sens. **41**, 320–338 (2020)
13. Z. Zhang, S. Liu, Joint sparse representation for multi-resolution representations of SAR images with application to target recognition. J. Electromagn. Waves Appl. **32**, 1342–1353 (2018)
14. L. Zhang, M. Yang, X. Feng, Sparse representation or collaborative representation: Which helps face recognition? in *Proceedings of 2011 International Conference on Computer Vision* (IEEE, 2011), p. 471–478
15. D.M. Vo, S.W. Lee, Robust face recognition via hierarchical collaborative representation. Inf. Sci. **432**, 332–346 (2018)
16. B. Liu, L. Jing, J. Li, J. Yu, A. Gittens, M.W. Mahoney, Group collaborative representation for image set classification. Int. J. Comput. Vis. **127**, 181–206 (2019)
17. H. Su, B. Zhao, Q. Du, P. Du, Z. Xue, Multifeature dictionary learning for collaborative representation classification of hyperspectral imagery. IEEE Trans. Geosci. Remote Sens. **56**, 2467–2484 (2018)
18. J. Geng, H. Wang, J. Fan, X. Ma, B. Wang, Wishart distance-based joint collaborative representation for polarimetric SAR image classification. IET Radar Sonar Navig. **11**, 1620–1628 (2017)
19. X. Zhang, Z. Tan, Y. Wang, SAR target recognition based on multi-feature multiple representation classifier fusion. J. Radars (2017)
20. J. Wang, X. Zhang, M. Liu, X. Tan, SAR target classification using multi-aspect multi-feature collaborative representation. Remote Sens. Lett. **11**, 720–729 (2020)
21. X. Feng, W. Haipeng, J. Yaqiu, Deep learning as applied in SAR target recognition and terrain classification. **6**, 136–148 (2017)
22. A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. **25**(2) (2012)
23. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 (2014)
24. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions. In *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), p. 1–9
25. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. In *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), p. 770–778
26. G. Huang, Z. Liu; L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks. In *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), p. 4700–4708
27. J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks. In *The Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), p. 7132–7141
28. S. Chen, H. Wang, F. Xu, Y.Q. Jin, Target classification using the deep convolutional networks for SAR images. IEEE Trans. Geosci. Remote Sens. **54**, 4806–4817 (2016)
29. F. Zhang, Z. Fu, Y. Zhou, W. Hu, W. Hong, Multi-aspect SAR target recognition based on space-fixed and space-varying scattering feature joint learning. Remote Sens. Lett. **10**, 998–1007 (2019)
30. R. Xue, X. Bai, F. Zhou, Spatial–temporal ensemble convolution for sequence SAR target classification. IEEE Trans. Geosci. Remote Sens. **59**, 1250–1262 (2020)
31. J.H. Cho, C.G. Park, Multiple feature aggregation using convolutional neural networks for SAR image-based automatic target recognition. IEEE Geosci. Remote Sens. Lett. **15**, 1882–1886 (2018)
32. F. Gao, T. Huang, J. Sun, J. Wang, A. Hussain, E. Yang, A new algorithm for SAR image target recognition based on an improved deep convolutional neural network. Cogn. Comput. **11**, 809–824 (2019)
33. P. Zhao, K. Liu, H. Zou, X. Zhen, Multi-stream convolutional neural network for SAR automatic target recognition. Remote Sens. **10**, 1473 (2018)
34. J. Ding, B. Chen, H. Liu, M. Huang, Convolutional neural network with data augmentation for SAR target recognition. IEEE Geosci. Remote Sens. Lett. **13**, 364–368 (2016)
35. D. A. Morgan Deep convolutional neural networks for ATR from SAR imagery. In: *Algorithms for Synthetic Aperture Radar Imagery XXII* vol. 9475 (SPIE, 2015), p. 116–128
36. J. Shao, C. Qu, J. Li, A performance analysis of convolutional neural network models in SAR target recognition. In *The Proceedings of 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA)* (IEEE, 2017), p. 1–6
37. X. Huang, Q. Yang, H. Qiao, Lightweight two-stream convolutional neural network for SAR target recognition. IEEE Geosci. Remote Sens. Lett. **18**(4), 667–671 (2020)
38. D. Zhang, Z.-H. Zhou, (2D)^{2}PCA: Two-directional two-dimensional PCA for efficient face representation and recognition. Neurocomputing **69**(1–3), 224–231 (2005). https://doi.org/10.1016/j.neucom.2005.06.004
39. Y. Li, L. Du, D. Wei, Multiscale CNN based on component analysis for SAR ATR. IEEE Trans. Geosci. Remote Sens. **60**, 1–12 (2021)

Publisher's Note