Decision fusion using virtual dictionary-based sparse representation for robust SAR automatic target recognition

Behrouz Mojarad Shafie\(1\), Payman Moallem\(1\)\(,\) Mohamad Farzan Sabahi\(2\)

\(1\)Department of Electrical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran
\(2\)E-mail: p_moallem@eng.ui.ac.ir

Abstract: Sparse representation displays remarkable characteristics when applied to image processing and classification. The critical point in the success of sparse representation-based classification is to learn an authentic dictionary. The present study proposes a virtual dictionary-based sparse representation for automatic target recognition. Based on the properties of the synthetic aperture radar (SAR) images, some low complexity modules including adding speckle noise, histogram equalisation mapping, and bicubic interpolation are applied to construct some virtual compact dictionaries using Fisher discriminative dictionary learning. These dictionaries have different discriminative information on targets, which are used independently in several sparse representation-based classifiers. The reconstruction error vectors of the latter classifiers are then combined to recognise the target using decision fusion. Based on experimental results obtained drawing upon moving and stationary target acquisition and recognition data set, the proposed method presents the highest accuracy in classification reported yet in the literature. Furthermore, the procedure improves the recognition robustness against most commonly extended operating conditions, e.g. speckle noise corruption, depression angle variation and reduced training set. Accordingly, the current study claims a robust parallel method of high real-time ability in the target recognition of SAR images applicable to practical situations.

1 Introduction

Synthetic aperture radar (SAR) has attracted considerable attention both in civilian and military applications. The interest arises from the fact that it can capture data from a large area of land despite all-weather conditions with high-resolution capabilities \([1, 2]\). As of late, a popular research area of SAR image processing and interpretation is being focused on automatic target recognition (ATR), i.e. applying signal processing methods to recognise unknown targets \([3]\). A SAR ATR system consists of three separate stages: detection, discrimination and classification. The detection stage locates the candidate targets in SAR images and detects regions of interest (ROI). The discrimination stage rejects the natural-clutter false alarms. The ROI data are sent to the classification stage to decide the target type \([4]\). The present paper concentrates on the third stage. In this regard, it is to be noted that the term ‘SAR ATR’ in the following parts indicates the classification stage.

As a benchmark for evaluating the SAR ATR algorithms, moving and stationary target acquisition and recognition (MSTAR) data set have been widely used in research works (a detailed description of the data set is supplied in Section 4). Experiments on MSTAR have been carried out under two conditions: standard operating condition (SOC), i.e. the test and the train samples are captured under similar conditions and extended operating condition (EOC), i.e. the test and the train samples are captured under different conditions \([5]\). While there have been considerable improvements in SOC over the last decade (the classification rate over 90%), EOC is still challenging to the researches. Thus, SAR ATR is not yet reliable enough to be used in practical situations.

Such problems as the properties of SAR images, multiplicative speckle noise, depression and aspect angle variation, target configurations and articulation, radar bandwidth, clutter type, camouflage and occlusion make EOC a difficult option to work with \([6]\). Various state-of-the-art machine learning methods have been utilised on MSTAR to improve the performance of SAR ATR, e.g. feature extraction methods like target's scattering centre \([7, 8]\), monogenic signal \([3, 9]\) binary target region \([10, 11]\) as well as classification methods like adaptive boosting \([12, 13]\), support vector machine \([14, 15]\) deep learning \([16–18]\), dictionary learning (DL) methods and sparse representation-based classification (SRC) \([19–21]\), in addition to decision-level fusion strategies \([22–24]\).

Neurologists developed a model for cells in mammalian primary visual cortex called ‘sparse coding’ or sparse representation (SR) \([25]\). Based on SR, SRC \([26]\) was introduced by Wright, et al. in 2009, turning out to be a remarkable classification scheme over the recent years as it exhibits such notable properties as inherent feature extraction and robustness to articulation while needing no processing and parameter estimation. Through this method, a dictionary is built by using training samples to encode a new test sample as a linear combination of them. This eventually leads up to solving an optimisation problem with sparsity constraint and finding the sparse coefficient vector \([26]\). There are two key questions in SRC. First, what approaches are exactly to be employed to construct the dictionary. Second, what type of algorithms are to be used for obtaining the sparse coefficient \([27]\). SRC needs sufficient samples in each class to build an over-complete dictionary. At the same time, it is sensitive to the quality of SAR images and the targets' changing conditions \([28]\). Considering the above facts, efforts have been made to develop SRC algorithms. A summary can be presented as follows: (i) using kernel method to transfer samples to new higher dimension spaces where classes can be linearly discriminated \([29–31]\), (ii) utilising manifold learning \([32–34]\), (iii) fusing SRC with other classification methods \([24, 35]\), (iv) using l2-norm \([36, 37]\) or other norms \([38, 39]\) instead of l1-norm in SRC, (v) acquiring a dictionary via DL methods instead of using training samples could be very effective in the SR and SRC results \([20, 40–42]\). Based on the latter facts, and making use of the Fisher criterion, Zhang and co-authors \([43, 44]\) introduced Fisher discriminative DL (FDDDL). Through this method, a discriminative and compact dictionary is learned so as to be used in SRC. Using this method, researchers have reported significant results in face recognition (FR) as well as digit and gender classification \([45]\).

This paper aims at improving the effectiveness and the robustness of SAR ATR, i.e. enhancing the classification accuracy under SOC and reinforcing the robustness against most common EOCs; namely, speckle noise corruption, depression angle variation, and a limited number of training samples available per class. The main idea is to construct some compact virtual
dictionaries, which can process different discriminative targets' information using different and simple modules in a parallel scheme. Thus, the present article proposes a virtual dictionary-based SR (VDSR). In this method, the FDDL method is adapted for DL for SAR images, eventually leading to the atoms of dictionaries needing fewer memory requirements to save in addition to providing intuitive targets' description for interpretation. Based on the characteristics of SAR images and the algorithms used in FDDL, some simple modules, e.g. adding speckle noise, histogram equalisation mapping, and bicubic interpolation are applied to construct some virtual dictionaries. These dictionaries are applied to SR-based classifiers. The decision-level fusion is then used to combine the classifiers' decision for target recognition. There are two main reasons for this procedure. Thus, the present article proposes a virtual dictionary-based SR (VDSR). In this method, the FDDL method is adapted for SAR images which can not only promote the ATR performance via generating discriminative and compact dictionaries but also provide intuitive targets' description for interpretation.

2 Related works

This section presents a brief review of SR, SRC and DL. The details of the three topics are in the given order provided in [26, 28, 54].

2.1 Sparse representation

The main idea of SR is to represent signals as sparse linear combination dictionary elements called 'atoms' indicated through the following relation:

\[ y \approx D\alpha. \quad D \in \mathbb{R}^{d \times n} \quad (1) \]

where \( n \) is the number of atoms, \( d \) is the dimension of atoms and \( D \) denotes an over-complete dictionary if \( n > d \), and \( \alpha \) is the coefficient vector that must be sparse. Thus, we have:

\[ \hat{\alpha} = \arg\min_{\alpha} \| y - D\alpha \|_2 < \varepsilon \quad (2) \]

Also, it yields fewer artifacts compared with other methods. This procedure reduces dimensions, hence, decreasing the computational complexities. Besides, it can increase the robustness of the SR module to noise as well as the robustness to resolution and depression angle variations [35].

The reasons why we used speckle noise, bicubic interpolation and histogram equalisation for generating dictionaries for classification purposes are explained in the following parts: (i) speckles are constructive and destructive interferences shown as bright and dark spots degrading the quality of SAR images consequently, creating problems for processing and interpretation of SAR images. Thus, before using SAR images in ATR, the effect of the speckle must be reduced. In spite of progresses made in several speckle reduction methods including adaptive or non-adaptive filtering, there is an inherent trade-off between speckle reduction and loss of useful information for attaining classification [49, 50]. Thus, for robustness vis-a-vis speckles, we do not aim at reducing speckles as pre-processing. Rather, multiplicative speckle noises are added to the original MSTAR images to construct a noised dictionary; (ii) The interpolated surfaces obtained by downsampling are smoother than those acquired through other methods. Thus, for robustness to nuisance factors, e.g. noise is enhanced [35].

This section presents a brief review of SR, SRC and DL. The main ideas of the three topics are in the given order provided in [26, 28, 54].

The reasons why we used speckle noise, bicubic interpolation and histogram equalisation for generating dictionaries for classification purposes are explained in the following parts: (i) speckles are constructive and destructive interferences shown as bright and dark spots degrading the quality of SAR images consequently, creating problems for processing and interpretation of SAR images. Thus, before using SAR images in ATR, the effect of the speckle must be reduced. In spite of progresses made in several speckle reduction methods including adaptive or non-adaptive filtering, there is an inherent trade-off between speckle reduction and loss of useful information for attaining classification [49, 50]. Thus, for robustness vis-a-vis speckles, we do not aim at reducing speckles as pre-processing. Rather, multiplicative speckle noises are added to the original MSTAR images to construct a noised dictionary; (ii) The interpolated surfaces obtained by downsampling are smoother than those acquired through other methods. Thus, for robustness to nuisance factors, e.g. noise is enhanced [35].

This procedure reduces dimensions, hence, decreasing the computational complexities. Besides, it can increase the robustness of the SR module to noise as well as the robustness to resolution and depression angle variations [35]. (iii) A process similar to histogram equalisations seems to occur in biological neural networks, i.e. maximisation of the output firing rate as a function of the input statistics of the neuron. This point has particularly been proved in the case of the fly retina [51]. Although it often produces unrealistic effects in photographs which might decrease the SNR due to an increase in the contrast in the background noise, it works to the benefit of scientific images [52, 53].

The main contributions of this paper can be summarised as follows:

(i) FDDL algorithms have been adapted for SAR images which can not only promote the ATR performance via generating discriminative and compact dictionaries but also provide intuitive targets' description for interpretation.

(ii) Some virtual dictionaries are proposed which utilise adding speckle noise, histogram equalisation mapping, and bicubic interpolation. Based on the experiments conducted on MSTAR data set, each of these dictionaries can significantly improve the classification accuracy under SOC and raise the robustness against most common EOCs; that is, speckle noise corruption, depression angle variation and a limited number of the training samples available per class.

(iii) A robust parallel scheme is suggested for SAR ATR exploiting decision fusion strategy which increases the generality and the accuracy of recognition. The search region of SAR ATR solution is expanded using combined classifier's results. The method advanced in this paper overcomes the weaknesses of each of the classifiers by using different discriminative targets' information existing in the different virtual dictionaries.

The rest of the paper is organised as follows. A brief review of the related research works is given in Section 2. Section 3 describes the proposed method in detail followed by experimental results on MSTAR data set and some discussions in Section 4. Section 5 concludes the research paper.
The choice of dictionary plays a vital part in the success of SR. An
over-complete dictionary is obtained using the training samples directly as dictionary atoms (SRC). It can also be constructed using transformation functions as pre-defined orthonormal bases, e.g. super wavelet transform, curvelet transform, contourlet transform etc. Also, the dictionary can be built up based on learning via DL methods [28, 54]. Utilising transformation functions is an attractive idea as they are inherently speedy and simple. This is not adequate enough for FR or SAR ATR, however. Further, the dictionary used in SRC is not effective enough due to noise and uncertain information in the original samples, which produces a huge dictionary leading to coding complexities and time wastages while the hidden discrimination information in the data is not completely exploited [43]. On account of these limitations, numerous DL methods have been proposed for image processing (representative DL) and classification (discriminative DL) [28].

2.3 Dictionary learning

The choice of dictionary plays a vital part in the success of SR. An over-complete dictionary is obtained using the training samples directly as dictionary atoms (SRC). It can also be constructed using transformation functions as predefined orthonormal bases, e.g. super wavelet transform, curvelet transform, contourlet transform etc. Also, the dictionary can be built up based on learning via DL methods [28, 54]. Utilising transformation functions is an attractive idea as they are inherently speedy and simple. This is not adequate enough for FR or SAR ATR, however. Further, the dictionary used in SRC is not effective enough due to noise and uncertain information in the original samples, which produces a huge dictionary leading to coding complexities and time wastages while the hidden discrimination information in the data is not completely exploited [43]. On account of these limitations, numerous DL methods have been proposed for image processing (representative DL) and classification (discriminative DL) [28].

3 Proposed method

This section presents a detailed description of the proposed method. The procedure adopted for the proposed method is illustrated through a flow diagram in Fig. 1.

The training samples are passed through six processing channels to generate six virtual data sets simultaneously. The total number of the samples and the label relevant to each sample in the six data sets have to correspond to the primary data set. The obtained virtual data sets are separately used for learning the six dictionaries via FDDL followed by target recognition schemes using six classifiers. The classifiers have to be SR-based; however, their algorithms or configurations can be different. A new testing sample is processed in the same channels. Nevertheless, it is to be used for classification after processing. Finally, the reconstruction error vectors of the classifiers are combined using decision fusion for reaching the final decision. The results obtained by applying this simple parallel scheme would be impressive.

3.1 Generating virtual data sets

The training and the testing image dimensions in MSTAR are 128 × 128 pixels. As a first step, in order to delete the edges and exclude the redundant backgrounds, the target images are cropped to 64 × 64 pixels, whereby the recognition accuracy is enhanced; hence, a considerable burden of computational complexities is removed [22, 62–64]. The cropped images are saved as the first data set. Likewise, the same procedure is applied to generate the second and the third data set executed by passing through adding speckle-noise module and histogram equalisation mapping, respectively. Multiplicative speckle noise is added to the original MSTAR images to generate a noised data set as follows:

\[
I(x, y) = I(x, y)u(x, y) + \delta
\]

where \(I(x, y)\) is the intensity of the original SAR image, \(u(x, y)\) denotes the uniformly distributed random noise with 0 as mean and \(\delta\) as variance and \(I(x, y)\) denotes the noised image. Thus, different SNRs of the noised SAR images can be obtained via changing \(\delta\). Histogram equalisation is a monotonic non-linear mapping technique which spreads out the most frequent pixel intensity values. Supposing that the probability of an occurrence of a level \(r_k\) pixel in the image is

\[
p_k(r_k) = p(r_k = r_k) = \frac{N_k}{MN} \quad k = 0, 1, \ldots, l
\]

where \(MN\) is the total number of pixels in the image, \(N_k\) is the number of pixels with the intensity \(r_k\), and \(l\) is the maximum possible intensity level. The plot of \(p_k(r_k)\) versus \(r_k\) is known as a histogram. The function \(T(r_k)\) which maps each pixel in the input image with intensity \(r_k\) into a corresponding pixel with a \(s_k\) level in the output image is known as histogram linearisation or histogram equalisation. This is represented through the following relation:

\[
s_k = T(r_k) = (l - 1) \sum_{j=0}^{k} p_j(r_j) = \frac{l - 1}{MN} \sum_{j=0}^{k} n_j, \quad k = 0, 1, \ldots, l
\]

This type of mapping causes the output image to have a uniform distribution of intensity values usually enhancing the contrast of the image with close contrast values spanning a wider range of intensity scales [68, 69]. The image \(I \in \mathbb{R}^{M \times N}\) used as a training and testing sample is a two-dimensional (2D) greyscale image corresponding to the amplitude of the complex SAR data. The original \(N \times N\) image basis could be remapped to an \(M \times M\) feature basis where \(M^2 = N^2\). There are several interpolations methods for down-sampling,

![Fig. 1 Proposed method flow diagram](image-url)
histogram equalisation on a down-sampled data set. Also, the fifth have been proposed to address the issue. The bicubic interpolation
learned, where

Based on Fisher discrimination criterion, Zhang and co-authors
3.2 Constructing virtual dictionaries using FDDL
coefficients having smaller within-class scatters, but bigger
between-class scatters. The following describes the FDDL model
investigate is to determine the 16 coefficients. Some approaches
method preserves the image detail better than the bilinear
algorithm. The fourth virtual data set is generated through
feature fusion (intermediate-level fusion) and decision-level fusion
strategies are suggested: information/data fusion (low-level fusion),
in this way, FDDL converges. Fig. 2 illustrates the convergence of FDDL by
plotting the total objective function value versus the iteration number on the
ten-class targets of MSTAR data set (see Table 1).

Compared with SRC, FDDL can arrive at a more compact and
effective dictionary, which feature enhances the accuracy of FR as
well as digit and gender classifications [45].

3.3 Decision fusion
Certain algorithms cannot extract some features of the targets
which are obtainable through other algorithms. To address the
problem, a decision fusion method can be used so that each
classifier executes its own classification exploiting subsets of
features having identical natures, then the classifiers’ outputs are
sent to be merged by means of various methods eventually leading
to a single final decision. Hence, decision fusion using multi-
classifier and decision-level fusion may expand the search region
for SAR ATR solutions, giving rise to improvement in accuracy
and generalisation in classification. Considering these facts, the
classification results obtained by different dictionaries are properly
fused to form a more precise decision on the intended target type.

Although there is not sufficient understanding as to why and under
what circumstances some decision fusion methods work better than others [47], the following three types of fusion
strategies are suggested: information/data fusion (low-level fusion),
feature fusion (intermediate-level fusion) and decision-level fusion
(high-level fusion). In this paper, the classifiers’ output is
considered error construction obtained over each class whose
product can be used as probabilistic confidence measures.

\[
\begin{align*}
J(D, \beta) &= \arg \min_{D, \beta} r(X, D, \beta) + \lambda_1 \| \beta \|_1 + \lambda_2 \| \beta \|_1^2 \\
&= \arg \min_{D, \beta} \| X - D \beta \|_F^2 + \lambda_1 \| \beta \|_1 + \lambda_2 \| \beta \|_1^2
\end{align*}
\]  

Eq. (13) is not jointly convex. Consequently, it could be
divided into two optimisation problems: updating \( \hat{\beta} \) by fixing \( D \)
then fixing \( \beta \) and updating \( D \) class by class. The procedure is
repeated iteratively until the maximum iteration number is covered.
Both these alternative sub-optimisations are compared; in this way,
FDDL converges. Fig. 2 illustrates the convergence of FDDL by
plotting the total objective function value versus the iteration number on the
ten-class targets of the MSTAR data set (see Table 1).

Compared with SRC, FDDL can arrive at a more compact and
effective dictionary, which feature enhances the accuracy of FR as
well as digit and gender classifications [45].

Based on Fisher discrimination criterion, Zhang and co-authors
introduced FDDL for classification. In their discriminative
method, as the within covariance of coefficients is minimised,
their between-class covariance is maximised. This results in coding
coefficients having smaller within-class scatters, but bigger
between-class scatters. The following describes the FDDL model

\[
f(\beta) = \text{tr}(SW(\beta)) - \text{tr}(SW_0(\beta)) + \eta \| \beta \|_F^2
\]

where \( \| \cdot \|_F \) is an elastic term, and \( \eta \) is a parameter which is added into (10) to be strictly convex. Also, \( SW(\cdot) \) and \( SW_0(\cdot) \) are the within-class and between-class scatters, respectively, defined as:

\[
SW(\beta) = \sum_{i=1}^{n} \sum_{\beta_k \in \beta_i} (\beta_k - m_i)(\beta_k - m_i)^T
\]

\[
SW_0(\beta) = \sum_{i=1}^{n} n_i (m_i - m)(m_i - m)^T
\]

The sparse coding coefficients of each class \( \beta_i \) could be used to compute the mean coefficient vector of that class denoted by \( m_i \).
So, \( m_i \) and \( m \) are the mean vectors of \( \beta_i \) and \( \beta \), respectively. Also, \( n_i \) is the number of samples in class \( i \) and \( C \) is the number of all classes. The final FDDL model is formulated as follows:

\[
J(D, \beta) = \arg \min_{D, \beta} \sum_{i=1}^{C} r(X, D, \beta) + \lambda_1 \text{tr}(SW(\beta) - SW_0(\beta)) + \eta \| \beta \|_F^2 + \lambda_2 \| \beta \|_1^2
\]

\[
+ \lambda_3 \left( \text{tr}(SW(\beta) - SW_0(\beta)) + \eta \| \beta \|_F^2 \right)
\]

\[
+ \lambda_4 \left( \| \beta \|_1 + \lambda_2 \| \beta \|_1^2 \right)
\]

Table 1 MSTAR training and testing samples

| Number | Class | Training (17°) | Testing (15°) |
|--------|------|---------------|---------------|
| 1      | BMP2 | 233(Sn_9563)  | 195(Sn_9563)  |
|        |      | 232(Sn_9566)  | 196(Sn_9566)  |
|        |      | 233(Sn_c21)   | 196(Sn_c21)   |
| 2      | BTR70| 233(Sn_c71)   | 196(Sn_c71)   |
| 3      | T72  | 232(Sn_132)   | 196(Sn_132)   |
|        |      | 231(Sn_812)   | 195(Sn_812)   |
|        |      | 233(Sn_S7)    | 191(Sn_S7)    |
| 4      | BTR60| 256           | 195           |
| 5      | ZS1  | 299           | 274           |
| 6      | BRDM2| 298           | 274           |
| 7      | D7   | 299           | 274           |
| 8      | T62  | 299           | 273           |
| 9      | ZIL131| 299         | 274          |
| 10     | ZSU23/4| 299         | 274          |

including (2D)1D nearest neighbour, (bi)linear, and (bi)cubic.
However, the advantages of the bicubic interpolation outweigh
the other methods. That is why it is preferred over other procedures.
Bilinear interpolation uses four pixels (2 × 2) whereas the
bicubic interpolation uses 16 pixels (4 × 4). So, the problem
to investigate is to determine the 16 coefficients. Some approaches
have been proposed to address the issue. The bicubic interpolation
method preserves the image detail better than the bilinear
algorithm. The fourth virtual data set is generated through
histogram equalisation on a down-sampled data set. Also, the fifth
and the sixth data sets are obtained through down-sampling on the
noised data set and the histogram equalised data set, respectively.

3.2 Constructing virtual dictionaries using FDDL

Based on Fisher discrimination criterion, Zhang and co-authors
introduced FDDL for classification. In their discriminative
method, as the within covariance of coefficients is minimised,
their between-class covariance is maximised. This results in coding
coefficients having smaller within-class scatters, but bigger
between-class scatters. The following describes the FDDL model

\[
J(D, \beta) = \arg \min_{D, \beta} r(X, D, \beta) + \lambda_1 \| \beta \|_1 + \lambda_2 \| \beta \|_1^2
\]
\[ DP(x) = \begin{bmatrix} d_{i,1}(x) & \ldots & d_{i,j}(x) & \ldots & d_{i,C}(x) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{L,1}(x) & \ldots & d_{L,j}(x) & \ldots & d_{L,C}(x) \end{bmatrix} \] (14)

where \( d_{i,j}(x) \) denotes the support that classifier \( D_i \) provides for the class label \( \phi_j \). Some fusion decision strategies perform a column-wise operation on \( DP(x) \) for calculating \( \mu(x) \) including sum, product, minimum and maximum, etc. Some other strategies use the entirety of \( DP(x) \) for obtaining \( \mu(x) \) e.g. [46, 70], decision templates [71] or Dumpster-Shafer combination [72, 73]. The present research work uses sum method as a fusion rule. The decision fusion based on sum rules can be formulated as follows:

\[ \mu_j(x) = \sum_{i=1}^{L} d_{i,j}(x), \quad j = 1, 2, \ldots, C \] (15)

4 Experimental results

The MSTAR data set is used as an experimental tool to evaluate the performance of the proposed method. The MSTAR data set comprises of ten classes of ground military targets collected with the help of a 10 GHz x-band HH polarisation SAR sensor in the spotlight mode with a resolution of 1 ft \( \times \) 1 ft. The target images were captured over a 0°–360° range of aspect angle at various depression angles (15°, 17°, 30° and 45°). The targets include BMP2 (tank), BTR70 (armored car), T72 (tank), BTR60 (armored car), 2S1 (rocket launcher), BRDM2 (armored car), D7 (bulldozer), T62 (tank), ZIL131 (truck) and ZSU23/4 (air defence unit) [16].

4.1 Ten-class target recognition under SOC

Table 1 lists the serial number, depression angle and the number of each target used for the purpose of training and testing in the ten-class SOC recognition [5]. Images at 17° depression angle are used as the training set, whereas the images at 15° depression angle are used as the testing set.

There are various target configurations for a ten-class recognition under SOC in the literature. The differences concern the use of varied serial numbers for BMP2 and T72 having variants with small structural modifications (see Table 1). The present paper uses all these three serial numbers of training and testing samples for recognition.

Figs. 3–8 displays the atoms of the learned dictionaries used for SR-based classifiers. In these Figures, each column is a sub-dictionary of a certain class. Consequently, the complete dictionary encompasses ten columns which in the order given from left to right correspond to the ten classes of targets in Table 1. Based on the number of experiments repeated, the number of atoms for each sub-dictionary in the learning algorithms is optimised to ten, which number is adequate for attaining satisfactory results. Hence, the number of atoms in each row is ten and each dictionary includes a hundred atoms (10 \( \times \) 10). Also, based on the experiments carried out (see Fig. 2), the maximum iteration number in the DL algorithm is optimised at five.
In the literature, the probability of correct classification \( \text{PCC} = \frac{n_c}{n_t} \times 100\% \) is commonly used to measure recognition performance. Where \( n_c \) is the number of samples correctly recognised and \( n_t \) is the total number of samples. In the ATR literature, the confusion matrix is commonly employed to illustrate the recognition performance. Each row of confusion matrix denotes the actual target (ground-truth) class, and each column represents the class recognised by the classifier. So, the diagonal entries of the confusion matrix correspond to the number of samples from a specific class correctly recognised (the PCC of that class). Also, the average PCC can be obtained by averaging the PCCs of the ten classes. Results of the ATRs obtained through different dictionaries are shown as confusion matrices in Figs. 9–14.

Fig. 15 exhibits the confusion matrix for the final results of the experiment obtained through decision fusion technique.
By comparing these six confusion matrices provided in Figs. 9–14, the following two conclusions can be drawn. First, using the proposed virtual dictionaries (Figs. 10–14), instead of the dictionary obtained via cropped images (Fig. 9), improves the PCC by more than six percent, which value can be regarded as a considerable achievement. Second, each dictionary recognises the specific targets better than the others. This can be attributed to the fact that there are different bits of information on the targets in each dictionary which represent the specifics of each target better than the others. Accordingly, as is shown in Fig. 15, using decision fusion of classifiers can improve the ATR performance.

4.2 Target recognition under EOC

The MSTAR images are captured under varying EOCs, e.g. different depression angles, aspect angles and serial numbers. Papers report some more EOCs, e.g. noise contamination including multiplicative speckle noise and additive white Gaussian noise, limited training resource, resolution variance, partial occlusion etc. [7–47]. In practice, however, it seems that most crucial EOCs are adding speckle noise, different depression angles and a limited number of the training samples. Hence, the robustness of the proposed method is evaluated under these three EOCs.

4.2.1 Three-class recognition under EOC (depression angle variance): In a real-world scenario, capturing SAR images in specific depression angels is too difficult. Also, the test target images might be obtained from quite different depression angles. Thus, the robustness of a classification algorithm relative to the depression angle variance is crucial. In this subsection, an experiment is designed to test the robustness of the proposed method as against varying depression angles.

In the MSTAR data set, there are images with different depression angles as EOC for evaluating the depression robustness. Table 2 lists the samples used in the experiment. The samples related to three different targets with depression angles of 17° are used for training, and samples with depression angles of 30° and 45° are utilised for testing. Table 3 shows the PCCs related to the algorithms for depression angle changing and Figs. 16–19 present the confusion matrices related to the experimental results.

Table 2 Number of targets with different depression angles

| Target type | 2S1 | BRDM2 | ZSU23/4 |
|-------------|-----|-------|---------|
| training (17°) | 299 | 298 | 299 |
| testing (30°) | 288 | 287 | 288 |
| testing (45°) | 303 | 303 | 303 |

Table 3 Experimental results for depression variations

| Depression angle | D1 | D2 | D3 | D4 | D5 | D6 | DF |
|------------------|----|----|----|----|----|----|----|
| % (30°) PCC      |    |    |    |    |    |    |    |
| 94.3 | 95.7 | 92.7 | 99.18 | 99.88 | 99.1 | 100 |
| % (45°) PCC      |    |    |    |    |    |    |    |
| 78.4 | 75.4 | 69.7 | 71.72 | 66.8 | 71.7 | 78.43 |

Specific targets better than the others. This can be attributed to the fact that there are different bits of information on the targets in each dictionary which represent the specifics of each target better than the others. Accordingly, as is shown in Fig. 15, using decision fusion of classifiers can improve the ATR performance.
In the first experiment, a change of 13° is observed in the depression angle (from 17° to 30°) between the available images for training and those for testing. As is shown in Table 3 and Figs. 16 and 17, all the PCCs obtained using different dictionaries are over 92%. Also, all the testing samples are correctly recognised using decision fusion of the classifiers (see Fig. 17c). In the second experiment, due to a drastic change of 28° in the depression angle (from 17° to 45°) between the testing and the training samples, there are significant changes in the global properties of the targets including the intensity values and the shapes. Therefore, the PCCs belonging to all methods drastically drop as well. This is shown in Table 3 and Figs. 18 and 19. The PCCs obtained by decision fusion of the classifiers is better than those obtained using the dictionaries (see Fig. 19c). The reason is that each dictionary recognises specific targets better than the others, which fact is demonstrated in Figs. 18 and 19.

4.2.2 Ten-class recognition under EOC (reduced training set): The SAR ATR resources for training samples are often expensive and scarce for SAR ATR. Besides, even after cropping the images to $64 \times 64$, the data dimension is 4096. However, the total number of SAR images per class for training in MSTAR is in the range of 190–300 (see Table 1). This creates a problem for learning high dimensional data under limited training set [74]. Thus, the performance of the proposed method has to be checked under the reduced training set. The samples’ proportions of 0.8, 0.6, 0.4, 0.2, and 0.1 are randomly selected from each of the ten classes. In the next stage, they are employed when the experiment is repeated. Figs. 20 and 21 display the PCC of the experiments using confusion matrices for comparison. Fig. 22 shows the PCCs plot of the ATR versus the proportion of the reduced training set.

As is indicated by Fig. 22, the PCC of the proposed method is over 90°. This occurs at a time when the size of training samples is reduced to 0.8, 0.6, 0.4, and 0.2 times of the original size. Nevertheless, when the size of the training samples is reduced to 0.1 times of the original samples, the PCC drastically drops from 93 to 76%. In general, the obtained results demonstrate that the proposed method is robust against small sample sizes.

4.2.3 Ten-class recognition under EOC (adding speckle noise): The speckle is the most common noise characteristically existing in SAR images due to the mechanism of data acquisition and the coherent nature of SAR. In this subsection, an experiment is designed to test the robustness of the proposed method to speckle corruption. The speckle noises are added to the testing samples by (5). Fig. 23 shows some examples of the noised samples. Performance of the ATR under different variances of the speckle noises are indicated in Fig. 24.

As shown in Fig. 24, even when the variance of the noise is four, the PCC of the proposed method is 70°. Thus, the proposed method is robust against noise corruptions.

4.3 Comparison with previous works

In this subsection, the results obtained through the proposed method are compared with those of some state-of-the-art ATRs on MSTAR. Table 4 compares the average PCCs of the proposed method under SOC with those of state-of-the-art researches.
Table 5 also compares the average PCCs of the proposed method under EOC (depression angle variance) with those of state-of-the-art researches.

Table 4 illustrates that all methods achieve very high PCCs under SOC while the achievements of the proposed method are highest of all. Table 5 also indicates that the PCC of decision fusion of the classifiers represents the highest percentage among all state-of-the-art methods for either 30° and 45° depression angle changes.

5 Conclusion
This paper is an attempt to introduce a method for target recognition in SAR images by using low computational complexity modules in parallel schemes. The feasibility of the proposed framework has successfully been tested on MSTAR data set. From the experimental results, the following conclusions can be drawn.

(i) Since different types of information on targets are extracted by means of constructing various virtual dictionaries, decision fusion can prove useful in improving the recognition process.

(ii) The proposed method has been able to attain the classification accuracy of over 99.8% under SOC, which results in comparison with other state-of-the-art methods is a considerable achievement. It shows the highest PCC reported in the literature as yet.

(iii) The suggested method improves the robustness to the most important SAR ATR problem, i.e. speckle noise corruption besides exhibiting excellent stability when the depression angle distances are varied or when the training samples are reduced.

(iv) The proposed method exhibits a high real-time capability as it employs simple algorithms together with parallel processing structures.

(v) Exploiting the proposed dictionaries rather than using samples directly for classification significantly reduces the memory requirements.

(vi) The proposed method can provide an intuitive description of targets for interpretation via constructing some virtual representative dictionaries.

(vii) As DL and SR-based classifiers are involved, the method needs no set initial parameters, feature extraction, or feature selection, which points a raise in the generalisation of the ATR.

Below are some suggestions for further research that might improve the performance of the proposed method.

(i) Using more data sets and dictionaries obtained through processing at different spaces.

(ii) Exploiting classifiers with different algorithms or configurations adapted to the properties of virtual dictionaries.
References

[1] Deng, S., Xu, L., Li, C., et al.: ‘SAR automatic target recognition based on Euclidean distance restricted autoencoder’, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 2017, 10 (7), pp. 3233–3233

[2] Dong, G., Kuang, G.: ‘Classification of the monogenic scale space: application to target recognition in SAR images’, IEEE Trans. Image Process., 2015, 24 (8), pp. 2527–2539

[3] Dong, G., Wang, N., Chang, G.: ‘Sparse representation of monogenic signal with application to target recognition in SAR images’, IEEE Signal Process. Lett., 2014, 21 (8), pp. 952–956

[4] E3-Daryniki, K., J. Remote Sens., S. Power, D., et al.: ‘Target detection in synthetic aperture radar imagery: A state-of-the-art survey’, J. Appl. Remote Sens., 2013, 7 (1), p. 71598

[5] Keydel, E.R., Lee, S.W., Moore, J.T.: ‘MSTAR extended operating conditions: A tutorial’, Algorithms for Synthetic Aperture Radar Imagery III, International Society for Optics and Photonics, 1996, 2757, pp. 228–242

[6] Ding, B., Wen, G.: ‘Combination of global and local filters for robust SAR target detection under various extended operating conditions’, Inf. Sci., 2019, 476, pp. 48–63

[7] Ding, B., Wen, G., Huang, X., et al.: ‘Target recognition in synthetic aperture radar images via classification of attributed scattering centers’, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 2017, 10 (7), pp. 3334–3347

[8] Ding, B., Wang, G., Zhong, J., et al.: ‘A robust similarity measure for attributed scattering center sets with application to SAR ATR’, Neurocomputing, 2017, 210–227

[9] Dong, G., Kuang, G.: ‘Target recognition via information aggregation through Dempster-Shafer's evidence theory’, IEEE Geosci. Remote Sens. Lett., 2015, 12 (6), pp. 1125–1129

[10] Zhan, X., Liu, Z., Liu, S., et al.: ‘Sparse coding of 2D-slice zernike moments for SAR ATR’, Int. J. Remote Sens., 2017, 38 (2), pp. 412–431

[11] Ding, B., Wen, G., Ma, C., et al.: ‘Target recognition in synthetic aperture radar images using statistical morphological operations’, J. Appl. Remote Sens., 2016, 10 (4), p. 1–14

[12] Lee, W., Jun, C.-H., Lee, J.-S.: ‘Instance categorization by support vector machines via self-weighting in Adaboost for imbalance data classification’, Inf. Sci., 2017, 381, pp. 912–923

[13] Sun, Y., Liu, Z., Todorovic, S., et al.: ‘Adaptive boosting for SAR automatic target recognition’, IEEE Trans. Aerosp. Electron. Syst., 2007, 43 (1), pp. 112–125

[14] Zhao, Q., Principe, J.C.: ‘Support vector machines for SAR automatic target recognition’, IEEE Trans. Aerosp. Electron. Syst., 2001, 37 (2), pp. 643–654

[15] Liu, J., Fang, N., Xie, Y.J., et al.: ‘SAR target classification using image vector machine and subspace methods’, IET Radar Sonar Nav., 2015, 9 (6), pp. 632–640

[16] Chen, S., Wang, H., Xu, F., et al.: ‘Target classification using the deep convolutional networks for SAR images’, IEEE Trans. Geosci. Remote Sens., 2016, 54 (8), pp. 4806–4817

[17] Cui, T., Zhang, G., Wang, X., et al.: ‘D-ATR for SAR images based on deep neural networks’, Remote Sens., 2019, 11 (8), pp. 1–17

[18] Zheng, C., Jiang, X., Liu, X., et al.: ‘Semi-supervised SAR ATR via multi-discriminator generative adversarial network’, IEEE Sens. J., 2019, 19, pp. 5252–7531

[19] Liu, M., Chen, S.: ‘SAR target recognition based on the Dempster-Shafer theory and sparse representation using a new classification criterion’, Int. J. Remote Sens., 2019, 40 (12), pp. 4604–4622

[20] Liu, M., Chen, S., Wang, X., et al.: ‘SAR target recognition via discriminative statistical dictionary learning’, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 2018, 11 (11), pp. 4218–4229

[21] Xiao, M., Shao, Y.: ‘SAR target recognition via sparse representation of multi-view SAR images with correlation analysis’, J. Electromagn. Waves Appl., 2019, 33 (7), pp. 897–910

[22] Liu, H., Li, S.: ‘Decision fusion of sparse representation and support vector machine for SAR image target recognition’, Neurocomputing, 2013, 113, pp. 97–104

[23] Ding, B., Wen, G., Ma, C., et al.: ‘Decision fusion based on physically relevant features for SAR ATR’, IET Radar Sonar Nav., 2016, 11 (4), pp. 682–690

[24] Liu, M., Wu, Y., Zhao, W., et al.: ‘Dempster-Shafer fusion of multiple sparse representations’, 2017, 219, pp. 130–145

[25] Ohlssenhaus, B.A., Field, D.J.: ‘Emergence of simple-cell receptive field properties by learning a sparse code for natural images’, Nature, 1996, 381 (6583), pp. 607–609

[26] Wright, J., Yang, A.Y., Ganesh, A., et al.: ‘Robust face recognition via sparse representation’, IEEE Trans. Pattern Anal. Mach. Intel., 2008, 30 (2), pp. 210–227

[27] Ding, B., Wen, G., ‘Target recognition of SAR images based on multi-resolution representation’, Remote Sens. Lett., 2017, 8 (11), pp. 1006–1014

[28] Zhang, Z., Xu, Y., Yang, J., et al.: ‘A survey of sparse representation: algorithms and applications’, IEEE Access, 2015, 3, pp. 490–530

[29] Wagner, S.A.: ‘SAR ATR by a combination of convolutional neural network and support vector machines’, IEEE Trans. Aerosp. Electron. Syst., 2016, 52 (6), pp. 2861–2872
[62] Cao, Z., Xu, L., Feng, J.: ‘Automatic target recognition with joint sparse representation of heterogeneous multi-view SAR images over a locally adaptive dictionary’, *Signal Process.*, 2016, **126**, pp. 27–34

[63] Yu, M., Zhang, S., Dong, G., et al.: ‘Target recognition in SAR image based on robust locality discriminant projection’, *IET Radar Sonar Navig.*, 2018, **12**(11), pp. 1285–1293

[64] Zhang, H., Nasrabadi, N.M., Zhang, Y., et al.: ‘Multi-view automatic target recognition using joint sparse representation’, *IEEE Trans. Aerosp. Electron. Syst.*, 2012, **48**(3), pp. 2481–2497

[65] Ding, B., Wen, G.: ‘Sparsity constraint nearest subspace classifier for target recognition of SAR images’, *J. Vis. Commun. Image Represent.*, 2018, **52**, pp. 170–176

[66] Ding, B., Wen, G.: ‘Target reconstruction based on 3D scattering center model for robust SAR ATR’, *IEEE Trans. Geosci. Remote Sens.*, 2018, **56**(7), pp. 3772–3785

[67] Dachasilaruk, S.: ‘Speckle noise reduction for SAR images using interscale multiplication and soft thresholding’, *Proc. Int. Conf. Wavelet Analysis and Pattern Recognition*, Hong Kong, China, 2008, pp. 188–193

[68] Gonzalez, R.C., Wintz, P.: ‘Digital image processing (applied mathematics and computations)’ (Addison-Wesley Publishing Co. Inc.: Reading, MA, USA, 1977), p. 451

[69] Kaur, H., Rani, J.: ‘MRI brain image enhancement using histogram equalization techniques’, *Proc. Int. Conf. Wireless Communications, Signal Processing and Networking (WISPNET)*, Chennai, India, 2016, pp. 770–773

[70] Lam, L., Suen, C.Y.: ‘Optimal combinations of pattern classifiers’, *Pattern Recognit. Lett.*, 1995, **16**(9), pp. 945–954

[71] Kuncheva, L.I., Bezdek, J.C., Duan, R.P.W.: ‘Decision templates for multiple classifier fusion: an experimental comparison’, *Pattern Recognit.*, 2001, **34**(2), pp. 299–314

[72] Sentz, K., Ferson, S.: ‘Combination of evidence in Dempster-Shafer theory’, (Sandia National Labs, Albuquerque, NM, USA, 2002), doi: 10.2172/800792

[73] Kuncheva, L.I.: ‘Combining pattern classifiers: methods and algorithms’ (John Wiley & Sons, Hoboken, New Jersey, 2014)

[74] Srinivas, U., Monga, V., Raj, R.G.: ‘SAR automatic target recognition using discriminative graphical models’, *IEEE Trans. Aerosp. Electron. Syst.*, 2014, **50**(1), pp. 591–606

[75] Dong, G., Kuang, G., Wang, N., et al.: ‘SAR target recognition via joint sparse representation of monogenic signal’, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 2015, **8**(7), pp. 3316–3328

[76] Dou, S.H., Smith, G.E., Baker, C.J.: ‘Aspect invariant features for radar target recognition’, *IET Radar Sonar Navig.*, 2016, **11**(4), pp. 597–604

[77] Cui, Z., Cao, Z., Yang, J., et al.: ‘Target recognition in synthetic aperture radar images via non-negative matrix factorisation’, *IET Radar Sonar Navig.*, 2015, **9**(9), pp. 1376–1385

[78] Liu, M., Chen, S., Lu, F., et al.: ‘SAR target configuration recognition via product sparse representation’, *Sensors*, 2018, **18**(10), pp. 1–15

[79] Tan, J., Fan, X., Wang, S., et al.: ‘Target recognition of SAR images via matching attributed scattering centers with binary target region’, *Sensors*, 2018, **18**(9), pp. 1–19

[80] Huang, P., Qiu, W.: ‘A robust decision fusion strategy for SAR target recognition’, *Remote Sens. Lett.*, 2018, **9**(6), pp. 507–514

[81] Jiang, C., Zhou, Y.: ‘Hierarchical fusion of convolutional neural networks and attributed scattering centers with application to robust SAR ATR’, *Remote Sens.*, 2018, **10**(6), pp. 1–21