Automatic target recognition (ATR) for military applications is one of the core processes toward enhancing intelligence and autonomously operating military platforms. Spurred by this and given that Synthetic Aperture Radar (SAR) presents several advantages over its counterpart data domains, this article surveys and assesses current SAR ATR algorithms that employ the most popular dataset for the SAR domain, namely the moving and stationary target acquisition and recognition (MSTAR) dataset. Specifically, we perform a direct comparison between current SAR ATR methods and highlight the strengths and weaknesses of each technique under both standard and extended operational conditions. Additionally, despite MSTAR being the standard SAR ATR benchmarking dataset, we also highlight its weaknesses and suggest future research directions.

INTRODUCTION

Automatic target recognition (ATR) for modern military applications has become a necessity to enhance intelligence, reduce collateral damage and fratricide, and ultimately support autonomously operating platforms. Thus, for the last decades, academia and industrial stakeholders have made various ATR attempts throughout several data domains including 2-D Infrared (IR-thermal) [1], [2], 3-D light detection and ranging (LIDAR) [3]–[5], 2-D synthetic aperture radar (SAR) [6], [7], [16]–[21], [8]–[15] or hybrid 3-D ATR originating from 2-D visual data [22]. The strengths and weaknesses of each data modality are presented in Table 1. Ideally, the sensor and thus the employed data domain should combine all the positive features of Table 1, with the most important ones to be compact, low cost, passive, operating from a long distance, and operating under almost all-weather night-and-day conditions, i.e., 24/7 conditions. However, the latter attribute is a game-changing advantage during military operations, and thus Synthetic Aperture Radar (SAR) SAR is overall an appealing ATR option as it combines the 24/7 advantage, a low ATR processing time, and a long operating range.

Spurred by the advantages of SAR, ATR employing this data domain has been attempted using various methods that either originate from the computer vision domain or techniques that are specifically designed for SAR data aiming at maximizing the target recognition performance. Examples of such methods are employing features, i.e., descriptive encodings of local regions in the SAR imagery, deep neural networks, remapping the SAR imagery in various manifolds, and then apply a target recognition strategy, etc. Additionally, SAR ATR algorithms have been evaluated on several datasets including ground-based targets or vessels. However, literature mostly concerns ground targets [23]–[25] with the vast majority employing the public release of the moving and stationary target acquisition and recognition (MSTAR) dataset [24]. Indeed, the MSTAR dataset is the most cited SAR ATR dataset with a continuously growing citation progression over the last years (see Figure 1).

Few surveys of SAR ATR are available, such as [26]–[32]. However, this article is the first to focus on SAR ATR regarding the popular MSTAR dataset. Besides, this article comprehensively covers different ATR schemes highlighting the various methods employed. Finally, our work extends the existing survey-type of literature with the major contributions of this article summarized as follows:

a) To the best of our knowledge, this is the first survey paper to comprehensively cover state-of-the-art and most recent SAR ATR methods that spread over a range of various disciplines.
b) Opposed to existing surveys [26]–[32], we specifically focus on the MSTAR dataset rather than presenting solutions involving other SAR datasets. This is important as the MSTAR has established itself as the major benchmarking dataset in the SAR ATR domain.

c) Despite the numerous papers exploiting the MSTAR database (see Figure 1), this article covers the most recent and advanced progress of SAR ATR. Therefore, it provides the reader with state-of-the-art methods.

d) Comprehensive comparisons of existing ATR algorithms with summaries, performance assessments, and future research directions are also presented.

The structure of this article is as follows. We present the MSTAR dataset in the next section, followed by the various SAR ATR methods. The section “Performance Analysis” compares and contrasts the performance of the presented ATR algorithms. Last, we conclude the article and propose future research directions.

**MOVING AND STATIONARY TARGET ACQUISITION AND RECOGNITION DATASET**

The MSTAR dataset comprises X-band SAR imagery with a 1x1 ft resolution, 360° articulation with a 1° spacing, and an image size of 128 x 128 pixels [32]. The commonly used targets in the literature are presented in Table 2 and Figure 2, while for a complete list of all target variants including the SLICY static target, the reader is referred to [23]. For completeness, SLICY is a precisely designed and machined engineering test target containing standard radar reflector primitive shapes such as flat plates, dihedrals, trihedrals, and top hats. The purpose of this target is to allow Image Understanding developers the ability to validate the functionality of their algorithm with a simple known target. However, it should be noted that despite the MSTAR dataset offers SLICY, the majority of the literature does not involve it during SAR ATR evaluation and thus in this article, it will be omitted.

The proposed SAR ATR methods are evaluated on two major target configuration sets namely the standard operation conditions (SOC) and the extended operation conditions (EOC). The former typically includes the ten target classes of Table 2 or a subset comprising of three targets. EOC involves the same ten-class targets but with various acquisition geometries, i.e., depression angle variations up to 45°, target state variations, i.e., articulation, alternative configurations, and intraclass variability, and finally local target deployment, i.e., various obscuration levels [33], [34].

However, despite the MSTAR dataset has established itself as the major SAR ATR benchmarking dataset, it is worth noting that it suffers from various deficiencies that can be grouped into dataset-linked and user-linked. Dataset-linked deficiencies include background correlation issues, and a priori known training and testing imagery, while user-linked deficiencies a variable target patch size extracted from the entire SAR image, and nonstandard SOC and EOC trials. An analysis of these shortages involves

a) **Background correlation.** This is one of the major dataset-linked flaws and is related to the strong background correlation among the SAR images allowing high true positive ATR rates even if the target is absent from the scene [35], [36].

b) **Nonstandard target patch size.** A number of papers aim to reduce the effectiveness of the background correlation either by cropping the target from the scene using a fixed mask size or by exploiting a target segmentation strategy. However, the final SAR patch size that contains the target is not fixed and thus may still contain background regions affecting the final ATR performance, and making a direct comparison between the available ATR methods not trivial.

c) **Nonstandard SOC and EOC trials.** Despite [33], [34] define the SOC and EOC target setups, over the years additional SOC and EOC setups have emerged by establishing various interclass (different classes of targets) and intraclass target configurations (variations of the same target class) from the standard MSTAR targets presented in Table 2. However, the multiple SOC and EOC setups prohibit an extended and direct comparison among the proposed SAR ATR algorithms, and evaluations are
constrained to common target subsets. In fact, the MSTAR related literature offers more than 17 types of 10-class and 3-class SOC trials, and more than 33 types of EOC.

d) A priori known training and testing imagery. Both training and testing target sets are a priori known making the evaluation unrealistic and questioning the true effectiveness of the proposed methods against unknown testing sets.

Though, despite these deficiencies, the MSTAR dataset is currently the standard SAR ATR benchmarking dataset with numerous citations (see Figure 1). The most common SOC and EOC configurations presented in the literature are shown in Table 3. However, for better readability, the performance comparison of various current techniques presented in the conclusion will involve only the most common MSTAR evaluation subsets presented in Table 3.

**ENCODING AND CLASSIFYING SAR IMAGERY**

SAR image encoding, i.e., a compact tensor-type of representation either describing the SAR image on a local or a global neighborhood basis, is the first major stage for a SAR ATR system. Conceptually, the simplest approach is to neglect any data domain remapping and directly encode the raw SAR reflectivity. Though, the simplicity of this strategy does not underestimate the complexity of the encoding process itself but rather reflects the immediacy of applying the encoding technique. On the other side, some methods initially transform the SAR data into a different data domain and then encode it. Hence, we define two major encoding schemes, i.e., the reflectivity attribute-based, and the transform-based, accordingly. However, regardless of the encoding scheme and data domain employed, the second major phase of ATR incorporates matching the template encodings against the target one and ultimately classifying the target. Despite the encoding and classification phases are discrete, the SAR ATR methods offered by the literature are presented as a complete algorithm where the two major phases are heavily associated and cross-tuned to gain the maximum ATR performance. Thus, it is more appropriate to present and evaluate each technique as a whole pipeline rather than two isolated schemes. Hence, for better readability, we group the current SAR ATR methods based on the data domain utilized and subgroup them depending on the governing method employed, i.e., the technique mostly affecting the classification performance among the ones used in a specific SAR ATR architecture.

**REFLECTIVITY ATTRIBUTE-BASED METHODS**

This type of technique exploits the raw SAR imagery without involving any preprocessing phase to alter the initial data domain. As described in the following subsections, this group of methods involves, feature-based methods, scattering centers calculation, low-rank matrix factorization, deep learning, sparse representation classification, and hybrid strategies combining two of the reflectivity methods.
FEATURE-BASED METHODS

These methods describe the SAR image by a set of attributes, i.e., handcrafted features, or moments, that are sufficiently descriptive to achieve appealing target classification under various nuisances. An overview of the methods is presented in Table 4.

Belloni et al. [36] initially segment the target from the scene to reduce any background correlation effects by employing a Gaussian Mixture Model (GMM) scheme [37]. Then, several features are individually used to encode the SAR image, where top performance is attained using the binary robust invariant scalable keypoints (BRISK) [38] features. Feature matching is based on a nearest neighbor distance ratio (NNDR) scheme utilizing the Hamming distance [39]. An example of BRISK features is presented in Figure 3. Amrani et al. [40] generated a graph-based visual saliency map [41] to extract the target from the SAR image, which is then encoded based on Gabor and Histogram of Oriented Gradients (HOG) [42] features. The latter ones are then fused and input to a two-level directed acyclic graph support vector metric learning (DAG-SVML). This strategy aims to reject weak classifiers via the DAG, emphasize the relevant features, and reduce the influence of noninformative features utilizing the Mahalanobis metric.

Bolourchi et al. [43] suggested a moment-fusion strategy to describe the input image and rank the moments in a descending order based on their Fisher score, which are then input to an SVM type of classifier. Experiments involve both Cartesian and radial moments, however, fusing the Zernike, pseudo-Zernike, Fourier–Mellin, Chebyshev–Fourier, radial Chebyshev, and radial harmonic Fourier moments afford a competitive ATR performance. It is worth noting that the involved moments oversample the central part of the SAR image where the target is placed, enhancing the overall performance. Finally, classification is based on the library support vector machine (LibSVM) [44].

Clemente et al. [45] extracted the discrete-defined Krawtchouk moments, which are then normalized and input to a $k$-NN classification scheme ($k = 3$). In contrast to current literature that is constrained to purely target classification, this article considers all three ATR phases, i.e., target recognition, identification, and classification.

A different strategy that extracts the target outlines is proposed by Tan et al. [46]. The latter ones are then segmented into several parts and matched independently with the template outlines to increase the robustness of the proposed method under target deformations. Finally, a variant of the Hausdorff distance metric is used as the similarity measure between the target and the template outline to

| Table 2. Standard MSTAR Database |
|----------------------------------|
| Target | BMP2 | BTR70 | T72 | BTR60 | 2S1 | BRDM2 | D7 | T62 | ZIL131 | ZSU23/4 |
| ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| serial N° | 9563 | 9566 | c21 | c71 | 132 | 812 | a7 | k10yt7532 | b01 | e71 | - | a51 | e12 | d08 |
| train | 15° | 233 | 232 | 233 | 233 | 232 | 231 | 228 | 256 | 299 | 298 | 299 | 299 | 299 | 299 |
| test | 30° | 195 | 196 | 196 | 196 | 195 | 191 | 195 | 274 | 274 | 274 | 274 | 273 | 274 | 274 |
| test | 45° | - | - | - | - | - | - | 288 | 287 | - | - | - | - | 288 |

Figure 2.
Examples of standard MSTAR targets. (a) Visual image. (b) SAR image.
determine the target’s class. Further methods include binary operations [47], sparse robust filters [48], and the azimuth and range target profile fusion [49].

**ATTRIBUTED SCATTERING CENTERS (ASC)**

Scattering centers or scatterers are the dominating high-frequency radar reflections from surfaces such as corners and flat plates. ATR techniques that exploit the scattering centers model the scattering behavior predicted by physical optics and the geometrical theory of diffraction and statistically estimate the object’s location, geometry, and polarization response. Each scattering center is a function of frequency $f$ and aspect angle $\varphi$, and the total backscattering field is the sum of these individual scattering centers [50], [51]:

$$E(f, \varphi; \Theta) = \sum_{i=1}^{N_{o}} E_i(f, \varphi; \Theta_i)$$  \hspace{1cm} (1)$$

$$E_i(f, \varphi; \Theta_i) = A_i \times \left( \frac{f}{f_c} \right)^{n_i} \times \exp\left[-\frac{4\pi f}{c} (x_i \cos \varphi + y_i \sin \varphi)\right] \times \sin \left(\frac{2\pi f}{c} L_i \sin \left(\varphi - \varphi_i\right)\right) \times \exp\left(-2\pi f y_i \sin \varphi\right)$$  \hspace{1cm} (2)$$

where $f_c$ is the radar center frequency, $c$ is the propagation velocity, $N_{o}$ is the number of scatterers with $\Theta = [\theta_1, \ldots, \theta_{1}, \ldots, \theta_{N_o}]$ incorporating the parameter matrices

---

**Table 3.**

| SOC          | SOC ID of most referenced target subsets (involved targets) | EOC          | Description                                                      |
|--------------|----------------------------------------------------------|--------------|------------------------------------------------------------------|
| 10-class     | 1 (all target serial N° for training and testing)          | Training targets | Testing at 15°, 30°, and 45° depression angles                   |
|              | 2 (2S1 - BMP2 (9563) - BTR60 - BTR70 - D7 - T62 - T72 (132) - ZIL131 - ZSU23/4 for training and testing) | Testing at 15° against training targets and BMP2 version variants |
| 3-class      | 1 (all BMP2 classes – BTR70 – all T72 classes, both for training and testing) | Testing at 15° with additive artificial Gaussian noise at $SNR = \{10, 5, 0, -5, -10\}dB$ |
|              |                                                          | EOC Training targets | Testing at 15° with additive occlusion level at \{10, 20, 30, 40, 50\}% |
|              |                                                          | Testing targets    | Testing at 15° with resolution variation at \{0.3 \times 0.3, 0.4 \times 0.4, 0.5 \times 0.5, 0.6 \times 0.6, 0.7 \times 0.7\}m |

**Table 4.**

| $N_{o}$ | Reference                          | SAR chip size | Pre-processing                                      | Main features                                                                 |
|---------|------------------------------------|---------------|----------------------------------------------------|-------------------------------------------------------------------------------|
| 1       | Belloni et al. [36]                | variable      | -                                                  | BRISK features, Hamming distance for classification                           |
| 2       | Bolourchi et al. [43]              | 128x128       | Histogram equalization, averaging filter, thresholding | LibSVM classification                                                         |
| 3       | Amrani et al. [40]                 | 128x128       | -                                                  | Gabor and HOG features, Mahalanobis distance for classification               |
| 4       | Tan et al. [46]                    | 88x88         | Histogram equalization, averaging filter, thresholding, morphological operations | Target outline “dictionary,” a variant of Hausdorff distance metric for classification |
| 5       | Clemente et al. [45]               | 128x128       | -                                                  | Krawtchouk moments, k-NN classification                                        |
of these scatterers. \( E_i(f;\varphi;\theta_i) \) is the response from the \( i \)th scatterer defined by the parameter tensor \( \theta_i = [A_i, x_i, y_i, L_i, \varphi_i, \alpha_i, \gamma_i]^T \), \( A_i \) is the relative amplitude, \( \alpha_i = \{ -1, -0.5, 0, 0.5, 1 \} \) is the frequency and geometry dependence, and \( x_i \) and \( y_i \) are the scatterer’s location. According to the length \( L_i \), a scattering center can be categorized into a localized scattering center (\( L_i = 0 \)) and a distributed scattering center with (\( L_i \neq 0 \)). \( \varphi_i \) is the orientation angle of a distributed scattering center, \( \gamma_i \) is the aspect dependence of a localized scattering center, and \( (\cdot)^T \) denotes the transpose operator. The scattering geometries are distinguishable by their parameters (\( \alpha, L \)), examples of which are presented in Table 5 [52], [53]. In the context of SAR ATR, the scattering centers create a 2-D attributed scattering center (ASC) model, which is used to characterize the target [54]. An example depicting the ASC model is presented in Figure 4, while an overview of the methods to be presented is shown in Table 6.

Ding et al. [54] proposed a SAR ATR algorithm based on the ASC modeling concept [55]. Initially, the SAR response of the target is segmented from the background by combining a Constant False Alarm Rate (CFAR) strategy with various morphological operations. Then, the processed image undergoes a multistaged processing pipeline that involves the highest SAR energy segmentation via the watershed algorithm, optimal parameter estimation for the scattering model via a novel quasi-Newton method [56], error estimation between the reconstructed SAR image based on the ASC, and the original SAR image via the CLEAN method [57]. The ASC parameters that present the smallest reconstruction error are the ones used to remap the SAR image into the ASC feature space. Finally, template matching in the ASC feature space is performed via the Hungarian algorithm [58], [59]. Ding et al. [60] extracted from the target image the ASC features employing the approximate maximum likelihood (AML) method [61], and create a one-to-one correspondence with the template images using the Hungarian algorithm. These correspondences are then fine matched through an amplitude and a structure similarity using the Dempster–Shafer (D–S) theory of evidence [62]. Accordingly, Zhang [63] also employs ASC features. Ding and Wen [64] fused global and local features where the former are the SAR image intensities of the entire image and the latter features are extracted from the ASC model. Ding et al. [65] enhanced the quality and completeness of the template dataset by suggesting a multilevel SAR imagery reconstruction using the ASC model. Each target is segmented from the background via a CFAR scheme. On the segmented target, a 9x9 grid is overlaid and the ASC model for each subregion is individually estimated based on the AML technique. The advantage of this multilevel method is a local ASC parameter setup, rather than a global which is the norm for ASC based SAR ATR. Ding et al. [66] segmented a region \( R \) that encloses the strongest ASC response utilizing the watershed algorithm. Then, the \( \theta \) parameter of that segment is iteratively optimized using a novel quasi-Newton method [56] until the ASC model subtracted from the original SAR image produces a residual error below a defined threshold, or when the number of maximum iterations is reached. Li and Du [53] suggested a multistaged method that initially convolves the ASC model with several filter sizes defined by a genetic

### Table 5.

| Scattering Geometry Interpretation of \( \alpha \) and \( L \) | \( \alpha \) | \( L \) |
|----------------------------------------------------------|-------|-------|
| Trihedral                                                | 1     | 0     |
| Top Hat                                                  | 0.5   | 0     |
| Corner diffraction                                       | -1    | 0     |
| Sphere                                                   | 0     | 0     |
| Edge broadside                                           | 0     | >0    |
| Edge diffraction                                         | -0.5  | >0    |
| Dihedral                                                 | 1     | >0    |
| Cylinder                                                 | 0.5   | >0    |
algorithm to create a set of low-level feature maps. The latter are sparsely encoded using a proposed locality constraint discriminative dictionary learning (LcLcDDL) that is created during the training phase. Then, the low-level local features are converted into high-level global features by employing a spatial pyramid matching scheme.

**Sparse Representation Classification (SRC)**

It has been proven that a test image $y$ can be adequately represented by intra-class training templates. Given $X = \{X_1, \ldots, X_k, \ldots, X_N\} \in R^{nxm}$ the $n$ templates of size $m$, $X_k = \{x_{k,1}, x_{k,2}, \ldots, x_{k,m}\} \in R^{mxn}$ a template subset containing $n_k$ templates from the $k$th class with $n = \sum n_k$, the SRC theory assumes that if $y$ belongs to the $k$th class, it will approximately lie in the linear span of $X_k$, $y \approx X_k a_k$, where $a_k = [a_{k,1}, a_{k,2}, \ldots, a_{k,m}]^T \in R^{mk}$ is the coefficient vector with entries being the weight, i.e., atoms, of the corresponding training samples in $X_k$, also known as the dictionary.

Given that, the test image class is not a priori known, SRC assumes that $y$ is a linear combination of the entire dataset $X$, the cardinality of the object classes is large, and the representation of $y$ is naturally sparse, SRC aims at finding the sparsest solution to $y = Xa$, which can be derived by solving the optimization problem

$$ a^* = \arg\min_a ||y - Xa||_F^2 + \gamma ||a||_0 $$

(3)

where $a^* = [a_1^*, \ldots, a_k^*, \ldots, a_N^*]$, $\gamma \geq 0$ and $|| \cdot ||_0$ is the $l_0$-norm counting the nonzero entries in $a$. Provided that the problem in (3) is NP-hard, compressive sensing theory [67] suggests that if $y$ is sparse enough the initial signal can be recovered by substituting the $l_0$-norm with the $l_1$-norm

$$ a^* = \arg\min_a ||y - Xa||_F^2 + \gamma ||a||_1, $$

(4)

However, advances in the SRC domain extend to several values within the $l_p - norm$ space $\gamma \in (0, \infty)$.

The atoms in $a$ are ideally related to samples that belong to the same class as $y$. Hence, during the SRC classification $y$ is assigned to the class that fulfills the following constraint:

$$ class(y) = \arg\min(e_k) = \arg\min(||y - X_k a_k^*||) $$

(5)

where $a_k^*$ is a vector with all entries zero except the ones associated with class $k$ and $e_k$ is the reconstruction error. In the context of SAR ATR, the various SRC methods utilize different dictionary constructions and $l_p - norm$ solutions. An overview of the current methods is presented in Table 7.

Wei et al. [68] suggested a fast discriminative dictionary learning (FaDDL) method that reduces the execution time without losing the classification accuracy, by adopting an $l_{1,\infty} - norm$ ball as a constraint to replace $l_0 - norm$ and $l_1 - norm$ that are typically used with the discriminative dictionary learning dictionaries. For completeness, it is worth noting that the general formulation of the $l_{1,\infty} - norm$ is; considering the projection of a matrix $A$ to the $l_{1,\infty}$ ball to find a matrix $B$ that solves the convex optimization problem [69]:

$$ l_{1,\infty} : \min_{B, \mu} \frac{1}{2} \sum_{i,j} (B_{i,j} - A_{i,j})^2 $$

(6)

s.t. $\forall i, j B_{i,j} \leq \mu_i$

$$ \sum_i \mu_i < C $$

(7)

where (6) is the Euclidean distance between $A$ and $B$, given that $B$ is in the boundary of $l_{1,\infty}$ ball of radius $C$. Variables $\mu_i$ constrain the coefficients of $B$ as of (7).
Zhang [70] proposed a multistream SRC configuration, where each stream utilizes the target image under different resolution. The SRC scheme employs a $l_{0,2} - \text{norm}$, which is a mixed $l_p - \text{norm}$ variant that first applies the $l_0 - \text{norm}$ on each row of the dictionary matrix, and then the $l_2 - \text{norm}$ on the resulting vector. Ding and Wen [71] also suggest a multiresolution method, but augment the SRC dictionary by creating multiple SAR images of lower resolutions originating from the original templates. Similarly, Zhang and Liu [72] augmented the templates with multiresolution representations, which then undergo a Principal Component Analysis (PCA) prior to the SRC scheme. Liu and Chen [73] weaken the influence of the target aspect angle sensitivity by normalizing the reconstruction error $e_k$ and thus enhance the contribution of the templates with small $e_k$ values. Finally, the normalized errors are fused via the Dempster–Shafer theory [62] and the final target classification relies on the maximum belief rule [74]. Chang and You [75] decoupled the three main contributors of the SAR imagery, i.e., target region, shadow, and background, and input the segmented target (target region and shadow) along with the original image into two distinct SRC pipelines. Then, both classification results are combined by a score-level fusion for target recognition.

Driven by the conclusion that sparse models learned from training samples have more advantages than the predefined ones [76], [77], Liu et al. [78] generated a discriminative statistical dictionary learning (DSDL) technique to enhance the SRC performance. The discriminative nature of the DSDL is further enhanced by manipulating the dictionary entries by minimizing the intraclass differences and maximizing the interclass differences. Similarly, Ren et al. [79] proposed a class-oriented dictionary learning scheme. The latter modifies (4) by adding the Tikhonov regularization matrix $F$:

$$a^* = \arg \min_{a} \| y - Xa \|_2^2 + \gamma \| \Gamma a \|_1$$

$$F = \begin{bmatrix}
d_{i,1}(x_{i,1}, a_{i,1}) & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & d_{i,n}(x_{i,n}, a_{i,n})
\end{bmatrix}.$$  

Li et al. [80] suggested a hybrid-structured synthesis dictionary to better reflect the interclass target differences that incorporate a learned dictionary $D^*$, which is the same for all templates and reduces the effect of common features, such as the high similarity caused by specular reflection. Rather than using the processing costly nonlinear $l_0 - \text{norm}$ and $l_1 - \text{norm}$ sparsity regularizers, the authors employ an analysis dictionary $P$ [81] to create the coding coefficients of the dictionary by linear projection.

Opposing to the previous methods that utilize a single target aspect, several SAR ATR pipelines exploit multiple target aspects. Joint SRC [82] for example, exploits three target views to increase the completeness of the target’s SAR signature and then utilizes a mixed $l_{0,2} - \text{norm}$. The reason for using multiple views is that these are highly

### Table 7.

| N* | Reference | SAR chip size | Pre-processing | Data augmentation | Main features |
|----|-----------|---------------|----------------|------------------|--------------|
| 1  | Wei et al. [68] | 128x128 | - | - | $l_{1,\infty} - \text{norm}$ |
| 2  | Miao and Shan [83] | 128x128 | - | - | multiple target views (6) |
| 3  | Liu and Chen [73] | 80x80 | normalization | - | $l_{2} - \text{norm}$, D-S theory of evidence, maximum belief rule |
| 4  | Ren et al. [79] | 64x64 | - | - | Tikhonov regularization matrix, learned dictionary |
| 5  | Ren et al. [85] | 64x64 | - | - | weighted regression strategy |
| 6  | Chang and You [75] | 128x128 | normalization, noise filtering, morphological operations | - | two-stream SRC, score-level fusion |
| 7  | Huang et al. [86] | 128x128 | - | - | multiple target views (3) |
| 8  | Liu et al. [78] | 80x80 | normalization | - | statistical dictionary learning |
| 9  | Zhang [70] | 128x128 | - | multi-resolution | $l_{0,2} - \text{norm}$ |
| 10 | Ding and Wen [71] | 128x128 | - | multi-resolution | $l_0 - \text{norm}$, $l_2 - \text{norm}$ minimum reconstruction error |
| 11 | Li et al. [80] | 80x80 | - | - | analysis dictionary learning |
| 12 | Zhang and Liu [72] | 128x128 | - | multi-resolution | $l_{0,2} - \text{norm}$ |
| 13 | Zhang et al. [82] | 64x64 | - | - | multiple target views (6) + $l_0 - \text{norm}$ |
correlated sharing the same response pattern within the dictionary and thus this conciseness can enhance the overall ATR performance. Miao and Shan [83] first clustered the SAR templates into several aspect sets based on an image correlation criterion. Then, the joint SRC (JSRC) [84] is applied utilizing six images of the same target but in various views. The JSRC problem is solved employing an \(l_{1,2} - \text{norm}\) (it first applies the \(l_2 - \text{norm}\) on each row of the matrix, and then the \(l_1 - \text{norm}\) on the resulting vector). Later, the reconstruction errors \(e_k\) from JSRC (one \(e_k\) per target view) are normalized and input to a Bayesian decision fusion scheme, where the classification decision relies on the maximum posterior probability. In the event a single target view is available, the proposed algorithm degrades to the typical SRC scheme. Ren et al. [85] proposed a method that learns a supervised sparse model from training single-view samples by utilizing sample label information. Additionally, during dictionary learning, a supervised classifier is jointly designed to enhance the classification performance and the classification error is backpropagated to the dictionary learning process to optimize the dictionary atoms. To enhance the classification performance, during testing, a multiview strategy is applied combined with a weighted regression strategy to assign the identity. Huang et al. [86] reduced the classification error imposed by speckle noise by generating a joint low-rank and sparse multiview denoising (JLSMD) dictionary. Creating the JLSMD dictionary requires combining several target views and exploiting the low-rank property of multiview target images and the sparsity of speckle-noise for SAR systems. The target identity is assigned based on an \(l_1 - \text{norm}\) scheme.

### LOW-RANK MATRIX FACTORIZATION

Nonnegative matrix factorization (NMF) and truncated singular value decomposition (TSVD) are two low-rank matrix factorization (LMF) methods that are used for SAR ATR. Given \(X = [x_1, \ldots, x_n] \in R^{m \times n}\) that includes \(n\) samples and \(m\) dimensions, NMF aims to find two non-negative matrices \(Z\) and \(H\) so that

\[
\min_{Z \geq 0, H \geq 0} \| X - ZH \|_F^2
\]

where \(Z = [z_1, \ldots, z_d] \in R^{m \times d}\) and \(H = [h_1, \ldots, h_n] \in R^{d \times n}\) are low-rank matrices, with \(Z\) mapping the templates \(X\) with their low-dimensional representation \(H\). For the TSVD case, the factorized matrix \(Z\) undergoes an orthogonality constraint and the corresponding optimization problem is defined as

\[
\min_{Z^T Z = I_d, H} \| X - ZH \|_F^2.
\]

The latest LMF techniques are mostly employing an NMF approach [11] with Dang et al. [87] suggesting an incremental NMF (INMF) that has sparse \(l_p\) constraints, where \(p = 1/2\) to employ a variety of norm-sparse constraints. In contrast to the typical NMF where existing and new samples are computed to retrain a new model, the INMF involves only the new samples to update the trained model incrementally. Hence, the proposed \(l_p - \text{INMF}\) applies an updating process under a generic sparse \(l_p\) constraint during the matrix decomposition of the INMF process. Similarly, in [15] Morgan suggested the \(l_{1,2} - \text{NMF}\) strategy, where the NMF features are input to an SRC scheme that employs an \(l_{1,2} - \text{NMF}\) optimization to identify the sparsest solution. Zhang et al. [88] first extracted three types of features, i.e., Gabor, PCA, and wavelet features, which then undergo an NMF process to enhance modeling data sparsity. Finally, the target and the template NMF-processed features are then input to a two-staged classification process that includes a coarse classification phase based on a voting process, and a fine classification phase relying on a Bayesian decision fusion process.

TSVD has also been employed with Yu et al. [89] augmenting the typical TSVD optimization (11) with a manifold regularization term to achieve better ATR performance under the EOC scenarios. The latter term includes a weighted pairwise similarity and a local linearity rule to capture the intrinsic manifold structure.

### Table 8.

| N | Reference | Type | SAR chip size | Main features |
|---|-----------|------|---------------|--------------|
| 1 | Dang et al. [87] | NMF | 64x64 | \(l_{1,2} - \text{norm}\) |
| 2 | Zhang et al. [88] | NMF | 64x64 | Gabor, PCA, and wavelet features, two-step classification: coarse using a voting scheme and fine using a Bayesian decision fusion |
| 3 | Dang et al. [15] | NMF | - | SRC \(l_{1,2} - \text{norm}\) classification |
| 4 | Yu et al. [89] | TSVD | 64x64 | - |
Deep Learning

During the last years, deep learning algorithms have been gaining large attention due to their high classification rates. An additional advantage is their end-to-end nature that includes all three major classification stages, i.e., feature detection, description, and matching. As described in the following subsections, in the context of SAR ATR deep learning includes convolutional neural networks, restricted Boltzmann machines, recurrent neural networks, stacked autoencoders, and hybrid strategies combining deep learning methods. An overview of the methods presented is shown in Table 8.

Convolutional Neural Networks (CNN). Gao et al. [90] enhanced the effectiveness of their CNN by employing a cost function during training that combines cross-entropy with the class separability information. The latter comprises of intraclass compactness and interclass separability. Yue et al. [91] proposed a CNN algorithm that combines both supervised and unsupervised training. Specifically, the unsupervised pipeline utilizes CNN to obtain the class probabilities of the unlabeled samples. Then, the class probabilities are optimized via a hard thresholding process, which are then used to calculate the scatter matrices of the linear discriminant analysis (LDA) method. Finally, the loss function of the supervised pipeline of the CNN is adapted by the scatter matrices created from the unsupervised pipeline. Chen et al. [6] and Wang et al. [92] proposed a CNN that comprises only of sparsely connected layers (convolutional and pooling layers) neglecting the typical fully connected layers. This strategy reduces the algorithm’s free parameters and thus overfitting due to limited training images. Their all-convolutional networks (A-ConvNets) are unarguably one of the top-performing SAR ATR techniques.

Wang et al. [93] realized the influence of speckle-noise to SAR ATR and reduce its impact by introducing a despeckling CNN as a preprocessing step applied on the entire SAR dataset. The latter CNN comprises five convolutional layers followed by a ReLU activation function while pooling layers are not used to preserve the feature map size through the entire network. For completeness, a ReLU activation function embedded within a ReLU layer aims at increasing the nonlinearity of a CNN by applying an individual truncation process on every input to that layer. The advantages of ReLU against the classic tanh activation function are the reduction in training time [94] and incorporating a purely supervised training scheme avoiding the need for unsupervised pretraining [95]. Kwak et al. [96] also considered the influence of speckle-noise and proposed a speckle noise-invariant CNN that employs a regularization term to minimize the feature variations caused by this type of noise. The regularization term minimizes the CNN features between the raw and the despeckled image variant and contributes to the final ATR classification score in a weighted manner along with the raw image classification output.

One subclass of CNN-based approaches involves CNN distillation to ultimately create a compact and shallow CNN that has the classification power of deeper CNNs. For example, Min et al. [97] extended the typical distillation scheme into the SAR ATR domain by adopting a student-teacher paradigm, where the teacher is a deep 18-layered CNN and the student a shallow two-layered ternary network (weights are -1, 0, 1). During training, the teacher’s output vector after the SoftMax processing is used to train the student’s CNN. The student’s loss function is the sum of the cross-entropy between the student network output and the ground truth target label, and the cross-entropy between the teacher network output and the ground truth target label. The student CNN is fully trained when its loss function minimizes. Another approach toward lightweight designs is proposed by Zhang et al. [98], which adapts the A-ConvNets [6], [92] to exploit a pruning and a knowledge distillation scheme. Their method is a three-stage process where initially the A-ConvNet is forced to a lossy compression by pruning the feature maps. Then, the CNN’s accuracy is recovered via a student–teacher type knowledge distillation concept, where the teacher is the original un-pruned A-ConvNet and the student its pruned variant. Trials demonstrated that the pruned A-ConvNets completely preserves its SAR ATR capability compared to the original A-ConvNets, despite being 68.7 times compressed and thus 2.5 faster to execute.

Another subclass of CNN-based approaches exploits data augmentation to enhance CNN’s classification performance by training it on more templates of various nuisances. Typical approaches create the additional templates manually, while more recent approaches rely on a Generative Adversarial Network (GAN) [99]-[101]. For example, Wagner [102] manually generated artificial training data that presents elastic distortion, rotation, and affine transformations simulating a changing depression angle or an incorrectly estimated aspect angle, aiming at increasing the robustness against these nuisances. Similarly, Ding et al. [103] suggested augmenting the training dataset by generating images that include speckle noise, several translations, and pose variations. Experimental results demonstrate that indeed, data augmentation enhances the ATR classification rates. Yan [104] extended the training images by manually generating speckle-noisy samples at different signal-to-noise ratios (SNRs), multisresolution representations, and partially occluded images. Furthermore, the training samples are preprocessed to reduce the interferences between the original and the augmented
| No. | Reference | Type | SAR chip size | Pre-processing | Data augmentation | Main features |
|-----|-----------|------|---------------|----------------|------------------|--------------|
| 1   | Zhang et al. [98] | CNN | 88x88 | - | CNN distillation of the A-ConvNets | |
| 2   | Huang et al. [120] | CNN | 88x88 | - | Random cropping | two-stream CNN |
| 3   | Gao et al. [90] | CNN | 64x64 | - | translation | cost function for inter-class and intra-class separability enhancement |
| 4   | Yue et al. [91] | CNN | 64x64 | - | - | combination of supervised and unsupervised learning |
| 5   | Min et al. [97] | CNN | 128x128 | - | - | CNN distillation, ternary student network |
| 6   | Cui et al. [105] | CNN | 64x64 | - | GAN | Employs Leaky ReLU layers |
| 7   | Zhang et al. [127] | CNN and LSTM | 68x68 | - | Random crop for translation invariance | |
| 8   | Sun et al. [108] | CNN | 64x64 | L₂ normalization | Angular GAN | residual network algorithm |
| 9   | Pei et al. [122] | CNN | 90x90 | orientation alignment | - | multiple target views (4) |
| 10  | Shang et al. [121] | CNN | 70x70 | - | Random clipping | Index table to store templates |
| 11  | Yan [104] | CNN | 72x72 | power function enhancement, normalization, downsampling | speckle noise, multi-resolution, and partial occlusion | manual data augmentation |
| 12  | Wang et al. [93] | CNN | 88x88 | noise de-speckling | - | five convolutional layers |
| 13  | Zhong et al. [118] | CNN | 88x88 | - | Random cropping | transfer learning on AlexNet |
| 14  | Kwak et al. [96] | CNN | 88x88 | - | speckle noise | speckle noise-invariant CNN |
| 15  | Tian et al. [119] | CNN | 96x96 | - | - | |
| 16  | Kechagias-Stamatis [114] | CNN | 128x128 | - | - | multidimensional feature tensor of a shallow VGG layer |
| 17  | Zhong and Ettinger [7] | CNN | 128x128 | - | Left-right image flipping | pose-aware regression function |
| 18  | Amrani and Jiang [117] | CNN | 128x128 | - | - | Concatenating three layers of various depths of the VGG network |
| 19  | Kechagias-Stamatis et al. [116] | CNN | 128x128 | - | - | clustering AlexNet layers, SVM for classification |
| 20  | Zhang et al. [129] | Bi-LSTM | 128x128 | - | - | Multiple views (4) |
| 21  | Chen et al. [6] | CNN | 88x88 | - | - | convolutional and pooling layers only |
| 22  | Huang et al. [126] | CNN and SAE | 128x128 | - | - | Transfer learning |
| 23  | Wang et al. [92] | CNN | 88x88 | - | - | convolutional and pooling layers only |
| 24  | Wagner [102] | CNN | 128x128 | elastic Distortion, affine transformation, rotation | manual data augmentation, SVM for classification |
| 25  | Huang et al. [125] | RBM | 90x90 | - | - | biologically inspired model |
| 26  | Kang et al. [14] | SAE | 128x128 | - | - | feature fusion |
| 27  | Ding et al. [103] | CNN | 128x128 | Pose, translation, Speckle noise | manual data augmentation |
| 28  | Deng et al. [13] | SAE | 128x128 | image normalization | - | - |
training imagery by cropping the target to exclude background and then enhancing the image via a power function. The resulting image is then normalized and downsampled before being input to the CNN. Zhong and Ettinger [7] augmented the CNN’s training process by utilizing a pose-aware regression function that aims at minimizing the weighted sum between the CNN estimated target class score, and the pose difference between the predicted and the ground truth rotation of the target.

On the contrary, Cui et al. [105] suggested data augmentation via a Wasserstein Generative Adversarial Nets scheme [106] with a gradient penalty (WGAN-GP) [107]. WGAN extends the typical GAN by incorporating a Wasserstein distance metric to measure the distance between the distribution of generated data and the real data. WGAN-GP further extends a WGAN by applying the Lipschitz constraints instead of the weight clipping of WGAN, affording faster convergence, and generating higher quality samples than the original WGAN. Once the SAR imagery is augmented, it is then input to a typical CNN algorithm that exploits Leaky ReLU activations. Similarly, Sun et al. [108] established an angular rotation GAN, where the attribute between the raw and the simulated SAR imagery is the target aspect angle. The synthetic images are then utilized to train a residual network (ResNet). Shi et al. [109] augmented the training dataset by generating superresolution SAR imagery via the GAN network. Up-sampling of the low-resolution images is created by introducing the two Pixel Shuffler layers after the residual block structure.

A number of approaches extend the usability of current CNNs trained for visual imagery into the SAR domain. These techniques commonly use the AlexNet [94], VGG [110], or ResNet [111], which are pretrained on the ImageNet dataset [112] that includes millions of optical images. The concept of this type of SAR ATR methods is to employ the knowledge of these CNNs on mid-level image representations such as the SAR imagery either by directly utilizing the raw weights of these pretrained CNNs or by exploiting the transfer learning technique (TL) [113] to fine-tune the raw weights incorporating knowledge from the SAR data domain. For the former case, Kechagias–Stamatis [114] exploited directly the weights of the pretrained VGG by clustering the VGG layers and extracting a multidimensional feature tensor of a shallow cluster. The reasoning is that shallow convolutional layers reveal the coarse object’s features and thus the shallow feature maps of the pretrained VGG are able to bridge the SAR—visual data domain gap. The target class is derived utilizing a near-est-neighbor feature matching strategy exploiting the cosine similarity metric between the target and the training imagery feature tensors. Similarly, Kechagias–Stamatis et al. [115] clustered AlexNet and combine it with a multiclass support vector machines (SVM) [116] for target classification. Amrani and Jiang [117] concatenated three layers of various depths of the VGG network features and create an index table containing these features along with the LDMLT similarity metric that measures the feature space of the training samples. Oppositely, Zhong et al. [118] applied the TL technique on the fifth convolutional layer of AlexNet and also amend AlexNet with an additional convolutional and a pooling layer. Their results demonstrated that TL attains a higher classification performance compared to completely training the same CNN. Finally, they further enhance the performance by employing data augmentation by randomly cropping the SAR chip.

In contrast to typical CNN pipelines, Tian et al. [119] suggested a CNN where the weights of each layer are manipulated from the previous layers by establishing a strong intrakernel relationship. Huang et al. [120] recommended a two-stream CNN to extract multilevel features, where the local stream uses global maximum pooling for significant local feature extraction and the global stream uses large-step convolution kernels for global feature extraction. Shang et al. [121] incorporated in their CNN algorithm an index table to store the training samples’ spatial features and suggest the M-Net algorithm. Then, for a query SAR image, this method calculates the spatial similarity between the query image features and the stored index table. To improve M-Net’s convergence, the authors proposed a two-staged training process. First, initial training is performed on solely the CNN module of the M-Net to initialize the CNN’s parameters. Second, the initialized parameters from step one are then applied to the M-Net to update CNN’s parameters and fine-tune the entire M-Net.

Opposing to the methods presented that employ a single target image, several solutions increase the classification rates by exploiting multiple aspects from the target. For example, Pei et al. [122] suggested a CNN topology that relies on multiple image views, where each aspect poses a distinct input. The CNN progressively fuses the feature maps from each image aspect to the last layer of the previous fusion, to ultimately classify the target. The fusion pipeline is presented in Figure 5.
Stacked Autoencoders (SAE). Stacked autoencoders (SAE) are sequentially placed autoencoders (AE) (see Figure 6), where the latter is a neural network that comprises an encoder and a decoder. The encoder remaps the input data to a hidden representation \( h \) according to

\[
h = f(x) = s_f(W_x + b_h)
\]

where \( s_f \) is a nonlinear activation function, e.g., sigmoid or ReLU, \( W \) is a weight matrix and \( b_h \) a bias vector. Accordingly, the decoder remaps \( h \) to a reconstructed version of the input data by minimizing a cost function. In terms of SAR ATR, only the encoding part on an AE is used, which is then input to a classification scheme, e.g., SVM. AEs are employed in a stacked fashion to increase descriptiveness.

Kang et al. [14] suggested a feature fusion process that involves 23 texture-based properties of the SAR image with features extracted via a three-patch local binary pattern (TPLBP) process. The former features include properties such as area, the bounding box, convex hull, etc., while the latter involves a set of classic local binary patterns (LBP) distributed uniformly on a ring of radius \( r \) and one LBP is placed at the center of the ring. Features are then cascaded and input to an SAE. Deng et al. [13] reduced overfitting by modifying the reconstruction error of SAE by adding a Euclidean distance restriction for the hidden layer features followed by a Dropout layer. Other autoencoder-based solutions are influenced by the human visual cortical system [123] or are combined with a Synergistic neural network concept [124].

Restricted Boltzmann Machine (RBM). Huang et al. [125] are inspired by the human cognition system and suggest a convolutional restricted Boltzmann machine (CRBM) algorithm that learns the episodic features which are then used to extract the semantic features of a SAR image. For the episodic feature extraction, a convolutional deep belief network is trained to segment the target’s SAR reflection and its shadow. These regions are then input to the semantic feature extraction module that considers area, length, slope, and gray value calculation. The episodic features are further utilized to enhance the ATR performance by estimating the aspect angle of the observed target. Training instances with aspect angles that exceed a \( \pm 10^\circ \) from the target aspect angle are discarded during the classification process.

Hybrid Deep Learning Techniques. Huang et al. [126] suggested an ATR algorithm that combines the CNN and SAE concepts along with the transfer learning technique. The CNN pipeline consists of a classification and a reconstruction module that is used as feedback to estimate the reconstruction loss. To avoid overfitting, the authors apply data augmentation and train the reconstruction module of the CNN pipeline with stacked convolutional autoencoders (SCAE). Chen and Wang [17] also combined the concepts of SAE and CNN by involving a single layer-CNN that is trained on randomly sampled image patches utilizing an unsupervised sparse autoencoder.

Zhang et al. [127] combined a residual (ResNet) network and a bidirectional long short-term memory (BiLSTM) network. Fusing these two types of networks aims at extracting the target’s scattering features from a single-aspect image (ResNet module) and extend their encoding over several adjacent multisaspect images through the recurrent learning (BiLSTM module). The latter network is a variant of the popular long short-term memory (LSTM) layer. For an analysis of the operating principles of LSTM layers, the reader is referred to [128].

**HYBRID REFLECTIVITY ATTRIBUTE-BASED METHODS**

Table 10 presents the hybrid reflectivity attribute-based algorithms. For example, Zhang et al. [129] combined the concepts of feature-based SAR ATR with deep learning. Specifically, the authors propose a multisaspect algorithm where Gabor filters and three-patch local binary patterns extract comprehensive spatial features, which are then input to a multilayer perceptron network followed by a bidirectional LSTM recurrent neural network. Lv and Liu [130] proposed a novel data augmentation algorithm, where the ASC are extracted using a sparse representation (SR) scheme [52], [131]. ASCs are created with various \( \theta \) parameters (2), the quality of which is evaluated based on a quality factor. Finally, the augmented dataset is used to train a CNN.

Spurred by the classification enhancement of decision fusion, current literature suggests various strategies of that type. For example, Karine et al. [132] combined the concepts of local features, saliency maps, and SRC. Specifically, given a SAR image, they create a saliency map inspired by the human visual system to establish a region of interest (ROI) type of area containing the target. Then, the scale-invariant feature transform (SIFT) [133] (a 2-D local keypoint detector and descriptor) is applied on the
ROI and for each detected keypoint a feature vector is created. The latter ones are converted into a matrix that is used as a dictionary during the SRC process. Given that for this method each SAR image provides a 2-D signature rather than a 1-D, which is the norm for SRC, the authors employ a multitask SRC, where the reconstruction error originates from a block of atoms. Sun et al. [134] used the original SAR image and the SIFT features, where the former representation incorporates intensity information and the latter gradient information. Both representations are used jointly [135], while the learned dictionary relies on the LC-KSVD algorithm [77]. Gishkori and Mulgrew [136] proposed a similar solution but utilize a set of rotation invariant Pseudo Zernike Moments (PZM) that are generated from the entire SAR image. Then, the PZM are employed to create a dictionary, which is used during the SRC process. Ding et al. [137] suggested a decision-level fusion of a two-stream SRC pipeline. The first SRC utilizes the SAR imagery, while the second SRC the azimuth sensitivity image (ASI) which reveals the target sensitivity at an azimuth angle. Finally, the target identity is defined by comparing the classification scores of both SRC streams. Zhang et al. [138] generated multilayer 2-D slices of the SAR image along with the amplitude direction and then calculate a set of Zernike moments (ZM) [2] for each slice. Finally, an SRC scheme is employed exploiting the ZM feature vector.

In [139], Kechagias-Stamatis and Aouf combined the strengths of SRC and CNN in a decision fusion scheme. For the SRC module, a novel adaptive elastic net type of optimization is proposed that balances the advantages of $l_1 - norm$ and $l_2 - norm$ depending on a Gaussian mixture model (GMM) [140] analysis of each SAR image. Regarding the CNN module, the authors suggest utilizing the AlexNet response map originating from a shallow layer, while target classification is relying on a multiclass SVM scheme. It should be noted that AlexNet is not retrained or any transfer learning is applied, but rather the pretrained weights of AlexNet are employed. Finally, the target identity is determined based on a decision-level scheme that adaptively changes its fusion weights. In [141], Ding et al. combine the concepts of SRC and ASC by hierarchically fusing the global and local features of a SAR image. Hence, in the first stage, the Gaussian random projection [142] features are used as global features that are then employed for an SRC process. If the decision is reliable, i.e., the reconstruction error does not exceed a threshold, then the classification process terminates, while if not, local features based on the ASC are extracted. The target identity is defined by the template ASC closest to the target ASC based on the Hungarian algorithm.

**SUMMARY**

Tables 4 to 10 give a summary of the reflectivity attributed-based methods. This category of techniques exploits the raw SAR image without remapping it into a different data domain.

- a) Most recent papers involve deep learning structures with the vast majority relying on convolutional neural networks.
- b) CNN-based architectures incorporate various schemes ranging from manual or GAN-based data augmentation to CNN distillation strategies, transfer learning and fully exploiting pretrained state-of-the-art deep networks.

| No | Reference | Type | SAR chip size | Data augmentation | Main features |
|----|-----------|------|---------------|-------------------|---------------|
| 1  | Zhang et al. [129] | features and deep learning | 128x128 | - | Gabor filters, LBP, LSTM |
| 2  | Lv and Liu [130] | CNN | 128x128 | ASC-based various orientations | - |
| 3  | Kechagias-Stamatis and Aouf [139] | CNN, SRC | 80x80 | - | decision fusion |
| 4  | Ding et al. [141] | ASC, SRC | 128x128 | - | feature fusion |
| 5  | Karine et al. [132] | features, SRC | 128x128 | - | decision fusion |
| 6  | Gishkori and Mulgrew [136] | Zernike moments, SRC | 96x96 | - | - |
| 7  | Ding et al. [137] | azimuth sensitivity, raw data | 128x128 | - | decision fusion |
| 8  | Zhang et al. [138] | Zernike moments, SRC | 70x70 | - | 2D slices of the SAR image along the amplitude direction |
| 9  | Sun et al. [134] | features, raw data | 65x65 | - | LC-KSVD learned dictionary |
c) Sparse classification representation for SAR ATR is also commonly used involving various $l_p - \text{norm}$ optimizations.

d) Hybrid reflectivity attribute-based methods are also of great interest as these aim at fusing the strengths of the involving disciplines.

**TRANSFORM-BASED METHODS**

Opposing to the reflectivity attribute-based methods, the transform-based strategies remap the raw SAR imagery into a different data domain on which the ATR pipeline is applied.

**FREQUENCY DOMAIN**

Dong et al. [143] remapped the SAR imagery into the frequency domain utilizing the discrete Fourier transform (DFT). Then, the authors apply the typical SRC employing an $l_1 - \text{norm}$.

**GRASSMANN MANIFOLD**

The monogenic signal is a multidimensional generalization of the analytic signal, where the latter is the combination of the signal itself with its Hilbert transformed version, i.e., a submanifold of a Grassmannian manifold [144]. For a 1-D signal $f(x)$, the analytic signal is

$$f_a(x) = f(x) + jH\{f\}(x) = A(x)\exp\{j\varphi(x)\}$$

(13)

where $H\{\cdot\}$ is the Hilbert transform, $A(x)$ the local amplitude, and $\varphi(x)$ the local phase. In a polar-coordinate frame, $A(x)$ and $\varphi(x)$ are linked to a local quantitative measure and qualitative information of a signal, respectively, [145].

Accordingly, the monogenic signal employs a scalar-to-vector extension of Hilbert transform, the Riesz transform. Hence, a 2-D monogenic signal is defined as

$$f_{m}(z) = f(z) - (i, j)f_z(z)$$

(14)

where $f(z)$ is the original signal, $i, j$ are imaginary units, $\{i, j, 1\}$ an orthonormal basis of $\mathbb{R}^3$, and $f_z(z)$ its Riesz transformed version with

$$f_z(z) = h_r(z) * f(z) = [h^1_r(z) * f(z), h^2_r(z) * f(z)]^T$$

(15)

where the Riesz kernel is

$$h_r(z) = -\frac{z}{2\pi|z|^3}$$

(16)

and $h^1_r(z), h^2_r(z)$ are the first and second-order Riesz kernels, respectively. However, the original signal has a finite length with periodic spectra and thus is infinitely extended via a band-pass filter

$$f_m(z) = (f * h_{bp})(z) - (i, j) * h_{bp}f_z(z)$$

(17)

where $h_{bp}$ is a band-pass filter. In the context of SAR ATR, the monogenic signal of (14) is decomposed into its real part, the $i$-imaginary part, and the $j$-imaginary part, which are then further processed in a target recognition pipeline.

Dong et al. [146] generated the monogenic signal utilizing a log-Gabor filter. The monogenic signal is then employed for a dictionary learning strategy that jointly considers SRC and low-rank representation, with the former aiming at restricting the representation and the latter limiting the feasible set of dictionaries. Dictionary learning incorporates several class-specific subdictionaries, which are then combined into a single dictionary. Zhou et al. [147] reduced the impact of the information redundancy between the scales and the heterogeneous nature of the three components of the monogenic signal, by proposing a scale selection method based on a weighted multitask joint sparse representation. The scale selection module relies on a Fisher score-based evaluation of each scale and monogenic component, where the scales and components with a high score are further utilized to produce a dictionary. The latter is then used in a multitask joint sparse representation classification framework. Dong et al. [148] embedded Grassmann manifolds into an implicit reproducing kernel Hilbert space (RKHS) and create an overcomplete dictionary. In this technique, the Grassmann manifolds are build using steerable wavelet coefficients.

**COMPRESSIVE SENSING (CS)**

Compressing sensing aims at recovering a signal that has been remapped from the originating domain to a domain where the signal is sparse, using a nonadaptive linear projection. Signal recovery is achieved via an $l_1$-norm optimization process. In the context of SAR ATR, multitask CS (MtCS) [149] exploits the statistical correlation among multiple target views to recover the target’s signature that is then used for target recognition under a compressive sensing scheme. Jin et al. [150] involved three features, namely the principal component analysis features (PCA) [151], elliptical Fourier descriptors (EFD) [152], and the azimuthal sensitivity image. The former two are extracted representing the intensity distribution and the target shape, respectively, while the latter is generated and describes the electromagnetic scattering characteristics of the target. The three features are then utilized in an MtCS scheme to derive the target identity. Liu and Yang [153] suggested a feature-fusion scheme that combines the PCA, the nonlinear extension of PCA namely the Kernel principal components analysis (KPCA) [154], and the NMF features. Finally, these features are input to an MtCS scheme.
Table 11 gives a summary of the transformed-based methods. In contrast to the reflectivity attribute techniques, this category remaps the original SAR imagery into a different data domain.

a) Most recent papers involve Grassmann manifolds and compressive sensing.

b) Quantitatively, this category is less popular mainly due to the imbalance between the data domain remapping process complexity and the ATR performance gain over the methods of the section “Encoding and Classifying SAR Imagery” (see also “Performance Analysis”).

c) Sparse classification representation for SAR ATR is also commonly used involving various $l_p - norm$ optimizations.

d) Hybrid reflectivity attribute-based methods are also of great interest as these aim at fusing the strengths of the involving disciplines. The majority of the proposed ATR pipelines rely on recovering the sparsest solution among the available ones, either by employing an SRC or a CS strategy.

| N° | Reference       | Type           | SAR chip size | Pre-processing | Main features                                      |
|----|-----------------|----------------|---------------|----------------|----------------------------------------------------|
| 1  | Dong et al. [148] | Grassmann manifold | 64x64        | -              | employing steerable wavelet coefficients            |
| 2  | Dong et al. [143] | frequency domain  | 128x128      | -              | SRC classification                                  |
| 3  | Dong et al. [146] | Grassmann manifold | 128x128      | Log-Gabor-filtering | SRC                                                  |
| 4  | Zhou et al. [147] | Grassmann manifold | 64x64        | Log-Gabor-filtering | multi-task joint SRC                                |
| 5  | Jin et al. [150]  | CS              | 128x128      | -              | PCA, elliptical Fourier descriptors, and azimuthal sensitivity input to Multitask CS |
| 6  | Liu et al. [149]  | CS              | 128x128      | -              | Multitask CS                                        |

As already noted (see section “Moving and Stationary Target Acquisition and Recognition Dataset”), despite the MSTAR being a typical SAR ATR benchmarking dataset, it has quite a few deficiencies, one of which is the strong background correlation between the template and the target images imposing high true positive ATR rates despite the target absence from the scene [35], [36]. Thus, several algorithms crop or segment the target from the scene leading to several SAR image patch sizes ranging from 128 x 128 (the original image size) down to 64 x 64 pixels. Given that background correlation, and thus the SAR image size may indeed affect the overall ATR performance of a method, comparing methods that employ different SAR image sizes is not fair. Additionally, simply manipulating the SAR images to meet the algorithm’s input image size constraints is still not an optimum choice as the entire parameter tuning of each method is linked to a specific image size. Finally, reimplementing and retuning each ATR technique is beyond the scope of this article. Therefore, for fairness in the following sections, we highlight the SAR patch size involved.

EVALUATING SOC SUBSETS

The most commonly used 10-class SOC trial is SOC 1 (see Table 3). Its main feature is utilizing during training and source code of the SAR ATR techniques challenged in this article is not available and reimplementing these ATR architectures may lead to solutions that underperform compared to their original implementation. Thus, given that the training and testing sets are well defined and are the same for each competitor method (as presented in Table 3), we compare the reported by the literature $acc$ of each SAR ATR method challenged. For completeness, it is worth mentioning that the papers that do not conform with the training and testing setup of Table 3 are not involved in this evaluation as such a comparison would not be valid.

For this work, each method will be evaluated based on the accuracy ($acc$) metric

$$acc = \frac{\sum Truepositive + \sum TrueNegative}{\sum TotalPopulation}.$$  \hspace{1cm} (18)

However, despite the original MSTAR dataset offers a decoy object, in the vast majority of the relevant literature, this object is omitted during the trials and therefore $\sum TrueNegative = 0$. It should be noted that the open-source code of the SAR ATR techniques challenged in this article is not available and reimplementing these ATR architectures may lead to solutions that underperform compared to their original implementation. Thus, given that the training and testing sets are well defined and are the same for each competitor method (as presented in Table 3), we compare the reported by the literature $acc$ of each SAR ATR method challenged. For completeness, it is worth mentioning that the papers that do not conform with the training and testing setup of Table 3 are not involved in this evaluation as such a comparison would not be valid.

As already noted (see section “Moving and Stationary Target Acquisition and Recognition Dataset”), despite the MSTAR being a typical SAR ATR benchmarking dataset, it has quite a few deficiencies, one of which is the strong background correlation between the template and the target images imposing high true positive ATR rates despite the target absence from the scene [35], [36]. Thus, several algorithms crop or segment the target from the scene leading to several SAR image patch sizes ranging from 128 x 128 (the original image size) down to 64 x 64 pixels. Given that background correlation, and thus the SAR image size may indeed affect the overall ATR performance of a method, comparing methods that employ different SAR image sizes is not fair. Additionally, simply manipulating the SAR images to meet the algorithm’s input image size constraints is still not an optimum choice as the entire parameter tuning of each method is linked to a specific image size. Finally, reimplementing and retuning each ATR technique is beyond the scope of this article. Therefore, for fairness in the following sections, we highlight the SAR patch size involved.

EVALUATING SOC SUBSETS

The most commonly used 10-class SOC trial is SOC 1 (see Table 3). Its main feature is utilizing during training and
testing all major model variants of BMP2 and T72. From Figure 7, it is obvious that the governing methods are CNN and SRC-based, with an exception of method 4-3 (i.e., Table 4, method 3). Feature-based methods are also available, but these tend to provide lower acc rates. CNN methods tend to be more appealing than their SRC counterparts, with the performance gain of the former being in the order of 4%. Specifically, for a SAR patch size of 128 x 128 pixels, methods 9-18 and 9-16 present the highest ATR performance demonstrating that partially exploiting structures of state-of-the-art deep networks used in the visual domain, i.e., the VGG for both these methods, can be highly beneficial. Interestingly, both 9-18 and 9-16 do not involve any Transfer Learning but employ the original VGG kernel weights. However, for smaller SAR images (96x96 pixels), pure CNN provides a relatively lower acc (method 9-15), while for an 80 x 80 pixel size, high ATR rates require more complex solutions such as a decision fusion scheme combining CNN and SRC (method 10-3). CNN-based methods offer a higher ATR performance than SRC because the latter downsamples the SAR image to create an overcomplete dictionary discarding some of the target’s features, and ultimately diminishes the descriptiveness of the SRC method. Regarding the feature-based techniques, these are highly dependent on the originating feature extraction and description concept of each method. However, since each feature description method was originally designed for images in the visual domain, the performance of the feature-based methods for SAR ATR depends on the capability of each technique to handle equally well the SAR and the visual data domain.

The next most common 10-class SOC configuration used in the literature is SOC 2. Compared to SOC-1, this subset employs only a single target model for BMP2 and T72 rather than all available ones and is therefore considered of relatively lower complexity. Indeed, regardless of the ATR method and SAR image size used, the average acc on SOC-1 is 96.5%, while on SOC-2 99.1%.

Figure 8 clearly shows that deep learning-based algorithms attain a minor performance gain over their counterparts. However, in contrast to the SOC-1 evaluation subset, here the acc difference is only 0.7% between the CNN and the SRC-based methods. Additionally, Figure 8 shows that acc is less related to the SAR image size, as various image sizes are equally spread throughout the acc performance plot. Hence, the ATR performance is less affected by the background correlation as top-performing solutions involve SAR image patches with various background correlation levels. However, as stated earlier, the complexity of this target subset is lower allowing higher acc rates.

Regarding the 3-class SOC trials, SOC 1 is the most evaluated configuration (for details see Table 3). This subset is similar to the 10-class SOC 1, in terms that for the BMP2, and T72 targets all major models are exploited both during training and testing. Thus, interestingly, the top-3 performing techniques are the same for both these SOC subsets, and indeed CNN-based algorithms are still performing relatively better than their competitors, with SRC and ASC to follow closely (see Figure 9). The CNN methods are more efficient as the fine details of the various training and testing models are properly encrypted by the convolutional layers. On the contrary, SRC-based pipelines...
downsample the raw SAR image during the creation of the overcomplete dictionary, discarding the fine details that differentiate the intraclass target variations. Similarly, ASC methods highlight the strong scatterers and discard the weak ones neglecting the fine details.

**EVALUATING EOC SUBSETS**

The vast majority of the SAR ATR papers also evaluate the effectiveness of the proposed methods against extended operation conditions. The most referenced EOC scenarios are the ones presented in Table 3 that involve depression angle variation, target version variation, and several SAR image nuisances such as Gaussian noise, occlusion, and resolution variation at various levels.

The most referenced EOC trial investigates the robustness under depression angle variation where the ATR methods are trained on the 2S1, BRDM2, and the ZSU23/4 targets at 17° depression angle and tested on the same targets at 30° and 45° while some methods are also challenged at 15°. Thus, for the sake of uniformity, in this article, we present the performance of current methods only at 30° and 45° depression angles. In contrast to the SOC scenarios, Figure 10 shows that CNN has only one representative solution, while the vast majority is SRC based. This is because altering the depression angle creates on the target different electromagnetic reflection geometries that ultimately affect the SAR image. However, SRC during the overcomplete dictionary creation smoothens the local radar reflectivity parameters due to the depression angle variation and ultimately increases the robustness of the SRC methods. In contrast, the scattering centers are very much affected by the depression angle and thus techniques relying on ASC’s are not suitable for this type of trial. Regarding the background correlation, from Figure 10 it is evident that various SAR patch sizes may perform equally well. This is because the SRC, CS, and LMF methods smooth the effectiveness of the background correlation. Additionally, it is also important to note that despite the average acc at 30° is 98.15%, at 45° it is only 76.32% highlighting the poor robustness of the majority of the methods to such an extended depression angle variation. Despite that, method 7-11 still manages a high acc rate at 45° (91.72%) due to its hybrid-structured synthesis. A detailed acc performance is presented in Table 12.

---

**Table 12.**

| acc on EOC Depression Angle Variation | Patch size (pix.) | Table – N° reference | average | 30° | 45° |
|--------------------------------------|------------------|----------------------|---------|-----|-----|
| Li et al. [80]                       | 80               | 7-11                 | 95.21   | 98.71 | 91.72 |
| Ding et al. [141]                    | 128              | 10-4                 | 91.23   | 99.07 | 83.39 |
| Dong et al. [148]                    | 64               | 11-1                 | 87.95   | 98.20 | 77.70 |
| Jin et al. [150]                     | 128              | 11-5                 | 87.54   | 98.73 | 76.35 |
| Ren et al. [85]                      | 64               | 7-5                  | 87.05   | 98.29 | 75.82 |
| Chang and You [75]                   | 128              | 7-6                  | 87.05   | 99.07 | 75.03 |
| Zhang [70]                           | 128              | 7-9                  | 86.65   | 99.05 | 74.25 |
| Lv and Liu [130]                     | 128              | 10-2                 | 86.54   | 98.61 | 74.48 |
| Ren et al. [79]                      | 64               | 7-4                  | 86.52   | 97.02 | 76.03 |
| Ding and Wen [71]                    | 128              | 7-10                 | 85.45   | 98.84 | 72.06 |
| Zhang and Liu [72]                   | 128              | 7-12                 | 85.38   | 98.92 | 71.85 |
| Kechagias-Stamatis and Aouf [139]    | 80               | 10-3                 | 85.24   | 99.61 | 70.87 |
| Ding et al. [137]                    | 128              | 10-7                 | 84.98   | 98.72 | 71.25 |
| Yu et al. [89]                       | 64               | 8-4                  | 84.53   | 91.29 | 77.78 |
| Ding and Wen [84]                    | 128              | 6-7                  | 84.46   | 96.76 | 72.12 |
| Miao and Shan [83]                   | 128              | 7-2                  | 84.25   | 97.12 | 71.38 |
| Liu and Yang [153]                   | 128              | 11-6                 | 84.07   | 98.84 | 69.31 |
| Karine et al. [132]                  | 128              | 10-5                 | 52.45   | 68.58 | 36.32 |

**Figure 10.**
EOC depression angle variation scenario – acc performance overview (best seen in color).
Figure 11 presents the ATR performance on additive Gaussian noise variation at SNR levels of $\{10, 5, 0, -5, -10\} \text{ dB}$. Similar to the previous EOC scenarios, SRC is still an appealing method with ASC being more robust. However, fusing these two methods provides the highest acc value. Nevertheless, it should be noted that the top-performing solutions have only a negligible performance difference.

Quite a few papers also evaluate the robustness on several resolution variations ranging from $0.7 \times 0.7 \text{ m}$ down to $0.3 \times 0.3 \text{ m}$ with a step size of $0.1 \times 0.1 \text{ m}$. The majority of the competitor techniques attain acc rates in the order of 95% (see Figure 12), with the 9-16 method (CNN-based) affording a 100% acc. It is worth noting that the vast majority of the evaluated methods utilize the entire SAR image (128 x 128 pixels).

Evaluating SAR ATR algorithms against version variants is also common. Figure 13 presents the performance attained for scenarios involving training on the BMP2, T72, BTR60, and T62 targets, and testing on various BMP2 variants. For this type of scenario, a CNN type of solution manages the highest acc, with the majority of techniques being SRC-based. Interestingly, background correlation is of less influence. However, the latter is purely due to the testing conditions of each method, which is related to the SAR image input size and does not necessarily directly link with the overall performance of each method as each technique is only evaluated on a single SAR patch size.

Finally, occlusion variation has also been evaluated, with occlusion levels ranging from 10% up to 50% of the target with a step size of 10%. The performance plot of Figure 14 highlights that most method categories are capable of an attractive SAR ATR performance.

**SUMMARY**

Table 13 presents an overview of the performance per method for the most commonly referenced SOC and EOC trials. From the analysis of this section, it can be succinctly interpreted that:
Overall, deep learning-based methods are more robust on the SOC scenarios, while the SRC techniques attain higher classification rates on the EOC scenarios. Despite CNN belong to deep learning methods, the CNN algorithms utilized for SAR ATR are relatively shallow with an average of four convolutional layers. This is related to the information content per pixel, which is much less than the corresponding one of the visual imagery.

a) Overall, deep learning-based methods are more robust on the SOC scenarios, while the SRC techniques attain higher classification rates on the EOC scenarios. Despite CNN belong to deep learning methods, the CNN algorithms utilized for SAR ATR are relatively shallow with an average of four convolutional layers. This is related to the information content per pixel, which is much less than the corresponding one of the visual imagery.

b) CNN and SRC are appealing choices for both SOC and EOC scenarios, while the remaining techniques tend to present attractive ATR performance only on a subset of the scenarios. Indeed, ASC is robust only on EOC scenarios, while CS and hybrid reflectivity attribute pipelines manage high classification rates on some of the SOC and EOC scenarios.

c) Transform-based methods and specifically frequency domain and Grassmann manifolds have not been evaluated yet on the commonly used SOC and EOC trials.

d) Generally, the methods employing a 128x128 pixel SAR image tend to attain higher ATR rates. Though, comprehensive experiments are required to demonstrate if background correlation is indeed the major factor for such performance or the same SAR ATR pipelines are still capable of high acc rates at smaller SAR image patch sizes. It should be noted that simply feeding a SAR ATR algorithm with an image size different from the one originally designed for is not fair as the entire parameter tuning of the algorithm is based on a specific image size.

CONCLUSION

This article presents a unique survey of current state-of-the-art SAR ATR algorithms that employ the MSTAR dataset. A comprehensive analysis and comparative performance evaluation on most common SOC and EOC trials are also presented, highlighting the merits and demerits of each technique.

Future research directions may involve:

a) Benchmark protocol: Major deficiencies of the MSTAR database are the background correlation between the training and the testing images, and that all testing images are a priori known affecting the fine-tuning process of each method and biasing the true effectiveness of each technique. Therefore, creating a fixed benchmarking dataset originating from the MSTAR dataset that involves a common target segmentation scheme and several SOC and EOC testing images with publicly unknown class labels, would reveal the true effectiveness of each SAR ATR method. The latter could be conceptually related to the evaluation platform for visual odometry as presented in [155]. Additionally, standardizing the SOC and EOC trials would make direct performance comparison easier.

b) Exploiting current synthetic SAR MSTAR-based datasets: Recently, the synthetic and measured paired labeled experiment (SAMPLE) dataset has been proposed, as a complementary SAR ATR dataset [156]. Future work could utilize this dataset, which however may present the same background constraints as the original MSTAR dataset.

c) Explainable deep learning: As presented in the conclusion, each class of methods has its strengths and weaknesses, with CNN-based techniques being more robust for the SOC scenarios and SRC on the EOC scenarios. Spurred by that, future work should include further research on explainable deep learning structures (XAI) to realize the underlying mechanisms of such networks and ultimately exploit that knowledge to further enhance the performance of CNN and expand their appealing performance to the EOC scenarios.

d) Data augmentation: Proven the performance enhancement of data augmentation, it would be interesting to evaluate the effectiveness of several data augmentation methods in conjunction with current ATR methods such as ASC and SRC.

REFERENCES

[1] G. J. Gray, N. Aouf, M. A. Richardson, B. Butters, and R. Walmsley, “An intelligent tracking algorithm for an imaging infrared anti-ship missile,” Proc. SPIE, vol. 8543, Nov. 2012, Art. no. 85430L.

[2] O. Kechagias-Stamatis, N. Aouf, and D. Nam, “Multimodal automatic target recognition for anti-ship missiles with imaging infrared capabilities,” in Proc. Sensor Signal Process. Defence Conf., Dec. 2017, pp. 1–5.
Automatic Target Recognition on Synthetic Aperture Radar Imagery: A Survey

[3] O. Keчhagiaς-Stamatis, N. Aouf, and M. A. Richardson, “3D automatic target recognition for future LIDAR missiles,” IEEE Trans. Aerosp. Electron. Syst., vol. 52, no. 6, pp. 2662–2675, Dec. 2016.

[4] O. Keчhagiaς-Stamatis and N. Aouf, “Evaluating 3D local descriptors for future LIDAR missiles with automatic target recognition capabilities,” Imag. Sci. J., vol. 65, no. 7, pp. 428–437, Aug. 2017.

[5] O. Keчhagiaς-Stamatis, N. Aouf, G. Gray, L. Chermak, M. Richardson, and F. Oudyi, “Local feature based automatic target recognition for future 3D active homing seeker missiles,” Aerosp. Sci. Technol., vol. 73, pp. 309–317, Feb. 2018.

[6] S. Chen, H. Wang, F. Xu, and Y.-Q. Q. Jin, “Target classification using the deep convolutional networks for SAR images,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 8, pp. 4806–4817, Aug. 2016.

[7] Y. Zhong and G. Ettinger, “Enlightening deep neural networks with knowledge of confounding factors,” in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Oct. 2017, pp. 1077–1086.

[8] A. Profeta, A. Rodriguez, and H. S. Clouse, “Convolutional neural networks for synthetic aperture radar classification,” Proc. SPIE, vol. 9843, May 2016, Art. no. 98430M.

[9] D. A. E. Morgan, “Deep convolutional neural networks for ATR from SAR imagery,” Proc. SPIE, vol. 9475, 2015, Art. no. 94750F.

[10] X. Zhang, J. Qin, and G. Li, “SAR target classification using Bayesian compressive sensing with scattering centers features,” Prog. Electromagn. Res., vol. 136, pp. 385–407, 2013.

[11] S. H. Doo et al., “Aspect invariant features for radar target recognition,” IET Radar, Sonar Navig., vol. 11, no. 4, pp. 597–604, Apr. 2017.

[12] K. El-Darymli, P. McGuire, E. W. Gill, D. Power, and C. Moloney, “Holism-based features for target classification in focused and complex-valued synthetic aperture radar imagery,” IEEE Trans. Aerosp. Electron. Syst., vol. 52, no. 2, pp. 786–808, Apr. 2016.

[13] S. Deng, L. Du, C. Li, J. Ding, and H. Liu, “SAR automatic target recognition based on euclidean distance restricted autoencoder,” IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens., vol. 10, no. 7, pp. 3323–3333, Jul. 2017.

[14] M. Kang, K. Ji, X. Leng, X. Xing, and H. Zou, “Synthetic aperture radar target recognition with feature fusion based on a stacked autoencoder,” Sensors, vol. 17, no. 12, Jan. 2017, Art. no. 192.

[15] Z. Cui, J. Feng, Z. Cao, H. Ren, and J. Yang, “Target recognition in synthetic aperture radar images via non-negative matrix factorisation,” IET Radar, Sonar Navig., vol. 9, no. 9, pp. 1376–1385, Dec. 2015.

[16] M. Wilmanski, C. Kreucher, and J. Lauer, “Modern approaches in deep learning for SAR ATR,” Proc. SPIE, vol. 9843, May 2016, Art. no. 98430N.

[17] S. Chen and H. Wang, “SAR target recognition based on deep learning,” in Proc. Int. Conf. Data Sci. Adv. Analytics, Oct. 2014, pp. 541–547.

[18] R. Paladini, M. Martorella, and F. Beririzzi, “Classification of man-made targets via invariant coherency-matrix eigenvector decomposition of polarimetric SAR/ISAR images,” IEEE Trans. Geosci. Remote Sens., vol. 49, no. 8, pp. 3022–3034, Aug. 2011.

[19] D. Perissin and A. Ferretti, “Urban-target recognition by means of repeated spaceborne SAR images,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 12, pp. 4043–4058, Dec. 2007.

[20] G. Dong, N. Wang, and G. Kuang, “Sparse representation of monogenic signal: With application to target recognition in SAR images,” IEEE Signal Process. Lett., vol. 21, no. 8, pp. 952–956, Aug. 2014.

[21] F. Zhou, L. Wang, X. Bai, and Y. Hui, “SAR ATR of ground vehicles based on LM-BN-CNN,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 12, pp. 7282–7293, Dec. 2018.

[22] O. Keчhagiaς-Stamatis and N. Aouf, “A new passive 3-D automatic target recognition architecture for aerial platforms,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 1, pp. 406–415, Jan. 2019.

[23] C. Belloni, A. Balleri, N. Aouf, T. J. Merlet, and J. -M. Le Caillec, “SAR image dataset of military ground targets with multiple poses for ATR,” Proc. SPIE, vol. 10432, 2017, Art. no. 104320N.

[24] AFRL and DARPA, “Sensor data management system website, MSTAR database.” 1995. [Online]. Available: https://www.sdms.afrl.af.mil/index.php?collection=mstar accessed May 13, 2020.

[25] N. J. S. Stacy, D. P. Badger, A. S. Goh, M. Preiss, and M. L. Williams, “The DSTO Ingara airborne X-Band SAR polarimetric upgrade: First results,” Int. Geosci. Remote Sens. Symp., vol. 7, no. C, pp. 4474–4476, 2003.

[26] K. El-Darymli, P. McGuire, D. Power, and C. Moloney, “Target detection in synthetic aperture radar imagery: A state-of-the-art survey,” J. Appl. Remote Sens., vol. 7, no. 1, 2013, Art. no. 071598.

[27] K. El-Darymli, E. W. Gill, P. McGuire, D. Power, and C. Moloney, “Automatic target recognition in synthetic aperture radar imagery: A state-of-the-art review,” IEEE Access, vol. 4, pp. 6014–6058, 2016.

[28] G. Dong, G. Liao, H. Liu, and G. Kuang, “A review of the autoencoder and its variants: A comparative perspective from target recognition in synthetic-aperture radar images,” IEEE Geosci. Remote Sens. Mag., vol. 6, no. 3, pp. 44–68, Sep. 2018.

[29] Y. Yang, Y. Qiu, and C. Lu, “Automatic target classification experiments on the MSTAR SAR images,” in Proc. 6th Int. Conf. Softw. Eng., Artif. Intell. Netw. Parallel/Distrib. Comput., 2005, pp. 2–7.
[30] J. R. Diemunsch and J. Wissinger, “Moving and stationary target acquisition and recognition (MSTAR) model-based automatic target recognition: Search technology for a robust ATR,” in Proc. Algorithms Synthetic Aperture Radar Imagery V, Sep. 1998, pp. 481–492.

[31] J. C. Mosing and T. D. Ross, “Evaluation of SAR ATR algorithm performance sensitivity to MSTAR extended operating conditions,” Algorithms Synthetic Aperture Radar Imagery V, International Society for Optics and Photonics, vol. 3370, 1998.

[32] E. Blasch, U. K. Majumder, E. G. Zelnio, and V. J. Velten, “Review of recent advances in AI/ML using the MSTAR data,” in Proc. SPIE, vol. 11393, May 2020, Art. no. 113930C.

[33] E. R. Keydel, S. W. Lee, and J. T. Moore, “MSTAR extended operating conditions: a tutorial,” Proc. SPIE, vol. 2757, pp. 228–242, 1996.

[34] T. D. Ross, L. A. Westerkamp, E. G. Zelnio, and T. J. Burns, “Extensibility and other model-based ATR evaluation concepts,” Proc. SPIE, vol. 3070, vol. 3070, pp. 213–222.

[35] R. Schumacher and K. Rosenbach, “ATR of battlefield targets by SAR–Classification results using the public MSTAR dataset compared with a dataset by QinetiQ, UK,” in Proc. RTO SET Symp. “Target Identification Recognition. Using RF Syst.,” 2004, pp. 11–13, [Online]. Available: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:ATR+of+of+BHofD+B-HoD:+A+lightweight+and+fast+binary+descriptor+for+3D+object+recognition+and+registration,” in Proc. IEEE 14th Int. Conf. Netw. Sens. Control, May 2017, pp. 67–74.

[36] C. Belloni, N. Aouf, J.-M. Le Caillec, and T. Merlet, “Comparison of descriptors for SAR ATR,” in Proc. IEEE Radar Conf., Apr. 2019, pp. 1–6.

[37] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction,” in Proc. 17th Int. Conf. Pattern Recognit., 2004, pp. 28–31.

[38] S. Leutenegger, M. Chli, and R. Y. Siegwart, “BRISK: Binary robust invariant scalable keypoints,” in Proc. Int. Conf. Comput. Vis., Nov. 2011, pp. 2548–2555.

[39] O. Kechagias-Stamatis, N. Aouf, and L. Chermak, “B-HoD: A lightweight and fast binary descriptor for 3D object recognition and registration,” in Proc. IEEE 14th Int. Conf. Netw., Sens. Control, May 2017, pp. 37–42.

[40] M. Amrani, F. Jiang, Y. Xu, S. Liu, and S. Zhang, “SAR-oriented visual saliency model and directed acyclic graph support vector metric based target classification,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 11, no. 10, pp. 3794–3810, Oct. 2018.

[41] J. Harel, C. Koch, and P. Perona, “Graph-based visual saliency,” in Proc. Adv. Neural Inf. Process. Syst., 2007, pp. 545–552.

[42] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., May 2005, pp. 886–893.

[43] P. Bolourchi, M. Moradi, H. Demirel, and S. Uysal, “Improved SAR target recognition by selecting moment methods based on Fisher score,” Signal, Image Video Process., vol. 14, no. 1, pp. 39–47, 2020.

[44] C.-C. Chang and C.-J. Lin, “LIBSVM—A library for support vector machines,” 2019. [Online]. Available: https://www.csie.ntu.edu.tw/~cjlin/libsvm/. Accessed: Jul. 06, 2020.

[45] C. Clemente, L. Pallotta, D. Gaglione, A. De Maio, and J. J. Soraghan, “Automatic target recognition of military vehicles with Krawtchouk moments,” IEEE Trans. Aerosp. Electron. Syst., vol. 53, no. 1, pp. 493–500, Feb. 2017.

[46] J. Tan et al. et al., “Target recognition of SAR images by partially matching of target outlines,” J. Electromagn. Waves Appl., vol. 5071, pp. 1–17, 2018.

[47] B. Ding, G. Wen, C. Ma, and X. Yang, “Target recognition in synthetic aperture radar images using binary morphological operations,” J. Appl. Remote Sens., vol. 10, no. 4, Oct. 2016, Art. no. 046006.

[48] S. Yang, M. Wang, H. Long, and Z. Liu, “Sparse robust filters for scene classification of synthetic aperture radar (SAR) images,” Neurocomputing, vol. 184, pp. 91–98, Apr. 2016.

[49] I. M. Gorovyi and D. S. Sharapov, “Efficient object classification and recognition in SAR imagery,” in Proc. 18th Int. Radar Symp., Jun. 2017, pp. 1–7.

[50] L. C. Potter and R. L. Moses, “Attributed scattering centers for SAR ATR,” IEEE Trans. Image Process., vol. 6, no. 1, pp. 79–91, Jan. 1997.

[51] J. B. Keller, “Geometrical theory of diffraction,” J. Opt. Soc. Amer., vol. 52, no. 2, pp. 116–130, Feb. 1962.

[52] H. Liu, B. Jiu, F. Li, and Y. Wang, “Attributed scattering center extraction algorithm based on sparse representation with dictionary refinement,” IEEE Trans. Antennas Propag., vol. 65, no. 5, pp. 2604–2614, May 2017.

[53] T. Li and L. Du, “SAR automatic target recognition based on attribute scattering center model and discriminative dictionary learning,” IEEE Sens. J., vol. 19, no. 12, pp. 4598–4611, Jun. 2019.

[54] B. Ding, G. Wen, J. Zhong, C. Ma, and X. Yang, “A robust similarity measure for attributed scattering center sets with application to SAR ATR,” Neurocomputing, vol. 219, pp. 130–143, Jan. 2017.

[55] M. J. Gerry, L. C. Potter, I. J. Gupta, and A. Van Der Merwe, “A parametric model for synthetic aperture radar measurements,” IEEE Trans. Antennas Propag., vol. 47, no. 7, pp. 1179–1188, Jul. 1999.

[56] M. Jing, X. Zhou, and C. Qi, “Quasi-Newton iterative projection algorithm for sparse recovery,” Neurocomputing, vol. 144, pp. 169–173, Nov. 2014.

[57] J. Tsao and B. D. Steinberg, “Reduction of sidelobe and speckle artifacts in microwave imaging: The CLEAN technique,” IEEE Trans. Antennas Propag., vol. 36, no. 4, pp. 543–556, Apr. 1988.
[58] H. W. Kuhn, “The Hungarian method for the assignment problem,” Naval Res. Logist. Quart., vol. 2, no. 1–2, pp. 83–97, Mar. 1955.
[59] D. Gerogiannis, C. Nikou, and A. Likas, “Registering sets of points using Bayesian regression,” Neurocomputing, vol. 89, pp. 122–133, Jul. 2012.
[60] B. Ding, G. Wen, J. Zhong, C. Ma, and X. Yang, “Robust method for the matching of attributed scattering centers with application to synthetic aperture radar automatic target recognition,” J. Appl. Remote Sens., vol. 10, no. 1, 2016, Art. no. 016010.
[61] H.-C. Chiang, R. L. Moses, and L. C. Potter, “Model-based classification of radar images,” IEEE Trans. Inf. Theory, vol. 46, no. 5, pp. 1842–1854, Aug. 2000.
[62] G. Shafer, A Mathematical Theory of Evidence. Princeton, NJ, USA: Princeton Univ. Press, 1976.
[63] X. Zhang, “Noise-robust target recognition of SAR images based on attribute scattering center matching,” Remote Sens. Lett., vol. 10, no. 2, pp. 186–194, Feb. 2019.
[64] B. Ding and G. Wen, “Combination of global and local filters for robust SAR target recognition under various extended operating conditions,” Inf. Sci. (Ny.), vol. 476, pp. 48–63, 2019.
[65] B. Ding, G. Wen, X. Huang, C. Ma, and X. Yang, “Data augmentation by multilevel reconstruction using attributed scattering center for SAR target recognition,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 6, pp. 979–983, Jun. 2017.
[66] B. Ding, G. Wen, X. Huang, C. Ma, and X. Yang, “Target recognition in synthetic aperture radar images via matching of attributed scattering centers,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 10, no. 7, pp. 3334–3347, 2017.
[67] E. J. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489–509, Feb. 2006.
[68] Y. Wei, L. Jiao, F. Liu, S. Yang, Q. Wu, and G. Sanga, “Fast DDL classification for SAR Images with L_inf constraint,” IEEE Access, vol. 7, pp. 68991–69006, 2019.
[69] A. Quattoni, X. Carreras, M. Collins, and T. Darrell, “An efficient projection for LI∞ regularization,” in Proc. 26th Int. Conf. Mach. Learn., 2009, pp. 857–864.
[70] Z. Zhang, “Joint classification of multiresolution representations with discrimination analysis for SAR ATR,” J. Electron. Imag., vol. 27, no. 04, 2018, Art. no. 043030.
[71] B. Ding and G. Wen, “Target recognition of SAR images based on multi-resolution representation,” Remote Sens. Lett., vol. 8, no. 11, pp. 1006–1014, 2017.
[72] Z. Zhang and S. Liu, “Joint sparse representation for multi-resolution representations of SAR images with application to target recognition,” J. Electron. Magn. Waves Appl., vol. 32, no. 11, pp. 1342–1353, 2018.
[73] M. Liu and S. Chen, “SAR target configuration recognition based on the Dempster–Shafer theory and sparse representation using a new classification criterion,” Int. J. Remote Sens., vol. 40, no. 12, pp. 4604–4622, 2019.
[74] A. Sarkar et al., “Landcover classification in MRF context using Dempster–Shafer fusion for multisensor imagery,” IEEE Trans. Image Process., vol. 14, no. 5, pp. 634–645, May 2005.
[75] M. Chang and X. You, “Target recognition in SAR images based on information-decoupled representation,” Remote Sens., vol. 10, no. 1, 2018, Art. no. 138.
[76] J. Mairal, F. Bach, and J. Ponce, “Task-driven dictionary learning,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 4, pp. 791–804, Apr. 2012.
[77] Z. Jiang, Z. Lin, and L. S. Davis, “Label consistent K-SVD: Learning a discriminative dictionary for recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 11, pp. 2651–2664, Nov. 2013.
[78] M. Liu, S. Chen, X. Wang, F. Lu, M. Xing, and J. Wu, “SAR target configuration recognition via discriminative statistical dictionary learning,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 11, no. 11, pp. 4218–4229, Nov. 2018.
[79] H. Ren, X. Yu, L. Zou, Y. Zhou, and X. Wang, “Class-oriented local structure preserving dictionary learning for SAR target recognition,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., 2019, pp. 1338–1341.
[80] M. Li, Y. Guo, M. Li, G. Luo, and X. Kong, “Coupled dictionary learning for target recognition in SAR images,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 6, pp. 791–795, Jun. 2017.
[81] S. Gu, L. Zhang, W. Zuo, and X. Feng, “Projective dictionary pair learning for pattern classification,” in Proc. 28th Annu. Conf. Neural Inf. Process. Syst., 2014, pp. 793–801.
[82] H. Zhang, N. M. Nasrabadi, Y. Zhang, and T. S. Huang, “Multi-view automatic target recognition using joint sparse representation,” IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 3, pp. 2481–2497, Jul. 2012.
[83] X. Miao and Y. Shan, “SAR target recognition via sparse representation of multi-view SAR images with correlation analysis,” J. Electron. Magn. Waves Appl., vol. 33, no. 7, pp. 897–910, 2019.
[84] S. Ji, D. Dunson, and L. Carin, “Multitask compressive sensing,” IEEE Trans. Signal Process., vol. 57, no. 1, pp. 92–106, Jan. 2009.
[85] H. Ren, X. Yu, L. Zou, Y. Zhou, and X. Wang, “Joint supervised dictionary and classifier learning for multi-view SAR image classification,” IEEE Access, vol. 7, pp. 165127–165142, 2019.
[86] Y. Huang, G. Liao, Z. Zhang, Y. Xiang, J. Li, and A. Nehorai, “SAR automatic target recognition using joint low-rank and sparse multiview denoising,” IEEE Geosci. Remote Sens. Lett., vol. 15, no. 10, pp. 1–5, 2018.
[87] S. Dang, Z. Cui, Z. Cao, and N. Liu, “SAR target recognition via incremental nonnegative matrix factorization,” Remote Sens., vol. 10, no. 3, 2018, Art. no. 374.

[88] X. Zhang, Y. Wang, D. Li, Z. Tan, and S. Liu, “Fusion of multifeature low-rank representation for synthetic aperture radar target configuration recognition,” IEEE Geosci. Remote Sens. Lett., vol. 15, no. 9, pp. 1402–1406, Sep. 2019.

[89] M. Yu, G. Dong, H. Fan, and G. Kuang, “SAR target recognition via local sparse representation of multi-manifold regularized low-rank approximation,” Remote Sens., vol. 10, no. 2, pp. 1–22, 2018.

[90] F. Gao, T. Huang, J. Sun, J. Wang, A. Hussain, and Yang, E, “A new algorithm for SAR Image target recognition based on an improved deep convolutional neural network,” Cognit. Comput., vol. 11, no. 6, pp. 809–824, 2019.

[91] Z. Yue et al., “A novel semi-supervised convolutional neural network method for synthetic aperture radar image recognition,” Cognit. Comput., Mar. 2019, doi: 10.1007/s12559-019-09639-x.

[92] H. Wang, S. Chen, F. Xu, and Y.-Q. Jin, “Application of deep-learning algorithms to MSTAR data,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., Jul. 2015, pp. 3743–3745.

[93] J. Wang, T. Zheng, P. Lei, and X. Bai, “Ground target classification in noisy SAR images using convolutional neural networks,” IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens., vol. 11, no. 11, pp. 4180–4192, Nov. 2018.

[94] A. Krizhevsky, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural networks,” in Proc. Adv. Neural Inf. Process. Syst., Dec. 2012, pp. 1097–1105.

[95] X. Glorot, A. Bordes, and Y. Bengio, “Deep Sparse Rectifier Neural Networks,” in Proc. 14th Int. Conf. Artif. Intell. Statist., 2011, pp. 315–323, [Online]. Available: http://proceedings.mlr.press/v15/glorot11a.html.

[96] Y. Kwak, W.-J. Song, and S.-E. Kim, “Speckle-noise-invariant convolutional neural network for SAR target recognition,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 4, pp. 549–553, Apr. 2019.

[97] R. Min, H. Lan, Z. Cao, and Z. Cui, “A gradually distilled CNN for SAR target recognition,” IEEE Access, vol. 7, pp. 42190–42200, 2019.

[98] F. Zhang, Y. Liu, Y. Zhou, Q. Yin, and H. C. Li, “A lossless lightweight CNN design for SAR target recognition,” Remote Sens. Lett., vol. 11, no. 5, pp. 485–494, 2020.

[99] J. Guo, B. Lei, C. Ding, and Y. Zhang, “Synthetic aperture radar image synthesis by using generative adversarial nets,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 7, pp. 1111–1115, Jul. 2017.

[100] B. Lewis, A. Wong, and J. Liu, “Generative adversarial networks for SAR image realism,” Proc. SPIE, vol. 10647, Apr. 2018, Art. no. 106470Q.

[101] B. Lewis, O. DeGuchy, J. Sebastian, and J. Kaminski, “Realistic SAR data augmentation using machine learning techniques,” Proc. SPIE, vol. 10987, May 2019, Art. no. 109870C.

[102] S. A. Wagner, “SAR ATR by a combination of convolutional neural network and support vector machines,” IEEE Trans. Aerosp. Electron. Syst., vol. 52, no. 6, pp. 2861–2872, Dec. 2016.

[103] J. Ding, B. Chen, H. Liu, and M. Huang, “Convolutional neural network with data augmentation for SAR target recognition,” IEEE Geosci. Remote Sens. Lett., vol. 13, no. 3, pp. 364–368, Mar. 2016.

[104] Y. Yan, “Convolutional neural networks based on augmented training samples for synthetic aperture radar target recognition,” J. Electron. Imag., vol. 27, no. 02, 2018, Art. no. 023024.

[105] Z. Cui, M. Zhang, Z. Cao, and C. Cao, “Image data augmentation for SAR sensor via generative adversarial nets,” IEEE Access, vol. 7, pp. 42255–42268, 2019.

[106] M. Arjovsky, S. Chintala, and L. Bottou, “GAN Wasserstein,” arXiv:1701.07875v3, Jan. 2017, [Online]. Available: http://arxiv.org/abs/1701.07875.

[107] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, “Improved training of Wasserstein GANs,” Mar. 2017. [Online]. Available: http://arxiv.org/abs/1704.00028.

[108] Y. Sun, Y. Wang, H. Liu, N. Wang, and J. Wang, “SAR target recognition with limited training data based on angular rotation generative network,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 11, pp. 1928–1932, Nov. 2020.

[109] X. Shi, F. Zhou, S. Yang, Z. Zhang, and T. Su, “Automatic target recognition for synthetic aperture radar images based on super-resolution generative adversarial network and deep convolutional neural network,” Remote Sens., vol. 11, no. 2, 2019, Art. no. 135.

[110] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” Inf. Softw. Technol., vol. 51, no. 4, pp. 769–784, Sep. 2014.

[111] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 770–778.

[112] J. Deng, W. Dong, R. Socher, L.-I. Li, K. Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 248–255.

[113] S. J. Pan and Q. Yang, “A survey on transfer learning,” IEEE Trans. Knowl. Data Eng., vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[114] O. Kechagias-Stamatis, “Target recognition for synthetic aperture radar imagery based on convolutional neural network feature fusion,” J. Appl. Remote Sens., vol. 13, no. 01, Dec. 2018, Art. no. 019902.

[115] O. Kechagias-Stamatis, N. Aouf, and C. Belloni, “SAR automatic target recognition based on convolutional neural networks,” in Proc. Int. Conf. Radar Syst., 2017, pp. 1–4.

[116] E. Mayoraz and E. Alpaydin, “Support vector machines for multi-class classification,” in Proc. Eng. Appl. Bio-Inspired Artif. Neural Netw., 1999, pp. 833–842.
Automatic Target Recognition on Synthetic Aperture Radar Imagery: A Survey

[117] M. Amrani and F. Jiang, “Deep feature extraction and combination for synthetic aperture radar target classification,” J. Appl. Remote Sens., vol. 11, no. 04, p. 1, 2017, Art. no. 042616.

[118] C. Zhong, X. Mu, X. He, J. Wang, and M. Zhu, “SAR target image classification based on transfer learning and model compression,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 3, pp. 412–416, Mar. 2019.

[119] Z. Tian, L. Wang, R. Zhan, J. Hu, and J. Zhang, “Classification via weighted kernel CNN: Application to SAR target recognition,” Int. J. Remote Sens., vol. 39, no. 23, pp. 9249–9268, Dec. 2018.

[120] X. Huang, Q. Yang, and H. Qiao, “Lightweight two-stream convolutional neural network for SAR target recognition,” IEEE Geosci. Remote Sens. Lett., to be published, doi: 10.1109/LGRS.2020.2983718.

[121] R. Shang, J. Wang, L. Jiao, R. Stolkin, B. Hou, and Y. Li, “SAR targets classification based on deep memory convolutional neural networks and transfer parameters,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 11, no. 8, pp. 2834–2846, Aug. 2018.

[122] J. Pei, Y. Huang, W. Huo, Y. Zhang, J. Yang, and T. S. Yeo, “SAR automatic target recognition based on multiview deep learning framework,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 4, pp. 2196–2210, Apr. 2018.

[123] J. C. Ni and Y. L. Xu, “SAR automatic target recognition based on a visual cortical system,” in Proc. 6th Int. Congr. Image Signal Process., Dec. 2013, pp. 778–782.

[124] Z. Sun, L. Xue, and Y. Xu, “Recognition of SAR target based on multilayer auto-encoder and SNN,” Int. J. Innov. Comput. Inf. Control, vol. 9, no. 11, pp. 4331–4341, 2013.

[125] X. Huang, X. Nie, W. Wu, H. Qiao, and B. Zhang, “SAR target configuration recognition based on the biologically inspired model,” Neurocomputing, vol. 234, pp. 185–191, Apr. 2017.

[126] Z. Huang, Z. Pan, and B. Lei, “Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data,” Remote Sens., vol. 9, no. 9, pp. 1–21, 2017.

[127] F. Zhang, Z. Fu, Y. Zhou, W. Hu, and W. Hong, “Multi-aspect SAR target recognition based on space-fixed and space-varying scattering feature joint learning,” Remote Sens. Lett., vol. 10, no. 10, pp. 998–1007, Oct. 2019.

[128] F. Gers, “Long short-term memory in recurrent neural networks,” EPFL, 2001.

[129] F. Zhang, C. Hu, Q. Yin, W. Li, H. C. Li, and W. Hong, “Multi-aspect-aware bidirectional LSTM networks for synthetic aperture radar target recognition,” IEEE Access, vol. 5, pp. 26880–26891, 2017.

[130] J. Lv and Y. Liu, “Data augmentation based on attributed scattering centers to train robust CNN for SAR ATR,” IEEE Access, vol. 7, pp. 25459–25473, 2019.

[131] Y. Cong, B. Chen, H. Liu, and B. Jiu, “Nonparametric Bayesian attributed scattering center extraction for synthetic aperture radar targets,” IEEE Trans. Signal Process., vol. 64, no. 18, pp. 4723–4736, Sep. 2016.

[132] A. Karine, A. Toumi, A. Khenchaf, and El M. Hassouni, “Radar target recognition using salient keypoint descriptors and multitask sparse representation,” Remote Sens., vol. 10, no. 6, 2018, Art. no. 843.

[133] D. G. Lowe, “Distinctive image features from scale invariant keypoints,” Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.

[134] Y. Sun, L. Du, Y. Y. Wang, Y. Y. Wang, and J. Hu, “SAR automatic target recognition based on dictionary learning and joint dynamic sparse representation,” IEEE Geosci. Remote Sens. Lett., vol. 13, no. 12, pp. 1777–1781, Dec. 2016.

[135] H. Zhang, N. M. Nasrabadi, Y. Zhang, and T. S. Huang, “Multi-observation visual recognition via joint dynamic sparse representation,” in Proc. Int. Conf. Comput. Vis., Nov. 2011, pp. 595–602.

[136] S. Gishkori and B. Mulgrew, “Pseudo-Zernike moments based sparse representations for SAR image classification,” IEEE Trans. Aerosp. Electron. Syst., vol. 55, no. 2, pp. 1037–1044, Apr. 2019.

[137] B. Ding, G. Wen, X. Huang, C. Ma, and X. Yang, X., “Target recognition in SAR images by exploiting the azimuth sensitivity,” Remote Sens. Lett., vol. 8, no. 9, pp. 821–830, 2017.

[138] X. Z. Zhang, Z. Liu, S. Liu, D. Li, Y. Jia, and P. Huang, “Sparse coding of 2D-slice Zernike moments for SAR ATR,” Int. J. Remote Sens., vol. 38, no. 2, pp. 412–431, 2017.

[139] O. Kechagias-Stamatis and N. Aouf, “Fusing deep learning and sparse coding for SAR ATR,” IEEE Trans. Aerosp. Electron. Syst., vol. 55, no. 2, pp. 785–797, Apr. 2019.

[140] G. McLachlan and D. Peel, Finite Mixture Models. New York, NY, USA: Wiley, 2004.

[141] B. Ding, G. Wen, C. Ma, and X. Yang, “An efficient and robust framework for SAR target recognition by hierarchically fusing global and local features,” IEEE Trans. Image Process., vol. 27, no. 12, pp. 5983–5995, 2018.

[142] B. Ravichandran, A. Gandhe, R. Smith, and R. Mehra, “Robust automatic target recognition using learning classifier systems,” Inf. Fusion, vol. 8, no. 3, pp. 252–265, Jul. 2007.

[143] G. Dong, H. Liu, G. Kuang, and J. Chanussot, “Target recognition in SAR images via sparse representation in the frequency domain,” Pattern Recognit., vol. 96, Dec. 2019, Art. no. 106972.

[144] R. H. Levene and S. C. Power, “Manifolds of Hilbert space projections,” Proc. London Math. Soc., vol. 100, no. 2, pp. 485–509, 2010.

[145] G. Dong, G. Kuang, G. Dong, and G. Kuang, “Classification on the monogenic scale space: Application to target recognition in SAR image,” IEEE Trans. Image Process., vol. 24, no. 8, pp. 2527–2539, Aug. 2015.
[146] G. Dong, N. Wang, G. Kuang, and H. Qiu, “Sparsity and low-rank dictionary learning for sparse representation of monogenic signal,” IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens., vol. 11, no. 1, pp. 141–153, Jan. 2018.

[147] Z. Zhou, M. Wang, Z. Cao, and Y. Pi, “SAR image recognition with monogenic scale selection-based weighted multi-task joint sparse representation,” Remote Sens., vol. 10, no. 4, 2018, Art. no. 504.

[148] G. Dong, G. Kuang, N. Wang, and W. Wang, “Classification via sparse representation of steerable wavelet frames on Grassmann manifold: Application to target recognition in SAR image,” IEEE Trans. Image Process., vol. 26, no. 6, pp. 2892–2904, 2017.

[149] S. Liu, R. Zhan, Q. Zhai, W. Wang, and J. Zhang, “Multi-view radar target recognition based on multitask compressive sensing,” J. Electromagn. Waves Appl., vol. 29, no. 14, pp. 1917–1934, 2015.

[150] L. Jin, J. Chen, and X. Peng, “Joint classification of complementary features based on multitask compressive sensing with application to synthetic aperture radar automatic target recognition,” J. Electron. Imag., vol. 27, no. 05, 2018, Art. no. 053034.

[151] A. K. Mishra, “Validation of PCA and LDA for SAR ATR,” in Proc. TENCON IEEE Region 10 Conf., Nov. 2008, pp. 1–6.

[152] G. C. Anagnostopoulos, “SVM-based target recognition from synthetic aperture radar images using target region outline descriptors,” Nonlinear Anal. Theory, Methods Appl., vol. 71, no. 12, pp. 2934–2939, Dec. 2009.

[153] S. Liu and J. Yang, “Target recognition in synthetic aperture radar images via joint multifeature decision fusion,” J. Appl. Remote Sens., vol. 12, no. 01, 2018, Art. no. 016012.

[154] A. K. Mishra and T. Motaung, “Application of linear and nonlinear PCA to SAR ATR,” in Proc. 25th Int. Conf. Radioelektronika, Apr. 2015, pp. 349–354.

[155] A. Geiger, P. Lenz, and R. Urtasun, “The KITTI vision benchmark suite,” Karlsruhe Inst. Technol., Karlsruhe, Germany, 2012.

[156] B. Lewis, T. Scarnati, E. Sudkamp, J. Nehrbass, S. Rosencrantz, and E. Zelnio, “A SAR dataset for ATR development: The synthetic and measured paired labeled experiment (SAMPLE),” Proc. SPIE, vol. 10987, 2019, Art. no. 109870H.