DOI. No. 10.1109/MAES.2021.3117369

Tutorial:

A Survey on the Applications of Convolutional Neural Networks for Synthetic Aperture Radar: Recent Advances

Amir Hosein Oveis, *Elisa Giusti*, *and Selenia Ghio*, **National Inter-University Consortium for Telecommunications, 56124 Pisa, Italy**
Marco Martorella, **National Inter-University Consortium for Telecommunications, 56124 Pisa, Italy, and also University of Pisa, 56124 Pisa, Italy**

In recent years, convolutional neural networks (CNNs) have drawn considerable attention for the analysis of synthetic aperture radar (SAR) data. In this study, major subareas of SAR data analysis that have been tackled by CNNs are systematically reviewed, such as automatic target recognition, land use and land cover classification, segmentation, change detection, object detection, and image denoising. Special emphasis has been given to practical techniques such as data augmentation and transfer learning. Complex-valued CNNs, which have been introduced to exploit phase information embedded in SAR complex images, have also been extensively reviewed. To conclude this review paper, open challenges and future research directions are highlighted.

INTRODUCTION

In our data-intensive world, machine learning (ML) techniques are becoming increasingly important. Features are extracted to reduce the dimension of input data to a more manageable scale by holding only informative and nonredundant parts. Feature extraction and feature classification are two steps of most ML problems. In the conventional ML algorithms, handcrafted features are extracted. However, the emphasis is currently on the automatic learning of features. Deep learning (DL) is undoubtedly the most popular trend in the ML domain. During the past decade, it has achieved impressive performance in the field of computer vision and image processing, with applications such as image classification, object detection, super-resolution restoration, etc. [1]. Based on the availability of ground truth or labels for the training set, learning procedures can be accomplished in supervised, unsupervised, or semi-supervised manners. In addition, reinforcement learning is another type of learning that is often discussed under the scope of semi-supervised and sometimes under unsupervised learning approaches [2]. Although supervised DL brings outstanding results, in case of insufficient labeled training samples [for example, in the remote sensing (RS) domain], other DL approaches can be employed to extract features from unlabeled images. Convolutional neural network (CNN) [3], known also as ConvNet, is one of the most popular DL approaches in computer vision and RS fields. A CNN, inspired by the human's visual perception mechanism, consists of different convolutional and pooling layers. Typically, features are extracted at first, and then they are fed to a multilayer perceptron (MLP) for classification or regression purposes. Other types of deep neural networks such as autoencoders (AEs) [4], deep belief networks (DBNs) [5], [6], and recurrent neural networks (RNN) [7] will be briefly introduced later in the section "Deep Neural Networks in Computer Vision and Remote Sensing."

In this article, applications of CNNs to SAR will be thoroughly analyzed. Moreover, some other DL approaches that incorporate convolution operation will be mentioned. We

Authors' current address: Amir Hosein Oveis, Elisa Giusti, Selenia Ghio, and Marco Martorella, Radar and Surveillance Systems, National Inter-University Consortium for Telecommunications, 56127, Pisa, Italy (e-mail: aoveis@cnit.it, elisa.giusti@cnit.it, selenia.ghio@cnit.it); Marco Martorella, Department of Information Engineering, University of Pisa, 56127, Pisa, Italy (e-mail: marco.martorella@iet.unipi.it).
Manuscript received September 30, 2020, revised September 15, 2021; accepted September 15, 2021, and ready for publication September 30, 2021.
Review handled by Dale Blair.
0885-8985/21/$26.00

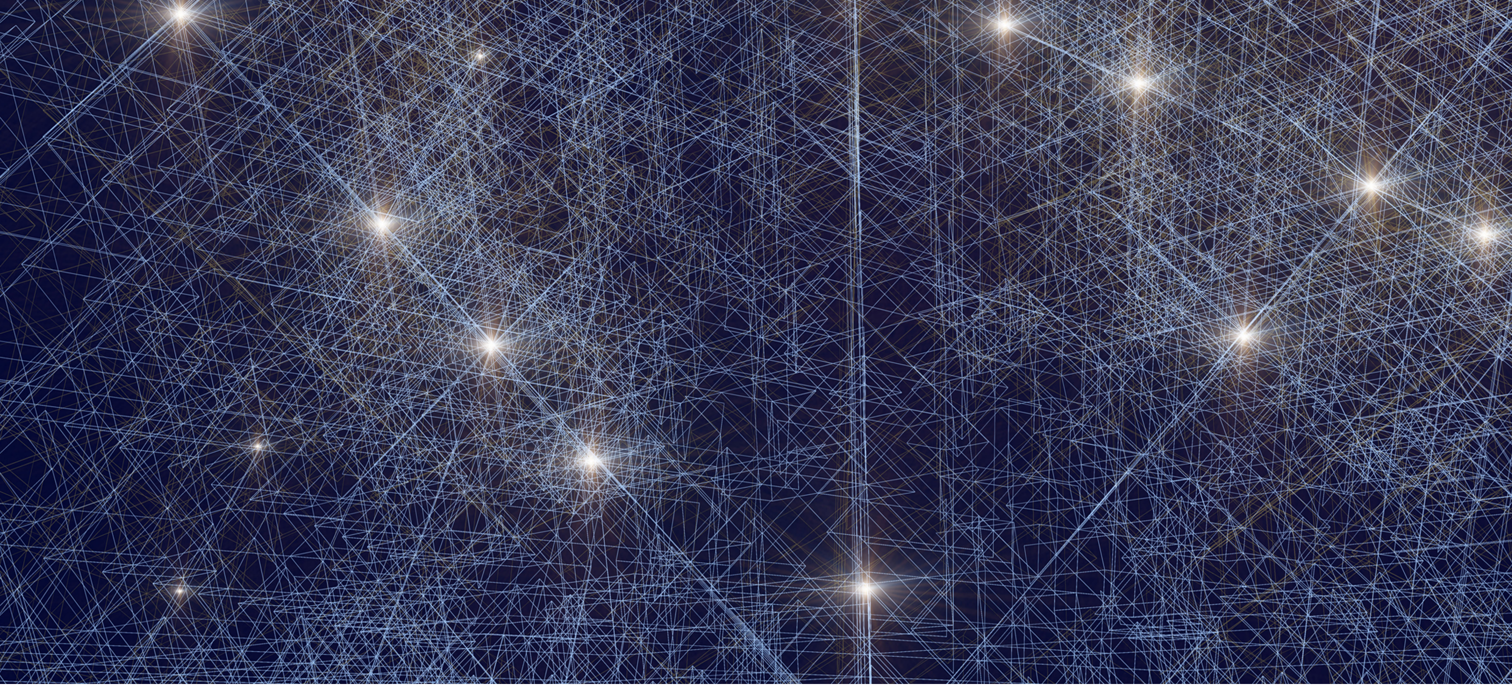


Image licensed by Ingram Publishing

summarized the work done by many researchers in this field including high-cited peer-reviewed articles in addition to some influential conference papers. The overview of this paper is presented in Figure 1 where in its left side the taxonomy of CNN applications to SAR is shown and in its right side the related topics to both CNN and SAR in order to be discussed thoroughly are shown. Moreover, corresponding key words to each item are also given for the ease of the reader and they will be referred to throughout this article.

SAR is an active imaging sensor installed on a moving platform such as an aircraft or a satellite and it offers high-resolution images in all-day and all-weather conditions. Despite the outstanding capabilities of SAR imagery, applying CNNs to SAR images is more challenging than optical images. Because of the unique features of SAR systems, there is a large discrepancy between SAR and optical images. In other words, radar illumination direction fundamentally changes how a target looks due to the self-shadowing phenomenon, noise is unique to this modality (multiplicative speckle), dynamic range is larger, etc. In addition, a large-scale SAR annotated dataset for training is scarce because of the high cost of SAR image acquisition. As mentioned in Figure 1, Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset [8], as the subject for a large body of published work in automatic target recognition (ATR) application [9] [10], is one of the few publicly available SAR labeled dataset.

The rest of this article is organized as follows. In the section "Deep Neural Networks in Computer Vision and Remote Sensing" we introduce different DL approaches, and then we concentrate on CNNs by addressing their role in both computer vision and RS fields and presenting a brief introduction about their typical structures. In "Applications of CNNs in SAR Data Analysis," subareas of SAR data analysis that are tackled by CNNs are reviewed. In the section "Related Topics to CNN and SAR," both CNN and SAR are thoroughly discussed. Last, we present "Conclusions." In addition, challenges and future research directions are highlighted.

DEEP NEURAL NETWORKS IN COMPUTER VISION AND RS

We start this section with a short introduction of different DL approaches in order to highlight their differences and then we concentrate on CNNs. A few applications to SAR data for each approach will also be mentioned.

An AE is a symmetrical neural network used for unsupervised learning from unlabeled data. In an undercomplete AE, the encoder aims to compress the data into a reduced dimension and the decoder aims to reconstruct it back. A stacked autoencoder (SAE) is a deep network model that makes use of multiple AEs, where the output of one AE becomes the input of the subsequent one. Convolutional AEs [11] have also been developed to provide AEs with spatial characteristics of convolutions. AEs have been widely used in SAR domain for polarimetric image classification [12]–[16], SAR change detection [17], [18], and so forth.

A DBN is formed by stacking two or more individual unsupervised networks such as restricted Boltzmann machines (RBMs) blocks [19] in such a way that each RBM layer can be connected to both previous layers as well as subsequent ones. In other words, a single RBM, which is an undirected model with a visible layer and a single hidden layer, has a limited power to extract features. Therefore, two or more RBMs are typically stacked together to form DBNs. It is noteworthy to mention that DBNs have been applied to SAR image classification [20]–[23].

An RNN is typically employed for the analysis of nonstationary signals. In an RNN, the input and output data can have different lengths. In the context of SAR, it has been used for image classification [24], as an inverse solver for image formation problems in passive SAR [25], for agricultural land cover mapping in satellite radar images [26] and similarly for winter vegetation quality mapping [27].

CNNs have made obsolete the paradigm of handcrafted feature extraction in the conventional computer

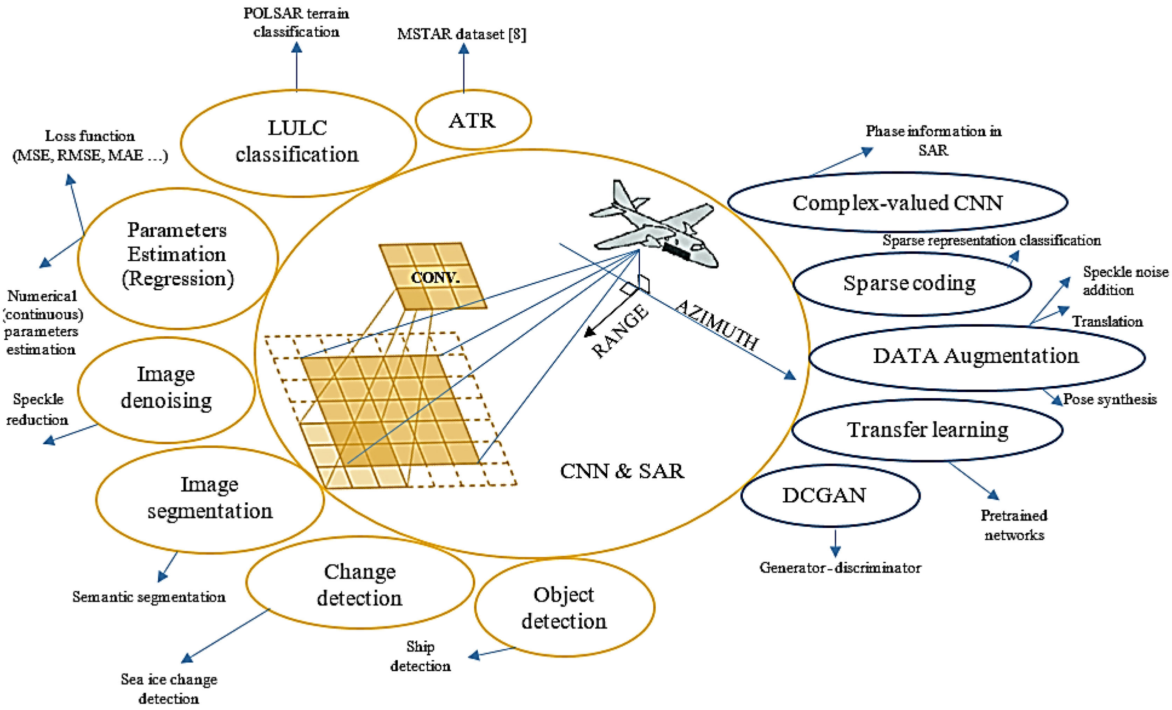


Figure 1.
Overview of the survey paper. CNN applications to SAR (left side) and related topics to CNN and SAR (right side) together with corresponding keywords to each item.

vision field. The main advantages of CNNs in comparison with their predecessors, i.e., classical regression and classification methods, are automatic feature extraction using hierarchical representations, weight sharing using convolution layers, and spatial invariance using pooling operations. CNNs have been proposed in the 1990s [3], however, due to limited computational and processing capacity and also the scarcity of annotated training data, they fell out of fashion. Thanks to the availability of large-scale dataset of optical images, CNNs have rapidly developed in recent years. In 2012, Krizhevsky *et al.* [28] succeeded to attract the attention to CNNs again by proposing AlexNet, as it won the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC-2012). AlexNet has five convolutional layers and three fully connected layers. Motivated by this, many other powerful CNN architectures, such as VGG-16 [29], GoogLeNet [30], and Res-Net [31] have been proposed to gradually decrease the top five error rates in the ImageNet challenge from 7.3%, 6.67% to 3.6%, respectively.

CNN MEETS SAR

CNNs have been extensively used to solve many SAR problems such as ATR, scene classification, noise removal, image formation, change and anomaly detection, land use and land cover (LULC) classification, and so forth. Moreover, CNNs showed their applications to different SAR techniques such as inverse SAR (ISAR), polarimetric SAR (PolSAR), and interferometric SAR (InSAR). Apart from the details of CNN architectures, such as the number of hidden layers and the kernel size of convolution layers, there are some noticeable differences between different state-of-the-art methods in SAR domain. As a prominent example, some studies exploited complex-valued architectures. Generally, both the input data (including training, validation, and test sets) and the network coefficients can be real or complex values. Hereafter, we use the terms RV and CV for real-valued and complex-valued CNN architectures. Complex input data can be considered as two independent channels of real and imaginary parts or magnitude and phase. Since usual computer vision applications seek for classification of real world images, like the well-known cats and dogs classification example, CV-CNNs are uncommon. However, they were introduced recently to exploit the phase information embedded for example in SAR images. In radar imaging scenarios, typically, only the amplitude information is used. However, phase information can be employed for a number of applications. We will thoroughly analyze the applications of complex-values CNNs to SAR systems in the section "Related Topics to CNN and SAR."

CNNs can be used for both regression and classification tasks. These two tasks are subsets of supervised ML. The main goal of a regression task is to predict continuous (numerical) variables, whereas classification is used for the prediction of discrete (categorical) variables. The majority of previous SAR studies that deal with CNNs are classification tasks and only a few studies have extended and applied CNNs for regression-type problems. Note that classification and regression problems with CNN do not necessarily have different networks. In fact, learning classification problems are also solved through regression and then the result is discretized. Here the main difference is the choice of the loss functions. In regression problems, mean squared error (MSE), root MSE (L2), mean absolute error (L1), and Huber loss are typically used, whereas, in classification problems, cross-entropy loss and Hinge loss are typically employed.

There are some survey papers on the applications of DL in RS such as [1], [7], and [32]–[37]. Main applications of DL to RS have been reviewed in the above-mentioned survey papers. However, we believe that the application of CNNs to SAR needs more attention. In fact, both RS and DL are general terms. In other words, RS include optical (multi and hyperspectral), LIDAR, and synthetic aperture radar (SAR) sensors, where the imaging geometries and content are completely different [37]. Similarly, there are several different types of DL architectures that have been utilized in RS where CNN is one of the most popular ones. Therefore, in this survey paper, we focus on the applications of CNNs to SAR and analyze recent advances in more detail. For instance, the concept of using complex-valued architectures has been mentioned in none of the above survey papers except for [37]. In addition, example CNN structures, their input and output shapes, and their performance in SAR applications will be covered.

BRIEF REVIEW ON THE STRUCTURE OF CNN

CNNs are typically formed from multiple basic feature extraction stages. Convolutional layer, activation function, and pooling layer are the main elements of each stage. These stages are followed by one or more fully connected layers and a classification module. A simple example of CNN with single-channel input is illustrated in Figure 2. The role of each element will be briefly introduced in the following parts.

CONVOLUTIONAL LAYER

A convolutional layer is the core part in each stage of CNN. It is a filter bank, say M filters, whose coefficients are to be learned by training. Let us assume that the input of the convolutional layer is a tensor ($W1 \times H1 \times C1$) where W1, H1, and C1 denote the width, the height, and the number of channels of the input, respectively. The first filter ($K \times K \times C1$), which is a small square kernel with side length of K pixels and the same number of channels (C1) as the input tensor, is slid over the width and the height of the input tensor. Using element-wise multiplication and summation, a new tensor ($W2 \times H2 \times C1$) is generated where W2 and H2 represent the width and the height of the new tensor, respectively. Next, different channels are added together to form the first feature map ($W2 \times H2 \times 1$). Employing M filters, M feature maps ($W2 \times H2 \times M$) are generated. The number of parameters to be trained in convolutional layers is far less than in traditional fully connected layers. The number of filters (M), dimension of each filter (K), stride, and zero-padding on the border are four hyper-parameters here and they are not changed during the training phase. Note that there are other types of convolutions such as three-dimensional (3-D), depthwise separable [38], spatial separable, dilated [39], etc., that have been defined in the DL literature. However, for the sake of conciseness, we briefly introduced the standard 2-D convolution, which is typically used in CNNs.

ACTIVATION FUNCTION

The activation function is a simple mathematical function that is applied to the output of a convolutional layer in an element-wise manner to introduce nonlinearity to the network. This nonlinearity allows the network to go deep, otherwise, no matter how complicated the architecture is, it would be equal to a single layer depth. Rectified Linear Units (ReLU), $y = \max(0,x)$ is often used in CNNs as it shows to be computationally efficient, to accelerate the

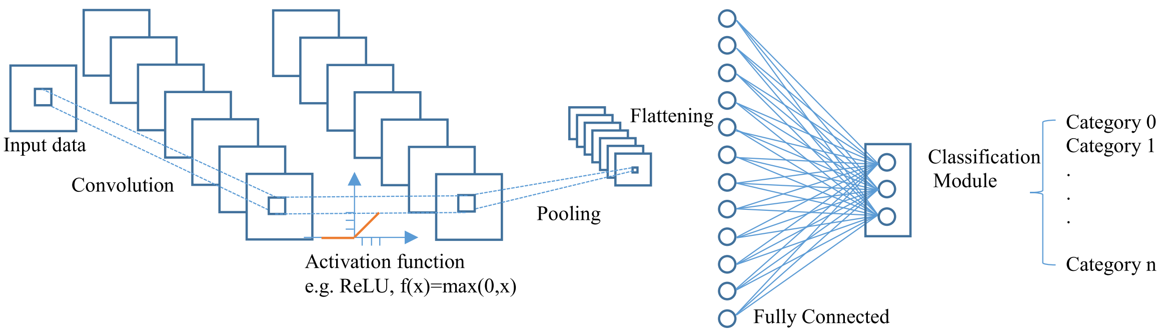


Figure 2.
Convolutional neural network.

convergence and to avoid the vanishing gradient problem. However, it also has some shortcomings. Since the derivative of the ReLU function is always zero when the input value is negative, it is likely that the phenomenon of neuronal necrosis happens [40]. Neuronal necrosis or dying ReLU problem refers to the condition in which ReLU gives zero for any input and neurons become dead or inactive. In order to solve this problem, Leaky ReLU was first introduced. It multiplies the negative input by a small value. Afterward, the Parametric ReLU (PReLU), was introduced where the slope in the negative region is learnt during the training phase.

POOLING LAYER

A pooling layer, typically used after convolutional and nonlinear layers, decreases the dimension of each feature map. Pooling is performed by sliding a window (typically square window) over each feature map, taking only one value out of each window. Note that the down sampling is only applied on the width and the height dimensions, therefore, the number of channels will not be changed after pooling. Max-pooling and average-pooling are two most common pooling mechanisms used in CNNs. Moreover, pooling helps to make the representation become approximately invariant to small shifts in the input [41].

FULLY CONNECTED LAYER

The output of the final pooling/convolutional layer is flattened to a vector and fed to a fully connected layer. A fully connected layer is to learn nonlinear combinations of previous high-level features. All the input elements of this layer are connected to its output elements similarly to a traditional MLP neural network. One or more fully connected layers can be used successively.

CLASSIFICATION MODULE

The classification layer, which is the last module in CNN, is used to map the output of the last fully connected layer to the probability domain. Softmax classifier with cross-entropy loss is widely used for multiclass classification. For the regression-type CNN applications, the classification module is removed and the output of the last fully connected layer is compared with the ground truth using the MSE loss function or other loss functions that are suitable for regression purposes.

APPLICATIONS OF CNN IN SAR DATA ANALYSIS

In this section, we categorize almost all the SAR tasks that were tackled by CNNs in recent studies. A few other DL techniques that incorporate convolution operation are also mentioned. Note that the different parts of this section are not semantically independent. For instance, we formed a separate category for image denoising to discuss it in detail. The emphasis of some papers is denoising, however, some other papers address the denoising as a preprocessing step for target recognition and other purposes. Similarly, there are some parts such as data augmentation and transfer learning that are related to CNNs and will be addressed in "Related Topics to CNN and SAR."

AUTOMATIC TARGET RECOGNITION

In the past few years, many SAR-ATR methods have been proposed that made use of CNNs. ATR based on SAR images plays a major role particularly in military applications such as friend and foe identification, battlefield surveillance, and disaster relief [42]. Ding *et al.* [42] utilized a simple CNN architecture based on three convolutional layers to implement a SAR-based ATR system. Since their emphasis is on the data augmentation, we will cover it in "Related Topics to CNN and SAR." Chen *et al.* [43] proposed a CNN architecture for ATR based on SAR images. Since the number of training SAR images for each target is limited, CNNs will suffer from severe overfitting. To address this problem, the authors of [43] proposed an all-convolutional network (A-ConvNet) with fewer degrees of freedom by removing the fully connected layers. These two papers, i.e., [42] and [43], are two of the most cited ones in the SAR-ATR context.

Some studies have concentrated on the classifier of their proposed CNNs for SAR-ATR. Wagner [44] proposed a combination of CNN and support vector machine (SVM), instead of a conventional softmax classifier, for the task of SAR-ATR. Since SVM can discriminate between two classes, they had to design ten SVMs for their problem. Inspired by [44], Gao *et al.* [45] have also addressed the combination of CNN and SVM for the SAR-ATR task with a different loss function. Zhang *et al.* [46] used an ensemble learning based classifier (linear combination of many weak classifiers) to replace the softmax layer in their CNN. Zhou *et al.* [47] utilized the large-margin (LM) softmax classifier [48] in the last layer to increase the divisibility of different target types for the SAR-ATR task. The LM-softmax classifier was introduced to increase the restriction of the traditional softmax classifier and has been also applied to the SAR-ATR task in [49]. It has also been noted by some studies that training using the LM-softmax can be unstable [50]–[53].

Some studies have focused on the compression and efficiency of CNNs for SAR-ATR tasks. A possible problem of lightweight CNNs is that in case of a new dataset they may lose their accuracies (in comparison with original CNNs) due to underfitting and decreased model capacity. Shao *et al.* [54] exploited depthwise convolution, which is a kind of model compression technique, to reduce the CNN parameters and accelerate the calculation for the

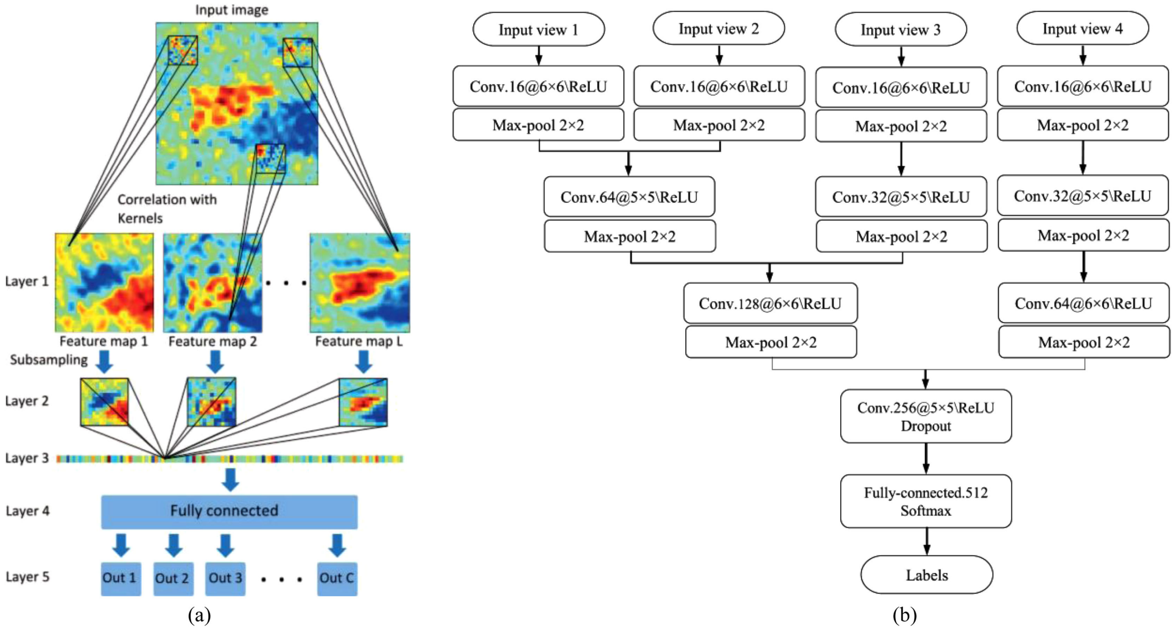


Figure 3.
Proposed SAR-ATR architectures. (a) Combining CNN and SVM [44]. (b) Multiview [59].

SAR-ATR applications. Chen *et al.* [55] used weight quantization and Huffman coding to compress a CNN for resource-constrained SAR-ATR tasks. Min *et al.* [56] proposed a micro CNN, which has only two layers with weights defined as "−1," "1," and "0," for decreasing the inference time of SAR-ATR tasks. Comparing with [55], the authors of [56] showed that their proposed method achieves a low error rate with a smaller model size and fewer parameters to process. The proposed technique in [56] is based on knowledge distillation [57], also known as "teacher-student paradigm," for training the micro-CNN under the supervision of a deeper network. Zhang *et al.* [58] have also studied knowledge distillation for real-time SAR-ATR tasks.

Some studies have discussed the applicability of CNNs for multiview SAR-ATR, which benefits from diverse recognition information by acquiring SAR images from different viewing angles. Pei *et al.* [59], [60] suggested a multiple-input CNN architecture, termed "multiview," to learn the features of targets at different azimuth angles. Their proposed architecture together with the one proposed by [44] are depicted in Figure 3. Similarly, Zhao *et al.* [61] proposed a framework named as multistream CNN for SAR-ATR tasks. Their suggested multi-input architecture is fed by multiple views of the same target. Afterward, a Fourier feature fusion layer has been added to fuse the features derived from multiple views.

In some SAR-ATR studies, feature fusion and aggregation have been addressed. Wang *et al.* [62] proposed a two-channel CNN to extract and fuse features from the intensity and edge information of SAR images for SAR-ATR tasks. Cho *et al.* [63] proposed a framework based on an aggregation of three feature extractor subnetworks for SAR-ATR application. Their work relies on the parallel use of max-pooling and average-pooling operations. Max-pooling extracts strong features from images. On the contrary, average-pooling is not often used because it makes features smooth. However, it can reduce the effect of noise in SAR images [63]. Bai *et al.* [64], proposed a CNN-based architecture for ATR of multipolarized ISAR images based on the fusion of the features from HH, HV, and VV-polarized ISAR images. This was done by adopting a double-layer spatial transformer network (STN) [65] module to adjust the image deformation of each polarimetric channel. STNs constitute a widely used end-to-end trainable solution for CNNs to learn invariance to image transformations [66].

Insufficient SAR labeled data has been mentioned in the major part of the papers in SAR-ATR. Apart from data augmentation methods**,** which will be covered in "Related Topics to CNN and SAR", there are some other alternatives to obviate this problem. Lin *et al.* [67] proposed a convolutional highway network to train deeper networks with limited SAR data. Highway networks [68], which are based on CNNs, can be trained with smaller datasets by exploiting skip connections, similar to Res-Net, for training deep models. Yue *et al.* [69] proposed a semi-supervised CNN method for SAR-ATR in which the information contained in the unlabeled samples was integrated into the loss function by means of a linear discriminant analysis method. In fact, they utilized a CNN to obtain the class probabilities of the unlabeled samples.

As previously mentioned, MSTAR benchmark dataset [8] is one of the few publicly available labeled dataset for

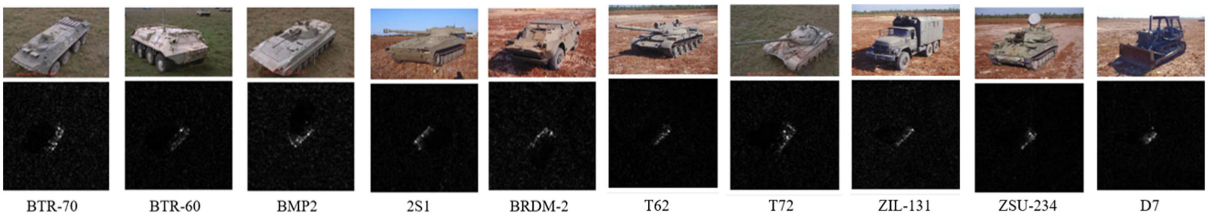


Figure 4.
MSTAR dataset including optical and SAR images of ten military vehicles.

SAR-ATR. It was collected and released by the U.S. Air Force Research Laboratory and Defense Advanced Research Projects Agency. This dataset has been widely used by so many studies for training and testing in SAR-ATR applications. MSTAR dataset, which contains ten classes of ground military vehicles as shown in Figure 4, has been collected by an X-band SAR sensor with 0.3 m $\times$ 0.3 m resolution and full 360° aspect angles coverage. Most of the target chips have the size of 128 $\times$ 128 pixels. MSTAR has been used in two different configuration sets: standard operation conditions (SOCs) and the extended operation conditions (EOC). In SOC, typically, the depression angle of 15° is used for training and the depression angle of 17° is used for test. Whereas, in EOC, there exists more dissimilarity between training and test sets such as larger depression angle variations and targets' intraclass variants. Furthermore, OpenSARShip [70] and SAR Ship Detection Dataset (SSDD) [71] are two other datasets typically used for marine surveillance and ship detection applications [72]. These two datasets will be discussed in "Applications of CNNs in SAR Data Analysis."

LAND USE AND LAND COVER CLASSIFICATION

LULC classification is a fundamental application in RS. SAR systems are generally employed for LULC classification as they produce unique features that cannot be obtained with electro-optical (EO) systems [73]. Land use refers to the purpose of the land (such as agricultural, residential, etc.), whereas land cover refers to the physical land type (such as water, ice, etc.). There have been so many studies for investigating the applicability of CNNs for LULC classification, such as buildings [74], [75], urban flood [76], [77] , and so forth.

Considering the input of proposed architectures, Chen and Tao [78] have addressed the problem of land cover classification using PolSAR data by exploiting null angle features derived from the matrix rotation [79] as the input of their CNN. Zhou *et al*. [80] proposed a CNN with six input channels for terrain classification (such as water, building, grass, etc.) using polarimetric SAR images. Although monostatic multilook complex POLSAR data can be represented by the Pauli-based polarimetric coherency matrix T [81] that is a 3 $\times$ 3 complex matrix, they proposed a new 6-D RV vector representation to feed their CNN. Liu *et al.* [82] have also proposed a polarimetric scattering coding matrix to feed a classifier based on a fully convolutional network (FCN) for PolSAR land cover classification. Wang *et al.* [83] proposed a fixed-feature size CNN, which makes some reference to Lenet-5 [3], to classify all pixels in a patch simultaneously using PolSAR data for land cover classification. The polarimetric data of all the pixels in a patch are used to generate a matrix as the input to their CNN.

Considering some modifications to the structure of the conventional deep CNNs, Ahishali *et al.* [84] proposed a compact and adaptive CNN for LULC classification of single-polarized COSMO-SkyMed and dual-polarized TerraSAR-X intensity data. Bi *et al.* [85], [86] integrated a graph model and CNN for PolSAR land cover classification. Xie *et al.* [87] employed convolutional AE with Wishart classifier, which uses the Wishart distance to measure similarity, for land cover classification in PolSAR images. Ren *et al.* [88] proposed an architecture for SAR land cover classification by integrating CNN and sparse AE for unsupervised feature learning.

The Flevoland[1] image has been widely used for many studies related to LULC PolSAR image classification. Acquired by the AIRSAR platform in 1989, this image consists of fifteen classes. The result of [80] together with the Flevoland image are shown in Figure 5.

CNN-BASED REGRESSION FOR PARAMETERS ESTIMATION

As previously mentioned, regression techniques are employed to predict continuous and numerical parameters. In CNN-based regression applications, the classification layer is removed and the output of the last fully connected layer is compared with the ground truth using loss functions that are suitable for regression purposes. Wang *et al.,* in [89] and [90], used CNN for the regression problem of ice concentration estimation using SAR images during melt and freeze-up, respectively. In both regression

[1]The dataset including the toolbox can be downloaded from: https://earth.esa.int/web/polsarpro/datasources/sample-datasets. https://earth.esa.int/documents/653194/658149/AIRSAR_Flevoland

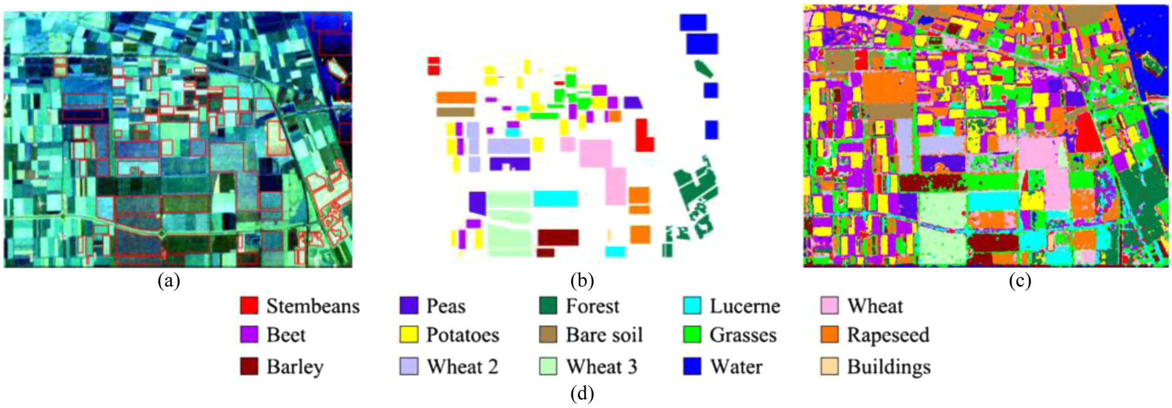


Figure 5.
Flevoland image. (a) Pseudo RGB. (b) Ground truth. (c) Classification result of [79]. (d) Color code for the ground truth.

methods [89] and [90], the MSE has been used as the loss function. As an example of regression with CNN, the result of ice concentration in [89] is illustrated in Figure 6. Estimating sea ice concentration from SAR images has also been addressed by Cooke *et al.* [91] where they used passive microwave data for training their CNN. Scarpa *et al.* [92] proposed a CNN to estimate normalized difference vegetation index through fusion of optical and SAR Sentinel data. They employed mean absolute error as the loss function. Estimation of ship orientation (angle regression) in SAR images, using a truncated VGG16 model as backbone, is addressed by Wang *et al.* [93]. All the convolutions in VGG16 have the kernel size of 3×3 and the number 16 in VGG16 implies that it has only 16 trainable layers. Smooth L1 is used in [93] as the loss function. Song *et al.* [94] studied regression with CNN for estimation of rough surface parameter from SAR images.

IMAGE DENOISING

Speckle, which is typically modeled as a multiplicative noise, is caused by constructive and destructive interference

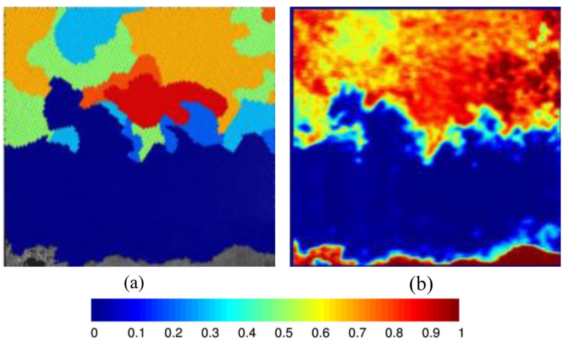


Figure 6.
Regression application for ice concentration presented in [89] together with color map (a) ground truth (b) regression result of CNN.

of the coherent returns scattered by small reflectors within one resolution cell [95]. Speckle noise hampers SAR image interpretability and many methods of speckle filtering have been proposed for SAR images in the past three decades. Wang *et al.* [95], [96] proposed a CNN architecture to reduce the speckle noise. By appropriate zero padding and avoiding any pooling layer, the output size of the last convolution layer is equal to that of the input. Subsequently, the input image, using a skip-connection, is divided by the output (speckle) to produce despeckled image. As an example of SAR image denoising with CNN, the model in [95] is illustrated in Figure 7.

Instead of employing residual learning to predict noise, Pan *et al.* [97] proposed another approach based on the FFDNet [98] for SAR image denoising. Fast and flexible denoising CNN (FFDNet), which uses downsampling–upsampling steps, takes a tunable noise level map as the input. In fact, a 2-D noise level map is fed to FFDNet during the training phase so the model parameters become more robust to the noise level. Consequently, only a single trained model is sufficient to denoise images with different noise levels.

Considering dilated convolutions, Zhang *et al.* [99] proposed a CNN by using residual learning and dilated convolutions for SAR image despeckling. In comparison with standard 2-D convolutions, dilated convolutions can deliver a larger receptive field with the same number of parameters. However, aggressive application of dilated convolution causes two problems, 1) spatial consistency between neighboring units becomes weak and 2) local structure cannot be extracted in higher layers [100]. Using dilated convolutions, Liu *et al.* [101] proposed a CNN framework to obtain a different denoised version of each SAR image. Afterward, they employed a guided filtering-based fusion algorithm to integrate them into a final-denoised image.

Mukherjee *et al.* [102] proposed a convolutional AE architecture for InSAR denoising purposes. Lattari *et al.* [103] proposed a modified version of U-Net, which is an

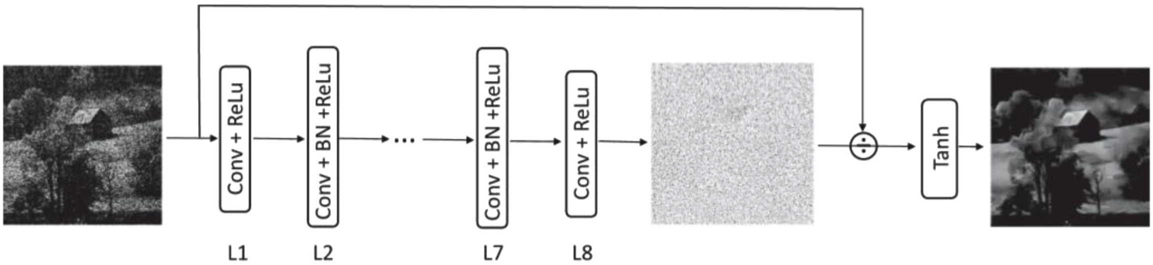


Figure 7.
Proposed *CNN* by [88] for speckle removal in SAR images.

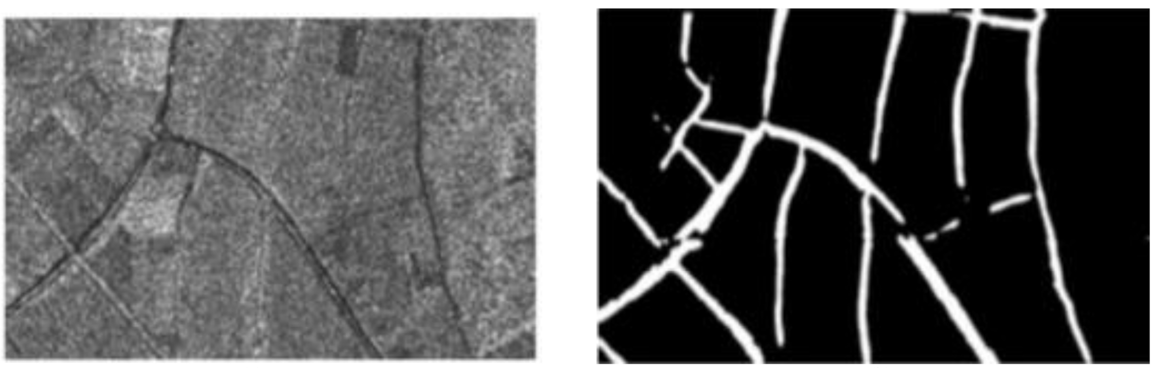
encoder–decoder CNN, for SAR image. despeckling. Li *et al.* [104], incorporated the convolutional block attention module (CBAM) [105] in their model to suppress the speckle noise in SAR images. CBAM has channel attention modules to concentrate on "what" is meaningful and spatial attention modules to concentrate on "where" the information is.

IMAGE SEGMENTATION

Image segmentation refers to the task of decomposing an image into different segments (regions) based on their characteristics. In semantic segmentation, meaningful labels are associated to each pixel of the image. In the RS community, the term "classification" is often preferred to the term "semantic segmentation" [106]. Moreover, terrain surface classification in SAR is very similar to the task of image segmentation in computer vision [37].

FCN [107] is the most remarkable architecture used for semantic segmentation. Instead of using one or more fully connected layers at the end of networks, FCNs exploit deconvolution layers. Deconvolution, in the context of CNN, means upsampling, therefore, for the sake of clarity the term "transposed convolution" is often preferred. Henry *et al.* [108] employed three CNN architectures of FCN-8s [107], DeepLabv3 [109], and Deep Residual U-Net [110] for road segmentation in SAR satellite images. The result of road segmentation with FCN-8 in [108] is illustrated in Figure 8. Roads, as they are seen as narrow objects in SAR images, are likely to be immersed in heavy clutter. Pixel-wise labels (1) are assigned to roads and (0) to the background. Mohammadimanesh *et al*. [111] proposed a new FCN architecture for semantic segmentation of polarimetric SAR imagery, especially for the classification of wetland complexes. Using inception modules [30] and skip connections in their suggested framework, input data is fed into a stack of convolutional filters (encoder) to extract high-level abstract features and a stack of transposed convolutional filters (decoder) to gradually up-sample the low-resolution output to the spatial resolution of the original input image. Inception modules [30] make parallel use of different kernel size (1×1, 3×3, and 5×5) of a feature map to improve the adaptability of the network to different scales.

U-Net [112] is an extension of the FCN and the letter "U" refers to symmetric encoder–decoder architecture. U-Net was initially proposed for biomedical image segmentation. DeepLabV3 segmentation model is also an encoder–decoder architecture that utilizes the dilated convolutions. Krestenitis *et al.* [113] addressed the applications of U-Net and DeepLabV3 for the segmentation of oil spills in satellite SAR images. A pixel with an oil label in natural image is somehow uncommon since large oil spills do not frequently occur. Cantorna *et al.* [114] have also discussed the problem of oil spills in satellite SAR images.



(a) (b)

Figure 8.
Road *segmentation* in [106]. (a) SAR image. (b) FCN-8 segmentation result.

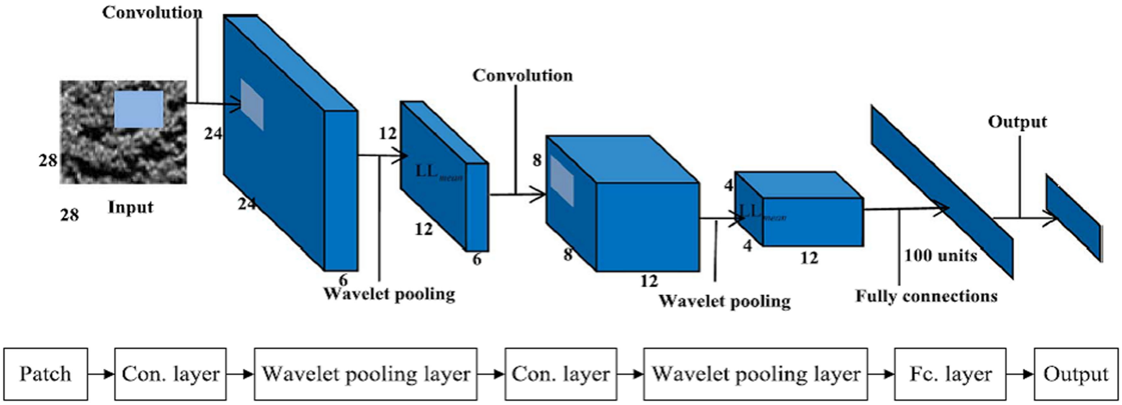


Figure 9.
SAR segmentation *architecture* using Wavelet pooling layers, proposed in [112].

Duan *et al.* [115] proposed a SAR Image segmentation method based on convolutional-wavelet neural network. They employed wavelet transform for pooling layers. In other words, a wavelet-constrained pooling layer is designed to substitute the conventional pooling in a CNN. Since conventional pooling layers in CNNs blindly take a maximum or average value from a rectangular window and this may lose the information of edges and corners, they employed wavelet transform for a more efficient pooling scheme. In fact, average brightness information is represented by low-frequency wavelet coefficients, and the textures and the edge information can be deduced by high-frequency ones. Their model is illustrated in Figure 9.

CHANGE DETECTION

The main aim of change detection is to distinguish changes between images that are taken from the same scene but at different times. The models for change detection typically take image pairs as input. Change detection is similar to a classification problem where pixels are divided into changed and unchanged categories. Change detection in general can be used for disaster monitoring [116], forest monitoring [117], video surveillance [118], etc.

Asokan *et al.* [119] reviewed the various change detection techniques in RS. Liu *et al.* [120] proposed a symmetric convolutional network for change detection between two input images, i.e., an optical and a SAR image with the same size. Gong *et al.* [121] proposed a change detection architecture by combining CNN and SAE for detecting changes from bi-temporal SAR images. They extracted features from the difference images by SAE and then these features were clustered into three classes as the pseudo labels (unchanged category, positive change, and negative change) for training CNN. The result of their ternary change detection is depicted in Figure 10. Iino *et al.* [122] proposed a short-term urban change detection method. They combined information contained in SAR satellite images with digital surface models to improve the accuracy of their suggested CNN-based network. Cao *et al.* [123] proposed a CNN-based denoising and change detection architecture for SAR images. Gao *et al.* [124] employed a convolutional-wavelet neural network, which was formally introduced in [115], for sea ice change detection from SAR images. Employing

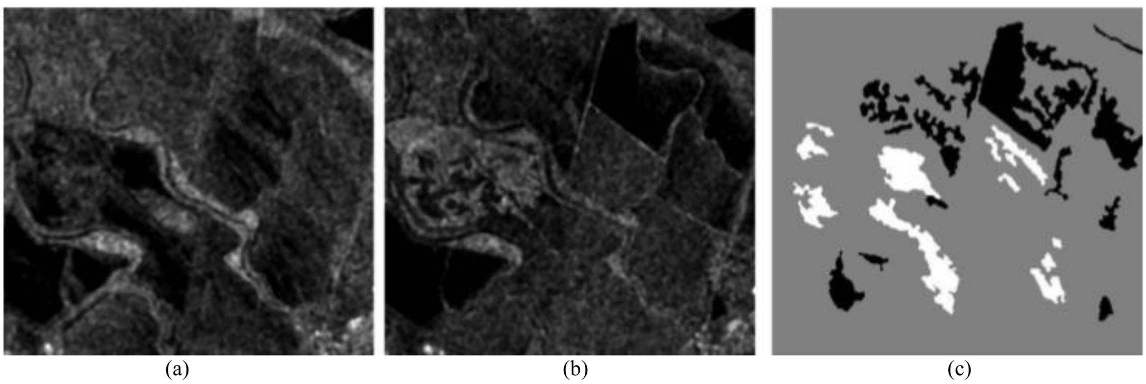


Figure 10.
Ternary change detection present in [118]. (a) Image acquired in 2003. (b) Image acquired in 2004. (c) Reference image.

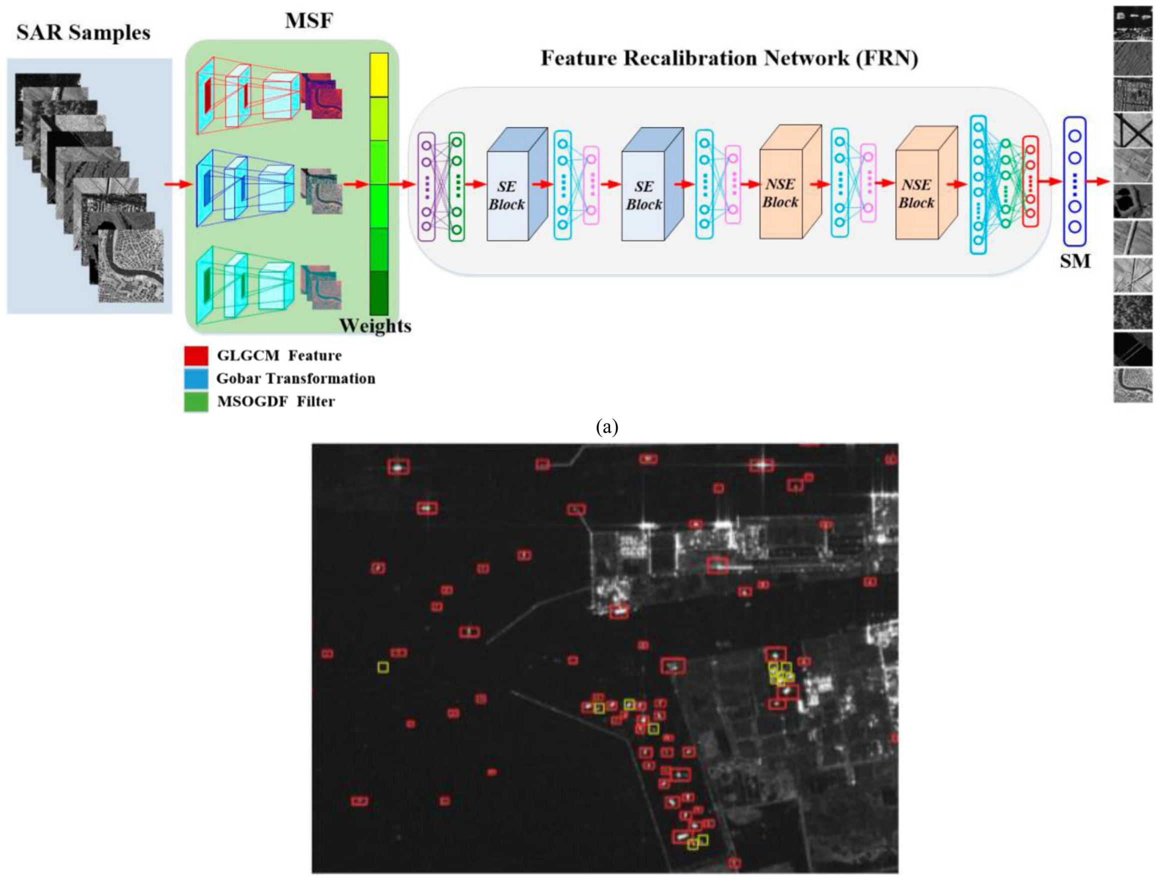


Figure 11.
Conceptual comparison between object detection and scene classification in SAR images. (a) Scene classification in [130]. (b) Object **(s**hip) detection using CNN in [131] where red and yellow rectangles show the detected and missed ships, respectively.

convolutional layers, Gao *et al*., [125] and [126], have also addressed the change detection problem in SAR images, by using a transfer learning-based multilevel fusion network and a channel weighting-based cascade network, respectively. Liu *et al*. [127] proposed a local-restricted CNN change detector for POLSAR data. Since neighboring pixels have similar characteristics, they formulated an extra regularization term, called local restriction, in the loss function and imposed it onto the output layer of their CNN.

Dong *et al*. [128] proposed a modified version of Siamese network for the problem of change detection in SAR images. The Siamese network [129] known as twin network, consists of two identical CNN-based subnetworks where all the weights and biases of the two subnetworks are tied together. The downside the of the Siamese network is that it can easily suffer from the interference of semantic distractors, in particular, the background [130], since it ignores the information of background [131]. The Siamese network can also be used to identify corresponding patches in SAR and optical images [132].

OBJECT DETECTION

It should be noted that SAR object detection is different from SAR scene classification. Scene classification aims at assigning only one label to a SAR image. On the contrary, object detection aims at assigning multiple labels to a SAR image where each label shows the location and the type of an object with a bounding box. An example to highlight the difference between scene classification [133] and object detection [134] in SAR images is illustrated in Figure 11.

Region-based CNNs (R-CNN) [135], Faster-RCNN [136], You Only Look Once (YOLO) [137] and SSD (Single Shot Multibox) [138] have been developed in the recent years for object detection tasks [139]. YOLO is a 1-stage-detector and it is significantly faster than Faster R–CNN since it performs region proposal and classification simultaneously. However, the detection accuracy of Faster R-CNN is better than that of YOLO. YOLO was implemented in a way similar to the human object recognition system [140]. Cui *et al*. [141] proposed a region-based CNN-based framework to locate ground targets on large scene SAR images by providing bounding boxes. In fact, they have randomly placed some targets from MSTAR dataset [8] in an image of fields, trees, and bushes to make up complex large scene SAR image. They used a fast sliding method to decompose the scene image into subimages to avoid dividing targets into different subimages.

An interesting topic of object detection in SAR images is ship detection. Kang *et al*. [134] proposed an R-CNN for ship detection that is based on contextual information and multilayer features. Zhao *et al*. [142], integrated the visual attention model in frequency domain into a CNN-based SAR ship detection framework to reduce the missed detections in small and densely clustered areas. Bentes *et al*. [143] proposed a framework in which ships in Terra-SAR-X images are detected by a standard CFAR algorithm and then classified by a CNN. For ship classification in SAR images, Wang *et al*. [144] studied transfer learning and fine-tuning of VGG16, VGG19, Xception [38], and InceptionV3 [145]. Lin *et al.* [146] exploited Squeeze and Excitation mechanism to improve the performance of Faster-RCNN for ship detection in SAR images. SENet [147] provides a weighting mechanism for each channel of the feature maps. Zhang *et al.* [148] proposed a grid-based CNN architecture that is inspired by YOLO [137] to decrease the computational time of ship detection problem in SAR images. Ai *et al.* [149] proposed a feature fusion scheme from two subnetworks for ship detection. They utilized the low-level Haar-like features [150] and high-level deep features of a CNN subnetwork. They showed that VGG-Net and Res-Net capture the high-level deep features of the ship targets. However, these networks ignore the local edge and texture information so their performance is poorer than the multi-level feature fusion-based discriminator of [149]. These low-level Haar-like features include horizontal and vertical edges, lines with different scales and rotations. Deng *et al.* [151] proposed a feature reuse strategy for their CNN-based ship detection. They used OpenSARShip dataset [70] that contains different types of ships collected from the Sentinel-1 satellite. SSDD has also been used in many ship detection studies. This dataset, which consists of 2358 ships in 1160 SAR satellite images, is labeled by Li *et al.* [71] by using LabelImg [152] as an image annotation software [153]. In [154]–[156], SAR images from the Chinese satellite Gaofen-3 that was launched in 2016 [157] are employed for the proposed CNN-based ship detection architectures. Other CNN-based SAR ship detection methods have been discussed in [158]–[160].

RELATED TOPICS TO CNN AND SAR

In this section, we address different aspects regarding the use of CNNs for SAR data processing. Different concepts such as CV architectures, data augmentation, and transfer learning will be discussed. Moreover, the use of convolutions in the generative adversarial network (GAN) and sparse coding will be addressed.

COMPLEX-VALUED CNN

Some researchers have attempted to exploit the phase inside CNNs because it is a well-known property of SAR that the phase contains very important information. One simple approach to transfer the phase information, which is embedded in complex SAR images, into the conventional RV-CNNs is to employ two channels where magnitude and phase (or real and imaginary parts) are taken as input. However, this is not an efficient approach and it is better to extract features in CV domain to maintain the integrality of information [161]. Mathematical formulations for CV-CNN have been thoroughly addressed in [162] and [163]. In CV-CNNs, all elements of CNNs including input – output layers, convolution layers, activation functions, and pooling layers are represented in complex form. Moreover, the complex backpropagation algorithm based on stochastic gradient descent has been studied. Zhang *et al.* [164] proposed a CV-CNN for polarimetric SAR image classification, where each pixel is classified into known terrain types, such as forest, grass, water, urban area, and sand. Their suggested network takes a tensor with six channels as input data similar to [80]. However, only the upper triangular part of the 3×3 PolSAR complex coherency matrix is used in [164]. Gao *et al.* [165] proposed a CV-CNN for ISAR image formation. They substituted conventional sigmoid function used in [164] by extending ReLU to its CV version. The input and output of this CV-CNN are radar echoes and expected images, respectively. They used the classical turntable model for generation of the training data. Sunaga *et al.* [166] proposed a CV-CNN for land form classification in InSAR data. Cao *et al*. [167] addressed CV-FCN for

PolSAR image segmentation. They introduced CV downsampling - CV upsampling to achieve pixel-wise classification. Moreover, they proposed a CV form of cross-entropy loss function and a weight initialization for their FCN network. Yu *et al.* [168] proposed CV-CNN for the SAR-ATR task in which they used only convolutions instead of pooling and fully connect layers to prevent overfitting. Wang *et al.* [169] proposed a CV-CNN for forest height mapping using polarimetric interferometric SAR (PolInSAR) data. Li *et al*. [170] combined CV-CNN with Contourlet filter bank to extract contour, edge and texture information for PolSAR image classification. Moving targets appear defocused and azimuthally displaced when using conventional SAR image formation algorithms. Mu *et al.* [171] addressed the problem of moving targets by using a CV-CNN, namely DeepImaging. Their proposed CV-CNN-based framework is depicted in

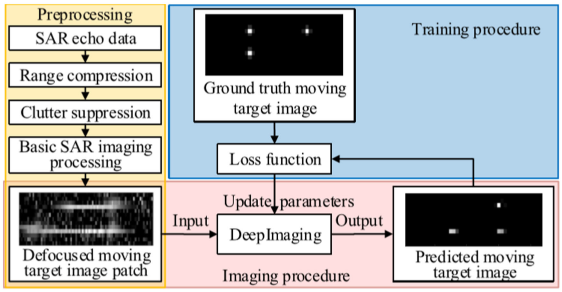


Figure 12.
The proposed CV-CNN based framework in [171] for moving targets imaging in SAR data.

Figure 12. Many defocused complex SAR images were used to train their CNN model. The loss function is defined to compare the real output with the ground truth image. One drawback of the model in [171] is that the authors neglected the geolocation task. In fact, they assumed that radial velocity of moving targets is zero and only refocusing of the moving targets is solved.

DATA AUGMENTATION

Data augmentation mechanism is widely employed to expand the diversity of training samples when the available dataset is small. In EO domain, this technique typically consists of scaling, cropping, padding, rotation, flipping, translation and so on. However, SAR data augmentation must be tied to the sensor itself meaning that rotation and flipping will not work as it does in EO domain because of the shadowing issues. Ding *et al*. [42] proposed a CNN for the SAR-ATR task with the emphasis on data augmentation techniques such as adding speckle noise, translation and pose synthesis. Kwak *et al.* [172] have also used speckle noise addition to make their CNN robust for SAR-ATR tasks. They adopted a similar approach to [42] i.e., using a filter to reduce speckle and then adding random speckle generated from a truncated exponential distribution. They employed Improved Lee sigma (ILS) filter [173] instead of median filtering that has been used in [42]. Du *et al.* [174] proposed a displacement- and rotation-insensitive CNN for data augmentation in MSTAR dataset. Their augmented images are shown in Figure 13 as an example of data augmentation in SAR images. Lv and Liu [175] used the concept of attributed scattering centers (ASC) for data augmentation. ASCs, which had already been employed for SAR-ATR tasks in 1990s by Potter and Moses [176], represent the electromagnetic scattering characteristics of targets. Note that the scattering centers are very much affected by the variation of the depression angle and thus the majority techniques relying only on ASC have poor robustness in experiments with extended depression angle variation [177]. Sparse representation was used in [175] to extract the ASCs of a single SAR image. Subsequently, some of the ASCs were selected for reconstruction of the image and many new training samples were generated by repeating this procedure. In fact, the overcomplete dictionary in sparserepresentation, smoothens the local radar reflectivity parameters and ultimately increases the robustness towards the depression angle variation [177]. Jiang and Zhou [178] have also proposed a hierarchical fusion of ASC matching and CNN for a more robust performance in SAR-ATR applications. Wang *et al*. [179] proposed a generative model based on Wasserstein AE [180] using convolutional layers, fully connected layers and residual modules that generate SAR images. Their model is useful for augmenting the dataset to increase the recognition accuracy.

SPARSE CODING AND CNN

Sparse coding (SC) is used for learning how to efficiently represent the input data by using a linear combination of

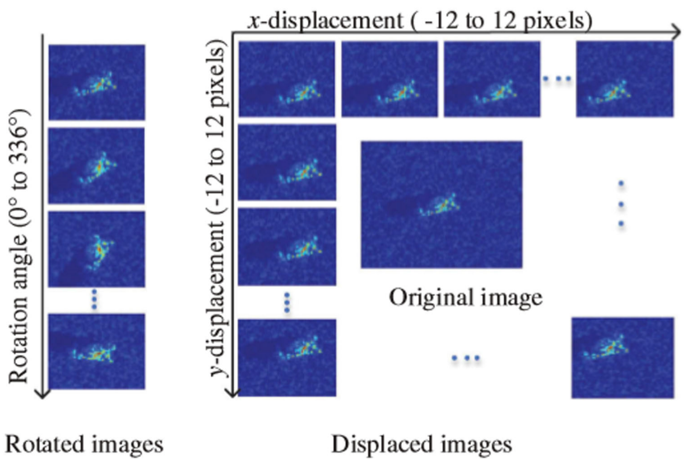


Figure 13
Data augmentation technique employed in [174].

atoms from an overcomplete dictionary. In other words, the main goal of sparse coding is to find a sparse representation from an overcomplete basis set [181]. Some studies addressed the combination of SC and CNN to solve a SAR task. SC-ATR searches for linear projections between the target and the feature spaces, while CNN-ATR searches for nonlinear projections [182]. Kechagias-Stamatis and Aouf [182] opened a new research area to improve SAR-ATR results by fusing these two concepts, i.e., SC and CNN. Instead of data-level and feature-level fusion schemes, they implemented a decision-level fusion. The reason why they ignored the data-level fusion is that both the CNN and the SC modules use the same input data in their architecture. Therefore, any fusion scheme in this sense is meaningless. Moreover, regarding the processing time requirement in near real-time applications, they have neglected feature-level fusion. Gao *et al*. [183] combined CNN and Joint Sparse Representation (JSR) for SAR target classification. They fused deep feature vectors from different convolution layers based on the multi-canonical correlation analysis (MCCA) [184] to maintain the relevance and eliminate the redundancy. Subsequently, labels were provided for targets based on the JSR error. JSR, which is a natural generalization of the single-task sparse representation, is a multi-task problem and it considers the inner correlations between different tasks [185], [186]. It is worthwhile to mention that Zhang *et al*. [187] had already addressed the concept of multi-view SAR-ATR using JSR in 2012, however, they did not exploit CNN. Lv *et al*. [188] have also employed feature maps from different convolution layers to generate multi-level deep features. Afterwards, they used JSR for the classification of generated multi-level deep features. Li *et al*. [189] employed sparse coding to reduce the computation burden and memory occupation of a sliding window FCN for PolSAR image classification.

TRANSFER LEARNING

Transfer learning is used to solve new tasks of which we do not have sufficient data for training, by transferring the knowledge from previous successful models. Transfer learning focuses on borrowing knowledge from one task (source domain) to another related task (target domain). One approach of transfer learning is to reuse a pretrained model. A pretrained model, as already trained on a large dataset, can be employed again to solve a new but similar problem. Transfer learning can decrease the required time for training procedures and increase the accuracy of the model when the labeled data is not sufficient for training. It is typically done by substituting the last layers of the network by new layers to fit the new problem. Finally, based on the size of the new dataset and its similarity to the older one, the entire model, or only some layers, will be fine-tuned. A typical case is the use of CNNs pretrained by optical images and then applied to SAR images. However, reusing well-known pretrained networks does not always achieve satisfactory performance for SAR applications since there exists a prominent discrepancy between SAR and optical images [190]. Instead, some studies followed different approaches such as transferring knowledge from simulated SAR data, unlabeled SAR images, and so on. Hansen *et al*. [191] showed that a CNN, pretrained on simulated data outperforms the one that is trained only on real data, especially when the labeled real data is not sufficient. They transferred knowledge from a simulated SAR dataset and fine-tuned it by using MSTAR dataset. Similarly, Wang *et al*. [192] utilized transfer learning between simulated SAR data and real SAR data to solve the problem of insufficient training samples. They used adversarial domain adaptation [193] to handle the problem of domains shift between the source and the target datasets. Huang *et al*. [194] proposed an assembled CNN architecture, including a reconstruction pathway and a classification pathway, for transferring the knowledge from a large number of unlabeled SAR images. In other words, they used unlabeled SAR images to train the reconstruction pathway and then reused the pretrained convolutional layers for SAR image-based classification. Their proposed framework is depicted Figure 14. Cui *et al*. [195] used unlabeled SAR images to feed both a CNN and an assistant classifier (SVM).

Afterwards, those unlabeled samples whose category recognition confidence is higher than a certain threshold are fed to the CNN again for fine tuning. Sun *et al*. [196] studied transferring knowledge about aspect angles of ground target from the source domain to the target domain. They proposed an angular rotation generative network to tackle the lack of training data at different aspect angles, which inevitably deteriorates the performance. Rostami *et al*. [197] proposed transferring knowledge from EO domain to the SAR domain using the Sliced Wasserstein Distance (SWD) [198] to measure and then minimize the discrepancy between the source - target domains. Some studies have also benefited from well-known CNN architectures that have been originally developed for natural images in the computer vision field. Zhong *et al*. [199] transferred the knowledge from CaffeNet [200], which is a modified version of Alexnet, to solve SAR classification tasks. Their suggested transfer-learning scheme is depicted in Figure 15 where they have employed MSTAR dataset [8] as well. Wang *et al*. [201] transferred a partial structure of VGGNet pretrained on the ImageNet data for the target detection in SAR images. Since the colorful natural images have three channels,

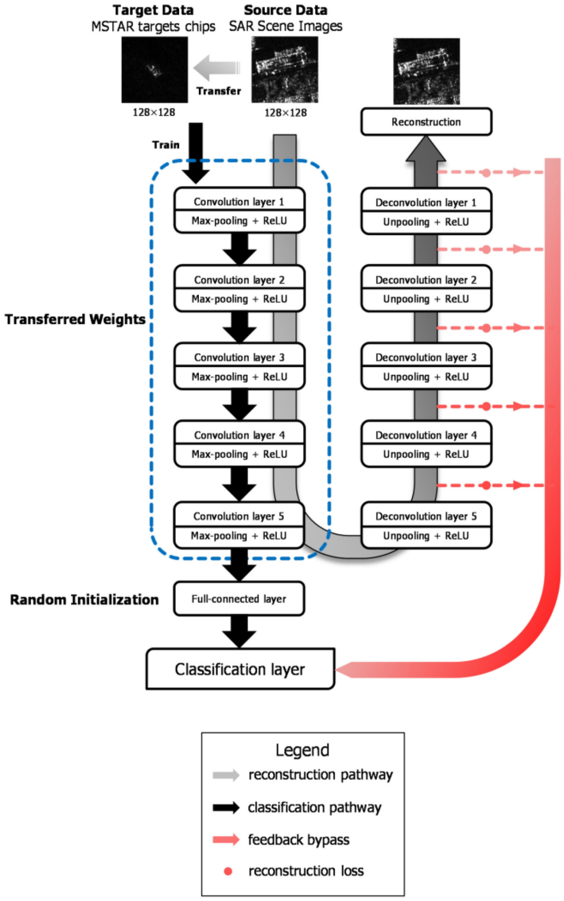


Figure 14.
Transferring knowledge from unlabeled SAR images proposed by [194].

they utilized the subaperture decomposition technique [202] to obtain three-channel subaperture SAR images. Utilizing a large pretrained network has also some downsides. First, since these pretrained networks contain millions or billions of parameters, the transferred networks tend to overfit on the scarce SAR data set. Second, SAR image classification are sometimes implemented in resource-constrained systems where storage, computation, and power are limited, such as embedded airborne SAR processing devices. In some scenarios, such as ground moving target indication (GMTI) in a terminal military guidance, it is required to have low-latency analysis, therefore, deep and large pretrained networks, which typically have considerable inference and storage costs, are not suitable [199]. Falqueto *et al*. [203] used VGGNet for the specific problem of oil rig recognition on Sentinel-1 SAR Images. Wurm *et al*. [204] studied the transfer learning capabilities of FCNs, from very high-resolution (VHR) optical satellite imagery of Quickbird to active SAR imagery of TerraSAR-X.

DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS

GANs [205] are DL architectures typically used for generating new instances of the input data that mimic the real data. They can also be used to distinguish between real and fake data. A GAN consists of a generator network and a discriminator network that compete against each other. The generator network tries to produce fake data and the discriminator tries to identify the real data from the fake one in order not to be fooled by the generator. At the end of this adversarial game, they reach the Nash equilibrium point. Radford *et al.* [206] introduced deep convolutional GANs (DCGANs) in 2015. Guo *et al*. [207] used DCGANs to implement a SAR image simulator. This simulator could be helpful to synthesize SAR images in a desired observation angle from a limited set of aspects. This is important when orbital geometry limitations and high maintenance costs are taken into consideration. Gao *et al*. [208] employed DCGANs to predict the labels of SAR samples by training the network with a

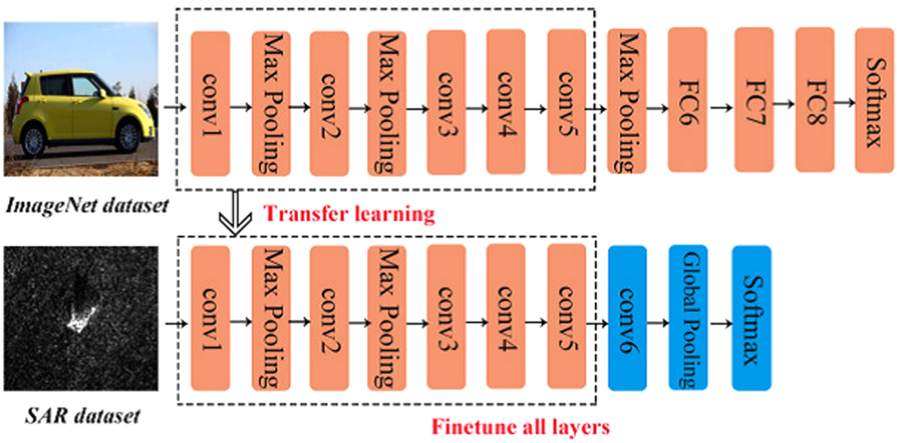


Figure 15.
Suggested transfer learning scheme by [195].

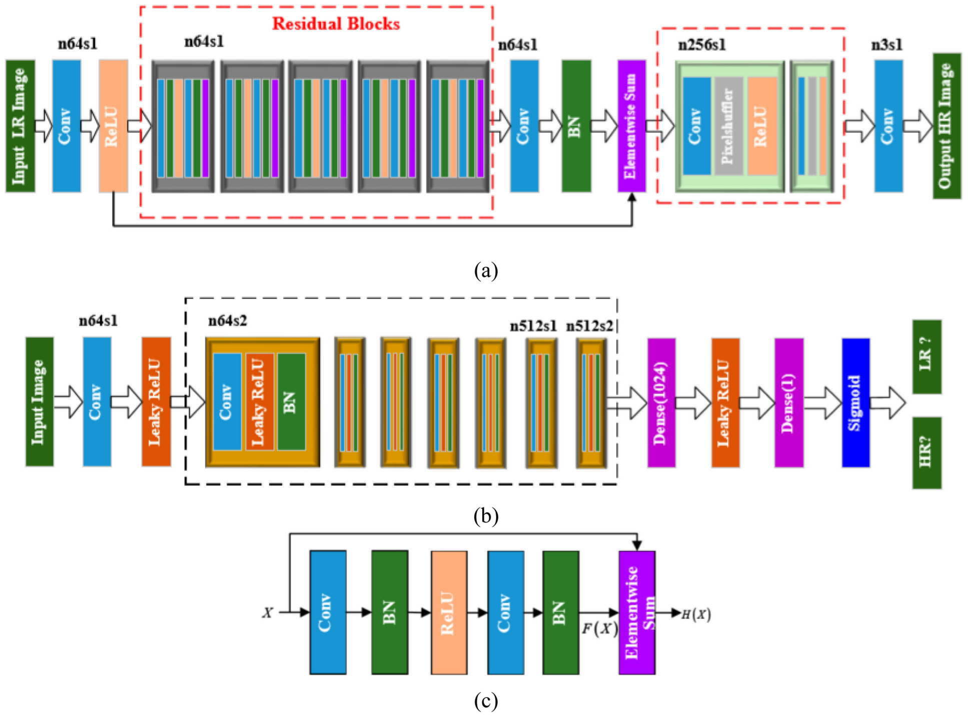


Figure 16.
Proposed GAN architecture of [206]. (a) Generator. (b) Discriminator. (c) Residual block.

small amount of labeled samples and then extending the labeled set iteratively. They used a cotraining [209] method to perform this task by utilizing a few labeled samples to predict the labels of the unlabeled samples at first. Afterward, samples with high confidence were chosen and added to the previous labeled set for the next training iteration. Shi *et al.* [210] employed DCGAN for SAR image enhancement and used the result for the SAR-ATR task. Their proposed architecture is depicted in Figure 16 as an example of DCGAN. Zhang *et al.* [211] used DCGAN for transferring knowledge from unlabeled SAR images. They trained a DCGAN with unlabeled samples to learn generic features of SAR images and reused the learned parameters for the SAR-ATR task.

CONCLUSION

CNN is undoubtedly the most successful DL technique in computer vision and it has attracted much attention in both civilian and military applications. In the same way, SAR plays a significant role in RS by offering high-resolution images with all-day and all-weather capabilities. In this survey, we have systematically reviewed the publications that applied CNN to different subareas of SAR data analysis such as ATR, scene classification with an emphasis on LULC, object detection with an emphasis on ships, segmentation, change detection, denoising, and regression for parameters estimation. Moreover, we discussed different topics related to CNN such complex architectures, transfer learning, and data augmentation. Despite the fact that CNN showed its promising role in tackling different SAR problems, we believe that this field is young and in the following years, more attention should be paid to address the remaining challenges. Some future research directions are listed below.

1) Taking the discrepancy between natural and SAR images into account, there is still potential to study the feasibility of transfer learning from well-known CNN architectures in the computer vision domain to SAR domain.

2) CV-CNNs have been developed to exploit the phase information embedded in SAR images. There are only a few studies that applied CV-CNNs to SAR images and it is expected that in the upcoming years more scholars will explore the capabilities of CV-CNNs.

3) The deeper the CNN is, the more complicated features can be extracted. However, more parameters have to be trained and more training samples are required to prevent overfitting. Due to the lack of large-scale annotated SAR images, effective data augmentation techniques, exploitation of unlabeled SAR images, and other possible solutions should be studied in more detail. Moreover,

extensive tests and validations of novel algorithms that are based on the use of DL are needed to confirm their applicability and effectiveness. To this purpose, appropriate metrics and suitable datasets must be developed and created to fulfill this task.

ACKNOWLEDGMENTS

The authors would like to express their sincere appreciation to the associate editor and the anonymous reviewers who carefully reviewed this manuscript and provided valuable and helpful remarks.

REFERENCES

[1] Y. Li, H. Zhang, X. Xue, Y. Jiang, and Q. Shen, "Deep learning for remote sensing image classification: A survey," *Wiley Interdiscipl. Rev. Data Mining Knowl. Discov.*, vol. 8, no. 6, 2018, Art. no. e1264.

[2] M. Z. Alom *et al.*, "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, 2019, Art. no. 292.

[3] Y. L. Cun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[4] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, Dec. 2010.

[5] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[6] G. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput*, vol. 18, no. 7, pp. 1527–1554, 2006.

[7] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 166–177, 2019.

[8] The air force moving and stationary target recognition database. 2014. [Online]. Available: https://www.sdms.afrl.af.mil/index.php?collection=mstar

[9] D. A. E. Morgan, "Deep convolutional neural networks for ATR from SAR imagery," *Algorithms Synth. Aperture Radar Imagery,* vol. 9475, 2015, Art. no. 94750F.

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

[11] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *Proc. Int. Conf. Artif. Neural Netw.*, 2011, pp. 52–59.

[12] R. Shang, Y. Liu, J. Wang, L. Jiao, and R. Stolkin, "Stacked auto-encoder for classification of polarimetric SAR images based on scattering energy," *Int. J. Remote Sens.,* vol. 40, no. 13, pp. 5094–5120, 2019.

[13] L. Zhang, W. Ma, and D. Zhang, "Stacked sparse autoencoder in PolSAR data classification using local spatial information," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 9, pp. 1359–1363, Jul .2016.

[14] J. Guo, H. Li, N. J., W. Han, Z. W., and Z. S. Zhou, "Feature dimension reduction using stacked sparse autoencoders for crop classification with multi-temporal, quad-pol SAR data," *Remote Sens,* vol. 12, no. 2, 2020, Art. no. 321.

[15] J. Wang, B. Hou, L. Jiao, and S. Wang, "POL-SAR image classification based on modified stacked autoencoder network and data distribution," *IEEE Trans. Geosci. Remote Sens.,* vol. 58 no. 3, pp. 1678–1695, Nov. 2019.

[16] B. Hou, H. Kou, and L. Jiao, "Classification of polarimetric SAR images using multilayer autoencoders and superpixels," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 7, pp. 3072–3081, Jul. 2016.

[17] P. Planinsic and D. Gleich, "Temporal change detection in SAR images using log cumulants and stacked autoencoder," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 297–301, Feb. 2018.

[18] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, "Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images," *IEEE Trans. Ind. Informat.*, vol. 14, no. 12, pp. 5530–5538, Dec. 2018.

[19] Y. Freund and D. Haussler, "Unsupervised learning of distributions on binary vectors using two layer networks," Univ. California, Santa Cruz, CA, USA, Tech. Rep. UCSC-CRL-94-25, 1994.

[20] Z. Zhao, L. Jiao, J. Zhao, J. Gu, and J. Zhao, "Discriminant deep belief network for high-resolution SAR image classification," *Pattern Recognit,* vol. 61, pp. 686–701, Jan. 2017.

[21] Z. Zhao, L. Guo, M. Jia, and L. Wang, "The generalized gamma-DBN for high-resolution SAR image classification," *Remote Sens,* vol. 10, no. 6, 2018, Art. no. 878.

[22] F. Liu, L. Jiao, B. Hou, and S. Yang, "POL-SAR image classification based on wishart DBN and local spatial information," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3292–3308, Jun. 2016.

[23] Q. Lv, Y. Dou, X. Niu, J. Xu, J. Xu, and F. Xia, "Urban land use and land cover classification using remotely sensed SAR data through deep belief networks," *J. Sensors*, vol. 2015, Jul. 2015, Art.no. 538063.

[24] J. Geng, H. Wang, J. Fan, and X. Ma, "SAR image classification via deep recurrent encoding neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2255–2269, Apr. 2018.

[25] B. Yonel, E. Mason, and B. Yazici, "Deep learning for passive synthetic aperture radar," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 1, pp. 90–103, Feb. 2018.

[26] E. Ndikumana, D. Ho Tong Minh, N. Baghdadi, D. Courault, and L. Hossard, "Deep recurrent neural network for agricultural classification using multitemporal SAR sentinel-1 for Camargue France," *Remote Sens,* vol. 10, no. 8, 2018, Art. no. 1217.

[27] D. H. T. Minh *et al.*, "Deep recurrent neural networks for winter vegetation quality mapping via multitemporal sar sentinel-1," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 464–468, Mar. 2018.

[28] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[30] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.

[31] 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.

[32] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: Theories tools and challenges for the community," *J. Appl. Remote Sens.*, vol. 11, no. 4, 2017, Art. no. 042609.

[33] J. E. Ball, D. T. Anderson, and C. S. Chan, "Feature and deep learning in remote sensing applications," *J. Appl. Remote Sens.*, vol. 11, no. 4, 2018, Art. no. 042609.

[34] J. Song, S. Gao, Y. Zhu, and C. Ma, "A survey of remote sensing image classification based on CNNs," *Big Earth Data*, vol. 3, no. 3, pp. 232–254, 2019.

[35] Y. Gu, Y. Wang, and Y. Li, "A survey on deep learning-driven remote sensing image scene understanding: Scene classification, scene retrieval and scene-guided object detection," *Appl. Sci.*, vol. 9, no. 10, 2019, Art. no. 2110.

[36] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[37] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

[38] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1800–1807.

[39] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, *arXiv:1511.07122*.

[40] Y. Wang, Y. Li, Y. Song, and X. Rong, "The influence of the activation function in a convolution neural network model of facial expression recognition," *Appl. Sci.*, vol. 10, no. 5, 2020, Art. no. 1897.

[41] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.

[42] 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.

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

[44] 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.

[45] F. Gao, T. Huang, J. Sun, J. Wang, A. Hussain, and E. Yang, "A new algorithm of SAR image target recognition based on improved deep convolutional neural network," *Cogn. Comput.*, vol. 5, pp. 1–16, Jun. 2018.

[46] F. Zhang, Y. Wang, J. Ni, Y. Zhou, and W. Hu, "SAR target small sample recognition based on CNN cascaded features and adaboost rotation forest," *IEEE Geosci. Remote Sens. Lett*, vol. 17, no. 6, pp. 1008–1012, Jun. 2020.

[47] 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. 99, no. 99, pp. 1–12, Jul. 2018.

[48] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-margin softmax loss for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1–1. [Online]. Available: http://proceedings.mlr.press/v48/liud16.pdf

[49] L. Wang, X. Bai, and F. Zhou, "SAR ATR of ground vehicles based on ESENet," *Remote Sens,* vol. 11, no. 11, 2019, Art. no. 1316.

[50] Y. Fathullah, C. Zhang, and P. C. Woodland, "Improved large-margin softmax loss for speaker diarisation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2020, pp. 7104–7108.

[51] Y. Liu, L. He, and J. Liu, "Large margin softmax loss for speaker verification," 2019, arXiv:1904.03479, [Online]. Available: http://arxiv.org/abs/1904.03479

[52] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-margin softmax loss for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2016*,* pp. 507–516.

[53] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "SphereFace: Deep hypersphere embedding for face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 6738–6746.

[54] J. Shao, C. Qu, J. Li, and S. Peng, "A lightweight convolutional neural network based on visual attention for SAR image target classification," *J. Sensors*, vol. 18, no. 9, 2018, Art. no. 3039.

[55] H. Chen, F. Zhang, B. Tang, Q. Yin, and X. Sun, "Slim and efficient neural network design for resource-constrained SAR target recognition," *Remote Sens.*, vol. 10, no. 10, 2018, Art. no. 1618.

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

[57] J. Ba and R. Caruana, "Do deep nets really need to be deep?," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2654–2662.

[58] 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.

[59] 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.

[60] J. Pei, Y. Huang, Z. Sun, Y. Zhang, J. Yang, and T. S. Yeo, "Multiview synthetic aperture radar automatic target recognition optimization: Modeling and implementation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no.no. 11, pp. 6425–6439, Nov. 2018.

[61] P. Zhao, K. Liu, H. Zou, and X. Zhen, "Multi-stream convolutional neural network for SAR automatic target recognition," *Remote Sens,* vol. 10, no. 9, pp. 1473, Sep. 2018.

[62] N. Wang, Y. Wang, H. Liu, Q. Zuo, and J. He, "Feature-fused SAR target discrimination using multiple convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1695–1699, Oct. 2017.

[63] J. H. Cho and C. G. Park, "Multiple feature aggregation using convolutional neural networks for SAR image-based automatic target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 12, pp. 1882–1886, Dec. 2018.

[64] X. Bai, X. Zhou, F. Zhang, L. Wang, R. Xue, and F. Zhou, "Robust Pol-ISAR target recognition based on ST-MC-DCNN," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9912–9927, Dec. 2019.

[65] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial transformer networks," *in Proc. Adv. Neural Inf. Process. Syst.*, 2015*,* pp. 2017–2025.

[66] L. Finnveden, Y. Jansson, and T. Lindeberg, "Understanding when spatial transformer networks do not support invariance, and what to do about it," in *Proc. 25th IEEE Int. Conf. Pattern Recognit.*, May 2020, pp. 3427–3434.

[67] Z. Lin, K. Ji, M. Kang, X. Leng, and H. Zou, "Deep convolutional highway unit network for SAR target classification with limited labeled training data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1091–1095, Jul. 2017.

[68] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Training very deep networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 2377–2385.

[69] Z. Yue *et al.*, "A novel semi-supervised convolutional neural network method for synthetic aperture radar image recognition," *Cogn. Comput.,* vol. 13, pp. 795–806, 2021.

[70] L. Huang *et al.*, "OpenSARShip: A dataset dedicated to sentinel-1 ship interpretation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 1, pp. 195–208, Jan. 2018.

[71] J. Li, C. Qu, and J. Shao, "Ship detection in SAR images based on an improved faster R-CNN," in *Proc. SAR Big Data Era, Models, Methods Appl.*, Nov. 2017, pp. 1–6.

[72] N. Liu, Z. Cao, Z. Cui, Y. Pi, and S. Dang, "Multi-scale proposal generation for ship detection in SAR images," *Remote Sens,* vol. 11, no. 5, pp. 526–546, Mar. 2019.

[73] M. Ohki and M. Shimada, "Large-area land use and land cover classification with quad compact and dual polarization SAR data by PALSAR-2," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5550–5557, Sep. 2018.

[74] Z. Xu, R. Wang, H. Zhang, N. Li, and L. Zhang, "Building extraction from high-resolution SAR imagery based on deep neural networks," *Remote Sens. Lett.*, vol. 8, no. 9, pp. 888–896, 2017.

[75] M. Shahzad, M. Maurer, F. Fraundorfer, Y. Wang, and X. X. Zhu, "Buildings detection in VHR SAR images using fully convolution neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1100–1116, Feb. 2019.

[76] Y. Li, S. Martinis, and M. Wieland, "Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 178–191, 2019.

[77] Y. N. Lin, S. H. Yun, A. Bhardwaj, and E. M. Hill, "Urban flood detection with sentinel-1 multi-temporal synthetic aperture radar (SAR) observations in a bayesian framework: A case study for hurricane matthew," *Remote Sens,* vol. 11, no. 15, 2019, Art. no. 1778.

[78] S. W. Chen and C. S. Tao, "PolSAR image classification using polarimetric-feature-driven deep convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 627–631, Apr. 2018.

[79] S. W. Chen, X. S. Wang, and M. Sato, "Uniform polarimetric matrix rotation theory and its applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4756–4770, Aug. 2014.

[80] Y. Zhou, H. Wang, F. Xu, and Y.-Q. Jin, "Polarimetric SAR image classification using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1935–1939, Dec. 2016.

[81] G. Singh, S. Mohanty, Y. Yamazaki, and Y. Yamaguchi, "Physical scattering interpretation of POLSAR coherency matrix by using compound scattering phenomenon," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2541–2556, Apr. 2019.

[82] X. Liu, L. Jiao, X. Tang, Q. Sun, and D. Zhang, "Polarimetric convolutional network for PolSAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 3040–3054, May 2019.

[83] L. Wang, X. Xu, H. Dong, R. Gui, and F. L. Pu, "Multi-pixel simultaneous classification of PolSAR image using convolutional neural networks," *Sensors*, vol. 18, no. 3, 2018, Art. no. 769.

[84] M. Ahishali, S. Kiranyaz, T. Ince, and M. Gabbouj, “Dual and single polarized SAR image classification using compact convolutional neural networks,” *Remote Sens.*, vol. 11, no. 11, 2019, Art. no. 1340.

[85] H. Bi, J. Sun, and Z. Xu, “A graph-based semisupervised deep learning model for PolSAR image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2116–2132, Apr. 2019.

[86] H. Bi, F. Xu, Z. Wei, Y. Xue, and Z. Xu, “An active deep learning approach for minimally supervised PolSAR image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9378–9395, Nov. 2019.

[87] W. Xie *et al.*, “PolSAR image classification via Wishart-AE model or wishart-CAE model,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3604–3615, Aug. 2017.

[88] Z. Ren, B. Hou, Z. Wen, and L. Jiao, “Patch-sorted deep feature learning for high resolution SAR image classification,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3113–3126, Sep. 2018.

[89] L. Wang, K. A. Scott, L. Xu, and D. A. Clausi, “Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: A case study,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4524–4533, Aug. 2016.

[90] L. Wang, K. A. Scott, and D. A. Clausi, “Sea ice concentration estimation during freeze-up from SAR imagery using a convolutional neural network,” *Remote Sens,* vol. 9, no. 5, 2017, Art. no. 408.

[91] C. L. Cooke and K. A. Scott, “Estimating sea ice concentration from SAR: Training convolutional neural networks with passive microwave data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4735–4747, Jul. 2019.

[92] G. Scarpa, M. Gargiulo, A. Mazza, and R. Gaetano, “A CNN-based fusion method for feature extraction from sentinel data,” *Remote Sens,* vol. 10, no. 2, 2018, Art. no. 236.

[93] J. Wang, C. Lu, and W. Jiang, “Simultaneous ship detection and orientation estimation in SAR images based on attention module and angle regression,” *Sensors*, vol. 18, no. 9, 2018, Art. no. 2851.

[94] T. Song, L. Kuang, L. Han, Y. Wang, and Q. H. Liu, “Inversion of rough surface parameters from SAR images using simulation-trained convolutional neural networks,” *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 7, pp. 1130–1134, Jul. 2018.

[95] P. Wang, H. Zhang, and V. M. Patel, “SAR image despeckling using a convolutional neural network,” *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1763–1767, Dec. 2017.

[96] 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 Observ. Remote Sens.*, vol. 11, no. 11, pp. 4180–4192, Nov. 2018.

[97] T. Pan, D. Peng, W. Yang, and H. C. Li, “A filter for SAR image despeckling using pre-trained convolutional neural network model,” *Remote Sens.*, vol. 11, no. 20, 2019, Art. no. 2379.

[98] K. Zhang, W. Zuo, and L. Zhang, “FFDNet: Toward a fast and flexible solution for CNN based image denoising,” *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.

[99] Q. Zhang, Q. Yuan, J. Li, Z. Yang, and X. Ma, “Learning a dilated residual network for SAR image despeckling,” *Remote Sens,* vol. 10, no. 2, 2018, Art. no. 196.

[100] R. Hamaguchi, A. Fujita, K. Nemoto, T. Imaizumi, and S. Hikosaka, “Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery,” in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2018, pp. 1442–1450.

[101] S. Liu *et al.*, “Convolutional neural network and guided filtering for SAR image denoising,” *Remote Sens,* vol. 11, no. 6, Mar. 2019, Art. no. 702.

[102] S. Mukherjee, A. Zimmer, N. K. Kottayil, X. Sun, P. Ghuman, and I. Cheng, “CNN-based InSAR denoising and coherence metric,” in *Proc. IEEE Sensors*, Dec. 2018, pp. 28–31.

[103] F. Lattari, B. G. Leon, F. Asaro, A. Rucci, C. Prati, and M. Matteucci, “Deep learning for SAR image despeckling,” *Remote Sens,* vol. 11, no. 13, 2019, Art. no. 1532.

[104] J. Li, Y. Li, Y. Xiao, and Y. Bai, “HDRANet: Hybrid dilated residual attention network for SAR image despeckling,” *Remote Sens.*, vol. 11, no. 24, 2019, Art. no. 2921.

[105] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, “CBAM: Convolutional block attention module,” in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 3–19.

[106] A. Romero, C. Gatta, and G. Camps-Valls, “Unsupervised deep feature extraction for remote sensing image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1349–1362, Mar. 2016.

[107] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Oct. 2015, pp. 3431–3440.

[108] C. Henry, S. M. Azimi, and N. Merkle, “Road segmentation in SAR satellite images with deep fully convolutional neural networks,” *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 12, pp. 1867–1871, Dec. 2018.

[109] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 833–851.

[110] Z. Zhang, Q. Liu, and Y. Wang, “Road extraction by deep residual U-net,” *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, May 2018.

[111] F. Mohammadimanesh, B. Salehi, M. Mahdianpari, E. Gill, and M. Molinier, “A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem,” *ISPRS J. Photogramm. Remote Sens.*, vol. 151, pp. 223–236, May 2019.

[112] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2015, pp. 234–241.

[113] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," *Remote Sens,* vol. 11, no. 15, Jul. 2019, Art. no. 1762.

[114] D. Cantorna, C. Dafonte, A. Iglesias, and B. Arcay, "Oil spill segmentation in SAR images using convolutional neural networks. A comparative analysis with clustering and logistic regression algorithms," *Appl. Soft Comput.*, vol. 84, 2019, Art. no. 105716.

[115] Y. Duan, F. Liu, L. Jiao, P. Zhao, and L. Zhang, "SAR image segmentation based on convolutional-wavelet neural network and Markov random field," *Pattern Recognit,* vol. 64, pp. 255–267, Apr. 2017.

[116] G. D. Martino, A. Iodice, D. Riccio, and G. Ruello, "A novel approach for disaster monitoring: Fractal models and tools," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1559–1570, Jun. 2007.

[117] T. Hame, I. Heiler, and J. S. Miguel-Ayanz, "An unsupervised change detection and recognition system for forestry," *Int. J. Remote Sens.*, vol. 19, no. 6, pp. 1079–1099, 1998.

[118] D.-M. Tsai and S.-C. Lai, "Independent component analysis-based background subtraction for indoor surveillance," *IEEE Trans. Image Process.,* vol. 18, no. 1, pp. 158–167, Jan. 2009.

[119] A. Asokan and J. Anitha, "Change detection techniques for remote sensing applications: A survey," *Earth Sci. Inform.*, vol. 12, no. 2, pp. 143–160, 2019.

[120] J. Liu, M. Gong, K. Qin, and P. Zhang, "A deep convolutional coupling network for change detection based on heterogeneous optical and radar images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 545–559, Mar. 2016.

[121] M. Gong, H. Yang, and P. Zhang, "Feature learning and change feature classification based on deep learning for ternary change detection in SAR images," *J. Photogramm. Remote Sens.*, vol. 129, pp. 212–225, Jul. 2017.

[122] S. Iino, R. Ito, K. Doi, T. Imaizumi, and S. Hikosaka, "CNN-based generation of high-accuracy urban distribution maps utilising SAR satellite imagery for short-term change monitoring," *Int. J. Image Data Fusion*, vol. 9, no. 4, pp. 302–318, 2018.

[123] X. Cao, Y. Ji, L. Wang, B. Ji, L. Jiao, and J. Han, "SAR image change detection based on deep denoising and CNN," *IET Image Process,* vol. 13, no. 9, pp. 1509–1515, 2019.

[124] F. Gao, X. Wang, Y. Gao, J. Dong, and S. Wang, "Sea ice change detection in SAR images based on convolutional-wavelet neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1240–1244, Aug. 2019.

[125] Y. Gao, F. Gao, J. Dong, and S. Wang, "Transferred deep learning for sea ice change detection from synthetic-aperture radar images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 10, pp. 1655–1659, Oct. 2019.

[126] Y. Gao, F. Gao, J. Dong, and S. Wang, "Change detection from synthetic aperture radar images based on channel weighting-based deep cascade network," *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.,* vol. 12, no. 11, pp. 4517–4529, Nov. 2019.

[127] F. Liu, L. Jiao, X. Tang, S. Yang, W. Ma, and B. Hou, "Local restricted convolutional neural network for change detection in polarimetric SAR images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 818–833, Mar. 2019.

[128] H. Dong, W. Ma, Y. Wu, M. Gong, and L. Jiao, "Local descriptor learning for change detection in synthetic aperture radar images via convolutional neural networks," *IEEE Access*, vol. 7, pp. 15389–15403, Dec. 2019.

[129] J. Bromley *et al.*, "Signature verification using a 'Siamese' time delay neural network," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 7, no. 4, pp. 669–688, 1993.

[130] W. Ruan *et al.*, "Ticnet: A target-insight correlation network for object tracking," *IEEE Trans. Cybern.*, early access, May 25, 2021, doi: 10.1109/TCYB.2021.3070677.

[131] H. Pang *et al.*, "Research on target tracking algorithm based on siamese neural network," *Mobile Inf. Syst.*, vol. 4, pp. 1–11, 2021.

[132] L. H. Hughes, M. Schmitt, L. Mou, Y. Wang, and X. X. Zhu, "Identifying corresponding patches in SAR and optical images with a pseudo-siamese CNN," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 784–788, May 2018.

[133] L. Chen *et al.*, "A new deep learning algorithm for SAR scene classification based on spatial statistical modeling and features re-Calibration," *Sensors*, vol. 19, no. 11, 2019, Art. no. 2479.

[134] M. Kang, K. Ji, X. Leng, and Z. Lin, "Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection," *Remote Sens,* vol. 9, no. 8, 2017, Art. no. 860.

[135] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2014, pp. 580–587.

[136] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.

[137] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified real-time object detection," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.

[138] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.

[139] U. Alganci, M. Soydas, and E. Sertel, "Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images," *Remote Sens.*, vol. 12, no. 3, 2020, Art. no. 458.

[140] J. A. Kim, J. Y. Sung, and S. H. Park, "Comparison of Faster-RCNN, YOLO, and SSD for real-time vehicle type recognition," in *Proc. IEEE Int. Conf. Consum. Electron.-Asia*, Nov. 2020, pp. 1–4.

[141] Z. Cui, S. Dang, Z. Cao, S. Wang, and N. Liu, “Sar target recognition in large scene images via region-based convolutional neural networks,” *Remote Sens,* vol. 10, no. 5, 2018, Art.no. 776.

[142] J. Zhao, Z. Zhang, W. Yu, and T.-K. Truong, “A cascade coupled convolutional neural network guided visual attention method for ship detection from SAR images,” *IEEE Access*, vol. 6, pp. 50693–50708, Sep. 2018.

[143] C. Bentes, D. Velotto, and B. Tings, “Ship classification in terrasar-x images with convolutional neural networks,” *IEEE J. Ocean. Eng.*, vol. 43, no. 1, pp. 258–266, Jan. 2018.

[144] Y. Wang, C. Wang, and H. Zhang, “Ship classification in high-resolution SAR images using deep learning of small datasets,” *J. Sensors*, vol. 18, no. 9, 2018, Art. no. 2929.

[145] C. Szegedy, V. Vincent, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2818–2826.

[146] Z. Lin, K. Ji, X. Leng, and G. Kuang, “Squeeze and excitation rank faster R-CNN for ship detection in SAR images,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 5, pp. 751–755, May 2019.

[147] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,”in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.

[148] T. Zhang and X. Zhang, “High-speed ship detection in SAR images based on a grid convolutional neural network,” *Remote Sens.*, vol. 11, no. 10, May 2019, Art. no. 1206.

[149] J. Ai, R. Tian, Q. Luo, J. Jin, and B. Tang, “Multi-scale rotation-invariant Haar-like feature integrated CNN-based ship detection algorithm of multiple-target environment in SAR imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10070–10087, Dec. 2019.

[150] R. Lienhart and J. Maydt, “An extended set of haar-like features for rapid object detection,” in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2002, vol. 1, pp. 900–903.

[151] Z. Deng, H. Sun, S. Zhou, and J. Zhao, “Learning deep ship detector in SAR images from scratch,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 4021–4039, Jun. 2019.

[152] *LabelImg. Git Code*, 2015 [Online]. Available: https://github.com/tzutalin/labelImg

[153] T. Zhang, X. Zhang, J. Shi, and S. Wei, “Depthwise separable convolution neural network for high-speed SAR ship detection,” *Remote Sens.*, vol. 11, no. 21, 2019, Art. no. 2483.

[154] M. Ma, J. Chen, W. Liu, and W. Yang, “Ship classification and detection based on CNN using GF-3 SAR images,” *Remote Sens,* vol. 10 no. 12, 2018, Art. no. 2043.

[155] Q. An, Z. Pan, and H. You, “Ship detection in Gaofen-3 SAR images based on sea clutter distribution analysis and deep convolutional neural network,” *Sensors*, vol. 18, no. 2, 2018, Art. no. 334.

[156] Y. Wang, C. Wang, H. Zhang, Y. Dong, and S. Wei, “Automatic ship detection based on retinanet using multi-resolution Gaofen-3 imagery,” *Remote Sens.*, vol. 11, no. 5, pp. 531–544, 2019.

[157] J. Liu, X. Qiu, and W. Hong, “Automated ortho-rectified SAR image of GF-3 satellite using reverse-range-Doppler method,” in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2015, pp. 4445–4448.

[158] C. Chen, C. He, C. Hu, H. Pei, and L. Jiao, “MSARN: A deep neural network based on an adaptive recalibration mechanism for multiscale and arbitrary-oriented SAR ship detection,” *IEEE Access,* vol. 7, pp. 159262–159283, Nov. 2019.

[159] X. Zhang *et al.*, “A lightweight feature optimizing network for ship detection in SAR image,” *IEEE Access,* vol. 7, pp. 141662–141678, Sep. 2019.

[160] C. Lu and W. Li, “Ship classification in high-resolution SAR images via transfer learning with small training dataset,” *Sensors*, vol. 19, no. 1, 2019, Art. no. 63.

[161] X. Tan, M. Li, P. Zhang, Y. Wu, and W. Song, “Complex-valued 3-D convolutional neural network for PolSAR image classification,” *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 1022–1026, Jun. 2020.

[162] N. Guberman, “On complex valued convolutional neural networks,” 2016. [Online]. Available: http://arxiv.org/abs/1602. 09046

[163] M. Tygert, J. Bruna, S. Chintala, Y. LeCun, S. Piantino, and A. Szlam, “A mathematical motivation for complex-valued convolutional networks,” *Neural Comput.*, vol. 28, no. 5, pp. 815–825, May 2016.

[164] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, “Complex-valued convolutional neural network and its application in polarimetric SAR image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7177–7188, Dec. 2017.

[165] J. Gao, B. Deng, Y. Qin, H. Wang, and X. Li, “Enhanced radar imaging using a complex-valued convolutional neural network,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 35–39, Jan. 2019.

[166] Y. Sunaga, R. Natsuaki, and A. Hirose, “Land form classification and similar land-shape discovery by using complex-valued convolutional neural networks,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 24–38, Oct. 2019.

[167] Y. Cao, Y. Wu, P. Zhang, W. Liang, and M. Li, “Pixel-wise PolSAR image classification via a novel complex-valued deep fully convolutional network,” *Remote Sens*, vol. 11, no. 22, Nov. 2019, Art. no. 2653.

[168] L. Yu, Y. Hu, X. Xie, Y. Lin, and W. Hong, “Complex-valued full convolutional neural network for SAR target classification,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no.no. 1, pp. 1752–1756, Oct. 2019.

[169] X. Wang and H. Wang. “Forest height mapping using complex-valued convolutional neural network,” *IEEE Access,* vol. 7, pp. 126334–126343, Sep. 2019.

[170] L. Li, L. Ma, L. Jiao, F. Liu, Q. Sun, and J. Zhao, "Complex contourlet-CNN for polarimetric SAR image classification," *Pattern Recognit.*, vol. 100, 2020, Art. no. 107110.

[171] H. Mu, Y. Zhang, C. Ding, Y. Jiang, M. H. Er, and A. C. Kot, "DeepImaging: A ground moving target imaging based on CNN for SAR-GMTI system," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 1, pp. 117–121, Jan. 2021.

[172] 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.

[173] J.-S. Lee, J.-H. Wen, T. L. Ainsworth, K.-S. Chen, and A. J. Chen, "Improved sigma filter for speckle filtering of SAR imagery," *IEEE Trans. Geosci. Remote Sens.,* vol. 47, no. 1, pp. 202–213, Jan. 2009.

[174] K. Du, Y. Deng, R. Wang, T. Zhao, and N. Li, "SAR ATR based on displacement-and rotation-insensitive CNN," *Remote Sens. Lett.*, vol. 7, no. 9, pp. 895–904, 2016.

[175] 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.

[176] 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.

[177] O. Kechagias-Stamatis and N. Aouf, "Automatic target recognition on synthetic aperture radar imagery: A survey," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 3, pp. 56–81, Mar. 2021.

[178] C. Jiang and Y. Zhou, "Hierarchical fusion of convolutional neural networks and attributed scattering centers with application to robust SAR ATR," *Remote Sens.*, vol. 10, no. 6, 2018, Art. no. 819.

[179] K. Wang, G. Zhang, Y. Leng, and H. Leung, "Synthetic aperture radar image generation with deep generative models," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 912–916, Jun. 2019.

[180] I. Tolstikhin, O. Bousquet, S. Gelly, and B. Schoelkopf, Wasserstein auto-encoders, 2018. [Online]. Available: https://arxiv.org/abs/1711.01558

[181] X. Lu, Y. Wang, and Y. Yuan, "Sparse coding from a Bayesian perspective," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 6, pp. 929–939, Jun. 2013.

[182] 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, Aug. 2019.

[183] H. Gao, S. Peng, and W. Zeng, "Recognition of targets in SAR images using joint classification of deep features fused by multi-canonical correlation analysis," *Remote Sens. Lett.*, vol. 10, no. 9, pp. 883–892, Sep. 2019.

[184] S. H. Lee and S. Choi, "Two-dimensional canonical correlation analysis," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 735–738, Oct. 2007.

[185] Y. Zhou, Y. Chen, R. Gao, J. Feng, P. Zhao, and L. Wang, "SAR target recognition via joint sparse representation of monogenic components with 2D canonical correlation analysis," *IEEE Access*, vol. 7, pp. 25815–25826, Feb. 2019.

[186] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *J. Signal Process.*, vol. 86, pp. 572–588, Apr. 2006.

[187] H. Zhang, N. 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. 2481–2495, Jul. 2011.

[188] J. Lv, "Exploiting multi-level deep features via joint sparse representation with application to SAR target recognition," *Int. J. Remote Sens.*, vol. 41, no. 1, pp. 320–338, 2020.

[189] Y. Li, Y. Chen, G. Liu, and L. Jiao, "A novel deep fully convolutional network for PolSAR image classification," *Remote Sens.*, vol. 10, no. 12, 2018, Art. no. 1984.

[190] Z. Huang, Z. Pan, and B. Lei, "What, where, and how to transfer in SAR target recognition based on deep CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 1–13, Apr. 2019.

[191] D. Malmgren-Hansen, A. Kusk, J. Dall, A. A. Nielsen, R. Engholm, and H. Skriver, "Improving SAR automatic target recognition models with transfer learning from simulated data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1484–1488, Sep. 2017.

[192] K. Wang, G. Zhang, and H. Leung, "SAR target recognition based on cross-domain and cross-task transfer learning," *IEEE Access*, vol. 7, pp. 153391–153399, Oct. 2019.

[193] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 7167–7176.

[194] 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, 2017, Art. no. 907.

[195] Z. Cui, C. Tang, Z. Cao, and S. Dang, "SAR unlabeled target recognition based on updating CNN with assistant decision," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 10, pp. 1585–1589, Oct. 2018.

[196] 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.

[197] M. Rostami, S. Kolouri, E. Eaton, and K. Kim, "Deep transfer learning for few-shot SAR image classification," *Remote Sens.*, vol. 11, no. 11, 2019, Art. no. 1374.

[198] J. Rabin, G. Peyré, J. Delon, and M. Bernot, “Wasserstein barycenter and its application to texture mixing,” in *Proc. Scale Space Variational Methods Comput. Vis.*, 2012, pp. 435–446.

[199] 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.

[200] Y. Jia *et al.*, “Caffe: Convolutional architecture for fast feature embedding,” in *Proc. ACM Int. Conf. Multimedia*, 2014, pp. 675–678.

[201] Z. Wang, L. Du, J. Mao, B. Liu, and D. Yang, “SAR target detection based on SSD with data augmentation and transfer learning,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 150–154, Jan. 2018.

[202] L. Ferro-Famil, A. Reigber, E. Pottier, and W.-M. Boerner, “Scene characterization using subaperture polarimetric SAR data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 10, pp. 2264–2276, Oct. 2003.

[203] L. E. Falqueto, J. A. Sá, R. L. Paes, and A. Passaro, “Oil rig recognition using convolutional neural network on Sentinel-1 SAR images,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1329–1333, Feb. 2019.

[204] M. Wurm, T. Stark, X. X. Zhu, M. Weigand, and H. Taubenböck, “Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks,” *ISPRS J. Photogramm. Remote Sens.*, vol. 150, pp. 59–69, Apr. 2019.

[205] I. Goodfellow *et al.*, “Generative adversarial nets,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[206] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” 2015. arXiv:1511.06434.

[207] 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.

[208] F. Gao, Y. Yang, J. Wang, J. Sun, E. Yang, and H. Zhou, “A deep convolutional generative adversarial networks (DCGANs)-based semi-supervised method for object recognition in synthetic aperture radar (SAR) images,” *Remote Sens*, vol. 10, no. 6, May 2018, Art. no. 846.

[209] A. Blum and T. Mitchell, “Combining labeled and unlabeled data with co-training,” in *Proc. 11th Ann. Conf. Comput. Learn. Theory*, 1998, pp. 92–100.

[210] 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, Jan. 2019, Art. no. 135.

[211] W. Zhang, Y. Zhu, and Q. Fu, “Semi-supervised deep transfer learning-based on adversarial feature learning for label limited SAR target recognition,” *IEEE Access*, vol. 7, pp. 152412–152420, Oct. 2019.

Amir Hosein Oveis was born in Tehran, Iran, in 1988. He is a Researcher with the Department of Information Engineering, the University of Pisa, Pisa, Italy. As a part of an artificial intelligence based research project, he was awarded a scholarship from CNIT, which is a National Inter-University Consortium for Telecommunications in Italy and joined to RaSS laboratory (Radar and Surveillance Systems) in Pisa, Italy, in November 2019. He has several years experience in radar signal processing and his current research interests include deep learning, radar signal processing, synthetic aperture radar, and ground moving target indication. He received the B.Sc. degree from K. N. Toosi University of Technology, Tehran, Iran, in 2011, the M.Sc. degree from Amir Kabir University of Technology, Tehran, Iran, in 2013, and the Ph.D. degree from K. N. Toosi University of Technology in 2018, all in electrical engineering and telecommunication field.

Elisa Giusti has been involved as a Researcher in several international projects funded by Italian ministries (Ministry of 1915 Defence, Ministry of Economic Development) and European organizations (EDA, ESA). She was a Researcher under contract with the Department of Information Engineering, the University of Pisa, Pisa, Italy, from June 2009 to November 2014. From November 2014 to December 2015, she was a Researcher under contract at the CNIT-RaSS National Laboratory. Since December 2015, she has been a permanent Researcher with the CNIT-RaSS National Laboratory, Pisa, Italy. She is a Co-Founder of a radar systems-related spin-off company, ECHOES Radar Technologies, located in Pisa. She has coauthored more than 60 papers and 3 book chapters. She is Editor of a book titled *Radar Imaging for Maritime Observation* (CRC Press), for which she was the recipient of the 2016 Outstanding Information Research Foundation Book publication award. Her research interests are mainly in the field of radar imaging, including active, passive, bistatic, multistatic, and polarimetric radar. She received the Telecommunication Engineering Laurea (cum laude) and Ph.D. degrees from the University of Pisa in 2006 and 2010, respectively.

Selenia Ghio (born in 1991) is currently a Researcher with the National Laboratory of Radar and Surveillance Systems (RaSS) of CNIT. Her major research interests are in the field of radar signal processing, SSA, objects' feature extraction for identification, time-frequency analysis, deep neural network applications in target classification, and 3-D radar imaging. During the work for her Ph.D. degree, she was hosted for 6 months at ESA\ESOC (Darmstadt, Germany) for a collaboration with the Space Debris Team. She has been involved in several project funded by Italian ministries (Ministry of Defence, Ministry of Economic Development, etc.) and European organizations (EDA). She received the bachelor's and master's (honors) degrees in telecommunication engineering from the University of Pisa, Pisa, Italy, in 2012 and 2015, respectively. In 2019, she received the Ph.D. degree in information engineering from the University of Pisa. In 2018, she was the recipient of the Amelia Earhart Fellowship.

Marco Martorella is currently an Associate Professor with the Department of Information Engineering, the University of Pisa and an external Professor with the University of Cape Town, where he lectures within the Masters in Radar and Electronic Defence. He is also Director of the CNIT's National Radar and Surveillance Systems Laboratory. He has authored more than 250 international journal and conference papers, 3 books, and 17 book chapters. He received the Laurea degree bachelor's and master's in telecommunication engineering (cum laude) and the Ph.D. degree in remote sensing, both from the University of Pisa, Pisa, Italy, in 1999 and 2003, respectively. He has presented several tutorials at international radar conferences, has lectured at NATO Lecture Series and organized international journal special issues on radar imaging topics. He is a member of the IEEE AES Radar Systems Panel, a member of the NATO SET Panel, where he sits as Chair of the Radio Frequency Technology Focus Group, and a member of the EDA Radar Captech. He has chaired several NATO research activities, including three research task groups, one exploratory team, and two specialist meetings. He was recipient of the 2008 Italy-Australia Award for Young Researchers, the 2010 Best Reviewer for the IEEE GRSL, the IEEE 2013 Fred Nathanson Memorial Radar Award, the 2016 Outstanding Information Research Foundation Book Publication Award for the book *Radar Imaging for Maritime Observation* and the 2017 NATO Set Panel Excellence Award. He is a Co-Founder of ECHOES, a radar systems-related spin-off company. His research interests mainly include the field of radar, with specific focus on radar imaging, multichannel radar and space situational awareness. He is a Fellow of the IEEE.