Engineering Applications of Artificial Intelligence 123 (2023) 106305

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Survey paper

Synthetic Aperture Radar image analysis based on deep learning: A review of a decade of research

Alicia Passah [a], Samarendra Nath Sur [b], Ajith Abraham [c], Debdatta Kandar [a,*]

[a] Department of Information Technology, North-Eastern Hill University, Shillong 793022, Meghalaya, India
[b] Department of Electronics and Communication Engineering, Sikkim Manipal Institute of Technology, Sikkim Manipal University, Majitar, Rangpo 737136, Sikkim, India
[c] Faculty of Computing and Data Science, FLAME University, Lavale, Pune 412115, Maharashtra, India

ARTICLE INFO

Keywords:
Deep learning
Remote sensing
Synthetic Aperture Radar
Target detection
Target recognition

ABSTRACT

Artificial intelligence research in the area of computer vision teaches machines to comprehend and interpret visual data. Machines can properly recognize and classify items using digital images captured by cameras and videos, deep learning models, and then respond to what they observe. Similarly, artificial intelligence has also been able to learn complex images captured by Synthetic Aperture Radar (SAR) that are widely used for various purposes but still leave room for improvements. Researchers have proposed numerous approaches in this field, from SAR target detection to SAR target recognition. This paper presents a survey on the different techniques and architectures proposed in the literature for various SAR image applications. The paper covers a survey on target detection models and target recognition models and their respective workflow to analyze the techniques involved and the performances of these models. This paper makes novel discussions, comparisons, and observations. It highlights the advantages and disadvantages of different approaches to give researchers the idea of how each technique can influence the performance for adoption in the future. The potential future directions along with hybrid models on each processing method are also highlighted based on the study.

1. Introduction

Synthetic Aperture Radar (SAR) is an imaging radar usually mounted on an unsteady platform. Being an active radar, SAR can operate in good as well as bad weather days, leading to the application of the acquired image in various fields such as disaster risk assessment, natural oil slick detection, civil and military defense (Chaturvedi, 2019). Many earth observation companies are launching satellites carrying SAR sensors; for example, Finnish company like ICEYE (ICEYE-Finland, 2019) is expecting to have around eighteen or more SAR sensors by the coming years. Also, Capella Space (Space, 2019) and other small satellite startups are planning their own SAR missions (Terrie, 2018). Therefore, with many SAR-carrying satellites coming up, enormous volumes of data pose immense challenges to archiving, processing, and analysis. The availability of SAR data does not imply their accessibility as they are not easily interpretable. Hence, researchers have applied various ways to process such images to extract meaningful information from such SAR data. However, there still lies multiple challenges in the process. The SAR modality has several characteristics that distinguish it from optical imagery, thereby challenging the interpretation and understanding of SAR images.

In recent years, methods using machine learning, particularly deep learning like convolutional neural networks, have effectively shown improvement in the performance of SAR image processing and their accuracies even exceed the conventional methods (Wang et al., 2017; Shang et al., 2018; Zhong et al., 2018; Pei et al., 2018; Liu et al., 2019a; Zeng et al., 2019; Huang et al., 2018; Küçük et al., 2016; Ndikumana et al., 2018). However, a severe drawback of SAR image is that it is polluted with speckle noise, making SAR image interpretation erroneous, even for post processing. Therefore, quite a number of methods are being proposed in the literature for improving the interpretation of SAR images to make it pertinent to real-world applications such as detection of forest fires, classification of sea and ice in water bodies, detection of oil slicks, reduction of disasters, detection of illegal mining, detection of changes in geographical areas, military applications, inspection of oceans, maritime detections and many more (Shang et al., 2014; Xiang et al., 2018; Biondi, 2019; Solberg et al., 1999; Mercier and Girard-Ardhuin, 2005; Solberg, 2012; Collard et al., 2005; Van Wimersma Greidanus, 2008). Despite the ability to function day and night and the capability to generate images irrespective of weather status, several issues need to be focused on using SAR images for such applications. Issues include improper recognition of targets due

* Corresponding author.
E-mail addresses: alicia.passah@gmail.com (A. Passah), samar.sur@gmail.com (S.N. Sur), ajith.abraham@ieee.org (A. Abraham), kdebdatta@gmail.com (D. Kandar).

https://doi.org/10.1016/j.engappai.2023.106305
Received 8 March 2022; Received in revised form 10 December 2022; Accepted 6 April 2023
Available online 17 April 2023

Table 1
Comparison table of the different survey papers related to SAR image target detection and recognition.

Paper	Pre processing issues	Detection			Recognition			Research issues challenges
		Survey covered	Architecture explained	Future discussion	Survey covered	Architecture explained	Future discussion	
El-Darymli et al. (2013)	✓	✓	×	✓	×	×	×	✓
El-Darymli et al. (2016)	✓	×	×	×	✓	✓	✓	✓
Zhu et al. (2017)	✓	✓	✓	✓	✓	✓	✓	✓
Hachicha and Chaabane (2014)	✓	✓	×	×	×	×	×	✓
Parikh et al. (2020)	✓	×	×	×	✓	×	×	✓
Our work	✓	✓	✓	✓	✓	✓	✓	✓

to the inability to determine features related to a specific target type properly. This is mainly due to the characteristics of SAR systems that generate very grainy images resulting from interference of signals that reached the SAR receiver (Tomiyasu, 1978; Meyer et al., 2013; Shimada et al., 2009; Rosen et al., 2008). Another issue that arises from this is removing the unwanted features called noise from SAR images to enhance image quality. This noise removal process is quite challenging as it may also remove the most critical features, further resulting in an imprecise outcome. Moreover, some detection-based applications faced issues such as detecting a non-target as a target and vice versa due to the inability to differentiate between the two.

Various works have been pondered in the literature, considering several SAR image interpretation issues while still leaving room for improvement. It may be noted that all techniques proposed in the literature are unique in their way. Their uniqueness needs to be explored and understood to enable researchers to adapt and enhance the same for future applications in the field of SAR image processing. Various effective surveys have been done in the literature regarding the processing of SAR images. Table 1 highlights the comparison between our paper and the existing survey papers related to SAR imaging. Khalid et al. have highlighted numerous research activities in the field of SAR target detection, including future perspectives (El-Darymli et al., 2013). On a different work, Khalid et al. have also assessed several state-of-the-art works related to SAR target recognition (El-Darymli et al., 2016). Zhu et al. have comprehensively reviewed numerous works, including the challenges faced when using deep learning in SAR image processing (Zhu et al., 2017). The paper includes a survey on detection, and recognition of SAR images. Therefore, most of the survey works cover only a part of the entire processing stages of SAR image analysis. The deep learning architectures used by recent SAR image processing are not explored in depth. Target detection and recognition are the vital processing steps in SAR image interpretation. SAR image detection is useful for detecting targets present on the ground through the captured images for monitoring purposes by intentionally investigating within a largely well-defined area. On the other hand, a target feature must be distinguishable from the backdrop in order for detection to take place, which is challenging in the case of SAR image analysis since SAR images are distorted with background clutters and noise. In contrast, SAR target recognition is useful for predicting the types of targets and scenes in specific SAR applications. For instance, in military applications, the types of military tanks are remotely determined with the help of target recognition algorithms. The major difference between target detection and target recognition is that in target detection, the presence and position of the targets are determined irrespective of the type and model of the targets, whereas, in target recognition, the description and type of the target are of significant concern. A study incorporating the works related to both these processing stages in a single survey paper would help and ease researchers and those new to SAR imaging, enabling them to be at par with recent advances in this area.

This has motivated us to conduct a survey on the different state-of-the-art techniques related to SAR image interpretation, including target detection and target recognition, by going into depth the techniques and architectures involved, particularly deep learning. Following are the main contributions of the paper:

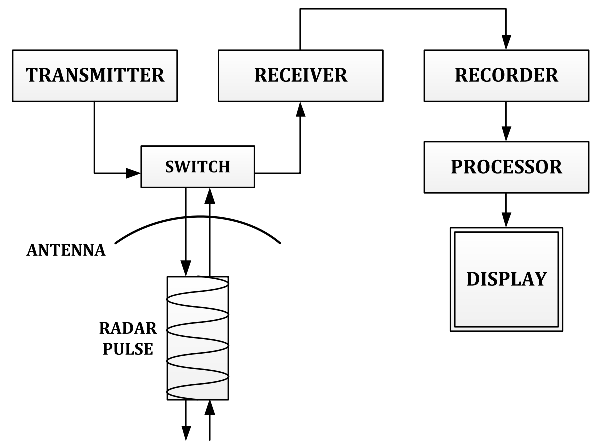

Fig. 1. The basic block diagram of a radar.

1. State-of-the-art techniques on SAR image processing comprising detection and recognition are discussed.
2. The different architectures and parameter settings of each technique are also examined.
3. The significant advantages and disadvantages of the existing techniques are also discussed to give researchers insights into each method.
4. Necessary observations on each processing technique are discussed based on the study.
5. Future approaches to improving the existing techniques for performance enhancement are also highlighted, along with few potential models.

This survey paper considers the performances of various processing approaches under SAR images, and a comparison of these methods has also been made. These genuine comparisons will help researchers understand and solve existing issues related to SAR image processing. The rest of the paper is organized as follows. Section 2 describes the general background study, which comprises SAR image processing, a brief about deep learning, and convolutional neural networks. The study and observations on various works associated with SAR target detection and SAR target recognition are discussed in Sections 3 and 4, respectively. Section 5 discusses various issues and challenges in SAR image processing. The future approach associated with each SAR image processing method is discussed in Section 6, followed by a conclusion in Section 7. Fig. 2 shows the organizational summary and framework of the paper.

2. Background

This section discusses the background study by introducing the working of SAR and how images are acquired from SAR. We also include the definition of various SAR image processing methods, followed by an explanation of deep learning and convolutional neural network.

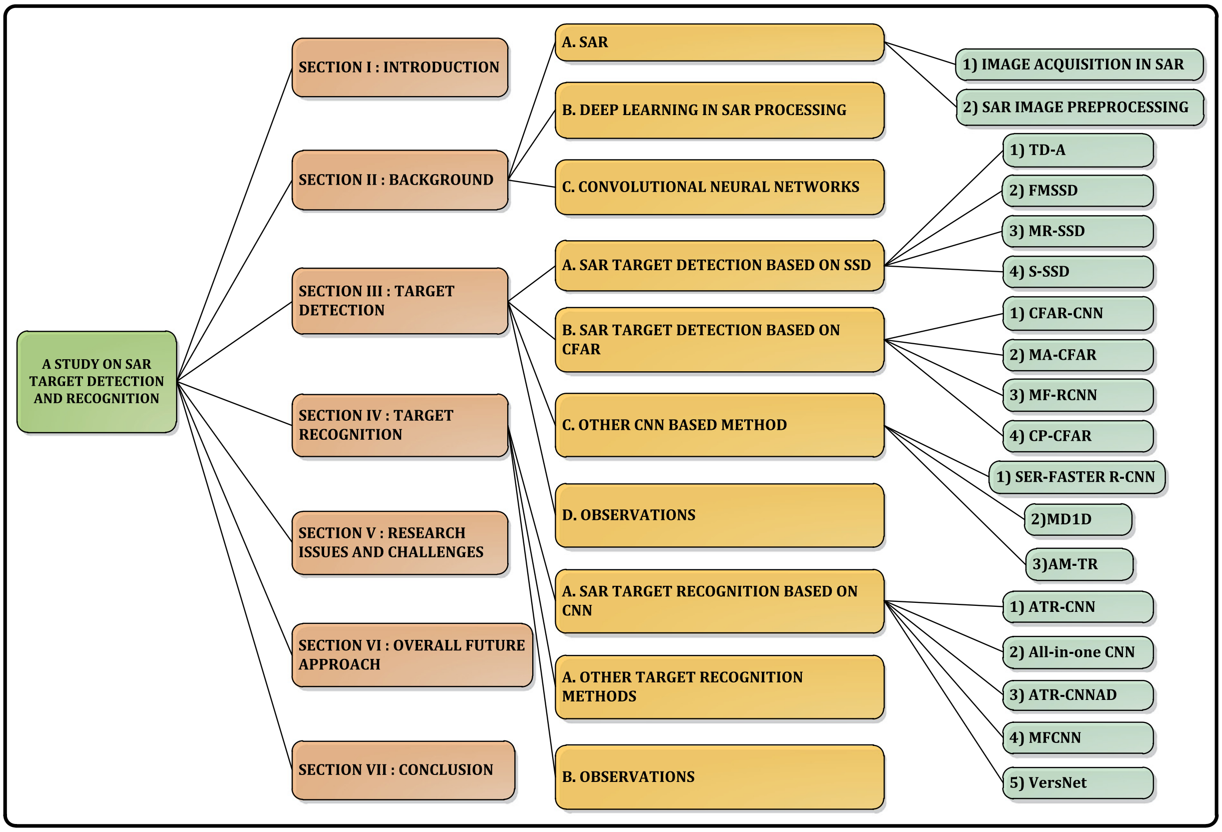

Fig. 2. Organization summary of the paper.

2.1. SAR

SAR is a sensing radar usually attached to an aircraft, spacecraft and even in guided airborne ranged missiles for sensing the different objects and scenes present on the ground. The primary characteristics of a self-illuminating SAR radar include the day and night imaging capabilities and the ability to function irrespective of the storm, rain, or mists (Ulaby et al., 1981; Hänsch et al., 2016; Tomiyasu, 1978). The characteristics of a SAR radar have made the processing of the generated images worthwhile for various applications (Chaturvedi, 2019). Processing SAR images involves recording the reflected electromagnetic waves of different ranges and azimuth, compressing the signal, and generating the image. Fig. 1 shows the basic block diagram of a typical radar (Chan and Koo, 2008). The SAR signal that is reflected back from the ground gets collected at the antenna and retrieved by the receiver for subsequent processing by the signal processing unit to obtain the SAR image (Tomiyasu, 1978). It may be mentioned that the images acquired by SAR are highly corrupted with undesirable features called noise, which reasons the degraded quality of SAR images, resulting in the loss of relevant details of the image. Fig. 3 shows how the relevant information that could contribute to a better understanding of the image for subsequent analysis, is being corrupted even after applying denoising to the noisy image. It is observed from the zoomed image shown in Fig. 3(d), that the image contains noticeable artifacts and most of the image details have been washed out. Therefore, SAR images are difficult to understand as can be seen in Fig. 3(b), and processing them is equally challenging.

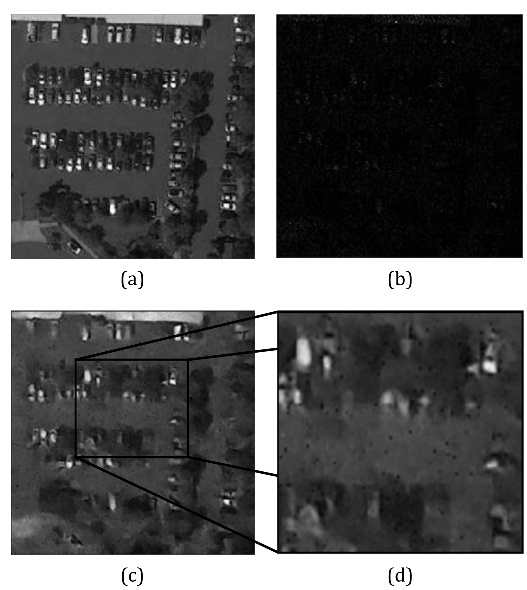

Fig. 3. Degradation of image quality after denoising. (a) Clean image, (b) Noisy Image, (c) Denoised image, (d) A portion of the denoised image when zoomed (Image: NWPU-RESISC Cheng et al., 2017).

2.1.1. Image acquisition in SAR

The electromagnetic waves that are transmitted from the SAR radar interact with the surface of the earth, and only a portion of the waves is back-scattered to the receiving antenna whereby the returned signals for the time T from position A to D as shown in Fig. 4 are stored in the SAR processor as amplitudes and phases for subsequent preprocessing. Wolff (2019), Inc. (1997) and for Remote Imaging Sensing and (CRISP). The data generated from SAR is raw data that needs high signal processing techniques to convert these raw data into interpretable images.

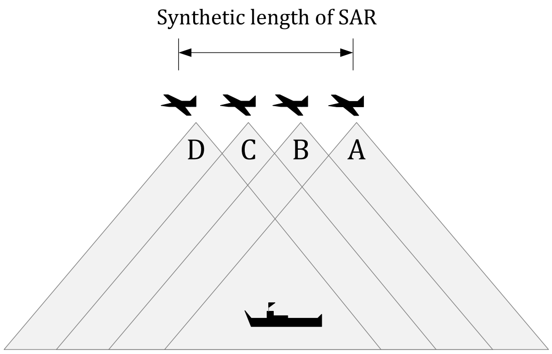

Fig. 4. The flight path of a SAR mounted platform (Wolff, 2019).

2.1.2. SAR image pre-processing

To enable SAR images to be applicable in several applications, they need to undergo different preprocessing stages to achieve the desired outcome. One of the most vital preprocessing stages is called denoising, or the noise removal stage that involves the removal of the unwanted granular features contained in SAR images. The noise that is present in SAR images is multiplicative in nature and is known to be one of the most catastrophic type of noise (Baltierra et al., 2022). Therefore denoising SAR images is one of the challenging pre-processing task. Image denoising can be achieved using different approaches, for example, by using statistical re-evaluation of the pixels that formed the image in an optimized manner or by automatically reconstructing the desired image using deep learning models. Various state-of-the-art works have been proposed in the literature for SAR image denoising that has helped improve the image quality (Passah et al., 2021; Wang et al., 2017; Zhang et al., 2018; Lattari et al., 2019). It may be mentioned that noise reduction algorithms may distort the signal to some degree (Singh et al., 2021). Therefore, radiometric correction is also one of the important pre-processing steps for SAR images as this step aims at precisely estimating the reflectance of an environment by correcting the distortions caused by the antenna, positioning glitch, and several other characteristics of the electronic components. It should also be noted that a SAR image can include several swaths or sectors. Therefore de-bursting is another preprocessing step wherein each swath is merged into a single image, thereby reducing redundant lines and black separation lines on the SAR image. Segmentation is another significant preprocessing stage of SAR image analysis. Segmentation involves the concentration of the image around a particular area or target by highlighting the outlines and separating the image group-wise based on their characteristics. Segmentation of targets in SAR images is an open field of research since segmentation models misinterpret noise features as part of the targets, resulting in erroneous outcomes. Numerous works available in the literature have been proposed to address the aforementioned issues (Feng et al., 2022; Sun et al., 2021; Yu et al., 2021).

2.2. Deep learning in SAR image processing

This section discusses one of the popular techniques under machine learning called deep learning. Deep learning is a sub-field of machine learning that deals with algorithms influenced by the structure and capacity of the mind called artificial neural networks (Brownlee, 2016). In other words, it reflects the working of human brains. Deep learning uses numerous layers to extract higher-level features from raw inputs progressively. For instance, in image processing, lower layers may distinguish edges, while higher layers may recognize human-significant things, for example, digits or letters or faces (Gu et al., 2018). The *deep* in *deep learning* refers to the number of layers through which the data is transformed. In most machine learning algorithms, the features are handcrafted, except for deep networks in which this tedious task is not required as it appears to be automatic and within the network. The deeper the network, the more data it can train on, and the better accuracy. The ability of deep networks to learn and process from extensive unlabeled data gave it a distinct advantage over previous algorithms (LeCun et al., 2015). Machine learning has recently been used in SAR image processing; particularly, deep networks are becoming an apple of the eye for various remote sensing applications (Zhong et al., 2018; Shang et al., 2018; Chen et al., 2016; Wang et al., 2018a). The applications of deep networks in SAR have shown considerable improvements in the results for target detection (Redmon et al., 2016), target recognition (Furukawa, 2018; Pei et al., 2018), denoising (Wang et al., 2017, 2018b), and other processing techniques. On the other hand, SAR images are also highly corrupted by speckle noises like any other images produced by coherent imagery systems. Speckle noise refers to the type of noise that is multiplicative in nature resulting from the impact of the surroundings on the imaging sensor during image acquisition. This fact deteriorates the applicability of SAR images in future fields of remote sensing. Also, this drawback of SAR images causes severe complications in automatic image interpretation. Therefore, researchers are still venturing into different ways to boost the performance of SAR image processing. Several existing works based on SAR image interpretation are discussed in the later sections. Next, we present one of the most used deep learning models proposed in the literature for SAR image analysis. However, speckle noise refers to the type of noise that is multiplicative in nature resulting from the impact of the surroundings on the imaging sensor during image acquisition.

2.3. Convolutional neural networks

Inspired by the structure of the visual cortex of a cat (Hubel and Wiesel, 1968; Fukushima, 1980), a Convolutional Neural Network, also known as CNN or ConvNet, is a class of neural networks that is one of the most popular types of deep learning algorithms used in image processing techniques. When an input image is fed to this network, learnable weights and biases are assigned to different forms of objects present in the image, resulting in the ability to distinguish one object from another. The CNN requires less pre-processing in comparison to other prior algorithms (Saha, 2018). CNN can automatically learn filters or characteristics with enough training, unlike primitive methods where filters are hand-engineered.

The CNN consists typically of the input, convolutional, subsampling, fully connected, and the output layer (Albawi et al., 2017). It is a two-stage process where the first stage comprising the convolutional and subsampling layers, is used mainly to extract relevant features from given inputs. In contrast, the second stage comprising the multi-layer perceptron or fully connected layer is a classifier. The CNN used in various applications consists of these two stages only. However, their configurations differ from one application to another, resulting in different performance accuracy. The proficiency of CNNs has enabled researchers to employ them in various SAR image analysis, whereby results therein have shown noticeable progress. This review paper mainly focuses on works based on CNNs and how they are being incorporated in various SAR image interpretations.

Few famous CNN architectures that are widely being inherited in various computer vision applications includes the LeNet-5 (Lecun et al., 1998), AlexNet (Krizhevsky et al., 2012), ZFNet (Zeiler and Fergus, 2014), InceptionVI (Szegedy et al., 2015), VGGNet (Simonyan and Zisserman, 2015), ResNet (He et al., 2016), written in the order of decreasing error rates and increasing number of layers as per the ILSVRC (Russakovsky et al., 2015). The aforementioned CNN models have shown substantial abilities in processing optical and medical images (Vogado et al., 2018; Yuan et al., 2019; Liu and Wu, 2016; Qin et al., 2019; Gong et al., 2019; Iqbal et al., 2021; Xu et al., 2021; Tang et al., 2021; Lawrence et al., 1997; Cheng et al., 2022). This has encouraged researchers to explore the models in various other research areas. As such, merging and incorporating these models in SAR applications is also limited, leaving space for research in this field. Considering the current state-of-the-art in the field of SAR image processing and its application, our paper aims to review the techniques, architectures, and models adopted by all the existing works to benefit even the fresh researchers that are new to SAR image analysis. However, the primary focus of this review is to highlight the ideas of different works related to the interpretation of Synthetic Aperture Radar(SAR) images, particularly the deep learning approaches. The models, architectures, and each work's respective workflow are discussed in the later section.

2.4. Evaluation metrics used for illustrating results of various works

The evaluation measures that are utilized to illustrate the outcomes of the various studies covered in this paper are discussed in this section. The widely used metrics such as the precision, recall, F1-score, accuracy, figure of merit (FOM), are discussed below.

- *Precision*: The precision metric P is utilized to calculate the overall number of true class predictions, out of all the predictions generated by the model to be that class. The precision calculation formula is written as.

$$P = \frac{T_p}{T_p + F_p} \tag{1}$$

  where, T_p is the number of correctly predicted classes, F_p is the number of other classes predicted to be that class.
- *Recall*: Recall R quantifies the proportion of all correctly predicted classes to all instances of that class in the dataset, and it is calculated as.

$$R = \frac{T_p}{T_p + F_n} \tag{2}$$

  where, F_n is the number of positive classes that are incorrectly being predicted as negative.
- *F1-score*: F1-score, uses harmonic mean to get the average of values by combining precision and recall. The F1-score is more suited to determining the average of ratios and can be expressed as.

$$F1\text{-}score = 2 * \frac{P * R}{P + R} \tag{3}$$

- *Accuracy*: The accuracy measure represents the proportion of accurate predictions made by the model out of all possible predictions and can be expressed as.

$$Accuracy = \frac{T_n + T_p}{T_n + F_p + T_p + F_n} \tag{4}$$

  where, T_n is the number of correctly predicted negative classes.
- *Figure of Merit*: The Figure of Merit (FOM), also known as the performance measure of interest for the purpose of target detection, is the percentage of target patterns that are detected when the false alarm rate is within acceptable bounds.

$$FOM = \frac{T_p}{F_p + T_n} * 100\% \tag{5}$$

2.5. Open source benchmarks for SAR image analysis

This subsection highlights some of the available benchmarks for SAR image analysis that are freely accessible for research purposes. Unlike optical images, SAR images are tough to interpret and annotate because of their highly complex nature. The SAR datasets are not readily available to the research community. Nevertheless, few datasets have been made available, which has benefited researchers in the field of SAR image analysis. A summary of the existing benchmarks is presented in Table 2.

3. Target detection

In this section, we focus on the detection of targets present in SAR images by reviewing the literature that incorporated CNN in their architectures with an aim to highlight the models and techniques adopted therein. Target detection is the ability to detect targets from a mixture of targets and non-targets present in an image (Smith, 1997; Changlin and Xuelian, 2007). The targets, however, depend from one application to another. In the case of SAR target detection, common applications include detecting oil spills, detecting illegal ship transportation, detecting targets on the ground in the presence of clutter, etc. However, with the presence of speckles in SAR imagery, target detection becomes a challenging task (Baltierra et al., 2022). Therefore, detecting targets rapidly and precisely from SAR images has also become a research topic at present (Steiniger et al., 2022). Fig. 5 depicts the concept of target detection using bounding boxes in a SAR image.

Researchers have proposed various algorithms and techniques to improve detection performance in SAR images. For instance, the authors in Wang et al. (2012) use the magnitude of a selected term in the coherence matrix to solve the issue of detecting non-reflecting targets in SAR image but degrades the detection of urban structures. Similarly, the authors in Sugimoto et al. (2013) suggested the use of model-based decomposition to block sea surface scattering from the sea target scattering. In a different work (Agrawal et al., 2015), the authors have first segmented the targets in SAR images followed by several processing stages such as normalization of the target position, histogram equalization, and dilation. Extraction of features was then carried out using scale-invariant feature extraction, and targets were then detected using these features. A similar work (Yu et al., 2016) was also proposed for SAR target detection that first segmented images to produce superpixels. These superpixels are helpful for clutter distribution parameters estimation. The two-parameter CFAR followed by superpixel clustering is then used to detect targets. The authors in Gao et al. (2017) have used shadow proposal, saliency analysis, and single class SVM to improve target detection in a cluttered SAR environment.

In the SAR target detection task, the samples used for training are usually SAR images with multiple targets and complicated background of an enormous scene. The most commonly used SAR detection mechanism is the CFAR (Gao et al., 2009). The popular two-parameter CFAR works by assuming the underlying clutter to agree on the Gaussian distribution. It is shown that this method performs efficiently for non-obstruction scenes, but the performance degrades on encountering complex scenes like SAR scenes (Wang et al., 2019).

Therefore, considering the need to overcome such issues, various works have been projected in the literature, many of which use deep learning networks such as CNN based on SSD and CFAR. Still, another issue arises in using CNNs for target detection in SAR images. Issues include the non-availability of a large number of SAR datasets for training a CNN leading to a few incorrect detection results. Hence the use of most recent techniques like CNN becomes challenging for SAR target detection, and several works that aim at solving such issues are discussed in the following subsections. Table 3 shows a comparison summary of the different target detection works.

3.1. SAR target detection based on SSD

The single-shot multibox detector (SSD) (Liu et al., 2016) is a real-time object detection model that excludes the region proposal network and uses multiscale convolutional bounding boxes connected to various feature maps over the network, allowing desirable shapes of output boxes. SSD employs lower resolution layers for the detection of large-scale objects. The SSD is known to perform better as compared to YOLO (Redmon et al., 2016). Other than detection capabilities in optical images, SSD could also attain better accuracy in SAR images. Hence various works in the literature have incorporated SSD in SAR target detection tasks, of which few are highlighted below.

Table 2
A summary of the existing benchmarks for SAR image analysis.

Dataset	Description	Coverage	No. of images	Labels	Multiclass	Links
SN6 MSAW (Shermeyer et al., 2020)	It is a multi-sensor all-weather mapping dataset that features a combination of Capella Space SAR imagery and Maxar World View2 electro-optical imagery	120 km^2 area of Rotterdam, Netherlands	48,000	Buildings, vehicles, boats, etc.	Yes	https://mediatum.ub.tum.de/1474000
MSTAR (Ross et al., 1998)	It is a dataset comprising of military vehicles and tanks belonging to ten different categories. Clutter images with no targets are also available in this dataset.	Military vehicles and clutter	17,658	Vehicles namely bulldozer, tanks, etc.	Yes	https://www.sdms.afrl.af.mil/index.php?collection=mstar
PolSF (Liu et al., 2019b)	It is also referred to as PolSAR images of the San Francisco region, and was obtained from various sensors, including ALOS-1, ALOS-2, Sentinel-1, Sentinel-2, GF-3, etc.	San Francisco region	3000	Annotations of land, sea, buildings, etc.	Yes	https://github.com/liuxuvip/PolSF
OpenSARUrban (Zhao et al., 2020)	It is captured by Sentinel-1.	High and low rise buildings, villas, industrial and vegetation areas covering 21 cities in China	33,358	Annotations of general residentials, high-rise buildings, villas, airports, etc.	Yes	https://ieee-dataport.org/documents/opensarurban
So2Sat LCZ42 (Zhu et al., 2020)	It is a dataset consisting of corresponding SAR and multispectral optical data captured by Sentinel-1 and Sentinel-2 satellites along with a corresponding label of local climate zones (LCZ)	42 urban cities	400,673	Annotations of local climate zones	Yes	https://mediatum.ub.tum.de/1483140
SARptical (Wang et al., 2017b)	It is a dataset comprising of very high-resolution SAR data extracted from TerraSAR-X images.	Dense urban agglomeration.	10,108	Corresponding optical representations of SAR images	–	https://syncandshare.lrz.de/dl/figixjRV9idETzPgG689dGB/SARptical_data.zip
OpenSARShip 2.0 (Li et al., 2017a)	The dataset consists of different ship image patches obtained from Sentinel-1.	Ship images.	34,528	Labels include ship details such as cargo, tanker, vessel, etc.	Yes	https://opensar.sjtu.edu.cn/
SAR-Ship-Dataset (Wang et al., 2019)	The dataset comprised of ship images and was created by involving the 108 Sentinel-1 images and 102 GF-3 images.	Ship images.	39,729	Ship details and types such as cargo, tanker, windmill, etc.	Yes	https://github.com/CAESAR-Radi/SAR-Ship-Dataset
SEN1-2 (Schmitt et al., 2018)	The SEN1-2 dataset contains corresponding pairs of Sentinel-1 and Sentinel-2 images covering seasons of different areas.	Different regions in all seasons	2, 82, 384	Corresponding optical images of SAR data	–	https://mediatum.ub.tum.de/1436631
SEN12MS (Schmitt et al., 2019)	The SEN12MS dataset is based on Sentinel-1 data, Sentinel-2 images, and MODIS (Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer) land cover.	Inhabited continents in all seasons	180,662	Geoinformation	Yes	https://mediatum.ub.tum.de/1474000

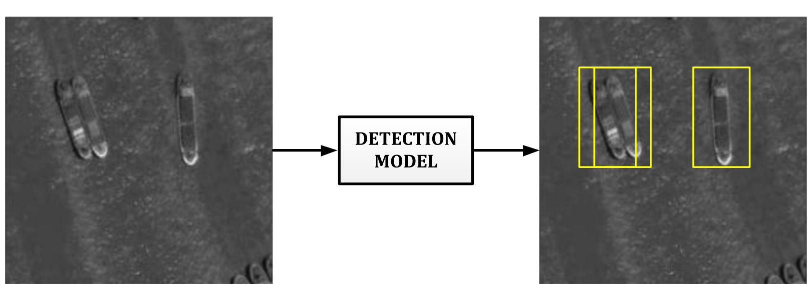

Fig. 5. Target detection concept: Detection model outputs detection results with the help of bounding boxes (Image: NWPU-RESISC Cheng et al., 2017).

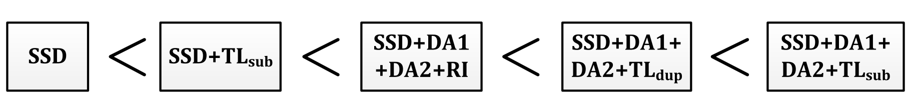

Fig. 6. F1 scores of different SSD based models in ascending order. RD: Random Initialization, DA1: Data Augmentation I, DA2: Data Augmentation II, TL: Transfer Learning, sub: Subimages as channels, dup: Duplicate images as channels.

Table 3
Summarizing the concept of different SAR target detection networks.

Paper	Dataset	Transfer learning	Augmentation	Base network	Pretrained network	Novelty	Architecture	Advantages	Disadvantages	Future
TD-A (Wang et al., 2019)	MSTAR	Yes	Yes	SSD +CNN	VggNet	Own augmentation methods	SSD +8ConvLayers	–	–	ResNet or GoogleNet can be explored to replace the pretrained network
FMSSD (Wang et al., 2020)	NWPU +DOTA	No	No	SSD	VggNet	Uses feature merged SSD	Atrous spatial feature pyramid with 5Conv layers +multiple atrous rates	Small objects determined due to new area-weighted loss function	Few missed detections	Develop feature fusion networks
MR-SSD (Ma et al., 2018)	Marine targets	Parameters adopted from SSD	Yes	VggNet	SSD	FC layers replaced by convolutional layers	Vgg16 layers+ SSD based layers	Could detect targets other than ships such as windmills etc.	Could not detect few targets	–
S-SSD (Du et al., 2020)	miniSAR +MSTAR	Yes	No	SSD	VggNet	Merging of saliency maps along with original SAR images	6 Dense networks in each subnetwork	Could suppress clutter	Results in false alarms, speed is slower than SSD	Explore parallel computing and optimization methods to improve speed and reduce false alarms
CFAR-CNN (Wang et al., 2018)	Maritime SAR images	No	Yes	CFAR +CNN	–	Two stage CFAR for local detection and global detection respectively	CFAR+ 6Conv Layer	Lower detection time with 82% accuracy	–	–
MF-RCNN (Kang et al., 2017)	Sentinel-1	Yes	No	Faster R-CNN +CFAR	VggNet	Use of CFAR to consider relevant features from false alarms	Vgg16+ FasterR-CNN+ CFAR	Could detect multiscale ships	Increase in false alarms	To perform data augmentation and explore other pretrained networks
CP-CFAR (Cui et al., 2018)	MSTAR	No	No	CFAR	No	Uses both horizontal and vertical sobel operators in CNN	Conv+ Pool+ CFAR+ median filtering	Outperforms CA-CFAR (Gao et al., 2009) and two-parameter CFAR (Ai et al., 2018)	–	–
SER Faster R-CNN (Lin et al., 2019)	Sentinel-1	Yes	No	Faster R-CNN	VggNet	Uses squeeze and excitation mechanism	5ConvLayer +SER block	SE mechanism played a role in improving the detection accuracy	Complex model	Simplify the architecture and use other pretrained networks

ConvLayers = Convolutional Layers.

3.1.1. TD-A

The work in TD-A (Target Detection with Augmentation) (Wang et al., 2019) aims at detecting vehicles from miniSAR images converted to three channeled using sub-aperture decomposition (Ferro-Famil et al., 2003). TD-A employs CNN-based SSD (Liu et al., 2016) as its base network whose parameters are initialized based on the learned VggNet model trained on the ImageNet data (Russakovsky et al., 2015). Different augmentation methods, namely DA1 and DA2, were also performed in TD-A to increase the training data. The steps involved in performing augmentations DA1 and DA2 are shown in Table 4. The images are then fed to eight convolutional layers pipelined before the finetuned SSD network. A convolutional predictor (Liu et al., 2016) was then applied to the extracted features and used Non-Maximum Suppression (NMS) (Ren et al., 2015; Girshick, 2015; Girshick et al., 2014) to remove redundant bounding boxes. TD-A outperforms the CFAR (Gao et al., 2009) and Faster RCNN (Ren et al., 2015). When including the transfer learning step and excluding the data augmentation process, or vice versa, this method still outperforms the randomly initialized SSD model. Fig. 6 shows the performance observations in ascending order which gave rise to the proposed method TD-A. The known network that emerged after VggNet called the ResNet can be explored and replaced with the VggNet used in the pretraining part of this work to improve accuracy. A possible increase in performance might be seen.

Table 4
Augmentation techniques used in TD-A (Wang et al., 2019).

Augmentation	Steps involved
DA1	(1) Sub-images are extracted from mini-SAR images. (2) Gaussian noise is added to sub-images. (3) Performs median filtering on sub-images. (4) Rotate the images by 270°.
DA2	(1) MSTAR images are padded with background pixels. (2) Images are resized to the size of miniSAR sub-images.

3.1.2. FMSSD

Another state-of-the-art work that uses SSD and pre-trained VggNet for SAR target detection is named the FMSSD (Wang et al., 2020) comprising a feature-merged SSD network that blends multiple and alike feature scale maps. FMSSD could detect objects of different classes. An area-weighted loss function was also developed to give equal importance to smaller targets in SAR images. FMSSD outperforms RICNN (Cheng et al., 2016), SSD (Liu et al., 2016) and DSSD (Fu et al., 2017). The latest trend on aggregating multiple features can be explored for improved detection models.

3.1.3. MR-SSD

This work aimed at classifying marine targets at patch level (Ma et al., 2018). An end-to-end detection technique was developed using SSD (Liu et al., 2016) with multi-resolution inputs. This detection technique is named MR-SSD, an acronym for Multi-Resolution Single Shot multibox Detector. Focusing on the detection part, which is of concern in this section, the MR-SSD is the first work that detects other marine targets such as windmills, iron towers, and platforms along with ships. MR-SSD first transformed the single channeled GF3 SAR images into a three-channeled RBG image using 2D Fourier transform and applying a low pass filter with different values of λ shown in Table 5, after which

Table 5
Hyper-parameters used in MR-SSD (Ma et al., 2018).

Parameters	Values
λ	$0.5, 0.25$
Learning rate	0.0001
Weight decay	0.005

Table 6
Comparison of average precisions with different algorithms that uses same dataset (Ma et al., 2018).

Target type	Faster-RCNN	SSD	MR-SSD
Cargo	89.47	89.37	89.77
Container	79.78	87.08	88.69
Tower	68.79	74.55	80.07
Platform	89.61	89.96	90.43
Tanker	86.70	86.46	87.28
Windmill	78.19	86.34	88.04
MaP	82.09	85.62	87.38

Table 7
Comparison results on various SAR image detection methods: Faster-RCNN (Ren et al., 2015), SSD (Liu et al., 2016), MR-SSD (Ma et al., 2018), TD-A (Wang et al., 2019), S-SSD (Du et al., 2020).

Method	Faster-RCNN	SSD	MR-SSD	TD-A	S-SSD
Recall (%)	92.97	94.53	95.31	93.45	96.00
Precision (%)	93.70	84.62	93.85	87.01	91.00
F1 (%)	93.33	89.30	94.57	90.11	94.00

the images are transformed to the time domain by inverse Fourier transform. The three channeled inputs are then used for training the MR-SSD model. This model was developed using VGGNet (Simonyan and Zisserman, 2014) architecture with fully connected layers replaced by convolutional layers having parameters adopted from the SSD. The MR-SSD was trained using the hyper-parameters shown in Table 5. Detection on large-scale SAR images was also done, where large-scale SAR image first undergoes a land masking method, after which the land parts are removed using the level set method (Li et al., 2010). Then small overlapping patches are extracted and created from them to feed the trained MR-SSD. The output is then mapped to a large image using coordinates. MR-SSD achieved highest mean average precision (Zhu, 2004) compared to Faster R-CNN (Ren et al., 2015) and SSD (Liu et al., 2016) as shown in Tables 6 and 7.

3.1.4. S-SSD

A similar work that is SSD-based is the S-SSD, acronym for saliency-guided single shot multibox detector (Du et al., 2020). This work appends the saliency information with the help of modified Itti's method (Itti et al., 1998) to its network, suggesting the model where to focus, resulting in advanced representation in complex scenes. In S-SSD, a dense network is used instead of SSD based simple network. S-SSD outperforms SSD (Liu et al., 2016), CFAR (Gao et al., 2009) and TD-A (Wang et al., 2019) by achieving an F1-score of 94. It may be mentioned that the S-SSD also uses truncated VggNet as part of the network. The MSTAR dataset was used in the implementation of the S-SSD.

3.2. SAR target detection based on CFAR

Constant False Alarm Rate (CFAR) (Gao et al., 2009) is a detection algorithm that works with the help of sliding windows and is used adaptively in radar systems in order to detect targets from SAR data contaminated with noise. The pixel-based CFAR is one of the most widely used detectors in SAR image analysis, but the time complexity is on the higher side because it is pixel-based. Researchers have made efforts to improve the overall performance of the CFAR detector, which is discussed in the following subsections.

Table 8
Hyper-parameters used in CFAR-CNN (Wang et al., 2018).

Parameters	Values
Learning rate	1
Batch size	2
Epoch	1

Table 9
Comparison results of multithreaded and multilevel CFAR (Wang et al., 2017a) and CFAR-CNN (Wang et al., 2018).

Method	#TOI	DT	MT	FA	RT	FOM	OD %	% DFA
MTML-CFAR	17	15	2	3	76	0.75	88.23%	20%
CFAR-CNN	17	14	3	0	59	0.82	82.35%	0%

*MTML-CFAR - Multi-threaded and Multilevel CFAR; #TOI - Total target of interest.
DT - Detected Target; MT - Missed Targets; FA - False Alarm, RT - Run Time (in secs).
FOM - Figure of Merit; OD- Overall Detection; DFA - Detected False Alarm.

3.2.1. CFAR-CNN

CFAR-CNN (Wang et al., 2018) aims at detecting ships from SAR images by incorporating CFAR (Gao et al., 2009) along with CNN. The ship dataset rotated by 60°, six times, used in CFAR-CNN consists of both targets and clutter alone. In CFAR-CNN, CFAR is used twice: once for global detection and next for local detection. A histogram probability distribution function is first obtained from the clutter points statistics of the entire SAR image during global detection. After setting the first level CFAR values, a threshold is obtained for comparison with each pixel on the image to differentiate between ship and clutter. This results in an index matrix from which a detection region with a target point as its center is detected and extracted using CFAR local detection and acts as inputs to the CNN trained using the hyper-parameters shown in Table 8, in order to determine whether the detected region is a ship or a clutter. The CNN used in this method is a six-layered network comprising three convolutional operations, two pooling layers, and one fully connected layer. The outputs of the CNN model are then merged into a single image, and ships are detected according to the areas of the four connected regions. The performance of CFAR-CNN outperforms the multi-threaded and multi-level CFAR (Wang et al., 2017a). The comparisons are shown in Table 9. Comparing the detection time, which is crucial in the case of target detection, CFAR-CNN achieved lower detection time compared to Wang et al. (2017a), with the ability to detect 82% of the targets. Therefore, this work, being a method that achieved lesser detection time, the detection accuracy might further be improved if the CNN used in this method is replaced with a pre-trained network such as VggNet or ResNet.

3.2.2. MA-CFAR

MA-CFAR, which is short for Manifold Adaptation for Constant False Alarm Rate (Schwegmann et al., 2015) is another SAR detection work based on CFAR that involves converting the original scalar threshold used by CFAR into a threshold manifold (Schwegmann et al., 2015) adjusted by simulated annealing (SA) algorithm (Kirkpatrick et al., 1983). This was done to enable the CFAR to adapt to information related to different characteristics of ship targets since the experiment was conducted on ship images. MA-CFAR attained a detection accuracy of 85.1%. It may be mentioned that precise selections of thresholds even avoid the computational shortcomings ordinarily connected with SA.

3.2.3. MF-RCNN

MF-RCNN aims at detecting multi-scale ships from SAR images (Kang et al., 2017) by modifying the Faster-RCNN. Hence we refer to this work as MF-RCNN (Modified Faster-RCNN). There are several ship detection methods in the literature such as CFAR (Gao et al., 2009), YOLO (Redmon et al., 2016), SSD (Liu et al., 2016) and Faster R-CNN (Ren et al., 2015), but the multi-scale attribute of ships in SAR images degrades the detection performances due to inability to learn accurate features of multi-scale ships (Marino et al., 2015; Iervolino et al., 2015; Iervolino and Guida, 2017). Therefore, MF-RCNN tried solving this issue by modifying the Faster R-CNN with the help of CFAR. The SAR images are first imported to a VggNet (Simonyan and Zisserman, 2014) based CNN to obtain feature mapping, which is then fed to a Faster R-CNN and a Region Proposal Network (RPN) (Ren et al., 2015). The RPN results in bounding boxes with respective scores for which the top n positive object proposals are fed to the ROI pooling of the faster R-CNN. The faster R-CNN then returns the classification scores and is finally fed into the CFAR detection after undergoing Non-Maximum Suppression (NMS) to remove overlapping bounding boxes. The CFAR detector also takes another input: the refined regions generated from the output of the faster R-CNN. The bounding boxes with higher classification scores are taken to be targets, whereas those with very low scores are considered false alarms. However, the boxes with relatively low scores (0.3–0.8) are chosen by the CFAR, from which it then determined the threshold value, where values exceeding this threshold are considered as targets. Compared to Faster R-CNN, MF-RCNN (Kang et al., 2017) achieved improved detection accuracy by incorporating CFAR in the Faster R-CNN. Therefore, CFAR contributed more towards the performance improvement by considering the relevant features irrespective of the target structure. On the other hand, MF-RCNN (Kang et al., 2017) gave rise to an increase in the false alarm, unlike Faster R-CNN with lesser false alarms.

3.2.4. CP-CFAR

The Convolution and Pooling CFAR (CP-CFAR) (Cui et al., 2018) is another CFAR based SAR target detection approach that uses convolutional layers and pooling layers for SAR target detection through parallel GPU processing in SAR images. CP-CFAR uses the MSTAR dataset for experimentation. Since CFAR is pixel-based, it consumes more time on large scenes even if GPU is used. Therefore, inputs are passed through the convolutional layer in CP-CFAR to distinguish between targets and background, enter the pooling layer to reduce images' dimensions, and finally enter the two-parameter CFAR, resulting in improved detection performance. The CP-CFAR outperforms the CA-CFAR and the two-parameter-CFAR (Ai et al., 2018) in terms of time and accuracy.

3.3. Other CNN based method

Apart from detectors using CNNs based on SSD and CFAR, other techniques involved encoder–decoder in the CNN networks and are discussed below.

3.3.1. SER-Faster R-CNN

This work proposed a faster R-CNN-based network architecture to improve detection performance in SAR images (Lin et al., 2019). The main contribution of this work is the proposal of the squeeze and excitation mechanism in the Faster R-CNN along with rank modification, hence the name Squeeze and Excitation Rank Faster R-CNN (SER-faster R-CNN). The squeeze and excitation mechanism is adopted in order to improve the representation power of a CNN. Also, the ranking of vectors in the SE block is done to suppress the redundant feature maps. The last three convolutional layers of a VggNet (Simonyan and Zisserman, 2014) trained on ImageNet is used as the first three layers of the SER Faster R-CNN model. The feature maps from each of these convolutional layers are normalized using L2 norm. The normalized results are concatenated and convolved, resulting in a 512 channeled input. This is then fed into RPN (Ren et al., 2015) to generate region proposals. The sub-feature maps from the first three convolutional layers of the SER-Faster R-CNN model are also extracted using ROI pooling (Ren et al., 2015). These sub-feature maps are concatenated and act as input to the SER block, which is convolved to 512 channels using convolutional layer and pooled to $512D$ vector with the help of global average pooling. This vector then goes through the encoding decoding block whose encoding dimension depends on a particular parameter r. This block outputs the top K values preserved and used to improve sub-feature mapping for the second classification stage. Hyperparameters used in the SER Faster R-CNN model is shown in Table 10. Detection performance realized from the concatenation of the first three layers is effective and hence chosen. Also, the performance of faster R-CNNs with SE block is effective. Performance on SE block with top 256-rank vectors is also evaluated on different r values, and the model with $r = 1$ achieved the highest F1 score.

Table 10
Hyperparameters used in SER faster R-CNN (Lin et al., 2019).

Hyperparameters	Values
Learning rate	1×10^{-4}
Epoch	20 000
Stride size	2
RPN anchor scales(r)	1, 2, 4, 8
Aspect ratios	0.125, 0.25, 0.5, 1

3.3.2. MD1D

Other works have also been proposed recently for various SAR image applications that involve SAR target detection. Wang et al. have proposed a detection technique that detects harbors from complex SAR images (Wang et al., 2021). In this work, the authors have incorporated multi-directional one-dimensional (MD1D) scanning to obtain the control points on the sea-land masking. The representation vectors are selected based on the control points. A one-dimensional CNN is then used to discriminate the harbor features, after which the features are finally merged to obtain the final detection results. It may be observed that the CNN used in MD1D consists of only three convolutional layers. This method shows significant results, and hence testing this technique in the detection of maritime targets other than harbor can be explored to benefit the surveillance of oceans using SAR images.

3.3.3. AM-TR

In another work, Wei et al. proposed a semi-supervised SAR target detection model using the attention mechanism (Wei et al., 2021). Hence we refer to this work as AM-TR (Attention Mechanism-based Target Recognition). AM-TR uses CNN models by dividing the process into four modules, the feature extraction module for extracting relevant features from SAR images, the attention module for obtaining the attention map, the scene recognition module for specifying whether the input SAR image contains targets or not, and the detection module for generating multiscale features. These multiscale features are used to predict the targets through convolutional predictors to finally generate the detection results. Though the work outperforms other SAR target detection works (Ren et al., 2015; Zhang et al., 2016), issues such as false alarms and missing targets prevail.

3.4. Observations

The different works discussed in the previous sections have shown their respective contributions in the field of SAR target detection, of which few are based on SSD and few on CFAR. It was observed from TD-A (Wang et al., 2019) that the use of parameters from pre-trained VggNET (Simonyan and Zisserman, 2014) in SSD (Liu et al., 2016) plays a role in detection performance. However, a network such as ResNet (He et al., 2016) outperforms VggNet. Therefore, ResNet can be studied in the future and used in pre-training SSD networks in TD-A (Wang et al., 2019). The detection method in CFAR-CNN (Wang et al., 2018) has used dual CFAR. The output then feeds a simple six-layered CNN architecture. This way, a lesser detection time is achieved, but few false alarms still exist. This is because CFAR results in false alarms (Wang et al., 2019). Parameters for CNN used in CFAR-CNN (Wang et al., 2018) can be adopted from pre-trained networks, and this can be explored in future works, which might reduce false alarms and improve the accuracy. The method in MF-RCNN (Kang et al., 2017) modifies the Faster R-CNN with the inclusion of a VggNet based CNN into its architecture before undergoing Faster R-CNN followed by a CFAR detector. This method has proved that using a CFAR detector for less scored regions has improved the detection performance compared to Faster R-CNN. On the other hand, choosing the range of scores to be included by the CFAR is challenging since the values chosen in Kang et al. (2017) result in a few false alarms. The MR-SSD (Ma et al., 2018) is similar to TD-A (Wang et al., 2019) in the sense that both have connections to the SSD and VggNet. In MR-SSD (Ma et al., 2018), the architecture used is the VggNet where the fully connected layers are being replaced with SSD-based convolutional layers, whereas in TD-A (Wang et al., 2019), parameters of VGGNet are taken from SSD. The novelty of MR-SSD (Ma et al., 2018) is that it can detect not only ships but also detect other maritime targets. It is also observed in MR-SSD (Ma et al., 2018) that the idea of transforming the single-channel SAR image into a multi-resolution three-channeled image, though depending on some parameter λ, has its role played in the outcome. However, the challenge lies in choosing the value of this parameter, as it differs based on application. The authors in Ma et al. (2018) have proved their detection method on large-scale SAR images, whereby it outperforms Faster R-CNN and SSD, while it still has few false alarms. Sea land masking may be inaccurate at times (Lin et al., 2019); therefore, this might have caused the false alarms. The SER Faster R-CNN (Lin et al., 2019) have used the last convolutional layers of a pre-trained VggNet to obtain shared feature maps for region proposal. It is observed in Lin et al. (2019) that the use of Squeeze and excitation mechanism to modify the RPN architecture along with the ranking of excitation vectors have helped in improving the detection performance.

In conclusion, we observe that the S-SSD (Du et al., 2020) and MR-SSD (Ma et al., 2018) give better results in their respective datasets. A comparison of the same can be made by studying them on the same dataset in the future. On the other hand, S-SSD gives better results compared to TD-A (Wang et al., 2019) on the same MSTAR dataset. This is because the saliency information helps the S-SSD network to accentuate relevant features, which further improves the overall detection capability of the model. It is also observed that almost all the methods have used VggNet as pre-trained networks. The ResNet, which is known to achieve better accuracy than VggNet in the ImageNet challenge, can be explored for application as pre-trained networks in place of VggNet, which might improve target detection results. Also, other deep learning networks like STDNet (Bosquet et al., 2020) that are used for detecting small targets can be used in the form of a hybrid network with those that detect larger objects to achieve better results, especially in SAR images. It is also worth mentioning that none of the work has preprocessed the SAR images, such as removing noise before undergoing target detection, which might be a requirement when all these techniques are applied to noisy SAR images. Also, crosschecking of the existing target detection techniques on a very noisy SAR image needs to be done in case we encounter such images in the future. State of the art denoising models such as IDCNN Wang et al. (2017) and Passah et al. (2021) can be studied and appended before the detection models for better performance Table 3 shows a comparison summary of the different target detection works.

4. Target recognition

This section discusses one of the most widely applied processing methods in SAR visions called target recognition, along with recent works and techniques. Target recognition is the ability to recognize a target from among a group of other targets (Tait, 2005). Fig. 7 depicts the concept of target recognition with four classes being correctly identified and one class being mistakenly identified. Recognizing targets from a given SAR image is challenging as the images produced from SAR are not easy to interpret as those produced optically. SAR images contain clutter along with noise; therefore, to be able to identify which target is within the scope of the application is also a difficult job; hence it has to undergo pre-processing such as suppression of noise and detection of the region of interest. These processes can corrupt the relevant features from images and decrease recognition accuracy. In practical situations where time and cost are significant concerns, target recognition becomes challenging. In other aspects, the unavailability of labeled images in SAR also restricts the development of SAR target recognition (Cui et al., 2018). Another challenge in SAR Automatic Target Recognition (SAR-ATR) is automatically recognizing the same target with a different pose. Thus, the SAR-ATR is challenging and has become one of the research hotspots for remote sensing technology. Several works on SAR target recognition have been projected in the literature for the past decade. The authors in Dong et al. (2014) have used a sparse representation of monogenic signal for SAR target recognition. The process also involves the augmentation of the feature vectors, which are then fed to a sparse representation-based recognition framework. The network could adapt to noisy images. In the extended work Dong and Kuang (2015a), the authors have also developed score level fusion for the sparse representation framework along with a hybrid kernel learning that resulted in improved target recognition performance. The authors in Dong and Kuang (2015b) uses Riemannian geometry for recognition. The work also adopted monogenic signals combined with the help of a covariance matrix to identify SAR images. The authors developed two classification schemes: first, that involved mapping the covariance matrix that feeds the sparse representation framework, and second, that involved embedding of Riemannian manifold.

Lately, deep learning has been known to attract researchers because of its tremendous ability to learn without human interventions. Enormous deep learning applications have been made in various computer vision tasks resulting in noticeable results. This has led researchers to explore deep learning for various SAR interpretation tasks (Li et al., 2017b; Soldin, 2018). Various existing works on SAR target recognition, using deep learning, have been proposed in the literature, of which a few state-of-the-art works that are based on CNNs are discussed in the following subsection.

4.1. SAR target recognition based on CNN

Since CNNs are the most widely used deep learning class for analyzing visual images, we review only those target recognition works based on CNN, thereby giving insights on the approach followed while using CNNs in SAR target recognition tasks. Few such works are highlighted below.

4.1.1. ATR-CNN

At the time when recognition of targets was achieved only by using handcrafted features, the work in Chen and Wang (2014) has experimented with the use of CNNs in SAR target recognition. We referred to the work as ATR-CNN (Automatic Target Recognition using CNN). ATR-CNN (Chen and Wang, 2014) first tries to learn the kernel and bias by training the inputs using a sparse autoencoder. Autoencoder usually learns the non-trivial identity function of the given input (Goodfellow et al., 2016). The whole process was carried out by first extracting patches from inputs and applying average value subtraction and Zone Component Analysis(ZCA) (Krizhevsky and Hinton, 2009) to each patch. The patches are then trained using sparse autoencoder with backpropagation to learn the kernels and bias. Overall, this was unsupervised learning with no labels fed during training. The model using labeled images is then trained using a single-stage CNN

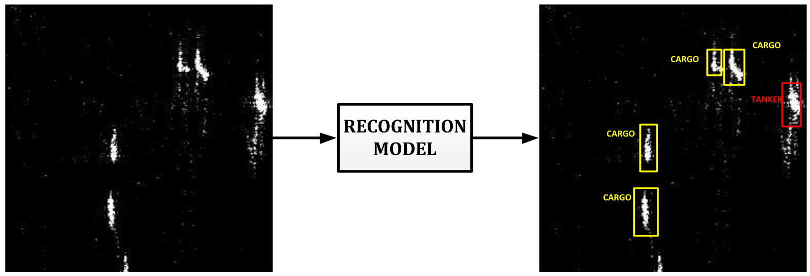

Fig. 7. Target recognition concept: Recognition model outputs recognition results (Image: SAR-ship dataset Wang et al., 2019).

tuned to the learned kernel and bias followed by a softmax regression classifier to generate the final output. The sparse autoencoder basically applies a sparsity constraint on the hidden layer, meaning that only a few neurons in the hidden layer will be activated. It may be mentioned that the sparse penalty used in this case is based on Kullback Leibler(KL) divergence (Kullback and Leibler, 1951). In the supervised learning stage, the single-stage CNN comprises a convolutional layer and a pooling layer with kernels and bias learned from the sparse autoencoder. ATR-CNN achieved 84.7% accuracy on ten types of MSTAR targets. Compared to techniques where features are extracted manually (Zhang and Huang, 2013), the accuracy of ATR-CNN (Chen and Wang, 2014) is only 5% less. Therefore ATR-CNN (Chen and Wang, 2014) have proved the capability of a single CNN to achieve accuracy close to that of handcrafted-based methods. This may also mean that if we use deeper CNN, the recognition performance might improve and outperform handcrafted recognition techniques.

4.1.2. All-in-one CNN

This technique aims at recognizing MSTAR targets by taking into primary consideration the advantages of augmented data while at the same time enlarging the datasets for training a CNN (Ding et al., 2016). Hence, it developed three data augmentation techniques: target translation, speckle noising, and pose synthesis. Target translation is achieved by the method of coordinate shifting with guard windows to restrict crossing the image boundary. An experiment on single augmentation based on target translation was performed by shifting each sample $6, 12, 21, 30$ and 45 times respectively and found that the recognition accuracy achieved was 93.30% compared to SVM (Hearst et al., 1998). It was also observed that the model trained on translation data does not work on noisy data. Therefore another augmentation on noising was carried out. This was achieved by using exponential distribution with certain parameter a set to $0.5, 0.5, 0.5, 1.0, 1.0, 1.0, 1.5, 1.5, 1.5$ respectively on each sample. An experiment on this single augmentation was also conducted and observed that CNN trained with noising augmentation data can perform better on noisy data than SVM. Lastly, for pose synthesis, K samples per class were randomly selected as base images to generate 5000 pose images with the help of a formula using azimuth angles. Experiments on different values of K were conducted wherein the highest accuracy achieved was when $K = 120$. Therefore the authors in Ding et al. (2016) combined all the three augmentations based on the results on each augmentation and trained a CNN model for target recognition. The model is named all-in-one CNN. These augmented data were used in training the CNN comprising three convolutional layers, three max-pooling layers, one fully connected layer, and a softmax layer. The recognition capability of the all-in-one CNN (Ding et al., 2016) was stronger compared to SVM and MINACE (Patnaik and Casasent, 2005). This is because MINACE can achieve translation invariance but not distortion invariance, especially with SAR images. On the other hand, SVM can handle distortions but not translation. Thus, the authors have proved that the inclusion of the three augmentation methods has contributed to improving SAR images' recognition accuracy in a unique way. A flowchart summarizing the technique in Ding et al. (2016) is shown in Fig. 8.

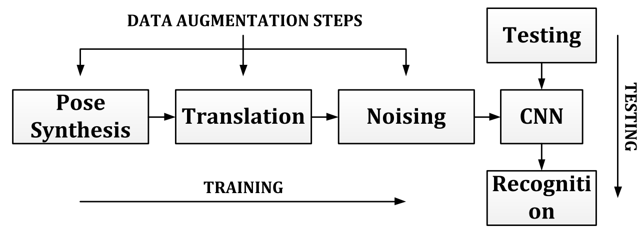

Fig. 8. Flowchart summarizing the technique in all-in-one CNN (Ding et al., 2016).

4.1.3. ATR-CNNAD

The unavailability of labeled SAR images makes CNN training difficult and challenging as CNN requires extensive training labeled data to learn more relevant features for future predictions. The work in Cui et al. (2018), hereafter referred to as ATR-CNNAD (ATR based on CNN with Assistant Decision), however, have tried to solve this issue for target recognition in unlabeled MSTAR images comprising of military vehicles. This was achieved by incorporating the concept of update learning with the help of a pre-trained CNN along with SVM (Hearst et al., 1998) as an assistant classifier. The CNN with three convolutional layers, a max-pooling layer and two fully connected layers were first pre-trained on a small labeled set called seed image set, along with SVM. The new unlabeled images are then fed iteratively into the pre-trained CNN in sync with the assistant classifier, and a decision was made based on their probability matrix output. Finally, in view of this decision, relevant images are added as new data with labels and retrain the CNN with this new set of data in order to update its parameters. This way, labeled data was increased while at the same time updating the learning ability of the model for upcoming new unlabeled data. It is also worth mentioning that each time the CNN is trained with a newly generated dataset, the error rates decrease gradually after every update, thereby improving recognition accuracy to 89%. Fig. 9 summarizes the technique in ATR-CNNAD (Cui et al., 2018). To further improve the recognition accuracy, improving the accuracy of each new training dataset can be studied. The update learning process can be optimized by improving the decision method, such as incorporating more than one assistant classifier for a better decision.

4.1.4. MFCNN

The presence of noise in SAR images has made target recognition challenging. This is because, to recognize a target in an image, the features associated with the target need to be learned. If features include noise features, the recognition performance gets degraded. To address this issue, several works have used pre-processing and pose information to lessen the effect of noise in target recognition (Srinivas et al., 2014; Dong et al., 2015; Deng et al., 2017). However, if proper pre-processing is not done, it may result in inaccurate prediction of targets because noise characteristics differ from image to image. Therefore, the authors in Cho and Park (2018) have proposed a multi feature-based convolutional neural network (MFCNN) for SAR target recognition without an extra pre-processing step. The MFCNN involves

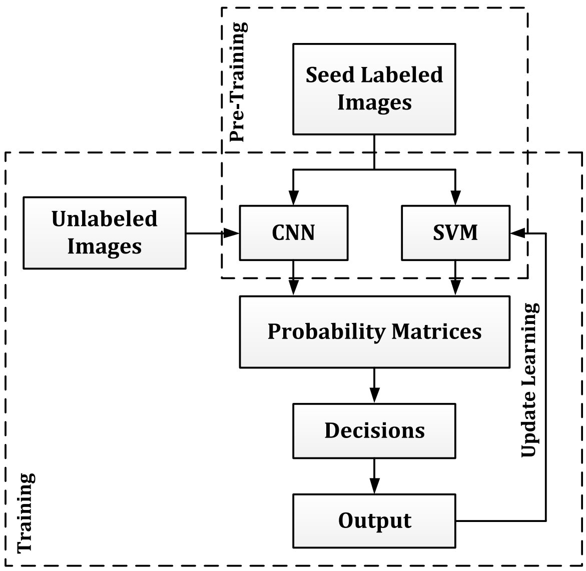

Fig. 9. Flowchart summarizing the technique in ATR-CNNAD (Cui et al., 2018).

three steps: multiple feature extraction, multiple feature aggregation, and target recognition. Multi-feature extraction comprises three stages FE1, FE2, FE3, whose details are highlighted in Table 12. The output of FE1, FE2 and FE3 are then aggregated in the next step called the multiple feature aggregation step forming a single vector that is finally fed to three fully connected layers, which is the final step for target recognition. The number of neurons in each fully connected layer is determined concerning VggNet (Simonyan and Zisserman, 2014), and the number of neurons in the final output layer is ten. Experimental results show that MFCNN has an average recognition rate of 95.52% on the MSTAR dataset (Ross et al., 1998; Laboratory, 2021) and it outperforms the conventional methods (Zhan et al., 2016; Ding et al., 2016; Srinivas et al., 2014; Dong et al., 2015; Deng et al., 2017), the reason being that in few approaches (Zhan et al., 2016; Ding et al., 2016), target features are extracted only from the final sub-sampling layer. The authors in Cho and Park (2018) additionally proved that MFCNN could attain a recognition accuracy of 94.85% even with only 50% of training data. In conclusion, the work in MFCNN (Cho and Park, 2018) can be extended to a multi-target environment wherein accuracy may be recorded, accordingly leaving room for improvement of target recognition in multiple target images.

4.1.5. VersNet

In most SAR target recognition works Ding et al. (2016), Cui et al. (2018) and Cho and Park (2018), single target chips are used in order to classify or recognize the object. The standard method usually involves three stages: detection, discrimination, and classification. The authors in Furukawa (2018) have proposed one such method by incorporating all these stages within a single CNN for end-to-end based SAR target recognition on the MSTAR dataset and named it the Verification Support Network (VersNet). This network takes arbitrary SAR image inputs with multiple targets of different classes and generates image output that specifies the class type, the target position, and pose information of each target on the image. The VersNet model consists of an encoder and a decoder, where the encoder is used to extract features from inputs. In contrast, the decoder converts the features to output the image recognition result. The encoder consists of ten blocks where each of the first four blocks is made up of two convolutional layers followed by a max-pooling layer. The fifth block consists of a convolutional layer followed by a dropout layer, and the last block consists of a single convolutional layer. On the other hand, the decoder stage consists of a single transposed convolutional layer which is finally used to generate the output. It is worth mentioning that VersNet does not contain fully connected layers, enabling the network to process images of arbitrary sizes. The segmentation labels of the MSTAR data are first generated for training the VersNet. As this method is expected to output the recognition results with the help of classification and the location and pose information results with the help of the segmentation, the experimental results are thus evaluated individually. For classification, the VersNet could attain an overall accuracy of 99.55 which means almost all targets are recognized correctly. On the other hand, the segmentation results show that the average Intersection over Union (IoU) on 10 classes along with 1 background and 1 foreground is 0.915. Finally, the Versnet model was also tested on multiclass and multitarget inputs where only visual results were displayed and seemed to be close to ground truth images. The VersNet (Furukawa, 2018) have proved the effectiveness of using encoder–decoder techniques in the CNN model without any fully connected layer. The advantage of Versnet is that it can accept random size inputs while improving the MSTAR target recognition performance. However, the challenge lies in including the distorted images, which were excluded from the testing set of this work.

4.2. Other target recognition methods

Several other SAR target recognition models have also been projected in the literature recently. The work in Jia et al. (2019) has proposed a recognition model that is flexible and less sensitive to noise and attitude angle. The deep features of the SAR image are first extracted using one-dimensional and two-dimensional SAR image data. For one-dimensional SAR image data, the stacked autoencoder is used for extracting deep features obtaining the one-dimensional feature vector. In contrast, CNN is used for extracting the features resulting in a two-dimensional feature vector for two-dimensional image data. The two feature vectors are then combined and passed through fully connected layers to generate the recognition output. As seen from this work, feature fusion plays a role in improving the performance of SAR recognition. The work in Gao et al. (2019) proposes a target recognition model by merging a basic CNN model with SVM. The CNN model was first trained by incorporating class separability as part of the cost function and uses softmax as a classifier, after which it was removed to train the SVM using only the top features of the trained model. Recognition accuracy on MSTAR targets has improved but still suffers misclassification.

4.3. Observations

This section briefly discusses the observations that can be concluded from the various SAR target recognition works reviewed in the previous section. Firstly, we observed from ATR-CNN (Chen and Wang, 2014) that the use of only a single CNN has also shown some effectiveness in the performance of SAR target recognition and whose kernels and bias are pre-trained on a sparse auto-encoder. However, the results of ATR-CNN have led to the urge to use more than one CNN to improve the recognition performances further. Therefore, most of the works are currently based on deep CNN networks. On the other hand, deep CNNs require more data to get trained to the desired results, which is challenging in SAR images. The all-in-one CNN (Ding et al., 2016) has included a technique for increasing the data by performing its own data augmentation techniques while at the same time improving the recognition performances but is not end-to-end based. A way to club these fragmented portions of pre-processing into a single model might help in the future. Again, considering the unavailability of labeled SAR images, the ATR-CNNAD (Cui et al., 2018) is able to undergo an improved target recognition method with the help of a seed amount of labeled set by employing an assistant classifier and iteratively training the model on new labeled data. In the future, the technique for improving the accuracy of each new training dataset can be studied to further improve the recognition accuracy. The update learning process can be optimized

Table 11
Summarizing the concept of different SAR target recognition networks.

Paper	Aim	Method	Architecture	Novelty	Advantage	Disadvantage	Results	Future
ATR-CNN (Chen and Wang, 2014)	To recognize target by using CNN rather than handcrafted features	Supervised + Unsupervised stage	Sparse autoencoder + 1CNN	Uses KL-divergence as sparse penalty	Accuracy at par with handcrafted features	When tested on non-deformation targets, accuracy is 84.7%	90% accuracy	Deeper CNNs may help improve accuracy further
All-in-one CNN (Ding et al., 2016)	To recognize targets in challenging situation such as translated targets and pose synthesis	End to end + posterior augmentation	Target-translation +Speckle-Noising +Pose-synthesis +Conv96@3x3, Conv96@3x3, Maxpool, Conv256@3x3, Maxpool, Fc, Softmax	Uses Three data augmentation techniques: target translation, speckle noising, pose synthesis	Outperforms SVM & MINACE and recognition accuracy is improved	Domain specific	93.16% on original data, 82.40% on translated data, 91.89% on noisy data	Deeper networks can be explored for the same
ATR-CNNAD (Cui et al., 2018)	To use unlabeled targets for training a CNN to recognize targets	Initial training + Update learning	Conv32@9x9, pool, Conv64@5x5, pool, Conv128@3x3, pool, fc1000, fc10, softmax	Uses iterative learning strategy with the help of pretrained CNN and assistant classifier	Dynamically generating labels	Inefficient update learning	89% accuracy	Update learning process can be optimized by incorporating more than one assistant classifiers.
MFCNN (Cho and Park, 2018)	To implement SAR-ATR without using a separate preprocessing process or pose information	Three stages: multiple feature extraction + feature aggregation + recognition	(Conv3x3)x2, avgPool, (Conv3x3)x2, maxpool. (Conv3x3)x2, maxpool, Conv3x3, maxpool, (Conv3x3)x2, avgpool, Conv3x3, avgpool. fc1,fc2,fc3	Parallel use of max-pool and avg-pool to simultaneously extract strong features and to reduce noise effects respectively	Could attain accuracy 94.85% with only 50% of training data	Test not extended to multi target environment	Avg. recognition rate is 95.52%	The model can be optimized and explored to extend its application in multi target environment
VersNet (Furukawa, 2018)	To perform SAR target recognition along with location and pose information from a multi target image based on end to end	End to end based CNN	(Conv32@3x3, ReLU) x2, maxpool, (Conv64@3x3, ReLU) x2, maxpool, (Conv128@3x3, ReLU) x2, maxpool, (Conv256@3x3, ReLU) x2, maxpool, Conv512@6x6, Dropout, Conv12@1x1, 32x32TransposedConv	Uses encoder decoder mechanism in the architecture with no fc layer	Can accept random size inputs because of no fc layers	The model was not tested on the distorted MSTAR images	99.5% overall accuracy	–

*Conv = Convolution, Fc = Fully Connected, **Dataset used MSTAR.

Table 12
Feature extraction methods used in MFCNN (Cho and Park, 2018).

Feature extraction	Methods	Aim
FE1	Features are extracted using both maxpooling and average pooling operations.	Aims at extracting mixed features.
FE2	Features are extracted using two max-pooling operations.	Aims at extracting target features.
FE3	Features are extracted using two average-pooling operations.	Aims at extracting features with reduced noise effect.

by improving the decision method, such as incorporating more than one assistant classifier for a better decision. In most of the works above, recognition performance may degrade when encountering noisy SAR images. To deal with such images, the MFCNN (Cho and Park, 2018) have used a technique that can recognize targets from noisy SAR images by extracting and aggregating target features as well as reducing noise features, but the work can further be extended to multiple target images in the future. Unlike ATR-CNN (Chen and Wang, 2014), all-in-one CNN (Ding et al., 2016), ATR-CNNAD (Cui et al., 2018) and MFCNN (Cho and Park, 2018) that deals with single target chip images, the Versnet (Furukawa, 2018) have considered multiple target input images of random sizes for recognition with the help of encoder–decoder fully convolutional technique along with segmentation that has led to performance improvement. A graph summarizing the average results of several works is shown in Fig. 10, and a comparison summary of the different target recognition works is shown in Table 11.

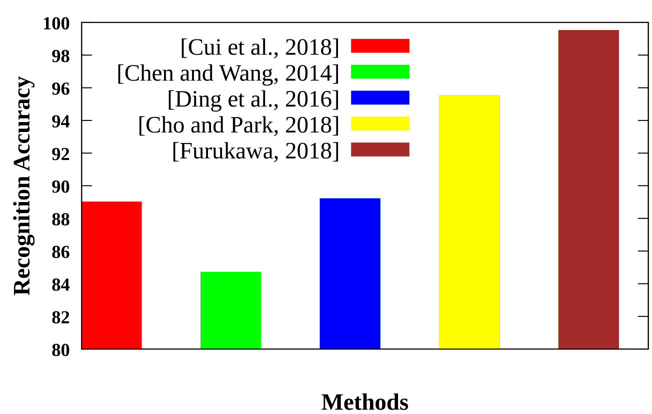

Fig. 10. Graph depicting the recognition accuracy of different methods on MSTAR targets.

5. Research issues and challenges

Several issues and challenges with respect to target detection and recognition from SAR images are of great concern. Detecting targets promptly and precisely from SAR images is currently an open area of research. With the occurrence of speckle noise in SAR images, target detection becomes a challenging job. Since in SAR complex scenes, clutter and targets resemble substantially, therefore identifying which target is within the scope of the application is difficult in SAR images. Even though noise removal techniques might be the remedy for the clutter-target issue, the process of suppressing noise can corrupt the most relevant features from images, resulting in the degradation of detection and recognition accuracy. Also, as observed from this study, the samples used for training are usually SAR images with numerous targets and sophisticated backgrounds of large scenes. Even the popular CFAR (Gao et al., 2009) technique which performs efficiently for target detection on non-obstructed scenes, degrades on encountering complicated scenes like SAR scenes (Wang et al., 2019). Because of this, SAR detection and recognition models experience false alarm issues. Deep learning has shown a tremendous impact on the improvement of various computer vision tasks (Zhu et al., 2017). However, the use of deep learning, such as convolutional neural networks for SAR image processing like detection or recognition, gave rise to issues like overfitting due to the non-availability of massive SAR datasets along with labeled images for training deep learning models resulting in incorrect predictions (Cui et al., 2018). Hence the use of deep learning also becomes challenging for SAR image detection and recognition. Therefore, research on the different data augmentation methods relevant to appropriately increasing the SAR image datasets is still undergoing. Data augmentation helps in increasing the population of the dataset making them relevant for training deep learning models. Data augmentation also improves the generalization capacity of the

Table 13
Summary of the different target detection and recognition approaches along with their significant outcome.

Ref	Image processing approach	Application environment	Significant outcome
Wang et al. (2019)	TD using SSD merged CNN	Detecting vehicles using MSTAR	Improved detection capability
Wang et al. (2020)	TD using feature merged SSD based on CNN and area-weighted loss function	Detecting multi scale objects	Small objects are also given importance
Ma et al. (2018)	TD using a no FC VggNet based model with SSD parameters	Detection of maritime targets	Other than ships, could detect other maritime targets
Du et al. (2020)	TD by mapping saliency information to SAR images to improve representation capability in complex scenes	Detection of targets in miniSAR images	Detection performance improved with fewer parameters
Wang et al. (2018)	TD using CFAR based CNN with CFAR at two places: local and global detection	Ship detection systems	Lesser detection time
Kang et al. (2017)	TD by modifying Faster R-CNN and giving importance to low score bounding boxes	Detecting multi scale ships	Reduces false alarms
Cui et al. (2018)	Uses sobel operator for convolutional kernel to improve contrast between target and background	Detecting military targets using MSTAR	Eliminates false alarms by undergoing morphological processing and median filtering
Chen and Wang (2014)	TR by using CNN based sparse auto encoder with KL divergence as sparse penalty	MSTAR target recognition	Performs better than hand crafted features even with a single CNN layer
Ding et al. (2016)	CNN trained by focusing on target translation, pose missing and speckle noising	MSTAR target recognition	Feasible for translated targets and missing pose
Cui et al. (2018)	Uses assistant classifier along with CNN for better decision regarding recognition	MSTAR target recognition	Minimum labeled data is enough to train the network
Cho and Park (2018)	TR using multiple extracted and aggregated features of SAR images with the help of CNN	MSTAR target recognition	Attain better accuracy even with only 50% of training data
Furukawa (2018)	Uses encoder–decoder based CNN with no FC layer	MSTAR target recognition	The network could process images with arbitrary sizes. It also specifies target position and pose information

TD = Target Detection, TR = Target Recognition, FC = Fully Connected.

model as it helps expand the dataset with some changes. Another challenge in SAR Automatic Target Recognition (SAR-ATR) is the ability to recognize similar targets with different poses in SAR images. For cases with mixed target size SAR images, detection performance deteriorates. The performance of SAR detection models on mixed-size targets can be improved if models such as STDNet (Bosquet et al., 2020) that are used to detect small targets are incorporated. It should also be noted that many other fields of research, such as astrophysics (Walmsley et al., 2020) and medical imaging (Gong et al., 2019), are also affected by some of the problems with deep learning applications, such as the lack of high-quality training data. Even the faulty interpretations due to the presence of noise in the data is another common issue faced by many other areas of research other than SAR, such as bio-medical applications (Gong et al., 2019; Qin et al., 2019; Zheng et al., 2016) (see Table 13).

6. Future approach

As observed from this survey, the deep learning-based approaches have shown improvement in the performances of various SAR images interpretation, from detection, recognition, and classification. However, all these methods have advantages and disadvantages, which can be further improved by adopting emerging ideas in the relevant field as future works. Concerning the detection of SAR targets, we observed that VggNet is the most used base architecture (Wang et al., 2019; Kang et al., 2017; Ma et al., 2018; Lin et al., 2019). Rarely could we find target detection models that incorporated other architectures such as ResNet, UNet, GoogleNet. As per the Imagenet competition (Russakovsky et al., 2015), ResNet (He et al., 2016) outperforms VggNet (Simonyan and Zisserman, 2015). Therefore, in future work, such architectures can be explored for SAR target detection by modifying the parameters and improving detection performances. For SAR images that include mixed target sizes, detection accuracy can be improved if models such as STDNet (Bosquet et al., 2020) that are used to detect small targets are incorporated, forming a potential hybrid detection network. Fig. 11 depicts the possible future hybrid model based on existing SAR detection works that may improve the detection performances in SAR images. Inspired by the work in CNN-MR (Bentes et al., 2018), the hybrid model takes in multiple resolution inputs in order to enable the model to learn more relevant features from SAR images. The incorporation of global and local CFAR in the model can reduce possible false alarms since global CFAR determines targets while local CFAR determines whether targets are actual targets or false targets. The CNN used in local CFAR can be pretrained using recent networks like EfficientNet (Tan and Le, 2019), GoogleNet, etc. STDNet (Bosquet et al., 2020) is known for its ability to detect small targets; therefore, to enable the model to detect both small and large targets, STDNet is used parallelly along with the model whose output is then concatenated with that of the prior model to get the final detection results. A-ConvNet (Chen et al., 2016) is known to perform better in classification problems using less number of parameters; it can therefore be adopted for recognition purposes as well. This may not only improve the performance but will also help to reduce the number of parameters in target recognition models in the future. Also, most of the recognition works reviewed in this paper have not explored deeper architectures (Chen and Wang, 2014; Ding et al., 2016; Cui et al., 2018). Therefore the performance of SAR target recognition methods using deeper networks can be implemented and analyzed in the future. Considering the advantages and disadvantages of all the works reviewed in this survey paper, an implementation of a hybrid

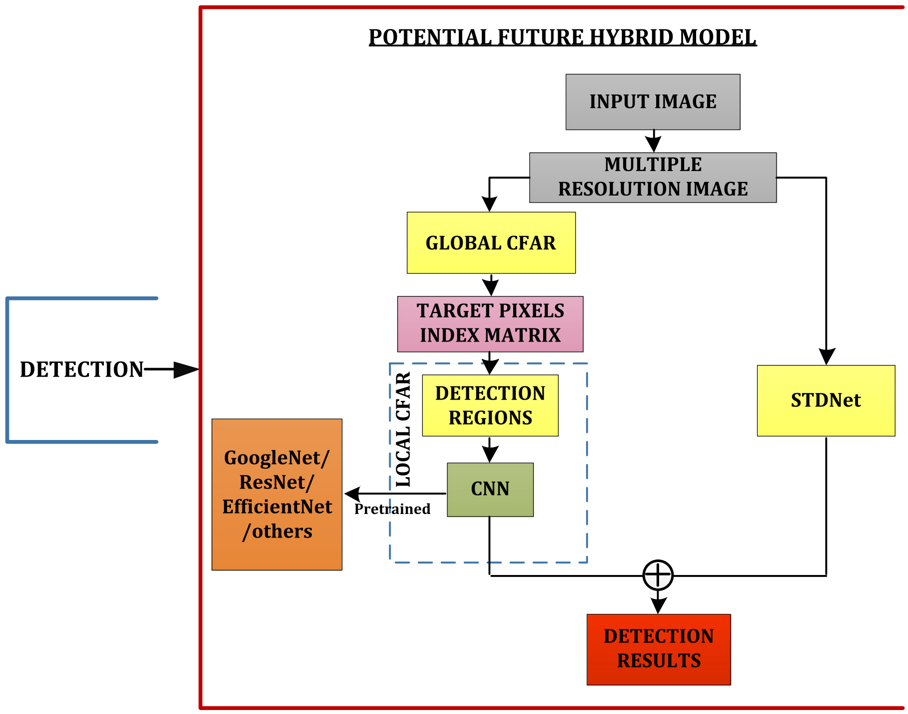

Fig. 11. A potential future model for SAR target detection based on existing works.

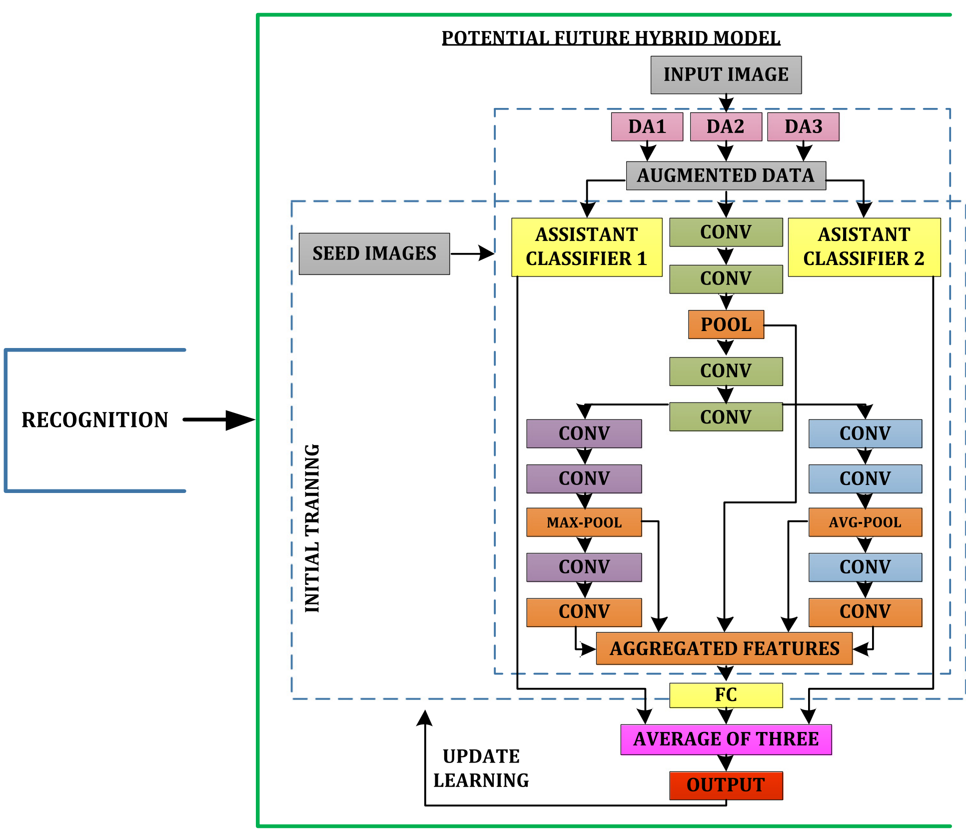

Fig. 12. A potential future model for SAR target recognition based on existing works.

architecture can be carried out in the future for recognition. It may be mentioned that the hybrid network may be trained initially by using the seed labeled images accompanied by the update learning process. This idea of incorporating seed images is motivated by the work in Shang et al. (2018), as this method leverages the need for a large set of labeled images which is challenging to obtain in the case of SAR images. In the hybrid architecture, encoder–decoder style CNN can be adopted along with two other CNN models because, as observed from the work in Furukawa (2018), encoder–decoder networks produce better representations of features. The use of three networks in a single

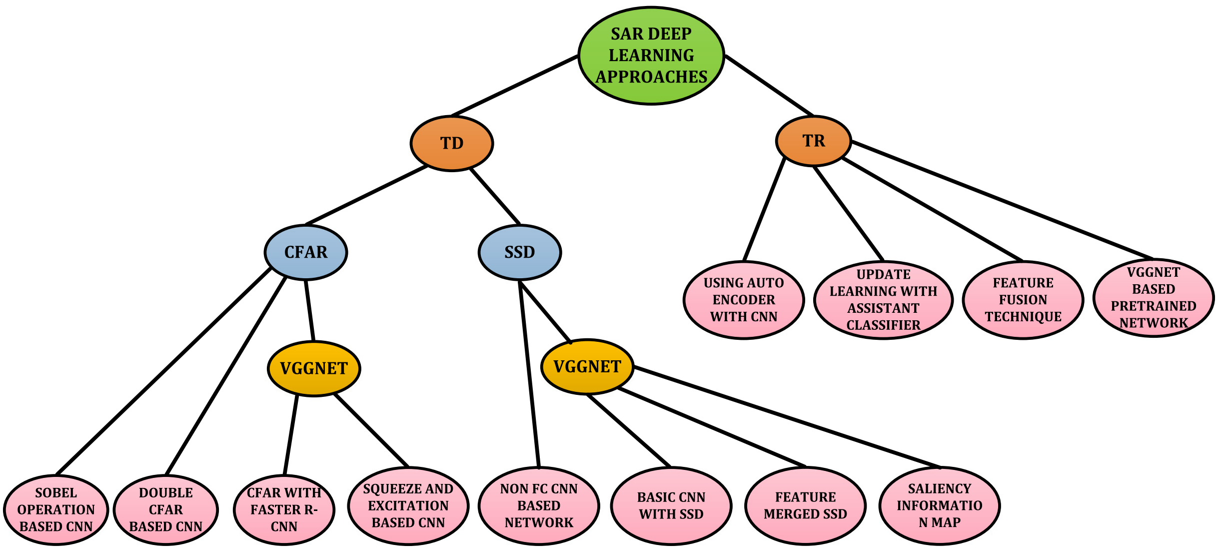

Fig. 13. Some of the existing deep-learning based approaches for SAR target detection and recognition. (TD: Target Detection, TD: Target Recognition, CFAR: Constant False Alarm Rate, SSD: Single Shot multibox Detector).

model can be examined wherein the features from each network will be aggregated to produce the recognition output. The multiple feature aggregation method is inspired by the method in Cho and Park (2018) that aims to gather the most relevant features for recognition purposes. Furthermore, assistant classifiers help ensure the most probable output (Cui et al., 2018); therefore, incorporating two other assistant classifiers as part of the hybrid architecture before producing the final output may help produce accurate results. The architecture of the model is highlighted in Fig. 12.

7. Conclusion

Various SAR image processing techniques have been presented in this paper, from SAR target detection to SAR target recognition. In target detection using SAR images, it can be observed that the detection of false alarms is still a significant issue. Also, false recognition of targets needs to be addressed when using SAR images for target recognition. This paper discusses some of the recent works on target detection and target recognition, highlighting their advantages and disadvantages followed by future scope based on the study. The tree diagram shown in Fig. 13 highlights the existing approaches adopted by various target detection and target recognition works in the literature. After analysis, it is observed that most works for SAR target detection have incorporated VggNet as part of the proposed architecture. Therefore, the performance of SAR target detection may be enhanced by adopting recent deep learning architectures such as Inception and EfficientNet. Further, the integration of features in deep learning models is less explored for detecting and recognizing targets in SAR images, and the same can be studied in the future. Future approaches have also been highlighted in this paper based on the study. Research issues and challenges concerning SAR target detection and recognition have also been discussed as depicted in Fig. 14. Possible future models built from the study for both SAR target detection and recognition are also presented in this paper. In conclusion, this paper encapsulates the significant SAR processing works comprising target detection and recognition and their advantages and disadvantages, intending to benefit the researchers and ease them to develop new algorithms in the future. The paper will also help researchers in other fields of image analysis and computer vision, as some of the ideas that were used for SAR image processing such as the integration of features to recognize targets, especially from noisy data, will also benefit other image processing fields, as this technique helps in reducing the noise features (Cho and Park, 2018) leading to a better understanding of the image information for better interpretations. Additionally, the blending of the local CFAR and global CFAR for target detection purposes in other fields apart from SAR also helps, as this is beneficial for reducing false alarms and hence can be globally adopted.

CRediT authorship contribution statement

Alicia Passah: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization. **Samarendra Nath Sur:** Conceptualization, Formal analysis, Investigation, Writing – original draft. **Ajith Abraham:** Validation, Writing – review & editing, Visualization, Supervision. **Debdatta Kandar:** Validation, Resources, Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

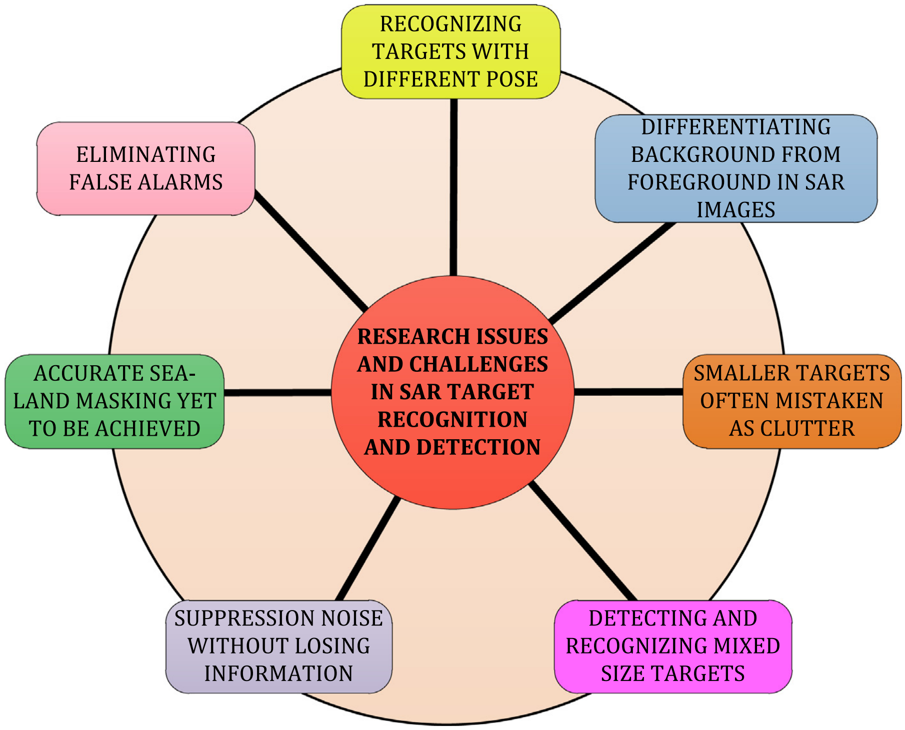

Fig. 14. Existing research issues concerning target detection and recognition using SAR images.

References

Agrawal, A., Mangalraj, P., Bisherwal, M.A., 2015. Target detection in SAR images using SIFT. In: 2015 IEEE International Symposium on Signal Processing and Information Technology. ISSPIT, pp. 90–94. http://dx.doi.org/10.1109/ISSPIT.2015.7394426.

Ai, J., Yang, X., Song, J., Dong, Z., Jia, L., Zhou, F., 2018. An adaptively truncated clutter-statistics-based two-parameter CFAR detector in SAR imagery. IEEE J. Ocean. Eng. 43 (1), 267–279.

Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology. ICET, pp. 1–6.

Baltierra, S., Valdebenito, J., Mora, M., 2022. A proposal of edge detection in images with multiplicative noise using the Ant Colony System algorithm. Eng. Appl. Artif. Intell. 110, 104715.

Bentes, C., Velotto, D., Tings, B., 2018. Ship classification in TerraSAR-X images with convolutional neural networks. IEEE J. Ocean. Eng. 43 (1), 258–266.

Biondi, F., 2019. A polarimetric extension of low-rank plus sparse decomposition and radon transform for ship wake detection in synthetic aperture radar images. IEEE Geosci. Remote Sens. Lett. 16 (1), 75–79.

Bosquet, B., Mucientes, M., Brea, V.M., 2020. STDnet: Exploiting high resolution feature maps for small object detection. Eng. Appl. Artif. Intell. 91, 103615.

Brownlee, J., 2016. Deep Learning and Artificial Neural Networks. URL: https://machinelearningmastery.com/what-is-deep-learning/. Last accessed 2021-08-27.

Chan, Y.K., Koo, V.C., 2008. An introduction to synthetic aperture radar (SAR). Prog. Electromagn. Res. 2, 27–60.

Changlin, W., Xuelian, Z., 2007. Analysis of man-made target detection in SAR imagery. In: Annual Meeting of the Envisat Symposium, Montreux, April. pp. 23–27.

Chaturvedi, S.K., 2019. Study of synthetic aperture radar and automatic identification system for ship target detection. J. Ocean Eng. Sci. 4 (2), 173–182.

Chen, S., Wang, H., 2014. SAR target recognition based on deep learning. In: 2014 International Conference on Data Science and Advanced Analytics. DSAA, pp. 541–547.

Chen, S., Wang, H., Xu, F., Jin, Y., 2016. Target classification using the deep convolutional networks for SAR images. IEEE Trans. Geosci. Remote Sens. 54 (8), 4806–4817.

Cheng, W.X., Gao, R., Suganthan, P., Yuen, K.F., 2022. EEG-based emotion recognition using random convolutional neural networks. Eng. Appl. Artif. Intell. 116, 105349.

Cheng, G., Han, J., Lu, X., 2017. Remote sensing image scene classification: Benchmark and state of the art. Proc. IEEE 105 (10), 1865–1883.

Cheng, G., Zhou, P., Han, J., 2016. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. IEEE Trans. Geosci. Remote Sens. 54 (12), 7405–7415.

Cho, J.H., Park, C.G., 2018. Multiple feature aggregation using convolutional neural networks for SAR image-based automatic target recognition. IEEE Geosci. Remote Sens. Lett. 15 (12), 1882–1886.

Collard, F., Ardhuin, F., Chapron, B., 2005. Extraction of coastal ocean wave fields from SAR images. IEEE J. Ocean. Eng. 30 (3), 526–533.

Cui, Z., Quan, H., Cao, Z., Xu, S., Ding, C., Wu, J., 2018. SAR target CFAR detection via GPU parallel operation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (12), 4884–4894.

Cui, Z., Tang, C., Cao, Z., Dang, S., 2018. SAR unlabeled target recognition based on updating CNN with assistant decision. IEEE Geosci. Remote Sens. Lett. 15 (10), 1585–1589.

Deng, S., Du, L., Li, C., Ding, J., Liu, H., 2017. SAR automatic target recognition based on euclidean distance restricted autoencoder. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (7), 3323–3333.

Ding, J., Chen, B., Liu, H., Huang, M., 2016. Convolutional neural network with data augmentation for SAR target recognition. IEEE Geosci. Remote Sens. Lett. 13 (3), 364–368.

Dong, G., Kuang, G., 2015a. Classification on the monogenic scale space: Application to target recognition in SAR image. IEEE Trans. Image Process. 24 (8), 2527–2539. http://dx.doi.org/10.1109/TIP.2015.2421440.

Dong, G., Kuang, G., 2015b. Target recognition in SAR images via classification on Riemannian manifolds. IEEE Geosci. Remote Sens. Lett. 12 (1), 199–203. http://dx.doi.org/10.1109/LGRS.2014.2332076.

Dong, G., Kuang, G., Wang, N., Zhao, L., Lu, J., 2015. SAR target recognition via joint sparse representation of monogenic signal. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (7), 3316–3328.

Dong, G., Wang, N., Kuang, G., 2014. Sparse representation of monogenic signal: With application to target recognition in SAR images. IEEE Signal Process. Lett. 21 (8), 952–956. http://dx.doi.org/10.1109/LSP.2014.2321565.

Du, L., Li, L., Wei, D., Mao, J., 2020. Saliency-guided single shot multibox detector for target detection in SAR images. IEEE Trans. Geosci. Remote Sens. 58 (5), 3366–3376.

El-Darymli, K., Gill, E.W., Mcguire, P., Power, D., Moloney, C., 2016. Automatic target recognition in synthetic aperture radar imagery: A state-of-the-art review. IEEE Access 4, 6014–6058.

El-Darymli, K., McGuire, P., Power, D., Moloney, C.R., 2013. Target detection in synthetic aperture radar imagery: A state-of-the-art survey. J. Appl. Remote Sens. 7 (1), 071598.

Feng, S., Ji, K., Ma, X., Zhang, L., Kuang, G., 2022. Target region segmentation in SAR vehicle chip image with ACM net. IEEE Geosci. Remote Sens. Lett. 19, 1–5. http://dx.doi.org/10.1109/LGRS.2021.3085188.

Ferro-Famil, L., Reigber, A., Pottier, E., Boerner, W.., 2003. Scene characterization using subaperture polarimetric SAR data. IEEE Trans. Geosci. Remote Sens. 41 (10), 2264–2276.

Fu, C.-Y., Liu, W., Ranga, A., Tyagi, A., Berg, A.C., 2017. Dssd: Deconvolutional single shot detector. arXiv preprint arXiv:1701.06659.

Fukushima, K., 1980. A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybernet. 36, 193–202.

Furukawa, H., 2018. Deep learning for end-to-end automatic target recognition from synthetic aperture radar imagery. arXiv preprint arXiv:1801.08558.
Gao, F., Huang, T., Sun, J., Wang, J., Hussain, A., Yang, E., 2019. A new algorithm for SAR image target recognition based on an improved deep convolutional neural network. Cogn. Comput. 11 (6), 809–824.
Gao, G., Liu, L., Zhao, L., Shi, G., Kuang, G., 2009. An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images. IEEE Trans. Geosci. Remote Sens. 47 (6), 1685–1697.
Gao, F., You, J., Wang, J., Sun, J., Yang, E., Zhou, H., 2017. A novel target detection method for SAR images based on shadow proposal and saliency analysis. Neurocomputing 267, 220–231. http://dx.doi.org/10.1016/j.neucom.2017.06.004, URL: https://www.sciencedirect.com/science/article/pii/S0925231217310597.
Girshick, R., 2015. Fast R-CNN. In: 2015 IEEE International Conference on Computer Vision. ICCV, pp. 1440–1448.
Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. pp. 580–587.
Gong, K., Guan, J., Kim, K., Zhang, X., Yang, J., Seo, Y., El Fakhri, G., Qi, J., Li, Q., 2019. Iterative PET image reconstruction using convolutional neural network representation. IEEE Trans. Med. Imaging 38 (3), 675–685.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning, Vol. 1. MIT Press Cambridge, Cambridge.
Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Chen, T., 2018. Recent advances in convolutional neural networks. Pattern Recognit. 77, 354–377.
Hachicha, S., Chaabane, F., 2014. On the SAR change detection review and optimal decision. Int. J. Remote Sens. 35 (5), 1693–1714.
Hänsch, R., Hellwich, O., Tu, X., 2016. Machine-learning based detection of corresponding interest points in optical and SAR images. In: 2016 IEEE International Geoscience and Remote Sensing Symposium. IGARSS, pp. 1492–1495.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 770–778.
Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13 (4), 18–28.
Huang, Y., Liao, G., Zhang, Z., Xiang, Y., Li, J., Nehorai, A., 2018. SAR automatic target recognition using joint low-rank and sparse multiview denoising. IEEE Geosci. Remote Sens. Lett. 15 (10), 1570–1574.
Hubel, D.H., Wiesel, T.N., 1968. Receptive fields and functional architecture of monkey striate cortex. J. Physiol. 195 (1), 215–243.
ICEYE-Finland, 2019. ICEYE. URL: https://www.iceye.com. Last accessed 2021-08-25.
Iervolino, P., Guida, R., 2017. A novel ship detector based on the generalized-likelihood ratio test for SAR imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (8), 3616–3630.
Iervolino, P., Guida, R., Whittaker, P., 2015. A new GLRT-based ship detection technique in SAR images. In: 2015 IEEE International Geoscience and Remote Sensing Symposium. IGARSS, pp. 3131–3134.
Inc., A.S., 1997. Theory of synthetic aperture radar. URL: http://www.geo.uzh.ch/fpaul/sar_theory.html. Last accessed 2021-05-10.
Iqbal, S., Siddiqui, G.F., Rehman, A., Hussain, L., Saba, T., Tariq, U., Abbasi, A.A., 2021. Prostate cancer detection using deep learning and traditional techniques. IEEE Access 9, 27085–27100.
Itti, L., Koch, C., Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. 20 (11), 1254–1259.
Jia, Z., Guangchang, D., Feng, C., Xiaodan, X., Chengming, Q., Lin, L., 2019. A deep learning fusion recognition method based on SAR image data. Procedia Comput. Sci. 147, 533–541, URL: https://www.sciencedirect.com/science/article/pii/S1877050919302510. 2018 International Conference on Identification, Information and Knowledge in the Internet of Things.
Kang, M., Leng, X., Lin, Z., Ji, K., 2017. A modified faster R-CNN based on CFAR algorithm for SAR ship detection. In: 2017 International Workshop on Remote Sensing with Intelligent Processing. RSIP, pp. 1–4.
Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.
Krizhevsky, A., Hinton, G., 2009. Learning Multiple Layers of Features from Tiny Images. Technical Report 0, University of Toronto, Toronto, Ontario.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 1097–1105.
Küçük, Ç., Taşkın, G., Erten, E., 2016. Paddy-rice phenology classification based on machine-learning methods using multitemporal co-polar X-band SAR images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (6), 2509–2519.
Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Ann. Math. Stat. 22 (1), 79–86.
Laboratory, A.F.R., 2021. MSTAR public targets. URL: https://www.sdms.afrl.af.mil/index.php?collection=mstar. Last accessed 2021-07-22.
Lattari, F., Gonzalez Leon, B., Asaro, F., Rucci, A., Prati, C., Matteucci, M., 2019. Deep learning for SAR image despeckling. Remote Sens. 11, 1532.
Lawrence, S., Giles, C., Tsoi, A.C., Back, A., 1997. Face recognition: a convolutional neural-network approach. IEEE Trans. Neural Netw. 8 (1), 98–113.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324.
Li, B., Liu, B., Huang, L., Guo, W., Zhang, Z., Yu, W., 2017a. OpenSARShip 2.0: A large-volume dataset for deeper interpretation of ship targets in Sentinel-1 imagery. In: 2017 SAR in Big Data Era: Models, Methods and Applications. BIGSARDATA, pp. 1–5. http://dx.doi.org/10.1109/BIGSARDATA.2017.8124929.
Li, Y., Wang, J., Xu, Y., Li, H., Miao, Z., Zhang, Y., 2017b. DeepSAR-Net: Deep convolutional neural networks for SAR target recognition. In: 2017 IEEE 2nd International Conference on Big Data Analysis. ICBDA, pp. 740–743. http://dx.doi.org/10.1109/ICBDA.2017.8078734.
Li, C., Xu, C., Gui, C., Fox, M.D., 2010. Distance regularized level set evolution and its application to image segmentation. IEEE Trans. Image Process. 19 (12), 3243–3254.
Lin, Z., Ji, K., Leng, X., Kuang, G., 2019. Squeeze and excitation rank faster R-CNN for ship detection in SAR images. IEEE Geosci. Remote Sens. Lett. 16 (5), 751–755.
Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C., 2016. SSD: Single shot MultiBox detector. In: Computer Vision – ECCV 2016. Springer International Publishing, Cham, pp. 21–37.
Liu, X., Jiao, L., Liu, F., 2019b. PolSF: PolSAR image dataset on San Francisco. arXiv:1912.07259.
Liu, S., Liu, T., Gao, L., Li, H., Hu, Q., Zhao, J., Wang, C., 2019a. Convolutional neural network and guided filtering for SAR image denoising. Remote Sens. 11 (6).
Liu, Y., Wu, L., 2016. Geological disaster recognition on optical remote sensing images using deep learning. Procedia Comput. Sci. 91, 566–575.
Ma, M., Chen, J., Liu, W., Yang, W., 2018. Ship classification and detection based on CNN using GF-3 SAR images. Remote Sens. 10 (12), 20–43.
Marino, A., Sanjuan-Ferrer, M.J., Hajnsek, I., Ouchi, K., 2015. Ship detection with spectral analysis of synthetic aperture radar: A comparison of new and well-known algorithms. Remote Sens. 7 (5), 5416–5439.
Mercier, G., Girard-Ardhuin, F., 2005. Oil slick detection by SAR imagery using support vector machines. In: Europe Oceans 2005, Vol. 1. IEEE, pp. 90–95.
Meyer, F.J., Nicoll, J.B., Doulgeris, A.P., 2013. Correction and characterization of radio frequency interference signatures in L-band synthetic aperture radar data. IEEE Trans. Geosci. Remote Sens. 51 (10), 4961–4972.
Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., Hossard, L., 2018. Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for Camargue, France. Remote Sens. 10 (8), http://dx.doi.org/10.3390/rs10081217, URL: https://www.mdpi.com/2072-4292/10/8/1217.
Parikh, H., Patel, S., Patel, V., 2020. Classification of SAR and polsar images using deep learning: a review. Int. J. Image Data Fusion 11 (1), 1–32.
Passah, A., Amitab, K., Kandar, D., 2021. SAR image despeckling using deep CNN. IET Image Process. 15 (6), 1285–1297.
Patnaik, R., Casasent, D., 2005. MINACE filter classification algorithms for ATR using MSTAR data. In: Automatic Target Recognition XV, Vol. 5807. International Society for Optics and Photonics, pp. 100–112.
Pei, J., Huang, Y., Huo, W., Zhang, Y., Yang, J., Yeo, T., 2018. SAR automatic target recognition based on multiview deep learning framework. IEEE Trans. Geosci. Remote Sens. 56 (4), 2196–2210.
Qin, C., Schlemper, J., Caballero, J., Price, A.N., Hajnal, J.V., Rueckert, D., 2019. Convolutional recurrent neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging 38 (1), 280–290.
Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 779–788.
for Remote Imaging Sensing, C., (CRISP), P., 2001. Microwave remote sensing. URL: https://crisp.nus.edu.sg/research/tutorial/mw.htm. Last accessed 2021-06-10.
Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1. NIPS '15, MIT Press, Cambridge, MA, USA, pp. 91–99.
Rosen, P.A., Hensley, S., Le, C., 2008. Observations and mitigation of RFI in ALOS PALSAR SAR data: Implications for the desdyni mission. In: 2008 IEEE Radar Conference. pp. 1–6.
Ross, T.D., Worrell, S.W., Velten, V.J., Mossing, J.C., Bryant, M.L., 1998. Standard SAR ATR evaluation experiments using the MSTAR public release data set. In: Zelnio, E.G. (Ed.), Algorithms for Synthetic Aperture Radar Imagery V, Vol. 3370. International Society for Optics and Photonics. SPIE, pp. 566–573. http://dx.doi.org/10.1117/12.321859.
Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. (IJCV) 115 (3), 211–252.
Saha, S., 2018. A comprehensive guide to convolutional neural networks. URL: https://towardsdatascience.com/a-comprehensive-guide-toconvolutional-neural-networks-the-eli5-way-3bd2b1164a53. Last accessed 2021-08-27.
Schmitt, M., Hughes, L.H., Qiu, C., Zhu, X.X., 2019. SEN12MS–A curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion. arXiv preprint arXiv:1906.07789.
Schmitt, M., Hughes, L.H., Zhu, X.X., 2018. The SEN1-2 dataset for deep learning in SAR-optical data fusion. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. IV-1, 141–146. http://dx.doi.org/10.5194/isprs-annals-IV-1-141-2018, URL: https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/IV-1/141/2018/.

Schwegmann, C.P., Kleynhans, W., Salmon, B.P., 2015. Manifold adaptation for constant false alarm rate ship detection in South African oceans. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (7), 3329–3337.
Shang, R., Qi, L., Jiao, L., Stolkin, R., Li, Y., 2014. Change detection in SAR images by artificial immune multi-objective clustering. Eng. Appl. Artif. Intell. 31, 53–67, Special Issue: Advances in Evolutionary Optimization Based Image Processing.
Shang, R., Wang, J., Jiao, L., Stolkin, R., Hou, B., Li, Y., 2018. SAR targets classification based on deep memory convolution neural networks and transfer parameters. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (8), 2834–2846.
Shermeyer, J., Hogan, D., Brown, J., Etten, A.V., Weir, N., Pacifici, F., Hansch, R., Bastidas, A., Soenen, S., Bacastow, T., Lewis, R., 2020. SpaceNet 6: Multi-sensor all weather mapping dataset. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. CVPRW, IEEE Computer Society, Los Alamitos, CA, USA, pp. 768–777. http://dx.doi.org/10.1109/CVPRW50498.2020.00106, URL: https://doi.ieeecomputersociety.org/10.1109/CVPRW50498.2020.00106.
Shimada, M., Isoguchi, O., Tadono, T., Isono, K., 2009. PALSAR radiometric and geometric calibration. IEEE Trans. Geosci. Remote Sens. 47 (12), 3915–3932.
Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556.
Singh, P., Diwakar, M., Shankar, A., Shree, R., Kumar, M., 2021. A review on SAR image and its despeckling. Arch. Comput. Methods Eng. 28 (7), 4633–4653.
Smith, S.W., 1997. The Scientist and Engineer's Guide to Digital Signal Processing. California Technical Publishing, California, URL: http://www.dspguide.com. Last accessed 2021-7-25.
Solberg, A.H.S., 2012. Remote sensing of ocean oil-spill pollution. Proc. IEEE 100 (10), 2931–2945.
Solberg, A., Storvik, G., Solberg, R., Volden, E., 1999. Automatic detection of oil spills in ERS SAR images. IEEE Trans. Geosci. Remote Sens. 37 (4), 1916–1924.
Soldin, R.J., 2018. SAR target recognition with deep learning. In: 2018 IEEE Applied Imagery Pattern Recognition Workshop. AIPR, pp. 1–8. http://dx.doi.org/10.1109/AIPR.2018.8707419.
Space, C., 2019. Capella space. URL: https://www.capellaspace.com. Last accessed 2021-08-25.
Srinivas, U., Monga, V., Raj, R.G., 2014. SAR automatic target recognition using discriminative graphical models. IEEE Trans. Aerosp. Electron. Syst. 50 (1), 591–606.
Steiniger, Y., Kraus, D., Meisen, T., 2022. Survey on deep learning based computer vision for sonar imagery. Eng. Appl. Artif. Intell. 114, 105157.
Sugimoto, M., Ouchi, K., Nakamura, Y., 2013. On the novel use of model-based decomposition in SAR polarimetry for target detection on the sea. Remote Sens. Lett. 4 (9), 843–852. http://dx.doi.org/10.1080/2150704X.2013.804220, arXiv: https://doi.org/10.1080/2150704X.2013.804220.
Sun, Y., Hua, Y., Mou, L., Zhu, X.X., 2021. Cg-net: Conditional gis-aware network for individual building segmentation in vhr sar images. IEEE Trans. Geosci. Remote Sens. 60, 1–15.
Szegedy, C., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 1–9.
Tait, P., 2005. Introduction to Radar Target Recognition, Vol. 18. IET, United Kingdom.
Tan, M., Le, Q., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning. PMLR, pp. 6105–6114.
Tang, Y., Sun, Y., Chiang, J.Y., Li, X., 2021. Research on multiple-instance learning for tongue coating classification. IEEE Access 9, 66361–66370.
Terrie, G., 2018. SAR presents Opportunities, Challenges for GEOINT. URL: http://trajectorymagazine.com/sar-presents-opportunities-challenges-for-geoint/. Last accessed 2021-08-25.
Tomiyasu, K., 1978. Tutorial review of synthetic-aperture radar (SAR) with applications to imaging of the ocean surface. Proc. IEEE 66 (5), 563–583.
Ulaby, F., Moore, R., Fung, A., 1981. Microwave Remote Sensing: Microwave Remote Sensing Fundamentals and Radiometry, Vol. 1. Addison-Wesley Publishing Company, Advanced Book Program/World Science Division, University of Minnesota.
Van Wimersma Greidanus, H., 2008. Satellite Imaging for Maritime Surveillance of the European Seas. Springer Science+Business Media B.V., Berlin (Germany), pp. 343–358, URL: http://www.springer.com/geosciences/book/978-1-4020-6771-6.
Vogado, L.H., Veras, R.M., Araujo, F.H., Silva, R.R., Aires, K.R., 2018. Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification. Eng. Appl. Artif. Intell. 72, 415–422.
Walmsley, M., Smith, L., Lintott, C., Gal, Y., Bamford, S., Dickinson, H., Fortson, L., Kruk, S., Masters, K., Scarlata, C., et al., 2020. Galaxy Zoo: probabilistic morphology through Bayesian CNNs and active learning. Mon. Not. R. Astron. Soc. 491 (2), 1554–1574.
Wang, Z., Du, L., Mao, J., Liu, B., Yang, D., 2019. SAR target detection based on SSD with data augmentation and transfer learning. IEEE Geosci. Remote Sens. Lett. 16 (1), 150–154.
Wang, R., Li, J., Duan, Y., Cao, H., Zhao, Y., 2018. Study on the combined application of CFAR and deep learning in ship detection. J. Indian Soc. Remote Sens. 46, 1413–1421.
Wang, R., Li, J., Hao, Y., Li, J., 2017a. An adaptive and fast CFAR algorithm based on multithreading for ship detection in SAR image. Int. J. Hybrid Inf. Technol. 10 (8), 67–78.
Wang, R., Pei, J., Xu, F., Huo, W., Huang, Y., Zhang, Y., Yang, J., 2021. A multidirectional one-dimensional scanning method for harbor detection from SAR images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 10003–10016.
Wang, N., Shi, G., Liu, L., Zhao, L., Kuang, G., 2012. Polarimetric SAR target detection using the reflection symmetry. IEEE Geosci. Remote Sens. Lett. 9 (6), 1104–1108. http://dx.doi.org/10.1109/LGRS.2012.2189548.
Wang, P., Sun, X., Diao, W., Fu, K., 2020. FMSSD: Feature-merged single-shot detection for multiscale objects in large-scale remote sensing imagery. IEEE Trans. Geosci. Remote Sens. 58 (5), 3377–3390.
Wang, Y., Wang, C., Zhang, H., Dong, Y., Wei, S., 2019. A SAR dataset of ship detection for deep learning under complex backgrounds. Remote Sens. 11 (7), http://dx.doi.org/10.3390/rs11070765, URL: https://www.mdpi.com/2072-4292/11/7/765.
Wang, P., Zhang, H., Patel, V.M., 2017. SAR image despeckling using a convolutional neural network. IEEE Signal Process. Lett. 24 (12), 1763–1767.
Wang, C., Zhang, H., Wang, Y., Zhang, B., 2018a. Sea ice classification with convolutional neural networks using sentinel-l scansar images. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. pp. 7125–7128.
Wang, J., Zheng, T., Lei, P., Bai, X., 2018b. Ground target classification in noisy SAR images using convolutional neural networks. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (11), 4180–4192.
Wang, Y., Zhu, X.X., Zeisl, B., Pollefeys, M., 2017b. Fusing meter-resolution 4-D InSAR point clouds and optical images for semantic urban infrastructure monitoring. IEEE Trans. Geosci. Remote Sens. 55 (1), 14–26. http://dx.doi.org/10.1109/TGRS.2016.2554563.
Wei, D., Du, Y., Du, L., Li, L., 2021. Target detection network for SAR images based on semi-supervised learning and attention mechanism. Remote Sens. 13 (14), URL: https://www.mdpi.com/2072-4292/13/14/2686.
Wolff, C., 2019. Radar tutorials. URL: http://www.radartutorial.eu/20.airborne/ab07.de.html. Last accessed 2021-08-28.
Xiang, Y., Wang, F., You, H., 2018. OS-SIFT: A robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas. IEEE Trans. Geosci. Remote Sens. 56 (6), 3078–3090.
Xu, Y., Zhao, B., Zhai, Y., Chen, Q., Zhou, Y., 2021. Maize diseases identification method based on multi-scale convolutional global pooling neural network. IEEE Access 9, 27959–27970.
Yu, H., Wang, C., Li, J., Sui, Y., 2021. Automatic extraction of green tide from GF-3 SAR images based on feature selection and deep learning. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 10598–10613.
Yu, W., Wang, Y., Liu, H., He, J., 2016. Superpixel-based CFAR target detection for high-resolution SAR images. IEEE Geosci. Remote Sens. Lett. 13 (5), 730–734. http://dx.doi.org/10.1109/LGRS.2016.2540809.
Yuan, J., Xiong, H.-C., Xiao, Y., Guan, W., Wang, M., Hong, R., Li, Z.-Y., 2019. Gated CNN: Integrating multi-scale feature layers for object detection. Pattern Recognit. 105, 107–131.
Zeiler, M.D., Fergus, R., 2014. Visualizing and understanding convolutional networks. In: European Conference on Computer Vision. Springer, pp. 818–833.
Zeng, T., So, H.K.-H., Lam, E.Y., 2019. Computational image speckle suppression using block matching and machine learning. Appl. Opt. 58 (7), B39–B45.
Zhan, R.-h., Tian, Z.-z., Hu, J.-m., Zhang, J., 2016. SAR automatic target recognition based on deep convolutional neural networks. DEStech Trans. Comput. Sci. Eng. (aita).
Zhang, F., Du, B., Zhang, L., Xu, M., 2016. Weakly supervised learning based on coupled convolutional neural networks for aircraft detection. IEEE Trans. Geosci. Remote Sens. 54 (9), 5553–5563.
Zhang, X.-Z., Huang, P.-K., 2013. SAR ATR based on Bayesian compressive sensing. Syst. Eng. Electron. 35 (1), 40–44.
Zhang, Q., Yuan, Q., Li, J., Yang, Z., Ma, X., 2018. Learning a dilated residual network for SAR image despeckling. Remote Sens. 10 (2), 196.
Zhao, J., Zhang, Z., Yao, W., Datcu, M., Xiong, H., Yu, W., 2020. OpenSARUrban: A Sentinel-1 SAR image dataset for urban interpretation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 187–203. http://dx.doi.org/10.1109/JSTARS.2019.2954850.
Zheng, S., Song, Y., Leung, T., Goodfellow, I., 2016. Improving the robustness of deep neural networks via stability training.
Zhong, C., Mu, X., He, X., Zhan, B., Niu, B., 2018. Classification for SAR scene matching areas based on convolutional neural networks. IEEE Geosci. Remote Sens. Lett. 15 (9), 1377–1381.
Zhu, M., 2004. Recall, Precision and Average Precision, Vol. 2. Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, p. 6.
Zhu, X.X., Hu, J., Qiu, C., Shi, Y., Kang, J., Mou, L., Bagheri, H., Haberle, M., Hua, Y., Huang, R., Hughes, L., Li, H., Sun, Y., Zhang, G., Han, S., Schmitt, M., Wang, Y., 2020. So2Sat LCZ42: A benchmark data set for the classification of global local climate zones [software and data sets]. IEEE Geosci. Remote Sens. Mag. 8 (3), 76–89. http://dx.doi.org/10.1109/MGRS.2020.2964708.
Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geosci. Remote Sens. Mag. 5 (4), 8–36.