

Explainable, Physics Aware, Trustworthy AI Paradigm Shift for Synthetic Aperture Radar

Mihai Datcu, *Fellow, IEEE*, Zhongling Huang[†], Andrei Anghel, Juanping Zhao, Remus Cacoveanu

Abstract—The recognition or understanding of the scenes observed with a SAR system requires a broader range of cues, beyond the spatial context. These encompass but are not limited to: imaging geometry, imaging mode, properties of the Fourier spectrum of the images or the behavior of the polarimetric signatures. In this paper, we propose a change of paradigm for explainability in data science for the case of Synthetic Aperture Radar (SAR) data to ground the explainable AI for SAR. It aims to use explainable data transformations based on well-established models to generate inputs for AI methods, to provide knowledgeable feedback for training process, and to learn or improve high-complexity unknown or un-formalized models from the data. At first, we introduce a representation of the SAR system with physical layers: i) instrument and platform, ii) imaging formation, iii) scattering signatures and objects, that can be integrated with an AI model for hybrid modeling. Successively, some illustrative examples are presented to demonstrate how to achieve hybrid modeling for SAR image understanding. The perspective of trustworthy model and supplementary explanations are discussed later. Finally, we draw the conclusion and we deem the proposed concept has applicability to the entire class of coherent imaging sensors and other computational imaging systems.

Index Terms—SAR image understanding, explainable artificial intelligence, deep neural networks, knowledge inspired data science

I. MOTIVATION AND SIGNIFICANCE

The Earth is facing unprecedented climatic, geomorphologic, environmental or anthropogenic changes, which require global scale, long term observation with Earth Observation (EO) sensors. SAR sensors, due to their observation capability during day and night and independence on atmospheric effects, are the only EO technology to insure global and continuous observations. Meanwhile, the SAR observations of Sentinel-1 satellites in the frame of the European Copernicus program, are worldwide freely and openly accessible. This is immensely enlarging the SAR Data Science and applications, covering a multitude of areas as: urbanization, agriculture, forestry, geology, tectonics, oceanography, polar surveys, or biomass estimation, only to enumerate a few. Copernicus Open Access Hub provides more than 457.59 PB data of satellites covering the Earth for more than 570,000 users all around the world.¹

SAR is a pioneer technology in the field of computational sensing and imaging, of which the imaging mechanism is totally different from optical sensors. A radar instrument carried by an airborne or spaceborne platform illuminates the scene by side-looking or forward-looking, which allows to discriminate objects in the range direction. As the platform moving along

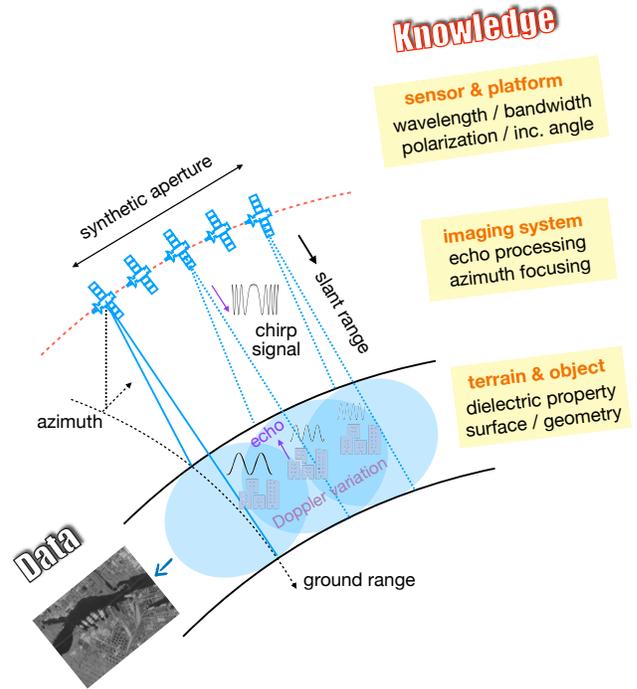


Fig. 1: A simple illustration of how SAR images the world (Stripmap Mode). SAR society is facing the big data challenge but with limited ground truth. In the meanwhile, the knowledge of SAR is equally important. This is also the motivation of the physical layers in this paper.

its track, the SAR sensor is constantly transmitting a sequence of chirp signals and receiving echos reflected from objects on the ground, as depicted in Fig. 1. When recording all individual acquisitions with a short physical antenna and mathematically combining them into a synthetic image, a much larger synthesized aperture is formed. This allows high capacity to distinguish objects in azimuth despite a physically small antenna [1]. A high resolution “image” can be processed by applying SAR focusing principle, e.g., matching filtering [2].

A deluge of SAR sensors have increased the data availability for various SAR applications. The allure of data-driven learning stems from the ability of automatically extracting abstract features from large data volumes [3]–[6], and therefore, many deep learning studies for SAR applications have been developed in recent years [7]–[10]. Current popular paradigm predominantly follows the blue part in Fig. 2 (a), where SAR image data is all that is required to operate an intelligent

[†] Corresponding author (huangzhongling@nwpu.edu.cn)

¹ <https://scihub.copernicus.eu/reportsandstats/>

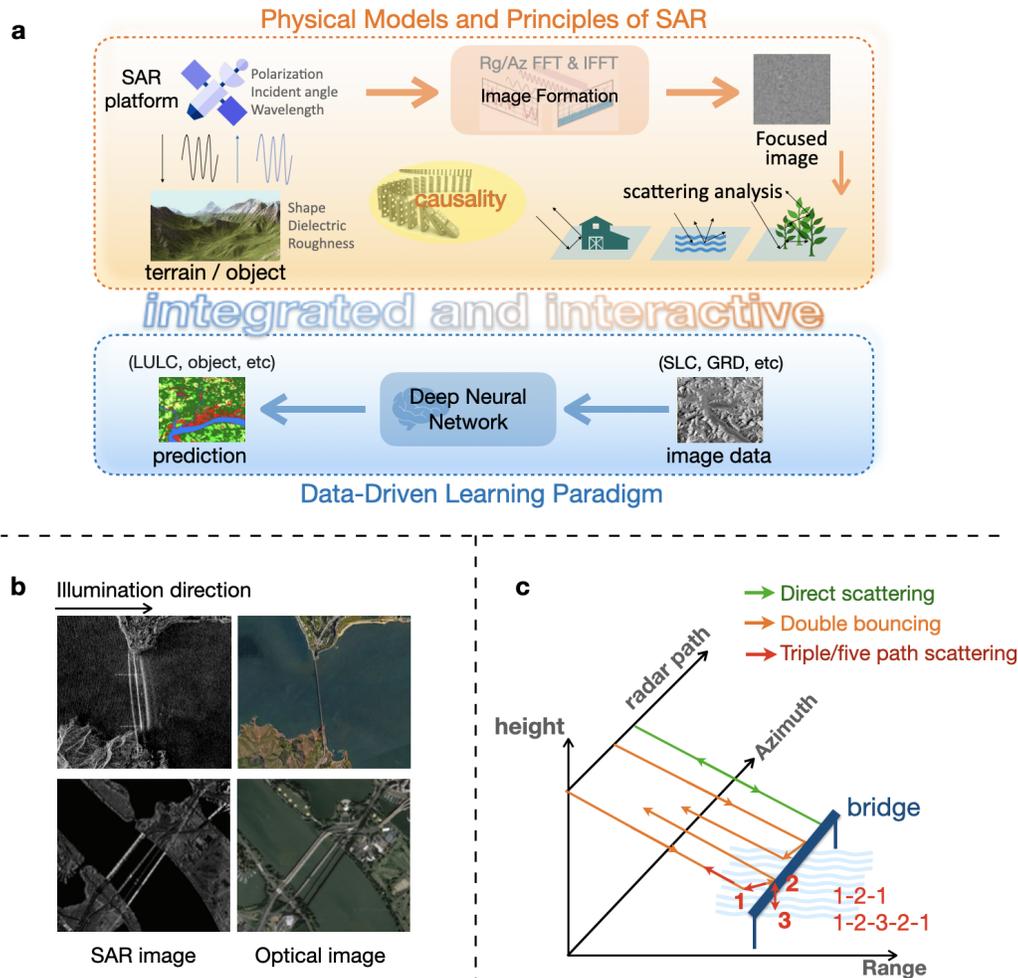


Fig. 2: **a** The conventional data-driven paradigm for intelligent SAR image understanding based on deep neural networks and the proposed paradigm shift integrated and interacted with physical knowledge of SAR. **b** A bridge can be imaged as multiple bright lines, similar as a couple of bridges imaged in the other SAR image, depending on the observation parameters and orientations. This positions the load and outmost difficulty of SAR image understanding. **c** The multipath scattering formation in the SAR image.

network. In addition to data, however, the physical model and principles of SAR sensor should not be neglected. In the upper example of Fig. 2 (b), A bridge over a placid river that is illuminated perpendicular to its primary orientation appears as many brilliant lines, resembling the lower SAR image in which several bridges are imaged from a different viewing angle. The phenomena can be explicable by multi-path scattering [11], [12], as illustrated in Fig. 2 (c). Apart from the direct scattering from the bridge, the double bounce reflection between the bridge and water or vice versa occurs at the corner reflector spanned from the smooth vertical bridge facets facing the sensor and the water surface. In addition, the triple-bounce reflection and maybe some five-path scattering would happen between the horizontal plane of bridge and water surface. Thus, SAR image implies the causality of multi-path scattering phenomena and object characteristics. This positions the load of SAR image understanding, and the outmost challenge of data science, as new and particular paradigm of Artificial Intelligence (AI).

So far, some researches have discussed the paradigm that attempts to bring scientific knowledge and data science models together, applied to a broad range of research themes such as partial differential equation solving [13] and Earth sciences [14]–[16]. In particular for SAR community, however, this topic has rarely been systematically analyzed and illustrated. Thus, we aim to prospect the hybrid modeling paradigm for intelligent SAR image understanding, where deep learning is integrated and interacted with SAR physical models and principles, to achieve explainability, physics awareness, and trustworthiness.

Explainable AI is a broad concept. A scientific understanding of explainability is the capacity to clarify the results in the context of domain knowledge. The algorithms still remain a black box. A different approach is the algorithmic explainability. This is constructed such that the results of the used model can be described algorithmically. To obtain a higher degree of explainability, we aim at the synergy of the paradigms: *algorithmic and scientific explainability*.

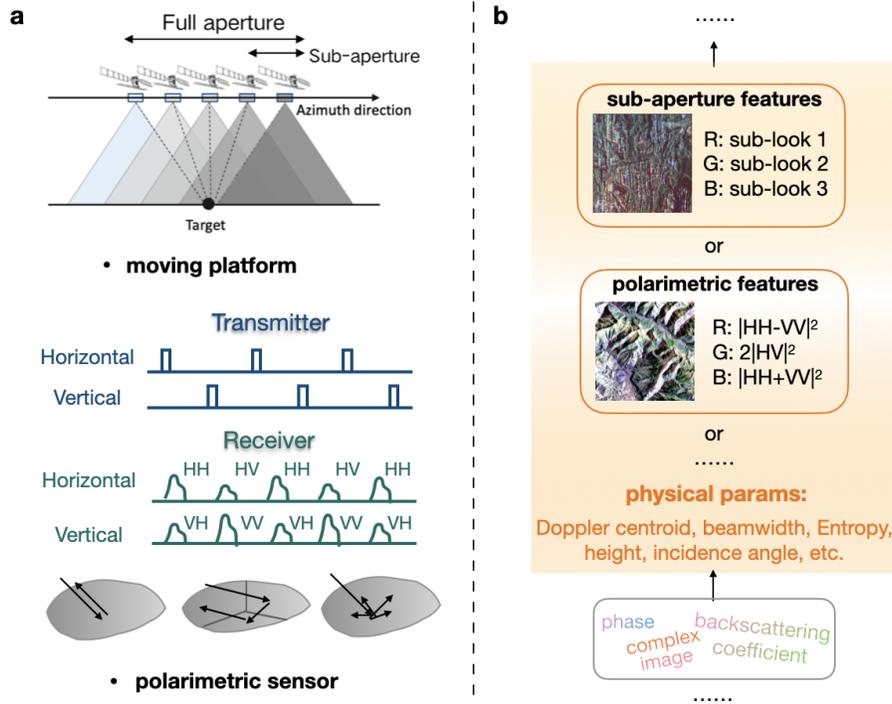


Fig. 3: Physical layer (i): Sensor and Platform. **a**: The moving platform creates Doppler variations and synthesizes large virtual aperture; PolSAR transmits and receives diverse polarized wave, and SAR polarimetric characteristics are depicted. **b**: Based on the physics behind the platform and sensor, the physical layer produces SAR specific representations such as sub-aperture synthesis image and polarimetric feature, with specified physical parameters.

Algorithmic explainability lies in the guarantee of transparency to understand how the machine learning algorithm works by participation of SAR physical models and principles. Scientific explainability ensures the physical consistency of AI output, as well as learning of trustworthy results with physical meaning.

To ground this, we first lay out a representation of SAR physical layer in the context of SAR domain knowledge, as presented in Section II. Further, we describe how to integrate and interact them with popular neural networks to build a hybrid and translucent model for SAR applications using illustrative examples, demonstrated in Section III. The perspective of trustworthy models and supplementary explanation for SAR community are discussed in Section IV and V. The conclusion and perspectives are finally given in Section VI.

II. SAR PHYSICAL LAYERS

Other than the neural network layers equipped with a number of learnable parameters, SAR physical layers are ones embedded with physical knowledge of SAR, well-established, interpretable, and supported by domain theories. The concept of "physical layer" apart from "neural network layer" arose in literature [16] to make the model more physically realistic. As motivated in Fig. 1, three SAR physical layers are highlighted specific for SAR applications in this paper, i.e., (i) sensor and platform: referring to antenna characteristics and moving satellite/aircraft, (ii) imaging system: figuring image formation with focusing process and (iii) scattering signature: reflecting the physical properties of terrain and objects.

A. Sensor and Platform

Fig. 3 demonstrates the physical layer of sensor and platform that indicates the physics behind the SAR acquisition principle, such as aperture synthesizing with moving platform and various characteristics of antenna.

Existing spaceborne EO SAR missions work in a monostatic or quasi-monostatic configuration. The simplest illumination mode of a SAR system is the stripmap mode in which the antenna pointing direction is constant throughout the acquisition, as shown in Fig. 3 **a**. The moving platform leads to a sliding Doppler spectrum that impacts the complex SAR image. Knowing the behaviour of the Doppler centroid to create sub-looks is essential for exploiting look angle diversity of the input data, especially for very high-resolution SAR images.

It is well-known that in high-resolution SAR image where the signals are performed over a broad bandwidth and wide angular aperture, the targets are no longer isotropic and non-dispersive. Instead, it is more plausible to infer that the target's backscattering is dependent on illumination angle and frequencies [17]. The sub-aperture processing can be applied to analyze the target scattering variations. Fig. 3 **b** gives an example of a synthesized pseudo color SAR image via sub-aperture processing. The complex-valued SAR image is first transformed to the azimuth spectral domain by a one-dimensional Fourier transform. Then, the full Doppler spectrum is equally split into three intervals, named sub-apertures or sub-looks, each containing 1/3 range of azimuth

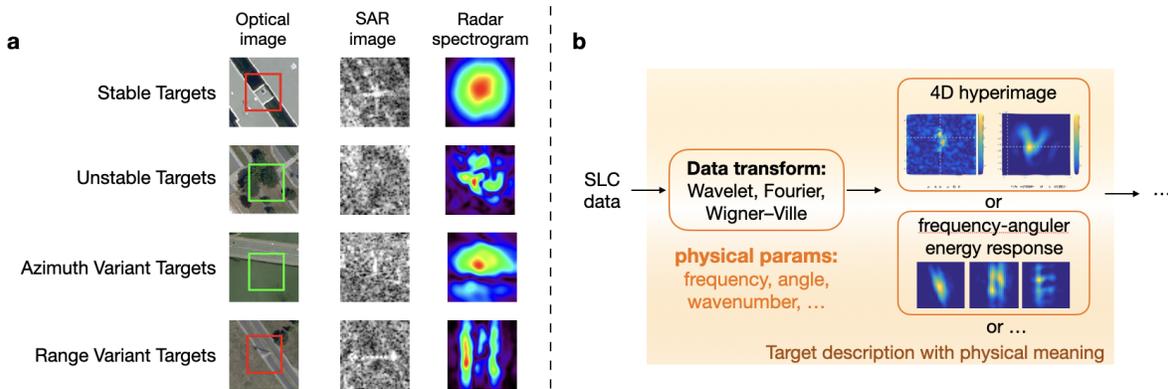


Fig. 4: Physical layer (ii): Image Formation. **a.** Targets are characterized by sliding bandpass filtering in the Fourier domain. **b.** On the basis of image formation principle and target scattering model, the physical layer generates the rich target description with physical meaning.

angles. Finally, the three sub-apertures are transformed back to time-domain using an inverse Fourier transform, coded as the R, G, and B channels, respectively. Red, Green, and Blue targets respond mainly on the first, second, and third sub-looks, respectively, whilst gray targets indicate that they respond equivalently in different sub-looks. The pseudo-colored image well demonstrates the particular behavior of some targets. Given the precise knowledge of the parameters related to Doppler variations (e.g., orbit, azimuth steering rate, radiation pattern, incidence angle), the physical layer can generate sub-look data deterministically and there is no need to design a neural network that should learn to create sub-looks from various types of SAR training data.

Sensor characteristics, such as polarization, interferometry and tomography, construct physical layer as well. Fig. 3 **b** presents a Pauli pseudo RGB image, where R, G, and B channels are formed with $|HH - VV|^2$, $2|HV|^2$, and $|HH + VV|^2$, respectively, indicating the polarimetric relation. Several physical layers can be stacked to represent rich physics of SAR sensor and platform. Early in literature [18], the diversity in the polarimetric features with the azimuthal look angle was exploited. Thus, the moving platform and polarimetric sensor are both characterized. Similarly, the stacked physical layers can represent polarimetric and interferometric properties of PolInSAR data, or any other combinations.

B. Imaging System

The second physical layer we suppose delineates the physics behind SAR image formation with an imaging system. The selected exemplars are illustrated in Fig. 4.

A pulse-based radar or a frequency modulated continuous wave (FMCW) radar is usually used in a SAR system, where a range profile is obtained for each transmitted/received waveform, either by range compression in the case of a pulse-based radar or by applying a Fourier transform to the beat signal in the case of an FMCW radar [19]. By a coherent processing of the range profiles, the azimuth focusing process outputs a SAR image representing a two-dimensional complex reflectivity map of the illuminated area. SAR processing,

taking a simple point target as example, aims to collect the dispersed signal energy in range and azimuth into a single pixel. Many traditional imaging algorithms are in terms of a Fourier synthesis framework [20], as such, Fourier transform provides a specific physical meaning for SAR image. This kind of physical layer assists AI model to better depict the target scattering beyond the "image" domain.

Fig. 4 (a) first shows a simple time-frequency analysis of target with short-time Fourier transform [21], [22], characterizing the backscattering intensity variations in 2-D range and azimuth frequency domain. Four kinds of backscattering behaviors observed in SAR were defined in literature [23], related to different objects shown in Fig. 4. In the high-resolution case (wide bandwidth chirp signal and broad angular aperture), the complex amplitude of a target is frequency and aspect dependent [17]. Thus, the image formation can be extended to four dimension (called hyperimage) with wavelet transform, providing a concise physically relevant description of target scattering. This frequency and angular energy response pattern is proved useful for discriminating different scatterers, offering valuable prior information to AI model, depicted in Fig. 4 **b**.

C. Scattering Signatures of Objects

Thirdly, we introduce the physical layer regarding the scattering signatures of objects, in which the causality of target characteristics and scattering behaviors is involved.

For optical images, what you see is what you receive, that is, the objects depicted on the optical image are in accord with human cognition. Targets in SAR images are reflected by scattering characteristics, yet they include a wealth of physical information that the human eye cannot immediately identify. Fig. 5 **a** shows example of two typical SAR targets of bridge and building. The scattering phenomenon that shows several parallel lines over the river can be interpreted as single, double, and multiple scattering of the bridge based on the domain knowledge. The building, with scattering signatures of layover, shadow, single and secondary scattering in the high-resolution SAR image, can also be reflected as only layover and shadow [24], depending on the building orientation and shape. Similar

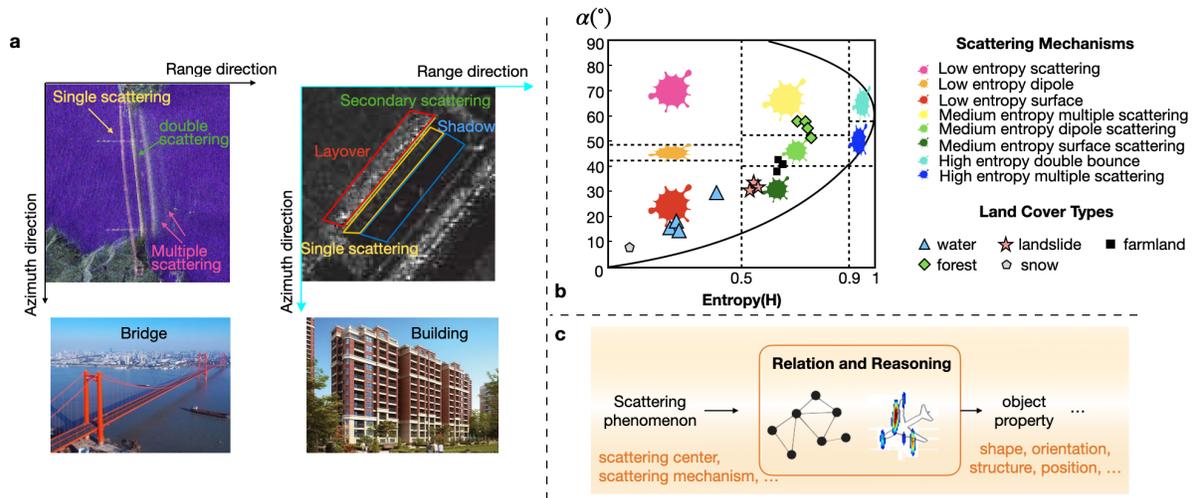


Fig. 5: Physical layer (iii): Scattering Signatures of Objects. **a.** The Golden Gate Bridge revealing multipath scattering characteristics in a Gaofen-3 quad-pol SAR image, and a typical single building representing different scattering regions in a high-resolution (1m) SAR image [24]. **b.** The scattering mechanisms indicated by the H/α plane for full-polarized SAR data point out the land-use and land-cover classes [25]. **c.** The physical layer describes the relationship and reasoning between the scattering characteristics seen in the SAR image and the object’s features, such as its shape, structure, or even semantics.

research by Ferro et al. [26] investigated the relationship between double bounce and the orientation of buildings in VHR SAR images. Fig. 5 **b** demonstrates the relations between the scattering mechanism of H/α plane and the semantics of land-cover and land-use classes [25]. Likewise, one can deduce the scattering center position and the specific shape of distributed target from a SAR image by applying some scattering models [27].

The conventional data-driven convolutional neural network can capture the image contents as we “see” in the SAR image, whereas it is not equipped with the ability to “interpret” the scattering phenomenon as we discussed before. This indicates the knowledge gap between SAR scattering signatures and human vision cognition. The physical layer delivering semantic understanding behind the SAR scattering signature permits a more thorough interpretation of the SAR image. As shown in Fig. 5 **c**, the physical layer defines the association between the scattering characteristics of a SAR image and the object’s qualities, such as shape, structure, or semantics. It can be written as an objective function or a regularization term that constrains the training of neural networks. This will improve the intelligence of AI model to master some causality between scattering signatures and the object nature.

III. HYBRID MODELING WITH SAR PHYSICAL LAYERS

The integration and interaction of neural network layers and physical layers construct the hybrid modeling for SAR image interpretation. In view of algorithmic explainability, the explainable physical models and domain knowledge improves the transparency. For scientific explainability, the hybrid modeling ensures the physical meaning of output in physical layers and the prediction can maintain the physical consistency. In this section, we demonstrate several hybrid modeling

approaches with SAR physical layer to achieve explainability and physics awareness.

A. Insert for Substitution

The introduced physical layer can be inserted in a deep neural network for substitution, extracting explainable and meaningful features, either as the input of a DNN or fused with DNN features in intermediate layers. A common way is to insert a physical layer into the input layer to obtain the polarimetric features for PolSAR image classification, including the elements of coherency matrix, Pauli decomposition features, etc [31], [32]. Similarly, the sub-aperture images are generated as the input for target detection [33]. The other usage of physical layer is for feature fusion, where the features obtained by well-established physical model and deep neural networks are combined [34], [35].

Our recent work, a deep learning framework named Deep SAR-Net (DSN) [28], addressed both aspects that inserts the physical layer into the input and the intermediate position of deep model. As shown in Fig. 6, DSN was proposed for classifying SAR images with complex values. Instead of the entire data-driven method, i.e. the complex-valued convolutional neural networks (CV-CNN), the designed DSN encompasses three shallow neural network modules and two physical layers. The first physical layer generates the high-dimensional radar spectrogram based on time-frequency analysis. The second one handles the features of the 2-D projection along the frequency axes [22] to maintain the location constraint, making it possible to be fused with spatial features from intensity image. DSN outperformed CV-CNN especially with limited labeled training data, and had a remarkable performance in discriminating the man-made target scenes compared with the traditional CNN. It demonstrates the Fourier process on single-look complex SAR image embedded the knowledge like synthesizing the antenna

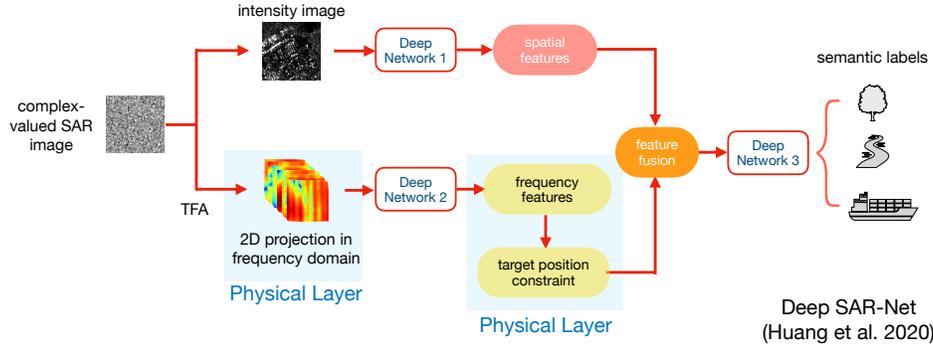


Fig. 6: Our recent work Deep SAR-Net (DSN) [28] for SAR image classification can be regarded as a typical example of inserting the physical layers into a deep model.

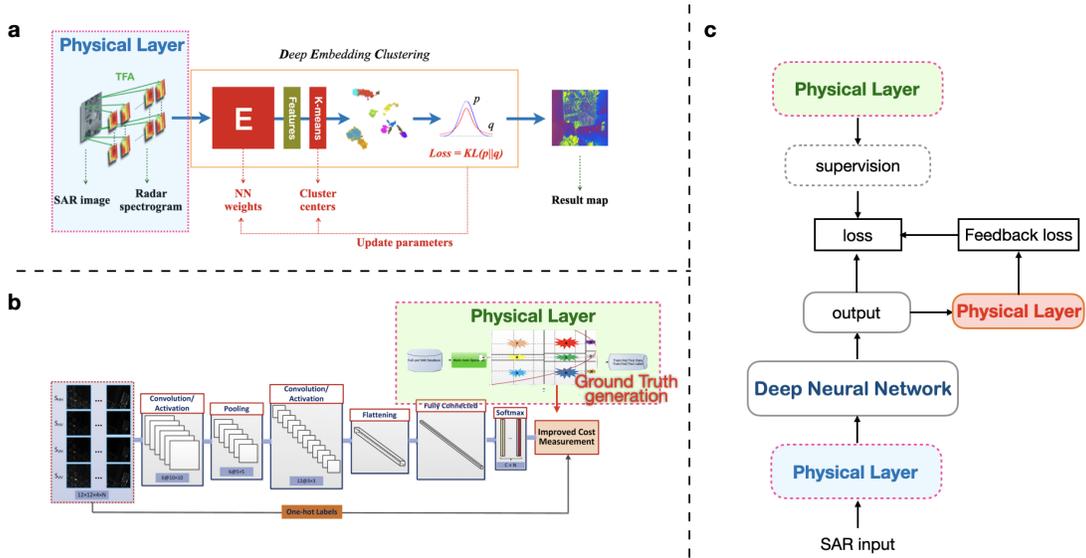


Fig. 7: **a.** Unsupervised HDEC-TFA method [29]. It automatically discovered the radar spectrogram patterns more than the four defined in [23] with deep neural networks. **b.** Learning the polarimetric features from single-polarized SAR image, supervised by Entropy-Alpha-Anisotropy generated from full-pol data [30]. **c.** The physical layers in **a** and **b** play the role of input transform (blue) and supervision generation (green) in hybrid modeling. In addition, the physical layer (red) can act as feedback to restrict learning and produce physically consistent outcomes.

well characterizes the physical property of SAR target, and the usages of physical layer cut down unnecessary parameters in neural network layers to improve the model performance with limited ground truth.

B. Compensation for Imperfect Knowledge with Feedback

In condition of unknown/inconclusive physical models or incomplete knowledge, it is difficult to extract perfect physical parameters or physical scattering characteristics of SAR via model-based methods. For instance, obtaining the polarimetric features from dual-pol, or even single-polarized SAR image. Thus, the physical layer interacted with deep neural network take effect.

1) Target Character Identification:

Some researches have analyzed the energy response pattern in frequency dimensions of target varied in SAR image,

and discussed the nonstationary targets [18], [36]. Spigai et al. [23] pointed out four canonical targets with a rough definition shown in Fig. 4 a. However, it remains unknown for many complicated scene and objects. Fig. 7 show our related work of using physical layer and deep neural network for compensation of imperfect knowledge. The first is the unsupervised hierarchical deep embedding clustering based on time-frequency analysis (HDEC-TFA) [29], which was proposed to automatically characterize the radar spectrogram (or the sub-band scattering pattern defined in [29]) basically in urban area, discovering the various scattering pattern more than the four specific classes defined in [23]. It offered a new perspective to describe the physical properties of single-polarized SAR. Furthermore, we used two stacked physical layers to obtain the polarimetric and time-frequency patterns and analyzed with deep neural network in reference [37].

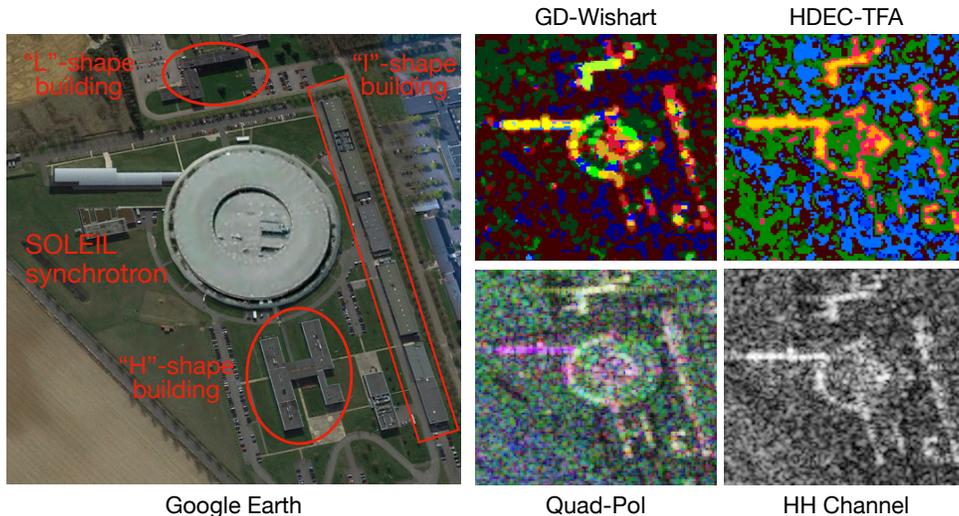


Fig. 8: The SOLEIL synchrotron in France and the surrounding buildings with different shapes are depicted in the Gaofen-3 SAR image. Both the GD-Wishart [38] result on Quad-Polarization SAR data and the HDEC-TFA [29] result on HH channel single-polarized SAR can capture the special scattering characteristics of the objects.

Fig. 8 demonstrates the result compared with the polarimetric physical model. The SOLEIL synchrotron in France, shown as the round building in the Google Earth remote sensing image, is surrounded by three different shapes of buildings. The HDEC-TFA method can capture the special characteristics of the architectures even in single HH channel SAR image, as much as the physical model based method GD-Wishart [38] on quad-pol SAR. Some other man-made targets examples characterized by time-frequency model with neural networks are given in [39]. Our experiments in [29] demonstrated the trained model varies with different imaging conditions since the sub-band scattering pattern is influenced by several imaging parameters, which should be taken into consideration when transferring the AI model to other situations.

2) Polarimetric Parameter Extraction:

By transmitting and receiving waves that are both horizontally and vertically polarized, the full-pol SAR image captures abundant physical characteristics of the imaged objects that can lead to various physical parameters. In contrast, single-pol and dual-pol SAR data are less informative for physical feature extraction. If only one polarization channel is obtained, one cannot derive the other polarization channels in principle. Once the objects are known, i.e., once the characteristics of targets such as geometry, surface roughness, etc. are identified, deep learning can be employed to transfer the knowledge learned from physical models to reconstruct the physical parameters of objects. As shown in Fig. 7 b, Zhao et al. [30] proposed a complex-CNN model to learn physical parameters (entropy H and α angle) with transfer learning from single-pol and dual-pol SAR data, supervised by features obtained with a physical layer. Some similar studies include but not limit to [40], [41]. Song et al. [40] addressed "radar image colorization" issue to reconstruct the polarimetric covariance matrix with a designed deep neural network, where the supervision was also generated with a physical layer.

When training a data-driven deep neural network, some

physical consistencies may not be guaranteed. The authors pointed that the reconstructed covariance matrix may not be positive semi-definite [40], and they proposed an algorithm to correct it. In this case, the additional physical layer embedded prior constraint acts as post-processing to revise the physically inconsistent result of DNNs. Furthermore, this type of physical layer is suggested to provide feedbacks during training, as demonstrated in Fig. 7 c, the red part. The feedback of physical layer aims to prevent the model from learning the physical inconsistency.

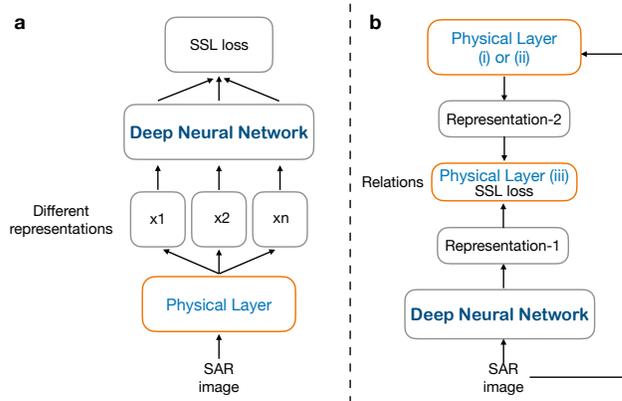


Fig. 9: The SAR physical layer can be integrated in a self-supervised learning framework to guide the neural network training without ground truth. **a**. The physical layer generates various modalities of SAR image using well-established physical models, such as sub-aperture images, different polarimetric features, etc. The self-supervised learning can be conducted with contrastive learning paradigm. **b**. The physical layer produces a physical representation of image, serving as a guided signal that drives the neural network to learn a similar representation.

3) SAR Image Generation/Simulation:

This paradigm can be popularized to other SAR applications. SAR target image generation (or simulation) based on deep generative model (such as variational auto-encoder [44] and generative adversarial network [45]) has attracted much attention in recent years. The generated SAR images are expected to be used as data supplements to support target identification. The authenticity and interpretability of current deep SAR image generation is a substantial obstacle that has a significant impact on subsequent tasks [46]. Many latest studies input important physical parameters into the deep generative model or use them as supervision at the output layer, such as depression angle and target orientation, that facilitated more reliable outcomes [47], [48]. We consider them physical layer as shown in Fig. 7 c, the green and blue part.

Coupling the physical layer as a feedback in neural network for SAR image generation has yet to be explored. When generative model produces a pseudo SAR image, a physical layer will be applied to verify whether it is consistent with the knowledge base of SAR, e.g. physical parameters derived from a well-established model. If not, the current generative model will revise the pseudo result to minimize the inconsistency. There are some examples to learn from in the field of fluid simulation [49], [50]. As such, the physical layer is used for constructing physical inconsistency as a feedback that explicitly constrain the generative model to fulfill some quantitative conditions, so as to guarantee authenticity. Referred to some latest studies in other fields, physical model as a feedback or constraint in the loop of deep learning is also applied to under water image enhancement [51] and seismic impedance inversion [52].

C. Self Supervised Learning Guidance

Self supervised learning has been attracted much attention in recent years, since it can help reduce the required amount of labeling. One can pre-train a model on unlabeled data and fine-tune it on a small labeled dataset. It offers great opportunity for SAR community where big data volume is available while the ground truth is usually difficult to obtain. There is a remarkable potential for SAR physical layer to apply for self-supervised learning.

As shown in Fig. 9, two self-supervised learning paradigms are given. The physical layer helps to establish a pretext task for SAR image. In Fig. 9 a, different SAR image representations are generated by physical layer, for instance, the sub-aperture images, various polarimetric feature images, etc. As similar to SimCLR [53] that conducted the contrastive learning based on data-augmentation, or NPID [54] that learned the optimal feature via instance-level discrimination, the surrogate task can be built to form a self-supervised learning. An illustrative example is in reference [55].

Fig. 9 b illustrates a second line of thought, which we refer to as physics guided learning. Firstly, the physical layer is used for generating meaningful physical representations, like scattering mechanisms (physical layer (i) and (ii) can both achieve this). Meanwhile, the neural network extracts hierarchical spatial features from SAR image. The crucial point is

how to establish a connection between physical properties and image features. We propose to exploit physical layer (iii) to reveal relationships and thereby design an objective function for self-supervised learning.

Our recent work [37], [42], [43] details the paradigm in Fig. 9 b. A physics guided network (PGN) for SAR image feature learning was proposed as shown in Fig. 10. First, a physical layer is deployed at the beginning, where the physical scattering properties are derived. Based on the crucial assumption that SAR image features and the abstract physical scattering mechanisms should share common attributes in semantic level, a surrogate task was established via the other physical layer that defines a loss function. The inspiration is from reference [56], which indicated the abstract topic mixture on scattering properties and the high-level image features are with similar semantics. Thus, we built the relation between the image semantics and SAR scattering characteristics. A novel objective function was designed to instruct self-supervised learning guided by physical scattering mechanisms.

The advantages of this kind of learning paradigm lie in two aspects. First, the training process takes all labeled and unlabeled data so that the learned features generalize well in test set. Second, the guidance of physical information leads to physics awareness of features learned by neural networks, i.e., the DNN feature maintains physical consistency. In a word, the prior physical knowledge is embedded in the neural network. The experiments in [43] verified this quantitatively and qualitatively.

Additionally, the outputs of the physically interpretable deep model can be further explained, which in turn inspires algorithm improvement. We illustrate with an example of sea-ice classification in polar area [42]. The physics guided learning is driven by physical signals that reflect the scattering properties of SAR image. The guided physical signals are visualized with t-sne in Fig. 11, where different colors in (a) represent semantic labels of sea-ice and each color in (b) indicates samples with similar physical scattering properties. One characteristic that can be seen is that young ice and water bodies have extremely similar physical representations, which would impede semantic discrimination. It can explain the physics guided learning result in [42] that about 23% test samples in water bodies class were predicted as young ice. The explanation will motivate us to improve the algorithm by, for instance, relaxing the physical constraints between the two classes.

Similarly, a very recent work [57] was proposed for SAR target recognition inspired by our work [43]. The authors proposed a CNN under the guidance of SAR target physical model, attributed scattering center (ASC), to extract the significant target features, that were successively injected into the classification network for more robust and interpretable results.

IV. TRUSTWORTHY MODELING

A. Why Trustworthy Modeling Needed

The results obtained by applying AI techniques in SAR processing can be validated using in situ measurements of

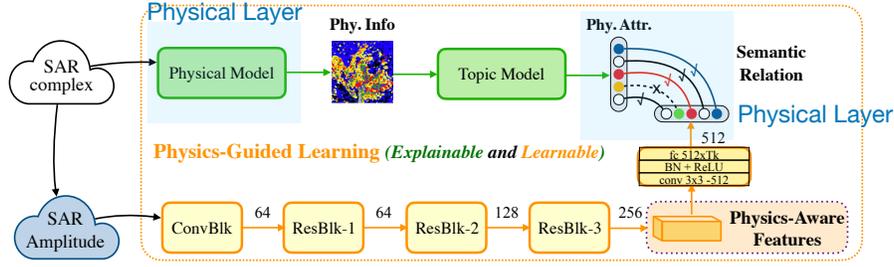


Fig. 10: A physics guided network was proposed [42], [43] where a novel deep learning paradigm and loss function were designed to associate the SAR scattering characteristics with image semantics.

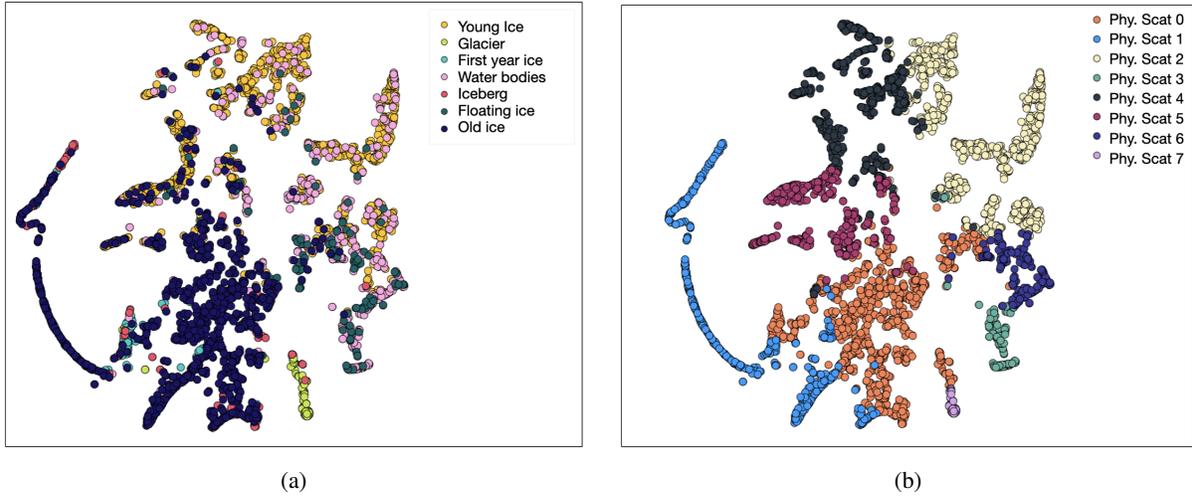


Fig. 11: Visualization of physics guided signals on test data by t-sne. (a) Different colors represent semantic labels of ice. (b) The physics guided signals are grouped into eight clusters, where each color indicates samples with similar physical scattering properties.

known targets. For example, a common approach for calibration/validation of SAR data is to employ an electronic target (transponder) that receives a signal, applies a controllable time delay, and transmits the delayed signal towards the receiver of the bistatic/monostatic system. Such a target can be used to validate results related to deformation measurements (e.g., atmospheric corrections) or polarimetric analysis.

Some real world applications of SAR requires the measurement of reliability and uncertainty. One example is the sea-ice classification in the untraversed polar regions where the ice is always promptly changeable, that would result in the difficulty for annotation and the lack of reference data. In this case, the predictions in unknown polar areas obtained by AI model need to be trusted by humans. Strong robustness and plausible degree of confidence of ML system prediction are equally as important as its accuracy.

Fig. 12 **a** indicates building orientations have a great impact on polarization orientation angles [58] and scattering mechanisms [38]. The zoomed-in region mainly contains ortho buildings where ϕ_1 is close to 0° and orientated buildings with a larger ϕ_2 . The polarization orientation angles of ortho buildings are obviously smaller than those of oriented buildings. Ortho built-up areas mainly depict double scattering

(DS) and mixed scattering (MS) where the double scattering dominates. The oriented buildings are with volume scattering (VS). Fig. 12 **b** shows limited robustness of recognition performance as the angle of test data varies when training with a small range of angles. A trustworthy model is expected to perceive SAR scattering variations with a variety of physical parameters and be perturbation-tolerant.

B. Trustworthy Modeling with Uncertainty Quantification

The development of Bayesian deep learning [59] has caught much attention in recent years, where the posterior distribution over parameters are obtained instead of the point estimation. A crucial property of the Bayesian method is its ability to quantify uncertainty, to the benefit of constructing trustworthy model.

In the case of Fig. 12 **b**, the performance of deep neural networks drops dramatically when testing SAR targets of very different orientation angles with training data. The model is over-confident about some uncertain data that cannot be perceived by frequentist method. Bayesian deep neural network, instead, is able to calibrate the output score and measure the uncertainty of the prediction. Some recent studies applied

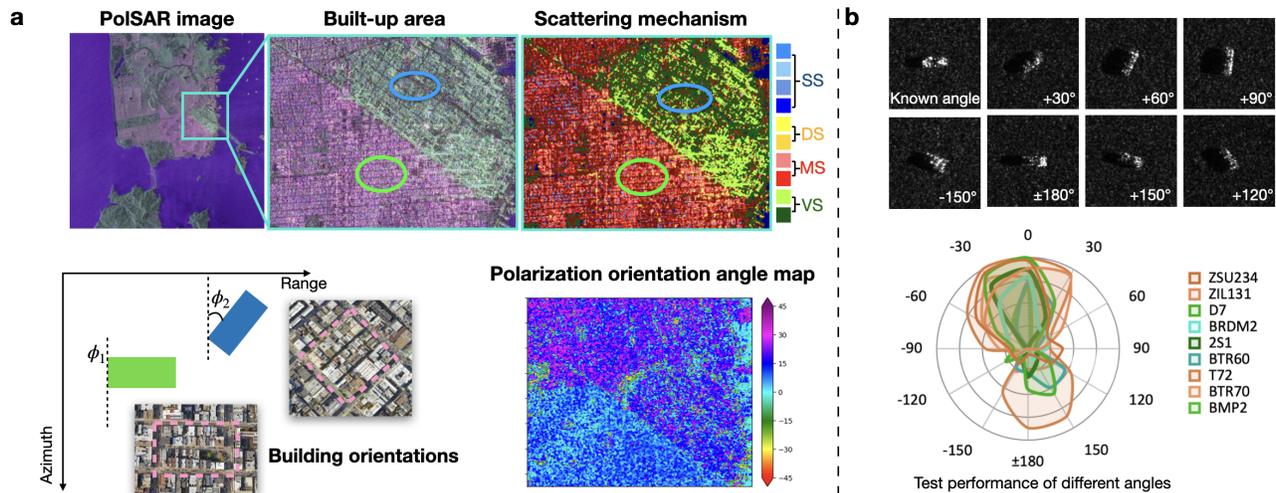


Fig. 12: A trustworthy model should perceive the SAR scattering variations with a variety of physical parameters and be perturbation-tolerant. **a** Differently oriented buildings reflect various polarization orientation angles and scattering mechanisms in a PolSAR image. **b** SAR targets vary violently with orientation angles. When training with a small range of angles, limited robustness of recognition performance is observed as the angle of test data varies.

Bayesian deep learning for SAR sea-ice segmentation [60]–[62], as well as target discrimination [63]. The generated uncertainty map can serve as a guideline for the experts in annotation and improve trust between users and the model. Some approximation strategies of Bayesian deep neural network, such as Monte Carlo Dropout [64] and Deep ensembles [65], are promising for different SAR applications.

We give an example of SAR ship detection for demonstration. The limited labeled training data, and the interference of complex scattering from target itself or the inshore background, would strongly restricted the detection performance. The ship detection result on some selected SAR images from AIR-SARShip-1.0 dataset [66], obtained by FCOS detection algorithm [67], are shown in the first row of Fig. 13. Compared with the ground truth annotation in the third row, the detection result appears many false alarms. It is crucial to estimate the model uncertainty, which is basically brought by inadequate training data, to evaluate the reliability of SAR ship detection model and provide more trustworthy predictions.

When we apply the Monte Carlo (MC) Dropout training strategy to approximate the Bayesian inference [64], it captures the uncertainty from the existing deep model for SAR ship detection. The results with high uncertainty and very low classification scores are discarded, with only the trustworthy predictions preserved. The results are shown in the second row of Fig. 13, where the false alarms are evidently reduced. In the fourth and fifth SAR image, the localization uncertainty of two large ships visualized with circles around the corner of the predicted bounding box is relatively high. Intuitively, we can infer the reason for the weak capability of the trained model in detecting such kind of targets is probably the lack of the large size ships in the training set. The feedback from the uncertainty estimation should further inspire the follow-up studies to improve the algorithm and build trustworthier models.

V. SUPPLEMENTARY EXPLANATIONS

Beyond the hybrid and trustworthy modeling, extra explanations and other interpretable models are as well required to assist with developing more transparent AI model for SAR. The explainable artificial intelligence (XAI) techniques, such as gradient based, attention based, and occlusion based explanation methods, are helpful to demonstrate the effectiveness of integrating physical layers to achieve explainability.

The transparent machine learning models, such as linear regression, decision trees, and Bayesian models, are interpretable [68]. The algorithm itself provides explanations, for example, Latent Dirichlet Allocation (LDA) builds a three-level hierarchical Bayesian model to describe the underlying relationship among document-topic-word. That is, the document can be explained with a set of topic, where each topic in turn, is represented by a distribution over words. Karmakar et al. [69] used the LDA model for SAR image data mining to generate the topic compositions and group them into semantic classes, which were fused with domain knowledge obtained by active learning from experts. The transparent model can be also integrated in a deep learning framework to approach the explainability. Huang et al. [42], [43] applied the LDA model to generate the physical attributes representation as the guided physics signals, rather than directly using the physical scattering characteristic labels to train the physics guided network. That is because the learned physics-aware features are expected to the benefit of semantic label prediction, but the semantic gap actually exists between the physical scattering characteristics and the semantic annotation. Consequently, the LDA model enables the guided signals to gain the abstract semantics and be explained with physical scattering properties.

The other purpose for approaching the explainability lies in the applications of transfer learning. The manual annotation in SAR domain is difficult and the deficiency of labeled data basically restricts the development of data-driven meth-

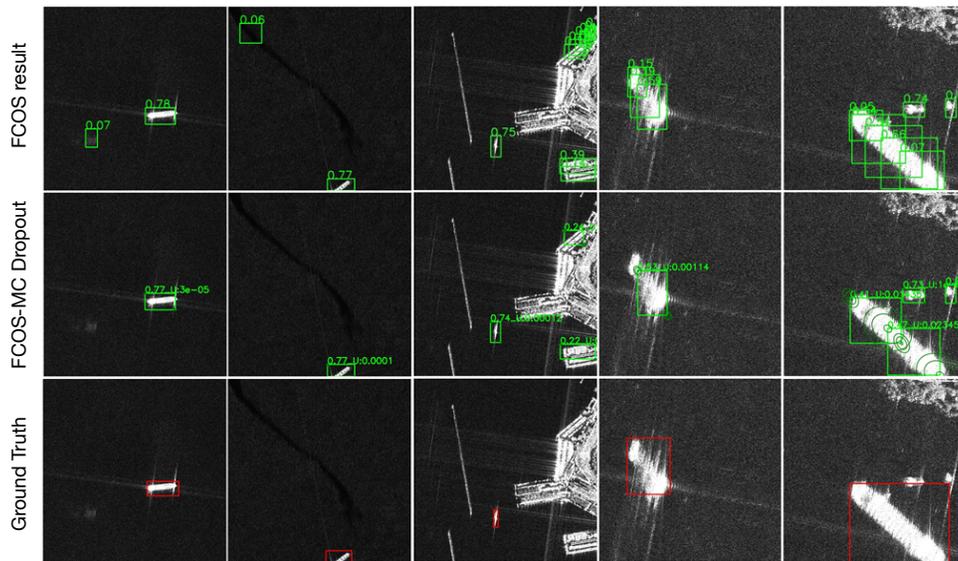


Fig. 13: The SAR ship detection result on AIR-SARShip-1.0 dataset [66], obtained by the detection deep learning algorithm FCOS [67]. Many false alarms appear in the detection result, due to the limited training data and the interference of complex scattering. The prediction uncertainty is estimated by MC-Dropout [64] and the uncertain results are discarded to achieve a better performance.

ods. Facing a wide variety of launched SAR platforms with various frequency bands and resolutions, as well as other multi-spectral, hyper-spectral, optical remote sensing sensors, it is of vital importance for elucidating the transferability of ML models among inhomogeneous data. Arrieta et al. [68] indicated the transferability is one of the goals toward reaching the explainability. Although many researches have explored different deep transfer learning methods in SAR domain [46], [70], [71], the inner transfer mechanisms of deep learning model still need explanation of insight. An insufficient understanding of the model may mislead the user toward inappropriate design of algorithm and fatal consequences, i.e. the negative transfer. Based on SAR target recognition, we proposed to analyze the transferability of features in DNN, which contributed to explaining what, where, and how to transfer more effectively for SAR images [72]. The inspiration also motivates the follow-up studies, including the SAR-specific pretrained model [73], the application in detection task [74], and the interpretability analysis of deep learning model in radar image [75].

VI. CONCLUSION AND PERSPECTIVES

In this paper, we prospect an AI paradigm shift for SAR applications that is explainable, physics aware and trustworthy. To ground this, SAR physical layers embedded with domain knowledge are introduced, which are supposed to be integrated and interacted with neural networks for hybrid modeling. Some illustrative examples are provided to demonstrate the general patterns, showing algorithmic and scientific explainability. In addition, we emphasize the importance and approaches of trustworthy modeling with Bayesian deep learning, as well as illustrating some other techniques such as interpretable machine learning method, explainable

techniques, and model transferability, that would assist with developing more transparent AI model for SAR. In fact, this field belonging to interdisciplinary research is still largely undeveloped. To our best knowledge, such approaches have not been formulated in the past years. So far, only some plain attempts have been made. Significant questions and challenges remain, e.g., the feasible representation of SAR physical layer, the optimized form of physical constraint, and hybrid modeling optimization.

Currently there are several smart sensing techniques in the SAR community that can be exploited as pre-processing steps of data fed into DNNs, e.g., multi-aperture focusing in bistatic configurations [76], monostatic/bistatic tomography, polarimetric decomposition, deformation time series. The outputs of these techniques can expose features that probably cannot be directly extracted by a DNN, especially when using a small training data set. The newly introduced AI paradigms can apply to the broad class of coherent imaging systems. A few examples can be enumerated: computer tomography, THz imaging, echographs in medicine or industrial applications, sonar or seismic observations in Earth sciences, or radio-telescope data in astrophysics.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 62101459, China Postdoctoral Science Foundation under Grant BX2021248, the Fundamental Research Funds for the Central Universities under Grant G2021KY05104, and a grant of the Romanian Ministry of Education and Research, CNCS - UEFISCDI, project number PN-III-P4-ID-PCE-2020-2120, within PNCDI III. We would like to thank the associate editor and the anonymous reviewers for their great contribution to this article.

REFERENCES

- [1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 1, pp. 6–43, 2013.
- [2] A. Meta, P. Hoogeboom, and L. P. Ligthart, "Signal processing for fmcw sar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3519–3532, 2007.
- [3] M. Amrani, A. Bey, and A. Amamra, "New sar target recognition based on yolo and very deep multi-canonical correlation analysis," *International Journal of Remote Sensing*, pp. 1–20, 2021.
- [4] M. Amrani, F. Jiang, Y. Xu, S. Liu, and S. Zhang, "Sar-oriented visual saliency model and directed acyclic graph support vector metric based target classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 10, pp. 3794–3810, 2018.
- [5] M. Amrani, K. Yang, D. Zhao, X. Fan, and F. Jiang, "An efficient feature selection for sar target classification," in *Pacific Rim Conference on Multimedia*. Springer, 2017, pp. 68–78.
- [6] M. Amrani and F. Jiang, "Deep feature extraction and combination for synthetic aperture radar target classification," *Journal of Applied Remote Sensing*, vol. 11, no. 4, p. 042616, 2017.
- [7] X. Zhu, S. Montazeri, M. Ali, Y. Hua, Y. Wang, L. Mou, Y. Shi, F. Xu, and R. Bamler, "Deep Learning Meets SAR: Concepts, Models, Pitfalls, and Perspectives," *IEEE Geoscience and Remote Sensing Magazine*, pp. 0–0, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9351574/>
- [8] F. Lattari, B. Gonzalez Leon, F. Asaro, A. Rucci, C. Prati, and M. Matteucci, "Deep learning for sar image despeckling," *Remote Sensing*, vol. 11, no. 13, 2019. [Online]. Available: <https://www.mdpi.com/2072-4292/11/13/1532>
- [9] L. H. Hughes, D. Marcos, S. Lobry, D. Tuia, and M. Schmitt, "A deep learning framework for matching of sar and optical imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 166–179, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924271620302598>
- [10] D. Xiang, Y. Xie, J. Cheng, Y. Xu, H. Zhang, and Y. Zheng, "Optical and sar image registration based on feature decoupling network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.
- [11] Y. Zhang, C. Ding, X. Qiu, and F. Li, "The Characteristics of the Multipath Scattering and the Application for Geometry Extraction in High-Resolution SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 8, pp. 4687–4699, Aug. 2015.
- [12] U. Soergel, A. Thiele, H. Gross, and U. Thoennessen, "Extraction of Bridge Features from high-resolution InSAR Data and optical Images," in *2007 Urban Remote Sensing Joint Event*. Paris, France: IEEE, Apr. 2007, pp. 1–6.
- [13] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, Jun. 2021.
- [14] A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar, "Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 10, pp. 2318–2331, Oct. 2017.
- [15] A. Karpatne, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar, "Machine Learning for the Geosciences: Challenges and Opportunities," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 8, pp. 1544–1554, Aug. 2019.
- [16] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019. [Online]. Available: <http://dx.doi.org/10.1038/s41586-019-0912-1>
- [17] M. Tria, J. Ovarlez, L. Vignaud, J. Castelli, and M. Benidir, "Discriminating real objects in radar imaging by exploiting the squared modulus of the continuous wavelet transform," *IET Radar, Sonar & Navigation*, vol. 1, no. 1, p. 27, 2007.
- [18] L. Ferro-Famil, A. Reigber, E. Pottier, and W.-M. Boerner, "Scene characterization using subaperture polarimetric sar data," vol. 41, no. 10, pp. 2264–2276. [Online]. Available: <http://ieeexplore.ieee.org/document/1237388/>
- [19] A. Anghel, G. Vasile, R. Cacoveanu, C. Ioana, and S. Ciochina, "Short-range wideband fmcw radar for millimetric displacement measurements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5633–5642, 2014.
- [20] C. Oliver, "Synthetic-aperture radar imaging," *Journal of Physics D: Applied Physics*, vol. 22, no. 7, p. 871, 1989.
- [21] M. Spigai, C. Tison, and J.-C. Souyris, "Time-frequency analysis of point target behaviour in high resolution single polarization sar images," in *7th European Conference on Synthetic Aperture Radar*. VDE, 2008, pp. 1–4.
- [22] J. Singh and M. Datcu, "Sar target analysis based on multiple-sublook decomposition: a visual exploration approach," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 2, pp. 247–251, 2011.
- [23] M. Spigai, C. Tison, and J.-c. Souyris, "Time-Frequency Analysis in High-Resolution SAR Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2699–2711, 2011.
- [24] J. Chen, X. Qiu, C. Ding, and Y. Wu, "CVCMMFF Net: Complex-Valued Convolutional and Multifeature Fusion Network for Building Semantic Segmentation of InSAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–14, 2021.
- [25] C. Yonezawa, M. Watanabe, and G. Saito, "Polarimetric Decomposition Analysis of ALOS PALSAR Observation Data before and after a Landslide Event," *Remote Sensing*, pp. 2314–2328, 2012.
- [26] A. Ferro, D. Brunner, L. Bruzzone, and G. Lemoine, "On the relationship between double bounce and the orientation of buildings in VHR SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 612–616, 2011.
- [27] L. C. Potter and R. L. Moses, "Attributed scattering centers for SAR ATR," *IEEE Transactions on Image Processing*, vol. 6, no. 1, pp. 79–91, 1997.
- [28] Z. Huang, M. Datcu, Z. Pan, and B. Lei, "Deep sar-net: Learning objects from signals," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 161, pp. 179 – 193, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924271620300162>
- [29] Z. Huang, M. Datcu, Z. Pan, X. Qiu, and B. Lei, "HDEC-TFA: An Unsupervised Learning Approach for Discovering Physical Scattering Properties of Single-Polarized SAR Image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 3054–3071, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9169671/>
- [30] J. Zhao, M. Datcu, Z. Zhang, H. Xiong, and W. Yu, "Contrastive-regulated cnn in the complex domain: A method to learn physical scattering signatures from flexible polsar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 10 116–10 135, 2019.
- [31] Z. Zhang, H. Wang, S. Member, F. Xu, and S. Member, "Complex-Valued Convolutional Neural Network and Its Application in Polarimetric SAR Image Classification," vol. 55, no. 12, pp. 7177–7188, 2017.
- [32] P. Vinayaraj, R. Sugimoto, R. Nakamura, and Y. Yamaguchi, "Transfer Learning With CNNs for Segmentation of PALSAR-2 Power Decomposition Components," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6352–6361, 2020.
- [33] S. Lei, X. Qiu, C. Ding, and S. Lei, "A Feature Enhancement Method Based on the Sub-Aperture Decomposition for Rotating Frame Ship Detection in SAR Images," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Jul. 2021, pp. 3573–3576, iSSN: 2153-7003.
- [34] J. Zhang, M. Xing, and Y. Xie, "FEC: A Feature Fusion Framework for SAR Target Recognition Based on Electromagnetic Scattering Features and Deep CNN Features," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–14, 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9127111/>
- [35] S. Feng, K. Ji, L. Zhang, X. Ma, and G. Kuang, "SAR Target Classification Based on Integration of ASC Parts Model and Deep Learning Algorithm," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–1, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9556630/>
- [36] J.-P. Ovarlez, L. Vignaud, J.-C. Castelli, M. Tria, and M. Benidir, "Analysis of SAR images by multidimensional wavelet transform," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 150, no. 4, p. 234, 2003.
- [37] Z. Huang, M. Datcu, Z. Pan, and B. Lei, "A Hybrid and Explainable Deep Learning Framework for SAR Images," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, sep 2020, pp. 1727–1730. [Online]. Available: <https://ieeexplore.ieee.org/document/9323845/>
- [38] D. Ratha, A. Bhattacharya, and A. C. Frery, "Unsupervised classification of PolSAR data using a scattering similarity measure derived from a geodesic distance," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 1, pp. 151–155, 2018.
- [39] Z. Huang and M. Datcu, "Urban Area Analysis in Single-polarized SAR Images Based on Unsupervised Deep Learning," *Proceedings of the*

- European Conference on Synthetic Aperture Radar, EUSAR*, vol. 2021-March, pp. 469–473, 2021.
- [40] Q. Song, F. Xu, and Y.-Q. Jin, “Radar Image Colorization: Converting Single-Polarization to Fully Polarimetric Using Deep Neural Networks,” *IEEE Access*, vol. 6, pp. 1647–1661, 2018. [Online]. Available: <http://ieeexplore.ieee.org/document/8141881/>
- [41] J. Qu, X. Qiu, and C. Ding, “A Study of Recovering PolSAR Information from Single-Polarized Data Using DNN,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Jul. 2021, pp. 812–815, iSSN: 2153-7003.
- [42] Z. Huang, C. O. Dumitru, and J. Ren, “Physics-aware feature learning of sar images with deep neural networks : A case study,” *IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1–4, 2021.
- [43] Z. Huang, X. Yao, Y. Liu, C. O. Dumitru, M. Datcu, and J. Han, “Physically explainable cnn for sar image classification,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 190, pp. 25–37, 2022.
- [44] X. Hu, W. Feng, Y. Guo, and Q. Wang, “Feature Learning for SAR Target Recognition with Unknown Classes by Using CVAE-GAN,” *Remote Sensing*, vol. 13, no. 18, p. 3554, Jan. 2021. [Online]. Available: <https://www.mdpi.com/2072-4292/13/18/3554>
- [45] J. Guo, B. Lei, C. Ding, and Y. Zhang, “Synthetic aperture radar image synthesis by using generative adversarial nets,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1111–1115, 2017.
- [46] D. Malmgren-Hansen, A. Kusk, J. Dall, A. A. Nielsen, R. Engholm, and H. Skriver, “Improving sar automatic target recognition models with transfer learning from simulated data,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1484–1488, 2017.
- [47] Q. Song, F. Xu, and Y.-Q. Jin, “SAR Image Representation Learning With Adversarial Autoencoder Networks,” in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2019, pp. 9498–9501, iSSN: 2153-7003.
- [48] J. Oh and M. Kim, “PeaceGAN: A GAN-Based Multi-Task Learning Method for SAR Target Image Generation with a Pose Estimator and an Auxiliary Classifier,” *Remote Sensing*, vol. 13, no. 19, p. 3939, Jan. 2021. [Online]. Available: <https://www.mdpi.com/2072-4292/13/19/3939>
- [49] M. Chu, N. Thuerey, H.-P. Seidel, C. Theobalt, and R. Zayer, “Learning meaningful controls for fluids,” *ACM Trans. Graph.*, vol. 40, no. 4, Jul. 2021. [Online]. Available: <https://doi.org/10.1145/3450626.3459845>
- [50] Y. O. U. Xie, E. Franz, M. Chu, and N. Thuerey, “tempoGAN : A Temporally Coherent , Volumetric GAN for Super-resolution Fluid Flow,” vol. 37, no. 4, 2018.
- [51] Y. Zhou, K. Yan, and X. Li, “Underwater Image Enhancement via Physical-Feedback Adversarial Transfer Learning,” *IEEE Journal of Oceanic Engineering*, vol. 47, no. 1, pp. 76–87, Jan. 2022.
- [52] Y. Wang, Q. Wang, W. Lu, and H. Li, “Physics-constrained seismic impedance inversion based on deep learning,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [53] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A simple framework for contrastive learning of visual representations,” in *Proceedings of the 37th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, H. D. III and A. Singh, Eds., vol. 119. PMLR, 13–18 Jul 2020, pp. 1597–1607. [Online]. Available: <https://proceedings.mlr.press/v119/chen20j.html>
- [54] Z. Wu, Y. Xiong, S. X. Yu, and D. Lin, “Unsupervised feature learning via non-parametric instance discrimination,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [55] B. Ren, Y. Zhao, B. Hou, J. Chanussot, and L. Jiao, “A mutual information-based self-supervised learning model for polsar land cover classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9224–9237, 2021.
- [56] R. Tănase, R. Bahmanyar, G. Schwarz, and M. Datcu, “Discovery of semantic relationships in polsar images using latent dirichlet allocation,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 2, pp. 237–241, 2017.
- [57] S. Feng, K. Ji, F. Wang, L. Zhang, X. Ma, and G. Kuang, “Electromagnetic scattering feature (esf) module embedded network based on asc model for robust and interpretable sar atr,” *IEEE Transactions on Geoscience and Remote Sensing*, 2022.
- [58] D. Xiang, T. Tang, Y. Ban, Y. Su, and G. Kuang, “Unsupervised polarimetric sar urban area classification based on model-based decomposition with cross scattering,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 86–100, 2016.
- [59] A. G. Wilson and P. Izmailov, “Bayesian deep learning and a probabilistic perspective of generalization,” *Advances in neural information processing systems*, vol. 33, pp. 4697–4708, 2020.
- [60] A. Hartmann, A. Davari, T. Seehaus, M. Braun, A. Maier, and V. Christlein, “Bayesian u-net for segmenting glaciers in SAR imagery,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 3479–3482.
- [61] N. Saberi, K. A. Scott, and C. Duguay, “Incorporating aleatoric uncertainties in lake ice mapping using radarsat-2 sar images and cnns,” *Remote Sensing*, vol. 14, no. 3, 2022. [Online]. Available: <https://www.mdpi.com/2072-4292/14/3/644>
- [62] N. Asadi, K. A. Scott, A. S. Komarov, M. Buehner, and D. A. Clausi, “Evaluation of a neural network with uncertainty for detection of ice and water in sar imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 247–259, 2021.
- [63] N. Blomerus, P. de Villiers, and W. Nel, “Improved explainability through uncertainty estimation in automatic target recognition of sar images,” in *2021 IEEE 24th International Conference on Information Fusion (FUSION)*, 2021, pp. 1–6.
- [64] Y. Gal and Z. Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning,” in *international conference on machine learning*. PMLR, 2016, pp. 1050–1059.
- [65] B. Lakshminarayanan, A. Pritzel, and C. Blundell, “Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles,” *arXiv:1612.01474 [cs, stat]*, Nov. 2017.
- [66] X. Sun, Z. Wang, Y. Sun, W. Diao, Y. Zhang, and K. Fu, “AIR-SARShip-1.0: High-resolution SAR Ship Detection Dataset (in English),” p. 852, 2019. [Online]. Available: <http://radars.ie.ac.cn/en/article/doi/10.12000/JR19097>
- [67] Z. Tian, C. Shen, H. Chen, and T. He, “Fcos: Fully convolutional one-stage object detection,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 9627–9636.
- [68] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bénéttot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, “Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai,” *Information Fusion*, vol. 58, pp. 82–115, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1566253519308103>
- [69] C. Karmakar, C. O. Dumitru, G. Schwarz, and M. Datcu, “Feature-free explainable data mining in sar images using latent dirichlet allocation,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020.
- [70] Z. Huang, Z. Pan, and B. Lei, “Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data,” *Remote Sensing*, vol. 9, no. 9, pp. 1–21, 2017.
- [71] M. Rostami, S. Kolouri, E. Eaton, and K. Kim, “Deep transfer learning for few-shot sar image classification,” *Remote Sensing*, vol. 11, no. 11, p. 1374, 2019.
- [72] Z. Huang, Z. Pan, and B. Lei, “What, Where, and How to Transfer in SAR Target Recognition Based on Deep CNNs,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 4, pp. 2324–2336, 2020.
- [73] Z. Huang, C. O. Dumitru, Z. Pan, B. Lei, and M. Datcu, “Classification of Large-Scale High-Resolution SAR Images With Deep Transfer Learning,” *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 1, pp. 107–111, Jan. 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/8966281/>
- [74] Q. An, Z. Pan, H. You, and Y. Hu, “Transitive transfer learning-based anchor free rotatable detector for sar target detection with few samples,” *IEEE Access*, vol. 9, pp. 24 011–24 025, 2021.
- [75] Y. L. X. L. C. Weijie LI, Wei YANG, “Research and exploration on interpretability of deep learning model in radar image,” *SCIENTIA SINICA Informationis*, pp. 1–21, 2021.
- [76] F. Rosu, A. Anghel, R. Căcoveanu, B. Rommen, and M. Datcu, “Multiaperture focusing for spaceborne transmitter/ground-based receiver bistatic sar,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5823–5832, 2020.

PLACE
PHOTO
HERE

Mihai Datcu (DLR - German Aerospace Center (DLR), Oberpfaffenhofen, Wesling 82234, Germany, and University POLITEHNICA of Bucharest (UPB), mihai.datcu@dlr.de). His research interests include explainable and physics aware AI, smart radar sensors design, and quantum machine learning with applications in Earth Observation (EO). He holds a Professor position in AI and information theory with UPB, and he is Senior Scientist with DLR. He elaborated AI4EO programs and systems for CNES, DLR, EC, ESA, NASA, ROSA. He was the recipient

of the Chaire d'excellence internationale Blaise Pascal 2017 for international recognition in the field of Data Science in Earth Observation. He is IEEE Fellow.

PLACE
PHOTO
HERE

Remus Cacoveanu (University POLIEHNICA of Bucharest (UPB), 313 Splaiul Independentei, Bucharest 060042, Romania, rcacoveanu at yahoo.com). His main field of expertise is in the wireless communication systems, antennas, radar sensors, propagation, and microwave circuits. He holds an Associate Professor position in telecommunications with UPB. For more than 10 years he was the technical lead of the Redline Communications' Romanian branch and between 2011-2015 he was technical consultant for Blinq Networks Canada.

One of the designed products received the "Best of WiMAX World EMEA 2008 Industry Choice Award", Munich Germany May 21st 2008.

PLACE
PHOTO
HERE

Zhongling Huang (Northwestern Polytechnical University, Xi'an, China. huangzhongling@nwpu.edu.cn). She received the B.Sc. degree in electronic information science and technology from Beijing Normal University, Beijing, China, in 2015, and the Ph.D. degree from the University of Chinese Academy of Sciences (UCAS) and the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2020. She served as a visiting scholar in German Aerospace Center (DLR) during

2018-2019, funded by UCAS. She is currently working in the BRAIN and Artificial Intelligence Lab (BRAIN LAB), School of Automation, Northwestern Polytechnical University, Xi'an, China. Her research interests include explainable deep learning for synthetic aperture radar (XAI4SAR), SAR image interpretation, deep learning, and remote sensing data mining.

PLACE
PHOTO
HERE

Andrei Anghel (University POLITEHNICA of Bucharest (UPB), 313 Splaiul Independentei, Bucharest 060042, Romania, andrei.anghelatmundepub.ro). His current research interests include remote sensing, smart radar, microwaves and signal processing. Between 2012 and 2015, he worked as a PhD researcher with Grenoble Image Speech Signal Automatics Laboratory (GIPSA-lab), Grenoble, France. Presently he is Associate Professor in telecommunications at UPB designing bistatic SAR

systems for ESA. He is IEEE Senior Member.

PLACE
PHOTO
HERE

Juanping Zhao (School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China, juanpingzhaoat-sjtu.edu.cn). Her research interests include AI for SAR and PolSAR image interpretation, pattern recognition, and machine learning. From 2018 to 2019, she was a visiting Ph.D. student with German Aerospace Center (DLR), Oberpfaffenhofen, Germany.