remote sensing

MDPI

Review

Physics-Aware Deep Learning for SAR and InSAR Remote Sensing: Models, Methods, and Open Challenges

Giorgio Taricco

Department of Electronics and Telecommunications, Politecnico di Torino (DET), Corso Duca degli Abruzzi 24, 10129 Turin, Italy; giorgio.taricco@polito.it

Highlights

What are the main findings?

- This review provides a systematic integration and extension of existing physical constraint dimensions for Synthetic Aperture Radar (SAR) and Interferometric Synthetic Aperture Radar (InSAR) learning, organized according to radar-specific constraints and algorithmic integration modes.
- The review shows that embedding radar physics improves interpretability, robustness, and consistency with sensing geometry, and two illustrative case studies—one for sparse SAR imaging and one for InSAR phase estimation—show how physics-aware designs compare with both black-box learning and classical structure-preserving baselines.

What are the implications of the main findings?

- Future learning methods for SAR and InSAR remote sensing should incorporate sensing physics, coherent scattering structure, and acquisition geometry rather than treating radar data as generic images.
- The paper provides a structured remote sensing perspective for developing more reliable, physically interpretable, and data-efficient learning methods for radar Earth observation.

Abstract

SAR and InSAR are fundamental sensing modalities for all-weather, day-and-night Earth observation because they operate independently of solar illumination and retain sensitivity to scene structure under conditions that often limit optical imaging. Recent Deep Learning (DL) methods have improved SAR image interpretation, inverse imaging, target recognition, and InSAR-based deformation analysis, but many purely data-driven pipelines still neglect the forward sensing model, coherent scattering physics, speckle statistics, phase structure, and acquisition geometry that shape radar observations. This paper develops a physics-aware perspective on learning for SAR and InSAR. The literature is organized along two complementary axes: the physical constraint dimensions that govern radar measurements, namely polarization, scattering, signal-domain structure, resolution, and interferometric phase/coherence, and the integration modes through which such structure enters modern learning systems, including representation design, model-guided architectures, and learning-assisted inverse problems. We further discuss representative applications, dataset and evaluation issues, domain-shift challenges, and trustworthy deployment. Finally, two constructive illustrative case studies are included: one on sparse SAR imaging with an explicit SAR sensing matrix, and one on InSAR phase-domain estimation under wrapped phase, coherence variation, and physics-aware regularization. Together, they illustrate how physics-aware learning compares with classical structure-preserving

Check for updates

Academic Editor: Georgia Koukiou

Received: 18 June 2026
Revised: 15 July 2026
Accepted: 22 July 2026
Published: 4 August 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

baselines and unconstrained black-box learning in representative amplitude-driven and phase-driven inverse settings.

Keywords: synthetic aperture radar; interferometric synthetic aperture radar; physics-aware learning; model-based deep learning; inverse problems; speckle; phase unwrapping; uncertainty quantification; multimodal fusion; Earth observation

1. Introduction

Synthetic Aperture Radar (SAR) has become a foundational remote sensing technology because it operates independently of solar illumination and is far less sensitive to cloud cover than optical sensors, while retaining sensitivity to scene structure, roughness, and dielectric properties through coherent electromagnetic backscattering [1–3]. Interferometric Synthetic Aperture Radar (InSAR) extends this capability by exploiting phase differences across repeated acquisitions to retrieve topography and surface deformation [2–5]. This has enabled a broad class of geophysical applications, including volcano and earthquake deformation monitoring [6–9]. It is also relevant to environmental and infrastructure-oriented mapping and monitoring problems, including surface-displacement analysis and building-scale monitoring [10–13]. Related developments in differential SAR tomography further extend this framework to more complex motion-retrieval settings [14].

At the same time, SAR and InSAR data are fundamentally different from ordinary optical images. They are produced by a coherent sensing process shaped by waveform design, acquisition geometry, polarization, phase history, resolution limits, and scattering physics. As a result, downstream inference is influenced not only by semantic scene content, but also by multiplicative speckle, geometric distortions, phase wrapping, decorrelation, and system-specific forward models [1,3,15,16]. These characteristics make radar remote sensing a particularly important testbed for learning methods that explicitly account for the sensing process rather than treating the data as generic images.

Recent years have witnessed a rapid expansion of DL across SAR and InSAR applications. Broad surveys now cover SAR target recognition, scene understanding, despeckling, super-resolution, sparse imaging, multimodal fusion, and radar-specific representation learning [15,17–19]. Related reviews focus more specifically on infrastructure monitoring with InSAR, automatic deformation detection, and physics-driven SAR target detection, further indicating that the literature is now large enough to warrant a more structured synthesis [16,20–22]. Recent InSAR-specific learning work has expanded rapidly across phase filtering and coherence estimation [23–31], phase unwrapping [32–42], and large-scale deformation or damage monitoring [43–45].

A first major line of work concerns SAR target recognition and classification. Early deep learning approaches adapted optical-style convolutional models to radar imagery and demonstrated competitive performance on benchmark datasets [46–49]. Related benchmark-oriented studies further explored target-recognition architectures, feature learning, and training protocols under controlled SAR recognition settings [50–53]. Subsequent work increasingly emphasized radar-specific information, including complex-valued processing, attributed scattering centers, electromagnetic feature transfer, manifold-aware representations, and physics-guided architectures [54–57]. Related developments also include physics-inspired attention and scattering-aware recognition strategies [58–60]. Additional recognition studies further examined architectural refinements and feature-learning variants in benchmark SAR target classification settings [61–63]. In the polarimetric SAR setting, physically informed and complex-valued networks have likewise been developed

for classification and segmentation [64–67]. Related Polarimetric Synthetic Aperture Radar (PolSAR)-oriented studies further explored deep feature learning, autoencoder-style representations, and segmentation/classification strategies tailored to polarimetric data [68–71]. Additional developments addressed more recent polarimetric classification and segmentation architectures [72–74].

A second major line of work concerns detection, especially ship and object detection in complex clutter. Here the literature likewise evolved from early convolutional ship- and object-detection pipelines toward more radar-specific detector designs [75–78]. Subsequent studies extended this line to sea–land segmentation, improved region-based detectors, and multiscale or multiscene detection settings [79–82]. Related detection work also explored additional detector variants and SAR-adapted localization strategies under complex clutter conditions [83,84]. More recent work also addresses cross-resolution adaptation and physically structured detection models [16,85–87]. One important lesson from this literature is that detection performance depends strongly on the interaction between target morphology and sensing conditions, so that physically meaningful constraints related to polarization, scattering structure, signal-domain behavior, and system resolution often matter as much as architectural scale alone [16]. Dataset construction has therefore become part of the methodological discussion as well. Widely used ship-detection resources include OpenSARShip and related Sentinel-1 or Gaofen-3 datasets, together with more complex-background benchmarks [88–90]. Additional datasets such as Air-SARShip further broaden evaluation conditions [91].

A third active area is image restoration and inverse imaging. Because classical SAR processing is already deeply rooted in forward modeling and regularized inversion, this area provides a particularly natural bridge between radar signal processing and modern hybrid learning. Classical despeckling work spans local statistics filters, wavelet-domain approaches, nonlocal methods, and patch-based probabilistic formulations [92–95]. Later developments extended this foundation through additional wavelet, nonlocal, and resolution-preserving formulations [96–99]. Closely related work also examined further nonlocal and restoration-oriented despeckling formulations [100–103]. Learning-based restoration then introduced deep denoisers, dilated residual networks, logarithmic-product models, multi-objective despeckling, and self-supervised or noisy reference formulations [104–107]. Related work further explored radar-specific neural despeckling and self-supervised restoration strategies [108–112]. The central issue nevertheless remains the same: how to suppress speckle without destroying scattering structure or phase-relevant information. For super-resolution and sparse imaging, recent work combines classical inverse formulations with trainable components, including complex-valued and unfolded architectures, thereby making the forward operator part of the learned reconstruction pipeline rather than merely a preprocessing step [113–117]. More broadly, plug-and-play priors, learned denoisers, and equilibrium-style implicit models illustrate how the interface between optimization and learning can be exploited in radar inverse problems while still preserving a transparent link to data consistency and prior structure [118–120].

A fourth important theme is that many useful radar learning problems are not purely single-modality problems. A substantial literature now addresses SAR–optical registration and matching, including geolocalization, feature descriptors, mutual-information alignment, and learned cross-modal descriptors [121–124]. Related studies also investigated additional matching and registration strategies for heterogeneous optical–SAR data [125,126]. Related work also explores optical–SAR matching, conditional adversarial models, and semi-supervised cross-modal alignment [127–134]. Complementary studies address semantic fusion and multimodal deep networks for paired radar-optical data [135–137]. Benchmark and dataset design has played an equally important role in this

area, with representative resources including Open-SARUrban, So2Sat LCZ42, SEN12MS-style data products, and SpaceNet 6 [138–142]. What makes this literature especially relevant here is that it exposes representation mismatch directly: a model must deal not only with semantic correspondence, but also with fundamentally different sensing geometries and image-formation mechanisms. This reinforces the broader argument of the present paper that learning performance depends strongly on whether the model respects radar-specific geometry and sensing structure rather than assuming interchangeable image modalities.

The same trend toward more radar-aware learning is now evident in InSAR. Deep models have been explored for phase restoration and singularity recovery in interferometric settings [143–145]. Related work also addresses volcano monitoring, synthetic deformation detection, deformation time-series interpretation, and neural treatment of seismic or geophysical interferometric observables [146–149]. Further studies consider synthetic-dataset-based volcano deformation detection, sustained deformation in InSAR time series, and neural processing for seismic source analysis [150–154]. More recent work has moved from generic restoration toward explicitly InSAR-oriented learning objectives, including joint phase/coherence estimation, nonlocal or sparse model-driven phase filtering, coherence-guided unwrapping, and large-scale phase-unwrapping frameworks [23,26,27,30,32,39,42]. Generative modeling of interferograms and recent review articles further highlight both the promise and the limitations of data-driven methods in phase-domain applications [20–22,24,155]. In InSAR, the most relevant information is often carried not by image texture but by wrapped phase, coherence, temporal geometry, and physically constrained deformation signatures. This makes naive transfer of conventional image architectures particularly problematic and motivates learning strategies that explicitly preserve interferometric structure.

These developments point to a broader methodological shift. The central issue is no longer simply whether DL can be applied to SAR or InSAR, but how physical structure should enter the learning system. A useful organizing principle is therefore to distinguish between the *physical constraint dimensions* that govern radar observations and the *integration modes* through which those constraints are embedded into modern learning pipelines. The physical dimensions include polarization, scattering, signal-domain structure, resolution, and interferometric phase/coherence, while the main integration modes include representation design, model-guided architectures, and learning-assisted inverse problems [16,57,58,115].

From this perspective, we adopt a focused definition of Physics-Aware Learning (PAL). In this paper, PAL refers to learning strategies that incorporate radar-specific sensing structure in a substantive way, such as coherent scattering behavior, electromagnetic priors, signal formation constraints, explicit forward models, interferometric geometry, phase topology, or coherence-aware objectives. Under this definition, generic transfer learning, data augmentation, or feature adaptation is not automatically physics-aware unless the adaptation itself preserves or enforces radar-specific physical structure. This distinction is important because a substantial part of the recent literature lies precisely at the boundary between convenient empirical adaptation and genuinely physics-informed modeling [15–17,19].

For operational clarity, we distinguish three levels of PAL. *Weak* PAL uses radar-specific representations or preprocessing, such as complex-valued channels, polarimetric features, coherence maps, or phase-derived inputs, but the learning rule itself remains largely generic. *Moderate* PAL injects physical structure into the model or objective, for example through scattering-aware branches, complex domain layers, coherence-weighted losses, equivariant or anisotropic operators, or phase-consistent penalties. *Strong* PAL keeps an explicit sensing, propagation, or interferometric forward model inside the inference loop, as in learned unrolling, differentiable simulators, plug-and-play inverse solvers, or uncertainty-aware formulations tied to the measurement likelihood. This weak/moderate/strong distinction is used only as an operational guide: a method is considered more strongly physics-aware when its predictions are constrained by radar measurement physics rather than merely trained on radar data.

Another recurring challenge is generalization under distribution shift. Differences in carrier frequency, incidence angle, polarization mode, bandwidth, resolution, orbit geometry, land cover composition, preprocessing chain, and temporal decorrelation can all alter the mapping from measurements to labels or reconstructions. As a result, strong performance on narrow benchmarks does not necessarily imply operational robustness. This is one reason why physically meaningful inductive bias is attractive not only for interpretability, but also for sample efficiency, regularization, and cross-sensor transfer [15,16,20].

This paper is therefore conceived as a review and perspective article on PAL for SAR and InSAR remote sensing. Its main contribution is not a single new architecture, but a systematic integration and extension of existing physical constraint perspectives that connects two viewpoints often treated separately: the physical constraint dimensions that define what information matters in radar observations, and the integration modes that determine how that structure enters modern learning systems. Within this framework, we review representative developments across target recognition, detection, restoration, sparse imaging, multimodal fusion, and interferometric analysis, and we also include two constructive illustrative case studies: one on sparse SAR imaging with an explicit sensing matrix and one on InSAR phase-domain estimation under wrapped phase, coherence variation, and physics-aware regularization. The aim is to provide a more structured and practically useful account of a literature that is currently broad, rapidly growing, and still methodologically fragmented.

Figure 1 summarizes the organization of the paper and the two complementary viewpoints used throughout: the physical dimensions of radar structure and the main integration modes through which that structure enters modern learning systems.

Figure 2 provides a schematic view of the sensing geometry that underlies many of the issues discussed in this paper. The two platform positions, denoted SAR_1 and SAR_2, represent repeated side-looking radar acquisitions separated by a baseline B. The line-of-sight paths emphasize that the measurements are geometry dependent, while the incidence angles θ_1 and θ_2 indicate the viewing geometry associated with the two acquisitions. In an interferometric setting, the quantity $\Delta\phi$ denotes the phase difference between the complex returns acquired at the two positions, which is the key observable for height and line-of-sight deformation retrieval. The scene labels further highlight that radar observations combine contributions from man-made targets, distributed clutter, vegetation, and deforming surfaces, each of which may require different forms of physics-aware modeling.

For engineering context, Table 1 reports a representative sensing configuration associated with the schematic geometry of Figure 2.

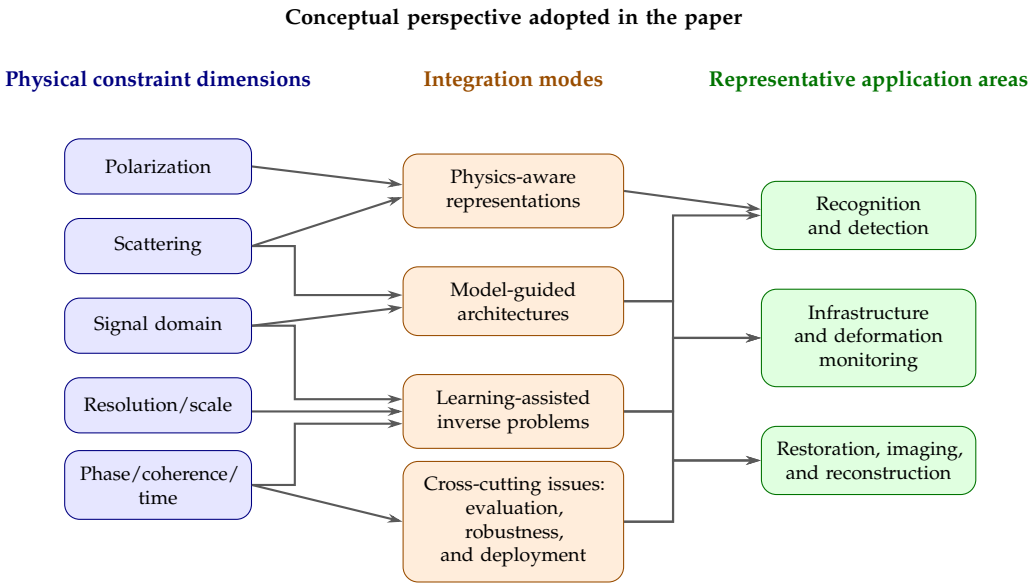


Figure 1. Conceptual framework used to organize the discussion in this paper. The figure highlights the main physical constraint dimensions emphasized in the review, the principal integration modes through which physical structure enters learning systems, and the representative application areas used to illustrate these ideas. It is intended as a guiding perspective rather than as an exhaustive taxonomy of all SAR/InSAR learning problems.

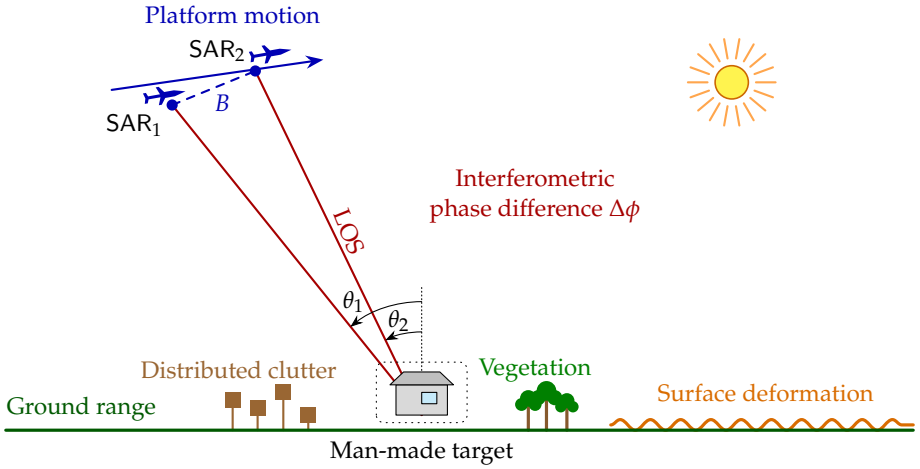


Figure 2. Illustrative SAR/InSAR sensing scenario. Two radar acquisitions with baseline B observe a scene containing a man-made target (stylized house in a dotted box), distributed clutter, vegetation, and a deforming surface patch. The geometry highlights the line of sight, the incidence angles θ_1 and θ_2, and the interferometric phase difference $\Delta\phi$. The presence of sunlight recalls that, unlike optical sensing, radar acquisition is independent of solar illumination.

Table 1. Representative SAR/InSAR sensing configuration associated with the scenario of Figure 2.

Parameter	Value
Carrier frequency	9.65 GHz (X-band)
Wavelength	3.11 cm
Signal bandwidth	300 MHz
Waveform modulation	Linear FM chirp
Pulse duration	20 µs
Pulse repetition frequency	2 kHz
Transmit polarization	HH
Receive polarization	HH

Table 1. *Cont.*

Parameter	Value
Platform altitude	6 km
Platform velocity	120 m/s
Slant range to scene center	10 km
Nominal incidence angle from vertical	53.1°
Interferometric baseline	20 m
Antenna length (azimuth)	1.5 m
Antenna width (range)	0.3 m
Antenna gain	30 dBi
Nominal range resolution	0.5 m
Nominal azimuth resolution	0.5 m

2. Physical Constraint Dimensions in SAR and InSAR Learning

A useful expansion of the physics-aware viewpoint is to identify the main types of physical structure that radar learning systems must preserve or exploit. Recent work on physics-driven SAR target detection organizes the problem around four physical constraint families: polarization characteristics, scattering mechanisms, signal-domain properties, and system resolution [16]. That taxonomy is helpful beyond detection because it highlights the fact that different radar tasks are limited by different types of information loss. Classification and detection may depend strongly on polarimetric and scattering descriptors, inverse problems often depend on signal-domain and resolution structure, and InSAR adds a further layer involving wrapped phase, coherence, and temporal geometry.

2.1. Polarization and Scattering Structure

Polarization is one of the most direct ways in which radar sensing differs from ordinary image acquisition. Unlike a single-channel intensity representation, polarimetric data carry information about the vector interaction between the incident electromagnetic wave and the target. As emphasized in recent physics-driven detection reviews, the complete scattering matrix and the derived coherency matrix provide physically interpretable access to surface, double-bounce, and volume scattering mechanisms, thereby improving target–background separability in heterogeneous scenes [16]. This is especially relevant when the scene contains a mixture of man-made structures and distributed clutter, since stable double-bounce or specular responses can be difficult to capture with intensity-only features.

A related but distinct notion is *scattering structure*. Here the emphasis is not on polarization state alone, but on the geometric and electromagnetic organization of the dominant scatterers. Attribute scattering center models, multiscale scattering decompositions, and electromagnetic feature maps translate raw radar returns into descriptors that are much closer to the physical composition of the target [59,60,156,157]. The value of these representations is twofold: first, they provide interpretable priors that can guide feature extraction, and second, they offer a way to build robustness against clutter and viewpoint variability. This is one reason why recent detection and recognition studies increasingly combine scattering-center information with deep learned features rather than relying on pixel amplitudes alone.

For InSAR, there is no direct analogue of full polarimetric target decomposition in many common workflows, yet the same logic applies to phase-derived structure. Coherence, phase gradients, local unwrap topology, and multi-temporal consistency all function as physically meaningful descriptors. In this framework, InSAR coherence and phase topology can be regarded as the interferometric counterpart of scattering-aware representation design.

2.2. Signal-Domain and Time–Frequency Structure

A second constraint family concerns *signal-domain structure*. Radar images are obtained only after a sequence of signal-processing operations, and some of the most informative cues are more naturally expressed before or alongside the final formed image. Recent surveys of physics-driven SAR detection stress that time–frequency representations and Range–Doppler (RD)-domain features expose micro-motion, Doppler shifts, and non-stationary scattering behavior that may be hidden or aliased in the image domain [16]. This is particularly relevant for moving-target detection, maritime surveillance, and high-clutter scenes, but the idea is broader: signal-domain features can also encode target-specific backscattering patterns even for nominally static objects.

The Deep SAR-Net line of work is an important illustration of this point. Rather than relying only on detected intensity images, it augments spatial features with time–frequency descriptors extracted from complex-valued SAR data, thereby preserving signal-domain information that is lost in amplitude-only pipelines [54,158]. This family of methods suggests that learning from radar should sometimes begin from a richer signal representation, especially when the task depends on subtle phase or frequency-domain variation. The same principle also underlies many modern inverse methods, where learned components are applied not only to images but also to partially processed or operator-domain quantities.

2.3. Resolution and Scale

A third constraint family is *resolution*. In radar, changes in bandwidth, aperture, sensor mode, or platform geometry do not simply rescale an image in the ordinary vision sense; they change the effective Point Spread Function (PSF), alter the visibility of scattering centers, and modify the appearance of target structure. Recent reviews emphasize that target representation is inherently resolution-dependent, and that many failures of cross-platform generalization are better understood as failures to preserve physically meaningful scale structure across acquisition regimes [16]. When the bandwidth is reduced, nearby scatterers merge, high-frequency target geometry becomes less visible, and both recognition and detection can deteriorate sharply.

This has direct consequences for PAL. Cross-resolution domain adaptation, structure-aware feature alignment, and radar-specific super-resolution should not be regarded as generic image enhancement problems. Rather, they are attempts to preserve scattering topology under changing system resolution [85,113,114]. In a broader review, this dimension can provide a bridge between target detection, ship detection, super-resolution, and sparse imaging, all of which depend on how the measurement system controls target separability.

2.4. Interferometric Phase, Coherence, and Temporal Geometry

For InSAR, the most distinctive physical constraints arise from interferometric phase and temporal acquisition structure. The key observables are not simply image intensities but wrapped phases, coherence estimates, baseline-dependent sensitivities, and multi-temporal phase histories. These quantities are governed by cyclic geometry, atmospheric contamination, temporal decorrelation, and platform revisit characteristics. As a result, many InSAR learning problems differ fundamentally from both optical vision tasks and single-image SAR interpretation [20–22].

This suggests that a complete physics-aware review should include a fifth dimension, specific to InSAR: interferometric phase, coherence, and temporal geometry. In practical terms, this includes respecting cyclic phase structure in phase filtering and unwrapping, incorporating coherence uncertainty into downstream decision making, and preserving deformation-related temporal signatures across stacks or time series. Even when a method

is presented as image-based, it should still be assessed in terms of whether it retains the relevant interferometric structure.

2.5. Coupling Between Physical Constraint Dimensions

The physical constraint dimensions introduced above are analytically useful, but they should not be interpreted as independent axes. In practical SAR and InSAR systems, they are strongly coupled through the sensing chain, and many recent learning methods succeed precisely because they exploit these couplings rather than isolating a single form of prior information.

A first important coupling is between polarization and scattering structure. Polarimetric measurements do not merely add channels; they provide a richer view of the underlying scattering mechanisms and therefore affect how surface, double-bounce, and volume responses can be separated or emphasized. Accordingly, scattering-aware learning and polarization-aware learning are often not distinct categories in practice, but two aspects of the same physical description of the radar return [16,57,59,60].

A second coupling links signal-domain structure and system resolution. Resolution is determined by waveform bandwidth, aperture extent, and sampling geometry, so it is inseparable from the signal model that generates the data. Changes in resolution therefore do not simply alter image sharpness; they modify the visibility, separability, and stability of scattering signatures. This is why cross-resolution adaptation, radar super-resolution, and sparse imaging are best understood not as generic image enhancement tasks, but as problems in preserving physically meaningful signal structure under altered measurement conditions [85,113–115].

A third and especially important coupling arises in InSAR, where phase, coherence, and temporal geometry are inseparable. Wrapped phase is interpreted through the acquisition geometry, coherence controls the reliability of phase observations, and multi-temporal baselines determine sensitivity to deformation and decorrelation. As a result, interferometric learning methods that treat phase, coherence, or temporal structure in isolation are often incomplete. The relevant physical object is the coupled interferometric measurement process itself, not any one of these quantities alone [20–22].

These couplings also appear at the architectural level. Some recent methods fuse scattering descriptors with learned deep features, thereby combining interpretable radar structure with data-driven representation learning [58,86,159]. Other methods combine complex-domain processing with electromagnetic feature transfer in order to preserve signal-domain and scattering information more explicitly [54–56]. Taken together, these examples show that many recent architectures do not rely on a single physical prior, but instead integrate several complementary constraints within the same learning pipeline. For this reason, the framework proposed in this paper should not be read as a list of parallel boxes. Rather, the physical dimensions identify the main sources of radar structure, while their couplings explain why effective PAL methods often draw on several of them at once.

From a design perspective, this observation has a practical implication: the most useful inductive biases are often those introduced at the interfaces between dimensions. Examples include coherence-weighted phase processing, signal-domain models that preserve resolution-dependent scattering behavior, and feature hierarchies that connect polarimetric observables to scattering topology. This is one reason why physics-aware learning in radar remote sensing should be formulated as a sensing-chain problem rather than as a simple extension of image-based deep learning.

3. Foundations and Integration Modes for Physics-Aware SAR/InSAR Learning

The early wave of DL in SAR was largely inspired by successes in optical computer vision. Standard Convolutional Neural Network (CNN)s were applied to amplitude imagery, often after converting complex SAR data into real-valued image-like representations. These methods achieved strong results on controlled benchmarks, particularly for Automatic Target Recognition (ATR), ship detection, and scene classification. Representative examples include complex data and feature fusion pipelines that sought to preserve more SAR-specific information than amplitude-only baselines [54–56].

The limitations of this first wave are now well understood. Performance often drops under sensor shift, geographic shift, incidence angle variation, reduced training labels, or operationally realistic clutter [15,16]. Moreover, many pipelines inherit implicit assumptions from optical vision that do not match coherent radar formation physics. Recent surveys explicitly note that future progress requires models that better exploit the physical characteristics of SAR data rather than treating the data as ordinary grayscale imagery [16,18,19].

PAL can therefore be interpreted broadly. In this study, we use the term to denote any learning strategy that incorporates one or more of the following:

(1) a physically meaningful data representation, such as complex-valued, polarimetric, scattering-center, coherence, or phase-based inputs;
(2) architectural constraints motivated by the forward model or scattering mechanism;
(3) algorithm unrolling or learning-assisted optimization for inverse problems;
(4) loss functions or training objectives that enforce physical consistency; and
(5) uncertainty modeling tied to acquisition conditions and measurement limitations.

This taxonomy is useful because it organizes methods by *where* physical structure enters the learning pipeline rather than only by application label. It also helps relate SAR and InSAR methods that are often discussed separately even when they share the same methodological logic. A concise way to structure the study is around three main integration modes: physics-aware representation design, model-guided architecture design, and learning-assisted inverse problems. Although these categories are not mutually exclusive, they cover a large share of the current literature and provide a clear way to organize the main technical design choices.

3.1. Physics-Aware Inputs and Representations

A natural route toward PAL is to retain more information from the sensing process itself. In SAR image interpretation, the choice of representation strongly influences what the network can learn. Amplitude-only inputs discard phase and often suppress polarization-dependent or scattering-structure information. By contrast, methods built on complex-valued or hybrid signal–image representations attempt to preserve radar-specific cues [54–56].

For example, Huang et al. [54] introduced a SAR-specific framework that learns from signal-derived representations rather than only post-processed image intensity. Liu and Lang [55] proposed multi-manifold feature fusion for complex-valued SAR target recognition, emphasizing the geometric structure of complex data. Liu et al. [56] further explored complex convolution and electromagnetic feature transfer learning for SAR target recognition. Taken together, these works illustrate a broader lesson: preserving domain-specific structure at the representation stage can improve both sample efficiency and robustness.

A related direction is the use of interpretable physics-derived descriptors as side information. Examples include scattering centers, polarimetric decompositions, coherence

maps, elevation cues (height-related descriptors, such as local topography, object height, or vertical scattering structure), or simulated electromagnetic attributes. Rather than replacing learned features, these variables can guide feature extraction or attention. Recent physics-guided classification work has shown that explicit incorporation of electromagnetic characteristics can improve explainability and performance, especially when labels are limited [57–60]. For InSAR, the corresponding issue is even sharper because the quantities of interest are often encoded in wrapped phase and coherence rather than directly visible spatial textures. This suggests that representation design for interferograms, coherence maps, stacks, and time series should be treated as a central topic rather than a preprocessing detail. A practical rule is to favor representation-level PAL when the main risk is information loss at the input stage: for example, when amplitude-only images discard phase, polarimetry, coherence, or complex-valued structure that is known to be relevant to the downstream task. For InSAR, recent phase filtering, coherence estimation, and unwrapping networks make this point concrete by using wrapped phase, coherence, amplitude, phase gradients, or quality maps as explicit inputs rather than treating interferograms as ordinary real-valued images [25,33,34,36,37,40].

3.2. Model-Guided Architectures

A second family of methods embeds physical knowledge directly into the network architecture. This can range from relatively mild design choices, such as complex-valued convolutions, to stronger physics-guided modules that inject scattering priors into the feature hierarchy. The main motivation is that appropriate inductive bias can reduce the amount of data required for generalization while preserving interpretability.

Huang et al. [57] proposed a physics-guided and injected learning framework for SAR image classification. Their approach combines explainable physics models, a physics-guided network, and a physics-injected prediction stage to preserve electromagnetic consistency in the learned representation. More recently, Huang et al. [58] introduced a physics-inspired hybrid attention mechanism that uses high-level semantic priors derived from physical information to reweight learned features for target recognition. The target detection literature reports a similar shift from feature-level augmentation toward deeper architectural integration, including local electromagnetic branches, scattering-guided fusion blocks, and manifold-aware complex-valued backbones [86,87,159].

This design philosophy aligns with broader trends in scientific machine learning, where inductive bias is used to narrow the hypothesis space. For SAR and InSAR, promising architectural priors include azimuth–range anisotropy, phase topology, scattering sparsity, multi-look statistics, and multitemporal consistency. These priors are best compared in a task-dependent way rather than treating "physics-aware" learning as a single recipe. Architectural PAL is most appropriate when the physical constraint is local or compositional enough to be embedded as a reusable network module, such as a scattering-aware branch, complex-valued convolution, phase-preserving block, or coherence-conditioned attention mechanism.

3.3. Learning-Assisted Inverse Problems

Many important SAR and InSAR applications are basically inverse problems rather than standard image-recognition tasks. Image formation, despeckling, super-resolution, tomography, sparse reconstruction, phase estimation, coherence estimation, and phase unwrapping are naturally placed at the interface between forward modeling and statistical inference [1,3].

This setting is especially well suited to hybrid model–data approaches. Learned unrolling, plug-and-play priors, score-based priors, and neural surrogates can be used to

accelerate or regularize iterative solvers while preserving a clear link to the underlying physics. Recent work on deep learning-based sparse SAR imaging via unrolling is one example of this trend [115]. In InSAR, recent overviews point to DL as a means of improving phase and coherence estimation as well as phase unwrapping under strong noise and decorrelation [22].

The conceptual advantage of this family of methods is that it avoids a false choice between handcrafted inverse algorithms and end-to-end black-box predictors. Instead, one starts from the forward operator or likelihood model and then learns the missing components: a prior, a proximal step, a denoiser, or an adaptive update rule. This is one of the clearest places where the connection between classical radar signal processing and modern machine learning becomes analytically useful. Inverse problem PAL is most appropriate when a credible forward operator or likelihood model is available and the main uncertainty lies in the prior, regularizer, stopping rule, or computational budget.

For additional context, recent work has also explored complex-valued and deep-unfolded formulations for radar imaging and 3-D reconstruction [115–117]. Related work on restoration has examined nonlocal learning, multi-objective formulations, and self-supervised or noisy reference strategies for SAR despeckling [105,109–111]. Noisy-reference deep restoration methods provide a further example of how reconstruction priors can be adapted to radar data without discarding the inverse problem viewpoint [112].

The observed sensitivity to forward-model inconsistency suggests a need for more flexible integration modes that decouple the learned prior from the specific sensing geometry. Recent advances in PAL suggest that frameworks such as Plug-and-Play (PnP) priors [118,119] or Deep Equilibrium (DEQ) models [120] may offer improved robustness against platform instabilities and atmospheric phase errors by avoiding fixed-depth unrolling tied to a single nominal operator.

4. Representative Applications, Datasets, and Evaluation Protocols

Among the goals of this study is showing how PAL affects performance in representative application settings, how such performance ought to be evaluated, and which unresolved questions are most likely to shape future progress. Three application domains are especially informative in this respect: target recognition and detection, infrastructure and deformation monitoring, and image restoration and reconstruction.

To improve the structural readability of the review, Table 2 summarizes representative contributions within the proposed PAL framework by mapping them to the physical constraint emphasized, the main integration mode, the target application domain, and the dataset or experimental setting. This synthesis is intended to complement the narrative discussion and to make cross-paper comparisons more transparent; consequently, the following text emphasizes interpretation rather than repeating every entry in the table.

Table 2. Synthesis of representative literature within the proposed PAL framework.

Representative Work	Physical Constraint Emphasized	Integration Mode	Application Domain	Dataset/Setting
Huang et al. [54]	Signal-domain structure, complex data	Physics-aware representation design	SAR target recognition	complex-valued SAR target data
Liu and Lang [55]	Complex data manifold structure	Physics-aware representation design	SAR target recognition	controlled recognition benchmarks
Huang et al. [57]	Electromagnetic/ scattering priors	Model-guided architecture	SAR classification	controlled recognition benchmarks
Huang et al. [58]	Physics-inspired semantic priors	Model-guided architecture	SAR target recognition	controlled recognition benchmarks

Table 2. *Cont.*

Representative Work	Physical Constraint Emphasized	Integration Mode	Application Domain	Dataset/Setting
Ji et al. [115]	Forward model, sparse-imaging physics	Learning-assisted inverse problem	Sparse SAR reconstruction	Synthetic/controlled sparse imaging
Qin et al. [85]	Resolution and scale	Representation plus adaptation	Cross-resolution SAR analysis	Cross-resolution benchmarks
Shen et al. [113]	Polarization and resolution	Learning-assisted restoration	PolSAR super-resolution	Polarimetric SAR imagery
Fontes et al. [20]	Interferometric phase and coherence	Survey of integration strategies	Infrastructure monitoring with InSAR	Multiple InSAR studies
Anantrasirichai et al. [21]	Interferometric phase and temporal geometry	Review of task pipelines	Deformation-pattern recognition	Multiple geohazard/ InSAR studies
Sun et al. [23], Li et al. [30]	Interferometric phase and coherence	Joint phase/ coherence estimation	Phase filtering and coherence estimation	Simulated and real InSAR data
Pu et al. [26], Wang et al. [27], Zhang et al. [29]	Phase noise, nonlocal structure, sparse filtering model	Model-guided/ feature-extraction architecture	Interferometric phase filtering	Simulated and Sentinel-1 InSAR data
Sica et al. [32], Mu et al. [39]	Coherence-guided phase topology	Coherence-conditioned learning	InSAR phase unwrapping	Synthetic and real interferograms
Yang et al. [38], Zhou and Yu [41], He et al. [42]	Phase gradients, discontinuities, scale	Deep/model-based unwrapping	Complex-event phase unwrapping	LiDAR-assisted, synthetic, and large-scene settings
Wang et al. [43], Stephenson et al. [44], Soni et al. [45]	Temporal coherence and deformation history	Detection-first/ sequence learning	Time-series deformation or damage monitoring	Large-scale InSAR stacks
Li et al. [16]	Polarization, scattering, signal domain, resolution	Cross-mode survey	SAR detection and recognition	Multiple detection benchmarks

4.1. Target Recognition and Detection

SAR target recognition remains a major testbed because it exposes issues of limited labels, viewpoint sensitivity, clutter, and operational robustness. The recent literature increasingly argues that physically meaningful information—such as attributed scattering centers, electromagnetic priors, or physically grounded attention—can improve both robustness and interpretability [16,57,58]. This makes target recognition a useful case study for the broader claim that suitable inductive bias can matter as much as raw model capacity.

The physics-driven target detection literature also provides a useful catalog of how different physical constraints matter in different scenarios. Polarimetric descriptors are very valuable in maritime scenes and clutter suppression; scattering-center models help with fine-grained recognition and target topology; signal-domain methods are especially relevant to moving-target scenarios; and resolution-aware strategies matter in cross-platform and multiscale detection [16]. Even though the present study has a broader scope than detection alone, these use cases illustrate how application requirements can determine which physical attributes should be emphasized.

4.2. Infrastructure and Deformation Monitoring

InSAR-based monitoring of built infrastructure and geophysical deformation is another natural application area for DL because interpretation can be labor-intensive and error-prone. A recent review on DL for InSAR-based infrastructure monitoring highlights opportunities for automated interpretation but also emphasizes persistent difficulties due

to noise, decorrelation, and false positives [20]. Similarly, recent work reviewing DL for volcanic and earthquake deformation detection underscores the importance of architecture design and dataset construction in InSAR workflows [21]. These applications also sharpen the need for uncertainty awareness because operational decisions may carry substantial practical consequences.

4.3. Image Restoration, Sparse Imaging, and Super-Resolution

Despeckling, super-resolution, and sparse imaging are attractive use cases for PAL because their objectives can be tied directly to observation models. In these tasks, the distinction between signal processing and learning is especially thin. The most convincing contributions in this area are likely to be those that compare learned components against classical regularized baselines under controlled degradation, reduced sampling, or changed sensing conditions rather than only on a single curated benchmark. Recent work on complex-valued sparse imaging, resolution-aware adaptation, and radar-specific super-resolution reinforces this point [85,113,115].

4.4. Datasets and Evaluation Protocols

A persistent difficulty in interpreting the recent literature is that reported gains are often obtained under specific dataset partitions, acquisition conditions, and preprocessing chains, so direct cross-paper comparison of absolute performance numbers should be treated with caution. This point is stressed explicitly in recent physics-driven target detection reviews, which note that variations in subsets, train/test partitions, augmentation strategies, and computational setup can strongly affect reported performance [16]. For detection-oriented tasks, standard benchmarks include Moving and Stationary Target Acquisition and Recognition (MSTAR), SAR Ship Detection Dataset (SSDD), High-Resolution SAR Images Dataset (HRSID), and SAR Rotated Ship Detection Dataset (SRSD); for reconstruction and imaging tasks, however, benchmark culture is much less standardized, and for InSAR deformation tasks the diversity of scenes and label sources makes comparison even harder [16,20,21].

Related benchmark and dataset efforts also underline the importance of careful empirical comparison in radar learning. For urban interpretation, land cover analysis, and multimodal fusion, widely used resources include Open-SARUrban, So2Sat LCZ42, and SEN12MS-style paired datasets [138–141]. For all-weather multimodal mapping, SpaceNet 6 has become an important benchmark [142]. For target recognition, MSTAR and its associated public evaluation protocols remain historically central reference points [160]. For ship detection, several dedicated datasets are now widely used, including OpenSARShip and related high-resolution resources [88–90]. Additional air- and land cover-oriented datasets further broaden the empirical base for robustness studies [91,161].

Across these domains, evaluation remains a major weakness in the current literature. Existing benchmark practice often overstates progress because many datasets are small, geographically narrow, or captured under limited acquisition diversity. Controlled benchmark performance does not necessarily translate to operational robustness. At least five evaluation gaps deserve explicit treatment: sensor and geometry shift, geographic and temporal shift, label scarcity and label quality, uncertainty calibration, and reproducibility. The value of such a synthesis lies in comparing these concerns across recognition, reconstruction, and deformation-analysis tasks rather than discussing them in isolation.

5. Illustrative Case Studies: Physics-Aware SAR Imaging and InSAR Phase Estimation

To make the preceding discussion concrete, this section presents two illustrative case studies in controlled but interpretable settings: one on sparse SAR imaging under controlled degradation, and one on InSAR phase-domain estimation with wrapped phase, coherence variation, and physics-aware regularization. The aim is not to establish new task-specific benchmarks, but to provide two complementary examples showing how physics-aware structure affects reconstruction quality, robustness, and failure modes in representative radar inverse problems. The two examples are deliberately placed on parallel footing. The first addresses a SAR inverse imaging problem in which the sensing matrix is explicit and reconstruction is limited by subsampling, noise, and finite reconstruction depth. The second addresses an InSAR phase-domain estimation problem in which wrapped phase, coherence, and unit modulus complex structure are central. Together, they illustrate how physics-aware learning interacts with forward modeling, inverse problems, and learning-based regularization across both amplitude-driven and phase-driven radar settings.

5.1. Common Inverse Problem Perspective

Both case studies can be understood within a common inverse problem framework. Let $\boldsymbol{x} \in \mathbb{C}^N$ denote the unknown signal of interest and let $\boldsymbol{y} \in \mathbb{C}^M$ denote the observed data. A simplified observation model may be written as

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{z}, \tag{1}$$

where $\boldsymbol{\Phi}$ represents the sensing or observation operator and $\boldsymbol{z}$ models measurement noise. When the problem is ill posed because of noise, subsampling, wrapping, or other information loss, recovery requires additional prior structure. A standard formulation is

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \left\{ \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + \lambda \mathcal{R}(\boldsymbol{x}) \right\}, \tag{2}$$

where $\mathcal{R}(\boldsymbol{x})$ encodes prior information and $\lambda > 0$ controls the tradeoff between fidelity to the observations and adherence to the prior [1,3].

In sparse imaging, a common choice is $\mathcal{R}(\boldsymbol{x}) = \|\boldsymbol{x}\|_1$, which promotes reconstructions with few dominant scatterers. More generally, one may impose sparsity after a transform $\boldsymbol{\Psi}$, so that $\mathcal{R}(\boldsymbol{x}) = \|\boldsymbol{\Psi}\boldsymbol{x}\|_1$, or use spatial regularity penalties such as total variation, $\mathcal{R}(\boldsymbol{x}) = \|\nabla_s \boldsymbol{x}\|_1$, where ∇_s denotes the discrete spatial gradient of the signal field. In interferometric phase estimation, the regularizer may instead encode local phase smoothness, complex-domain consistency, coherence weighting, or explicit geometric structure of the unit modulus phasor field. In both settings, the forward model specifies the physically grounded part of the problem, while the prior represents the structure that must be supplied analytically or learned from data.

This viewpoint is useful because it separates the components of the inverse problem that are fixed by the sensing physics from those that may be adapted. In particular, the forward operator $\boldsymbol{\Phi}$ enforces consistency with the acquisition model, while the prior term expresses assumptions about scene or phase structure. Hybrid methods retain the physical model while replacing part of the inference procedure with trainable components; for example, unrolled iterative schemes preserve the structure of classical optimization algorithms but learn step sizes, thresholds, proximal maps, or blending coefficients. From this perspective, the key issue is not whether a method is model-based or learned, but which parts of the inverse problem remain explicit and which are adapted from data.

For example, a standard iterative update takes the form

$$\boldsymbol{x}^{(k+1)} = \mathcal{S}_{\tau^{(k)}}\Big(\boldsymbol{x}^{(k)} + \rho^{(k)}\boldsymbol{\Phi}^{\mathsf{H}}\Big(\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}^{(k)}\Big)\Big), \tag{3}$$

where $\rho^{(k)}$ is a step size, $\tau^{(k)}$ is a threshold parameter, $(\cdot)^{\mathsf{H}}$ denotes the Hermitian transpose, and $\mathcal{S}_\tau$ is a shrinkage or correction operator. A conventional reconstruction algorithm fixes these ingredients analytically or heuristically. A physics-aware learned variant instead keeps the data-consistency term tied to the sensing operator while allowing some update components to be learned. This structure narrows the hypothesis space relative to fully black-box predictors, preserves a direct mechanism for enforcing measurement consistency, and makes failure analysis easier because errors can be attributed separately to forward-model mismatch, prior mismatch, and limited expressive power in the learned module.

It is important to clarify the role of Section 5 within the present paper. The sparse-SAR equations above are not introduced as new inverse problem formulations, nor is the goal to review deep unfolded SAR imaging as a field in itself. Rather, this case study illustrates one specific integration mode within the broader PAL framework developed in this review: the sensing operator remains explicit, data consistency is enforced through the SAR forward model, and only selected update components are learned. In this sense, the example serves as a concrete instance of model-guided reconstruction inside a wider review that also covers recognition, detection, multimodal fusion, InSAR phase-domain learning, and evaluation issues. This broader scope distinguishes the present paper from more task-specific surveys, for example recent reviews focused primarily on physics-driven SAR detection [16], because the present paper explicitly connects representation design, model-guided architectures, inverse problems, interferometric phase-domain learning, and evaluation issues within a common physics-aware synthesis.

5.2. Illustrative SAR Case Study: Sparse Imaging with an Explicit Sensing Matrix

We first consider a controlled sparse SAR imaging experiment in which the forward model is an explicit SAR sensing matrix derived from a simplified monostatic aperture geometry. The goal is to identify a concrete operating regime in which a physics-aware learned update is useful. In particular, we focus on the finite-depth setting, where reconstruction quality must be achieved with a small number of iterations or layers.

Let $\boldsymbol{\Phi}_{\mathsf{SAR}} \in \mathbb{C}^{M\times N}$ denote the SAR sensing matrix associated with the chosen acquisition geometry, where M is the number of collected measurements and $N = N_x N_y$ is the number of pixels in the discretized scene. If the scene grid is $\{(x_n, y_n)\}_{n=1}^{N}$ and the radar platform location at measurement index m is $(x_{a,m}, y_{a,m}, z_{a,m})$, then the slant range from the platform to the nth pixel is $R_{m,n} = \sqrt{(x_{a,m}-x_n)^2 + (y_{a,m}-y_n)^2 + z_{a,m}^2}$. With carrier frequency f_c and propagation speed c, the wavenumber is $k = 4\pi f_c/c$. Using a reference range $R_{0,m}$ for phase centering, the unnormalized matrix entries are defined as

$$\widetilde{\Phi}_{\mathsf{SAR}}(m,n) = \frac{1}{\max(R_{m,n}^{\gamma}, \varepsilon)} \exp\!\left(-\frac{1}{2}\left(\frac{x_n - x_{a,m}}{\sigma_b}\right)^2 - \mathrm{j}\,k(R_{m,n} - R_{0,m})\right), \tag{4}$$

where σ_b is the beam-spread parameter, γ is the range-decay exponent, and $\varepsilon > 0$ prevents division by zero. The actual sensing matrix is obtained by row normalization (Hereafter, $\boldsymbol{X}(m,:)$ denotes the m-th row of the matrix $\boldsymbol{X}$):

$$\boldsymbol{\Phi}_{\mathsf{SAR}}(m,:) = \frac{\widetilde{\boldsymbol{\Phi}}_{\mathsf{SAR}}(m,:)}{\|\widetilde{\boldsymbol{\Phi}}_{\mathsf{SAR}}(m,:)\|_2}, \qquad m = 1,\ldots,M. \tag{5}$$

Hence, $\mathbf{\Phi}_{\mathsf{SAR}}$ is not a generic random matrix, but a structured matrix determined explicitly by the SAR geometry and propagation model. The measurements obey

$$\boldsymbol{y} = \mathbf{\Phi}_{\mathsf{SAR}}\boldsymbol{x} + \boldsymbol{z}, \tag{6}$$

where $\boldsymbol{x} \in \mathbb{C}^N$ is the scene reflectivity vector and $\boldsymbol{z}$ is circular complex Gaussian noise.

In the present experiment the scene size is 32×32, so that $N = 1024$, while the measurement dimension is $M = 819$, corresponding to a sampling ratio $M/N = 0.8$. The SAR geometry is specified by $d_x = d_y = 1$ m, carrier frequency $f_c = 10$ GHz, altitude $H = 200$ m, platform speed $v = 100$ m/s, scene center $(x_c, y_c) = (0, 1000)$ m, pulse repetition interval $\tau_{\mathsf{PRI}} = 0.02$ s, beam-width parameter $\sigma_b = 80$, and unit range exponent. The scene contains 20 active complex scatterers, the input Signal-to-Noise Ratio (SNR) is 20 dB, and no reconstruction operator mismatch is applied in the main example. All reported results use the same simulation seed, so that the displayed scene is fixed across reconstructions and robustness sweeps.

Before turning to reconstruction, it is useful to examine the singular-value distribution of the explicit SAR sensing matrix. Figure 3 compares it with that of a Gaussian random matrix of the same size, included as a standard idealized compressed-sensing reference [162,163]. The matrix $\mathbf{\Phi}_{\mathsf{SAR}}$ is instead constructed directly from the SAR acquisition geometry and propagation model, in line with physically grounded SAR forward modeling and recent physics-aware learning approaches [1,15,16,115]. The comparison shows that the SAR matrix is not spectrally pathological, but it is distinctly more structured than the random reference. In the present configuration, both matrices are full row rank, while the SAR matrix has the larger condition number (≈27.7 versus ≈17.3). This confirms that the forward model considered here is informative yet non-generic, which motivates reconstruction rules that explicitly incorporate SAR physics.

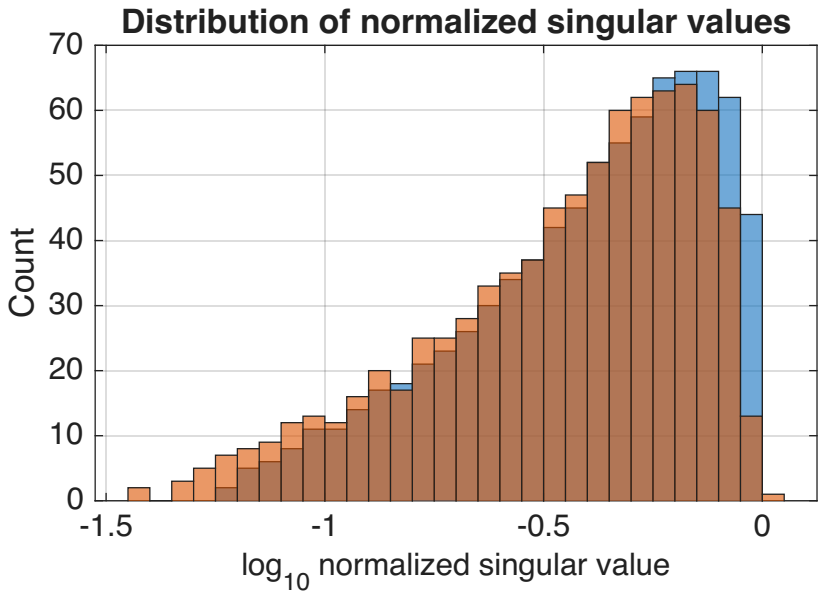


Figure 3. Histogram of normalized singular values for the explicit SAR sensing matrix (orange), compared with a complex Gaussian reference matrix of the same dimensions with i.i.d. entries distributed as $\mathcal{CN}(0, 1/N)$ (blue), with some overlap between the two distributions (brown). A larger fraction of the normalized singular values lies near zero for the SAR matrix than for the Gaussian reference, which indicates stronger deterministic structure, higher coherence, and poorer conditioning. The Gaussian matrix is included as a diagnostic benchmark for matrix structure and is not used as a forward model in the reconstruction experiment.

We compare six methods. The matched-filter baseline is

$$\hat{\boldsymbol{x}}_{\mathrm{MF}} = \boldsymbol{\Phi}_{\mathrm{SAR}}^{\mathsf{H}} \boldsymbol{y}. \tag{7}$$

The classical sparse baselines solve

$$\hat{\boldsymbol{x}}_{\ell_1} = \arg\min_{\boldsymbol{x}} \left\{ \|\boldsymbol{y} - \boldsymbol{\Phi}_{\mathrm{SAR}} \boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_1 \right\} \tag{8}$$

using a finite-depth ℓ_1-Iterative Soft-Thresholding Algorithm (ISTA) run and two stronger optimization references: long-run ℓ_1-ISTA and long-run FISTA. These additional baselines are included to separate finite-depth unrolling effects from the behavior of more nearly converged variational solvers. The black-box Neural Network (NN) learns a direct map

$$\hat{\boldsymbol{x}}_{\mathrm{BB}} = f_{\boldsymbol{\theta}}(\boldsymbol{y}). \tag{9}$$

The physics-aware network unrolls a model-based iteration and keeps the SAR matrix explicit in every layer:

$$\boldsymbol{x}^{(k+1)} = \mathcal{S}_{\tau^{(k)}}\left(\boldsymbol{x}^{(k)} + \rho^{(k)} \boldsymbol{\Phi}_{\mathrm{SAR}}^{\mathsf{H}} (\boldsymbol{y} - \boldsymbol{\Phi}_{\mathrm{SAR}} \boldsymbol{x}^{(k)})\right), \tag{10}$$

where the trainable parameters are the layer-wise step sizes $\rho^{(k)}$ and thresholds $\tau^{(k)}$. Thus, the learned model is not a generic black-box denoiser grafted onto the inverse problem; it is a finite-depth SAR reconstruction rule whose data-consistency term remains tied to the physical sensing matrix.

The point of the experiment is to compare these methods under the same limited computational budget. Both learned models are trained for 50 epochs with batch size 64. The physics-aware model uses $T = 8$ unrolled layers. The finite-depth sparse baseline uses 60 ℓ_1-ISTA iterations with $\lambda = 0.02$, while the stronger classical references use 500 iterations of ℓ_1-ISTA and 500 iterations of FISTA. Thus, the experiment distinguishes two questions: whether a physics-aware network can improve over a finite-depth model-based baseline, and how it compares with substantially longer classical optimization when computational latency is less constrained. Additional implementation details for the illustrative numerical case studies are collected in Appendix A. Algorithm 1 summarizes the simulation protocol.

Algorithm 1 Constructive sparse SAR case study with explicit SAR matrix

Input: Scene dimensions N_x, N_y, SAR sensing matrix $\boldsymbol{\Phi}_{\mathrm{SAR}}$, training and test set sizes, nominal input SNR, sampling-ratio set $\mathcal{S}$, input-SNR set $\mathcal{Q}$, mismatch set $\mathcal{M}$
Output: Reconstruction metrics and representative reconstructions for six methods
1: Generate sparse complex scenes $\{\boldsymbol{x}_i\}$ on the discretized SAR grid
2: Form noisy training measurements $\boldsymbol{y}_i = \boldsymbol{\Phi}_{\mathrm{SAR}} \boldsymbol{x}_i + \boldsymbol{z}_i$ at the nominal SNR
3: Train the black-box NN from pairs $(\boldsymbol{y}_i, \boldsymbol{x}_i)$
4: Train the physics-aware unrolled NN using $\boldsymbol{\Phi}_{\mathrm{SAR}}$ and the same training scenes
5: **for** each test scene **do**
6: Form noisy and possibly masked measurements from $\boldsymbol{\Phi}_{\mathrm{SAR}}$
7: Reconstruct using (i) matched filtering, (ii) finite-depth ℓ_1-ISTA, (iii) long-run ℓ_1-ISTA, (iv) long-run FISTA, (v) black-box NN, and (vi) physics-aware unrolling
8: Compute Normalized Mean Square Error (NMSE), Peak Signal-to-Noise Ratio (PSNR), and support recovery metrics
9: **end for**
10: Repeat the evaluation while varying sampling ratio, input SNR, and reconstruction operator mismatch
11: Report representative reconstructions, training curves, and robustness plots

Figure 4 shows a representative reconstruction example. The matched-filter output is dominated by artifact-like fluctuations and poor contrast, and the black-box NN remains visually unstable. The finite-depth physics-aware network produces a structured reconstruction, but the newly added long-run ℓ_1-ISTA and FISTA baselines are stronger classical references and achieve the best Normalized Mean Square Error (NMSE) in this controlled sparse reconstruction setting. Accordingly, the revised interpretation is deliberately conservative: the physics-aware network is useful as a fixed-depth model-guided reconstruction rule, but it is not claimed to dominate converged variational solvers.

The corresponding test set metrics are summarized in Table 3; the training convergence is shown in Figure 5, and the robustness analysis is presented in Figure 6. The main conclusion is now separated into two parts. First, explicit model guidance is essential because the unconstrained black-box network remains poor. Second, fair comparison against stronger classical optimization shows that fully converged iterative solvers remain highly competitive and should be included whenever the claim concerns reconstruction accuracy rather than finite-depth inference.

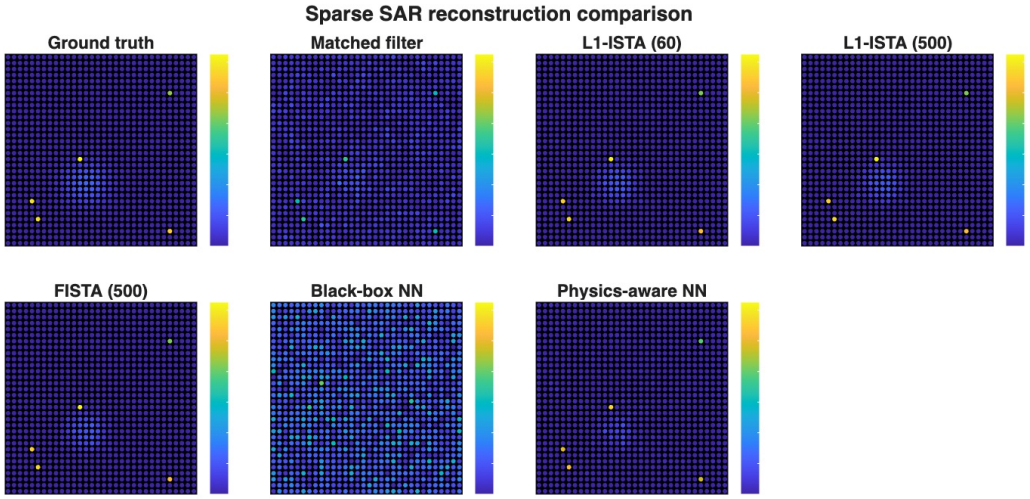


Figure 4. Sparse SAR reconstruction comparison on a representative synthetic scene. All panels display the normalized magnitude $|x|/\max|x_{\text{true}}|$ using a common color scale, with dark colors corresponding to values near 0 and bright yellow corresponding to values near 1. In addition to the finite-depth ℓ_1-ISTA baseline, long-run ℓ_1-ISTA and FISTA are included as stronger classical optimization references. The results show that fully converged iterative solvers remain highly competitive in NMSE, while the physics-aware network should be interpreted as a fixed-depth model-guided reconstruction strategy rather than as a replacement for converged variational optimization.

Table 3. Representative nominal-point results for the constructive sparse SAR experiment with the explicit SAR sensing matrix. The values are taken from the revised MATLAB (version R2026A) simulation output. Lower NMSE is better, higher Peak Signal-to-Noise Ratio (PSNR) and support scores are better.

Method	NMSE (dB)	PSNR (dB)	Precision	Recall	F1
Matched filter	−1.15	24.18	0.964	0.718	0.809
ℓ_1-ISTA (60)	−16.55	39.65	1.000	0.725	0.829
ℓ_1-ISTA (500)	−16.62	39.72	1.000	0.725	0.829
FISTA (500)	−16.62	39.72	1.000	0.725	0.829
Black-box NN	10.15	13.00	0.009	0.438	0.017
Physics-aware NN	−10.49	33.55	1.000	0.704	0.811

In addition to NMSE and Peak Signal-to-Noise Ratio (PSNR), Table 3 reports support precision, support recall, and support F1 score from the revised evaluation. These metrics

are included because sparse SAR imaging is not only an image fidelity problem: correct localization of dominant scatterers is often more relevant than average pixel error. For a thresholded support estimate $\widehat{\mathcal{S}}$ and true support $\mathcal{S}$, precision, recall, and F1 are defined as

$$P = \frac{|\widehat{\mathcal{S}} \cap \mathcal{S}|}{|\widehat{\mathcal{S}}| + \varepsilon}, \qquad R = \frac{|\widehat{\mathcal{S}} \cap \mathcal{S}|}{|\mathcal{S}| + \varepsilon}, \qquad F_1 = \frac{2PR}{P + R + \varepsilon}. \tag{11}$$

These additional metrics do not change the qualitative conclusion that long-run variational solvers are strong references in NMSE and PSNR. They also show that the physics-aware network retains high support precision and a support F1 score close to the sparse iterative baselines, making the comparison less dependent on a single image fidelity score.

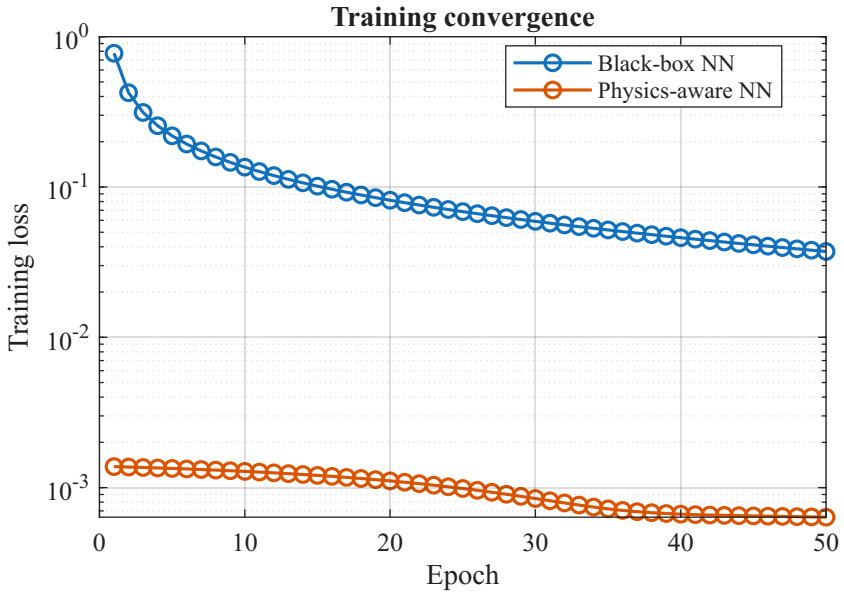


Figure 5. Training convergence for the black-box and physics-aware networks used in the constructive sparse SAR experiment. This diagnostic reports optimization behavior during learning and is not used as evidence of reconstruction superiority over the long-run classical baselines.

To move beyond a single operating point, we further evaluate the methods under controlled variations of sampling ratio, input SNR, and reconstruction operator mismatch. Figure 6 summarizes the corresponding robustness study. A key point is that the main conclusion of the SAR case study is local rather than universal: the physics-aware method is useful in a specific finite-budget SAR regime, not as a blanket replacement for sparse reconstruction.

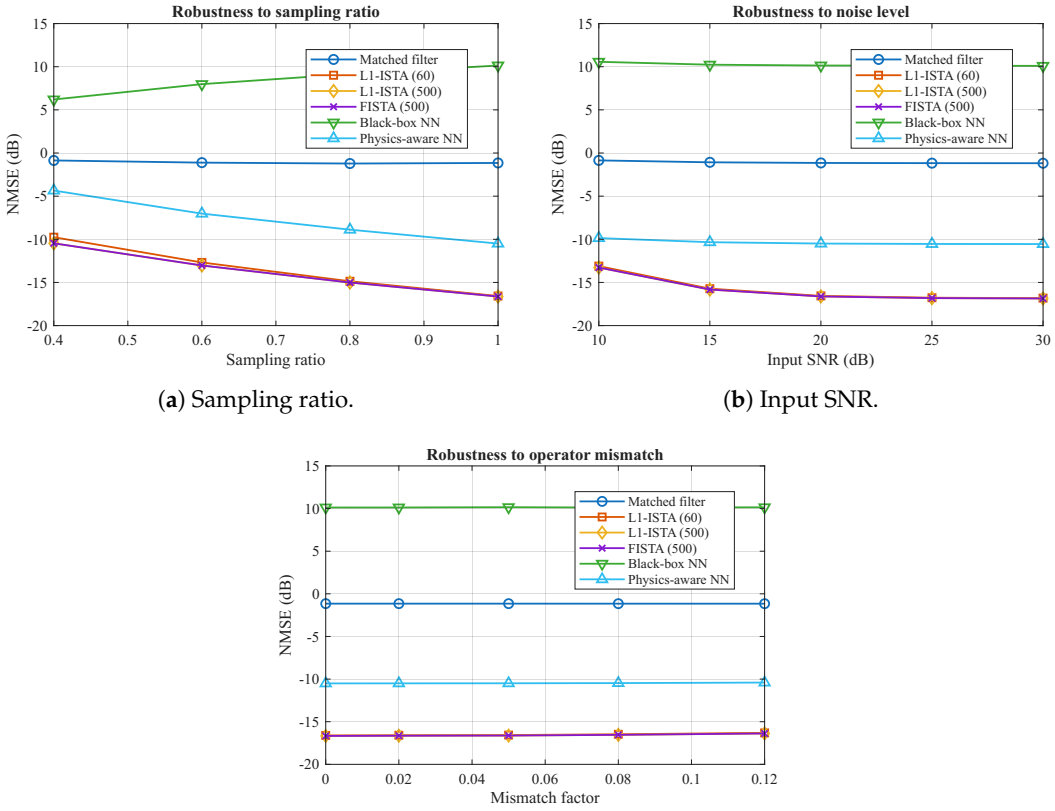
As the sampling ratio increases, the long-run ℓ_1-ISTA and FISTA baselines achieve the lowest NMSE, while the physics-aware network remains substantially better than the black-box network and the matched filter. This behavior clarifies the scope of the finite-depth result: learned model guidance can be useful under a fixed inference budget, but enough iterations of a well-matched sparse solver remain a very strong baseline.

The noise sweep leads to the same qualitative conclusion. Across the tested input SNR range, long-run ℓ_1-ISTA and FISTA are the strongest methods in NMSE; the physics-aware network provides a structured fixed-depth reconstruction; the matched filter is weak; and the black-box NN is essentially noncompetitive. Thus, the revised experiment supports physics-aware modeling as a way to avoid unconstrained black-box failure, not as a universal replacement for converged sparse optimization.

The mismatch sweep is comparatively flat for the plotted NMSE curves under the tested perturbation range, but it still serves as a useful diagnostic because it evaluates all

methods under an explicitly perturbed reconstruction operator. The figure should therefore be read as a controlled operator-perturbation check rather than as proof of broad out-of-distribution robustness. Therefore, the degradation is not explained by rank collapse or catastrophic ill-conditioning of the perturbed matrix. Instead, it is a genuine forward-model mismatch effect: the learned reconstruction rule is sensitive to inconsistency between the SAR matrix used during training and the one used during reconstruction.

Taken together, these results support a more cautious conclusion. In the present sparse SAR setup, explicit model guidance is clearly preferable to unconstrained black-box regression, and a finite-depth physics-aware network provides structured reconstructions within a fixed computational budget. However, long-run ℓ_1-ISTA and FISTA achieve the lowest NMSE, so converged variational optimization remains a strong reference whenever latency is less constrained. This is exactly the kind of regime-specific conclusion that physics-aware learning should aim to expose.



(**a**) Sampling ratio. (**b**) Input SNR.

(**c**) Reconstruction-matrix mismatch.

Figure 6. Robustness analysis for the constructive sparse SAR experiment. Each panel includes the matched filter, finite-depth ℓ_1-ISTA, long-run ℓ_1-ISTA, long-run FISTA, black-box NN, and physics-aware NN, with the legend identifying the line style used for each method. The stronger classical baselines provide the lowest NMSE, while the learned physics-aware model should be interpreted as a fixed-depth reconstruction strategy.

5.3. Illustrative InSAR Case Study: Phase-Domain Estimation Under Wrapping and Coherence Variation

To complement the sparse-SAR imaging example and align the numerical material with the full SAR/InSAR scope of the paper, we next consider a compact phase-domain case study based on synthetically generated interferometric data. The goal is again illustrative rather than benchmark oriented. Here the emphasis is on wrapped phase, coherence variation, and local complex-domain regularity, so that the key question becomes how well different estimators respect phase geometry and interferometric structure.

The synthetic scene consists of a smoothly varying unwrapped phase field with superimposed localized structures and a spatially varying coherence map. A noisy complex interferogram is generated as

$$z = \gamma\, \mathrm{e}^{\mathrm{j}\phi} + w, \tag{12}$$

where ϕ denotes the target phase, γ is the coherence, and w is circular complex Gaussian noise. Thus, unlike the sparse-SAR case, the quantity of interest is not a sparse reflectivity field but a phase-bearing complex phasor field whose interpretation depends directly on wrapping and coherence. The reported example includes the true unwrapped phase, the true coherence, the observed wrapped phase, and the observed interferogram magnitude, together with three estimation strategies: a boxcar multilook baseline, a black-box neural network, and a physics-aware iterative estimator.

For the black-box estimator, let $\mathcal{N}_i$ denote the square patch of radius $r = 1$ centered at pixel i. From the observed interferogram z, we form the real-valued input vector

$$\boldsymbol{u}_i = \begin{bmatrix} \Re\{z_j\}_{j\in\mathcal{N}_i} \\ \Im\{z_j\}_{j\in\mathcal{N}_i} \end{bmatrix} \in \mathbb{R}^{2(2r+1)^2}. \tag{13}$$

The black-box model is a one-hidden-layer multilayer perceptron

$$\boldsymbol{h}_i = \sigma(\boldsymbol{W}_1\boldsymbol{u}_i + \boldsymbol{b}_1), \qquad \boldsymbol{q}_i = \boldsymbol{W}_2\boldsymbol{h}_i + \boldsymbol{b}_2, \tag{14}$$

where $\sigma(\cdot)$ is the ReLU activation. The output $\boldsymbol{q_i}$ is projected onto the unit circle through its two real components,

$$\hat{\boldsymbol{p}}_i = \frac{\boldsymbol{q}_i}{\|\boldsymbol{q}_i\|_2 + \varepsilon} = \begin{bmatrix} \cos\hat{\phi}_i \\ \sin\hat{\phi}_i \end{bmatrix}, \tag{15}$$

so that the prediction corresponds to the unit modulus phasor at the center pixel. Training minimizes the mean-squared phasor loss

$$\mathcal{L}_{\mathrm{BB}} = \frac{1}{N_{\mathrm{tr}}} \sum_i \left\| \hat{\boldsymbol{p}}_i - \begin{bmatrix} \cos\phi_i \\ \sin\phi_i \end{bmatrix} \right\|_2^2. \tag{16}$$

The physics-aware estimator instead operates directly on complex unit modulus phasors. Let

$$u_i^{(0)} = \frac{z_i}{|z_i| + \varepsilon} \tag{17}$$

denote the normalized observed phasor. At iteration k, a local complex average is formed as

$$m_i^{(k)} = \frac{1}{|\mathcal{N}_i|} \sum_{j\in\mathcal{N}_i} u_j^{(k)}, \tag{18}$$

and a discrete Laplacian-like smoothing correction is computed as

$$\Delta u_i^{(k)} = \sum_{j\in\mathcal{N}_i\setminus\{i\}} \left(u_j^{(k)} - u_i^{(k)}\right). \tag{19}$$

The unnormalized update is then

$$v_i^{(k+1)} = (1-\beta)\left[(1-\rho)u_i^{(k)} + \rho\, m_i^{(k)} + \tau\, \Delta u_i^{(k)}\right] + \beta\, u_i^{(0)}, \tag{20}$$

where ρ controls local averaging, τ controls the smoothing correction, and β enforces data consistency with the observed interferogram. The phase estimate is constrained to remain on the complex unit circle by projection,

$$u_i^{(k+1)} = \frac{v_i^{(k+1)}}{|v_i^{(k+1)}| + \varepsilon}, \tag{21}$$

and the final phase estimate is $\hat{\phi}_i = \arg u_i^{(K)}$ after K iterations. Its hyperparameters are selected by grid search over the update coefficients.

In the present implementation the phase grid is 48×48, the training and test sets contain 120 and 24 synthetic scenes, respectively, the patch radius is 1, the hidden-layer size is 40, the black-box model is trained for 12 epochs with batch size 256 and learning rate 5×10^{-3}, and the physics-aware estimator uses 8 iterations. The parameter grids are $\rho \in \{0.10, 0.20, 0.35, 0.50\}$, $\tau \in \{0.25, 0.40, 0.55, 0.70\}$, and $\beta \in \{0.20, 0.35, 0.50\}$.

Figures 7–9 summarize the resulting observations, reconstructions, and robustness trends. For the displayed example, the boxcar multilook baseline attains an RMSE of about 4.3°, the black-box neural network attains about 8.5°, and the physics-aware estimator attains about 6.6°. The revised evaluation also computes circular phase mean absolute error and coherence-stratified phase RMSE, separating performance in low- and high-coherence regions. These metrics are more appropriate than a single global RMSE when phase wrapping and coherence variation are central to the task. Thus, in this example, the learned methods do not outperform the simple phase-preserving baseline, although the physics-aware design improves substantially over the black-box model. This is precisely the kind of result that is valuable in a perspective paper: embedding physics can regularize learning, but naive learning does not automatically surpass classical structure-exploiting estimators, especially in low-data and phase-sensitive settings.

The training and tuning curves further clarify the behavior of the two learned approaches. The black-box training loss decreases rapidly during optimization, while the physics-aware phasor loss decreases across the unrolled iterations, indicating stable convergence of the model-driven updates. Robustness sweeps with respect to noise level and coherence show the expected degradation as observation quality worsens, again highlighting that phase-aware learning should be judged not only by nominal accuracy but also by sensitivity to physically meaningful perturbations.

Overall, the InSAR example reinforces one of the main messages of the paper: in coherent imaging problems, useful inductive bias often lies less in generic image statistics than in respecting the geometry of the signal itself, here through wrapped phase, coherence, and unit modulus complex structure. At the same time, the example also warns against overstating the benefits of learned models in the absence of carefully matched priors and robust uncertainty handling. This case study is intentionally limited to a synthetic single-interferogram phase estimation setting: it validates phase geometry and coherence-aware priors, but it does not cover real InSAR complications such as time-series processing, nonlinear deformation constraints, atmospheric and orbital phase screens, unwrapping ambiguities across long temporal stacks, or sensor-specific preprocessing artifacts.

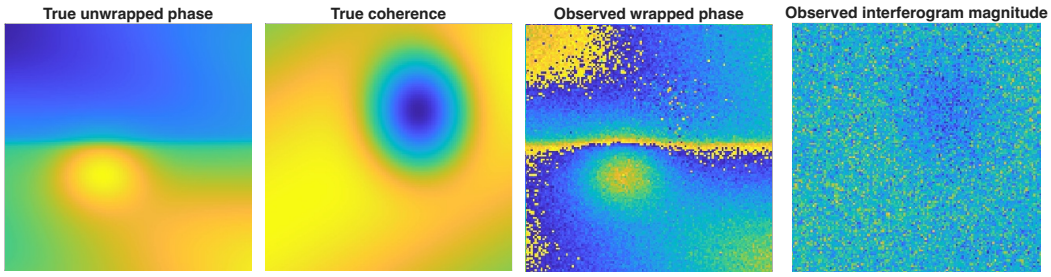


Figure 7. Observation-side quantities for the illustrative InSAR phase-domain case study. The synthetic example includes a smoothly varying unwrapped phase field, a nonuniform coherence map, the corresponding wrapped phase observation, and the observed interferogram magnitude. Colors follow the MATLAB parula colormap, where dark blue corresponds to lower phase values and yellow to higher phase values in the phase panels.

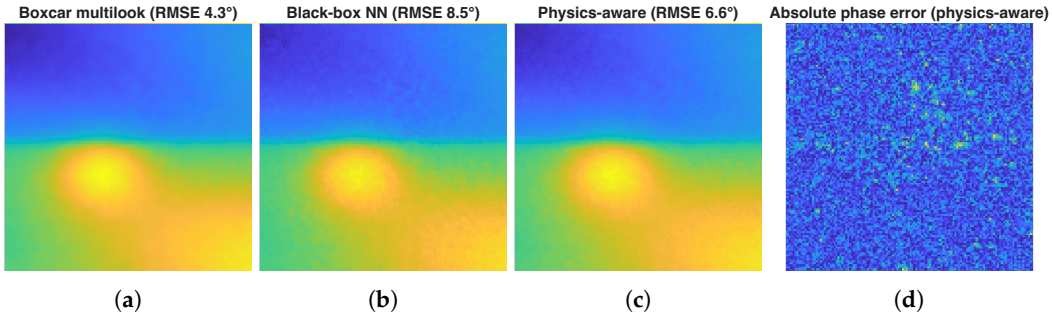


Figure 8. Representative reconstructions for the illustrative InSAR case study. In the shown scene, the classical boxcar multilook baseline remains strongest, while the physics-aware estimator improves over the black-box neural network but does not surpass the phase-preserving baseline. Colors follow the MATLAB parula colormap, with dark blue indicating lower estimated phase values and yellow indicating higher values.

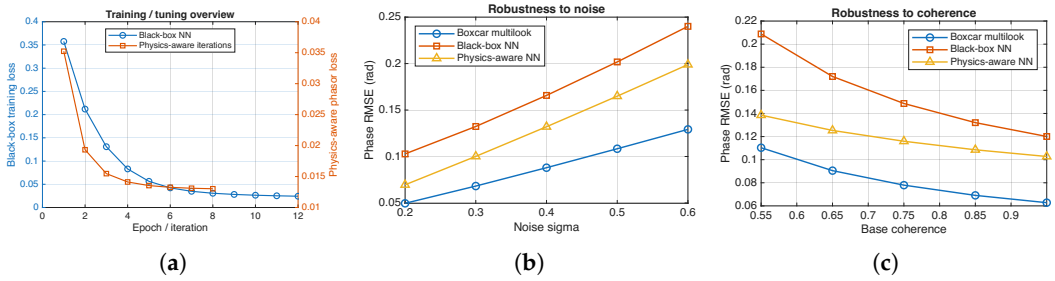


Figure 9. Optimization and robustness summaries for the illustrative InSAR phase-domain case study. The black-box loss decreases during training, the physics-aware phasor loss decreases across iterations, and both learned approaches degrade as observation quality worsens.

5.4. Comparative Interpretation of the Two Examples

Taken together, the two examples support a more balanced message than either one would in isolation. In the sparse-SAR case, the physics-aware learned update clearly outperforms the black-box model, but the added long-run ℓ_1-ISTA and FISTA baselines show that converged variational optimization remains stronger in NMSE. In the InSAR phase-domain case, the physics-aware estimator again improves on the black-box model, but a simple structure-preserving baseline remains strongest. This contrast is useful. It shows that the value of physics-aware learning depends on the task regime, the quality of the forward model, and the strength of the classical baseline.

The two case studies therefore play complementary roles in the paper. The SAR example demonstrates that explicit integration of sensing physics can avoid black-box failure and produce structured finite-depth reconstructions, while also showing that stronger classical baselines are necessary for fair accuracy claims. The InSAR example demonstrates that, in phase-sensitive coherent estimation problems, embedding physics is still beneficial relative to naive learning, but does not automatically guarantee superiority over carefully chosen classical estimators. This is exactly the kind of regime-dependent conclusion that a physics-aware perspective should expose.

6. Open Challenges and Future Directions

The following points are perspective-level research directions rather than direct empirical conclusions from the two illustrative case studies. The examples above evaluate reconstruction quality, support recovery, and phase estimation behavior in controlled settings; they do not by themselves validate operational trustworthiness, calibrated uncertainty, or out-of-distribution deployment. Several open problems should therefore anchor the perspective portion of the paper. First, there is still no consensus on the best interface between physics and learning. Possible choices include explicit parametric quantities, simulated labels, complex-domain operations, differentiable forward models, and constraint-aware losses, and the right choice is likely task dependent.

Second, it remains unclear how much physics is enough. Strong inductive bias can improve data efficiency but may also reduce adaptability when the assumed model is incomplete. A useful future research agenda is therefore not simply "more physics," but *adaptive* PAL in which the model can decide when to trust prior structure and when to deviate from it. The multi-constraint detection literature already points in this direction by proposing adaptive fusion strategies that select among polarimetric, scattering, signal-domain, and resolution features according to imaging conditions [16]. This idea can be generalized beyond detection to broader SAR/InSAR learning systems.

Third, InSAR-specific topology remains under-theorized relative to amplitude-image interpretation. Wrapped phase, discontinuities, coherence loss, and multitemporal consistency create learning problems that differ substantially from optical vision benchmarks and deserve a more explicit conceptual treatment. The recent InSAR learning literature already points in this direction, ranging from phase/coherence estimation and filtering to unwrapping networks and time-series deformation models [23,25,30,38,41,43,45]. A particularly promising avenue is to integrate phase geometry, uncertainty, and temporal consistency into architectures that remain interpretable at the interferometric level.

Finally, trustworthy operational deployment requires more than average benchmark accuracy. For high-impact applications such as geohazard monitoring or infrastructure surveillance, useful systems should provide calibrated uncertainty, out-of-distribution awareness, physically meaningful failure analysis, and careful human-in-the-loop validation. These requirements are not peripheral; they are central to whether PAL in SAR and InSAR can move from promising demonstrations to dependable operational use.

7. Conclusions

Deep learning has already become a major methodological force in SAR and InSAR remote sensing, but the field is moving beyond straightforward transfers from optical computer vision. The central argument of this paper is that the most effective learning systems in radar remote sensing are neither purely model-based nor purely data-driven. Instead, they combine flexible representation learning with physical structure derived from the sensing process, including scattering behavior, signal-domain information, resolution effects, and interferometric geometry. A useful way to understand this development

is to connect two complementary perspectives: the physical constraint dimensions that define what information matters in radar observations, and the integration modes that determine how that structure enters the learning system. From this viewpoint, physics-aware representations, model-guided architectures, and learning-assisted inverse methods are not isolated trends, but parts of a broader methodological shift toward more robust, interpretable, and trustworthy radar learning. The two illustrative case studies support this interpretation from complementary perspectives: the sparse-SAR example shows that physically structured learning can avoid unconstrained black-box failure but should be compared against strong long-run ISTA/FISTA baselines, whereas the InSAR phase-domain example shows that physics-aware learning can improve over naive black-box estimation while also clarifying regimes in which simple structure-preserving baselines remain difficult to surpass. Overall, the synthesis developed here is intended to organize currently fragmented lines of work and to clarify the practical conditions under which PAL can contribute most effectively to future SAR/InSAR systems.

Funding: This research received no external funding.

Data Availability Statement: This statement refers to the manuscript entitled "Physics-Aware Deep Learning for SAR and InSAR Remote Sensing: Models, Methods, and Open Challenges" by Giorgio Taricco. The data presented in this study are synthetically generated within the computational framework described in the manuscript. The raw data supporting the conclusions of this article will be made available by the author on request. In particular, the numerical experiments are based on computer-generated data produced from explicitly defined simulation procedures, controlled parameter settings, and fixed random seeds, so that reproducibility has been carefully taken into account in the designof the study. Requests to access the datasets should be directed to the corresponding author.

Conflicts of Interest: The author declares no conflicts of interest.

List of Acronyms

ATR	Automatic Target Recognition
CNN	Convolutional Neural Network
DL	Deep Learning
DEQ	Deep Equilibrium
FRFT	FRactional Fourier Transform
HRSID	High-Resolution SAR Images Dataset
InSAR	Interferometric Synthetic Aperture Radar
ISTA	Iterative Soft-Thresholding Algorithm
MSTAR	Moving and Stationary Target Acquisition and Recognition
MoE	Mixture of Experts
NMSE	Normalized Mean Square Error
NN	Neural Network
PAL	Physics-Aware Learning
PolSAR	Polarimetric Synthetic Aperture Radar
PnP	Plug-and-Play
PSF	Point Spread Function
PSNR	Peak Signal-to-Noise Ratio
RD	Range–Doppler
SAR	Synthetic Aperture Radar
SNR	Signal-to-Noise Ratio
SRSD	SAR Rotated Ship Detection Dataset
SSDD	SAR Ship Detection Dataset

Appendix A. Implementation Details of the Numerical Case Study

The numerical case study uses a simplified complex-valued forward model, but the learning components are implemented in real arithmetic. Accordingly, each complex measurement vector $\boldsymbol{y} \in \mathbb{C}^M$ and scene vector $\boldsymbol{x} \in \mathbb{C}^N$ is converted into a stacked real form,

$$\tilde{\boldsymbol{y}} = \begin{bmatrix} \Re\{\boldsymbol{y}\} \\ \Im\{\boldsymbol{y}\} \end{bmatrix} \in \mathbb{R}^{2M}, \qquad \tilde{\boldsymbol{x}} = \begin{bmatrix} \Re\{\boldsymbol{x}\} \\ \Im\{\boldsymbol{x}\} \end{bmatrix} \in \mathbb{R}^{2N}, \tag{A1}$$

so that the black-box model can be trained as a standard real-valued regressor.

The black-box NN is a feedforward network with five hidden layers and rectified linear unit activations. For a mini-batch of size B, let $\boldsymbol{A}_{\ell-1} \in \mathbb{R}^{n_{\ell-1} \times B}$ denote the matrix whose columns are the activations entering layer ℓ, where $n_{\ell-1}$ is the width of the previous layer and n_ℓ is the width of the current layer. Here, ℓ indexes the trainable layers of the network, with $\ell = 1, \ldots, 6$ in the present implementation. Let $\boldsymbol{W}_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ be the corresponding weight matrix and let $\boldsymbol{b}_\ell \in \mathbb{R}^{n_\ell}$ be the bias vector. The affine part of layer ℓ is

$$\boldsymbol{Z}_\ell = \boldsymbol{W}_\ell \boldsymbol{A}_{\ell-1} + \boldsymbol{b}_\ell \boldsymbol{1}^{\mathsf{T}}, \tag{A2}$$

where $\boldsymbol{1} \in \mathbb{R}^B$ is the all-ones vector. If $\hat{\boldsymbol{X}} \in \mathbb{R}^{2N \times B}$ denotes the batch output and $\boldsymbol{X} \in \mathbb{R}^{2N \times B}$ the corresponding target batch, the training loss is the mean squared error

$$\mathcal{L}_{\mathsf{BB}} = \frac{1}{2NB} \|\hat{\boldsymbol{X}} - \boldsymbol{X}\|_F^2. \tag{A3}$$

The gradient with respect to a matrix is defined by

$$[\nabla_{\boldsymbol{W}_\ell} \mathcal{L}_{\mathsf{BB}}]_{ij} = \frac{\partial \mathcal{L}_{\mathsf{BB}}}{\partial [\boldsymbol{W}_\ell]_{ij}}. \tag{A4}$$

With this convention, backpropagation gives

$$\nabla_{\boldsymbol{W}_\ell} \mathcal{L}_{\mathsf{BB}} = \nabla_{\boldsymbol{Z}_\ell} \mathcal{L}_{\mathsf{BB}} \boldsymbol{A}_{\ell-1}^{\mathsf{T}}, \qquad \nabla_{\boldsymbol{b}_\ell} \mathcal{L}_{\mathsf{BB}} = \nabla_{\boldsymbol{Z}_\ell} \mathcal{L}_{\mathsf{BB}} \boldsymbol{1}, \tag{A5}$$

and the parameters are updated by gradient descent,

$$\boldsymbol{W}_\ell \leftarrow \boldsymbol{W}_\ell - \eta \, \nabla_{\boldsymbol{W}_\ell} \mathcal{L}_{\mathsf{BB}}, \qquad \boldsymbol{b}_\ell \leftarrow \boldsymbol{b}_\ell - \eta \, \nabla_{\boldsymbol{b}_\ell} \mathcal{L}_{\mathsf{BB}}, \tag{A6}$$

where η is the learning rate. The physics-aware unrolled model is trained in a similar mini-batch fashion, but its architecture retains the explicit SAR sensing matrix in each iteration. Using the update

$$\boldsymbol{x}^{(k+1)} = \mathcal{S}_{\tau^{(k)}}\big(\boldsymbol{x}^{(k)} + \rho^{(k)} \boldsymbol{\Phi}_{\mathsf{SAR}}^{\mathsf{H}} (\boldsymbol{y} - \boldsymbol{\Phi}_{\mathsf{SAR}} \boldsymbol{x}^{(k)})\big), \tag{A7}$$

the trainable parameters are the layer-wise step sizes $\rho^{(k)}$ and thresholds $\tau^{(k)}$. Thus, the present physics-aware model learns only a small number of scalar parameters while keeping the data-consistency term fully tied to the SAR forward model. They are learned by minimizing the same mean squared reconstruction loss over the training set,

$$\mathcal{L}_{\mathsf{PA}} = \frac{1}{2NB} \|\hat{\boldsymbol{X}}_{\mathsf{PA}} - \boldsymbol{X}\|_F^2, \tag{A8}$$

with gradients propagated backward through the unrolled iterations.

Different sampling ratios are modeled by applying a binary mask to the full measurement vector: retained samples correspond to mask entries equal to one, discarded samples

to entries equal to zero. In this way, the number of effective measurements changes while the input dimension of the learned models remains fixed. Reconstruction-matrix mismatch is modeled by perturbing the SAR reconstruction matrix relative to the matrix used to generate the data through geometry-dependent changes in altitude, carrier frequency, scene center, and platform speed. These choices allow the robustness curves reported in the paper to be generated without changing the dimensionality of the learned architectures.

References

1. Cumming, I.G.; Wong, F.H. *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*; Artech House: Norwood, MA, USA, 2005.
2. Rosen, P.A.; Hensley, S.; Joughin, I.R.; Li, F.K.; Madsen, S.N.; Rodriguez, E.; Goldstein, R.M. Synthetic Aperture Radar Interferometry. *Proc. IEEE* **2000**, *88*, 333–382. [CrossRef]
3. Hanssen, R.F. *Radar Interferometry: Data Interpretation and Error Analysis*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2001.
4. Zebker, H.; Werner, C.; Rosen, P.; Hensley, S. Accuracy of topographic maps derived from ERS-1 interferometric radar. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 823–836. [CrossRef]
5. Abdelfattah, R.; Nicolas, J. Topographic SAR interferometry formulation for high-precision DEM generation. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2415–2426. [CrossRef]
6. Massonnet, D.; Briole, P.; Arnaud, A. Deflation of mount Etna monitored by spaceborne radar interferometry. *Nature* **1995**, *375*, 567–570. [CrossRef]
7. Ruch, J.; Anderssohn, J.; Walter, T.; Motagh, M. Caldera-scale inflation of the Lazufre volcanic area, South America: Evidence from InSAR. *J. Volcanol. Geotherm. Res.* **2008**, *174*, 337–344. [CrossRef]
8. Trasatti, E.; Casu, F.; Giunchi, C.; Pepe, S.; Solaro, G.; Tagliaventi, S.; Berardino, P.; Manzo, M.; Pepe, A.; Ricciardi, G.P.; et al. The 2004–2006 uplift episode at Campi Flegrei caldera (Italy): Constraints from SBAS-DInSAR ENVISAT data and Bayesian source inference. *Geophys. Res. Lett.* **2008**, *35*, L07308. [CrossRef]
9. Massonnet, D.; Rossi, M.; Carmona, C.; Adragna, F.; Peltzer, G.; Feigl, K.; Rabaute, T. The displacement field of the Landers earthquake mapped by radar interferometry. *Nature* **1993**, *364*, 138–142. [CrossRef]
10. Peltzer, G.; Rosen, P. Surface displacement of the 17 May 1993 Eureka valley, California, earthquake observed by SAR interferometry. *Science* **1995**, *268*, 1333–1336. [CrossRef] [PubMed]
11. Ketelaar, V.B.H.G. Satellite Radar Interferometry. In *Remote Sensing and Digital Image Processing*; Springer: Assen, The Netherlands, 2009; Volume 14.
12. Gernhardt, S.; Bamler, R. Deformation monitoring of single buildings using meter-resolution SAR data in PSI. *ISPRS J. Photogramm. Remote Sens.* **2012**, *73*, 68–79. [CrossRef]
13. Montazeri, S.; Zhu, X.X.; Eineder, M.; Bamler, R. Three-dimensional deformation monitoring of urban infrastructure by tomographic SAR using multitrack TerraSAR-x data stacks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6868–6878. [CrossRef]
14. Zhu, X.X.; Bamler, R. Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography. *IEEE Geosci. Remote Sens. Lett.* **2010**, *8*, 735–739. [CrossRef]
15. Zhu, X.X.; Montazeri, S.; Ali, M.; Hua, Y.; Wang, Y.; Mou, L.; Shi, Y.; Xu, F.; Bamler, R. Deep Learning Meets SAR: Concepts, Models, Pitfalls, and Perspectives. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 143–172. [CrossRef]
16. Li, X.; Liu, L.; Wan, G.; Zheng, F.; Guo, S.; Sun, G.; Wang, Z.; Liu, X. Physics-Driven SAR Target Detection: A Review and Perspective. *Remote Sens.* **2026**, *18*, 200. [CrossRef]
17. Zhang, T.; Zeng, T.; Zhang, X. Synthetic Aperture Radar (SAR) Meets Deep Learning. *Remote Sens.* **2023**, *15*, 303. [CrossRef]
18. Passah, A.; Sur, S.N.; Abraham, A.; Kandar, D. Synthetic Aperture Radar image analysis based on deep learning: A review of a decade of research. *Eng. Appl. Artif. Intell.* **2023**, *121*, 106305. [CrossRef]
19. Peter, E.; Ang, L.M.; Seng, K.P.; Srivastava, S. Recent Advances in Deep Learning for SAR Images: Overview of Methods, Challenges, and Future Directions. *Sensors* **2026**, *26*, 1143. [CrossRef] [PubMed]
20. Fontes, M.; Bakon, M.; Cunha, A.; Sousa, J.J. Deep Learning Meets InSAR for Infrastructure Monitoring: A Systematic Review of Models, Applications, and Challenges. *Sensors* **2025**, *25*, 7169. [CrossRef] [PubMed]
21. Liu, X.; Zhang, Y.; Shan, X.; Wang, Z.; Gong, W.; Zhang, G. Deep Learning for Automatic Detection of Volcanic and Earthquake Deformation Signals from InSAR Images. *Remote Sens.* **2025**, *17*, 686. [CrossRef]
22. Sica, F. Deep Learning for InSAR Processing. In *Deep Learning for Synthetic Aperture Radar Remote Sensing*; Elsevier: Amsterdam, The Netherlands, 2026; pp. 147–173. [CrossRef]
23. Sun, X.; Zimmer, A.; Mukherjee, S.; Kottayil, N.K.; Ghuman, P.; Cheng, I. DeepInSAR: A Deep Learning Framework for SAR Interferometric Phase Restoration and Coherence Estimation. *Remote Sens.* **2020**, *12*, 2340. [CrossRef]

24. Mukherjee, S.; Zimmer, A.; Sun, X.; Ghuman, P.; Cheng, I. An Unsupervised Generative Neural Approach for InSAR Phase Filtering and Coherence Estimation. *arXiv* **2020**, arXiv:2001.09631.
25. Murdaca, G.; Rucci, A.; Prati, C. Deep Learning for InSAR Phase Filtering: An Optimized Framework for Phase Unwrapping. *Remote Sens.* **2022**, *14*, 4956. [CrossRef]
26. Pu, L.; Zhang, X.; Zhou, L.; Li, L.; Shi, J.; Wei, S. Nonlocal Feature Selection Encoder–Decoder Network for Accurate InSAR Phase Filtering. *Remote Sens.* **2022**, *14*, 1174. [CrossRef]
27. Wang, N.; Zhang, X.; Zhang, T.; Pu, L.; Zhan, X.; Xu, X.; Hu, Y.; Shi, J.; Wei, S. A Sparse-Model-Driven Network for Efficient and High-Accuracy InSAR Phase Filtering. *Remote Sens.* **2022**, *14*, 2614. [CrossRef]
28. Yang, W.; He, Y.; Yao, S.; Zhang, L.; Cao, S.; Wen, Z. An InSAR Interferogram Filtering Method Based on Multi-Level Feature Fusion CNN. *Sensors* **2022**, *22*, 5956. [CrossRef] [PubMed]
29. Zhang, X.; Peng, C.; Li, Z.; Zhang, Y.; Liu, Y.; Wang, Y. MOMFNet: A Deep Learning Approach for InSAR Phase Filtering Based on Multi-Objective Multi-Kernel Feature Extraction. *Sensors* **2024**, *24*, 7821. [CrossRef] [PubMed]
30. Li, H.; Wang, J.; Ai, C.; Wu, Y.; Ren, X. NBDNet: A Self-Supervised CNN-Based Method for InSAR Phase and Coherence Estimation. *Remote Sens.* **2025**, *17*, 1181. [CrossRef]
31. Sica, F.; Pulella, A.; Schmitt, M. Beyond Backscatter: InSAR Coherence from Detected SAR Images. *arXiv* **2026**, arXiv:2606.07374.
32. Sica, F.; Calvanese, F.; Scarpa, G.; Rizzoli, P. A CNN-Based Coherence-Driven Approach for InSAR Phase Unwrapping. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 4003705. [CrossRef]
33. Li, L.; Zhang, H.; Tang, Y.; Wang, C.; Gu, F. InSAR Phase Unwrapping by Deep Learning Based on Gradient Information Fusion. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 4502305. [CrossRef]
34. Yu, H.; Zhang, X.; Shi, J.; Wei, S. A Novel Quality-Guided Two-Dimensional InSAR Phase Unwrapping Method via Deep Learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7840–7856. [CrossRef]
35. Zhou, L.; Yu, H.; Pascazio, V.; Xing, M. PU-GAN: A One-Step 2-D InSAR Phase Unwrapping Based on Conditional Generative Adversarial Network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5221510. [CrossRef]
36. Yang, T.; Zhao, W.; Jin, S.; Liu, G. A Robust InSAR Phase Unwrapping Method via Improving the pix2pix Network. *Remote Sens.* **2023**, *15*, 4885. [CrossRef]
37. Kumar, S.V.; Sun, X.; Wang, Z.; Goldsbury, R.; Cheng, I. A U-Net Approach for InSAR Phase Unwrapping and Denoising. *Remote Sens.* **2023**, *15*, 5081. [CrossRef]
38. Yang, W.; He, Y.; Zhu, Q.; Zhang, L.; Jin, L. Unwrap-Net: A Deep Neural Network-Based InSAR Phase Unwrapping Method Assisted by Airborne LiDAR Data. *ISPRS J. Photogramm. Remote Sens.* **2024**, *218*, 510–529. [CrossRef]
39. Mu, J.; Wang, Y.; Zhan, S.; Yao, G.; Liu, K.; Zhu, Y.; Wang, L. A Coherence-Guided InSAR Phase Unwrapping Method with Cycle-Consistent Adversarial Networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 2690–2704. [CrossRef]
40. Yang, K.; Yuan, Z.; Xing, X.; Chen, L. Deep-Learning-Based Mask-Cut Method for InSAR Phase Unwrapping. *IEEE J. Miniaturization Air Space Syst.* **2023**, *4*, 221–230. [CrossRef]
41. Zhou, L.; Yu, H. MoDL-PU: Model-Based Deep Learning for InSAR Phase Unwrapping. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5206711. [CrossRef]
42. He, Y.; Yang, W.; Zhu, Q.; Zhang, L.; Jin, L. An InSAR Phase Unwrapping Framework for Large-Scale and Complex Events. *arXiv* **2026**, arXiv:2603.21378.
43. Wang, T.; Zhang, Q.; Wu, Z. A Deep-Learning-Facilitated, Detection-First Strategy for Operationally Monitoring Localized Deformation with Large-Scale InSAR. *Remote Sens.* **2023**, *15*, 2310. [CrossRef]
44. Stephenson, O.L.; Köhne, T.; Zhan, E.; Cahill, B.E.; Yun, S.H.; Ross, Z.E.; Simons, M. Deep Learning-Based Damage Mapping with InSAR Coherence Time Series. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5207917. [CrossRef]
45. Soni, R.; Alam, M.S.; Vishwakarma, G.K. Prediction of InSAR Deformation Time-Series Using Improved LSTM Deep Learning Model. *Sci. Rep.* **2025**, *15*, 5333. [CrossRef] [PubMed]
46. Chen, S.; Wang, H.; Xu, F.; Jin, Y. Target classification using the deep convolutional networks for SAR images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4806–4817. [CrossRef]
47. Ding, J.; Chen, B.; Liu, H.; Huang, M. Convolutional neural network with data augmentation for SAR target recognition. *IEEE Geosci. Remote Sens. Lett.* **2015**, *13*, 364–368. [CrossRef]
48. Du, K.; Deng, Y.; Wang, R.; Zhao, T.; Li, N. SAR ATR based on displacement-and rotation-insensitive CNN. *Remote Sens. Lett.* **2016**, *7*, 895–904. [CrossRef]
49. Zhang, F.; Hu, C.; Yin, Q.; Li, W.; Li, H.C.; Hong, W. Multi-Aspect-Aware Bidirectional LSTM Networks for Synthetic Aperture Radar Target Recognition. *IEEE Access* **2017**, *5*, 26880–26891. [CrossRef]
50. Wilmanski, M.; Kreucher, C.; Lauer, J. Modern approaches in deep learning for SAR ATR. In *Algorithms for Synthetic Aperture Radar Imagery XXIII*; SPIE: Baltimore, MD, USA, 2016; Volume 9843, p. 98430. [CrossRef]
51. Wagner, S. SAR ATR by a combination of convolutional neural network and support vector machines. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2861–2872. [CrossRef]

52. Gao, F.; Huang, T.; Sun, J.; Wang, J.; Hussain, A.; Yang, E. A new algorithm for SAR image target recognition based on an improved deep convolutional neural network. *Cogn. Comput.* **2019**, *11*, 809–824. [CrossRef]
53. Gao, F.; Huang, T.; Wang, J.; Sun, J.; Yang, E.; Hussain, A. Combining deep convolutional neural network and SVM to SAR image target recognition. In Proceedings of the 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter, UK, 21–23 June 2017; pp. 1082–1085. [CrossRef]
54. Huang, Z.; Datcu, M.; Pan, Z.; Lei, B. Deep SAR-Net: Learning Objects from Signals. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 179–193. [CrossRef]
55. Liu, Q.; Lang, L. MMFF: Multi-Manifold Feature Fusion Based Neural Networks for Target Recognition in Complex-Valued SAR Imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *180*, 151–162. [CrossRef]
56. Liu, J.; Xing, M.; Yu, H.; Sun, G. EFTL: Complex Convolutional Networks with Electromagnetic Feature Transfer Learning for SAR Target Recognition. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5209811. [CrossRef]
57. Huang, Z.; Yao, X.; Liu, Y.; Dumitru, C.O.; Datcu, M.; Han, J. Physically Explainable CNN for SAR Image Classification. *ISPRS J. Photogramm. Remote Sens.* **2022**, *190*, 25–37. [CrossRef]
58. Huang, Z.; Wu, C.; Yao, X.; Zhao, Z.; Huang, X.; Han, J. Physics Inspired Hybrid Attention for SAR Target Recognition. *ISPRS J. Photogramm. Remote Sens.* **2024**, *208*, 12–29. [CrossRef]
59. Li, C.; Du, L.; Li, Y.; Song, J. A Novel SAR Target Recognition Method Combining Electromagnetic Scattering Information and GCN. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 4508705. [CrossRef]
60. Li, Y.; Du, L. Design of the Physically Interpretable Sar Target Recognition Network Combined with Electromagnetic Scattering Characteristics. In Proceedings of the IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 4988–4991. [CrossRef]
61. Kazemi, S.; Yonel, B.; Yazici, B. Deep learning for direct automatic target recognition from SAR data. In Proceedings of the 2019 IEEE Radar Conference (RadarConf), Boston, MA, USA, 22–26 April 2019; pp. 1–6. [CrossRef]
62. Rostami, M.; Kolouri, S.; Eaton, E.; Kim, K. Deep transfer learning for few-shot SAR image classification. *Remote Sens.* **2019**, *11*, 1374. [CrossRef]
63. Huang, Z.; Pan, Z.; Lei, B. What, where, and how to transfer in SAR target recognition based on deep CNNs. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 2324–2336. [CrossRef]
64. Zhang, Z.; Wang, H.; Xu, F.; Jin, Y. Complex-valued convolutional neural network and its application in polarimetric SAR image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 7177–7188. [CrossRef]
65. Mullissa, A.G.; Persello, C.; Stein, A. PolSARNet: A deep fully convolutional network for polarimetric SAR image classification. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2019**, *12*, 5300–5309. [CrossRef]
66. Li, L.; Ma, L.; Jiao, L.; Liu, F.; Sun, Q.; Zhao, J. Complex contourlet-CNN for polarimetric SAR image classification. *Pattern Recognit.* **2019**, *100*, 107110. [CrossRef]
67. Cao, Y.; Wu, Y.; Zhang, P.; Liang, W.; Li, M. Pixel-wise Polsar image classification via a novel complex-valued deep fully convolutional network. *Remote Sens.* **2019**, *11*, 2653. [CrossRef]
68. Chen, S.; Tao, C. PolSAR image classification using polarimetriC-feature-driven deep convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 627–631. [CrossRef]
69. He, C.; Tu, M.; Xiong, D.; Liao, M. Nonlinear manifold learning integrated with fully convolutional networks for PolSAR image classification. *Remote Sens.* **2020**, *12*, 655. [CrossRef]
70. Dong, H.; Zhang, L.; Zou, B. PolSAR image classification with lightweight 3D convolutional networks. *Remote Sens.* **2020**, *12*, 396. [CrossRef]
71. Xie, W.; Ma, G.; Zhao, F.; Liu, H.; Zhang, L. PolSAR image classification via a novel semi-supervised recurrent complex-valued convolution neural network. *Neurocomputing* **2020**, *388*, 255–268. [CrossRef]
72. Yu, P.; Qin, A.; Clausi, D. Unsupervised polarimetric SAR image segmentation and classification using region growing with edge penalty. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 1302–1317. [CrossRef]
73. Hoekman, D.; Vissers, M. A new polarimetric classification approach evaluated for agricultural crops. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2881–2889. [CrossRef]
74. Yang, W.; Dai, D.; Wu, J.; He, C. Weakly supervised polarimetric SAR image classification with multi-modal Markov aspect model. In Proceedings of the ISPRS TC VII Symposium—100 Years ISPRS, Vienna, Austria, 5–7 July 2010.
75. Cozzolino, D.; Martino, G.D.; Poggi, G.; Verdoliva, L. A fully convolutional neural network for low-complexity single-stage ship detection in Sentinel-1 SAR images. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 886–889. [CrossRef]
76. Schwegmann, C.; Kleynhans, W.; Salmon, B.; Mdakane, L.; Meyer, R. Very deep learning for ship discrimination in synthetic aperture radar imagery. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 104–107. [CrossRef]

77. Bentes, C.; Frost, A.; Velotto, D.; Tings, B. Ship-iceberg discrimination with convolutional neural networks in high resolution SAR images. In Proceedings of the EUSAR 2016: 11th European Conference on Synthetic Aperture Radar, Hamburg, Germany, 6–9 June 2016; pp. 1–4.
78. Ødegaard, N.; Knapskog, A.; Cochin, C.; Louvigne, J. Classification of ships using real and simulated data in a convolutional neural network. In Proceedings of the 2016 IEEE Radar Conference (RadarConf), Philadelphia, PA, USA, 2–6 May 2016; pp. 1–6. [CrossRef]
79. Liu, Y.; Zhang, M.; Xu, P.; Guo, Z. SAR ship detection using sea-land segmentation-based convolutional neural network. In Proceedings of the 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP), Shanghai, China, 18–21 May 2017; pp. 1–4. [CrossRef]
80. Li, J.; Qu, C.; Shao, J. Ship detection in SAR images based on an improved faster R-CNN. In Proceedings of the 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA), Beijing, China, 13–14 November 2017; pp. 1–6. [CrossRef]
81. Kang, M.; Ji, K.; Leng, X.; Lin, Z. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection. *Remote Sens.* **2017**, *9*, 860. [CrossRef]
82. Jiao, J.; Zhang, Y.; Sun, H.; Yang, X.; Gao, X.; Hong, W.; Fu, K.; Sun, X. A Densely Connected End-to-End Neural Network for Multiscale and Multiscene SAR Ship Detection. *IEEE Access* **2018**, *6*, 20881–20892. [CrossRef]
83. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), ICCV '15, Santiago, Chile, 7–13 December 2015; pp. 1440–1448. [CrossRef]
84. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 1137–1149. [CrossRef] [PubMed]
85. Qin, J.; Zou, B.; Li, H.; Zhang, L. Cross-Resolution SAR Target Detection Using Structural Hierarchy Adaptation and Reliable Adjacency Alignment. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5221816. [CrossRef]
86. Xiong, X.; Zhang, X.; Jiang, W.; Liu, T.; Liu, Y.; Liu, L. Lightweight Dual-Stream SAR–ATR Framework Based on an Attention Mechanism-Guided Heterogeneous Graph Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 537–556. [CrossRef]
87. Zhang, H.; Wang, W.; Deng, J.; Guo, Y.; Liu, S.; Zhang, J. MASFF-Net: Multiazimuth Scattering Feature Fusion Network for SAR Target Recognition. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 19425–19440. [CrossRef]
88. Li, B.; Liu, B.; Huang, L.; Guo, W.; Zhang, Z.; Yu, W. 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)*; IEEE: New York, NY, USA, 2017; pp. 1–5. [CrossRef]
89. Wang, Y.; Wang, C.; Zhang, H.; Dong, Y.; Wei, S. A SAR dataset of ship detection for deep learning under complex backgrounds. *Remote Sens.* **2019**, *11*, 765. [CrossRef]
90. Hou, X.; Ao, W.; Song, Q.; Lai, J.; Wang, H.; Xu, F. FUSAR-ship: A high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition. *Sci. China Inf. Sci.* **2020**, *63*, 140303. [CrossRef]
91. Sun, X.; Wang, Z.; Sun, Y.; Diao, W.; Zhang, Y.; Fu, K. Air-sarship-1.0: High resolution SAR ship detection dataset. *J. Radars* **2019**, *8*, 852–862.
92. Lee, J. Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans. Pattern Anal. Mach. Intell.* **1980**, *PAMI-2*, 165–168. [CrossRef] [PubMed]
93. Kuan, D.; Sawchuk, A.; Strand, T.; Chavel, P. Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Trans. Pattern Anal. Mach. Intell.* **1985**, *PAMI-7*, 165–177. [CrossRef] [PubMed]
94. Frost, V.; Stiles, J.; Shanmugan, K.; Holtzman, J. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Trans. Pattern Anal. Mach. Intell.* **1982**, *PAMI-4*, 157–166. [CrossRef] [PubMed]
95. Xie, H.; Pierce, L.; Ulaby, F. SAR speckle reduction using wavelet denoising and Markov random field modeling. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2196–2212. [CrossRef]
96. Argenti, F.; Alparone, L. Speckle removal from SAR images in the undecimated wavelet domain. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2363–2374. [CrossRef]
97. Achim, A.; Tsakalides, P.; Bezerianos, A. SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1773–1784. [CrossRef]
98. Argenti, F.; Lapini, A.; Bianchi, T.; Alparone, L. A tutorial on speckle reduction in synthetic aperture radar images. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–35. [CrossRef]
99. Deledalle, C.A.; Denis, L.; Tupin, F. Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. *IEEE Trans. Image Process.* **2009**, *18*, 2661–2672. [CrossRef] [PubMed]
100. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; pp. 60–65. [CrossRef]
101. Su, X.; Deledalle, C.A.; Tupin, F.; Sun, H. Two-step multi-temporal nonlocal means for synthetic aperture radar images. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 6181–6196. [CrossRef]

102. Deledalle, C.A.; Denis, L.; Tupin, F.; Reigber, A.; Jäger, M. NL-SAR: A Unified Nonlocal Framework for Resolution-Preserving (Pol)(In)SAR Denoising. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2021–2038. [CrossRef]
103. Tupin, F.; Denis, L.; Deledalle, C.A.; Ferraioli, G. Ten years of patch-based approaches for SAR imaging: A review. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 5105–5108. [CrossRef]
104. Parrilli, S.; Poderico, M.; Angelino, C.V.; Verdoliva, L. A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 606–616. [CrossRef]
105. Cozzolino, D.; Verdoliva, L.; Scarpa, G.; Poggi, G. Nonlocal CNN SAR image despeckling. *Remote Sens.* **2020**, *12*, 1006. [CrossRef]
106. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* **2017**, *26*, 3142–3155. [CrossRef] [PubMed]
107. Zhang, Q.; Yuan, Q.; Li, J.; Yang, Z.; Ma, X. Learning a dilated residual network for SAR image despeckling. *Remote Sens.* **2018**, *10*, 196. [CrossRef]
108. Yue, D.X.; Xu, F.; Jin, Y.Q. SAR despeckling neural network with logarithmic convolutional product model. *Int. J. Remote Sens.* **2018**, *39*, 7483–7505. [CrossRef]
109. Vitale, S.; Ferraioli, G.; Pascazio, V. Multi-objective CNN based algorithm for SAR despeckling. *arXiv* **2020**, arXiv:2006.09050v4.
110. Baier, G.; He, W.; Yokoya, N. Robust nonlocal low-rank SAR time series despeckling considering speckle correlation by total variation regularization. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 7942–7954. [CrossRef]
111. Lehtinen, J.; Munkberg, J.; Hasselgren, J.; Laine, S.; Karras, T.; Aittala, M.; Aila, T. Noise2noise: Learning image restoration without clean data. *arXiv* **2018**, arXiv:1803.04189.
112. Ma, X.; Wang, C.; Yin, Z.; Wu, P. SAR image despeckling by noisy reference-based deep learning method. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8807–8818. [CrossRef]
113. Shen, H.; Lin, L.; Li, J.; Yuan, Q.; Zhao, L. A Residual Convolutional Neural Network for Polarimetric SAR Image Super-Resolution. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 90–108. [CrossRef]
114. Wang, C.; Sertel, E. Multi-frame super-resolution of remote sensing images using attention-based GAN models. *Knowl. Based Syst.* **2023**, *266*, 110387. [CrossRef]
115. Ji, Z.; Li, L.; Bi, H. Deep Learning-Based Approximated Observation Sparse SAR Imaging via Complex-Valued Convolutional Neural Network. *Remote Sens.* **2024**, *16*, 3850. [CrossRef]
116. Gao, J.; Ye, Y.; Li, S.; Qin, Y.; Gao, X.; Li, X. Fast super-resolution 3D SAR imaging using an unfolded deep network. In Proceedings of the IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, 11–13 December 2019; pp. 1–5. [CrossRef]
117. Wu, C.; Zhang, Z.; Chen, L.; Yu, W. Super-resolution for MIMO array SAR 3-D imaging based on compressive sensing and deep neural network. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2020**, *13*, 3109–3124. [CrossRef]
118. Venkatakrishnan, S.V.; Bouman, C.A.; Wohlberg, B. Plug-and-Play Priors for Model Based Reconstruction. In Proceedings of the 2013 IEEE Global Conference on Signal and Information Processing, Austin, TX, USA, 3–5 December 2013; pp. 945–948. [CrossRef]
119. Zhang, K.; Zuo, W.; Gu, S.; Zhang, L. Learning Deep CNN Denoiser Prior for Image Restoration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2808–2817. [CrossRef]
120. Bai, S.; Kolter, J.Z.; Koltun, V. Deep Equilibrium Models. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; Volume 32.
121. Schmitt, M.; Zhu, X.X. On the challenges in stereogrammetric fusion of SAR and optical imagery for urban areas. *Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B7*, 719–722. [CrossRef]
122. Wang, Y.; Zhu, X.X. The SARptical dataset for joint analysis of SAR and optical image in dense urban area. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 6840–6843. [CrossRef]
123. Wang, S.; Quan, D.; Liang, X.; Ning, M.; Guo, Y.; Jiao, L. A deep learning framework for remote sensing image registration. *ISPRS J. Photogram. Remote Sens.* **2017**, *145*, 148–164. [CrossRef]
124. Merkle, N.; Luo, W.; Auer, S.; Müller, R.; Urtasun, R. Exploiting deep matching and SAR data for the geolocalization accuracy improvement of optical satellite images. *Remote Sens.* **2017**, *9*, 586. [CrossRef]
125. Wang, Y.; Zhu, X.X.; Montazeri, S.; Kang, J.; Mou, L.; Schmitt, M. Potential of the 'SARptical' system. In Proceedings of the FRINGE, Helsinki, Finland, 4–28 August 2017.
126. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]
127. Suri, S.; Reinartz, P. Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas. *IEEE Trans. Geosci. Remote Sens.* **2009**, *48*, 939–949. [CrossRef]
128. Dellinger, F.; Delon, J.; Gousseau, Y.; Michel, J.; Tupin, F. SAR-SIFT: A SIFT-like algorithm for SAR images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *53*, 453–466. [CrossRef]

129. Abulkhanov, D.; Konovalenko, I.; Nikolaev, D.; Savchik, A.; Shvets, E.; Sidorchuk, D. Neural network-based feature point descriptors for registration of optical and SAR images. In Proceedings of the Tenth International Conference on Machine Vision (ICMV 2017), Vienna, Austria, 13–15 November 2017; Volume 10696, p. 106960. [CrossRef]
130. Merkle, N.; Auer, S.; Müller, R.; Reinartz, P. Exploring the potential of conditional adversarial networks for optical and SAR image matching. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2018**, *11*, 1811–1820. [CrossRef]
131. Hughes, L.H.; Merkle, N.; Bürgmann, T.; Auer, S.; Schmitt, M. Deep Learning for SAR-Optical Image Matching. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 4877–4880. [CrossRef]
132. Reyes, M.F.; Auer, S.; Merkle, N.; Henry, C.; Schmitt, M. SAR-to-optical image translation based on conditional generative adversarial networks-optimization, opportunities and limits. *Remote Sens.* **2019**, *11*, 2067. [CrossRef]
133. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2242–2251. [CrossRef]
134. Hughes, L.H.; Schmitt, M. A semi-supervised approach to SAR-optical image matching. *ISPRS Ann. Photogram. Remote Sens. Spat. Inform. Sci.* **2019**, *IV-2*, 71–78. [CrossRef]
135. Yao, W.; Marmanis, D.; Datcu, M. Semantic segmentation using deep neural networks for SAR and optical image pairs. In Proceedings of the Big Data from Space, Toulouse, France, 28–30 November 2017.
136. Audebert, N.; Saux, B.L.; Lefevre, S. Semantic segmentation of earth observation data using multimodal and multi-scale deep networks. In *Computer Vision—ACCV 2016. ACCV 2016. Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2017; Volume 10111, pp. 180–196.
137. Schmitt, M.; Hughes, L.; Körner, M.; Zhu, X.X. Colorizing Sentinel-1 SAR images using a variational autoencoder conditioned on Sentinel-2 imagery. *Int. Arch. Photogram. Remote Sens. Spat. Inform. Sci.* **2018**, *42*, 1045–1051. [CrossRef]
138. Zhao, J.; Zhang, Z.; Yao, W.; Datcu, M.; Xiong, H.; Yu, W. Open-SARUrban: A Sentinel-1 SAR image dataset for urban interpretation. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2019**, *13*, 187–203. [CrossRef]
139. Zhu, X.X.; Hu, J.; Qiu, C.; Shi, Y.; Kang, J.; Mou, L.; Bagheri, H.; Haberle, M.; Hua, Y.; Huang, R.; et al. So2Sat LCZ42: A Benchmark Data Set for the Classification of Global Local Climate Zones [Software and Data Sets]. *IEEE Geosci. Remote Sens. Mag.* **2020**, *8*, 76–89. [CrossRef]
140. Schmitt, M.; Hughes, L.H.; Qiu, C.; Zhu, X.X. SEN12MS-A curated dataset of georeferenced multi-spectral Sentinel-1/2 imagery for deep learning and data fusion. *ISPRS Ann. Photogram. Remote Sens. Spat. Inform. Sci.* **2019**, *IV-2*, 153–160. [CrossRef]
141. Schmitt, M.; Hughes, L.H.; Zhu, X.X. The SEN1-2 dataset for deep learning in SAR-Optical data fusion. In Proceedings of the ISPRS TC I Mid-Term Symposium "Innovative Sensing—From Sensors to Methods and Applications", Karlsruhe, Germany, 10–12 October 2018; pp. 141–146.
142. Shermeyer, J.; Hogan, D.; Brown, J.; Van Etten, A.; Weir, N.; Pacifici, F.; Hänsch, R.; Bastidas, A.; Soenen, S.; Bacastow, T.; et al. SpaceNet 6: Multi-Sensor All Weather Mapping Dataset. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Virtual Conference, 14–19 June 2020; pp. 768–777. [CrossRef]
143. Ichikawa, K.; Hirose, A. Singular unit restoration in InSAR using complex-valued neural networks in the spectral domain. *IEEE Trans. Geosci. Remote Sens.* **2016**, *55*, 1717–1723. [CrossRef]
144. Yamaki, R.; Hirose, A. Singular unit restoration in interferograms based on complex-valued markov random field model for phase unwrapping. *IEEE Geosci. Remote Sens. Lett.* **2008**, *6*, 18–22. [CrossRef]
145. Oyama, K.; Hirose, A. Adaptive phase-singular-unit restoration with entire-spectrum-processing complex-valued neural networks in interferometric SAR. *Electron. Lett.* **2017**, *54*, 43–44. [CrossRef]
146. Valade, S.; Ley, A.; Massimetti, F.; D'Hondt, O.; Laiolo, M.; Coppola, D.; Loibl, D.; Hellwich, O.; Walter, T.R. Towards Global Volcano Monitoring Using Multisensor Sentinel Missions and Artificial Intelligence: The MOUNTS Monitoring System. *Remote Sens.* **2019**, *11*, 1528. [CrossRef]
147. Costante, G.; Ciarfuglia, T.; Biondi, F. Towards monocular digital elevation model (DEM) estimation by convolutional neural networks-application on synthetic aperture radar images. *arXiv* **2018**, arXiv:1803.05387.
148. Schwegmann, C.; Kleynhans, W.; Engelbrecht, J.; Mdakane, L.; Meyer, R. Subsidence feature discrimination using deep convolutional neural networks in synthetic aperture radar imagery. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 4626–4629. [CrossRef]
149. Anantrasirichai, N.; Albino, F.; Hill, P.; Bull, D.; Biggs, J. Detecting volcano deformation in InSAR using deep learning. *arXiv* **2018**, arXiv:1803.00380.
150. Anantrasirichai, N.; Biggs, J.; Albino, F.; Bull, D. A deep learning approach to detecting volcano deformation from satellite imagery using synthetic datasets. *Remote Sens. Environ.* **2019**, *230*, 111179. [CrossRef]
151. Anantrasirichai, N.; Biggs, J.; Albino, F.; Bull, D. The application of convolutional neural networks to detect slow, sustained deformation in InSAR time series. *Geophys. Res. Lett.* **2019**, *46*, 11850–11858. [CrossRef]

152. Frate, F.D.; Picchiani, M.; Schiavon, G.; Stramondo, S. Neural networks and SAR interferometry for the characterization of seismic events. In *SAR Image Analysis, Modeling, and Techniques X*; SPIE: Baltimore, MD, USA, 2010; p. 78290. [CrossRef]
153. Picchiani, M.; Frate, F.D.; Schiavon, G.; Stramondo, S.; Chini, M.; Bignami, C. Neural networks for automatic seismic source analysis from DInSAR data. In *SAR Image Analysis, Modeling, and Techniques XI*; SPIE: Baltimore, MD, USA, 2011; p. 81790. [CrossRef]
154. Stramondo, S.; Frate, F.D.; Picchiani, M.; Schiavon, G. Seismic source quantitative parameters retrieval from In-SAR data and neural networks. *IEEE Trans. Geosci. Remote Sens.* **2010**, *49*, 96–104. [CrossRef]
155. Rongier, G.; Rude, C.; Herring, T.; Pankratius, V. Generative Modeling of InSAR Interferograms. *Earth Space Sci.* **2019**, *6*, 2671–2683. [CrossRef] [PubMed]
156. Zhang, J.; Xing, M.; Xie, Y. FEC: A Feature Fusion Framework for SAR Target Recognition Based on Electromagnetic Scattering Features and Deep CNN Features. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2174–2187. [CrossRef]
157. Zhang, J.; Xing, M.; Sun, G.-C.; Bao, Z. Integrating the Reconstructed Scattering Center Feature Maps with Deep CNN Feature Maps for Automatic SAR Target Recognition. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 4009605. [CrossRef]
158. Huang, Z.; Datcu, M.; Pan, Z.; Qiu, X.; Lei, B. HDEC-TFA: An Unsupervised Learning Approach for Discovering Physical Scattering Properties of Single-Polarized SAR Image. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 3054–3071. [CrossRef]
159. Ni, P.; Xu, G.; Pei, H.; Qiao, Y.; Yu, H.; Hong, W. Dual-Stream Manifold Multiscale Network for Target Recognition in Complex-Valued SAR Image with Electromagnetic Feature Fusion. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5217215. [CrossRef]
160. Ross, T.; Worrell, S.; Velten, V.; Mossing, J.; Bryant, M. Standard SAR ATR evaluation experiments using the MSTAR public release data set. In *Algorithms for Synthetic Aperture Radar Imagery V*; SPIE: Baltimore, MD, USA, 1998. [CrossRef]
161. Dumitru, C.O.; Schwarz, G.; Datcu, M. SAR image land cover datasets for classification benchmarking of temporal changes. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2018**, *11*, 1571–1592. [CrossRef]
162. Candes, E.J.; Tao, T. Decoding by Linear Programming. *IEEE Trans. Inf. Theory* **2005**, *51*, 4203–4215. [CrossRef]
163. Donoho, D.L. Compressed Sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.