ALIGNSAR: AN OPEN-SOURCE PACKAGE OF SAR BENCHMARK DATASET CREATION FOR MACHINE LEARNING APPLICATIONS

Ling Chang[1*], *Xu Zhang*[1], *Anurag Kulshrestha*[1,5], *Serkan Girgin*[1], *Alfred Stein*[1], *José Manuel Delgado Blasco*[2,6], *Angie Catalina Carrillo Chappe*[2], *Andrea Cavallini*[2], *Marco Uccelli*[2], *Milan Lazecky*[3], *Andy Hooper*[3], *Wojciech Witkowski*[4], *Magdalena Lucka*[4], *Artur Guzy*[4]

[1]Unversity of Twente, The Netherlands, [2]RHEA System S.p.A., Italy,
[3]University of Leeds, United Kingdom, [4]AGH University of Krakow,
Poland, [5]ICEYE company, Finland, [6]University of Jaén, Spain

ABSTRACT

As many dedicated SAR missions routinely and freely deliver SAR images, the community of SAR users is also expanding. Nevertheless, SAR images are less widely utilized than optical images for machine learning applications due to their complex nature and the limited availability of labeled/reference SAR datasets. Open-access SAR benchmark datasets, along with detailed specifications that can facilitate such applications, are therefore needed. In this regard, AlignSAR project aims 1) to design and demonstrate a generic procedure for the creation of SAR benchmark datasets; 2) to develop methods to align all available geospatial observations from e.g. SAR and LiDAR into a common reference; 3) to define a specification of the SAR signatures and their associated descriptors so that they can be easily indexed and programmatically searched and retrieved; and 4) to provide a relevant open-source software package with associated documentation. In this paper we tested our package using seven Sentinel-1a SAR acquisitions in VV and VH, covering the city of Groningen, the Netherlands, and extracted fourteen SAR signatures for further ANN-based land use and land cover classification. The average $F1$-score for the four classes of buildings, roads, railways and water is 0.93, and kappa coefficient κ is 0.86. We conclude that our AlignSAR package facilitates machine learning applications, and lowers the barriers to entry for users with limited knowledge of SAR.

Index Terms— AlignSAR, SAR, SAR benchmark datasets, Machine learning

1. INTRODUCTION

SAR (Synthetic Aperture Radar) techniques, including InSAR (Interferometric SAR) [1] and PolSAR (Polarimetric SAR) [2], are well-established and are employed in natural and anthropogenic hazard monitoring, as well as land use and land cover (LULC) classification. The community of SAR users, especially for machine learning applications, is continuously expanding, thanks to SAR's inherent merits like global observation coverage, cloud insensitivity, regular updates and day-night operability, and free-of-charge accessibility. Yet, SAR benchmark datasets (SARbd) that can be treated as references, are still far from sufficient for machine learning applications. This is mainly attributed to 1) the complex nature of SAR data that follow a circular Gaussian distribution for SAR complex numbers confined between $-\pi$ and $+\pi$; 2) single or less than four polarimetric SAR data that capture limited Earth information; 3) speckle noise that downgrades the quality of SAR data and is impossible to be entirely eliminated; 4) lack of corresponding ground-truth data and a way to link it with SAR data at SAR pixel level that makes the creation of high-quality SARbd impossible; 5) no open-source tools that are tailored and ready for these applications. This research addresses this issue, and designs and demonstrates a generic procedure for the creation of SARbd, and provides a processing package primarily based upon Python language. This package has been released on GitHub AlignSAR, along with three use cases over the Netherlands, Poland and India for LULC classification and oil-spill object detection, using an Artificial Neural Network (ANN) with Dense layers and Yolov8 models, respectively. In this paper, we present the case over the Netherlands, using seven Sentinel-1a SAR acquisitions in VV and VH, to demonstrate the feasibility and operability of this AlignSAR package.

2. METHODOLOGIES

We designed a generic processing workflow to create SAR benchmark datasets, and released all related tools on GitHub https://github.com/AlignSAR/alignSAR. In general, this work-

[*]Thanks to ESA for funding Open SAR Library 4000140600/23/I-DT and to all AlignSAR members. Email: ling.chang@utwente.nl

flow, as shown in Fig. 1, is composed of three main blocks: Input, Processing, and Output.

Input, SAR data and other (reference) data are collected. SAR data includes (partially) freely accessible medium resolution SAR images, such as Sentinel-1a&b SAR data, and others have other geospatial observations from e.g. GPS observations, AHN (actual height of Netherlands), Topographic map, and high-resolution TanDEM-X SAR and Copernicus DEM.

Processing, all required Python-based open-source tools are created. To facilitate analysis and ensure reproducibility, we created a Docker that contains the pre-installed necessary software tools, i.e. SNAP 9.0, modified Doris-5 and Python 2/3 environment as well as all the other developed scripts. By doing so, the end users can directly follow the procedure to create SAR benchmark datasets without manually installing the required software tools individually. To realize the alignment of all observations from SAR and other data, radarcoding based on Doris-5 or GMTSAR is implemented. Fig. 2 illustrates the radarcoding procedure (based on Doris-5) for other (reference) data, adapted from [3]. The basic idea is to first coregister and resample all SAR images to a common master grid using differential SAR interferometry (DInSAR) with Doris-5, next rasterize other data which may contain point-line-polygons or classification product using GDAL, and then crop this rasterized data based on SAR coverage boundary information and later assign corresponding radar coordinates to it as well as the geo-coordinates in the end. For details refer to [3]. A number of representative SAR signatures, for instance, amplitude in different polarimetric channels, interferometric phase, coherence, intensity difference/summation/ratio between different channels, coherence, and entropy, are selected, cf. [4]. Denoising is to reduce noise impact on SAR images, using e.g. boxcar filtering, spatio-temporal filtering [5], MONet [6]. Some other data can be employed as reference data to assess the quality of the extracted SAR signatures of every pixel.

Output, based on the quality level, we extract the pixels with high quality, namely SAR benchmark datasets (SARbd). We categorize SARbd into three levels in terms of the quality level of SARbd. Level 1, SARbd-L1, is merely based on SAR statistics; Level 2, SARbd-L2, is based on external geospatial reference data; Level 3, SARbd-L3, is based on both SAR statistics and external geospatial reference data. SARbd-L1 is selected based on some of SAR signatures thresholding and visual interpretation. Such a selection is analogous to data labeling in optical imagery. For instance, one can use amplitude/coherence (SAR signatures) / amplitude dispersion index thresholding to categorize the SAR pixels and label them as e.g. buildings, roads, crops, and water. The threshold values are defined mainly based on visual interpretation or apriori knowledge. SARbd-L2 are defined based on additional reference information. For instance, having radarcoded or geocoded LULC products, some SAR pixels would inherit LULC signatures from this additional reference information, and are treated as SARbd-L2. Acknowledging additional reference information may contain errors, SAR signatures thresholding can be used to exclude unreliably labeled SARbd-L2, especially the SARbd pixels on the boundary (buffer zone) of two LULC classes, in order to obtain the highest-quality SARbd, i.e. SARbd-L3. The prototype library of the SARbd is delivered, along with the open-source tools and use case demonstration. Ultimately, SARbd library is integrated with EO-TDL (Earth Observation Training Data Lab) platform for e.g. further training data creation, training data augmentation, and machine learning applications. Note that the SARbd output is stored in netCDF (.nc) and COG (Cloud-optimized GeoTIFFs) format with STAC (SpatioTemporal Asset Catalogs, containing global and local attributes, cf. stac-extensions), which equates to EO-TDL Q1 data.

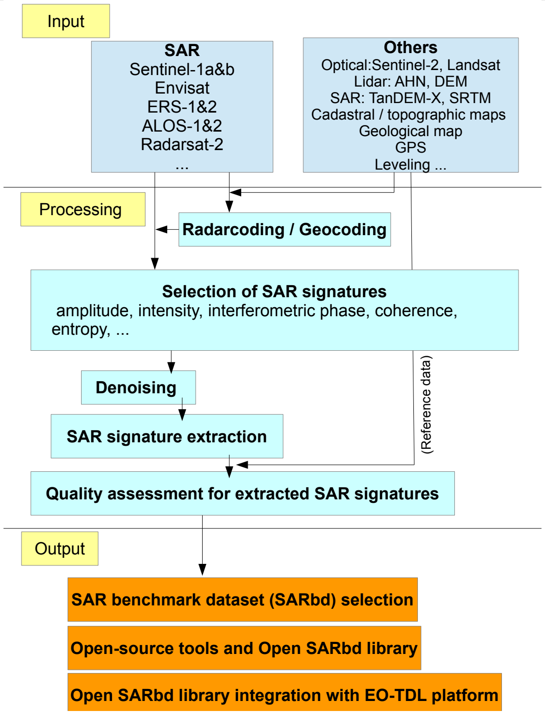
Input
SAR
Sentinel-1a&b
Envisat
ERS-1&2
ALOS-1&2
Radarsat-2
...
Others
Optical:Sentinel-2, Landsat
Lidar: AHN, DEM
SAR: TanDEM-X, SRTM
Cadastral / topographic maps
Geological map
GPS
Leveling ...
Processing
Radarcoding / Geocoding
Selection of SAR signatures
amplitude, intensity, interferometric phase, coherence, entropy, ...
Denoising
SAR signature extraction
(Reference data)
Quality assessment for extracted SAR signatures
Output
SAR benchmark dataset (SARbd) selection
Open-source tools and Open SARbd library
Open SARbd library integration with EO-TDL platform

Fig. 1. Processing flowchart for the SARbd creation

3. TEST SITE AND DATA DESCRIPTION

The test site is situated in Groningen, a city in the northern part of the Netherlands, highlighted in green in Figs. 3(a) and 3(b). This site is relevant as serious land deformation has been reported that has resulted in damage to houses and other buildings. Seven C-band Sentinel-1a acquisitions (Path 15, Frame 169) in ascending orbit were collected, outlined in red in Fig. 3(a). They have VV and VH polarization channels, a spatial resolution of 20 × 5 m, and were acquired separately on 09 Jan, 21 Jan, 02 Feb, 14 Feb (master), 26

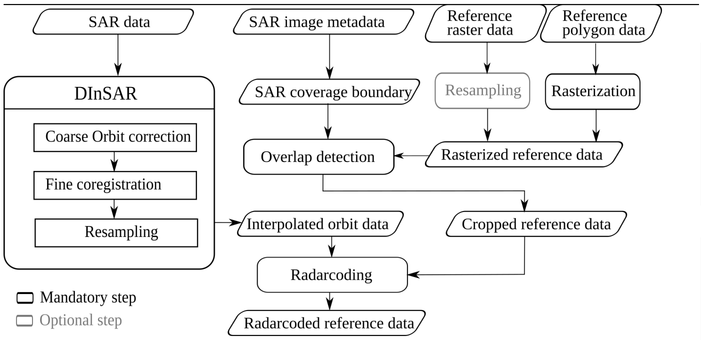

Fig. 2. Radarcoding procedure, adapted from [3]

Feb, 10 Mar and 22 Mar 2022. The topographic base map – TOP10NL was used as an additional geospatial observation dataset with about 1 m resolution, including LULC features (buildings, roads, railways, and water). Fig. 3(b) illustrates the TOP10NL product over Area1 ($\sim 10 \times 13$ km^2) in green.

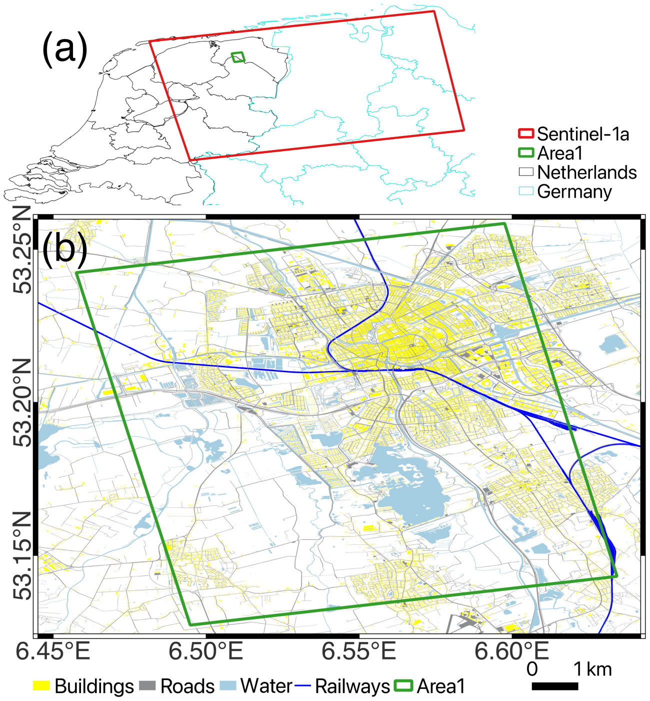

Fig. 3. (a) Sentinel-1a image coverage. (b) TOP10NL-based LULC map over Area1.

4. RESULTS

According to Section 2, we employed our modified Doris-5 to generate coregistered and resampled SAR images all in the master (14 Feb 2022) grid. As an example, the first ten radar signatures for the SAR image acquired on 09 Jan 2022 shown in Fig. 4 were directly created from the coregistered and resampled SAR images themselves. Note that coregistering a stack of SAR images will pave the way for SARbd time series signatures creation and analysis in the near future. TOP10NL data were also aligned to the master grid using our radarcod-

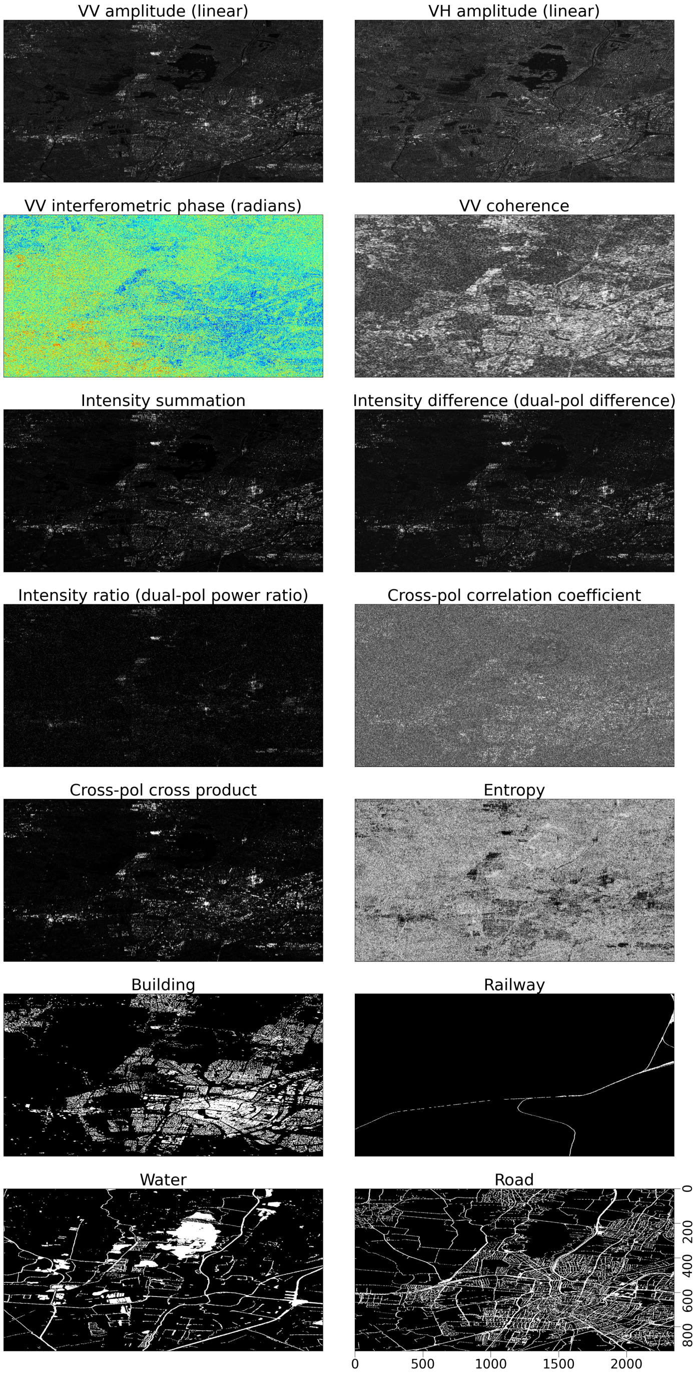

Fig. 4. An example of fourteen radar signatures for SAR acquisition on 09 Jan 2022 in the master's radar coordinates. Note that all images are upside down w.r.t. WGS84 coordinates (in Fig. 3) due to the satellite's ascending orbit.

ing tool, see the last four signatures (semi binary maps), buildings, roads, railways, water in white in Fig. 4. In total, fourteen signatures were created, but our tool is extendable and allows one to create more relevant signatures. Here we showcase the use of SAR-L2 for machine learning based LULC classification on Area1 with a size of 940 lines and 2350 pixels. The last four signatures inherited from TOP10NL were treated as the reference, and 28770, 51233, 1141, and 37989 pixels were separately labeled as buildings, roads, railways, and water class. The rest pixels are uncharted. To avoid potential biases in the models when training, we applied filtering and spatial resampling to have each class at the same proportion (i.e. 1000). We then utilized an artificial neural network (ANN) with six dense layers and two dropout layers, and trained the model on 50000 epochs with a batch size of 100 using an Adam optimized with a default learning rate of 0.01 and a categorical cross entropy loss. Fig. 5(a) depicts the target label map for SAR image acquired on 09 Jan 2022, which was generated by using the last four signatures (i.e. buildings, roads, railways, and water in yellow, gray, blue and cyan, respectively), while Fig. 5(b) shows the predicted label map provided by the ANN. Table 1 lists the quality metric values of each class and their average. We found that the average $F1$-score is 0.93 and kappa coefficient κ is 0.86.

Table 1. Quality metrics of the LULC result

	Buildings	Roads	Railways	Water	Average
Recall	0.95	0.89	0.94	0.97	0.94
Precision	0.93	0.93	0.91	0.96	0.93
$F1$-score	0.94	0.91	0.92	0.97	0.93

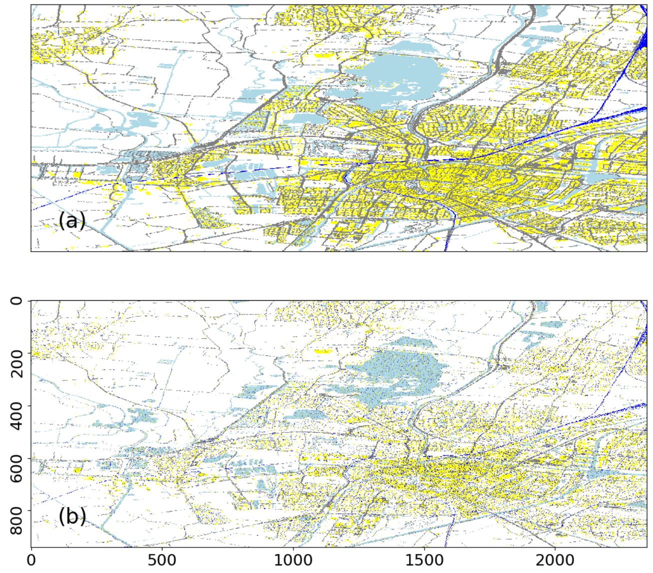

Fig. 5. (a) True label map derived from the last four signatures, for SAR image acquired on 09 Jan 2022. (b) Predicted label map. Buildings, roads, railways, and water are denoted in yellow, gray, blue and cyan, respectively.

5. CONCLUSIONS

We have developed AlignSAR package as an open-source tool that standardizes SAR benchmark dataset creations by combining relevant existing software and tools such as SNAP, Doris-5, GMTSAR and GDAL. The tools within the package can run and be customized independently as long as the required input (format) is prepared, and they also can be encapsulated and run in Docker. The demonstration of using the Groningen case for ANN-based LULC classification confirms the feasibility and operability of the AlignSAR package. In addition, this case study shows the fourteen SAR signatures are representative and relevant for such an analysis, and the quality of LULC classification using SAR-L2 is convincing with the average $F1$-score of 0.93, and κ of 0.86. For future work, we plan to further develop a reference, quality-controlled, documented, open database of InSAR time series spatial and temporal signatures of complex real-world targets, with a focus on time series SARbd creation and demonstration, to serve a vast number of machine learning applications and enable users to easily employ time series SAR data for machine learning analysis.

6. REFERENCES

[1] R. Bamler and P. Hartl, “Synthetic aperture radar interferometry,” *Inverse problems*, vol. 14, no. 4, pp. R1, 1998.

[2] J.S. Lee and E. Pottier, *Polarimetric radar imaging: from basics to applications*, CRC press, 2017.

[3] A. Kulshrestha, L. Chang, and A. Stein, “Radarcoding reference data for SAR training data creation in radar coordinates,” *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[4] L. Chang, A. Kulshrestha, B. Zhang, and X. Zhang, “Extraction and analysis of radar scatterer attributes for PAZ SAR by combining time series InSAR, PolSAR, and land use measurements,” *Remote Sensing*, vol. 15, no. 6, pp. 1571, 2023.

[5] S. Quegan and J. Yu, “Filtering of multichannel SAR images,” *IEEE Transactions on geoscience and remote sensing*, vol. 39, no. 11, pp. 2373–2379, 2001.

[6] S. Vitale, G. Ferraioli, V. Pascazio, and G. Schirinzi, “InSAR-MONet: Interferometric SAR phase denoising using a multiobjective neural network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.