

AI-BAQ: Deep Learning for Adaptive SAR Raw Data Quantization

Nicola Gollin¹, Michele Martone¹, Ernesto Imbembo¹, Max Ghiglione¹, Stefan Knoll, Gerhard Krieger¹, *Fellow, IEEE*, and Paola Rizzoli¹, *Member, IEEE*

Abstract—Next-generation (NG) synthetic aperture radar (SAR) systems will be capable of performing high-resolution, wide-swath acquisitions at frequent revisit times. The overcoming of conventional SAR limitations will also lead to the generation of very large volumes of onboard data, which need to be stored and managed by the system and downlinked to the ground. This poses severe constraints in terms of onboard memory requirements and downlink capacity, and in this challenging scenario, the onboard quantization of SAR raw data represents a crucial aspect, acting as a tradeoff between the achievable product quality and the resulting onboard volume of data. State-of-the-art (SoA) quantization schemes allow for enhanced data rate allocation; however, the optimization is directly performed on raw data, without targeting a desired performance on the final higher level SAR/interferometric SAR (InSAR) product. In this article, we investigate the use of artificial intelligence (AI), and in particular of deep learning (DL), for developing a flexible onboard SAR raw data quantization method, with the aim of deriving an optimized and fully adaptive data rate allocation given a set of desired performance metrics and requirements in the resulting focused SAR and InSAR products, without relying on a priori information on the acquired scene. Different performance parameters are considered, such as the signal-to-quantization-noise ratio (SQNR), the phase errors, the InSAR coherence loss, as well as the resulting noise equivalent sigma zero (NESZ), extending the capabilities of the architecture to provide multiple bitrate estimations for a single input scene at the same time, depending on the desired application case. We use experimental TanDEM-X bistatic SAR data, both for the training of the DL model as well as for the validation and demonstration of the suitability of the proposed method. In view of a potential onboard implementation, a possible hardware architecture for the proposed compression scheme is investigated as well.

Index Terms—Convolutional neural network (CNN), deep learning (DL), performance-optimized quantization, synthetic aperture radar (SAR) raw data compression.

Received 27 November 2024; revised 30 July 2025; accepted 3 September 2025. Date of publication 8 September 2025; date of current version 6 October 2025. This work was supported by European Space Research and Technology Centre (ESTEC), European Space Agency (ESA), through the Project “ARTISTE: Adaptive SAR Signal Compression Through Artificial Intelligence” under Contract ESA AO/1-11419/22/NL/GLC/my. (*Corresponding author: Nicola Gollin.*)

Nicola Gollin, Michele Martone, Gerhard Krieger, and Paola Rizzoli are with the Microwaves and Radar Institute, German Aerospace Center (DLR), 82234 Weßling, Germany (e-mail: nicola.gollin@dlr.de; michele.martone@dlr.de; gerhard.krieger@dlr.de; paola.rizzoli@dlr.de).

Ernesto Imbembo and Max Ghiglione are with the Radio Frequency Payloads and Technology Division, European Space Agency (ESA), 2200 AG Noordwijk, The Netherlands (e-mail: ernesto.imbembo@esa.int; max.ghiglione@esa.int).

Stefan Knoll is with Microwave Instruments, Airbus Defence and Space GmbH, 88039 Friedrichshafen, Germany (e-mail: stefan.s.knoll@airbus.com).

Digital Object Identifier 10.1109/TGRS.2025.3607089

I. INTRODUCTION

NEXT-GENERATION (NG) synthetic aperture radar (SAR) systems will allow for a huge leap in performance, by exploiting large bandwidths and digital beamforming (DBF) techniques, in combination with multichannel [1], [2] and/or multistatic configurations [3]. This will allow for the realization of high-resolution, wide-swath interferometric SAR (InSAR) and polarimetric SAR (PolSAR) acquisitions [4]. Moreover, present spaceborne SAR missions, such as Sentinel-1 [5], or future ones, such as NISAR [6], [7], and especially ROSE-L [8] and Sentinel-1 NG [9], [10], will acquire data either globally or over selected areas with a temporal sampling in the order of one week or less. As a consequence, an unprecedented volume of data will be generated on such systems, which will have to be stored and then downlinked to the ground. This aspect represents a crucial, high-demanding requirement for the design of the SAR system itself [1].

In this scenario, the definition of efficient onboard SAR raw data quantization methods is of paramount importance. On the one hand, the specific data compression strategy defines the resulting volume of onboard data, which is linked to the required onboard memory and downlink capacity, and on the other hand, it directly affects the quality of the generated SAR products. These two aspects must be carefully considered due to the limited acquisition capacity and onboard resources of the system and, at the same time, to allow for the achievement of the specified higher level SAR products requirements and quality. Moreover, the nonlinear contribution of the quantization error in the raw data representation has strong implications in the final SAR and InSAR image (and to higher level products such as digital elevation models (DEMs) or biophysical parameters), due to its dependency on both system parameters and local backscatter characteristics [11]. A widely used method for SAR raw data quantization is the block-adaptive quantization (BAQ) [12], originally proposed for the Magellan SAR mission to Venus. It represents a good tradeoff between signal representation quality and resulting data volume, as it adapts the quantization levels to the signal statistics of the input raw data samples within a certain block, aiming at minimizing the quantization error. In the last decade, novel algorithms have been proposed starting from the principle of BAQ, allowing for an improved and more targeted performance optimization and resource allocation. Typically, these are acquisition-dependent compression

schemes, as, e.g., for the case of the flexible dynamic BAQ (FDBAQ), where the bitrate is selected depending on the local signal-to-thermal-noise ratio (STNR) [13], [14]. For its computation, the STNR requires the knowledge of the backscatter within the illuminated scene, which is estimated on board from the raw data power. The quantization settings are then selected based on a series of lookup tables (LUTs), which are stored on board. In turn, these are derived on the ground by considering large-scale backscatter statistics obtained from available global backscatter maps. However, the FDBAQ carries out the bitrate allocation without considering the actual performance degradation in the resulting high-level SAR products and applications. Indeed, the local variability and inhomogeneities in the backscatter distribution severely impact the resulting quantization degradation [11], requiring a direct link between the quantization settings and the focused SAR domain to be properly handled.

An attempt to close this gap is represented by the performance-optimized BAQ (PO-BAQ) [15], which is based on the estimation of a 2-D spatial-variant bitrate allocation map in the raw data domain depending on the final performance requirement defined on the higher level SAR and InSAR products, such as the signal-to-quantization-noise ratio (SQNR) or the interferometric phase errors. In order to estimate the local distribution of the SAR intensity and, in particular, its degree of homogeneity, the PO-BAQ exploits a priori knowledge on the SAR backscatter statistics of the imaged scene. This information allows for deriving 2-D bitrate maps (BRMs), which must be available on board (stored or uplinked) before commanding. For these reasons, the PO-BAQ is not fully adaptive to the acquired scene since the quantization settings are derived from prior considerations and do not directly account for the local conditions at the time of the SAR survey. In addition, different quantization techniques have been proposed in the literature, investigating their potential for specific SAR configurations. This is the case, e.g., of the FFT-BAQ [16], the principal components block quantization [17], the multichannel block-adaptive quantization (MC-BAQ) [18], and predictive coding for range compressed [19], [20] and staggered SAR raw data [21], [22], as well as transform coding [23] for the frequency scan (FScan) acquisition mode [24].

In this challenging scenario, the advent of artificial intelligence (AI), and in particular of deep learning (DL), represents one of the most promising approaches for an effective exploration of the data information content [25], [26], [27]. In the field of remote sensing, convolutional neural networks (CNNs), and especially convolutional autoencoders (CAEs), have already demonstrated a high potential for the compression and restoration of optical images, thanks to their ability in handling 2-D spatial patterns [28], [29], [30]. Regarding SAR data, several DL models have been proposed for the compression of fully focused images [31], [32], [33], which, for a real utilization in a spaceborne mission scenario, would require the computationally expensive implementation of a complete onboard focusing chain. In the context of SAR raw data digitization and compression, the topic has remained highly unexplored, mainly due to the lack of spatial correlation

and self-similarity among samples typically observed in the raw data domain, which complicates the task of pattern recognition. A first attempt to utilize machine learning (ML) was proposed by Hay et al. [34], where they utilized shallow ML methods to dynamically select different types of quantizers, i.e., BAQ or FFT-BAQ, depending on the characteristics of the recorded SAR raw data. Differently, Pilikos et al. [35] proposed the use of a CAE for the compression of the real and imaginary parts of the SAR raw data matrix and quantized the output of the encoder using a learned vector quantization. A high compression rate was achieved at the cost of severe distortion of the reconstructed signal. This work was further extended in [36], with the development of a vector-quantized variational autoencoder (VQVAE). Finally, Asiyabi et al. [37] proposed a complex-valued CAE with side information to simultaneously compress both real and imaginary parts of the raw data matrix. To the best of our knowledge, these approaches currently remain at early stages of development and their robustness and operational applicability still need to be thoroughly assessed.

In this work, we propose a novel DL method for performing a dynamic and adaptive onboard bitrate allocation to feed a space-varying BAQ. The core idea comes from the intuition that a direct link between the raw data and the focused domains can be implemented through a DL model, without the need for a complete SAR focusing (a preliminary proof of concept of this idea was presented in [38] and [39]). This allows for achieving a certain desired performance in the final focused SAR product thanks to a dynamic allocation of quantization bits, which only depends on the raw data characteristics and on the desired quality of the output SAR/InSAR products. To do so, we tackle the problem as a fully supervised DL regression task, where the model is optimized with respect to various target performance criteria in the focused SAR domain. As exemplary optimization criterion, we consider targeting a certain uniform performance in the resulting focused SAR product, similar to what is done by the PO-BAQ. This allows for achieving an application-driven raw data quantization method, which features a high degree of flexibility in terms of target bit allocation criteria as well as adaptation to the local conditions (e.g., backscatter) of the acquired scene, without the need for additional information and data required on board. This is particularly crucial in real-world applications, where different scenarios need multiple tradeoffs between data quality and compression efficiency. For instance, urban areas and high-resolution DEMs demand higher bitrate requirements. Conversely, repetitive acquisitions of vast land cover areas for monitoring deforestation or mapping require a more relaxed quality setting to extract the essential information while reducing data transmission rates and enhancing the acquisition capacity of the system [11], [15]. As the DL model is used for parameter estimation (i.e., bitrate) and the actual quantization is performed by means of a BAQ, the proposed method enhances the performance of traditional fixed-rate BAQ quantizers without altering their operation principle.

This article is structured as follows. A brief overview of SAR raw data quantization background concepts is recalled

in Section II. The description of the proposed method, named AI-BAQ, including the model architecture, the used dataset, and the training strategy, is presented in Section III. Section IV details the obtained results, including the performance analysis on the final SAR products, while Section V illustrates a hardware feasibility assessment for a potential onboard implementation. A discussion summarizing the main outcomes and findings of the present research is provided in Section VI, and finally, conclusions and outlook are drawn in Section VII.

II. BACKGROUND CONCEPTS

In this section, the basics of SAR quantization are recalled (Section II-A) together with the considered SAR performance parameters (Section II-B).

A. Background Concepts on SAR Raw Data Quantization

Quantization represents a fundamental step in SAR signal processing and data reduction, involving the representation of continuous or analog signals with a limited set of discrete values. In the process, a quantization error is introduced and denoted as

$$q = s - s_q \quad (1)$$

with q representing the difference between the original signal s and its quantized version s_q . This error arises due to the finite precision used for the representation of continuous ranges with a discrete set of values, leading to inevitable inaccuracies in the quantized signal. The proper handling and minimization of the quantization error is crucial for maintaining signal fidelity and achieving efficient data reduction. If the input s has a random nature, q can be modeled as an additive and signal-correlated random noise source [40]. Quantization errors are typically considered as the joint contribution of granular and overhead noise, where the number of decision levels (i.e., the number of bits used to represent the samples) and the signal clipping values directly impact the former and the latter error term, respectively. In particular, when considering SAR scenes featuring a high degree of inhomogeneity in the backscatter spatial distribution, an additional signal-dependent performance degradation is introduced (e.g., in urban areas, which are typically characterized by a high dynamic range in backscatter). This effect is also referred to as *low-scatter suppression* [11], [41], which, differently from the granular and overhead errors, can be detected only after focusing on the SAR images.

In an SAR acquisition, the responses of the illuminated scatterers overlap in the raw data domain within an area A_{SAR} of dimensions

$$A_{\text{SAR}} = L_{\text{ch}} \times L_{\text{s}} \quad (2)$$

where L_{ch} is the ground range-projected chirp length and L_{s} indicates the azimuth synthetic aperture [15]. The two quantities are defined as

$$L_{\text{ch}} = \frac{c\tau_{\text{ch}}}{2 \cdot \sin(\theta_{\text{inc}})} \quad (3)$$

$$L_{\text{s}} = \lambda \frac{r}{L_{\text{a}}} \quad (4)$$

where c is the speed of light, τ_{ch} is the chirp pulse duration, r is the slant range distance, θ_{inc} is the incidence angle, and L_{a} is the azimuth antenna length. Moreover, the responses from the scatterers are weighted by the azimuth and elevation antenna patterns (APs) depending on their location with respect to the transmitting and receiving beams.

SAR raw data quantization is typically carried out using a block-based scheme, such as the BAQ, where the decision levels and the clipping thresholds for the compression are derived from the mean power of the defined raw data block. For this reason, if two overlapping targets have different magnitude responses and are separated in range by d_{r} and in azimuth by d_{a} with

$$d_{\text{r}} \ll L_{\text{ch}}, \quad d_{\text{a}} \ll L_{\text{s}} \quad (5)$$

the quantizer will adapt its decision levels based on the power level associated with the stronger target, which will be better reconstructed, while the lower signal will appear strongly distorted [11]. According to this, in addition to the compression rate N_{b} used for data digitization, the degree of inhomogeneity in the SAR backscatter distribution crucially impacts the quantization performance on the final SAR and InSAR product. According to these considerations, quantization errors must be treated as a nonlinear and signal-dependent noise source affecting the SAR data and, for this reason, the encoding performance of a certain quantization method in SAR imaging has to be evaluated within windows of size A_{SAR} as in (2).

B. Quantization Performance Parameters

In order to assess the impact of quantization errors and to target specific performance assessments, we consider the following quality measures, evaluated on the final SAR and InSAR products.

1) *Signal-to-Quantization-Noise Ratio*: The SQNR is a standard quantization performance metric, which can be defined for a given SAR image pixel as

$$\text{SQNR} = \frac{|x|^2}{|x - x_q|^2} \quad (6)$$

where x is the nonquantized SAR image after digitization through the analog-to-digital converter (ADC) and x_q is the quantized one. The SQNR can be related to the total signal-to-noise ratio (SNR) as follows:

$$\begin{aligned} \text{SNR} &= \frac{\sigma^0}{\sigma_{\text{N}}^2} = \frac{\sigma^0}{\sigma_{\text{TN}}^2 + \sigma_{\text{QN}}^2} \rightarrow \text{SNR}^{-1} \\ &= \text{STNR}^{-1} + \text{SQNR}^{-1} \end{aligned} \quad (7)$$

where σ_{TN}^2 and σ_{QN}^2 represent the thermal and the quantization noise power, respectively, and STNR is the signal-to-thermal-noise ratio. In order to exclude potential outliers in the corresponding A_{SAR} region in the SAR image, we consider the median value in each A_{SAR} window for the present performance assessment. Hence, the resulting bitrate is not significantly affected by the presence of sparse very high or very low-performance pixels (clearly assuming that they do not dominate in the area of interest).

2) *Quantization Decorrelation* γ_{Quant} : In the presence of InSAR acquisitions, the impact of SAR raw data quantization on the total coherence γ is assessed by considering the factorization proposed in [42] and [43]. Here, γ_{Tot} can be decomposed into the following components:

$$\gamma_{\text{Tot}} = \gamma_{\text{Sys}}\gamma_{\text{Vol}}\gamma_{\text{Temp}}\gamma_{\text{Quant}} \quad (8)$$

where the different terms on the right-hand side of the equation represent the decorrelation caused by: system parameters (γ_{Sys}) which includes SNR, ambiguities, baseline decorrelation, and spectral shift; volumetric scattering (γ_{Vol}); temporal changes (γ_{Temp}); and quantization (γ_{Quant}). Specifically, γ_{Quant} can be computed as

$$\gamma_{\text{Quant}} = \frac{\gamma_{x_q}}{\gamma_x} \quad (9)$$

where γ_{x_q} and γ_x represent the coherence of the quantized and nonquantized images, respectively. Typically, the mean value of γ_{Quant} computed within the area A_{SAR} is associated with its central pixel.

3) *Phase Error* φ_{err} : The phase error represents the absolute difference between the phases of the quantized (φ_{x_q}) and unquantized (φ_x) InSAR images

$$\varphi_{\text{err}} = |\varphi_{x_q} - \varphi_x| \bmod 2\pi \quad (10)$$

which is wrapped between 0 and 2π . For the performance assessment, the median of the distribution of the absolute value of φ_{err} is considered within the A_{SAR} integration window for a better filtering of outliers, similarly as for the SQNR discussed above.

4) *Total Noise Equivalent Sigma Zero* NESZ_{Tot} : The NESZ_{Tot} quantifies the minimum backscatter signal strength, which can be detected by an SAR system, and it can be decomposed into the following two terms:

$$\text{NESZ}_{\text{Tot}} = \text{NESZ}_{\text{th}} + \text{NESZ}_{\text{q}} \quad (11)$$

where NESZ_{th} and NESZ_{q} are the thermal and quantization noise equivalent sigma zero (NESZ) contribution, respectively. The first term represents the system thermal noise floor and varies along the range depending on the system parameters and the second one considers a further degradation introduced by quantization noise. NESZ_{th} on a given swath position can be theoretically computed by knowing specific system parameters [43], [44], while the additional impact given by the quantization can be expressed in terms of quantization noise power as

$$\text{NESZ}_{\text{q}} = |x - x_q|^2. \quad (12)$$

Also, in this case, the median value is considered within the A_{SAR} integration window.

III. METHOD

We address the problem of optimizing the onboard bitrate allocation as a deep supervised regression task. For a given input block of raw data, we estimate the target bitrate, which ranges within a continuous interval of possible values. The choice of a supervised learning approach lies in the possibility of generating reliable labeled data, as detailed in the following. In Fig. 1, the complete high-level framework of the proposed

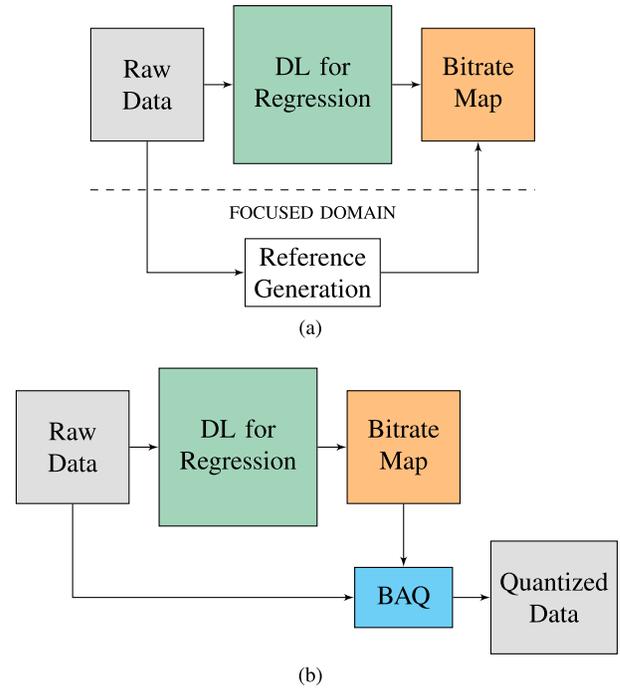


Fig. 1. Flowchart of (a) training and (b) inference for the proposed AI-BAQ method. The top plot highlights that the architecture is trained in order to link uncompressed (i.e., not quantized) raw data and reference BRMs generated in the focused domain. In the estimation task depicted in the bottom figure, the trained architecture is able to provide the required BRM, which fulfills a specific requirement (in the SAR image domain) directly from the raw data without specific a priori knowledge or information. An SoA adaptive quantizer, such as, in our case, a standard BAQ, performs the actual raw data encoding exploiting the estimated BRM.

approach is depicted (a detailed description of the DL model architecture is provided later on in Section III-B). Fig. 1(a) characterizes the training phase of the DL model for the regression task, which aims at estimating a 2-D BRM starting from the SAR raw data. In particular, the reference BRMs used during this supervised learning phase are derived from target performance parameters computed in the fully focused image domain, requiring a prior alignment between the considered portion of SAR raw data and the corresponding focused SAR image. In this way, a direct link between the uncompressed raw data and a certain expected performance of the derived higher level SAR products can be established without the need for nearly real-time onboard SAR focusing. Moreover, it is important to remark that the DL model training can be performed on ground, prior to its operational deployment on board.

Fig. 1(b) shows the structure of the complete onboard quantization framework, called AI-BAQ, where the trained DL model is used to dynamically estimate a BRM, given a certain target performance parameter, which is finally used as input to a BAQ for performing the actual quantization of the input raw data.

In Section III-A the dataset generation used to train and test the proposed AI-BAQ is described, followed by details on the DL architecture in Section III-B and training strategy in Section III-C; then, the performance evaluation framework is outlined in Section III-D.

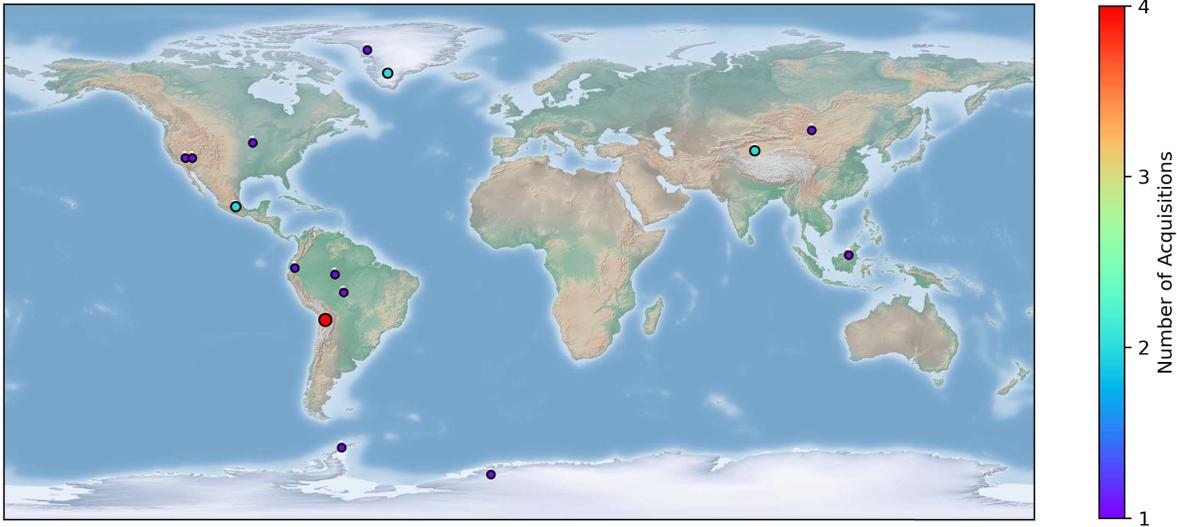


Fig. 2. Considered dataset for training and testing the proposed DL model. Each marker indicates the approximate location of the experimental TanDEM-X data taken commanded with BAQ bypass (i.e., 8-bit ADC).

TABLE I

TEST SITES' DESCRIPTION, DATE, AND THEIR APPROXIMATE LOCATION FOR TRAINING (UPPER GROUP) AND TEST (LOWER GROUP)

Test Site	Date	Land Cover	Location
Amazon (Brazil) 1	21.06.2012	Rainforest	4°S, 64°W
Amazon (Brazil) 2	17.01.2012	Rainforest/Flat Terr.	10°S, 61°W
Andes (Ecuador)	24.11.2022	Rugged Terrain	2°S, 78°W
Antarctica 1	05.06.2012	Snow/Ice	74°S, 9°W
Antarctica 2	23.10.2012	Water, Snow/Ice	64°S, 62°W
Borneo (Malaysia)	05.06.2012	Tropical Forest	2°N, 114°E
Death Valley (USA)	09.06.2012	Soil/Rocks	36°N, 116°W
Gobi (Mongolia)	05.04.2012	Sandy Desert	46°N, 102°E
Greenland (Greenland) 1	29.03.2012	Snow/Ice	65°N, 45°W
Greenland (Greenland) 1	30.03.2012	Snow/Ice	65°N, 45°W
Iowa (USA)	10.11.2010	Urban/Agric. Area	41°N, 93°W
Mexico City (Mexico)	07.02.2012	Urban/Topography	19°N, 99°W
Taklamakan (China)	30.01.2012	Sandy Desert	38°N, 82°E
Taklamakan (China)	08.11.2016	Sandy Desert	38°N, 82°E
Uyuni (Bolivia)	17.06.2012	Salty lake	20°S, 67°W
Uyuni (Bolivia)	26.03.2017	Salty lake	20°S, 67°W
Uyuni (Bolivia)	06.10.2019	Salty lake	20°S, 67°W
Greenland (Greenland) 2	05.04.2012	Snow/Ice	74°N, 53°W
Las Vegas	13.08.2020	Desert/Urban Area	36°N, 115°W
Mexico City (Mexico)	29.02.2012	Urban/Topography	19°N, 99°W
Uyuni (Bolivia)	28.10.2019	Salty lake	20°S, 67°W

A. Dataset Generation

For the generation of a representative and consistent dataset to be used to train, validate, and test the proposed DL model, we utilize 21 TerraSAR-X and TanDEM-X bistatic data acquired in bypass configuration, i.e., raw data are digitized with a uniform 8-bit ADC and no further quantization is applied. The acquisitions cover a variety of land cover types, including deserts, snow- and ice-covered regions, forests, and urban areas, characterized by variable local topography. Their location is shown in Fig. 2 and further detailed in Table I. Seventeen acquisitions are used for training and validation, while the remaining four acquisitions are used for testing. We are aware that this could lead to some correlation between the datasets caused by the illumination of the same area on the ground; nevertheless, given the relatively small amount of

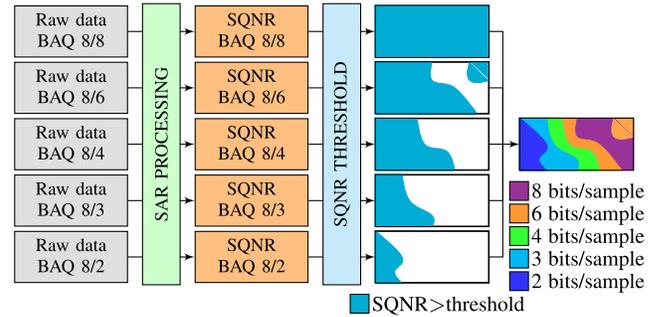


Fig. 3. Approach used to derive the reference BRMs for training the DL architecture based on thresholding on a given performance requirement. In this case, the SQR is selected as a performance parameter; however, the same method can be applied to derive other metrics as well (e.g., phase errors and coherence loss).

available acquisitions, we decided to also include these two areas in testing.

For the generation of the reference BRMs to be used during the supervised training, we requantize the bypass acquisitions on ground using different BAQ rates [i.e., 2, 3, 4, and 6 bits/sample (bps)] and then perform the complete SAR and InSAR processing [45], allowing for the derivation of different SAR and InSAR products for each quantization rate. In order to achieve more granularity in the reference data, even if only integer (BAQ) bitrate values are available, we perform an interpolation on the obtained performance such that we are able to define a fractional bitrate, which satisfies a specific requirement, as proposed in [15]. Afterward, we derive a binary mask for each requantized raw data by setting a threshold on the specific target performance parameter. An overall reference BRM is then computed by selecting the minimum number of bits that ensures a certain performance in the corresponding area of the focused SAR/InSAR product. In this way, we generate a reference map that correctly targets a certain uniform performance in the focused SAR product. An example of this method is depicted in Fig. 3, where the SQR is considered as the target performance

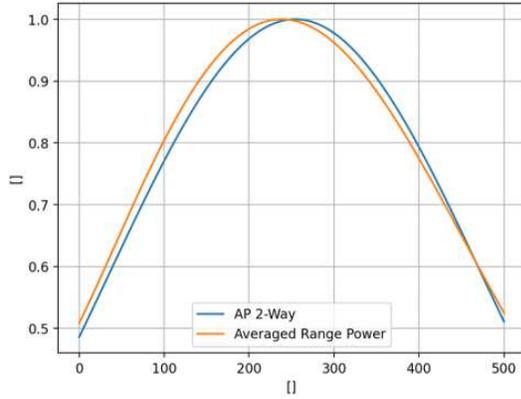


Fig. 4. Comparison between the normalized mean range power for a homogenous TanDEM-X acquisition over Greenland (orange curve) and the two-way normalized elevation AP (blue curve) utilized for this specific acquisition.

metric. Moreover, if a target performance cannot be achieved even in the original bypass acquisitions (e.g., because of particular acquisition geometries or system parameters), we arbitrarily set the number of quantization bits to the minimum value of 2 bps. An example of that is the case in which the $NESZ_{th}$ is larger than the required one due to system parameters (e.g., the used range beam). This criterion could, of course, be changed depending on the objective of the estimation. Since quantization errors in SAR images are integrated within a large area on the ground [see (2)], as introduced in Section III-A [15], the derived BRM typically shows a smooth spatial variability.

Here, $NESZ_{th}$ and $NESZ_q$ are the thermal and quantization $NESZ$ contribution, respectively. The first term represents the system thermal noise floor and varies along the range depending on the system parameters and the second one considers a further degradation introduced by quantization noise. The input to our DL architecture consists of patches of 128×128 samples of uncompressed raw data amplitude, noted as x_k . In order to link this information to the corresponding reference bitrate value, the derived reference BRM is averaged within a window of the same size as the corresponding raw data patch (128×128 samples), centered around the patch center sample and noted as y_k . In this way, a single reference bitrate y_k value is associated with the entire input raw data patch x_k . The achieved granularity (one bitrate value per patch) does not cause a significant loss of information, because of the smooth spatial variability of the original reference BRM.

Furthermore, we observed that preliminary BRMs estimated on homogeneous scenes (such as, e.g., Greenland, where both the backscatter distribution and, consequently, the reference BRM are rather uniform for a given requirement) showed an elevation angle-dependent error caused by the weighting of the AP. Indeed, a good match between the mean range power and the AP range beams can be observed, as it is shown in Fig. 4. Here, the slight shift between the two curves is due to the fact that the comparison (i.e., range pattern compensation) is normally performed in the range-compressed domain and not in the raw data domain, in order to have a precise overlap between the range beam and the received echo. This effect

needs to be compensated in order to remove the impact of the AP on the estimation of the BRM through the DL model.

As part of preprocessing, we therefore consider the following operations to define the input x to the DL architecture:

$$x = \frac{|s_{raw}|}{k_{s,max}} \oslash \tilde{P}_{range} \quad (13)$$

where $|s_{raw}|$ is the amplitude of the raw data. $k_{s,max}$ is a normalization factor, derived by considering raw data statistics on high backscatter areas, which scales the overall distribution of all the input raw data amplitudes between 0 and 1. Finally, \tilde{P}_{range} represents the normalized elevation AP considered for the acquisition, which is applied as element-wise division (\oslash) for each single range line.

Overall, the complete dataset for training the DL model consists of about 11 million independent data patches, derived from 17 TanDEM-X SAR images, out of which 80% are used for training and 20% are used for validation. This deliberate split ensures that the DL model is exposed to a substantial portion of the overall dataset during training, hence allowing for robust learning. Simultaneously, the reserved 20% serves as a distinct validation subset, allowing for the monitoring of the optimization process by assessing the model's generalization capability to previously unseen data.

Finally, for testing the trained DL model, four acquisitions (not included in the training dataset) are selected, representing: Greenland (snow/ice), Uyuni (soil and rock), Las Vegas (urban areas), and Mexico City (urban areas and high-relief topography). The main goal is to evaluate the performance of the method when applied to different scenarios. In fact, homogeneous scenes featuring both high (Greenland) and low (Uyuni) backscatter values are meant to highlight estimation biases throughout the scene; differently, the scenes over Las Vegas and Mexico City are chosen to test the robustness of the method as they feature a high variability in both backscatter and topography. The considered backscatter images (β^0) are depicted in Fig. 5, together with their mean value and standard deviation.

B. DL Model Architecture

The proposed DL architecture for the regression of BRMs is presented in Fig. 6. It is composed of a sequence of three convolutional layers (with 64, 128, and 256 3×3 kernels) with rectified linear unit (ReLU) activation function, interleaved by max-pooling layers that halve the dimensions of the input features at each layer. This first part of the network mirrors the structure of a typical fully CNN (FCNN). This configuration allows the model to extract hierarchical features from the input raw data, capturing both low- and high-level representations. Such a structure is in theory capable of emulating a pseudo-focusing operation as a result of multiple subsequent convolutions, similar to the principle of the unfocused SAR processing, which allows for deriving a noisy, partially focused image by simply applying boxcar filters instead of matched filters in both range and azimuth directions [46]. Moreover, the use of convolutional layers is particularly beneficial for extracting informative spatial relationships and

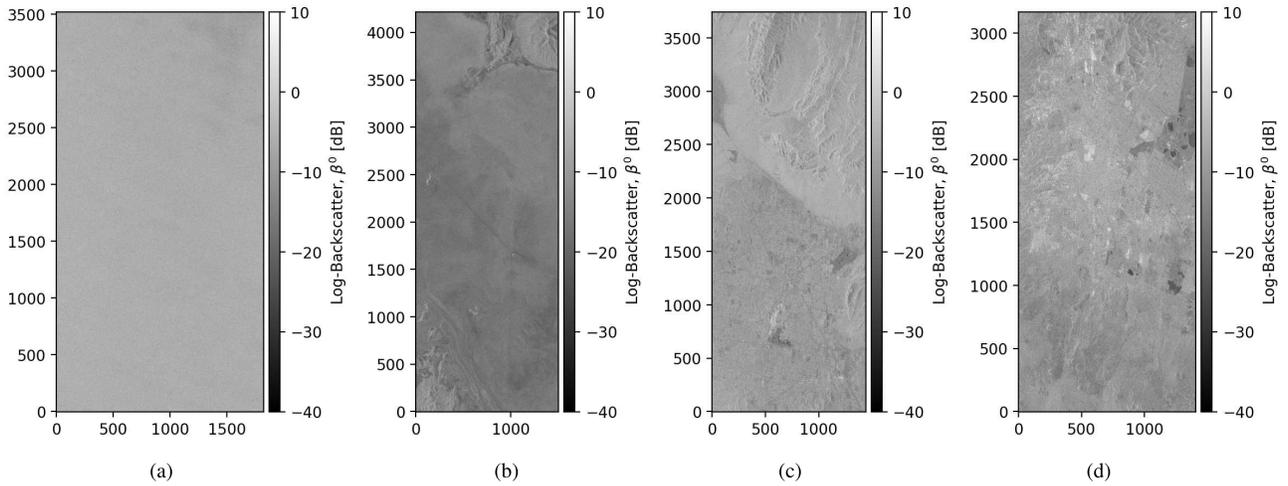


Fig. 5. Backscatter (β^0) for the four test images. (a) Greenland (-6.3 ± 5.6 dB), (b) Uyuni (-9.8 ± 7.6 dB), (c) Las Vegas (-8.6 ± 7.0 dB), and (d) Mexico City (-9.8 ± 7.6 dB). The corresponding mean μ and standard deviation σ are reported at the bottom of each image in the form of $\mu \pm \sigma$.

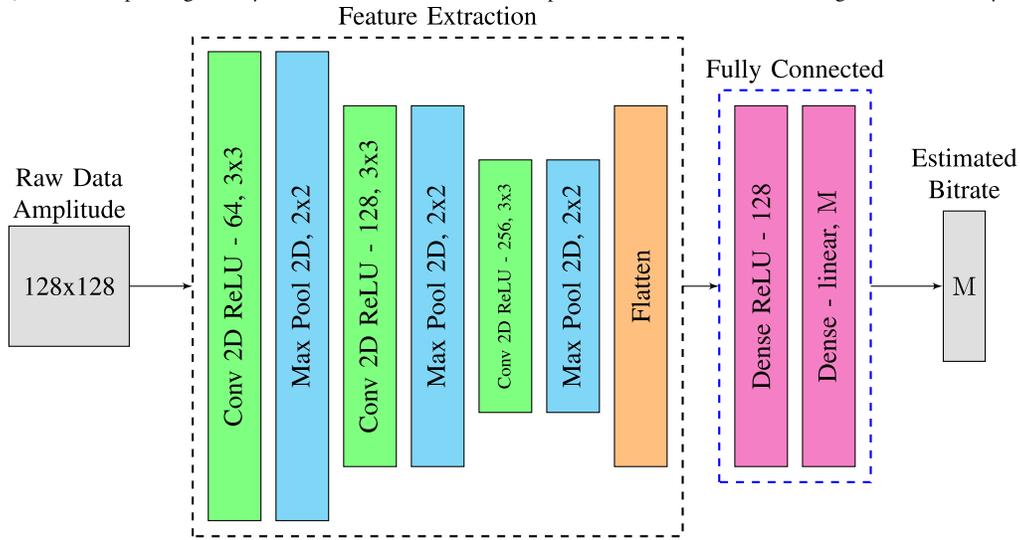


Fig. 6. DL model architecture for estimating the BRM as a regression task. Three consecutive 3×3 convolutional layers with 2×2 max-pooling layers are followed by a flattening operation in order to feed the data to the fully connected (dense) layer responsible for the regression operation. As output, a single value (bitrate) is estimated for each of the M desired optimization cases (i.e., different target performance parameters).

patterns in image-based data [47], [48]. Following the convolutional blocks, a flattening layer [with a dropout to improve convergence (dropout rate of 0.2)] followed by a 128-element fully connected layer is introduced. This latter layer serves as a transition point where the learned features are aggregated and processed for the estimation. A final linear regression layer returns an $M \times 1$ vector of bitrate values, where M represents the number of optimization parameters considered during the training process.

The architecture's hyperparameters (number of layers, number of kernels, size of the dense layer, and size of the input patches) are selected through empirical hyperparameter tuning, as a tradeoff between achievable performance and onboard computational complexity, in a direct synergy with the hardware feasibility assessment presented in Section V.

The input raw data patch of size 128×128 samples (in range and azimuth dimensions) implies the storage in the onboard memory of 128 azimuth lines, which is still a manageable size with current hardware components for

spaceborne SAR. Moreover, 128 range samples represent the standard range block size for the application of the BAQ quantizer in current spaceborne SAR missions.

The output vector is crafted to provide a continuous value representing the predicted bitrate within the range between 2 and 6 bps, which are typical values for SAR applications. The dimension of the output vector M is a deliberate choice, corresponding in size to the number of considered target performance parameters. This allows the architecture to build a one-to-many relationship between a single input patch x_k and a vector of M possible output values y_k^j , where $j = [1, \dots, M]$, hence enhancing the flexibility in the choice of the number of optimization parameters. The notation considers the subscript k as the patch index (i.e., its position), while the superscript j is the target performance case (e.g., for a network optimizing for $\text{SQNR} = [10, 15, 20, 25]$, y_k^1 represents the bitrate required on patch x_k to achieve an SQNR of 10 dB). This principle is depicted in Fig. 7, where an optimization of multiple SQNR target values is considered; the output vector results

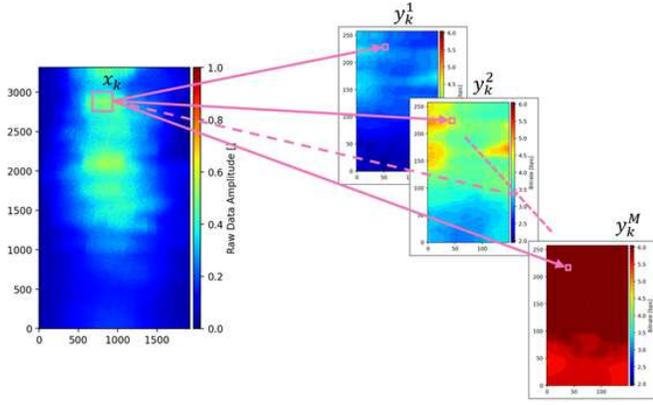


Fig. 7. Example of the flexible approach considered for the input–output relationship of the DL method. For each considered input raw data amplitude patch x_k , M reference outputs y_k^1, \dots, y_k^M values are used during the training phase, allowing the architecture to estimate the required bitrate for multiple optimization parameters at once.

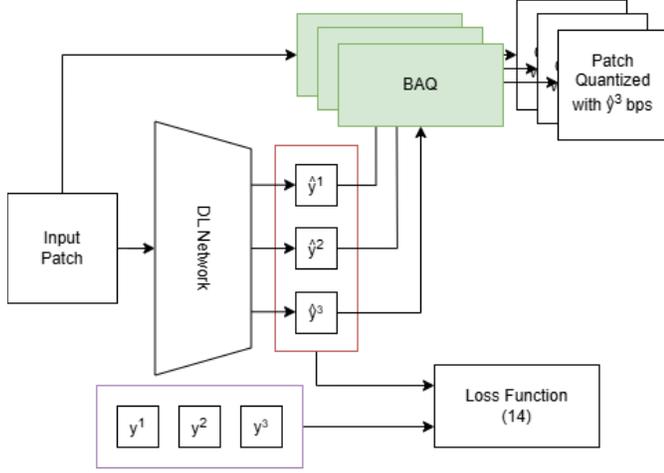


Fig. 8. Flowchart depicting the relationship between true bitrate values y and estimated bitrate values \hat{y} . During the training, both y and \hat{y} are used in the loss function computation, while during inference, the predicted bitrate values \hat{y} are passed to the BAQ to perform the quantization of the input patch. In this example, all the quantization values are used, while in an operative context, the required performance target will define which value must be used.

in a 1-D estimated bitrate vector of size $M \times 1$, where each value corresponds to the required bitrate to achieve a different performance target (e.g., SQNR = [10, 15, 20, 25] dB). This means that, at inference time, a specific BAQ rate can be estimated and applied to blocks of 128×128 pixels within the input raw data in order to achieve a user-defined performance in the higher level SAR product. The practical realization of nonuniform bitrates estimated by the network is done through azimuth-switched quantization (ASQ) within each input patch [49]. Here, by toggling integer quantization rates along the azimuth dimension, it is possible to effectively implement a fractional bitrate.

C. Training Strategy

For training the proposed DL model, we design a multiobjective loss function \mathcal{L} as a combination of two terms, defined as follows. As the goal of the quantization rate regression task is to produce an estimation as close as possible to the reference value, ideally matching it, the first term of the loss

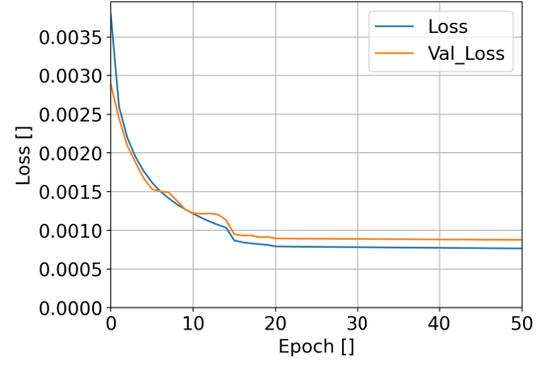


Fig. 9. Evolution of the loss function during training (blue) and validation (orange).

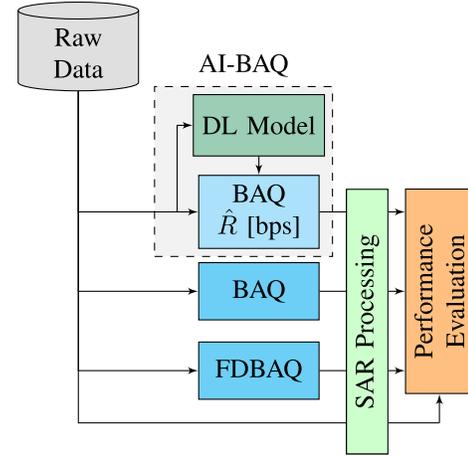


Fig. 10. Flowchart of the performance evaluation framework. The input raw data are fed into the trained DL model, which predicts the required bitrate \hat{R} to be used by a BAQ. The same raw data are also quantized by means of a fixed-rate BAQ and an FDBAQ. The full SAR processing is then performed for all the resulting encoded images, and the performance evaluation is carried out in the focused SAR image domain by comparison with the uncompressed focused SAR data (bypass case).

function is the mean square error (mse) between the true y and estimated values \hat{y} , serving as the core optimization criterion. Additionally, in order to impose further control over the resulting mean quantization rate for all optimization parameters, we introduce as a second term the estimation \hat{y} itself, resulting in

$$\mathcal{L} = \alpha \left[\frac{1}{N} \frac{1}{M} \sum_{k=1}^N \sum_{j=1}^M (y_k^j - \hat{y}_k^j)^2 \right] + (1 - \alpha) \left[\frac{1}{N} \frac{1}{M} \sum_{k=1}^N \sum_{j=1}^M (\hat{y}_k^j)^2 \right] \quad (14)$$

where N represents the batch size used during training, M is the number of desired performance target (i.e., network output size), and α serves as a weighting term for balancing the two different contributions. Since larger bitrate values \hat{y} reflect into higher values of \mathcal{L} , a proper setting of α allows for optimizing a DL model, which, besides the quantization performance, can favor a more stringent constraint on the overall bitrate. As also depicted in Fig. 7, the pedix k refers to the patch index, while the apex j represents the output value for a

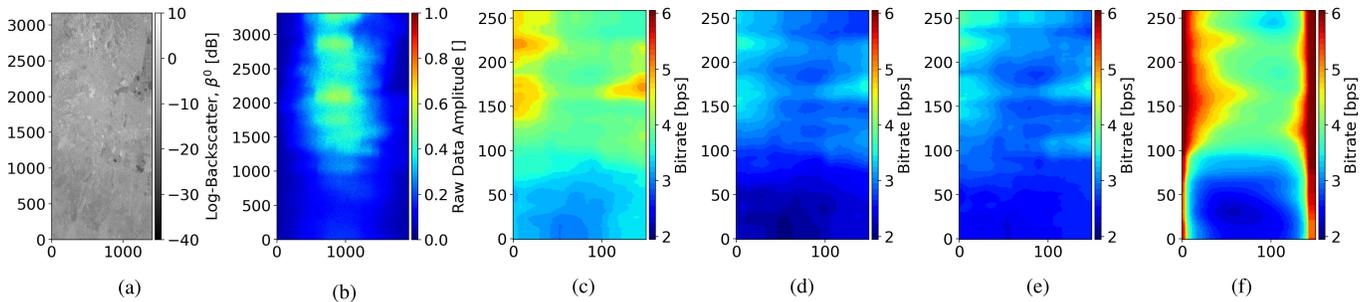


Fig. 11. Comparison of reference BRM on the Mexico City test case. (a) Backscatter (β^0), (b) normalized raw data amplitude ($(|s_{\text{raw}}|/k_{\text{s,max}})$), and (c)–(f) different reference BRMs. (c) Reference BRMs related to a target SQNR = 15 dB, (d) $\gamma_{\text{Quant}} = 0.95$, (e) $\varphi_{\text{err}} = 10^\circ$, and (f) $\text{NESZ}_{\text{tot}} = -19$ dB. For the different requirements, the reference BRMs are varying in terms of both bitrate and spatial pattern.

specific performance requirement. To better describe the flow of information during training and testing, we provide in Fig. 8 how the true y and predicted \hat{y} bitrate values are used as input to (14) and during inference to the BAQ.

To improve convergence and mitigate the effects of vanishing gradients during training, we normalize the input as described in (13), while for the reference output y , we introduce the following linear conversion formula to scale the dynamic between 0 and 1:

$$y = \frac{R}{4} - 0.5 \quad (15)$$

where R stands for the quantization rate, spanning between 2 and 6 bps. In the same manner, the predicted value from the architecture \hat{y} in $[0, 1]$ is converted to a 2–6 bps value representation by inverting (15) into

$$\hat{R} = 4 \cdot (\hat{y} + 0.5). \quad (16)$$

Before the subdivision in training and validation, the dataset is randomly shuffled to mitigate biases associated with sequential data arrangement, contributing to a more representative and unbiased learning for the model. The training is carried out using the ADAM optimizer [50] with an initial learning rate of 10^{-3} , further reduced by a factor of 10 at epochs 15 and 20 to improve convergence. The loss function evolution for both the training (blue) and validation (orange) is shown in Fig. 9. We apply early stopping with a patience of four epochs based on the validation loss. This means that the training was halted if the validation loss did not improve for three consecutive epochs. The choice of patience was determined through empirical verification, balancing training efficiency with model performance to prevent overfitting.

D. Performance Evaluation Framework

The testing phase of the DL model consists of two separate operations. The first step involved testing the BRM regression capabilities, i.e., assessing the estimation accuracy in terms of bps, while the second step consists in the evaluation of the resulting performance metrics in the SAR image domain. For the first task, the estimated BRM is compared with the “true” BRM derived by exploiting a priori knowledge of the scene (with the same approach used for deriving the reference BRMs for the training phase). The comparison consists of measuring

the difference ΔR between the reference and test BRM, where each pixel is evaluated as

$$\Delta R = \hat{R} - R. \quad (17)$$

The estimation error ΔR is considered as mean error (μ) and standard deviation (σ) on the entire test BRM, providing information about the estimation bias and standard deviation. The second part of the test phase consists of the performance evaluation in the final SAR image domain. For this, we consider the framework depicted in Fig. 10. The raw data are quantized with the estimated BRM, according to the proposed AI-based quantization scheme, and the complete SAR processing is carried out, resulting in a focused SAR image. In addition to AI-BAQ, other quantization methods are considered on the same test acquisition as well (i.e., the direct BAQ and FDBAQ) in order to assess the performance of all the methods against the uncompressed (i.e., nonquantized) one. The quantization performance parameters (introduced in Section II-B) are calculated on the final SAR images, providing valuable insight into the respective quantization method, and its settings (i.e., bitrate for the BAQ or performance target for the AI-BAQ) used for encoding the data. Concerning the FDBAQ, we implemented the method as in [13], considering, for the quantization rate derivation, a threshold on the NESZ of -19 dB (corresponding to the requirement on the TerraSAR-X satellite) and taking into account typical TanDEM-X system parameters [44].

IV. EXPERIMENTS AND RESULTS

In this section, we describe the conducted experiments and the achieved results. Each experiment considers different target performance parameters or a combination of them, aiming at achieving a certain uniform performance in the SAR focused domain. Moreover, for each scenario, the same input SAR raw dataset is used during training, while the reference BRMs are separately computed according to the different optimization targets. An example is presented in Fig. 11, where, for a single SAR raw data matrix, the corresponding focused image and different BRMs are depicted, depending on the specific target performance parameter. It then becomes clear how, even if the same optimization criterion is considered (uniform performance), the DL model needs to learn different spatial patterns depending on the reference BRM.

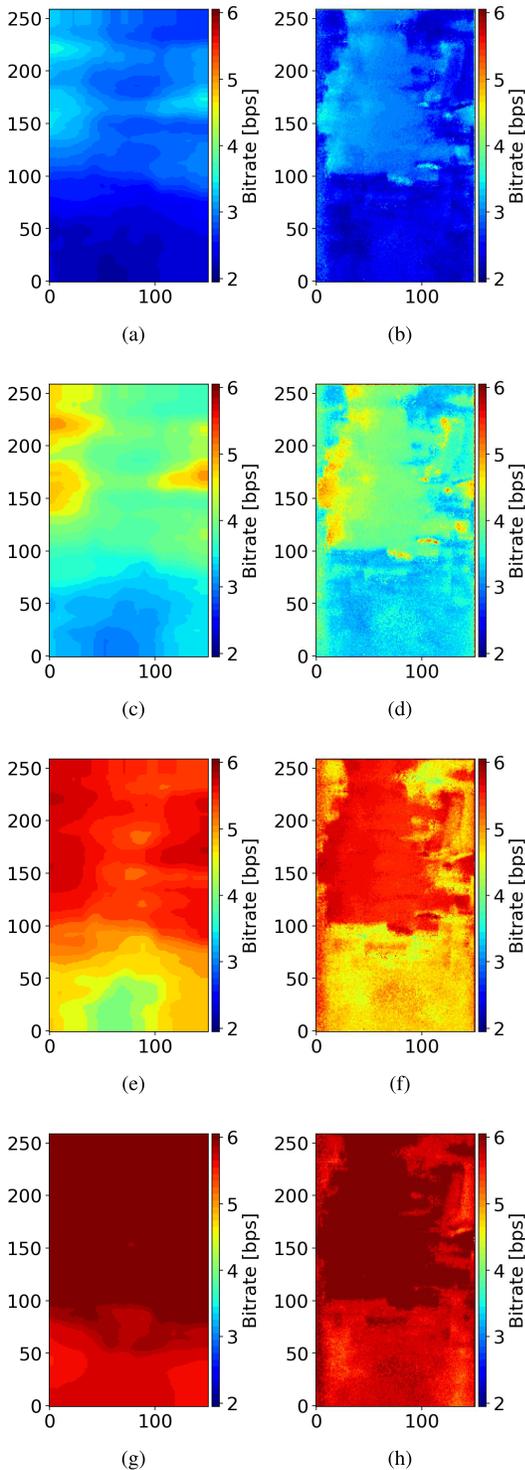


Fig. 12. BRM estimation results for case A_1 on the Mexico City test scene. (Left column) Reference and (right column) estimated BRM for the cases of (a) and (b) SQNR = 10 dB, (c) and (d) SQNR = 15 dB, (e) and (f) SQNR = 20 dB, and (g) and (h) SQNR = 25 dB.

In Section IV-A, we consider a scenario in which a uniform performance in the final SAR product is targeted, considering multiple SQNR values. InSAR-related performance parameters, i.e., the phase error φ_{err} and the coherence degradation γ_{Quant} , and the total NESZ are analyzed in Sections IV-B and IV-C, respectively. In Section IV-D, we concentrate on the combination of multiple performance parameters using

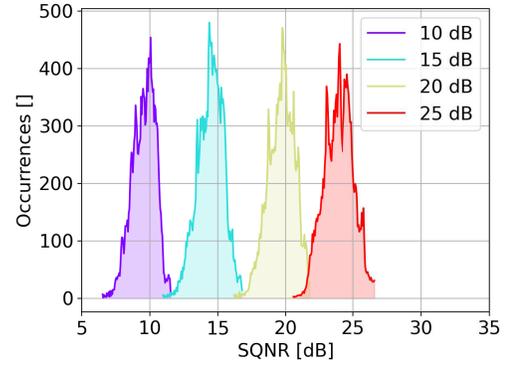


Fig. 13. SAR performance results on the Mexico City test site for the experiment considering a uniform SQNR performance throughout the scene and overall four different SQNR target values.

a single DL model. This case represents a very challenging scenario since it significantly increases the level of complexity of the regression task. The impact of the α parameter [introduced in (14)] is detailed in Section IV-E and, finally, a comparison with reference state-of-the-art (SoA) quantization algorithms is presented in Section IV-F. According to this, it is worth noting that since the optimization criteria adopted by the FDBAQ [13] and by the methods proposed in this article are different, a direct performance comparison is not straightforward. However, we wish to point out that our approach allows for implementing a multiple performance and data rate optimization using a single DL architecture, hence allowing for higher flexibility depending on the specific application requirements.

A. SQNR Optimization

As the first case, we consider the BRM optimization task on a uniform target SQNR. In the specific, we aim at estimating bitrate values for achieving an SQNR of 10, 15, 20, and 25 dB (i.e., $M = 4$). For this case, the loss function considers an α parameter equal to 0. As an example, the reference BRMs and the estimation results over the Mexico City test scene are depicted in Fig. 12 for all four different SQNR cases. Regarding the SAR performance results, the corresponding SQNR histograms computed in the focused SAR domain are presented in Fig. 13. Here, one can appreciate how the distribution of the SQNR correctly varies depending on the target input requirement. The different level of granularity between the reference BRMs and the estimated ones is caused by the fact that each value of the reference BRMs is estimated using a large integration window A_{SAR} in the SAR image, as in (2). Differently, each single pixel in the estimated BRMs represents the output prediction obtained from an input patch of only 128×128 raw data samples. The complete performance for all test acquisitions is reported in Tables II and III (A_1 case), showing both the BRM estimation error (i.e., ΔR in bps) and the resulting performance in the focused domain (SAR performance column). In all cases, the performance of the BRM estimation is almost unbiased (mean $\Delta R \in [-0.4, 0.3]$ bps) with a standard deviation always lower than 0.4 bps. Regarding the SQNR performance, the distribution mean val-

TABLE II

PERFORMANCE RESULTS FOR TWO OUT OF THE FOUR TEST ACQUISITIONS (GREENLAND AND UYUNI TEST AREAS) FEATURING HOMOGENEOUS BACKSCATTER. EACH ARCHITECTURE CONFIGURATION (FIRST COLUMN) IS NAMED AFTER THE SUBSECTION IN WHICH IT IS INTRODUCED. FOR EACH CASE, THE TARGETED PERFORMANCE PARAMETERS ARE REPORTED (SECOND COLUMN), AND FOR EACH ESTIMATION (LINES), WE REPORT ESTIMATION RESULTS IN TERMS OF BITRATE DIFFERENCE ($\Delta R = \hat{R} - R$) AND RESULTING RATE IN THE CYAN COLUMN, AS WELL AS THE FINAL PERFORMANCE AFTER SAR PROCESSING IN THE ORANGE COLUMN. BOTH THE ESTIMATION DIFFERENCE AND PERFORMANCE RESULTS ARE EXPRESSED AS MEAN AND STANDARD DEVIATION ($\mu \pm \sigma$). THE ACHIEVED PERFORMANCE OF SOA METHODS (BAQ AND FDBAQ) ON THE SAME SCENES IS REPORTED FOR ALL THE INVESTIGATED PARAMETERS (SQNR, NESZ, φ_{err} , AND γ_{Quant}) IN THE LOWER TABLE

	$\beta^0 (\mu \pm \sigma)$	Greenland -6.3±5.6 dB				Uyuni -9.8±7.6 dB			
		Est. [bps]		SAR Perf.		Est. [bps]		SAR Perf.	
		\hat{R}	ΔR	-	-	\hat{R}	ΔR	-	-
A ₁	SQNR=10 dB	2.2	0.1±0.1	10.7±0.1 dB	2.2	0.1±0.1	10.2±0.5 dB		
	SQNR=15 dB	3.2	0.2±0.1	15.5±0.2 dB	3.1	0.1±0.2	15.3±0.6 dB		
	SQNR=20 dB	4.1	-0.2±0.2	18.7±0.6 dB	4.2	0.1±0.2	20.4±0.5 dB		
	SQNR=25 dB	4.6	-0.1±0.3	22.5±1.1 dB	5.1	0.0±0.2	25.0±0.6 dB		
B ₁	$\varphi_{\text{err}}=10^\circ$	2.2	0.1±0.1	9.8±0.1 °	2.1	0.1±0.2	9.6±0.5 °		
	$\gamma_{\text{Quant}}=0.95$	2.5	0.0±0.1	0.94±0.0	2.4	0.0±0.1	0.94±0.0		
C ₁	NESZ=-15 dB	2.5	-0.2±0.1	-14.0±0.1 dB	2.0	0.0±0.1	-18.7±0.6 dB		
	NESZ=-17 dB	3.3	-0.7±0.3	-15.9±0.6 dB	2.0	0.1±0.2	-18.7±0.5 dB		
	NESZ=-18 dB	3.7	-0.9±0.5	-16.6±0.7 dB	2.2	0.1±0.3	-18.8±0.5 dB		
	NESZ=-19 dB	2.1	1.1±0.3	-12.9±0.4 dB	2.8	0.2±0.7	-19.2±0.4 dB		
D ₁	SQNR=10 dB	2.2	0.0±0.1	10.4±0.2 dB	2.2	0.1±0.2	10.2±0.5 dB		
	SQNR=15 dB	3.3	0.1±0.2	15.3±0.3 dB	3.1	0.1±0.2	15.3±0.6 dB		
	SQNR=20 dB	4.0	-0.2±0.2	18.4±0.8 dB	4.2	0.1±0.2	20.5±0.5 dB		
	NESZ=-15 dB	2.3	-0.3±0.1	-13.5±0.1 dB	2.0	0.0±0.1	-18.7±0.6 dB		
	NESZ=-17 dB	2.9	-1.1±0.3	-15.2±0.6 dB	2.1	0.1±0.2	-18.7±0.5 dB		
	NESZ=-18 dB	3.2	-1.4±0.5	-15.7±0.7 dB	2.2	0.1±0.3	-18.8±0.4 dB		
E ₁ $\alpha=0.95$	SQNR=10 dB	2.1	0.0±0.1	10.3±0.1 dB	2.1	0.0±0.1	9.8±0.5 dB		
	SQNR=15 dB	3.1	0.1±0.2	15.3±0.3 dB	3.0	0.0±0.2	14.7±0.6 dB		
	SQNR=20 dB	4.0	-0.3±0.2	18.1±0.8 dB	4.1	0.0±0.2	19.9±0.5 dB		
	SQNR=25 dB	4.5	-0.3±0.3	21.6±1.1 dB	5.0	-0.1±0.2	24.6±0.5 dB		
E ₂ $\alpha=0.75$	SQNR=10 dB	2.0	-0.1±0.3	9.8±0.2 dB	2.0	-0.1±0.1	9.5±0.3 dB		
	SQNR=15 dB	2.8	-0.3±0.2	13.4±0.6 dB	2.6	-0.4±0.2	12.4±0.6 dB		
	SQNR=20 dB	3.6	-0.7±0.2	16.7±0.4 dB	3.7	-0.4±0.2	17.6±0.6 dB		
	SQNR=25 dB	4.1	-0.6±0.3	19.2±1.0 dB	4.7	-0.5±0.2	22.7±0.7 dB		

State-of-the-Art Quantization Schemes					
	case	R	SAR Perf.	R	SAR Perf.
		bps	SQNR; NESZ; φ_{err} ; γ_{Quant}	bps	SQNR; NESZ; φ_{err} ; γ_{Quant}
F ₁	BAQ 8:2	2.0	9.3±0.2 dB; -12.7±0.3 dB; 10.4±0.1 °; 0.92±0.0	2.0	9.5±0.2 dB; -18.7±0.6 dB; 10.0±0.3 °; 0.92±0.0
	BAQ 8:3	3.0	15.1±0.2 dB; -15.8±0.2 dB; 5.6±0.1 °; 0.97±0.0	3.0	15.0±0.4 dB; -19.8±0.4 dB; 5.4±0.2 °; 0.98±0.0
	BAQ 8:4	4.0	18.7±0.4 dB; -17.1±0.2 dB; 3.3±0.1 °; 0.99±0.0	4.0	19.8±0.6 dB; -20.1±0.4 dB; 3.0±0.1 °; 0.99±0.0
	BAQ 8:6	6.0	58.2±4.3 dB; -18.4±0.2 dB; 0.0±0.0 °; 1.0±0.0	6.0	31.3±5.0 dB; -20.3±0.4 0.8±0.3 °; 1.0±0.0
F ₂	FDBAQ	3.3	14.8±0.4 dB; -15.4±0.2 dB; 6.1±0.3 °; 0.97±0.0	3.1	15.1±0.3 dB; -19.7±0.5 dB; 5.9±0.2 °; 0.97±0.0

ues correctly follow the target performance, with a spread that varies depending on the backscatter characteristics of the scene.

B. InSAR-Related Parameter Optimization

A different scenario is investigated for the experiment reported at line B₁ of Tables II and III. Here, we consider the optimization for a uniform requirement on a phase error φ_{err} of 10° and a quantization decorrelation γ_{Quant} of 0.95 (where γ_{Quant} represents the quantization decorrelation factor in the factorization of the interferometric coherence presented in [43]). Also, in this case, the parameter α in the loss function is set to 0. Being γ_{Quant} an interferometric parameter, the required rate is derived from the interferometric pairs, while

during inference, we perform the prediction on a single image (reference image) and apply the resulting rate on both raw data (i.e., references and secondary images). The performance is finally evaluated after the interferometric processing. The performance remains consistent and shows similar values as in the previous case, demonstrating that the DL model can correctly manage different optimization parameters. Remarkably, the targeted phase error is correctly achieved in all test cases, including the more inhomogeneous ones, which represent very challenging scenarios due to the high dynamic variation of the backscatter.

C. Total NESZ Optimization

As the thermal NESZ for TerraSAR-X and TanDEM-X typically spans in the range of about -17 up to -25 dB

TABLE III

PERFORMANCE RESULTS FOR THE TWO TEST ACQUISITIONS (LAS VEGAS AND MEXICO CITY TEST AREAS) FEATURING INHOMOGENEOUS BACKSCATTER AND HIGH-RELIEF TOPOGRAPHY. EACH ARCHITECTURE CONFIGURATION (FIRST COLUMN) IS NAMED AFTER THE SUBSECTION IN WHICH IT IS INTRODUCED. FOR EACH CASE, THE TARGETED PERFORMANCE PARAMETERS ARE REPORTED (SECOND COLUMN), AND FOR EACH ESTIMATION (LINES), WE REPORT ESTIMATION RESULTS IN TERMS OF BITRATE DIFFERENCE ($\Delta R = \hat{R} - R$) AND RESULTING RATE IN THE CYAN COLUMN, AS WELL AS THE FINAL PERFORMANCE AFTER SAR PROCESSING IN THE ORANGE COLUMN. BOTH THE ESTIMATION DIFFERENCE AND PERFORMANCE RESULTS ARE EXPRESSED AS MEAN AND STANDARD DEVIATION ($\mu \pm \sigma$). THE ACHIEVED PERFORMANCE OF SOA METHODS (BAQ AND FDBAQ) ON THE SAME SCENES IS REPORTED FOR ALL THE INVESTIGATED PARAMETERS (SQNR, NESZ, φ_{err} , AND γ_{Quant}) IN THE LOWER TABLE

		Las Vegas -8.6±7.0 dB				Mexico City -9.8±7.6 dB					
		$\beta^0(\mu \pm \sigma)$		Est. [bps]		SAR Perf.		Est. [bps]		SAR Perf.	
		Target		\hat{R}	ΔR	-	-	\hat{R}	ΔR	-	-
A ₁	SQNR=10 dB	2.5	0.0±0.3	9.7±1.3 dB	2.7	-0.1±0.3	9.6±0.8 dB				
	SQNR=15 dB	3.5	0.0±0.4	14.7±1.3 dB	3.7	-0.1±0.3	14.5±0.9 dB				
	SQNR=20 dB	4.5	0.0±0.3	20.0±1.3 dB	4.7	-0.1±0.3	19.7±1.0 dB				
	SQNR=25 dB	5.4	-0.2±0.4	23.8±1.3 dB	5.6	-0.2±0.3	24.0±1.1 dB				
B ₁	$\varphi_{err}=10^\circ$	2.5	0.0±0.4	10.4±1.6 °	2.7	-0.1±0.3	10.7±1.2 °				
	$\gamma_{Quant}=0.95$	2.7	0.1±0.2	0.93±0.01	2.8	-0.1±0.3	0.92±0.01				
C ₁	NESZ=-15 dB	2.8	0.3±0.3	-16.1±0.9 dB	2.6	0.0±0.3	-15.5±1.2 dB				
	NESZ=-17 dB	3.5	0.4±0.4	-17.8±0.8 dB	3.2	0.1±0.5	-17.1±0.9 dB				
	NESZ=-18 dB	3.8	0.5±0.5	-18.7±0.8 dB	3.6	0.1±0.5	-18.2±0.9 dB				
	NESZ=-19 dB	4.3	0.4±0.7	-19.5±0.8 dB	4.0	0.2±0.7	-19.2±1.0 dB				
D ₁	SQNR=10 dB	2.5	0.0±0.3	9.8±1.3 dB	2.6	-0.1±0.3	9.6±0.8 dB				
	SQNR=15 dB	3.5	0.0±0.3	14.6±1.4 dB	3.7	-0.1±0.3	14.5±1.0 dB				
	SQNR=20 dB	4.5	-0.1±0.4	19.9±1.3 dB	4.6	-0.1±0.3	19.6±1.0 dB				
	NESZ=-15 dB	2.8	0.3±0.3	-16.1±1.0 dB	2.6	0.0±0.3	-15.5±1.1 dB				
	NESZ=-17 dB	3.4	0.4±0.4	-17.8±0.9 dB	3.2	0.1±0.4	-17.2±0.9 dB				
	NESZ=-18 dB	3.8	0.4±0.5	-18.7±0.8 dB	3.6	0.1±0.5	-18.2±0.9 dB				
E ₁ $\alpha=0.95$	SQNR=10 dB	2.4	-0.1±0.3	9.3±1.2 dB	2.6	-0.2±0.3	9.1±0.8 dB				
	SQNR=15 dB	3.4	-0.1±0.3	14.2±1.3 dB	3.6	-0.2±0.3	14.0±0.9 dB				
	SQNR=20 dB	4.4	-0.1±0.3	19.4±1.3 dB	4.6	-0.2±0.3	19.2±1.0 dB				
	SQNR=25 dB	5.3	-0.3±0.4	23.5±1.3 dB	5.5	-0.3±0.3	23.7±1.1 dB				
E ₂ $\alpha=0.75$	SQNR=10 dB	2.1	-0.4±0.3	8.2±1.1 dB	2.3	-0.5±0.3	7.6±0.9 dB				
	SQNR=15 dB	3.0	-0.4±0.3	12.3±1.3 dB	3.2	-0.5±0.3	12.0±0.9 dB				
	SQNR=20 dB	4.1	-0.4±0.3	17.6±1.3 dB	4.2	-0.6±0.3	17.2±0.9 dB				
	SQNR=25 dB	5.0	-0.6±0.4	22.0±1.4 dB	5.1	-0.6±0.3	21.6±0.9 dB				

State-of-the-Art Quantization Schemes							
		R	SAR Perf.		R	SAR Perf.	
case		bps	SQNR; NESZ; φ_{err} ; γ_{Quant}		bps	SQNR; NESZ; φ_{err} ; γ_{Quant}	
F ₁	BAQ 8:2	2.0	7.7±1.2 dB; -13.3±0.8 dB; 13.3±2.5 °; 0.88±0.02		2.0	6.6±1.4 dB; -12.7±2.9 dB; 15.6±3.1 °; 0.84±0.04	
	BAQ 8:3	3.0	12.9±1.5 dB; -17.3±0.8 dB; 7.3±1.5 °; 0.96±0.01		3.0	11.6±1.8 dB; -16.8±2.3 dB; 8.7±1.9 °; 0.94±0.02	
	BAQ 8:4	4.0	17.8±1.6 dB; -19.9±0.8 dB; 4.1±0.8 °; 0.99±0.0		4.0	16.5±1.8 dB; -19.7±1.5 dB; 4.8±1.0 °; 0.98±0.0	
	BAQ 8:6	6.0	27.2±2.6; -21.7±1.0 1.3±0.3 °; 1.0±0.0		6.0	26.5±2.8 dB; -22.0±1.2 1.5±0.4 °; 1.0±0.0	
F ₂	FDBAQ	3.5	15.4±1.6 dB; -18.7±0.7 dB; 5.7±1.0 °; 0.97±0.0		3.4	13.2±1.0 dB; -18.1±1.3 dB; 7.2±0.9 °; 0.95±0.01	

TABLE IV

PERFORMANCE COMPARISON VGG16 AND THE PROPOSED CNN

VCK190 1* DPUCVDX8G 192 AIEs (C32B6CU1L2S2) @ 1250MHz		
Parameter	VGG16	Proposed Arch.
E2E fps Single Thread	505.43 fps	25271 fps
E2E fps Multi Thread	621.19 fps	31060 fps

depending on the acquisition settings [44], we select four target performance values (i.e., also in this case a DL model with $M = 4$) of NESZ equal to -15, -17, -18, and -19 dB (with again $\alpha = 0$). An example of reference and estimated BRM for the test acquisition of Mexico City is presented in Fig. 14 by considering all different NESZ requirements. Here, it is particularly interesting to note the increase in quantization

bits at the borders of the swath (near and far range), which are characterized by higher (i.e., worse) NESZ values due to a lower gain in the AP. The estimated BRMs also show different spatial patterns with respect to the ones in Fig. 12, accounting for the different optimization task. The complete results over all test acquisitions are reported at line C₁ in Tables II and III. The BRM mean estimation error ΔR always remains below 0.4 and 0.9 bps for the test sites characterized by homogeneous and inhomogeneous backscatter, respectively, with a standard deviation confined between 0.1 and 0.7 bps. A consistent performance is also achieved when computing the total NESZ in the focused SAR domain, showing a good agreement with the target performance. This can be seen, e.g., in Fig. 15, where the histograms of the NESZ for the Mexico City test acquisitions are presented for all different target

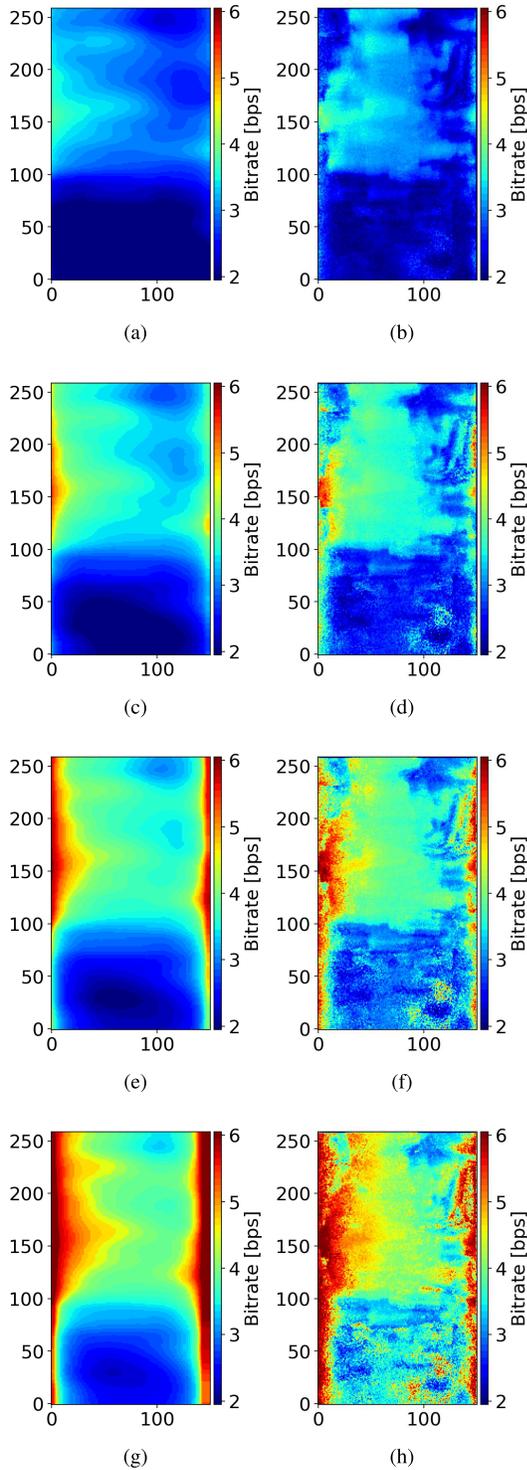


Fig. 14. Estimation results plots for the case C_1 on the Mexico City test scene. (Left column) True and (right column) estimated BRMs for the cases of (a) and (b) $NESZ = -15$ dB, (c) and (d) $NESZ = -17$ dB, (e) and (f) $NESZ = -18$ dB, and (g) and (h) $NESZ = -19$ dB. The bitrate allocation increases at near and far range, accounting for the lower gain of the AP (i.e., larger thermal NESZ).

requirements. The only test case, which severely deviates from the expected performance, is the case of $NESZ = -19$ dB for the Greenland scenario (reported in line C_1 of Table II). This is due to the fact that this specific acquisition was performed

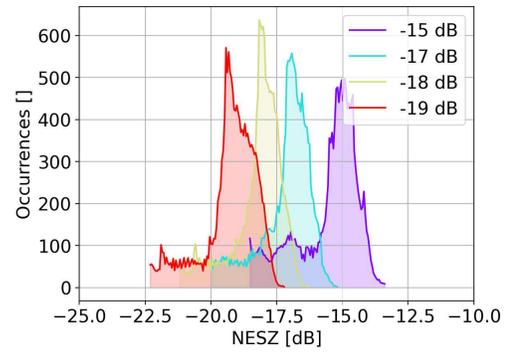


Fig. 15. SAR performance results from the experiment B_1 on the Mexico City test scene for each targeted case of NESZ.

using a very-far-range beam, which did not allow for achieving a sufficient NESZ even in the unquantized bypass data. In this case, the resulting estimated BRM correctly assigns two quantization bits per sample to the entire scene, as learned during training.

D. Combination of Different Performance Parameters

The scenarios investigated in the previous test cases have shown the capability of the DL model to correctly regress different BRMs depending on a specific target performance parameter. In this section, we now investigate a scenario in which different performance parameters, namely, SQNR and NESZ, are jointly estimated by a single DL model. The aim is to test the capacity of the model to correctly recognize different spatial patterns at the same time, through the minimization of a multiobjective loss function. Specifically, we combine three requirements of SQNR and three of NESZ, for a total of $M = 6$ output predicted values. The complete results for this experiment are reported at line D_1 of Tables II and III. The results are extremely consistent with the previous experiments, which considered a single optimization parameter (lines A_1 , B_2 , and C_1 in Tables II and III), confirming the capability of the network to correctly manage different estimation scenarios at a time.

E. Role of the α Parameter

The correct prediction of a certain target performance might, in some cases, be of less relevance in the presence of stringent constraints in the downlink or onboard memory resources. Differently from, e.g., the BAQ quantizer, with the proposed method, the resulting overall data rate is not known a priori, meaning that the required memory storage is also unknown before the encoding is performed. Even though this figure might be empirically derived from large-scale simulations, in this section, we investigate the possibility of tuning the BRM estimation itself by acting on the α parameter in the loss function [introduced in (14)]. To do so, we consider two identical cases with respect to A_1 (uniform SQNR optimization), and we train the DL model with $\alpha = 0.95$ and $\alpha = 0.75$ (identified in Tables II and III as cases E_1 and E_2 , respectively). The first case represents a scenario in which a given performance is required with a small constraint on the resulting data volume (i.e., the architecture will tend to

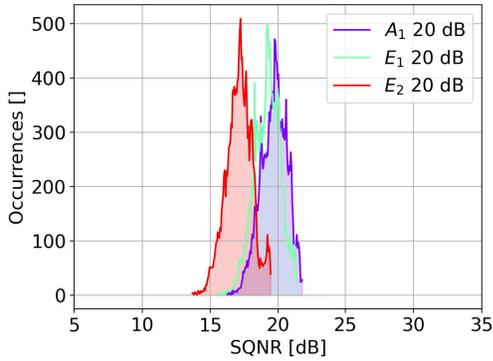


Fig. 16. SAR performance results from the experiment A_1 , E_1 , and E_2 for a target SQNR of 20 dB on the Mexico City test scene. The constraint on the α parameter for the cases E_1 ($\alpha = 0.95$) and E_2 ($\alpha = 0.75$) results in a lower data rate penalizing the performance (as it is confirmed by the histograms shifting toward the left-hand side).

slightly underestimate the bitrate), while the latter is a scenario in which onboard resources are crucial, aiming at significantly reducing the overall bitrate, at the cost of penalizing the final target performance. The effects of these choices for the Mexico City test site are visible in Fig. 16, which depicts the resulting histograms of the SQNR when considering a target SQNR of 20 dB and an α parameter equal to 0 (A_1), 0.95 (E_1), and 0.75 (E_2). As expected, an increase of the α parameter results in a general underestimation of the SQNR, which is marginal for E_1 and more evident for E_2 . Nevertheless, this also allows for a consistent reduction of the overall quantization rate \hat{R} , which decreases from 3.1 bps (A_1) down to 2.6 bps (E_2), leading to an overall reduction of 0.5 bps. A similar trend can be observed for all the other considered test cases.

F. Comparison With SoA Algorithms

In order to provide a first comparison in terms of data rate and SAR performance with SoA methods, we now consider the BAQ and FDBAQ quantization algorithms, summarizing the performance in Tables II and III at lines F_1 and F_2 , respectively. The BAQ can be applied by setting a constant quantization rate throughout the acquisition. We consider the cases of 2, 3, 4, and 6 bps, reporting the resulting performance in terms of SQNR, NESZ, φ_{err} , and γ_{Quant} , respectively. As an example, Fig. 17 depicts the performance in terms of SQNR and NESZ for the Mexico City test area obtained by applying the BAQ and the FDBAQ, respectively. The corresponding results on the same scene achieved by the proposed AI-BAQ framework for different target performance parameters are presented in Fig. 17(c) and (d). Considering the SQNR plots [Fig. 17(a) and (c)], it is possible to appreciate how the AI-BAQ is able to target a specific performance parameter, which is confirmed by the limited dispersion of the distributions with respect to SoA methods. When considering the NESZ, one can note how the AI-BAQ is able to overall adapt to different requirements on the minimum acceptable NESZ, shifting the resulting distribution accordingly. Moreover, it is also worth noting that, when considering the AI-BAQ requirement of NESZ = -19 dB, the corresponding NESZ distribution is similar to the FDBAQ case, showing an overall comparable performance.

Finally, it is of interest to compare the different methods in terms of both target performance and quantization rate, as presented in Fig. 18. As it can be seen, all methods are aligned on achieving a similar mean performance when considering comparable quantization rates. As expected, the AI-BAQ presents a much lower SQNR dispersion with respect to the BAQ and is correctly centered around the desired target performance. For better comparison, we have included the performance for the BAQ at the same rate of the AI-BAQ by means of ASQ. For a high-performance target, the AI-BAQ shows a better performance than the BAQ as the optimization criterion (i.e., PO-BAQ) considers the minimum rate in each area of the scene to satisfy the requirement. To further support the direct comparison of the methods, in Appendix 2, the performance figures for the cases discussed in Fig. 18 are presented as a visual comparison. The precision of the AI-BAQ in targeting a given performance requirement is clearly recognizable by the level of SQNR homogeneity in the final SAR image.

V. PRELIMINARY HARDWARE FEASIBILITY ASSESSMENT

In this section, we investigate the feasibility of a potential hardware architecture for the proposed DL-based method for the estimation of BRMs.

Regarding the details of the proposed DL architecture, the total number of trainable parameters (i.e., network weights) is strictly related to the size of the dense layer (128 elements), to the input patch size (128×128 samples), and to the number of filtering kernels. Therefore, the total number of trainable parameters considering $M = 5$ is $\sim 8.76 \times 10^6$. If we assume saving in single precision, the memory needed for storing this network setup is about 34 MB. This value represents a worse case scenario, as it could be reduced by performing architecture storage optimization (e.g., pruning and weights quantization [51]).

For efficient FPGA implementations, it is essential to represent the SAR raw data in fixed-point number format at the DL model input; this has to be considered as part of the preprocessing operation in (13). Given that the initial layers involve convolutions (more specifically, correlations), a feasible architecture for performing 2-D/3-D convolutions across all feature inputs (channels) of the preceding layer is depicted in Fig. 19 and represents the steps of loading the image from external double-data-rate (DDR) Synchronous Dynamic Random-Access Memory (SDRAM) into the input buffer, loading the weights from DDR into the input buffer, and performing the calculations and storing the results from the output buffer back to DDR. To enhance the performance of convolution operations, it is necessary to employ multiple blocks with different kernel weights, operating in an interleaved mode that synchronizes memory reads/writes and calculations. The proposed real-time operation on the data from the instrument represents one possible solution, while other approaches could be considered (e.g., applying the method as part of offline encoding in the mass memory).

For CNN-based AI applications, various hardware accelerators are available for NG DPUs. In our case, we consider Xilinx Versal, which provides an excellent foundation and

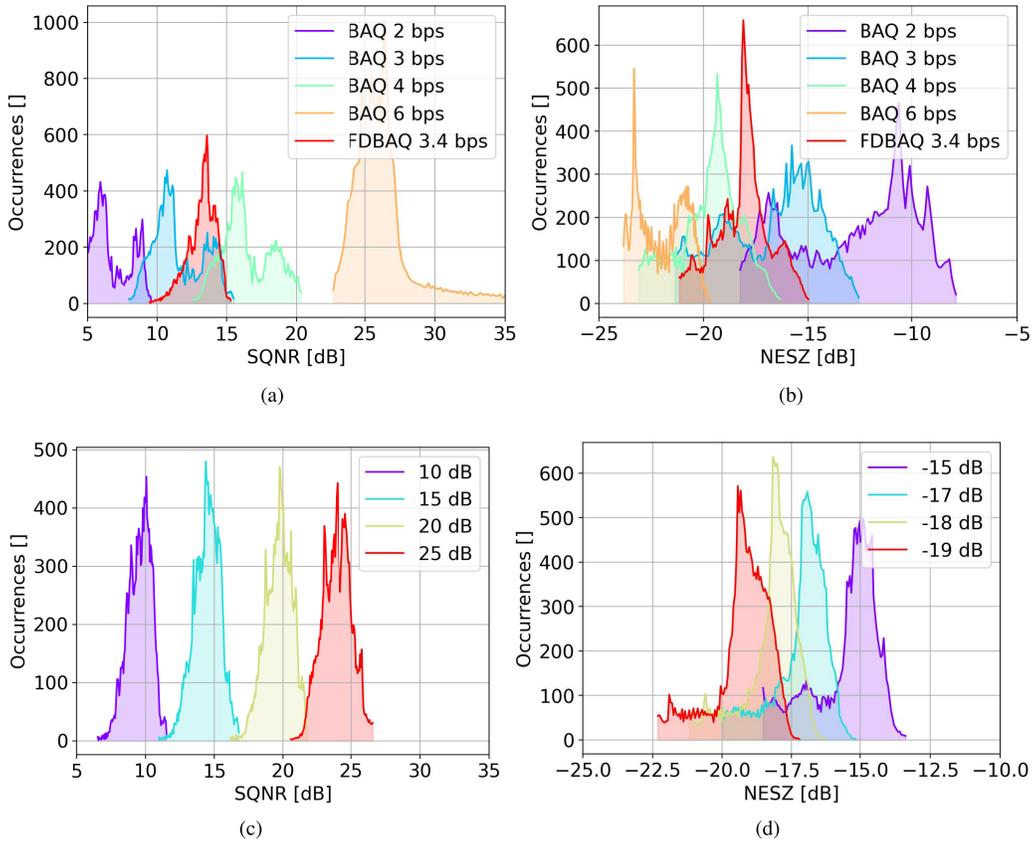


Fig. 17. (a) and (b) SAR performance results in terms of SQNR and NESZ for the SoA cases F_1 and F_2 on the Mexico City test scene. (c) and (d) Achieved results of the AI-BAQ are reported and highlight the greater flexibility of the method.

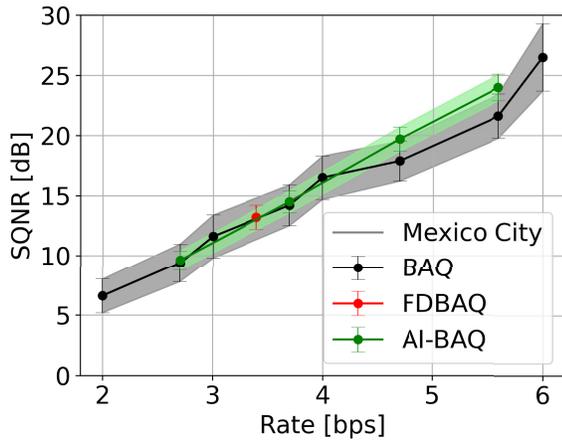


Fig. 18. Performance results in terms of SQNR for the SoA cases and the AI-BAQ on the Mexico City test scene with respect to the quantization rate. Solid dots represent the performance mean value, while the error bars (and colored background parts) represent its standard deviation. For the sake of comparison, in addition to the BAQ integer rates (2, 3, 4, and 6 bps), the resulting rates achieved by the AI-BAQ (2.7, 3.7, 4.7, and 5.6 bps) have been implemented by means of ASQ for the BAQ.

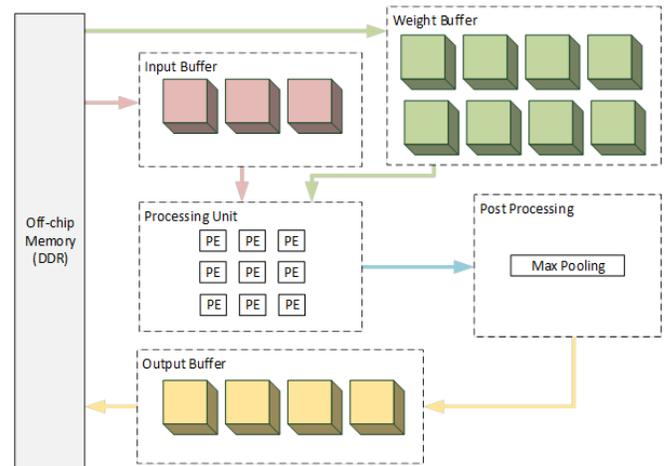


Fig. 19. Block diagram for convolution implementation.

figures, the number of multiply and accumulate operations (MAC OPs) required in a convolution layer can be determined by the following equation:

$$\text{MAC OPs}_{\text{Conv}} = K_h \cdot K_w \cdot F_{\text{IN}} \cdot F_{\text{OUT}} \cdot R_h \cdot R_w \quad (18)$$

where K_h and K_w denote the kernel's height and width, respectively; F_{IN} and F_{OUT} represent the input and output features, respectively; and R_h and R_w correspond to the resulting height and width, respectively. For the architecture

can serve as a baseline for the hardware implementation of this compression method [52]. The toolchain supports various AI functions, such as 2-D/3-D convolution, ReLU, max pooling, flattening, and fully connected layers, along with their architectural interconnections. To provide performance

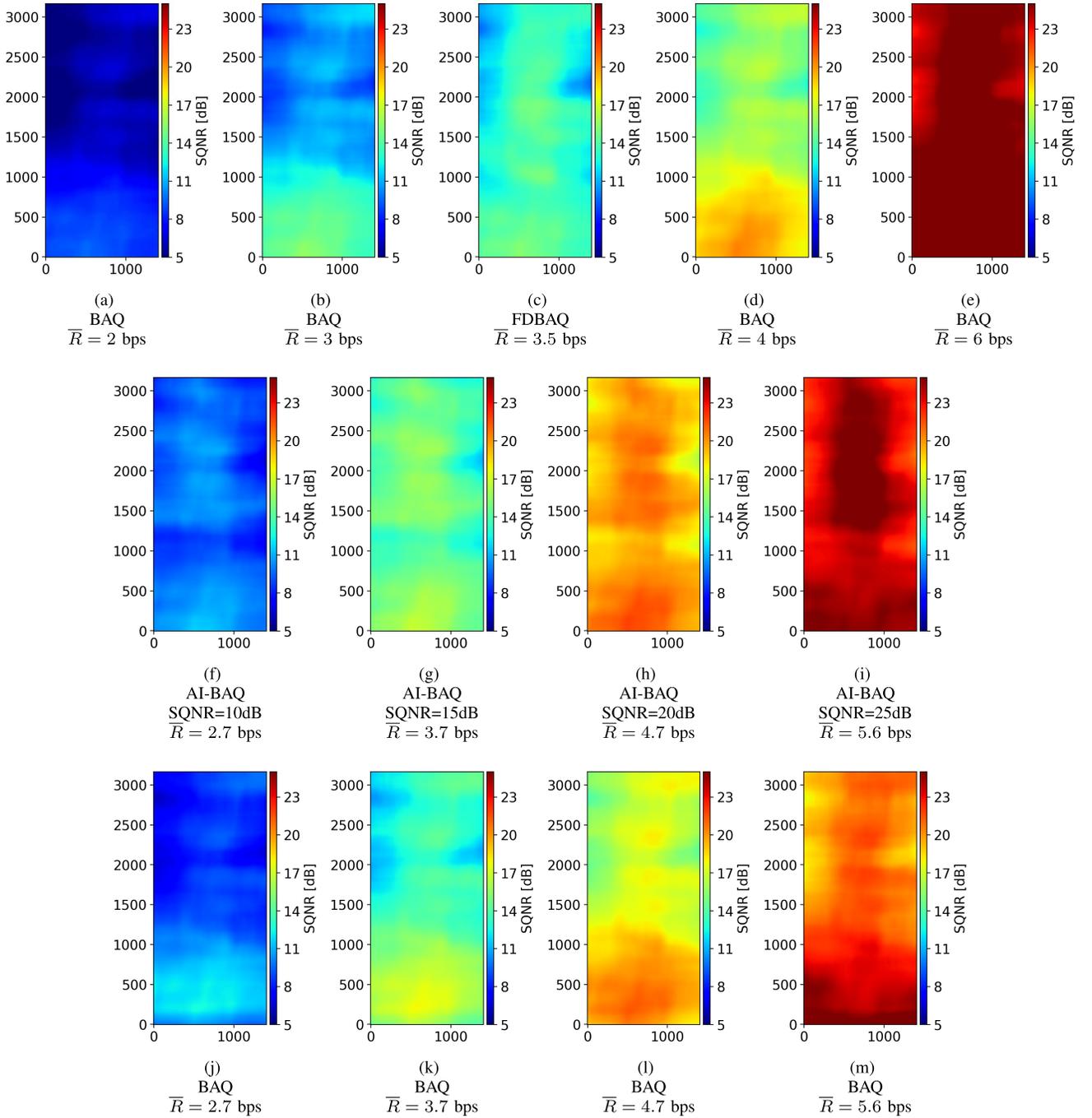


Fig. 20. Visual comparison of SQNR in the SAR image domain over the Mexico City test scene between (a)–(e) SoA BAQ and FD-BAQ, (f)–(i) AI-BAQ of case A_1 targeting uniform SQNR values, and (j)–(m) BAQ with noninteger rates matching the average bitrate of the AI-BAQ. For each case, the average rate \bar{R} is indicated. The proposed method results in a much more homogeneous performance with respect to the SoA methods, even at the same bitrate allocation (last row), despite the highly heterogeneous backscatter.

under consideration, the total number of MAC OPs for all convolutional layers amounts to approximately 613.5 million. Similarly, for a fully connected dense layer, the MAC OPs can be calculated as

$$\text{MAC OPs}_{\text{Dense}} = F_{\text{IN}} \cdot F_{\text{OUT}} \quad (19)$$

which results in 8.4 million operations, leading to a total of 621.9 million MAC OPs for the entire architecture.

Looking ahead to future SAR missions, expected data rates are projected to reach approximately 3000 Mb/s. Assuming an 8-bit sample size and a patch size (frame) of 128×128 pixels, this equates to 131 072 bits per frame. The real-time performance requirement in frames per second (fps_r) can be calculated as

$$\text{fps}_r = \frac{3000 \frac{\text{Mbit}}{\text{s}}}{131072 \frac{\text{bit}}{\text{frame}}} = 22\,888. \quad (20)$$

To estimate the performance, we use the Xilinx CNN benchmark with the VGG16 model (Vitis-AI Model Zoo Name: tf_vgg16_imagenet_224_224_30.96G) as a comparison. The selected hardware platform is the VCK190, featuring 1xDPUCVDX8G 192 AIEs (C32B6CU1L2S2) operating at 1250 MHz with fixed-point calculations. The VGG16 CNN uses the same kernel size (3×3) and downsampling (stride) size (2×2), sharing a similar basic structure with respect to the proposed architecture.

Table IV provides a detailed performance comparison between the VGG16 model, which has 30.96 GOPs, and the proposed CNN architecture with 0.62 GOPs. To qualitatively analyze the performance, the fps_r value is multiplied by a factor of approximately 50 ($30.96/0.62$).

Given the assumptions made, the investigation of the feasibility for $\text{fps}_r = 22\,888$ from (20) suggests that the proposed architecture can successfully meet the hardware requirements of NG SAR missions.

VI. DISCUSSION

The results obtained from the application of the AI-BAQ with different settings and over different scenarios demonstrate that an effective link between the SAR raw data domain and the focused SAR image domain can be established using a pure DL architecture. This allows for the derivation of an application-oriented quantization scheme, targeting a specific performance in higher level SAR products. As an example, one could imagine a possible future mission scenario to plan an acquisition for differential InSAR applications. By knowing this information a priori, the data could be quantized onboard, minimizing the phase distortion and maximizing the coherence, providing the user with the best possible data quality.

The proposed DL model for the estimation of BRMs is extremely compact (only about 34 MB are needed to store the parameters onboard) and, at the same time, can jointly implement a set of different BAQ-based quantizers without requiring any additional computational complexity or the storage onboard of a priori information in the form, e.g., of LUTs. This represents a considerable step forward with respect to the SoA, leading to a quantization scheme that can automatically adjust its settings simply depending on the characteristics of the input SAR raw data and on the desired output performance. Moreover, in this work, we concentrated on the investigation of a single optimization criterion, i.e., achieving a uniform performance in the focused SAR products, applied to different performance parameters. Since the DL model showed itself to be able to learn different spatial patterns depending on the reference BRM used in training, this opens up new possibilities to extend this concept to further optimization criteria by changing the generation process of the reference BRMs. These could be oriented, e.g., toward the joint maximization of the SNR and the minimization of the bit allocation in the presence of certain types of targets, always keeping the flexibility of having multiple application-oriented settings available. Additionally, if needed, the entire DL model could be easily modified by performing a new

training on ground and then simply updating the network coefficients onboard.

Certainly, the proposed methodology represents a starting point for the development of an operational quantization scheme that could be used in a real spaceborne SAR mission scenario. Several aspects will require further investigation in order to derive a robust DL architecture for operational deployment. In particular, the training dataset needs to be extended in order to better represent different acquisition scenarios and backscatter conditions, and further investigations on the DL architecture itself could lead to a general performance improvement as well. Moreover, a crucial aspect is represented by the overall data rate considering a certain orbit duty cycle and acquisition scenario, which, at the current stage of development, cannot be a priori defined or controlled. Promising insights can be drawn from the analysis on the α parameter; by acting on the loss function during the model optimization phase, one can impact the overall data volume. A proper tuning of the α parameter could be performed by vectorizing it with respect to the different output performance targets as well as by considering global acquisition scenarios with determined data rate requirements. This aspect could be combined with a learning phase characterized by the use of different reference BRMs, oriented toward a more efficient data reduction, keeping in mind the desired high-level performance.

Finally, the positive outcome of the preliminary hardware feasibility assessment suggests that the proposed method is suitable for onboard implementation in present-generation spaceborne SAR systems, representing a first step toward the use of onboard AI in SAR missions.

VII. CONCLUSION AND OUTLOOK

In this article, we investigated a novel approach to perform a dynamic BAQ bitrate allocation for the onboard quantization of SAR raw data by means of a DL regression model. The main advantages of the proposed method rely on the fact that no a priori information is required by the system for its onboard implementation, hence allowing for an adaptive bitrate allocation only depending on the characteristics of the SAR raw data and on the considered target performance in the corresponding higher level SAR product.

We presented the relevant aspects and details of the model design as well as the definition of the training, validation, and testing datasets and strategies, together with an assessment of the estimation performance for a set of independent test acquisitions. We investigated different optimization scenarios, which confirmed that an accurate BRM estimation can be achieved by the proposed DL model, which remains consistent when the performance parameters are evaluated on the final SAR and InSAR products. The proposed method is capable of providing multiple optimization methods at once, being the first bitrate allocation method to achieve this without a priori information on the investigated scene. We addressed the issue of the potential uncertainty in the resulting memory storage requirement (caused by a variable rate allocation) by performing a preliminary customization

of the loss function, which introduces a further constraint on the resulting bit allocation. The comparison with the SoA BAQ and FDFAQ algorithms highlights the flexibility of the proposed method to meet the desired performance on different scenes. Finally, the introduction and evaluation of a possible hardware implementation of the proposed DL architecture also resulted in a positive outcome for a future onboard implementation using current spaceborne-ready technology.

While DL offers strong potential in Earth observation, especially where large training datasets are available, the lack of generalization and the prediction unreliability in the presence of out-of-distribution values, i.e., not seen during training, represent a major drawback for large-scale operational deployment. In this context, the proposed framework represents an initial investigation into the use of DL for onboard SAR raw data quantization, confirming promising potential.

As an outlook to future activities, we intend to further develop the methodology toward a realistic joint optimization of both performance and data rate requirements in a global SAR mission scenario. To efficiently validate the method on a global scenario, the dataset will be enriched with more experimental and synthetic generated, aiming at improving the network generalization capabilities, covering multiple polarizations and including different optimization criteria as well as additional SAR acquisition configurations, such as, e.g., wide-beam, multichannel, and multistatic systems.

APPENDIX A PERFORMANCE TABLES

This appendix contains the detailed results of the performance analysis conducted over the four considered test images: Greenland and Uyuni (Table II), characterized by homogeneous backscatter, and Las Vegas and Mexico City (Table III), characterized by inhomogeneous backscatter and by the presence of high-relief topography.

APPENDIX B VISUAL COMPARISON

This appendix provides an additional visual comparison of SQNR performance calculated in the SAR image domain. In particular, the cases discussed in Fig. 18 (i.e., SoA BAQ with integer rate, FDFAQ, AI-BAQ, and BAQ at the rate achieved by the AI-BAQ) are presented in the form of 2-D SQNR maps [according to (6)] in Fig. 20. Here, it is possible to observe the high degree of adaptivity of the proposed method to the scene characteristics, resulting in a uniform performance, and in targeting the requirement without the need for a priori knowledge.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their helpful comments that improved this paper. The method is patented under the German patent (No. 10 2022 123 897) and European patent pending (No. 23 192 942.3).

REFERENCES

- [1] A. Moreira et al., "Tandem-L: A highly innovative bistatic SAR mission for global observation of dynamic processes on the Earth's surface," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 2, pp. 8–23, Jun. 2015.
- [2] F. Queiroz de Almeida, M. Younis, G. Krieger, and A. Moreira, "Multichannel staggered SAR azimuth processing," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2772–2788, May 2018.
- [3] J. Mittermayer et al., "MirrorSAR: An HRWS add-on for single-pass multi-baseline SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–18, 2022.
- [4] M. Nogueira Peixoto, G. Krieger, A. Moreira, C. Waldschmidt, and M. Villano, "On the exploitation of CubeSats for highly accurate and robust single-pass SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–16, 2023.
- [5] D. Geudtner, R. Torres, P. Snoeij, M. Davidson, and B. Rommen, "Sentinel-1 system capabilities and applications," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 1457–1460.
- [6] P. Rosen et al., "The NASA-ISRO SAR (NISAR) mission dual-band radar instrument preliminary design," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 3832–3835.
- [7] K. Kellogg et al., "NASA-ISRO synthetic aperture radar (NISAR) mission," in *Proc. IEEE Aerosp. Conf.*, Mar. 2020, pp. 1–21.
- [8] N. Pierdicca, "The Copernicus L-band SAR mission ROSE-L (radar observing system for Europe)(conference presentation)," *Proc. SPIE*, vol. 11154, Sep. 2019, Art. no. 111540E.
- [9] D. Geudtner, M. Tossaint, M. Davidson, and R. Torres, "Copernicus Sentinel-1 next generation mission," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2021, pp. 874–876.
- [10] M. Zonno et al., "Sentinel-1 next generation: Main mission and instrument performance of the phase 0," in *Proc. 13th Eur. Conf. Synth. Aperture Radar*, Mar. 2021, pp. 1–5.
- [11] M. Martone, B. Bräutigam, and G. Krieger, "Quantization effects in TanDEM-X data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 583–597, Feb. 2015.
- [12] R. Kwok and W. T. K. Johnson, "Block adaptive quantization of Magellan SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 27, no. 4, pp. 375–383, Jul. 1989.
- [13] E. Attema et al., "Flexible dynamic block adaptive quantization for Sentinel-1 SAR missions," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 766–770, Oct. 2010.
- [14] P. Guccione, M. Belotti, D. Giudici, A. M. Guarnieri, and I. Navas-Traver, "Sentinel-1A: Analysis of FDFAQ performance on real data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6804–6812, Dec. 2015.
- [15] M. Martone, N. Gollin, P. Rizzoli, and G. Krieger, "Performance-optimized quantization for SAR and InSAR applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–22, 2022.
- [16] U. Benz, K. Strodl, and A. Moreira, "A comparison of several algorithms for SAR raw data compression," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 5, pp. 1266–1276, May 1995.
- [17] P. Guccione, M. Scagliola, and D. Giudici, "Principal components dynamic block quantization for multichannel SAR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 2090–2093.
- [18] M. Martone, M. Villano, M. Younis, and G. Krieger, "Efficient onboard quantization for multichannel SAR systems," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1859–1863, Dec. 2019.
- [19] E. Magli and G. Olmo, "Lossy predictive coding of SAR raw data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 5, pp. 977–987, May 2003.
- [20] T. Ikuma, M. Naraghi-Pour, and T. Lewis, "Predictive quantization of range-focused SAR raw data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1340–1348, Apr. 2012.
- [21] M. Martone, N. Gollin, M. Villano, P. Rizzoli, and G. Krieger, "Predictive quantization for data volume reduction in staggered SAR systems," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5575–5587, Aug. 2020.
- [22] N. Gollin, J. Giez, M. Martone, P. Rizzoli, R. Scheiber, and G. Krieger, "Dynamic predictive quantization for staggered SAR: Experiments with real data," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–5, 2023.
- [23] A. Moreira, "Real-time synthetic aperture radar (SAR) processing with a new subaperture approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 4, pp. 714–722, Jul. 1992.
- [24] N. Gollin, R. Scheiber, M. Martone, P. Rizzoli, and G. Krieger, "SAR imaging in frequency scan mode: System optimization and potentials for data volume reduction," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–20, 2023.

- [25] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [26] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.
- [27] X. X. Zhu et al., "Deep learning meets SAR: Concepts, models, pitfalls, and perspectives," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 4, pp. 143–172, Dec. 2021.
- [28] J. Kuester, W. Gross, S. Schreiner, W. Middelmann, and M. Heizmann, "Adaptive two-stage multisensor convolutional autoencoder model for lossy compression of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–22, 2023.
- [29] G. Guerrisi, F. D. Frate, and G. Schiavon, "Artificial intelligence based on-board image compression for the Φ -Sat-2 mission," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 8063–8075, 2023.
- [30] Y. Guo, Y. Chong, and S. Pan, "Hyperspectral image compression via cross-channel contrastive learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–18, 2023.
- [31] Q. Xu et al., "Synthetic aperture radar image compression based on a variational autoencoder," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [32] C. Fu, B. Du, and L. Zhang, "SAR image compression based on multi-resolution and global context," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–5, 2023.
- [33] N. Foix-Colonier, J. Amao-Oliva, and F. Sica, "Deep learning-based compression and despeckling of SAR images," in *Proc. 15th Eur. Conf. Synth. Aperture Radar*, 2024, pp. 237–242.
- [34] C. Hay, C. Anderson, L. Donnell, M. Ireland, and M. Yaghoobi, "Adaptive raw SAR data compression using machine learning enhanced block adaptive quantization," in *Proc. Small Satell. Conf.*, 2024, pp. 1–17.
- [35] G. Pilikos, M. Azcueta, R. Camarero, and N. Floury, "Raw data compression for synthetic aperture radar using deep learning," in *Proc. 8th Int. Workshop Payload Data Compress. (OBPDC)*, 2022.
- [36] G. Pilikos, M. Azcueta, R. Camarero, and N. Floury, "Raw SAR data compression with deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2024, pp. 2546–2549.
- [37] R. M. Asiyabi, A. Anghel, P. Rizzoli, M. Martone, and M. Datcu, "Complex-valued autoencoder for multi-polarization SLC SAR data compression with side information," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2023, pp. 1787–1790.
- [38] N. Gollin, M. Martone, G. Krieger, and P. Rizzoli, "AI-based performance-optimized quantization for future SAR systems," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2023, pp. 1771–1774.
- [39] N. Gollin, M. Martone, E. Imbembo, S. Knoll, G. Krieger, and P. Rizzoli, "AI for performance-optimized quantization in future SAR systems," in *Proc. 15th Eur. Conf. Synth. Aperture Radar*, 2024, pp. 231–236.
- [40] N. S. Jayant and P. Noll, *Digital Coding of Waveform (Prentice-Hall 1206 Signal Processing Series)* Englewood Cliffs, NJ, USA: Prentice-Hall, 1984.
- [41] S. Huber, M. Younis, and G. Krieger, "The TanDEM-X mission: Overview and interferometric performance," *Int. J. Microw. Wireless Technol.*, vol. 2, nos. 3–4, pp. 379–389, Aug. 2010.
- [42] M. Martone, B. Brütigam, P. Rizzoli, C. Gonzalez, M. Bachmann, and G. Krieger, "Coherence evaluation of TanDEM-X interferometric data," *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 21–29, Sep. 2012.
- [43] P. Rizzoli, L. Dell'Amore, J.-L. Bueso-Bello, N. Gollin, D. Carcereri, and M. Martone, "On the derivation of volume decorrelation from TanDEM-X bistatic coherence," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3504–3518, 2022.
- [44] G. Krieger et al., "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3317–3341, Nov. 2007.
- [45] P. Prats et al., "Taxi: A versatile processing chain for experimental TanDEM-X product evaluation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 4059–4062.
- [46] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar*, vol. 11. New York, NJ, USA: Wiley, 1991.
- [47] F. Sica, G. Gobbi, P. Rizzoli, and L. Bruzzone, " Φ -Net: Deep residual learning for InSAR parameters estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3917–3941, May 2021.
- [48] D. Carcereri, P. Rizzoli, D. Ienco, and L. Bruzzone, "A deep learning framework for the estimation of forest height from bistatic TanDEM-X data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 8334–8352, 2023.
- [49] M. Martone, B. Brütigam, and G. Krieger, "Azimuth-switched quantization for SAR systems and performance analysis on TanDEM-X data," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 181–185, Jan. 2014.
- [50] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [51] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio, "Quantized neural networks: Training neural networks with low precision weights and activations," *J. Mach. Learn. Res.*, vol. 18, no. 187, pp. 1–30, 2018.
- [52] A. Ushiroyama, M. Watanabe, N. Watanabe, and A. Nagoya, "Convolutional neural network implementations using vitis AI," in *Proc. IEEE 12th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2022, pp. 365–371.



Nicola Gollin received the M.Sc. degree in information and communication engineering from the Università degli Studi di Trento, Trento, Italy, in 2018. His M.Sc. thesis was on predictive quantization for data volume reduction in staggered synthetic aperture radar (SAR) systems at the Microwaves and Radar Institute, German Aerospace Center (DLR), Weßling, Germany.

Since 2018, he has been with the Radar Science Group, Microwaves and Radar Institute, DLR, where he is currently a Project Engineer. He contributed to multiple international SAR project activities in cooperation with research and industry partners. His research interests include the analysis and design of onboard data volume reduction methods in future SAR systems.

Mr. Gollin was a recipient of the Best Paper Award at German Microwave Conference (GeMiC) in 2019 and the ARGUS Science Award in 2019. He received the IRS Young Scientist Award at the International Radar Symposium (IRS) in 2024 and one of the best student paper awards at European SAR Conference (EUSAR) in 2024.

Mr. Gollin was a recipient of the Best Paper Award at German Microwave Conference (GeMiC) in 2019 and the ARGUS Science Award in 2019. He received the IRS Young Scientist Award at the International Radar Symposium (IRS) in 2024 and one of the best student paper awards at European SAR Conference (EUSAR) in 2024.



Michele Martone received the B.Sc. and M.Sc. degrees in telecommunication engineering from the University of Naples "Federico II," Naples, Italy, in 2006 and 2009, respectively, and the Ph.D. degree from Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2019.

Since 2009, he has been with the Microwaves and Radar Institute, German Aerospace Center (DLR), Weßling, Germany, where he has contributed to the development and optimization of the TerraSAR-X and TanDEM-X spaceborne synthetic aperture radar

(SAR) systems and has been involved in multiple projects in collaboration with research and industry partners on the design of future SAR missions. He has authored or co-authored more than 100 research papers in peer-reviewed journals and international conference proceedings and holds three patents in the field of SAR. His research interests include signal processing and machine learning for SAR interferometry, and the design of innovative SAR systems, with particular focus on efficient quantization and data volume reduction methods.

Dr. Martone regularly serves as a member of the technical program committees for several international conferences and workshops and is a member of the Consultative Committee for Space Data Systems (CCSDS). He received the DLR Science Award in 2018 and the Best Paper Award at German Microwave Conference in 2019.



Ernesto Imbembo received the master's degree in telecommunication engineering from the University of Naples "Federico II," Naples, Italy, in 2010. His master's thesis was carried out at German Aerospace Center (DLR), Weßling, Germany.

In 2010, he was with Fraunhofer FHR, Wachtberg, Germany, as a Radar Engineer for the "CropSense" Project. In 2011, he joined Metasensing BV, Munich, Germany, where he worked on the signal and data processing of airborne synthetic aperture radar (SAR) data. He worked on the PanelSAR Project during his experience at SSBV, Noordwijk, The Netherlands, and joined Airbus Defence and Space GmbH, Friedrichshafen, Germany, as an Instrument Performance Engineer for the ESA BIOMASS Project. Since 2019, he has been at ESTEC, European Space Agency (ESA), Noordwijk, where he currently works as an SAR Instrument Principal Engineer of the Harmony Project. His main activity relates to all aspects concerning SAR, from concept and design to signal processing and implementation.



Max Ghiglione received the B.Sc. and M.Sc. degrees in electronics engineering from the Polytechnic University of Turin, Turin, Italy, in 2016 and 2018, respectively.

From 2018 to 2023, he was with Airbus Defence and Space GmbH, Munich, Germany, as a Senior Onboard Processing Engineer, where he led projects on artificial intelligence/machine learning and field-programmable gate array (FPGA)-based systems. He is currently a Digital Payload Engineer with European Space Agency (ESA), Noordwijk, The Netherlands, specializing in digital onboard processing for RF payloads. His research interests include FPGA design, digital signal processing, payload processor architectures, and digital RF equipment.



Stefan Knoll received the Dipl.-Ing. degree in electrical and information engineering from the University of Stuttgart, Stuttgart, Germany, in 2007, with a focus on simulation and development of an adaptive decision feedback equalizer for an optical transmission system for an implementation on a field-programmable gate array (FPGA).

In 2008, he joined EADS Deutschland GmbH, Ottobrunn, Germany, and worked as an FPGA and Board Design Development Engineer for a digital wideband receiver (SIGINT) with RADAR signal analysis. Since 2011, he has been with Airbus Defence and Space GmbH, Friedrichshafen, Germany, as a Development Engineer at several projects for space electronics, e.g., Radar Instrument Control Units, Mass Memories, and Power and Motor Control Units with the focus on data handling subsystems on FPGAs.



Gerhard Krieger (Fellow, IEEE) received the Dipl.-Ing. (M.S.) and Dr.Ing. (Ph.D.) (Hons.) degrees in electrical and communication engineering from the Technical University of Munich, Munich, Germany, in 1992 and 1999, respectively.

From 1992 to 1999, he was with the Ludwig Maximilian University of Munich, Munich, where he conducted multidisciplinary research on neuronal modeling and nonlinear information processing in biological and technical vision systems. Since 1999, he has been with the Microwaves and Radar Institute, German Aerospace Center (DLR), Weßling, Germany, where he started as a Research Associate developing signal processing algorithms for a novel forward-looking radar system employing digital beamforming on receive. From 2001 to 2007, he led the New SAR Missions Group, which pioneered the development of advanced bistatic and multistatic radar systems, such as TanDEM-X, as well as innovative multichannel synthetic aperture radar (SAR) techniques and algorithms for high-resolution wide-swath SAR imaging. Since 2008, he has been the Head of the Radar Concepts Department, DLR, which currently hosts about 60 scientists focusing on new SAR techniques, missions, and applications. He has been serving as a Mission Engineer for TanDEM-X, and he has also made major contributions to the development of the Tandem-L mission concept, where he led the Phase-0 and Phase-A studies. Since 2019, he has been holding a professorship at Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen, Germany, and he is the author or co-author of more than 100 peer-reviewed journal articles, nine invited book chapters, about 500 conference papers, and more than 30 patents.

Prof. Krieger received several national and international awards, including two best paper awards at European Conference on Synthetic Aperture Radar, two Transactions Prize Paper Awards of the IEEE Geoscience and Remote Sensing Society, and the W.R.G. Baker Prize Paper Award from the IEEE Board of Directors. In 2014 and 2024, he served as the Technical Program Chair for European Conference on Synthetic Aperture Radar, and in 2014, he served as a Guest Editor for IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He has been an Associate Editor of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING since 2012.



Paola Rizzoli (Member, IEEE) received the B.Sc. and M.Sc. degrees in telecommunication engineering from the Politecnico di Milano (Polimi), Milan, Italy, in 2003 and 2006, respectively, and the Ph.D. degree (Hons.) in electrical engineering and information technology from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2018.

From 2006 to 2008, she was with the Politecnico di Milano and Aresys s.r.l., Milan, a PoliMi spin-off company, as a Scientific Researcher and a Project Engineer. In 2008, she joined the Microwaves and Radar Institute, German Aerospace Center (DLR), Weßling, Germany, as a Project Engineer, where she has been involved in the development and optimization of the TerraSAR-X and TanDEM-X missions, concentrating, in particular, on the generation of the TanDEM-X global digital elevation model. Since 2016, she has been leading the Radar Science Group, Satellite SAR Systems Department, DLR, being responsible for the final performance assessment of the global TanDEM-X DEM and the generation of the global TanDEM-X Forest/Non-Forest map. Her main research interests include synthetic aperture radar (SAR) systems design, data reduction techniques, estimation theory, signal processing, and artificial intelligence algorithms.

Dr. Rizzoli received the DLR Science Award in 2018, the Best Paper Award at German Microwave Conference in 2019, and the Best Poster Award at the ESA Worldcover Conference in 2017. She regularly serves as a reviewer for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.