Pattern Recognition 134 (2023) 109082

Contents lists available at ScienceDirect

Pattern Recognition

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

A goal-driven unsupervised image segmentation method combining graph-based processing and Markov random fields

Marco Trombini*, David Solarna, Gabriele Moser, Silvana Dellepiane

Electrical, Electronics, and Telecommunication Engineering and Naval Architecture Department (DITEN), Università degli Studi di Genova, Via all'Opera Pia 11a, Genoa I-16145, Italy

ARTICLE INFO

Article history:
Received 19 December 2021
Revised 4 August 2022
Accepted 27 September 2022
Available online 30 September 2022

Keywords:
Graph signal processing
Segmentation
Markovian modeling
Parametric model estimation
Pattern recognition
Synthetic aperture radar
Magnetic resonance imagery

ABSTRACT

Image segmentation is the process of partitioning a digital image into a set of homogeneous regions (according to some homogeneity criterion) to facilitate a subsequent higher-level analysis. In this context, the present paper proposes an unsupervised and graph-based method of image segmentation, which is driven by an application goal, namely, the generation of image segments associated with a user-defined and application-specific goal. A graph, together with a random grid of source elements, is defined on top of the input image. From each source satisfying a goal-driven predicate, called seed, a propagation algorithm assigns a cost to each pixel on the basis of similarity and topological connectivity, measuring the degree of association with the reference seed. Then, the set of most significant regions is automatically extracted and used to estimate a statistical model for each region. Finally, the segmentation problem is expressed in a Bayesian framework in terms of probabilistic Markov random field (MRF) graphical modeling. An ad hoc energy function is defined based on parametric models, a seed-specific spatial feature, a background-specific potential, and local-contextual information. This energy function is minimized through graph cuts and, more specifically, the alpha-beta swap algorithm, yielding the final goal-driven segmentation based on the maximum a posteriori (MAP) decision rule. The proposed method does not require deep a priori knowledge (e.g., labelled datasets), as it only requires the choice of a goal-driven predicate and a suited parametric model for the data. In the experimental validation with both magnetic resonance (MR) and synthetic aperture radar (SAR) images, the method demonstrates robustness, versatility, and applicability to different domains, thus allowing for further analyses guided by the generated products.

1. Introduction

The analysis and interpretation of digital images is consistently aided by their partition into homogeneous areas, which possibly correspond to meaningful regions of interest (ROIs) in the scene. This partition process is called image segmentation. In particular, according to its classical definition [1], the segmentation result is the partition of an image into disjoint, non-empty, and connected subregions, for which some predicate of homogeneity is satisfied. In addition, the same predicate of homogeneity must not be valid for the union of any such subregions. As a natural development of such a definition, the purpose of segmentation can often be oriented to the detection of image regions that are meaningful within a particular application domain.

In applicative domains such as the medical field [2], this fundamental activity is often performed manually, resulting in time consuming endeavors. Indeed, spurred by the recent growth in data availability, automatic image segmentation techniques occupy an important time-saving role in the image processing domain [3].

State-of-the-art segmentation techniques can be divided into two types of models, deformable and parametric. The methods in both categories have many advantages in terms of usability, as they can be completely unsupervised or they may not require a large amount of data.

Deformable models originate from [4] and [5] and are defined by curves or surfaces that move under the influence of internal forces coming from within the model itself and external forces computed from the image data. Later in the development, an ensemble of their geometrical features gave birth to models such as [6]. They are still commonly used in multiple domains, such as in the biomedical field [7] and in remote sensing [8].

* Corresponding author.
E-mail address: marco.trombini@edu.unige.it (M. Trombini).

https://doi.org/10.1016/j.patcog.2022.109082

Conversely, when dealing with parametric models, the segmentation problem is often formulated in terms of a Bayesian probabilistic graphical model [9]. The Bayesian statistical framework allows for exploiting data and prior distributions [10], and aims at linking image segmentation with ROI detection and fostering weakly supervised or interactive approaches [11]. In such a context, Markovian modeling is often adopted as a spatial regularization approach [12–14].

In addition to the aforementioned techniques, the availability of wide varieties of image datasets and the recent developments documented in the deep learning literature have produced effective advancements for semantic segmentation [15,16]. These advancements are actually different with respect to the image segmentation addressed in this manuscript, because semantic segmentation aims to partition the image to obtain class labeling with the use of a training set.

The proposed method combines the parametric framework with a graph-based formulation. Indeed, graph-based approaches have been widely used in the signal processing context [17]. For instance, images (and volumes) can be associated with weighted graphs via the one-to-one correspondence between the pixels (and voxels) and the nodes of the graphs. Relations and similarity distances between pixels are generally expressed in terms of the minimum-length path between nodes. As an example, the method described in [18] starts from a set of points, called seeds, and it looks for maximally connected paths and assigns a cost to each pixel related to each seed. Such a procedure is equivalent to the fuzzy connectivity computation proposed in [19], which integrates intensity similarity with topological connectedness. In [18], a graph cut is performed after the cost-computation phase to obtain the segmentation. One might prove that such a method can be formally described in terms of the image-foresting transform (IFT) approach [20]. In [21], the IFT was shown to be equivalent to the minimum-spanning forest (MSF) cuts constrained by seeds; hence, some analogies between the graph-cut in [18] and other methods, e.g., the threshold in the watershed hierarchy in [22], might be investigated.

A graph-based approach was also proposed in [23], where the IFT transform was extended to support superpixel computations. However, according to the work in [24], the generation of superpixels is the result of an over-segmentation of the input image. As a consequence, the predicate of homogeneity is also satisfied by the union of subregions, in contrast with the image segmentation definition. Indeed, the detection of meaningful image parts requires a further merging step.

In the present work, the definition of image segmentation is equipped with a goal-oriented component, which is included in the predicate of homogeneity that the desired segments are meant to satisfy. Three operational phases can be identified (see Fig. 1). They combine the mapping of the images into a weighted graph with the definition of a Bayesian probabilistic graphical framework. Phase 1 refers to the cost-computation process and takes into consideration the numerical/homogeneity and topological/connectivity properties of the nodes with respect to the chosen seeds. Such a step is also responsible for defining the predicate of homogeneity with the intent of integrating the application-specific and user-defined goal into the segmentation process. However, actual segmentation is not performed at this step. Conversely, the set of the most significant regions of interest is automatically identified on the basis of their homogeneity and geometric properties, so that a probabilistic model can be fitted on them. Then, in Phase 2, the parameters of the probabilistic model are estimated on the samples identified in Phase 1. The only a priori knowledge required in Phase 2 is related to the type of data under analysis, which affects the choice of the parametric model to be fitted on the data. Finally, in Phase 3, a Markov random field (MRF) model [25,26] is defined together with a suitable energy function, whose minimization yields the segmentation. A graph-cut algorithm is applied in this third phase. Thanks to the contextual information brought about by Markovian modeling, no merge step is necessary as a post-processing operation before generating the final product.

The goal-oriented characteristic proposed here determines one of the novel contributions of the proposed approach. Indeed, the homogeneity predicate is conditioned by the application domain and allows for splitting the image into ROIs and background (i.e., reject regions), rather than finding the complete image partitioning. In addition, in standard graph-based approaches, the cost computation and the graph-cut are strictly linked, yielding the final result. Conversely, the three phases proposed here separate the cost computation used for finding ROI samples from the final MRF formulation associated with the graph cut by using an intermediate statistical procedure. As a result, the presented graph-based approach addresses numerical, syntactic, and contextual phases, thus capturing the topological, statistical, and goal-related properties of the input images. Hence, the first phase is data-driven, and exploits the informative content of the image; the second phase is goal-driven and discloses the information of the application domain in which the method is applied; and the third phase collects and combines the information from the previous steps to obtain the final result. Furthermore, in the MRF formulation, two novel contributions are included in the pixelwise term, which enclose the topological information and distinguish ROIs and background according to the goal-driven objective.

In particular, two different versions of the method are proposed, which differ in the definition of the goal-oriented predicate and correspond to images with additive Gaussian noise (e.g., passive cameras and magnetic resonance imaging) and with multiplicative speckle (e.g., synthetic aperture radar, sonar systems, and ultrasound imaging). These two cases encompass a variety of scenarios associated with data collected by both active coherent and passive imaging instruments. Among these scenarios, the experimental results achieved by the proposed method are analyzed in two specific contexts, i.e., biomedical and remote sensing imagery. In particular, applications for low-intensity field magnetic resonance (MR) images of wrist bones from the dataset in [2] and synthetic aperture radar (SAR) images acquired by both the TerraSAR-X and COSMO-SkyMed satellite missions are described. The proposed method is proven to be effective in such frameworks, regardless of the possible issues that typically affect the considered types of image data, e.g., MRI artifacts [27] and speckle in the radar imagery [28]. In addition, it is remarkable that recent works in these contexts prove the interest towards methods that can help when the amount of annotated data is insufficient [29,30].

In summary, the novelties of the present work are: (i) the definition of a method that combines graph-based processing and Markov random fields to integrate local and global properties of the image based on a specific a priori task to be fulfilled; (ii) an experimental analysis showing the effectiveness and versatility of the proposed method, whose applicability has been examined with respect to two different domains and in comparison with multiple state-of-the-art solutions; and (iii) the development of an unsupervised formulation not requiring any labeled dataset.

The paper is organized as follows. Section 2 gives an overview of the proposed methodological framework and introduces the novel definition of goal-driven image segmentation. In particular, Section 2.3 describes the graph construction procedure, while Section 2.4 describes the included parametric models, together with the parameter estimation strategies, and Section 2.5 presents both the Markovian framework and the adopted energy minimization strategy. Finally, Section 3 reports the experimental analysis, with details about the performances achieved by the proposed method in two different application domains and compar-

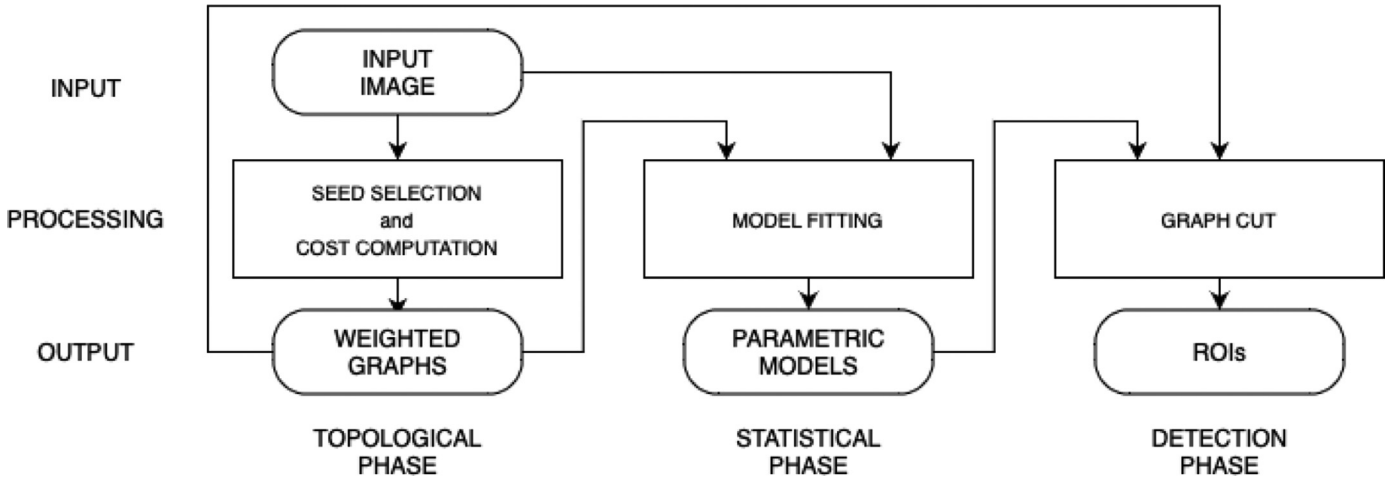


Fig. 1. Method flow chart.

isons with state-of-the-art solutions. In addition, for the sake of clarity and ease of reading, appendix Appendix A contains a table summarizing the entire set of notations adopted within the manuscript.

2. The proposed method

2.1. The goal-driven approach to segmentation

The present section is devoted to the definition of the goal-driven formulation of image segmentation that is adopted here. Let $I \subset \mathbb{Z}^2$ be the pixel lattice, let $\mathcal{X} = \{x_i\}_{i \in I}$ be the collection of the observed pixel intensities, and let the image be formally represented as $X = I \times \mathcal{X}$. With this formalization, each individual pixel is conveniently associated with the pair (i, x_i) so that both the spatial location $i \in I$ within the imaged area and the intensity x_i are specified.

According to the classical definition [1], the region segmentation of an image X into disjoint non-empty subregions $X_1, \ldots, X_M$ has to satisfy the following requirements:

i) $\bigcup_{j=1}^{M} X_j = X$;
ii) X_j is connected for all $j = 1, \ldots, M$;
iii) $P(X_j) = \texttt{true}$ for all $j = 1, \ldots, M$;
iv) $P(X_j \cup X_k) = \texttt{false}$ for all $j \neq k$,

where P is a predicate of homogeneity. To focus the aforementioned classical definition with particular emphasis on the goal to be achieved, one can consider a goal-driven definition of image segmentation, which is aimed at extracting some ROIs that are interesting for the application itself. Here, instead of searching for an image partition, the goal is to split the image X into a finite collection of disjoint ROIs $X_1, \ldots, X_M$ and a background (also denoted as reject-region) $B = \left(\bigcup_{j=1}^{M} X_j \right)^c$, where $(\cdot)^c$ denotes the complementary set. Therefore, each ROI has to satisfy the following requirements:

i) $\bigcup_{j=1}^{M} X_j = X \setminus B$;
ii) X_j is connected for all $j = 1, \ldots, M$;
iii) $P(X_j) = \texttt{true}$ for all $j = 1, \ldots, M$;
iv) $P(X_j \cup X_k) = \texttt{false}$ for all $j \neq k$;
v) $P(B) = \texttt{false}$,

where P is again a goal-driven predicate of homogeneity, while the background B does not satisfy this homogeneity predicate (when B is not connected, all the partitions do not satisfy the predicate). Here, the goal-driven predicate P encompasses the information related to the topological, morphological, and statistical properties of data, and depends on the user-defined goal to be achieved.

The predicate is defined by the user and reflects their expertise (e.g., a medical doctor, a photointerpreter, an ultrasound technician, etc.). Indeed, as in actual practice, it is important to define what is the aim of the segmentation procedure. The presented method aims to emulate such an approach by including some prior-knowledge expressed in terms of the application goal to be achieved. Such a goal is formally expressed in mathematical terms and is deployed by coding the corresponding control functions.

Within the goal-driven definition of image segmentation, the function P drives the processing, as described in Section 2.2. In particular, the goal is defined as a collection of properties that are desired for the ROIs and not for the background. These properties are determined by the application for which the segmentation task is addressed. Among the possible options, in this paper, two properties are always deployed. They refer to the intensity and the distribution of the pixels within each ROI. Indeed, the former is related to some visual features of the regions to detect (e.g., bright or dark target structures), while the latter is aimed at properly modeling the distribution of the data within the regions taking into account the type of data under analysis (e.g., modeling the type of noise affecting the available acquisition). In addition, other properties, including positional, topological, or geometric properties, can be integrated in the proposed framework as well. In the following, the image is modeled as a realization of a two-dimensional stochastic process.

To formally define the aforementioned predicate P, let h_g and h_s be two predicates that are used to represent the desired pixel-intensity and statistical properties. Namely, denoting as 2^X the power set of X (i.e., the collection of all subsets of X), they are Boolean-valued functions $h_g, h_s : 2^X \to \{\texttt{true}, \texttt{false}\}$.

Specifically, let $A \subseteq X$; the predicate h_g is defined as:

$$h_g(A) = \text{true} \iff \mathbb{E}\{x_i\} \in J_m \wedge \text{Var}\{x_i\} \in J_v \quad \forall (i, x_i) \in A, \tag{1}$$

where J_m and J_v are two intervals that specify the values of the mean and the variance that the user deems feasible for their target ROIs, and where $\wedge$ indicates a logical "and."

Then, the predicate h_s addresses the statistical properties of the ROIs. Indeed, h_s holds true on a set $A \subseteq X$ if the intensities x_i of the pixels in A, when conditioned to their membership in the same region, are independent and identically distributed (i.i.d.) and their probability density function (pdf) belongs to a given parametric pdf family f (e.g., Gaussian, Gamma).

In addition, other properties that may be considered can be similarly taken into account. For instance, a positional property can be considered if the application requires that the target regions are located roughly in a certain area of the image (e.g., in the center of the image). For this purpose, let h_p be a further predicate such that:

$$h_p(A) = \texttt{true} \quad \Longleftrightarrow \quad i \in \bar{I}, \quad \forall (i, x_i) \in A, \tag{2}$$

where $\bar{I} \subset I$ indicates the area of the imaged scene where the target ROIs are supposed to be located (e.g., the central area of the image, without a border of a predefined size).

Now, let H be the finite collection of predicates involved:

$$H = \{h_g, h_s, h_{other}\}, \tag{3}$$

with h_{other} collectively indicating possible optional predicates (e.g., h_p) encoding further desired properties specified by the user. Then, the predicate $P : 2^X \rightarrow \{\texttt{true}, \texttt{false}\}$ is defined as:

$$P(A) = \bigwedge_{h \in H} h(A). \tag{4}$$

In the proposed approach, the predicate and its components are used to guide the behavior of the presented segmentation method. Yet, it is not necessarily ensured that the output segmentation result will satisfy all the related Boolean conditions. Nevertheless, P drives the computation, i.e., it is used to encode the goal within the processing steps, and it is applied for specific operational needs. In particular, due to the seed-growing nature of the proposed approach, the predicate will be conveniently applied to either the seed points or the ROI samples.

2.2. Overview of the proposed approach

The proposed segmentation method[1] is aimed at extracting, in an unsupervised fashion, elements of interest from 2-D images, taking into consideration the user-defined goal and the application domain which it is applied to.

The resulting output would be a label image $L = \{l_i\}, i \in I$, defined over the same (pixel) lattice, whose pixels are associated with labels identifying a particular element, i.e. $l_i \in \Lambda$, with $\Lambda = \{0, 1, 2, \ldots, M\}$ and M being the number of different elements, i.e., of the different ROIs, and the label $l_i = 0$ standing for a background pixel.

The rationale of the method is to combine weighted graphs, parametric density modeling, and Markov random fields to benefit from both the topological and the statistical properties of the input image in the generation of an output map made of the regions of interest to the application. The flowchart is presented in Fig. 1 and highlights three different phases. It is worth noting that the proposed method acts driven by the predicate P defined in the previous section. Indeed, the predicate, and thus the goal-driven philosophy of the segmentation method, come into play in different parts of the processing pipeline summarized in Fig. 1.

In the first phase, an initial set of seeds is placed within the pixel lattice I. More in details, a random grid is defined over the pixel lattice, whose vertices are randomly shifted in the vertical and horizontal directions according to a uniform distribution (note that the maximum amount of shift is set equal to half the size of the grid in each direction). The resulting vertices correspond to the supernumerary initial set of seeds S_0.

Following the goal-driven philosophy of the method, the seeds in such an initial set are filtered according to the predicate. Specifically, the predicates of H referring to the seed-selection phase are verified for each seed $s \in S_0$, and s is retained if and only if they are `true`. In particular, the predicate h_g prevents the selection of an isolated noisy pixel as one of the seed points; it is applied to $\tilde{X}_s$, with s being a candidate seed point and $\tilde{X}_s$ being a small region around s identified via a windowing operator. The choice of the window size does not critically affect the method. Nevertheless, for the experiments in this paper, it is set equal to 5×5. In contrast, h_p may force the seeds to be placed in a specific region of interest (e.g., to discard seeds randomly placed on the border). In the following, the subset $S \subset S_0$ of the original set of seeds that are selected according to P is denoted as $S = \{s_n\}, n = 1, 2, \ldots, N$.

The spacing of the grid can be considered a hyperparameter depending on the size of the object of interest. On the one hand, a small spacing ensures that a seed is placed in each object but yields many computations. On the other hand, a large spacing may lead to neglecting some objects within the region. Since the seed selection process deletes redundant seeds, small spacings are generally preferable to large spacings. In the experimental phase, based on the size of the ROIs, such a hyperparameter was assumed to be the 10% of the image width. Small changes in the spacing did not change the final output from a visual point of view.

Then, the propagation algorithm in [31] is used to compute a set of cost functions $F = \{F_n\}$, with n identifying each of the seeds $s_n \in S$. Each cost function is defined over the entire image such that $F_n : I \rightarrow \mathbb{R}$ (Phase 1). In contrast to the graph-based methods mentioned in the introduction [18,20–23], no graph-cut is actually performed at this level. Hence, no segmentation is obtained in the first phase. For additional details, the reader is referred to Section 2.3.

On the contrary, a set of samples $X_n \subset X$, defining a region associated with the particular seed s_n, is extracted from the image based on each cost function $F_n, n = 1, 2, \ldots, N$. They can be considered fuzzy syntactic primitives, whose properties depend on the application domain and that influence the homogeneity predicate. The complementary region (i.e. the set of image pixels not assigned to any set of samples X_n) is defined as the initial background $B_0 = X \setminus \{\cup_{n=1}^{N} X_n\}$. Moreover, according to the graph-based processing, it is not forbidden for the same pixel to be assigned to multiple sets (e.g., when two random seeds are close and inside the same region). Nevertheless, the Markovian framework characterizing the further processing steps is aimed at taking care of such a situation.

Finally, a set of parametric models, each associated with each seed $s_n \in S$, is estimated using the samples collected in the regions X_n (Phase 2), as described in detail in Section 2.4.

The segmentation problem is then formulated in a Bayesian probabilistic graphical framework by defining a Markov random field (MRF) model. Accordingly, the maximum a posteriori (MAP) decision rule is formulated as the minimization of a suitable energy function [26]. The energy function is composed of two terms, a unary term $D_i(\cdot)$ and a pairwise term $V(\cdot)$ (see Eq. (5)). The unary pixelwise contribution is related to: i) the parametric model associated with the set of samples corresponding to each seed; ii) a background-specific potential associated with the set of samples in the initial background B_0; and iii) a spatial feature related to the distance, in the image lattice, between the location of a given pixel and the location of each seed. Conversely, the pairwise contribution brings about local-contextual information. Indeed, the energy function is defined as:

$$U(L|X) = \sum_{i \in I} D_i(x_i|l_i) + \beta \sum_{i \sim j} V(l_i, l_j), \tag{5}$$

[1] The Python code of the proposed method along with a demo test is available at https://github.com/DavidSolarna/Goal-Driven-Segmentation.

where $i \sim j$ denotes pixels i and j being neighbors according to a first-order neighborhood system, $i, j \in I$ (Phase 3).

These final steps of the processing pipeline are again driven by the application goal through the predicate h_s. It is directly related to the chosen model (i.e., the pdf density function used to model the set of samples $X_n \subset X$ extracted via the graph-based computation). In contrast to the selection of the seeds, the choice of the pdf does not directly imply the elimination of some candidate ROIs in the proposed method. Nevertheless, the samples that poorly fit the chosen model will provide a negligible contribution to the energy function of the Markovian formulation of the problem, influencing the removal of the candidate ROI from the final map.

Because of the Bayesian formulation of the problem, it is possible that the final label map L does not contain all the labels from the seeds list. It is worth recalling that each seed is associated with a particular label from the preliminary set of labels $\Lambda' = \{0, 1, 2, \ldots, N\}$. Indeed, this filtering behavior is actually beneficial, as it allows the proposed method to autonomously remove, in the output result, the contributions of spurious seeds (i.e. those seeds that do not correspond to well-defined regions). Additionally, such a Bayesian formulation also solves possible conflicts, such as the cases where multiple seeds are placed inside the same region. The minimization process assigns a unique label to such a region. In other words, it allows the method to automatically determine the number of most relevant regions, whereas N plays the role of an upper bound on such number. For additional details on the energy function and the energy minimization problem the reader is referred to Section 2.5.

To conclude, according to the novel definition of goal-driven segmentation in Section 2.2, the resulting final map L is composed of the extracted ROIs X_j $(j = 1, 2, \ldots, M)$ and the background B.

As a final remark, since the method is mainly composed of three steps (i.e., the graph-based processing, the estimation of the parametric models, and the Markovian modelling), the overall computational complexity can be expressed as the sequential combination of the complexities of the three parts. Anticipating that the input image is represented with a graph $G(I, E)$, with I being the vertices and E being the edges (for the details refer to Section 2.3), the computational complexity of the first step is $O(n\_seeds \times I)$, with the flood filling applied for each seed and each operation involving the need to traverse the whole graph once. Then, the computational complexity of the second processing step is $O(I)$. Indeed, the estimation of the model parameters requires the computation of the moments of the image (refer to Section 2.4). Such moments are computed in parallel while visiting all the nodes in the graph. Finally, the final step is related to the minimization of the Markovian energy function via graph-cuts and the alpha-beta swap technique. As reported in [32], in case of binary classification problems, the computational complexity of this step is $O(EI^2)$. In the multiclass case, a set of binary classification problems are solved sequentially until convergence is reached. Moreover, [33] shows that the complexity scale linearly with the number of labels.

2.3. Graph-based cost computation

The graph-based approach proposed here is an extension of the method described in [31] and is based on the computation of the set of cost functions $F = \{F_n\}$, each associated with a specific seed $s_n \in S$.

For the sake of simplicity and ease of notation, this section describes the cost computation process focusing on a specific seed, assuming that the extension to the case of multiple seeds is straightforward and only requires replicating the process for each seed. Therefore, only the index n will be reported, without specifying $n = 1, 2, \ldots, N$. In addition, the input image is assumed to be a scalar value, while the extension to vector valued images is straightforward and only requires a minor mathematical adjustment.

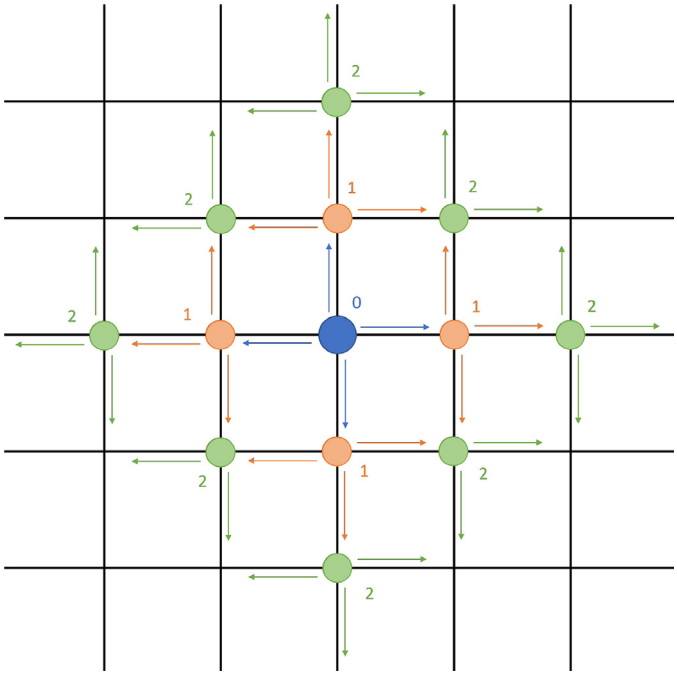


Fig. 2. Schematic representation of the flooding process starting from the seed node (blue circle) and using a 4-connected neighborhood model (for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Let the image X be mapped into a non-empty, fully-connected, undirected and vertex-weighted graph $G = (I, E)$, where I is the set of vertices representing the pixels in the image (as already specified in Section 2.2) and $E \subseteq \{(i, j) \in I \times I \mid i \neq j\}$ is the set of edges. The goal of this section is thus to describe the computation of $F_n(i), \forall i \in I$, given the seed $s_n \in S$.

For each pair of nodes i and j, let w be a difference function such that:

$$w : I \times I \to \mathbb{R}, \quad w(i, j) = |x_i - x_j|, \tag{6}$$

The function in Eq. (6) represents the difference between each pair of vertices i and j. The higher the value of $w(i, j)$, the more different the two nodes are in terms of gray level.

By fixing a specific node s_n in the graph (i.e., the seed), each point in the image, and thus each node in the graph, can be associated with a value representing its dissimilarity with respect to s_n. Indeed, it is possible to compute the difference function with respect to the seed according to $w_n(i) = w(i, s_n)$ $\forall i \in I$. The proposed cost computation process is based on the computation of $w_n(i)$ for each node in the graph and according to a flooding scheme [18].

At the beginning, all of the nodes are in the *unvisited* status except for the seed, whose cost is zero. A set T, initially containing only the seed node, is defined to keep track of all the visited nodes. Then, the flooding process is started from the seed and toward the connected neighbors (a 4-connected neighborhood model is chosen). Among the neighboring nodes, the one providing the minimum value of w_n is chosen and added to T. Consequently, the value of the cost function associated with the considered node is computed. Then, the flooding process is iterated until the set T of visited nodes coincides with the entire set of nodes I (see Fig. 2).

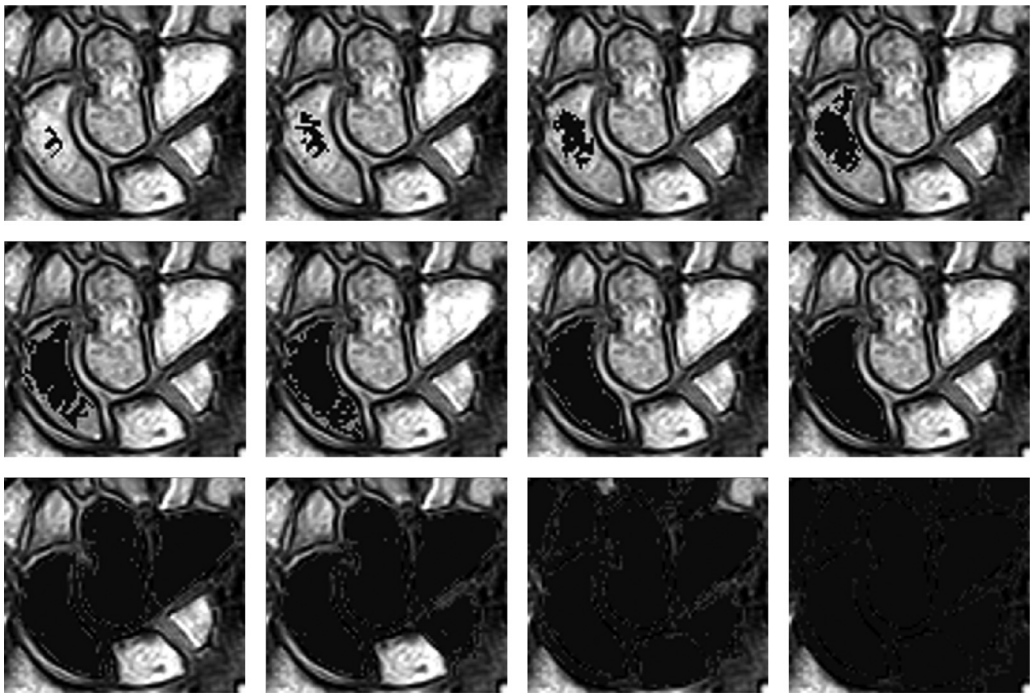


Fig. 3. Assignment of the cost values to the nodes in the graph. The figures from left to right and from top to bottom show the assignment of such costs following the order sorting of the pixels.

Concerning the computation of the cost function, at each iteration two nodes are taken into consideration: i) the node i being considered; and ii) its father node i^-, which is the node that, according to the flooding scheme, led to i. In particular, if the weight $w_n(i)$ of i with respect to the seed is larger than the cost of i^-, then the cost associated with i is set equal to $w_n(i)$; otherwise, the cost value of $w_n(i^-)$ is inherited (see Eq. (7)). In this way, the cost function never decreases with the flooding scheme:

$$F_n(i) = \begin{cases} w_n(i) & \text{if } w_n(i) > F_n(i^-), \\ F_n(i^-) & \text{if } w_n(i) \leq F_n(i^-). \end{cases} \tag{7}$$

Leveraging such a propagation algorithm, a cost value is assigned to all nodes and the process defines an optimal path from the seed to each vertex. To better understand the process, Fig. 3 shows the assignment of these cost values to all the nodes of the graph built on top of a magnetic resonance image. Looking at the images from left to right and from top to bottom, it is possible to appreciate the pixel sorting order during the process of cost value assignment, starting from a given seed point. As an example, a limited number of steps of the iterative process are depicted. At each step, the nodes already assigned with a cost value are painted black; moving away from the seed, the evaluated cost values are always increasing. Such a procedure yields the set of region samples X_n associated with the nth seed. In particular, the cost function is thresholded according to the method in [31]. The set of pixels associated with costs lower than such a threshold are assigned to the set X_n. The resulting region is not disjoint due to the non-decreasing formulation of the cost function F_n.

As mentioned at the beginning of this section, repeating the process for each seed s_n yields the set of region samples $\{X_n\}, n \in \{1, 2, \ldots, N\}$. Indeed, due to the definition of the graph and the cost function in Eq. (7), such regions are characterized by homogeneity and low granularity. Consistent with the discussion in Section 2.2, the set of regions $\{X_n\}, n \in \{1, 2, \ldots, N\}$ obtained from the present phase of the method generally does not cover the entire image, i.e., it leaves out the initial background defined as B_0.

2.4. Parametric model estimation

This section presents the different model estimation strategies adopted in the proposed method for the observations in each set of samples $X_n,\ n \in \{1, 2, \ldots, N\}$. Here, the focus is on the radiometric information contained in $\mathcal{X}$, and not on the entire image X conveniently defined as observations and pixel locations. Therefore, this section will refer to the set $\mathcal{X}_n$, which is the collection of the pixel intensities contained in the region X_n.

Similar to the format adopted in Section 2.3, the estimation strategies are described here without considering the various regions $\mathcal{X}_n \subset \mathcal{X}$, but taking into consideration a generic random variable Z and a set $\mathcal{Z} = \{z_i\}_{i=1}^q$ of independent and identically distributed (i.i.d.) samples drawn from Z. In the proposed method, the approach described here is separately applied to each region obtained in the previous phase from the nth seed by using $\mathcal{Z} = \mathcal{X}_n,\ n \in \{1, 2, \ldots, N\}$.

Due to the possibility of applying the proposed method in different domains, depending on the user-defined goal, different parametric models may be embodied in the technique based on prior information on the characteristics of the considered image data. In general, any parametric model for which a corresponding parameter estimation approach is available can be integrated in the proposed approach. In the following, two specific models are considered that encompass a significant number of image processing applications.

First, leveraging the fact that the proposed methodology is aimed at isolating homogeneous and low-granularity regions, the case of a Gaussian model was taken into consideration. Such a model is known to be well suited for modeling, as examples, images collected by passive cameras, for which additive Gaussian noise is usually an appropriate model, as well as MRI instruments [34,35]. In this case, the estimation of the underlying parameters is addressed, in a maximum-likelihood fashion, via the sample mean and the sample variance.

Second, the case of images affected by a multiplicative noise-like component is addressed. This is a situation that intrinsically

occurs when active radar or sonar instruments are used because of the speckle phenomenon. A variety of parametric models have been explored in the literature, including Gamma, Weibull, log-normal, K, symmetric α-stable, $\mathcal{G}_0$, generalized Gaussian-Rayleigh, Fisher, and generalized Gamma distributions [36,37].

It is worth noting that the models characterized by more than two parameters (Fisher, K, generalized Gamma, etc.) have been mostly designed for modeling non-homogeneous regions, where texture or granularity phenomena are highly visible. Here, the Gamma distribution has been considered, leveraging again the homogeneity and low granularity of the regions identified by the previous phase of the proposed methodology. Such a distribution is a well-known model for the statistics of homogeneous non-textured regions of image data affected by multiplicative speckle, such as in radar imagery [28], sonar systems [38], and ultrasound imagery [39], and its PDF is supported on $(0, +\infty)$:

$$p_Z(z) = \frac{1}{\Gamma(\alpha)}\left(\frac{\alpha}{u}\right)^{\alpha} z^{\alpha-1}\exp\left(-\frac{\alpha z}{u}\right), \tag{8}$$

where u is the mean of the distribution and α is its shape parameter ($\alpha, u > 0$).

Parameter estimation is performed via the method of log-cumulants (MoLC) [40,41], which is a technique developed in the SAR literature for distributions of positive-valued data characterized by multiplicative components (such as the aforementioned speckle) [42–44]. MoLC estimation follows a principle similar to the well-known method of moments (MoM), which makes use of the Laplace transform (in terms of moment generating function) to define relations between the moments and the parameters of the underlying distribution. In the case of MoLC, the Mellin transform is used to relate the parameters of the distribution to its logarithmic cumulants [43]. In the application to several of the aforementioned distributions, MoLC has led to advantages over MoM in terms of lower estimation variance and over maximum likelihood in terms of easier analytical or numerical solution [40].

In particular, if Z is a positive-valued random variable, the first- and second-order log-cumulants κ_1 and κ_2 can be proven equivalent to the logarithmic mean and variance of the distribution of Z:

$$\begin{aligned}\kappa_1 &= \mathbb{E}\{\ln Z\},\\ \kappa_2 &= \mathbb{E}\{[\ln Z - \kappa_1]^2\}.\end{aligned} \tag{9}$$

Leveraging the use of the Mellin transforms and the definition of the Gamma distribution (see Eq. (8)), these logarithmic cumulants can be related to their parameters $[u, \alpha]$ through the following MoLC equations [45]:

$$\begin{cases}\kappa_1 = \Psi(0,\alpha) - \ln u - \ln\alpha,\\ \kappa_2 = \Psi(1,\alpha),\end{cases} \tag{10}$$

where $\Psi(0,\alpha)$ is the digamma function (i.e., the logarithmic derivative of the Gamma function) and $\Psi(1,\cdot)$ is the first order polygamma function (i.e., the derivative of $\Psi(0,\alpha)$) [46].

Given the i.i.d. samples in $\mathcal{Z}$, the sample estimates of κ_1 and κ_2 can be obtained as:

$$\begin{aligned}\hat{\kappa}_1 &= \tfrac{1}{q}\sum_{i=1}^{q}\ln z_i,\\ \hat{\kappa}_2 &= \tfrac{1}{q}\sum_{i=1}^{q}[\ln z_i - \hat{\kappa}_1]^2.\end{aligned} \tag{11}$$

Then, substituting the sample estimates into Eq. (10), it is possible to write the MoLC equations for the Gamma distribution incorporating the MoLC estimated parameters $\hat{u}$ and $\hat{\alpha}$ to $\hat{\kappa}_1$ and $\hat{\kappa}_2$:

$$\begin{cases}\hat{\kappa}_1 = \Psi(0,\hat{\alpha}) - \ln\hat{u} - \ln\hat{\alpha},\\ \hat{\kappa}_2 = \Psi(1,\hat{\alpha}),\end{cases} \tag{12}$$

The parameter $\hat{\alpha}$ is first obtained by numerically solving the second equation via the Newton-Raphson method [47], and then $\hat{u}$ is retrieved from the first equation by substituting $\hat{\alpha}$. The former solution step is numerically simple thanks to the strict monotonicity of $\Psi(1,\cdot)$.

2.5. Energy minimization through graph cut

In the proposed method, the energy function in Eq. (5) is composed of a unary term and a pairwise term. The unary pixel-wise term $D_i(x_i|l_i)$ is defined differently in case it refers to the seed-specific labels (i.e., $l_i = 1,\dots,N$) or to the background B_0 (i.e., $l_i = 0$). In the former case, it is related to the seed-specific parametric models and to a term measuring the distance between the considered pixel x_i and the seeds. Conversely, in the latter case, it is related to a piecewise constant background-specific potential. The pairwise potential is defined as the Potts model [26]. Indeed, the two terms are as follows:

$$\begin{aligned}D_i(x|n) &= \begin{cases} d_i(x|n) & n = 1,2,\dots,N\\ c_1 & n = 0, i \in B_0\\ c_2 & n = 0, i \in I - B_0\end{cases}\\ V(n,m) &= 1 - \delta(n,m),\end{aligned} \tag{13}$$

where $\delta(v,w)$ represents the Kronecker delta, whose value is 1 if and only if v equals w, and zero otherwise, $d_i(x|n)$ is the seed-specific potential, and c_1 and c_2 are the values of the piecewise constant background-specific potential.

Concerning the seed-specific potential, let $\hat{p}(\cdot|n)$ be the PDF estimate obtained as described in the previous section from the set of samples X_n, $n \in \{1,2,\dots,N\}$ and according to either a Gaussian or a Gamma model. Then, the potential can be written as:

$$d_i(x|n) = -\ln\hat{p}(x|n) - \gamma[\Delta(i,s_n)]^{-1}, \tag{14}$$

where $\Delta(a,b)$ is the Euclidean distance between pixel locations a and b in the image plane ($a, b \in I$), and γ is a positive coefficient balancing the two contributions.

Concerning the piecewise-constant background potential, the two values c_1 and c_2 are automatically chosen according to the pth and $(1-p)$th percentiles of the distribution of the seed-specific potentials $d_i(x|n)$. The rationale is to balance the background-specific and the seed-specific potentials to favor the goal-driven result while also not censoring any possible outcome in the output label map Y. Both the parameters γ (Eq. (14)) and β (Eq. (5)) are determined via a trial-and-error procedure.

In particular, the term $\Delta(i,s_{l_i})$ represents the spatial distance between the location of the pixel $i \in I$ in the image lattice and the location of the seed s_{l_i} that corresponds to its label l_i. While the first component of the unary potential is traditionally related to the likelihood of the data, the integration of the second contribution in the unary term is inherited from a family of segmentation methods that explicitly integrate a spatial feature in the formulation of the unary term [48]. To the best of the authors' knowledge, the integration between the MRF model and the seed location is proposed for the first time in the present manuscript. This integration is also responsible for linking the intermediate statistical procedure with the initial topological cost computation and the final cut yielding the results.

The minimization of a Markovian energy function, such as the one defined in Eq. (5), has been dealt with using several techniques in the last few decades. A major difference between such techniques resides in the trade-off between computational complexity and optimality of the solution, with deterministic methods, such as iterated conditional mode (ICM) [49], requiring a short convergence time but reaching only local minima, and stochastic methods, such as simulated annealing [25], providing more effective results but requiring a very long time to converge to global minima.

Here, the graph cut approach [33] is used, which is based on the reformulation of the energy minimization problem as a max-flow/min-cut problem over a suitable graph. In the case of binary labeling, graph cut approaches are also proven to converge in polynomial time to the global minimum. In the case of more than two labels, ad hoc techniques have been formulated. Here, the chosen technique is the alpha-beta swap technique, which reformulates the problem as a sequence of binary subproblems and, for each subproblem, a global energy minimum is reached through the max-flow/min-cut formulation. Convergence to a local minimum with strong optimality properties is guaranteed in this case [33].

3. Experimental results

This section presents and discusses the results obtained by the proposed method in three real case scenarios. The first is related to the biomedical field and, more specifically, the segmentation of the wrist bones in an MR image. The other cases deal with SAR data and are related to the segmentation of rural areas and an oil slick in the Mediterranean Sea. According to the above considerations and to Section 2.4, in the former case, the adopted parametric model is the Gaussian distribution, while in the other cases, the adopted parametric model is the Gamma distribution. In the case of the experiments with SoA methods, the SAR data has been preprocessed with the SRAD filter [50] in order to reduce the impact of speckle noise. The declared goal in the former case is to detect the wrist bones, which appear as bright and compact regions. Conversely, in the case of SAR images, the goal is to segment dark and bright agricultural fields and the dark oil spilled in the sea.

The experimental analysis, together with the selected accuracy figures, is designed to address the capability of the proposed method to partition the image in homogeneous regions. Moreover, to quantitatively assess the achieved accuracy, the segmentation results are also compared with ground truth data related to specific objects of interest in the scene (e.g., the wrist bones and the agricultural fields or the oil slick).

Finally, for the sake of completeness, all the experiments have been run on a desktop computer equipped with an Intel® Core™ i7-4790 CPU, working at a frequency of 3.60GHz, and 24GB of RAM memory. Also, the programming language used in all the experiments is Python.

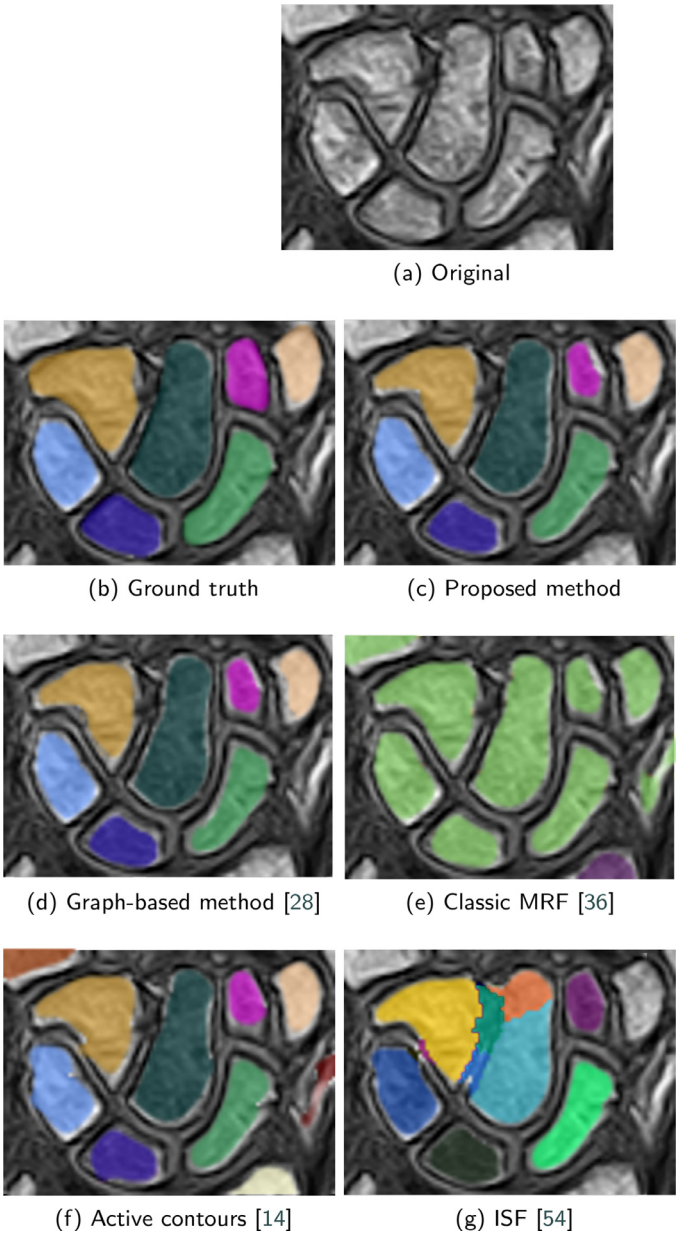
3.1. Magnetic resonance image

The proposed method was tested on ten MR T1-weighted images from the database in [2]. They are images of the wrist district depicting seven carpal bones (i.e. capitate, hamate, lunate, scaphoid, trapezium, trapezoid, and triquetrum), in the coronal plane. The acquisition was performed via the 0.2 Tesla Artoscan (Esaote Spa, Genova, Italy). The dimension of each image is 80×100 pixels, the maximum gray-level is 255, and the ground truth segmentation was performed by medical operators with extensive experience in rheumatoid artritis evaluations. The goal to be achieved is the detection of bright regions corresponding to the trabecular part of the wrist bones.

Specifically:

$$h_g(X_n) = \text{true} \iff \mathbb{E}\{x_i\} \in (m_{MR}, M_{MR}) \wedge \text{Var}\{x_i\} \in (0, v_{MR}) \quad \forall (i, x_i) \in \tilde{X}_{s_n}, \tag{15}$$

with $\mathbb{E}\{\cdot\}$ and $\text{Var}\{\cdot\}$ being the sample mean and sample variance, and m_{MR}, M_{MR}, and v_{MR} being thresholds that can be set either directly by the user (e.g., the expert selecting those values based on prior knowledge) or automatically, e.g., in a supervised fashion. Here, the former strategy has been adopted, with the values assigned as $m_{MR} = 100, M_{MR} = 255, v_{MR} = 25$.



(a) Original

(b) Ground truth

(c) Proposed method

(d) Graph-based method [28]

(e) Classic MRF [36]

(f) Active contours [14]

(g) ISF [54]

Fig. 4. Carpal bones in MR: 4a original data; 4b ground truth [2]; 4c results obtained by the proposed method; 4d results obtained by the baseline graph-based method in [31]; 4f results obtained by the classical formulation of MRF models; 4e results obtained by the method in [51]; 4g results obtained by the method in [23] (for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Then, the pdf adopted for h_s is a Gaussian distribution. Finally

$$h_p(X_n) = \texttt{true} \iff i \in \bar{I}, \quad \forall (i, x_i) \in X_n, \tag{16}$$

where $\bar{I} \subset I$ indicates the central area of the imaged scene where the target ROIs are supposed to be located. Here, it corresponds to the imaged area I without the border whose size was chosen as the 10% of the smallest image size. Finally, the weight coefficients related to the Markovian modeling have been set equal to $\beta = 5, \gamma = 1$.

Fig. 4 shows one of the original images, the corresponding ground truth, and the segmentation maps obtained by the proposed method and by four other state-of-the-art techniques.

Based on the ground truth in [2], the results are compared with those obtained by the methods in [51], in [31], in [23], and with a

Table 1
DICE score comparison. MRI experiment involving the image reported in Fig. 4.

	Proposed Method	Graph-based method [31]	Classic MRF [26]	Active contours [51]	ISF [23]
Capitate	**0.903 ±0.032**	0.826 ± 0.043	0.877 ± 0.036	0.878 ± 0.045	0.830 ± 0.078
Hamate	**0.875 ±0.037**	0.785 ± 0.049	0.843 ± 0.078	0.874 ± 0.046	0.850 ± 0.049
Lunate	**0.850 ±0.050**	0.765 ± 0.061	0.819 ± 0.039	0.826 ± 0.057	0.737 ± 0.117
Scaphoid	**0.877 ±0.027**	0.789 ± 0.030	0.842 ± 0.034	0.857 ± 0.034	0.805 ± 0.084
Trapezium	**0.880 ±0.037**	0.832 ± 0.086	0.735 ± 0.174	0.854 ± 0.068	0.482 ± 0.289
Trapezoid	0.799 ± 0.086	**0.813 ±0.069**	0.740 ± 0.067	0.792 ± 0.062	0.614 ± 0.295
Triquetrum	**0.911 ±0.031**	0.695 ± 0.038	0.864 ± 0.031	0.900 ± 0.037	0.665 ± 0.296
Total	**0.871**	0.786	0.817	0.854	0.712

Table 2
Performance comparison on the three available SAR dataset. For each method, the number of identified fields and the average dice score computed on such fields are reported. The three dataset correspond to the images shown in Fig. 5, 6, and 7.

Dataset		Proposed	Graph-based [31]	Classic MRF [26]	Active contours [51]	ISF [23]	RSLC [56]
Case 1	Fields	9	6	3	4	6	9
	DICE	**0.910**	0.548	0.816	0.600	0.501	0.785
Case 2	Fields	14	7	4	3	5	14
	DICE	**0.879**	0.424	0.421	0.320	0.517	0.734
Case 3	Fields	6	6	6	5	4	6
	DICE	**0.787**	0.203	0.663	0.340	0.357	0.782

classical MRF model (i.e., without the term $[\Delta(i, s_{l_i})]^{-1}$ in Eq. (14)), taking into account the dice score as a performance indicator [52].

The dice score is one of the most frequently used metrics in the evaluation of image segmentation results [53,54]. Let S_{GT} be the ground truth segmentation and S_{res} be the output of a segmentation method, then the dice score is computed as:

$$DSC = \frac{2(S_{GT} \cup S_{res})}{|S_{GT}| + |S_{res}|} \tag{17}$$

where $|\cdot|$ computes the number of pixels in the segmentation. Therefore, the dice score is twice the area of the overlap between S_{GT} and S_{res} divided by the total number of pixels in the two segmentation. The range of values is between 0 and 1, with the former indicating the worst-case scenario (no overlap), and the latter indicating a perfect overlap between the obtained segmentation and the ground truth.

The dice score is highly correlated with another frequently used metric, the Jaccard index [55], whose definition and computation are very similar to the dice score. Indeed, the Jaccard index is computed as the intersection of S_{GT} and S_{res} over their union. Due to their similarity and based on the result of the research, only the dice score is reported in the experimental analysis. Furthermore, it is more intuitive to interpret, as it can be seen as the percentage of overlap between the two sets.

As mentioned in Section 2.3, the method in [31] is the baseline for the current work. Both the graph-based approach and the propagation mechanism are present, but no Bayesian formulation is included. As a consequence, each ROI is extracted one at a time, or an optimal criterion for seed placement has to be defined. In particular, the segmented bones are smaller than those obtained via the proposed method; the false negative pixels are reduced through the developed approach.

The other starting point for the proposed method is an MRF model in its classical formulation, i.e., without the distance term, which is present in Eq. (13). Here, not only the carpal bones but also other regions are extracted. Furthermore, one can notice how the distance term in Eq. (13) allows labeling each bone with a different label, meaning that each ROI is represented by a different seed.

Then, the method in [51] is an example of a deformable model. It consists of a flexible active contour model that makes use of thresholding or gradient-based methods. Even though it is widely used in the medical field, in this case, the shapes of the segmented ROIs are less regular and compact than those obtained by the proposed method. Moreover, the method in [51] also provides a significant number of false positive pixels, resulting in a set of small and spurious ROIs.

Finally, the method presented in [23] is designed for superpixel computation, and hence, it results in an over-segmentation of the image. Indeed, in many cases, the bones were oversegmented. Thus, to enable the comparison with the proposed method, the output superpixels have been manually merged. In some cases, the method in [23] was not able to detect some of the carpal bones, thus affecting the final mean scores. For a quantitative comparison, the accuracy measured by using such a method does not mirror its actual capability. Indeed, the performances are positively biased by the manual merging operations that have been performed on the output segmentation to allow the comparison.

The performance measures reported in Table 1, together with the qualitative results reported in Fig. 4, demonstrate the capability of the proposed method to provide a segmentation result characterized by regular shapes, achieving effective performances that are also consistent across the different bones, and avoiding spurious and noisy ROIs in the output map. The regularization of the MR artifacts and noise is guaranteed by the MRF model, which allows segmenting all of the bones that are present in the considered image as connected components that can be easily extracted and identified.

3.2. Synthetic aperture radar image

The second domain in which the method was tested is the case of radar remote sensing image segmentation. In particular, four SAR intensity images were considered. The first three images were acquired by COSMO-SkyMed in the HH polarization and depict rural areas in the northern Italy, while the fourth image is a TerraSAR-X image, SC mode, acquired in the Mediterranean Sea near the French coast. The COSMO-SkyMed image sizes are 361×671, 504×458, and 619×417 pixels, respectively; all of them are at a spatial resolution of 5m, while the size of the TerraSAR-X image is 1500×2700 pixels at a resolution of 18m.

Concerning the first images, the visible agricultural fields were mainly of three types, corresponding to the lower, higher, and intermediate values of the backscattering coefficient and do not exhibit appreciable texture. The proposed segmentation method was applied with the purpose of segmenting the two types of field

Table 3
DICE score comparison. SAR experiment involving the image reported in Fig. 5 (Case 1).

	Proposed Method	RSLC [56]
Field 1	**0.925**	0.862
Field 2	**0.922**	0.818
Field 3	**0.921**	0.586
Field 4	0.902	**0.921**
Field 5	**0.901**	0.894
Field 6	**0.898**	0.386
Field 7	0.903	**0.944**
Field 8	**0.913**	0.837
Field 9	**0.906**	0.819
Total	**0.910**	0.785

corresponding to brighter and darker areas (i.e., higher and lower backscattering). In this case, two kinds of h_g conditions were deployed to consider the seeds in both the bright and dark regions.

In particular:

$$h_g(X_n) = \text{true} \iff \mathbb{E}\{x_i\} \in (m_{SAR}, M_{SAR}) \wedge \text{Var}\{x_i\} \in (0, v_{SAR}) \quad \forall (i, x_i) \in \widetilde{X}_{s_n}, \tag{18}$$

where, analogously to the medical application, m_{SAR}, M_{SAR}, and v_{SAR} are thresholds that can be set either directly by the user or automatically. Here, the former strategy has been adopted. In the case of bright fields, the parameters have been chosen according to $m_{SAR} = 140, M_{SAR} = 220, v_{SAR} = 35$. Conversely, in case of dark fields, the parameters have been chosen according to $m_{SAR} = 0, M_{SAR} = 120, v_{SAR} = 25$.

Then, the pdf adopted for h_s is a Gamma distribution, while the MRF-related parameters have been set equal to the previous case.

The proposed method was compared with a state-of-the-art approach for SAR image segmentation [56], which will be hereinafter referred to as RSLC (i.e., region smoothing and label correction), and with the previously introduced segmentation methods. However, the results of such methods are not as effective as the proposed and the RSLC methods, despite the SRAD-based preprocessing step aimed at reducing the impact of the speckle noise. Moreover, the resulting segmentation maps did not include all the fields, with some of them being fused into regions associated to the background, and some being merged into a single region. The reason is twofold, with the residual speckle noise on one side and the not well-defined borders in the SAR data with respect to the MRI images above. Nevertheless, a performance indicator has been computed also in the case of such methods by averaging the dice scores obtained on the fields that were correctly identified and not considering the ones that were merged or not detected at all. Table 2 reports a summary of such analysis and indicates, for each method, the number of fields being identified and the average dice score on such fields. In the following, the thorough analysis of the performances, with details on the single fields and on the resulting segmentation maps, will take into consideration the best performing methods, being the proposed one and the RSLC.

Figs. 5, 6, and 7 show the original images, the available and manually annotated ground truth data, and the segmentation results obtained via the two aforementioned cases in the considered trials. In addition, Tables 3, 4, and 5 summarize the performance of the methods by showing the dice score. It is worth mentioning that when referring to RSLC, the ROIs included in the ground truth have been manually selected in the final segmentation map (i.e., by merging oversegmented areas) to enable a quantitative comparison. This emphasizes once more the novelty of the proposed method to provide a goal-driven result, which is not possible by the compared solution.

Similar to the case of the MR images, the use of the MRF model allowed to achieve spatial regularization in the segmentation map.

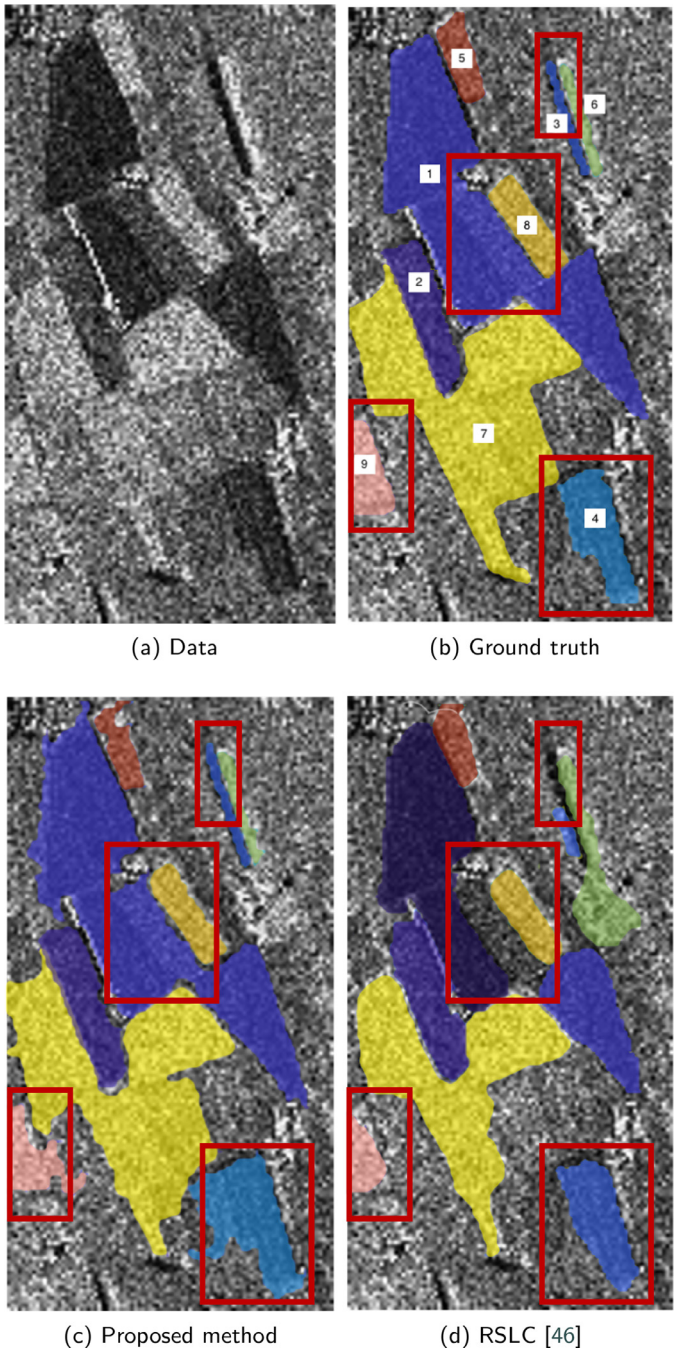


Fig. 5. Agricultural field segmentation in SAR imagery, Case 1. Original image, ground truth with corresponding enumeration, the result obtained via the proposed method, and via the RSLC method in [56] (for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Indeed, the polygonal shapes of the fields are preserved, despite the speckle [28] that may lead to errors in identifying the borders.

As described in Section 2.4, the adopted parametric model for SAR imagery is the Gamma distribution. Fig. 8 depicts the gray level histogram of one of the ROIs in the COSMO-SkyMed dataset and the corresponding MoLC-fitted Gamma distribution. On the one hand, the figure shows the accuracy of the MoLC estimation for the parameters of the Gamma distribution, and on the other hand, it also visually confirms the Gamma distribution to be a valid model for the statistics of the pixel intensities in the regions detected by the proposed method when applied to this type of image.

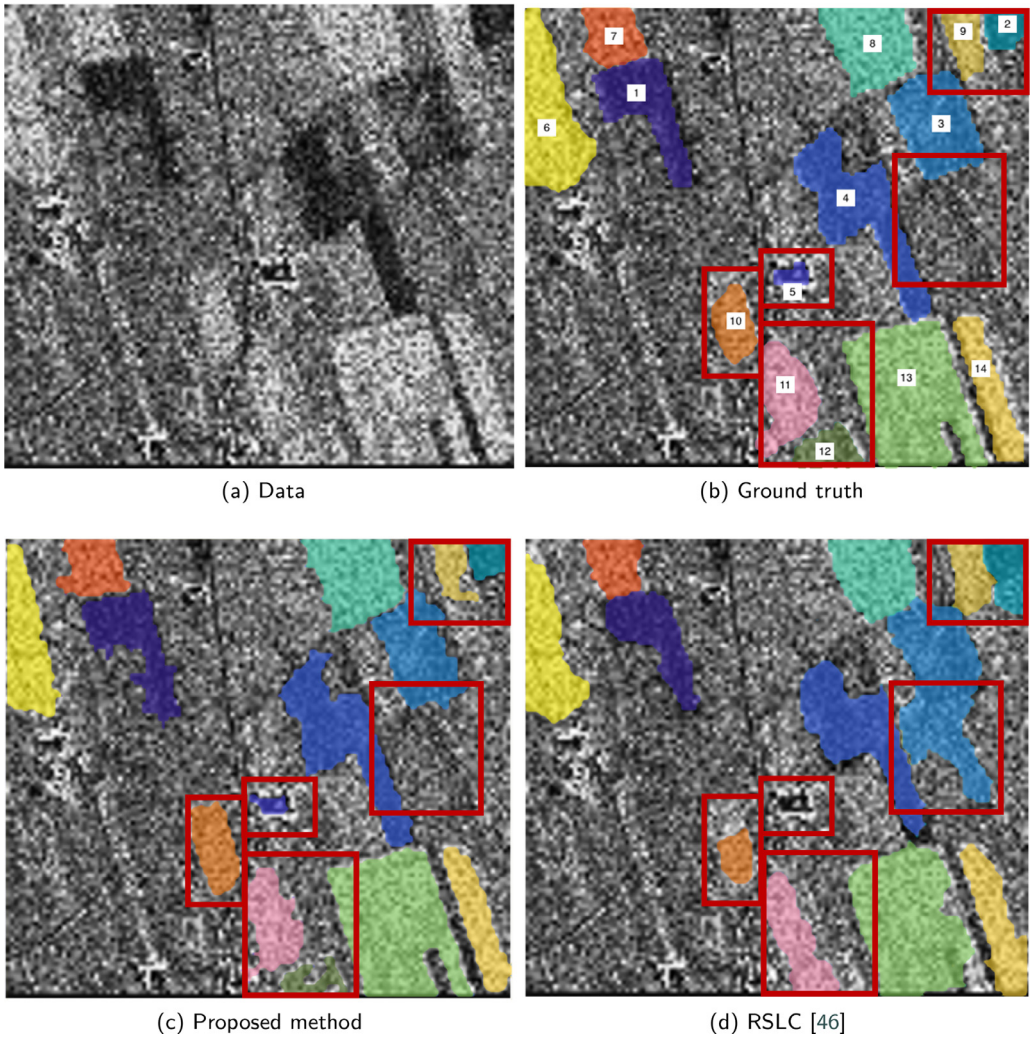


Fig. 6. Agricultural field segmentation in SAR imagery, Case 2. Original image, ground truth with corresponding enumeration, the result obtained via the proposed method, and via the RSLC method in [56] (for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Table 4
DICE score comparison. SAR experiment involving the image reported in Fig. 6 (Case 2). *When referring to the RSLC method in [56], fields 11 and 12 are merged.

	Proposed Method	RSLC [56]
Field 1	0.881	**0.901**
Field 2	**0.931**	0.712
Field 3	**0.893**	0.674
Field 4	**0.928**	0.924
Field 5	**0.757**	0
Field 6	0.840	**0.869**
Field 7	**0.923**	0.907
Field 8	**0.952**	0.940
Field 9	0.819	**0.866**
Field 10	**0.852**	0.742
Field 11*	**0.871**	0.645
Field 12*	**0.765**	0.335
Field 13	**0.969**	0.907
Field 14	**0.923**	0.853
Total	**0.879**	0.734

Table 5
DICE score comparison. SAR experiment involving the image reported in Fig. 7 (Case 3).

	Proposed Method	RSLC [56]
Field 1	**0.916**	0.890
Field 2	**0.892**	0.865
Field 3	0.576	**0.715**
Field 4	**0.887**	0.787
Field 5	**0.825**	0.814
Field 6	**0.629**	0.622
Total	**0.787**	0.782

The same comment holds with regard to the statistics conditioned on other regions in the image.

Fig. 9 shows an oil slick located in the Mediterranean Sea, along with the segmentation results. The goal is to detect the oil spill, which appears as a dark region over a bright background, to enable the quantification of its spread. In contrast to the previous case, no ground truth was available in this case study. However, qualitative results are provided in order to give a visual feedback on the performances achieved by the proposed method in this new scenario. Indeed, even though the comparison is not quantitative, it gives an intuition on the behaviors of the two solutions and allows to choose the best option with respect to the application at hand.

The goal specifications are:

$$h_g(X_n) = \text{true} \iff \mathbb{E}\{x_i\} \in (m_{Oil}, M_{Oil}) \wedge \text{Var}\{x_i\} \in (0, v_{Oil}) \quad \forall (i, x_i) \in \tilde{X}_{s_n}, \tag{19}$$

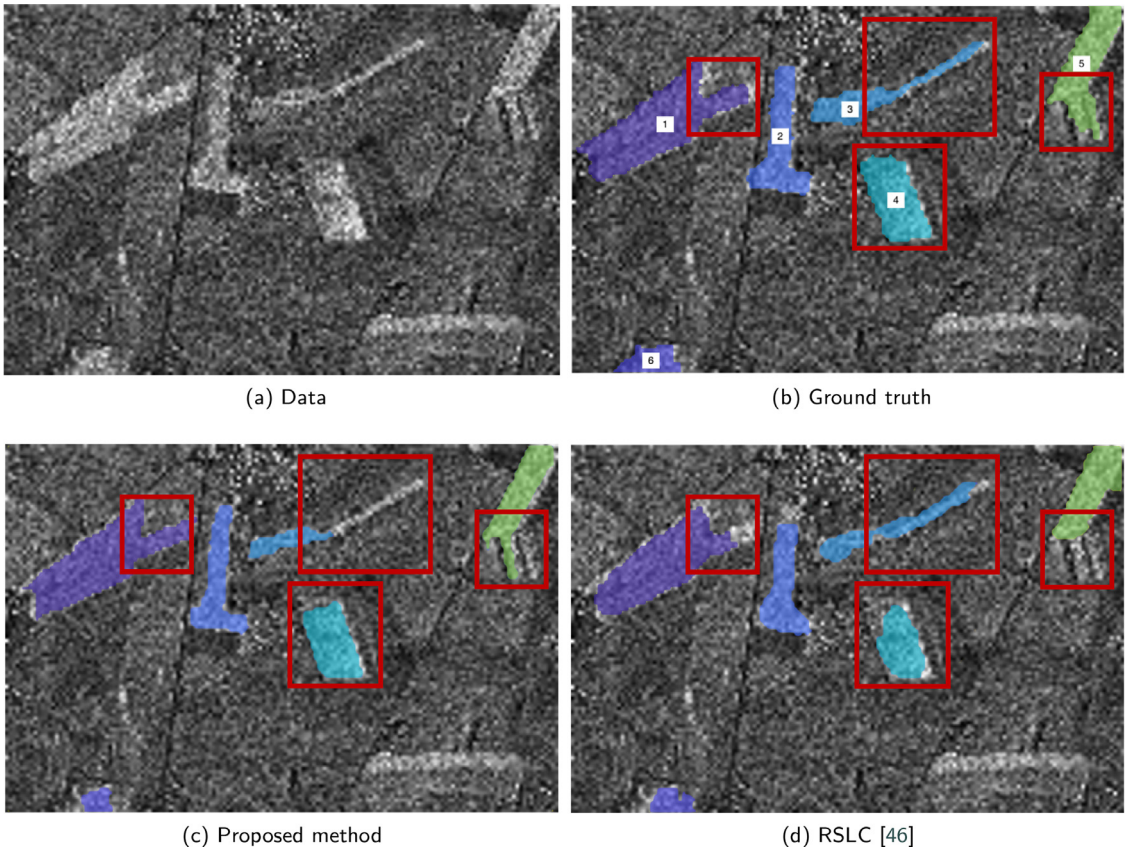


Fig. 7. Agricultural field segmentation in SAR imagery, Case 3. Original image, ground truth with corresponding enumeration, the result obtained via the proposed method, and via the RSLC method in [56] (for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

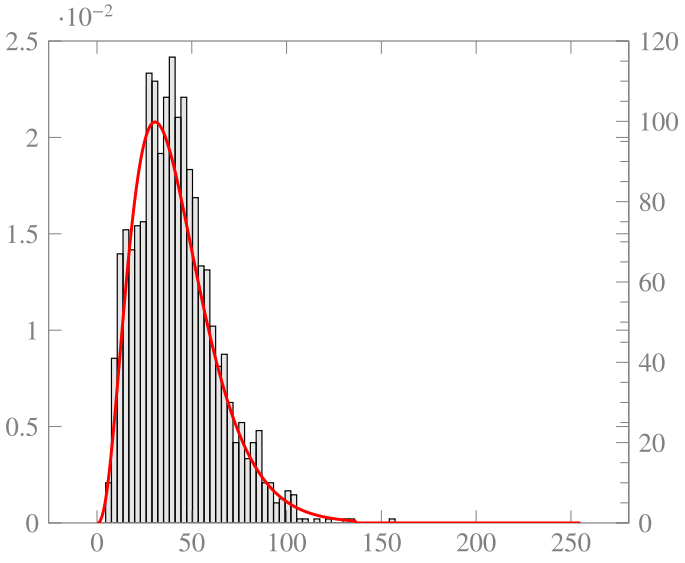


Fig. 8. Gray level histogram of one of the ROIs in the SAR image depicting the agricultural fields and the MoLC-fitted gamma distribution.

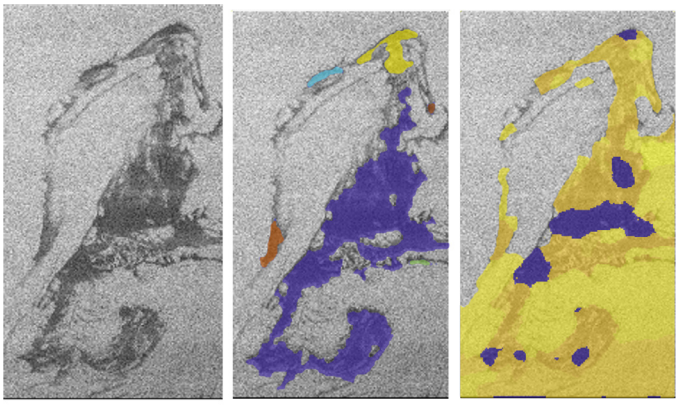


Fig. 9. Oil spill segmentation in SAR images: original image, segmented ROIs obtained by the proposed method, and the result obtained via the RSLC method [56] (for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

where m_{Oil}, M_{Oil}, and ν_{Oil} are set in the same fashion as in the dark field segmentation task (i.e., $m_{Oil} = 0, M_{Oil} = 120, \nu_{Oil} = 25$; the pdf adopted for h_s is a Gamma distribution, and the parameters inside the Markovian energy function are set equal to the previous cases.

Also in this case, the segmentation method was able to achieve satisfactory results. In particular, the segmentation was able to separate the homogeneous regions characterized by the oil spill from the surrounding area, yet the fine details of the oil spill in the upper left part of the images were not identified. This is due to the random grid strategy in the seed placement step. Due to the thin shape of such details, none of the seeds was placed inside the region of interest, and as a consequence, such fine details were not assigned to the homogeneous region associated with the oil spill. The reader may also notice the difference with respect to the output of method [56], where no particular ROI is actually visible.

In addition, the method was also compared with two techniques based on deep learning that are presented in [16]. The

Fig. 10. Visual comparison with the supervised methods reported in [16].

Table 6
Results from the comparison with respect to the methods reported in [16].

	Black Fields	White Fields	Kappa
Proposed Method	**0.965**	**0.998**	0.929
CWNN [16]	0.905	0.955	**0.935**
CNN [16]	0.853	0.960	0.895

data used for such experiments are provided by the authors (i.e., the publicly released "Noerdlinger Ries" dataset), as well as the adopted performance metrics, i.e., the percentage of pixels that are correctly segmented with respect to the considered classes, and the Kappa coefficient, computed from the confusion matrix. Fig. 10 and Table 6 report the qualitative and quantitative results. With respect to the qualitative results, all the methods are effective in discriminating the two types of field, with the proposed method producing the least noisy maps, and the method based on a convolutional neural network and on the wavelet transform being able to better discriminate the boundaries between elements. The latter was the most effective in terms of Kappa coefficient, while the proposed method outperformed the others with respect to the percentage of correctly classified pixels. The method based on the sole CNN was the least effective with respect to the quantitative analysis. This result is indeed very promising for the proposed method,

Table 7
Computation times needed by the proposed method to generate the segmentation maps with two datasets while varying the step size of the random grid, and hence the number of seeds.

	Step	Size [w x h]	Seeds	Time [m]
MRI	4	84 x 75	67	02:21
	5		44	01:00
	6		28	00:24
	8		18	00:10
SAR	20	498 x 442	74	12:33
	25		54	05:50
	30		37	02:29
	35		30	01:45
	40		21	00:58

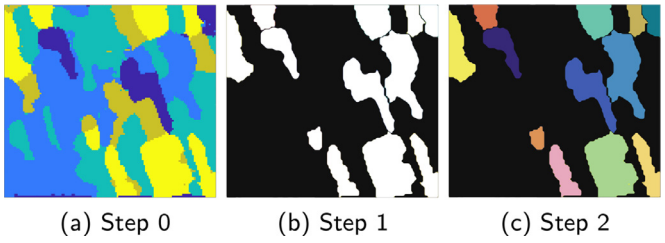


Fig. 11. Results from the steps of the post processing operations. They are required to obtain an output similar to the proposed goal-driven method (for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

as the two DL-based solutions required a labeled training set and a training procedure to produce the maps reported in Fig. 10 and evaluated in Table 6. In detail, the full "Noerdlinger Ries" dataset was split into a set of patches, 60% of which were used for training the two neural networks. Conversely, the proposed method required no training data and was directly applied to the image at hand.

Finally, Table 7 reports an analysis on the computation time needed by the proposed method to generate the segmentation map. Such an analysis involves two different datasets: (i) the MR image reported in Fig. 4; and (ii) the SAR image of Fig. 10. The analysis correlates the step size used for placing the initial set of seeds in the image, and thus the number of seeds used in the processing, with the time needed by the method to generate the maps. Moreover, it is possible to appreciate the impact of the number of seeds on the computational requirements of the method. Finally, such a comparison is also intended as a sensitivity analysis, as the experiments reported in the table resulted in accuracies comparable with those reported in the previous tables (i.e., Table 1 for the MR image and Table 5 for the SAR image). Using a larger and a smaller step size than the ones reported caused the method to reduce the accuracy due to an over-segmentation of the ROIs or due to the impossibility of identifying all the elements, respectively.

3.3. The advantage of the goal-driven formulation

Since classical segmentation methods (i.e., satisfying the first definition of Section 2.2) are aimed at partitioning the whole image into disjoint subregions, their outputs often require some post processing to enable comparisons with the one presented here. Hence, the aim of the present section is to underline the post processing steps that can be avoided by leveraging the goal-driven formulation of the presented approach. For the sake of simplicity, an example of the post processing procedure with respect to the second SAR image is provided.

Fig. 11 a depicts the output of the method presented in [56], which is denoted as Step 0 of the post processing procedure. Rather than the extraction of some particular regions, here, the final map is a segmentation map in the classical sense. Hence, the regions to be considered need to be selected and separated (even manually). For the present purpose, several operations may be required, including binarization (in favor of connected component extraction), manual correction (e.g., for properly separating objects of interest), and morphological operations (for regularizing the shape of the extracted regions). The result of the present process (Step 1) is shown in Fig. 11b. Finally, the obtained ROIs are labeled (Step 2), producing the final map in Fig. 11c.

All of these operations yield the final output, which is similar to the one produced by the novel method presented above. Furthermore, these steps are necessary to have a final map that can be evaluated with respect to the available ground truth. However, these post processing steps result in a time-consuming task that must be adapted to the specific output of the segmentation method being considered.

Conversely, the goal-driven formulation encompasses the task to be fulfilled, thus avoiding further post processing endeavors, which represents an advantage in terms of usability with respect to the compared methods. Within the proposed method, some user interaction/user expertise is still needed in the definition of the goal. However, fewer operations are required for such a phase than in the post processing described above. Moreover, the definition of the goal is generally done once for each application domain. Once the goal is well defined, no interaction is required for the proposed method to run in an end-to-end fashion.

4. Conclusion

This paper presents a novel goal-driven unsupervised image segmentation method that, thanks to a novel definition of image segmentation taking into consideration a user-defined goal, is capable of partitioning the input image into a set of homogeneous regions of interest and a background area. By combining weighted graphs, parametric density modeling, and Markov random field modeling, the proposed method has been experimentally demonstrated to be effective in two different domains of applicability, i.e. medical magnetic resonance images and remote-sensing SAR imagery.

The proposed method has been experimentally evaluated with respect to other state-of-the-art solutions. Indeed, the experimental comparisons highlighted the capability of the proposed solution to take advantage of the user-defined goal and produce an accurate segmentation map that is able to discriminate the objects of interest from the background. Hence, the comparison with the previous methods was carried out by manually post-processing the results obtained by the existing methods, as the separation of the ROIs from the background in a goal-driven fashion is not typical of standard segmentation scenarios.

Moreover, thanks to the possibility of being integrated with multiple parametric models, the proposed method was effective in different applicability domains. Indeed, the experimental analysis was carried out with respect to MR and SAR images. In the former case, the adopted model was the Gaussian distribution, and the underlying parameter estimation strategy was addressed in a maximum-likelihood fashion via the sample mean and the sample variance. Conversely, in the latter case of active radar imagery, usually affected by a multiplicative noise-like component due to the speckle phenomenon, the Gamma distribution was chosen as an appropriate parametric model, and the parameter estimation was performed via the method of logarithmic cumulants.

Because of the definition of an underlying graph and the formulation of a Bayesian probabilistic graphical model using Markov random field modeling, the proposed method was proven able to integrate the local and global properties of the image, preventing artifacts either due to the intrinsic noise of the input image, or due to spurious regions relative to misplaced seeds, so that no merge step was necessary for generating the output segmentation. Moreover, the experiments conducted with two different application-specific scenarios associated with image data, with different statistical and noise characteristics, highlighted the effectiveness and versatility of the proposed method, which was also enriched by the unsupervised formulation not requiring any labeled dataset.

However, of course, the proposed method presents some limitations. The most relevant one is that it is likely to fail when ROIs are very similar and close, such as in the case of the dataset “Noerdlinger Ries”. Here, the dark fields and the city are very close and similar in terms of gray levels, and indeed, in such an area, the method performs poorly. In addition, in this situation, streets are not distinguishable from other regions; thus, they may be included in the fields.

Among the possible future developments, the proposed method could be reformulated to work with volumetric data [57]. Indeed, it would be necessary to reformulate the graph-based processing, together with the Markov random field model, to take into account the geometry of such 3D input data. In this case, the MR wrist bone dataset could be segmented by considering the entire volume at once and not each single slice sequentially. Moreover, the experimentation could be expanded to other data types, and therefore to different domains and user-defined goals. Therefore, to deal with these different sources, additional parametric models with appropriate parameter estimation strategies could be integrated in the proposed framework.

Additionally, the proposed segmentation method was specifically designed to identify homogeneous and non-granular regions. Indeed, in the experimental analysis, the Gaussian and the Gamma distributions were the adopted parametric model. Therefore, another possible development could be the extension to the case of textured and granular regions. In these cases, possible parametric models would be mixture models of the aforementioned distributions. Indeed, the model fitting phase would require solutions such as the expectation-maximization (EM) algorithm [58]. This could also fix the issue raised in the case of very similar and close ROIs.

Finally, a further interesting future development could be the integration of graph-based and MRF formulations with deep learning techniques. Such an hybrid methodology could indeed balance the weaknesses of the two approaches when used on their own. In the recent literature, some work dealing with possible combinations between the two approaches can be found in [59]. In fact, the proposed approach could also be combined with supervised methods and semantic segmentation, to be used, for instance, in a preprocessing phase. Indeed, semantic information is usually manually delivered by medical experts, thus being affected by intra-rater and inter-rater variability and by the experience of annotators. As a result, semantic segmentation would significantly benefit from starting from unsupervised image segmentation; for instance, allowing the expert to focus on peculiar areas where they could use their knowledge in the labeling phase. The deployment of the proposed approach, which not only acts in an unsupervised fashion but also encompasses the information related to the application context and the goal, could help for this purpose. It would make manual work easier for the experts, reduce raters variability, and bridge the experience gap among several experts.

Declaration of Competing Interest

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

CRediT authorship contribution statement

Marco Trombini: Conceptualization, Methodology, Software, Validation, Writing – original draft. **David Solarna:** Conceptualization, Methodology, Software, Validation, Writing – original draft. **Gabriele Moser:** Conceptualization, Methodology, Supervision, Writing – review & editing, Project administration. **Silvana Dellepiane:** Conceptualization, Methodology, Supervision, Writing – review & editing, Project administration.

Appendix A. Table of Notation

Symbol	Meaning
I	Pixel lattice over which an intensity image $\mathcal{X}$ is defined.
$\mathcal{X}$	Intensity image, i.e., the collection of the observed pixel intensities.
X	Two-dimensional image defined as $X = I \times \mathcal{X}$.
i	Pixel location on the image lattice I.
x_i	Single pixel of an intensity image $\mathcal{X}$, with $i \in I$.
$X_1, \ldots, X_M$	M disjoint, non-empty, connected subregions of an image X (ROIs).
M	Number of ROIs composing the image X.
P	Predicate of homogeneity.
B	Background of an image X.
h_g	Pixel intensity-related predicate.
h_s	Statistical property-related predicate.
A	General subset of the image X, $A \subseteq X$.
J_m	Feasible values interval for the mean in h_g.
J_v	Feasible values interval for the variance in h_g.
f	Feasible probability density function.
h_p	Position-related predicate.
$\bar{I}$	Area of the imaged scene where the target ROIs are supposed to be located.
H	Collection of all the involved predicates.
h_{other}	Term collectively indicating possible optional predicates.
$Bool$	Boolean set.
L	Label image composed of the M ROIs and the background B.
l_i	Single pixel of a label image L, with $i \in I$.
Λ	Set of labels assigned to the M subregions of X. $\Lambda = \{0, 1, 2, \ldots, M\}$, with the value 0 indicating the background.
S_0	Initial set of seeds.
s	Single seed belonging to the set S_0 (i.e., $s \in S_0$) and located inside the image X.
$\tilde{X}_s$	Small circular region of the image X around the seed s.
S	Filtered set of seeds.
N	Number of seeds belonging to the filtered set S.
s_n	Single seed belonging to the set $S = \{s_n\}, n \in \{1, 2, \ldots, N\}$ and located inside the image X.
F	Set of cost functions, each associated with a seed in S.
F_n	Single cost function belonging to the set $F = \{F_n\}$, with n identifying one of the seeds $s_n \in S$. It is defined over the whole image X such that $F_n : I \to \mathbb{R}$.
X_n	A region of the image X associated with the seed s_n and extracted via the function F_n.
B_0	The initial background. The set of image pixels not assigned to any region X_n, i.e., $B_0 = X \setminus \{\cup_{n=1}^{N} X_n\}$.
$U(\cdot)$	Markovian energy function.
$U(L\|X)$	Markovian energy function associated with the posterior distribution of the label map L conditioned on the image X.
$D_i(x_i\|l_i)$	Pixelwise energy term composing the Markovian energy function $U(\cdot)$ and corresponding to the pixel i.
$i \sim j$	Notation for the pixels i and j being neighbors according to a first-order neighboring system.
$V(l_i, l_j)$	Pairwise energy term composing the Markovian energy function $U(\cdot)$ and corresponding to the pixels i and j, with $i \sim j$.
β	Positive coefficient weighting the pixelwise and pairwise terms inside the Markovian energy function.
Λ'	Preliminary set of labels $\Lambda' = \{0, 1, 2, \ldots, N\}$. The value 0 is associated with the background, and the values from 1 to N are associated with each of the N seeds.
$G = (I, E)$	Non-empty, fully-connected, undirected, and vertex-weighted graph onto which the image X is mapped. I is the set of vertices (i.e., the pixel lattice of the image), while E is the set of edges connecting neighboring nodes.

(*continued on next page*)

Symbol	Meaning
$w(i, j)$	Difference function defined as $w : I \times I \to \mathbb{R}, w(i, j) = \|x_i - x_j\|$.
$w_n(i)$	It is equivalent to $w(i, s_n), \forall i \in I$.
T	Set of all the visited nodes.
i^-	Father node of i according to a flooding scheme.
$F_n(i)$	Value of the cost function F_n for a generic seed s_n and applied to the pixel i.
Z	Positive-valued random variable.
$\mathcal{Z}$	Set of independent and identically distributed (i.i.d.) samples drawn from Z.
z_i	An i.i.d. sample drawn from Z and belonging to the set $\mathcal{Z}$.
$p_Z(z)$	Probability density function of the random variable Z modeled according to the Gamma distribution.
Γ	Gamma function.
u	Mean of the Gamma distribution $p_Z(z)$.
α	Shape of the Gamma distribution $p_Z(z)$.
κ_i	ith order logarithmic cumulant of Z.
$\Psi(0, \alpha)$	Digamma function, i.e., the logarithmic derivative of the Gamma function.
$\Psi(1, \cdot)$	First-order polygamma function, i.e., the derivative of $\Psi(0, \alpha)$.
$\hat{\kappa}_i$	Sample estimate of the ith order logarithmic cumulant κ_i.
$\hat{u}$	MoLC estimate of u.
$\hat{\alpha}$	MoLC estimate of α.
$d_i(x\|n)$	Pixelwise potential corresponding to one of the seeds.
c_1, c_2	Values of the piecewise constant background-specific potential. They are automatically chosen according to the pth and $(1-p)$th percentiles of the distribution of the seed-specific potentials $d_i(x\|n)$.
$\hat{p}(x\|n)$	Estimate of the Gamma probability density function modeling the set of samples X_n.
$\Delta(i, s_n)$	Euclidean distance between the position of the pixel i and the position of the seed s_n within the pixel lattice I.
γ	Positive coefficient weighting the two terms composing the pixelwise potential $d_i(x\|n)$.
m_{MR}	Threshold used in the pixel intensity predicate for the MR imagery application.
M_{MR}	Threshold used in the pixel intensity predicate for the MR imagery application.
v_{MR}	Threshold used in the pixel intensity predicate for the MR imagery application.
m_{SAR}	Threshold used in the pixel intensity predicate for the SAR imagery applied to field segmentation.
M_{SAR}	Threshold used in the pixel intensity predicate for the SAR imagery applied to field segmentation.
v_{SAR}	Threshold used in the pixel intensity predicate for the SAR imagery applied to field segmentation.
m_{Oil}	Threshold used in the pixel intensity predicate for the SAR imagery applied to oil spill segmentation.
M_{Oil}	Threshold used in the pixel intensity predicate for the SAR imagery applied to oil spill segmentation.
v_{Oil}	Threshold used in the pixel intensity predicate for the SAR imagery applied to oil spill segmentation.

References

[1] K.-S. Fu, J. Mui, A survey on image segmentation, Pattern Recognit. 13 (1) (1981) 3–16.

[2] V. Tomatis, M.A. Cimmino, F. Barbieri, G. Troglio, P. Parascandolo, L. Cesario, G. Viano, L. Vosilla, M. Pitikakis, A. Schiappacasse, et al., A database of segmented MRI images of the wrist and the hand in patients with rheumatic diseases, in: International Conference on Image Analysis and Processing, Springer, 2015, pp. 143–150.

[3] J. Włodarczyk, W. Wojciechowski, K. Czaplicka, A. Urbanik, Z. Tabor, Fast automated segmentation of wrist bones in magnetic resonance images, Comput. Biol. Med. 65 (2015) 44–53.

[4] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, Int. J. Comput. Vis. 1 (4) (1988) 321–331.

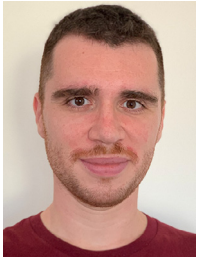
[5] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, Int. J. Comput. Vis. 22 (1) (1997) 61–79.

[6] J. Cardinale, G. Paul, I.F. Sbalzarini, Discrete region competition for unknown numbers of connected regions, IEEE Trans. Image Process. 21 (8) (2012) 3531–3545.

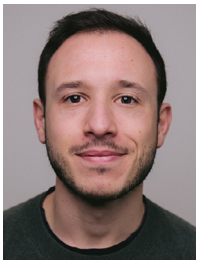
[7] D. Baltissen, T. Wollmann, M. Gunkel, I. Chung, H. Erfle, K. Rippe, K. Rohr, Comparison of segmentation methods for tissue microscopy images of glioblastoma cells, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 396–399.

[8] S. Yin, Y. Zhang, S. Karim, Large scale remote sensing image segmentation based on fuzzy region competition and gaussian mixture model, IEEE Access 6 (2018) 26069–26080.

[9] C.A. Bouman, M. Shapiro, A multiscale random field model for Bayesian image segmentation, IEEE Trans. Image Process. 3 (2) (1994) 162–177.
[10] Z. Tu, S.-C. Zhu, Image segmentation by data-driven Markov chain monte carlo, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002) 657–673.
[11] C. Panagiotakis, H. Papadakis, E. Grinias, N. Komodakis, P. Fragopoulou, G. Tziritas, Interactive image segmentation based on synthetic graph coordinates, Pattern Recognit. 46 (11) (2013) 2940–2952.
[12] H. Bi, L. Xu, X. Cao, Y. Xue, Z. Xu, Polarimetric SAR image semantic segmentation with 3D discrete wavelet transform and Markov random field, IEEE Trans. Image Process. 29 (2020) 6601–6614.
[13] H. Filali, K. Kalti, Image segmentation using MRF model optimized by a hybrid ACO-ICM algorithm, Soft Comput. 25 (15) (2021) 10181–10204.
[14] Q.-h. Zhao, X.-l. Li, Y. Li, X.-m. Zhao, A fuzzy clustering image segmentation algorithm based on Hidden Markov random field models and Voronoi tessellation, Pattern Recognit. Lett. 85 (2017) 49–55.
[15] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2017) 834–848.
[16] Y. Duan, F. Liu, L. Jiao, P. Zhao, L. Zhang, SAR image segmentation based on convolutional-wavelet neural network and Markov random field, Pattern Recognit. 64 (2017) 255–267.
[17] A. Ortega, P. Frossard, J. Kovačević, J.M. Moura, P. Vandergheynst, Graph signal processing: overview, challenges, and applications, Proc. IEEE 106 (5) (2018) 808–828.
[18] L. Gemme, S. Dellepiane, A new graph-based method for automatic segmentation, in: International Conference on Image Analysis and Processing, Springer, 2015, pp. 601–611.
[19] S. Dellepiane, F. Fontana, Extraction of intensity connectedness for image processing, Pattern Recognit. Lett. 16 (3) (1995) 313–324.
[20] A.X. Falcão, J. Stolfi, R. de Alencar Lotufo, The image foresting transform: theory, algorithms, and applications, IEEE Trans. Pattern Anal. Mach. Intell. 26 (1) (2004) 19–29.
[21] J. Cousty, G. Bertrand, L. Najman, M. Couprie, Watershed cuts: minimum spanning forests and the drop of water principle, IEEE Trans. Pattern Anal. Mach. Intell. 31 (8) (2008) 1362–1374.
[22] J. Cousty, L. Najman, Incremental algorithm for hierarchical minimum spanning forests and saliency of watershed cuts, in: International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing, Springer, 2011, pp. 272–283.
[23] J.E. Vargas-Muñoz, A.S. Chowdhury, E.B. Alexandre, F.L. Galvão, P.A.V. Miranda, A.X. Falcão, An iterative spanning forest framework for superpixel segmentation, IEEE Trans. Image Process. 28 (7) (2019) 3477–3489.
[24] D. Stutz, A. Hermans, B. Leibe, Superpixels: an evaluation of the state-of-the-art, Comput. Vision Image Understanding 166 (2018) 1–27.
[25] Z. Kato, J. Zerubia, et al., Markov random fields in image segmentation, Foundations Trends® Signal Process. 5 (1–2) (2012) 1–155.
[26] S.Z. Li, Markov random field modeling in image analysis, Advances in Computer Vision and Pattern Recognition, Springer-Verlag, London, 2009.
[27] K. Krupa, M. Bekiesińska-Figatowska, Artifacts in magnetic resonance imaging, Polish J. Radiol. 80 (2015) 93.
[28] F.T. Ulaby, R.K. Moore, A.K. Fung, Microwave remote sensing: active and passive. Volume 3-from theory to applications (1986). https://ntrs.nasa.gov/citations/19860041708.
[29] P. Sasmal, M.K. Bhuyan, S. Dutta, Y. Iwahori, An unsupervised approach of colonic polyp segmentation using adaptive Markov random fields, Pattern Recognit. Lett. 154 (2022) 7–15.
[30] C. Zheng, Y. Chen, J. Shao, L. Wang, An MRF-based multigranularity edge-preservation optimization for semantic segmentation of remote sensing images, IEEE Geosci. Remote Sens. Lett. 19 (2021) 1–5.
[31] L. Gemme, S.G. Dellepiane, An automatic data-driven method for SAR image segmentation in sea surface analysis, IEEE Trans. Geosci. Remote Sens. 56 (5) (2018) 2633–2646.
[32] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, IEEE Trans. Pattern Anal. Mach. Intell. 26 (9) (2004) 1124–1137.
[33] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, IEEE Trans. Pattern Anal. Mach. Intell. 23 (11) (2001) 1222–1239.
[34] H. Greenspan, A. Ruf, J. Goldberger, Constrained Gaussian mixture model framework for automatic segmentation of MR brain images, IEEE Trans. Med. Imaging 25 (9) (2006) 1233–1245.
[35] Z. Ji, Y. Xia, Q. Sun, Q. Chen, D. Xia, D.D. Feng, Fuzzy local Gaussian mixture model for brain mr image segmentation, IEEE Trans. Inf. Technol. Biomed. 16 (3) (2012) 339–347.
[36] D. Solarna, G. Moser, S.B. Serpico, Multiresolution and multimodality SAR data fusion based on Markov and conditional random fields for unsupervised change detection, in: 2019 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2019, pp. 29–32.
[37] D.-X. Yue, F. Xu, A.C. Frery, Y.-Q. Jin, Synthetic aperture radar image statistical modeling: part one-single-pixel statistical models, IEEE Geosci. Remote Sens. Mag. 9 (1) (2020) 82–114.
[38] A.N. Gavrilov, I.M. Parnum, Fluctuations of seafloor backscatter data from multibeam sonar systems, IEEE J. Oceanic Eng. 35 (2) (2010) 209–219.
[39] P. Sudeep, P. Palanisamy, J. Rajan, H. Baradaran, L. Saba, A. Gupta, J.S. Suri, Speckle reduction in medical ultrasound images using an unbiased non-local means method, Biomed. Signal Process. Control 28 (2016) 1–8.
[40] J.-M. Nicolas, S.N. Anfinsen, Introduction to second kind statistics: application of log-moments and log-cumulants to the analysis of radar image distributions, Trait. Signal 19 (3) (2002) 139–167.
[41] J.-M. Nicolas, F. Tupin, Gamma mixture modeled with "second kind statistics": application to SAR image processing, in: IEEE International Geoscience and Remote Sensing Symposium, vol. 4, IEEE, 2002, pp. 2489–2491.
[42] L. Bombrun, S.N. Anfinsen, O. Harant, A complete coverage of log-cumulant space in terms of distributions for polarimetric SAR data, in: POLinSAR 2011-5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry, 2011, pp. 1–8.
[43] J.-M. Nicolas, F. Tupin, Statistical models for SAR amplitude data: a unified vision through Mellin transform and Meijer functions, in: 2016 24th European Signal Processing Conference (EUSIPCO), IEEE, 2016, pp. 518–522.
[44] C. Tison, J.-M. Nicolas, F. Tupin, H. Maître, A new statistical model for Markovian classification of urban areas in high-resolution SAR images, IEEE Trans. Geosci. Remote Sens. 42 (10) (2004) 2046–2057.
[45] V.A. Krylov, G. Moser, S.B. Serpico, J. Zerubia, On the method of logarithmic cumulants for parametric probability density function estimation, IEEE Trans. Image Process. 22 (10) (2013) 3791–3806.
[46] M. Abramowitz, I. Stegun, Handbook of mathematical functions: with formulas, graphs, and mathematical tables, National Bureau of Standards Applied Mathematics Series, Dover, New York, NY, USA, 1964.
[47] M. Hazewinkel, Newton Method, Encyclopedia of Mathematics, Springer, 2001.
[48] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, SLIC superpixels compared to state-of-the-art superpixel methods, IEEE Trans. Pattern Anal. Mach. Intell. 34 (11) (2012) 2274–2282.
[49] J. Besag, Spatial interaction and the statistical analysis of lattice systems, J. R. Stat. Soc. Ser. B (Methodological) 36 (2) (1974) 192–236.
[50] S.G. Dellepiane, E. Angiati, A new method for cross-normalization and multitemporal visualization of SAR images for the detection of flooded areas, IEEE Trans. Geosci. Remote Sens. 50 (7) (2012) 2765–2779.
[51] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process. 10 (2) (2001) 266–277.
[52] H.-H. Chang, A.H. Zhuang, D.J. Valentino, W.-C. Chu, Performance measure characterization for evaluating neuroimage segmentation algorithms, Neuroimage 47 (1) (2009) 122–135.
[53] T. Eelbode, J. Bertels, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, M.B. Blaschko, Optimization for medical image segmentation: theory and practice when evaluating with dice score or jaccard index, IEEE Trans. Med. Imaging 39 (11) (2020) 3679–3690.
[54] M.M. Petrou, C. Petrou, Image Processing: The Fundamentals, John Wiley & Sons, 2010.
[55] K. McGuinness, N.E. O'connor, A comparative evaluation of interactive segmentation algorithms, Pattern Recognit. 43 (2) (2010) 434–444.
[56] R. Shang, J. Lin, L. Jiao, Y. Li, SAR image segmentation using region smoothing and label correction, Remote Sens. 12 (5) (2020) 803.
[57] E.M.A. Anas, A. Rasoulian, A. Seitel, K. Darras, D. Wilson, P.S. John, D. Pichora, P. Mousavi, R. Rohling, P. Abolmaesumi, Automatic segmentation of wrist bones in CT using a statistical wrist shape + pose model, IEEE Trans. Med. Imaging 35 (8) (2016) 1789–1801.
[58] T.K. Moon, The expectation-maximization algorithm, IEEE Signal Process. Mag. 13 (6) (1996) 47–60.
[59] M. Brudfors, Y. Balbastre, J. Ashburner, G. Rees, P. Nachev, S. Ourselin, M.J. Cardoso, An MRF-UNet product of experts for image segmentation, in: Medical Imaging with Deep Learning, PMLR, 2021, pp. 48–59.



Marco Trombini received the BSc degree and MSc degree (cum laude) in Applied Mathematics from the Università degli Studi di Genova in 2016 and 2018 respectively, and the PhD degree in Science and Technology for Electronic and Telecommunication Engineering from the same University in 2022. From January to June 2015, he was a Visiting Student with the Universitetet i Bergen, Bergen, Norway. He is Member of the IEEE. His research topics cover signal processing, machine learning, and pattern recognition.



David Solarna received the BSc degree (cum laude), the MSc degree (cum laude), and the PhD degree from the University of Genoa, Italy, in 2014, 2016, and 2021, respectively. In 2016, he was a Visiting Researcher with the NASA Goddard Space Flight Center, Greenbelt, MD, USA. From 2017 to 2020, he was a Research Fellow with the NATO STO CMRE, La Spezia, Italy. His area of interest covers data fusion, with particular focus on image registration techniques, machine learning, and pattern recognition methods applied to remote sensing and underwater sonar data, and the combined use of modeling and simulation and autonomous systems in the underwater domain. Dr. Solarna received the Interactive Symposium

Best Paper Award of the 2016 IEEE International Geoscience and Remote Sensing Symposium.

Gabriele Moser is a Full Professor of Telecommunications at the University of Genoa. Since 2001, he has cooperated with the Image Processing and Pattern Recognition for Remote Sensing laboratory of the University of Genoa. Since 2013, he has been the Head of the Remote Sensing for Environment and Sustainability laboratory at the Savona Campus of the University of Genoa. From January to March 2004, he was a Visiting Student with the Institut National de Recherche en Informatique et en Automatique (INRIA), Sophia Antipolis, France. In 2016, he spent a period as Visiting Professor at the Institut National Polytechnique de Toulouse, France. His research activity is focused on pattern recognition and image processing methodologies for remote sensing and energy applications. He has been an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS since 2008. He was an Area Editor of PATTERN RECOGNITION LETTERS (PRL) from 2015 to 2018, an Associate Editor of PRL from 2011 to 2015. He is a Senior Member of the IEEE.

Silvana Dellepiane received the Laurea (MSc equivalent) degree (cum laude) in Electronic Engineering and the PhD in Electronic and Computer Engineering in 1986 and in 1990 respectively. She became university researcher in 1992 at Universitá degli Studi di Genova, where she is currently Associate Professor. She has been teaching Signal Theory, Recognition Theories and Techniques, Statistical Methods. She is currently teaching Electrical Communications and Digital Image Processing in the courses of Telecommunications Engineering, Biomedical Engineering and History of Art and Preservation of Cultural Heritage. She has been responsible for signal analysis during rehabilitation in the Rehab@home EU project. She is member of IEEE BISP Technical Committee and Honorary Member of SISS Societá Italiana per lo Studio dello Stroke.