A survey of recent interactive image segmentation methods

Hiba Ramadan[1] (✉)**, Chaymae Lachqar**[2]**, and Hamid Tairi**[1]

Abstract Image segmentation is one of the most basic tasks in computer vision and remains an initial step of many applications. In this paper, we focus on interactive image segmentation (IIS), often referred to as foreground–background separation or object extraction, guided by user interaction. We provide an overview of the IIS literature by covering more than 150 publications, especially recent works that have not been surveyed before. Moreover, we try to give a comprehensive classification of them according to different viewpoints and present a general and concise comparison of the most recent published works. Furthermore, we survey widely used datasets, evaluation metrics, and available resources in the field of IIS.

Keywords interactive image segmentation; user interaction; label propagation; deep learning; superpixels

1 Introduction

The main goal of image segmentation is to divide an image into homogeneous regions according to common characteristics such as spatial position, color, shape, texture, and motion (in the case of video segmentation). Emulating the human perceptual system's ability to segment or divide an image into meaningful regions is still challenging and has been widely studied since the early days of computer vision. While many works, including recent ones, have been presented in the literature to review segmentation algorithms [1, 2], semantic segmentation techniques

1 Department of Informatics, Faculty of Sciences Dhar El Mahraz, University of Sidi Mohamed Ben Abdellah, Fez 30000, Morocco. E-mail: hiba.ramadan@usmba.ac.ma (✉).

2 Faculty of Medicine and Pharmacy, University of Sidi Mohamed Ben Abdellah, Fez 30000, Morocco.

[3], and medical image segmentation [4, 5], few works have been dedicated to interactive image segmentation (IIS) methods, although research in this area is very active and a recent periodic overview remains necessary. In fact, except for the comparative evaluation in Ref. [6] of some IIS techniques proposed before 2010 and the work of Ref. [7] in 2013, the only work to do so, as far as we are aware, is Ref. [1] which briefly addressed IIS as part of its survey.

IIS, "supervised segmentation", and "semi-automatic segmentation" all mean the task of extracting an image region or object of interest from the background (BG) using prior knowledge provided by user interaction. This interaction, either in the form of some points or scribbles to mark the object of interest and/or the BG, either using a bounding box (BB) or polygon to delimit the region of interest (ROI), allows the user to provide good constraints (on size, color, location, objectness ...) to guide the segmentation process. This can improve results as well as reducing runtime compared to automatic segmentation methods [7]. In fact, many computer vision applications (medical imaging, image editing, object recognition, and object tracking) need such user intervention to obtain accurate segmentation results, which are then used as input for other high-level processing.

IIS methods can be classified in different ways depending on the criteria used.

User interaction: the type of user interaction required can be used to divide methods into seed-based and ROI-based approaches [8] or into active and passive interaction-based approaches [9].

Methodology: the methodology used to segment the desired object can be based either on contours or label propagation [1]. In this work, IIS methods are divided into: contour, graph cut (GC), random walk (RW), and region merging (RM)/region growing

(RG) based methods. In addition to these categories, with the recent great success of convolutional neural networks (CNN) [10] in image segmentation and classification, a new family of CNN-based IIS methods has appeared, providing high accuracy. IIS techniques that do not adopt any of the above methodologies are classified as "other methods".

Processing level: IIS techniques can be categorized into pixel-wise, superpixel-wise, or hybrid (pixel/superpixel) methods according to the type of image units used in segmentation [7].

Figure 1 summarizes the different categories of IIS methods according to these criteria.

Due to the importance and progress of research in the field of IIS, the aim of this work is to update the literature in this area and to present a recent reference for researchers. We do not tackle the task of interactive video object segmentation (VOS) which usually takes motion information into consideration to segment the desired object. For a good survey of VOS methods, the reader can refer to Ref. [11]. Furthermore, the application of IIS to medical imaging is not addressed in this survey; it has already been discussed in Refs. [4, 5].
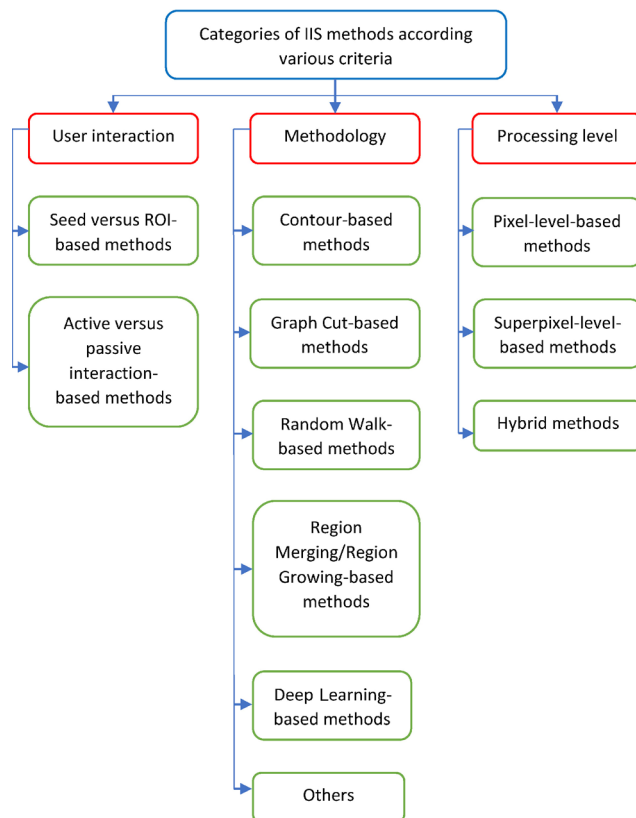

Fig. 1   Categories of IIS methods according to various criteria.

In the rest of this paper, we will review the different families of IIS methods classified according to user interaction (Section 2), the adopted methodology (Section 3), and the level of processing (Section 4). Moreover, the most frequently used datasets and evaluation metrics will be reviewed in Sections 5 and 6 respectively. Existing evaluations of IIS methods in the literature and a general comparison of the most recent works are presented in Section 7. Before concluding, Section 8 provides links to code and software for different IIS methods.

2   Classification of IIS methods based on user interaction

To define the image content to be segmented, IIS methods need some kind of user interaction. Many kinds of interaction can be involved, for example some points, line segments, or strokes to mark the object and/or the BG. These interactions provide seeds. Another popular type of interaction is to delimit the desired object with a BB, polygon, or any closed contour to define the ROI. According to the nature of the user interaction, IIS techniques can be divided into seed-based and ROI-based methods. From another perspective, some IIS methods provide the possibility for active intervention or online assistance from the user during the segmentation process [9]. Such methods can be called active interaction-based approaches, while others are considered to be passive.

In this section, we consider different IIS methods according to the type of interaction without giving details of the algorithms; these will be addressed in the methodology classification section.

2.1   Seed-based methods versus ROI-based methods

Seed-based methods.   Early seed-based IIS approaches include intelligent scissors [12] and live wire [13]. These methods can be classified as boundary seed-based methods since they require accurate seed points (or anchor points) on the boundary of the desired object. In the same subcategory, Ref. [14] is a variant of Ref. [13] which improves its speed, and Riverbed [15] requires fewer seeds to segment the object. The other subcategory of seed-based IIS methods is region seed-based approaches. These methods do not impose hard constraints on user interaction like the first family,

but the result of segmentation is very sensitive to the number of seeds, and different interactions may give different results. The seeds may be just a few points [16–26] or segments of points (scribbles, strokes) [27–48]. Examples of region seed-based methods are interactive GC (IGC) [49], random walk (RW) [50], and their many variants [51–71]. These approaches exploit priors from object and BG regions marked by the user with other constraints via an energy minimization framework to segment the object of interest. Many other works using active contours (AC) [72–74] or region merging (RM) [75, 76] are also region seed-based methods. An example of a hybrid that combines boundary seeds and region seeds is proposed in Ref. [77]. Based on initial boundary seeds, the object contour is traced then a set of foreground (FG) and BG region seeds are selected from adjacent contour points, and the final segmentation is achieved using GC [49] or an image foresting transform [78].

ROI-based methods. Another popular way to guide segmentation is to draw an ROI, and then prior knowledge about FG and BG is extracted from pixels inside and outside the ROI, respectively. GrabCut [79] is the most widely used IIS method which allows user to draw a BB separating the object of interest from the BG. Many works have been presented in the literature to improve GrabCut's performance, for instance, One-Cut [80], Loose-Cut [81], Point-Cut [82], Super-Cut [83], Dense-Cut [84], Deep-Cut [85], Deep GrabCut [86], and Neutro Connectedness Cut (NC-Cut) [8]. Also using a provided BB, the authors in Refs. [87, 88] solved the object segmentation as a figure-ground classification problem. As discussed in Ref. [81], the performance of many BB-based IIS techniques degrades when the input BB does not tightly cover the FG object. To overcome this issue, some works introduce the concept of a *tightness* prior such as Pin-Point [89] and Mil-Cut [90]. Since ROI-based methods use a simpler form of interaction, there are many such approaches, our list is not exhaustive [91–96]. Some methods in literature accept either seeds or ROI as methods of user interaction [53, 97–100].

Figure 2 gives some examples of the different modalities of user input for IIS.

In general, it is easier for users to indicate the candidate object using an ROI by making some mouse clicks to specify a BB or a polygon. However,

due to the complexity of the object boundary and its appearance, segmentation accuracy is usually limited by how tightly the ROI is delimited [81]. Therefore, many ROI-based approaches integrate iterative refinement steps [8, 79] or/and added new constraints to their framework to achieve accurate object segmentation [8, 81].

On the other hand, seed-based algorithms can tackle complex-shaped objects as long as sufficient user inputs are given. Sometimes more rounds of interaction are needed than for ROI-based algorithms [101]. However, some authors have fused multiple cues (edges, regions, and geometric cues) in their framework to segment an object of interest driven by a single touch [102]. Another work [103], minimizes use of user provided seeds by first generating many segmentations of the input image automatically, and then allows the user to click on the boundary of the desired object to refine the final segmentation. Others have exploited all modalities of user interaction by integrating BB-based, seed/scribble-based, and query-based (see Section 2.2) mechanisms in an unified scheme for fast IIS [104].

Regardless of the interaction modality, the key challenge of any IIS is to minimize the amount of effort by the user while accurately segmenting the desired object. Recently, this goal has been ever more achieved by DL-based techniques [17, 23, 105–109] which have demonstrated high performance with minimum user inputs.

2.2 Active interaction-based methods versus passive interaction-based methods

The principle of active IIS is to give an initial estimate of the image segmentation, and based on this human–machine interaction is required to provide the user's intent. The authors in Ref. [110] propose an active IIS method which suggests uncertain regions to the user based on non-local uncertainty measurements, and then watershed cut [111] is applied to get the final segmentation, guided by user selection. Another method using this kind of segmentation is introduced in Ref. [112], which asks the user some binary questions based on the probability distribution over a set of sampled segmentations, and then based on the user's responses, the computer estimates label of image regions. In Ref. [9] and its extended version [113], IIS from 1-Bit Feedback is presented using superpixels, entropy, and transductive inference to

Fig. 2 Different modalities of user input in IIS. (a) Boundary seeds. (b) Region seeds, FG, and BG points. (c) Region seeds, FG, and BG strokes/scribbles. (d) ROI, drawn. (e) ROI, loose BB. (f) ROI, tight BB.

propose a sequence of informative yes–no questions to the user. Segmentation is then performed progressively according to the user's answers until the desired result is obtained. The work presented in Ref. [114] incorporates the user's feedback via a constrained spectral clustering method to achieve IIS. The strategy used alternates between updating the segmentation and requesting the user to select pairwise constraints. The process finishes when the user is satisfied with the segmentation result or iterations. More recent work in Ref. [104] starts with a user ROI covering the desired FG object. Then, the proposed system provides an active user-assistance mechanism. The most uncertain region is presented as a query for the user who responds with a true-or-false answer to label it.

A few works use online query-based human–machine interaction to guide segmentation. Other related works have used this concept for image co-segmentation [115], video segmentation [116], and 3D reconstruction [117]. In passive interaction-based methods, the majority of IIS techniques, it is the user's responsibility to choose the scribbles to guide

segmentation, these provide the input used by the computer to perform segmentation. Results are updated when the user modifies the input strokes.

Although passive interaction is widely used not just for the task of image segmentation, but in many computer vision applications, active interaction presents a very important alternative in some specific scenarios where input devices that receive binary signals can be used to collect user responses and exploit them without touching the computer. For example, in a sterilized environment, a physically touched computer control for medical image segmentation is inappropriate, or, for instance, when tiny screens on wearable computers have limited interface capabilities [9, 113].

3 Classification of IIS methods based on methodology

Based on the adopted methodology, IIS methods can be classified in different ways [7]. Here, we divide IIS techniques into contour-based methods, GC-based methods, RW-based methods, RG/RM-

based methods, deep learning (DL) techniques, and others. We try to discuss the majority of the methods in these categories, especially recent works that have not been covered in previous surveys. Our proposed classification is not strict because many IIS methods employ a variety of techniques in their algorithms, so can belong to more than one category.

3.1 Contour-based methods

The main principle of contour-based methods is to extract object contours using edge features and prior knowledge provided by the initial user interaction.

For example, intelligent scissors [12], which can be seen as an implementation of Live wire [13], has been integrated successfully as a tool in GIMP [118], and is one of the earliest contour-based methods. The object contour is extracted by calculating the shortest path linking the seed points using Djikstra's algorithm. A variant of Ref. [13] is proposed in Ref. [14] using a faster shortest path algorithm for improved speed. Riverbed [15] is another IIS based on boundary seeds which requires fewer user interactions and uses an optimum boundary tracking process to extract the object contour.

The supervised active contours model (ACM) is the best known family of contour-based IIS methods. For instance, interactive convex active contours (CAC) [72] takes as input user FG and BG strokes and the segmentation result obtained by any other ISS technique such as Ref. [53] or [50]. Then the energy equation of the convex active contours in Ref. [119] is reformulated to take into account prior knowledge from the user interaction and the probability map of the initial segmentation. Finally, evolution of the contours is performed by minimizing the energy using the split Bregman method [120]. To improve the performance of the CAC method, the authors in Ref. [121] introduce a geodesic energy term in their model; a seed refinement technique is used to update the segmentation iteratively. Recently, to tackle the case of noisy images with inhomogeneous intensity, the authors in Ref. [122] proposed a multi-phase level-set model by combining denoising a constrained surface with a denoising fidelity term. In Ref. [74], an interactive ACM with kernel descriptor is presented. First, a color kernel descriptor is employed to compute image patch features, and these patches are grouped into clusters. Then, template feature sets are used to formulate an energy functional and evolution is

performed via the level set method based on the initial contour provided by user interaction.

Using an ACM, the work of Ref. [123] presents a selective segmentation method based on the Chan–Vese model [124] and the geodesic ACM [125], which can segment noisy images given an initial contour and some boundary object seeds. Another supervised ACM based on self-organizing-map (SOM) is presented in Ref. [126]. It combines a variational level set method with the weights of the neurons of two SOMs to preserve the image intensity distribution. The authors in Refs. [127] and [128] respectively use nonparametric kernel density estimation (KDE) and a parametric Gaussian mixture model (GMM) in the level set framework to guide evolution of the contour. Minimum paths in the ACM framework have demonstrated a good performance for boundary extraction in many works such as the geodesically linked ACM introduced in Refs. [129] and [130] and the Finsler minimal path model [131–133]. The method in Ref. [134] incorporates discriminative classification models and distance transforms with the level set to avoid local minima and extract accurate object boundaries. Another IIS framework proposed in Ref. [135] uses level sets and Dempster–Shafer theory of evidence. Recently, a novel local region-based ACM for supervised segmentation uses Bayes theorem [136].

Contour-based methods for IIS are very efficient for accurate boundary extraction and very suitable for deformable object segmentation. Several methods based on ACM have been introduced in recent years to deal with intensity inhomogeneity and the presence of noise [122, 126]. Common shortcomings of this category of methods are the need for manual adjustment of the initial parameters and a lenground-truthy processing time [137].

3.2 GC-based methods

The first GC-based IIS (IGC) [49, 138] solved the problem of image labelling in the Markov random field (MRF) framework [139] by optimizing the following energy functional using min-cut/max-flow algorithm [140]:

$$E(L_b) = \sum_i U_i(l_i) + \sum_{(i,j) \in N_s} V_i(l_i, l_j) \qquad (1)$$

where $L_b = \{l_i\}$ is the final segmentation of the image and $l_i \in \{0, 1\}$ such that, $l_i = 1$ if the ith pixel belongs to the FG and $l_i = 0$ otherwise.

TSINGHUA UNIVERSITY PRESS | Springer

The first term of the energy equation U_i is a unary potential calculated from the intensity histogram. The second term V_i is the pairwise potential encouraging spatial coherence. i and j index pixels and N_s is the set of pairs of adjacent pixels.

GrabCut [79] extends IGC by replacing the unary potential in Eq. (1) by Gaussian mixture models (GMM) computed from BG and FG regions instead of the monochrome histogram. Furthermore, GrabCut applies an iterative framework by alternating between optimization and GMM estimation until convergence. Lazy snapping (LS) [59] was proposed to improve both speed and accuracy of IGC by processing image regions instead of pixels using the watershed algorithm for over-segmentation [141]. GMMRF [142] is an adaptive Gaussian mixture Markov random field model based on a *pseudo-likelihood* algorithm [143] that aims to learn GC parameters from image data before optimizing the energy functional. PinPoint [89] employed convex continuous optimization and GC to tackle the task of IIS under hard constraints with respect to the tightness of the provided BB. The authors in Ref. [144] applied the same segmentation scheme as GMMRF except that an initial optimization step is included. This initial step combines the standard k-means algorithm [145] with swap moves [146] to find a clustering that leads to histograms which better separate the FG region from the BG GrabCut in One-Cut [80] proposed a new energy term using the L_1 distance between FG and BG appearance models that can be optimized in a single GC step assuring good segmentation and fast running time. LooseCut [81] followed the MRF used in GrabCut [79], to handle cases where the BB only loosely covers the object of interest. To do so, the authors added a label consistency term to the energy function to give the same label to pixels with similar appearance. Furthermore, a global similarity constraint is imposed in the iterative process to distinguish FG and BG according to the difference between the appearance models. The work of Ref. [61] proposed discriminative Gaussian mixtures (DGMs) to boost the performance of Ref. [49]. It uses an automatically selects features, number of models in the mixture, and parameter estimation to maximize the discriminant power of the models.

In Ref. [64], a diffusive likehood based GC method is proposed to perform accurate object segmentation from input seeds using a likehood diffusion strategy and perceptual learning. First, initial probabilities of both pixels and superpixels are estimated using a GMM. Then, a diffusion technique is applied to explore global similarity relationships. Finally, the segmentation is obtained by optimization of an energy function at pixel and superpixel levels using the GC algorithm. The method of Ref. [147] combines both color and texture information in the GC model and incorporates AC in the segmentation process to address the case of textured images. Using pixel-level and patch-level information, Ref. [148] generates structural features of image patches using a GMM. Then, patch-level and pixel-level information are integrated in the GC framework, allowing preservation of details around boundary regions and improving the accuracy of the segmentation result. Superpixel-guided IIS based on GC is presented in Ref. [31]; it achieves segmentation in two stages. An initial segmentation is obtained at the superpixel level using GC, and then a narrow band is constructed along object contours using a morphology operator. Secondly, pixel-level GC segmentation provides accurate segmentation of pixels around the edges. Super-Cut [83] was proposed to improve GrabCut. It first over-segments the input image into superpixels, and then clusters them using a novel local similarity constraint. Finally, segmentation is finished in one cut using the clustered results.

GC-based methods with geodesic priors exploit the compactness of objects as a spatial constraint and use geodesic distance to compute the lowest cost path between two points instead of Euclidian distance. For example, in Ref. [54] the authors integrate the geodesic distance in the data term of the energy function; Ref. [53] extends the same principle to perform soft segmentation. Geos [149] is another geodesic-based method that computes the geodesic distance to obtain a set of restricted image segments, and obtains the final segmentation by finding the solution minimizing the energy cost. Geodesic star convexity proposed in Ref. [55] employs the geodesic distance in addition to shape constraints to perform robust IIS. Recent work in Ref. [150] takes the advantages of Refs. [54] and [80] to develop a geodesic appearance overlap GC framework for IIS.

GC-based methods with shape priors impose shape

constraints for IIS. For example, Refs. [55, 57, 151, 152] take into consideration the fact that most objects are convex and use star convexity constraints on IGC to improve the connectivity of segmented FG objects either using single star convexity [140] or multiple stars [55]. To incorporate prior shape knowledge, Ref. [153] integrates graph edge-weights containing information about a level-set function of a template, in addition to the boundary and region terms in the GC formulation. In Ref. [97], an adaptive optimal shape constraint selection system uses a non-rigid shape registration technique and local shape consistency evaluation to optimize the GC model. To handle objects with compact shape, Ref. [154] includes a compact shape prior in the GC framework to achieve robust segmentation with a minimum of user seeds.

GC-based methods with connectivity priors exploit topological properties of the image. Connectivity in digital topology [155] defines adjacency between points and it is usually used to solve the *shrinking bias* problem in GC: see Refs. [29, 57]. Recently, Ref. [8] introduced the concept of *neutro-connectedness* (NC) from classic logic to generalize the notion of *fuzzy connectedness* (FC) proposed earlier in Ref. [156], used for medical image segmentation in Refs. [157, 158]. NC-Cut models the topology of image regions with indeterminacy, and then segmentation is performed using both pixel-wise appearance models and region-based NC in the GC framework. EISeg [19] improves the effectiveness of user interaction by computing the NC between image regions and image boundary and provides visual cues so that the user can guide the segmentation process with a minimum number of seeds.

GC-based methods present the most popular approaches for IIS. Our list of works discussed is incomplete and many other ideas have been proposed in the literature to extend or improve the efficiency of GC [28, 39, 51, 52, 56, 63, 92, 96, 159].

3.3 RW-based methods

The original random walk model [50] starts by constructing an undirected graph $G = (V, E)$ to represent the input image. V is the set of nodes; node v_i designates pixel i. E is the set of edges; each edge connects two neighboring pixels (i, j) and has a pairwise weight $w_{i,j}$ reflecting the probability of a random walk stepping between these two

nodes. The principle of RW is to find the set x of probabilities where x_i is the probability of a random walk remaining at node i. The solution is obtained by minimizing the following function:

$$E(x) = x^{\mathrm{T}} L x \tag{2}$$

where L is the combinatorial Laplacian matrix given by

$$L = \begin{cases} \sum_j w_{i,j}, & \text{if } i = j \\ -w_{i,j}, & \text{if } i \neq j \text{ and } i \text{ and } j \text{ are adjacent nodes} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Let V_m and V_u be the sets of nodes marked by user scribbles and unmarked nodes respectively: $V_m \cup V_u = V$ and $V_m \cap V_u = \emptyset$. Let f_j be a label vector of dimension $|V_m| \times 1$ defined as

$$f_j = \begin{cases} 1, & \text{if } j \text{ belongs to } fg \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Letting the matrix L be divided into blocks:

$$L = \begin{bmatrix} L_m & B \\ B^{\mathrm{T}} & L_u \end{bmatrix} \tag{5}$$

The result of the minimization of Eq. (2) can be obtained by solving a sparse linear system:

$$L_u x_u = -B^{\mathrm{T}} f \tag{6}$$

To handle images with complex texture, RW with restart (RWR) [160] introduces the restart probability that the RW will return to the starting node or walk out to an adjacent node. In Ref. [68] image content is modelled using a directed hypergraph adapted to semi-supervised segmentation using an RW process. Sub-Markov RW (SRW) [161] is based on a traditional RW on the graph but adds some new auxiliary nodes. The proposed sub-RW method with label priors solves the twig segmentation problem (handling long thin objects). Graph-driven diffusion and RW for IIS [66, 162] incorporates a degree term into the original RW formulation to take into consideration the centrality of every adjacent node and measures its contribution in the diffusion process. In Ref. [163], a new energy functional is used to generalize the RW to semi-local and nonlocal frameworks. The Laplacian coordinates (LC)-based IIS method [58] uses a simple formulation of RW that improves the diffusion process by keeping pixels with similar attributes close to each other and imposing big jumps on the boundary of image regions. Another RW-based model using constrained Laplacian optimization is proposed

in Ref. [164] that incorporates the constraints provided by user scribbles into the energy function as Laplacian energy and applies an acceleration strategy to reduce the runtime of the proposed algorithm. Iterative boundary RW [165] presented two approaches, iterative RW and boundary RW, for segmentation potential, in order to develop a feedback system and produce an intuitive segmentation with reduced input. In addition to these approaches, other IIS algorithms [32, 62, 100, 108, 111, 166, 167] have proposed to extend the original Grady RW [50].

3.4 RG/RM-based methods

The main idea of region growing and region merging methods is to start from provided labels given by user seeds and to merge similar adjacent regions according to a homogeneity criterion. The original seeded RG method is presented in Ref. [16] and many works have been proposed to improve its performance such as Refs. [25, 168–172]. Maximal similarity-based RM (MSRM) in Refs. [75, 76] applies a pre-processing step by over-segmenting the input image into regions or superpixels, and then the user marks the FG and BG. The RM operation is then performed iteratively until all image regions are labelled according to user intention. Other work in Ref. [93] presented an object extraction method from a BB using a double sparse reconstruction strategy, and then applied the RM process of Ref. [75] to obtain the final segmentation result. In Ref. [98], a fast IIS method is presented using discriminative clustering and RM. After over-segmenting the image using the mean-shift algorithm [173], discriminative clusteringbased RM is performed to classify the unmarked regions using color features omitting the spatial information. Then, a pruning step is applied by considering local neighborhood information and using a connected component algorithm. In addition to these works, other proposed RM techniques for IIS include Refs. [174, 175].

Despite the conceptual simplicity of RG/RM-based approaches, the limitation of such methods is that different merging orders can produce different results [16]. Furthermore, accurate FG extraction requires sufficient user input to cover the main feature regions, especially when parts of the FG are very similar to the BG and in the presence of challenges such as shadows, low-contrast edges, and textured objects [75].

3.5 Deep learning (DL)-based methods

In recent years, CNNs have achieved great success in many vision applications [10] including image segmentation. A detailed recent survey of image segmentation using DL can be found in Ref. [176].

Recently, many IIS methods using deep architectures have been proposed achieving higher performance than classical models. Deep Extreme Cut (DEXTR) [17] presents a CNN architecture for IIS using extreme points at top, bottom, left, and right of the object of interest. The annotated extreme points are given as input to the network in addition to the RGB channels of the input image; ResNet-101 [177] is utilized as the backbone of the proposed architecture. Deep interactive object selection [23] transforms positive and negative clicks provided by the user into separate Euclidean distance maps. These maps are combined with the RGB channels of the input image to create an (image, user interactions) pair used by a fine tuned fully convolutional network (FCN). Finally, a refinement step is applied using GC optimization. Deep GrabCut for Object Selection [86] is inspired by Ref. [23] and takes a user-provided BB as input instead of user clicks. Furthermore, the proposed model is trained end-to-end and does not need a post-processing step to refine the FCN outputs. Deep interactive region segmentation and captioning [178] uses a hybrid deep architecture that allows detection, segmentation and captioning of the user region of interest from a few clicks using a new deep architecture named Lyncean FCN (LFCN). Iteratively trained IIS [22] followed the work of Ref. [23] and uses an iterative model that receives user clicks as input for the CNN then during training, further clicks are added iteratively based on the errors in the current segmentation. Regional IIS networks [179] presented a new deep framework that exploits local information surrounding provided inputs to capture local region information. Then, multiscale global information is used to improve the feature representation. IIS using a fully convolutional two-stream fusion network [180] has a new deep architecture consisting of two sub-networks: a two-stream late fusion network that estimates the FG at a reduced resolution, and a multi-scale refinement network that refines the FG at full resolution. SeedNet [21] provides an

interactive segmentation agent which assists a user to segment an object accurately using an automatic seed generation framework with deep reinforcement learning. The work in Ref. [181] investigated the role of guidance maps in IIS using FCNs. A scale aware guidance map is generated using hierarchical image information, which leads to a significant reduction in the average number of clicks (NoCs) required to extract a desirable FG mask. IIS with latent diversity [24] presents a composite architecture. The first module is a single FCN trained to takes the user's input and the image representation to synthesize a diverse set of solutions. The second module is a network trained to select one of the synthesized segmentations. The authors of Polygon-RNN [182] and Polygon-RNN++ [183] solve IIS as a polygon prediction problem using a recurrent neural network (RNN) to sequentially predict the vertices of the polygons delimiting the object of interest. 100× faster than Polygon-RNN++, a new framework alleviates the sequential nature of Polygon-RNN by predicting all vertices simultaneously using a graph convolutional network (GCN) [184]. Furthermore, the proposed Curve-GCN enables both a polygon and a spline representation of an object contour. DeepCut [85] presents a DL-based extension of GrabCut and formulates object segmentation as an energy optimization problem via a densely connected CRF. It applies an iterative process to update the training models and obtain the final segmentation. The work in Ref. [185] combines powerful CNN models with level set optimization in an end-to-end fashion to tackle the task of IIS. The proposed model employs a multi-branch architecture that learns to predict level set evolution parameters conditioned on a given input image, and evolves a predicted initial contour to extract the object. Both extreme points and motion vectors from annotators dragging and dropping erroneous points, have been incorporated in the interactive framework. Recently, Ref. [106] proposed a backpropagation scheme-based IIS algorithm. To segment a target object, an FCN is trained, and in the test phase, the forward pass in the proposed network is performed using an input image and user-annotation. A backpropagating refinement scheme (BRS) which constrains user-specified locations to have correct labels and refines

the segmentation result of the forward pass, is also developed in this work. Following Ref. [106], the authors in their newly proposed work [186] have developed a feature-BRS (f-BRS) that solves the optimization problem with respect to auxiliary variables instead of the network inputs, and requires running forward and backward passes just for a small part of a network. An IIS method that considers all regions jointly has been proposed in Ref. [107] based on Mask R-CNN [187]. The authors have adapted their architecture to deal with single object segmentation as well as full image segmentation. In Ref. [188], a scale-diverse IIS network based on Resnet-101 is proposed, incorporating a set of two-dimensional scale priors into the model to generate a set of scale-varying proposals that conform to the user input. New work in Ref. [189] has explored and demonstrated the importance of the first click for IIS. The authors have developed a deep framework, named First Click Attention Network (FCA-Net), which adds a simple module to the basic segmentation network to shift more attention to the first click.

With the popularization of DL and the high quality of the segmentation results obtained by DL-based IIS techniques with just a few clicks, research in this field remains very promising. However, a common problem of most DL-based techniques that some works [24, 179, 181, 186, 189] have tried to overcome is the that excellent results are achieved on objects present in the training set, while poor performance is achieved for unseen object classes. Furthermore, some authors [33] found that the large amounts of data needed to train a large number of model parameters makes the practicality of applying such models to real applications questionable. In contrast with this statement, and due to the emergence of increasingly available data sets and the use of data augmentation to boost the performance of the models [176], DL-based approaches have been successfully used in industry and smart factories [190].

3.6 Other methods

Other IIS techniques apply different strategies to propagate initial labels provided by the user to the rest of the image in order to extract the object of interest. In this subsection, we present some of these

methods.

IIS by matching attributed relational graphs (MARG) [191]: This method solves the segmentation problem by matching two graphs, one which represents the over-segmented image and one representing only the regions labelled by user strokes. The optimization step is performed based on deformed graphs [192].

Kernel propagation cut (KP-Cut) [193]: Based on kernel propagation [194], KP-Cut starts by generating a small-size seed-kernel matrix that is propagated into a full-kernel matrix for the whole image. FG–BG separation is effectively performed during the kernel propagation process.

Multiple instance learning cut (MIL-Cut) [90]: By imposing and exploring the property of tightness of the BB covering the object of interest, Mil-Cut solves the segmentation problem as a multiple instance learning task by generating positive bags from pixels of sweeping lines inside the BB and negative bags from pixels outside the box.

IIS via graph-based manifold ranking [195]: A graph-based semi-supervised learning framework ranks similarities of unlabeled data points to labeled ones by exploiting global and local consistency of all the data. A three-stage strategy is used to generate the segmentation. First, a k-regular graph is built to model spatial relationships. Then, user provided seeds are integrated to enhance the graph structure. Finally, to overcome instability due to sensitivity to the hyper-parameter, a content based locally adaptive kernel width parameter is used to provide graph edge weights.

Robust affinity diffusion (RAD) for IIS [33]: Unlike algorithms that construct segmentation models based on local affinity graphs such as GC and RW extended approaches, this work proposes iterative diffusion of the local affinity graph to explore global affinity across the whole image. The segmentation model is then constructed on the global graph via an energy function obtained by multiplication of an affinity matrix and a prior probability vector estimated from user seeds. Diffusion is performed as in Ref. [196], and the segmentation is obtained by solving a linear system.

Adaptive constraint propagation-cut (ACP-Cut) [43]: To effectively exploit the small quantity of user information, this work utilizes ACP for

semi-supervised kernel matrix learning which allows adaptive seed propagation and complexity reduction. Finally, for FG–BG separation a global k-means [197] algorithm is applied.

Parametric pseudo bound cuts (pPBC) [94]: This work proposed a new general pseudo-bound optimization paradigm for approximate iterative minimization of high-order and non-submodular binary energies. The pPBC algorithm improved the stateof-the-art in many energy minimization problems, in particular IIS using GrabCut [79].

IIS using sample reconstruction and Fisher's linear discriminative analysis (SR-FLDA) [46]: This IIS method generates image superpixels in an initial step and builds a dictionary using labelled ones. Then, a classification strategy using a discriminative projection matrix through the FLDA algorithm [198] is employed to group unlabeled superpixels into FG or BG by calculating their minimal norm.

Adaptive figure-ground classification [87, 88]: This work extracts the FG region from the BG using a BB in multiple steps. First, an adaptive mean-shift algorithm is applied to over-segment the image into regions. Then, BG and FG priors are explored to gradually refine these regions. Various distance measures and score functions are computed to generate multiple hypotheses and the final segmentation is obtained using a weighted combination or a voting scheme.

Constrained dominant sets for IIS [99, 199]: This work is based on the notion of *dominant sets* [200] which is a graph-theoretic concept and can be seen as a generalization of a maximal clique to edge-weighted graphs. This algorithm can deal with many types of constraints and input modality, and solves the IIS problem in a quadratic optimization framework.

IIS via cascaded metric learning [44]: The proposed approach starts by generating image superpixels and represents them in a space of extracted features (color, intensity, and texture). Positive and negative samples are selected using provided seeds and an optimal classification metric is computed to classify unlabeled samples. Then metric learning and the classification process are performed again using new training samples obtained from the previous iteration; this is repeated until convergence.

Click carving for IIS [103]: Whereas conventional IIS methods start from user provided seeds to segment objects, click carving enables accurate segmentation by precomputing possible segmentation hypotheses. Then, the user clicks on the object boundary to carve away erroneous parts, and the segmentation is refined iteratively until the user is satisfied.

IIS using label propagation through complex networks [201]: This provides a simple graph-based method for IIS with two stages. In the first stage, pixels are connected to their k-nearest neighbors to build a complex network with the small-world property to spread labels quickly. In the second stage, a regular network in grid format is used to refine the segmentation.

4 Classification of IIS methods based on processing level

With the great success of superpixel over-segmentation algorithms in many computer vision applications, a large number of IIS methods operate on image regions instead of pixels to achieve FG–BG separation while consuming less time and memory. For example, MSRM [75] segments the object of interest by merging homogenous regions obtained by initial over-segmentation using the meanshift algorithm [173]. Meanshift has also been adopted by other works such as ACP-Cut and Refs. [87, 88, 98, 202] to generate image entities in an initial step. The LS [59] method can be seen as a fast version of GC which processes image superpixels obtained by the watershed algorithm [141] instead of pixels. Similarly, the watershed has been employed in Ref. [191]. The active method in Ref. [113] partitions the input image into superpixels using the SLIC method [203] and takes them as input of their algorithm. Since SLIC over-segmentation is fast to compute, it has been used as a pre-processing step in many IIS methods [44, 90, 93, 104, 175, 204–206]. Other works such as Loose-Cut use a multiscale superpixel algorithm [207] to generate input regions and reduce the complexity of their systems. The method of Ref. [199] chose the ultrametric contour map algorithm (UCM) [208] to over-segment the input image. SR-FLDA and [209] followed [210] by applying two over-segmentation algorithms: meanshift and UCM [208]. As well

as classic IIS methods that operate directly on pixels and those that operate on superpixels (already mentioned above), some hybrid algorithms process both pixels and superpixels. For example, NC-Cut starts by computing connectedness between image regions then the final segmentation is found via a pixel-level optimization framework. In the same way, recent work in Ref. [211] proposes a coarse-to-fine method from region-level segmentation to pixel-level segmentation. In Ref. [212], an interactive multi-label segmentation approach combines both pixels and superpixels via robust multi-layer graph constraints. Based on graph theory, a hybrid IIS method using pixels and superpixels is presented in Ref. [31]. The authors in Ref. [202] use game theory to optimize the combinational energy functional related to both pixels and superpixels for IIS. Super-Cut stars its algorithm by applying the superpixel over-segmentation of Ref. [213], and then a pixel-level clustering is applied to obtain the final segmentation.

The main goal of using superpixels in IIS methods is to create visually meaningful entities while heavily reducing the number of primitives for subsequent processing steps [214], and therefore reducing the time consumed by the segmentation method [59]. The price of the speed of most superpixel-based methods compared to pixel-based ones is that these methods are very sensitive to the quality of the initial over-segmentation and may result in failure, especially in the presence of shadows, low-contrast edges and similar BG and FG regions [75]. To achieve a balanced compromise between speed and segmentation accuracy, it has been shown that a combination of pixel and superpixel level segmentation always outperforms a single level segmentation [83, 202].

5 Datasets

To evaluate IIS methods, many datasets containing images with associated ground-truth have been proposed in the literature. The most widely used datasets for the task of IIS are presented and described in Table 1. We note that some datasets were designed for other computer vision applications such as object recognition, salient object detection, and image co-segmentation; those used to evaluate IIS are also presented in Table 1.

Table 1 Different datasets used in the evaluation of IIS methods

Dataset	Description
Berkeley segmentation data set (BSDS) [215]	Large dataset of natural images with human annotation to serve as ground-truth. Benchmark to evaluate different contour detection and image segmentation algorithms. BSDS500 is the latest version and contains 500 natural images with their ground-truth.
Microsoft GrabCut dataset [79]	Dataset for IIS containing 50 color images. Ground-truth is stored as tri-maps identifying FG, BG, and mixed pixels (unknown).
MSRA10K dataset [216]	Dataset consisting of 10,000 images with pixel accurate salient object labeling. Proposed to evaluate salient object detection and segmentation methods.
Pascal VOC datasets [217]	Data sets from the VOC challenges[(1)] providing standardized image data sets for object class recognition; includes image annotations.
Alpha matting dataset [218]	High-quality matting database that extends the dataset of Ref. [195] by adding challenging images from natural scenes.
Icoseg dataset [115]	Large challenging dataset with 643 images proposed to evaluate image co-segmentation methods. It contains 38 groups of images from real scenes. Each group consists of instances of similar objects.
Weizmann dataset [219]	Segmentation evaluation dataset containing 200 gray level images with ground-truth segmentations.
Microsoft COCO [220]	Dataset for detecting and segmenting objects containing a vast collection of object instances with a total of 2,500,000 labeled instances in 328,000 images.
Cityscapes [221]	Dataset for semantic urban scene understanding containing 5000 finely annotated images of driving scenes, including 2975 images for training, 500 for validation, and 1525 for testing. Eight object classes are provided with per-instance annotation.
Kitti [222]	Suite of vision components for an autonomous driving platform. The object detection dataset contains 7481 training images annotated with 3D bounding boxes. A full description of the annotations can be found in the readme of the object development kit readme on the Kitti homepage[(2)]. Pixel-level annotation of a subset of images from the dataset is provided by Ref. [223].

(1) http://host.robots.ox.ac.uk/pascal/VOC/
(2) http://www.cvlibs.net/datasets/kitti/

6 Evaluation metrics

For objective quantitative evaluation of image segmentation techniques including IIS ones, a comparison between the results obtained by a segmentation method and a human-labeled segmentation—the ground-truth—must be performed by measuring their similarity according to some metric [1]. In this section, we report the commonly used different measures for such evaluation as described in various works [1, 6, 180, 198, 199]. Table 2 summarizes these measures by describing them, giving their mathematical formulation or a reference to it, and alternative names. In Table 2, we use S and G to designate the output result of the segmentation to be evaluated and the ground-truth, respectively.

7 General comparison

7.1 Evaluation of IIS methods in the literature

Some works in the literature focus on the evaluation of

IIS methods. For example, the authors in their survey of graph-based approaches to image segmentation [229], dedicated an experimental section to compare four selected IIS methods: IGC, Refs. [65], [59] and [50] using the segmentation accuracy and the number of interactions required for segmentation. They concluded in their study that IGC outperforms all the other approaches by producing a high segmentation accuracy while requiring the least amount of user interaction. Earlier, an evaluation study of four popular algorithms: IGC, Refs. [27], [16], and [42] was presented in Ref. [6]. In this study, by exploring two measures to compute object and boundary accuracy, the experiments showed the most effective techniques to be IGC and Ref. [42]. More extensive evaluation of IIS techniques has been reported in Ref. [7] including ICG, CAC, MSRM, RWR, and MARG. Taking into consideration the accuracy and the robustness of segmentation results, the experiments concluded that each method has its own limitations such as sensitivity to user inputs and noise, sensitivity to the initial over-segmentation, etc. For more details of the results of

Table 2 Different measures used in the literature to evaluate IIS methods

Metric	Description	Formulae	Alias
Error rate (ERR)	Fraction or percentage of wrongly classified pixels in ROI	$Err = |S \cup G - G|/(nb.\text{of pixels in ROI})$	Mis-segmentation rate
True Positive Rate (TPR) False Negative Rate (FNR) False Positive Rate (FPR) True Negative Rate (TNR) Precision (P) F-measure (F)	Measures of FG and BG assignment accuracy	$TPR = |S \cap G|/|G|$ $FNR = 1 - TPR$ $FPR = |S \cup G - G|/|G|$ $TNR = 1 - FPR$ $P = |S \cap G|/|S|$ $F = (2 \times R \times P)/(R + P)$	Recall F-score
Intersection over Union (IoU)	Number of correctly labelled FG/BG pixels to the number of pixels labelled as FG or BG in either G or S	$IoU = |S \cap G|/|S \cup G|$	Jaccard index, overlap coefficient
Dice Similarity Coefficient (DSC)	Overlap between the two segmented object volumes G and S	$DSC = 2 \times |S \cap G|/(|S| + |G|)$	—
Rand Index (RI)	Fraction of pixel pairs whose labels are consistent between G and S	[224]	—
Global Consistency Error (GCE)	Measure of the extent to which one segmentation can be viewed as a refinement of the other	[215]	—
Boundary Displacement Error (BDE)	Average displacement error of boundary pixels between G and S	[225]	—
Variation of Information (VI)	Measure of the distance between the two segmentations G and S using average conditional entropy	[226]	—
Segmentation Covering (SC)	Quality measure of the covering of the segmentation S by G	[208]	—
Modified Hausdorff Distance (MHD)	Measure of the displacement between the segmentation S and G	[227]	—
Number of Clicks ($NoCs$)	Mean number of clicks required to achieve a certain IoU	[23]	—
F-Boundary Score (FB)	Precision/recall for the boundaries of G and S	[228]	—

these evaluation studies, the reader can refer to the original works Refs. [229], [6], and [7].

7.2 General comparison of recent works

In this paper, we select some recent works (about 40 papers) that have been published in the four last years, i.e., between 2016 and 2020, and based on the results reported by their authors, we give a general comparison of these methods grouped by the adopted methodology.

GC-based IIS models are very popular in practical applications because of their solid theoretical foundation and good performance. Thus, the research into improving and extending GC methods is always in progress. For example to overcome the drawbacks of GC with ROI-based methods, i.e., GrabCut extents, Super-Cut [83], NC-Cut [8], LooseCut [81], and

Ref. [211] have been proposed. Loose-Cut handles robustly the cases of images with a loose input BB while other methods fail because of the inaccurate FG appearance model estimation in this case. Super-Cut tackles also the issue of loose user ROI definition while consuming less time than Loose-Cut. NC-Cut computes an NC map to reduce the sensitivity of user input by exploiting topological properties of the image. All these methods achieve good performance against other methods [79, 80, 90, 94] for different datasets using several evaluation metrics. Point-Cut and the works in Refs. [64] and [150] are other GC-based methods but they are seed-based models. The work in Ref. [64] outperforms many methods including ICG, the original RW, Refs. [58] and [37] as well as other recent approaches such as Refs. [161] and [202]. Furthermore, the results obtained by Point-Cut

are very competitive despite the use of only one input user seed point. The authors in Ref. [150] exploited the advantages of both of geodesic GC [54] and One-Cut [80] with little additional runtime and reduced user interaction.

RW-based methods present another category of widely used approaches to perform IIS based on graph theory. Belonging to this category, SRW [161], Refs. [100] and [165] have been recently proposed to boost the performance of RW models. SRW focuses on the improvement of the segmentation of complex textured images which other RW-based methods [50, 160, 230] fail to solve. However the time consuming of SRW is greater than for the compared methods. The model presented in Ref. [100] achieves good results in spite of the presence of inaccurate initial labels, but some failures occur when initial FG and BG labels have similar color distributions. Recently Ref. [165], a feedback system allows the computer to exploit and understand the intent of limited user input using an iterative boundary.

To solve IIS for deformable objects in the presence of challenges, some recent contour-based methods have been proposed. For instance, the model in Ref. [122] tackles the case of images with intensity inhomogeneity and a high level of noise, producing accurate segmentation compared to other level set methods. In the same way, the interactive ACM in Ref. [74] generates better segmentation results for objects in heterogeneous and cluttered images than CAC [72] and GrabCut. Other recent work [136] has demonstrated its superiority over other well-known ACM methods with respect to the balance of segmentation accuracy and speed, in the presence of different types of noise and with the use of various initializations.

As described in Section 3.5, recent approaches use CNNs to perform effective IIS. They achieve the state of the art in both accuracy and limited user interaction. The first DL-based IIS method [23] requires only a few positive and negative clicks to mark FG and BG respectively. Its results demonstrate superiority by requiring the fewest *NoCs* to achieve a certain *IoU* accuracy when compared to other conventional works on different datasets. Following Ref. [23], many DL-based approaches

have been presented to extent its principles and improve its performance using different strategies such as region proposals [179], extreme points [17], iterative training [22], coupled CNN architecture [24, 180], backpropagation refinement [106, 186], polygon [182, 183], and curve-graph [184] based networks, deep extreme level sets [185], and other strategies [21, 107, 181, 188, 189]. In general, all DL-based IIS approaches have the ability to perceive complex global and local image features and achieve better performance than conventional methods. Although direct comparison between the different DL-based techniques cannot be performed due to differences in frameworks, datasets and hardware used, the results reported in two recent works [186, 189] reached the state-of-the-art for the datasets used.

In addition to the abovementioned methods, many recent works have been published in the last years and achieved promising results for the task of IIS.

Table 3 presents a general comparison of recent IIS methods grouped by the methodology followed. For each method, the key words, the level of processing: pixels (Px) or superpixels (Spx), the type of user interaction, the evaluation metrics, the advantages, and the limitations (if known) are given.

8   Available resources

The aim of this section is to provide resources such as links to available code and software for IIS research. Table 4 presents the collected resources grouped by methodology. We note that the last access of the links is the time of writing this article.

9   Conclusions

In this paper, we have reviewed the literature covering IIS methods. We have classified them according to different criteria: the type of human interaction, the adopted methodology, and the level of processing. We focused on recent works including DL-based methods. We have given a general comparison of them and collected available resources including datasets, evaluation metrics, and on line code sources. The main goal of this work is to present a recent survey to serve as reference to researchers in the field of IIS.

Table 3 General comparison of recent IIS works. Evaluation metrics are described in Table 2

Methodology	Work	Year	Keywords	Processing level	User interaction	Evaluation metrics	Advantages	Limitations
Contour-based methods	[122]	2018	Level set; multi-phase formulation	Px	Seeds	IoU	Successful segmentation of images with intensity inhomogeneity and a high level of noise	—
	[74]	2018	Level set; kernel descriptor	Px	Seeds	RI, CGE, VoI, SC	Robustness in heterogeneous and cluttered images	—
	[136]	2020	Intensity inhomogeneity; Bayesian criterion; Markov random	Px	Seeds	IoU, runtime	Robustness to different types of noise and initializations; superior to other well-known ACM with respect to the balance of segmentation accuracy and speed	Need to experimentally set several control parameters
GC-based methods	Point-Cut [82]	2016	One point; object proposals	Px/Spx	Seeds	DSC	Competitive results using only one input seed point	Dependence on the quality of object proposals; failure in case of highly textured FG objects
	NC-Cut [8]	2016	NC; topology	Px/Spx	ROI	IoU, Err, RI, GCE, BDE	Much less sensitive to a loose ROI than other methods	Failure when an image has high measure of average indeterminacy
	Loose-Cut [81]	2017	Label consistency; global similarity constraint	Px/Spx	ROI	Err, F-measure	Outperforms other methods given a loose BB	—
	Super-Cut [83]	2017	Superpixel clustering; local similarity constraint	Px/Spx	ROI	IoU, F-measure	Reduced time cost due to the one cut segmentation; accurate labelling due to clustering of pixels and superpixels	Failure when the input BB is much tighter than the ground-truth
	[211]	2018	Two phase; region-level- and pixel-level-segmentation	Px/Spx	ROI	Err	Improvement of GrabCut extents methods using two phases	Sensitivity of region-level segmentation phase to the tightness of the input ROI
	[64]	2018	Diffusion likelihood; superpixel-based grouping	Px/Spx	Seeds	IoU, Err, RI, VoI, runtime	Accurate prediction of label probabilities from few seeds; enforcing continuity for the object segmentation by using Spx grouping	
	[150]	2019	Geodesic GC; one cut	Px	Seeds	IoU, DSC, Err	Improvement of geodesic GC and One cut with little additional runtime and low user interaction	—
RW-based methods	SRW [161]	2016	RW; complex texture	Px	Seeds	Err, IoU, runtime	Accurate segmentation of images with complex texture	Slower than other methods

(Continued)

Methodology	Work	Year	Keywords	Processing level	User interaction	Evaluation metrics	Advantages	Limitations
	[100]	2017	Structure-aware labeling; occurrence and co-occurrence probability	Px	Seeds	DSC	Robust to inaccurate initial labels	Failure when initial labels have similar color distributions
	[165]	2018	Iterative boundary RW; feed-back system	Px	Seeds	Err	Feedback system that allows the computer to exploit and understand the intent of limited user input	—
	[23]	2016	FCN; GC	Px	Seeds	IoU, NoCs	First work solving IIS using DL model; high quality segmentations compared to other methods for few clicks	Failure to correct the generated prediction by producing similar outputs regardless of additional clicks added
	[179]	2017	ROI proposals	Px	Seeds	IoU, NoCs	Reduction of the amount of user interaction required for accurate segmentation	Failure in cases of occluded or hairy objects
	[17]	2017	Extreme points	Px	Seeds	Err, NoCs	Reduced user input; achieves state-of-the-art results for various benchmarks and datasets	Restriction to using four clicks to generate segmentations; difficult to refine results with additional clicks when the quality of the segmentation is low
Deep learning-based methods	[182]	2017	Polygon-RNN	Px	Seeds	IoU, NoCs	Facilitation of the annotation task by treating object segmentation as a polygon prediction task; easy to incorporate user corrections to improve segmentation	Output with low resolution (28×28) producing blocky polygons for large objects
	[22]	2018	Iterative training	Px	Seeds	IoU, NoCs	Robust to variations in clicks; significant improvement over state of the art; needs 0.2 clicks fewer than DEXTR [17]	Some failures to remove unwanted objects with a few clicks
	[21]	2018	Seed generation; deep Reinforcement learning	Px	Seeds	IoU, NoCs	Significant reduction of human effort for the IIS task	—
	[24]	2018	Coupled CNN architecture; clicks	Px	Seeds	IoU, Err, NoCs, runtime	2.3 times faster than [21]; reduction of user-guided segmentation to a forward pass in a CNN	—

(Continued)

Methodology	Work	Year	Keywords	Processing level	User interaction	Evaluation metrics	Advantages	Limitations
	[180]	2018	FCN; two-stream network	Px	Seeds	IoU, NoCs	Improved performance over single-stream networks	—
	[183]	2018	Polygon-RNN++; RNN; reinforcement learning	Px	Seeds	IoU, NoCs	Effective tool for fast and accurate object annotation in challenging datasets; improvement of the polygon-RNN model [182]	Annotation of a single object and failure in cases of multi-component objects
	[106]	2019	Backpropagating refinement scheme (BRS)	Px	Seeds	IoU, NoCs	Better performance for hard cases due to the use of optimization-based backpropagation refinement; possibility to apply backpropagating refinement to other computer vision tasks	Computational cost per click is high due to running forward and backward passes through a deep network multiple times; not practical for many end user scenarios
	[107]	2019	Mask R-CNN	Px	Seeds	IoU, NoCs	Efficient for both single object and full image segmentation	—
	[188]	2019	Scale-diversity; non-maximum suppression	Px	Seeds	IoU, NoCs, runtime	Capability to produce multiple diverse and semantically meaningful segmentations	—
	[181]	2019	Content-aware guidance maps	Px/Spx	Seeds	IoU, NoCs	Using guidance maps, even the most basic FCNs are able to outperform state-of-the-arts DL-based methods	Failure for some challenging scenarios (objects with holes, occlusion)
	[184]	2019	Curve-GCN	Px	Seeds	IoU, NoCs, FB	Efficient annotation for both line-based and curved objects, by either polygons or splines; 100x faster than Polygon-RNN++ [183]	
	[185]	2019	Deep extreme level set	Px	Seeds	IoU, NoCs, FB	Robust boundary object extraction even in the presence of noise	—
	[189]	2020	First click attention network; FCN	Px	Seeds	IoU, NoCs	Superiority of the proposed method for various datasets	Failure when the desired object may not be clicked by the user due to structure or occlusion
	[186]	2020	Feature-BRS	Px	Seeds	IoU, NoCs, runtime	Improvement of the original BRS by running forward and backward passes just for a small part of a network; high performance in terms of speed and accuracy	—

(Continued)

Methodology	Work	Year	Keywords	Processing level	User interaction	Evaluation metrics	Advantages	Limitations
	ACP-Cut [43]	2016	Discriminative learning; seed propagation	Spx	Seeds	*TPR, TNR*, runtime	Improved segmentation results	Failure when Spx share similar characteristics in both FG and BG
	[202]	2016	Multi-layer graph; non-parametric learning; game theory	Px/Spx	Seeds	*Err, IoU*, runtime	Improved segmentation accuracy by exploring relationships between Px/Spx and labels	Less robust for practical applications because of the requirement to set several parameters experimentally
	Dominant Sets [199]	2016	Dominant sets; quadratic optimization	Spx	Seeds/ROI	*IoU, Err, DSC*	Flexible to input modalities (BB, loose BB, scribbles only on the ROI and scribbles with error)	Pre-processing step using the original UCM [180] is very slow for a practical tool
	[46]	2017	Samples reconstruction; FLDA	Spx	Seeds	*Err, TPR, FPR*	Reduction of sensitivity to number of seeds and locations, captures long range grouping cues	Dependete on high quality superpixel over-segmentation; failure when image contains hairy objects or other objects with complex edges
Others	[33]	2018	Global affinity; graph diffusion	Px	Seeds	*Err*, runtime	Better results when FG and BG have similar appearances than for other local affinity graph-based models (GC and RW-based methods)	Less robust for practical applications because of need to set the diffusion control parameter experimentally
	[104]	2018	Manifold ranking; machine assisted	Spx	Seeds/ROI	*DCS*	Flexible to different user interactions (ROI/seeds/true-or-false feed-back); practical for real-time use	—
	[231]	2019	Likelihood learning; probabilistic estimation	Px	Seeds	*RI, VoI, Err*, seed quantity and location	Improvement over results of latest IIS methods including some DL-based one	—
	[201]	2019	Label propagation; complex networks	Px	Seeds	*Err*, runtime	Simplicity; low time; multi-class	
	[103]	2019	Region proposals; boundary click; objectness score; user votes	Spx	Seeds	*IoU, NoCs*, runtime	Suitable for both static and dynamic scenes; only a few clicks are needed for accurate object segmentation; competitive results against a recent DL-based method [183]	Dependent on the quality of the generated region proposals

Table 4 Available resources in the field of IIS

GC-based methods

[29, 49, 80, 152]	https://vision.cs.uwaterloo.ca/code
[79]	https://GrabCut.weebly.com/code.html
[59]	http://www.cs.cmu.edu/~mohitg/segmentation.htm
[55]	http://www.robots.ox.ac.uk/~vgg/research/iseg/
[19]	Code available upon request to author
[84]	https://mmcheng.net/densecut/
[70]	par http://coopcut.berkeleyvision.org/
[28]	http://www1.icsi.berkeley.edu/~cnieuwe/code/multisegpub.zip
[69]	https://github.com/meng-tang/KernelCut
[71]	https://github.com/aosokin/coopCuts_CVPR2013
[52, 232]	https://github.com/meng-tang/KernelCut_ICCV15/
[205]	https://github.com/Borda/pyImSegm

RW-based methods

[50, 111]	http://leogrady.net/software/
[62]	http://www3.ntu.edu.sg/home/asjfcai/Tool_ImageComposite.rar
[58]	https://sites.google.com/site/wallacecoc/publications/lapcoordinates.zip?attredirects=0
[48]	https://github.com/ahmdtaha/SL-icip
[66, 162]	https://github.com/christosbampis/NRW_release
[161]	https://github.com/shenjianbing/SubMarkov-Random-Walk-for-Image-Segmentation-

RM-based methods

[75]	http://www4.comp.polyu.edu.hk/~cslzhang/MSRM/PR_MSRM_website.htm
[66]	http://cvsp.cs.ntua.gr/research/GraphClustering/
[206]	https://github.com/Borda/pyImSegm

Contour-based methods

[12]	http://www.cs.technion.ac.il/~cs234326/projects/scissors/Scissors.html
[126]	http://uk.mathworks.com/matlabcentral/fileexchange/47586-soac-zip
[72]	http://www.ntu.edu.sg/home/asjfcai/Anh_Image%20segmentation.zip
[122]	Code available upon request to author
[123]	https://www.liverpool.ac.uk/~cmchenke/softw/TV-ML-Noise-2006.htm
[129, 130]	https://github.com/julien-mille/comb-pgeo-paths
[26]	http://pages.cs.wisc.edu/~jiaxu/projects/euler-seg/euler-seg-src.zip

DL-based methods

[17]	http://www.vision.ee.ethz.ch/~cvlsegmentation/dextr/
[22]	https://github.com/sabarim/itis
[106]	https://github.com/wdjang/BRS-Interactive_segmentation
[186]	https://github.com/saic-vul/fbrs_interactive_segmentation
[182, 183]	https://github.com/fidler-lab/polyrnn-pp-pytorch
[184]	https://github.com/fidler-lab/curve-gcn
[185]	https://github.com/fidler-lab/delse

Others

[191]	http://structuralsegm.sourceforge.net/
[110]	http://ilastik.org/
[90]	Demo available upon request to author
[35]	http://www.eng.alexu.edu.eg/~mtorki/Research/Codes/MMFT%20-%20BMVC%202015.zip
[34]	https://sites.google.com/site/feipingnie/file/slrw.rar?attredirects=0
[94]	https://github.com/meng-tang/PBO/
[30]	https://sites.google.com/site/feipingnie/file/interactive_image_segmentation_for_web.zip?attredirects=0
[195]	https://codeload.github.com/valhongli/GMR_Seg/zip/master
[201]	https://github.com/fbreve/LPCN

References

[1] Zhu, H. Y.; Meng, F. M.; Cai, J. F.; Lu, S. J. Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation. *Journal of Visual Communication and Image Representation* Vol. 34, 12–27, 2016.

[2] Jain, S.; Laxmi, V. Color image segmentation techniques: A survey. In: *Proceedings of the International Conference on Microelectronics, Computing & Communication Systems. Lecture Notes in Electrical Engineering, Vol. 453.* Nath, V. Ed. Springer Singapore, 189–197, 2017.

[3] Yu, H. S.; Yang, Z. G.; Tan, L.; Wang, Y. N.; Sun, W.; Sun, M. G.; Tang, Y. D. Methods and datasets on semantic segmentation: A review. *Neurocomputing* Vol. 304, 82–103, 2018.

[4] Suri, J. S.; Setarehdan, S. K.; Singh, S. *Advanced Algorithmic Approaches to Medical Image Segmentation: State-of-the-Art Applications in Cardiology, Neurology, Mammography and Pathology.* Springer-Verlag London, 2001.

[5] Chen, X. J.; Pan, L. J. A survey of graph cuts/graph search based medical image segmentation. *IEEE Reviews in Biomedical Engineering* Vol. 11, 112–124, 2018.

[6] McGuinness, K.; O'Connor, N. E. A comparative evaluation of interactive segmentation algorithms. *Pattern Recognition* Vol. 43, No. 2, 434–444, 2010.

[7] He, J.; Kim, C. S.; Kuo, C. C. J. Interactive image segmentation techniques. In: *Interactive Segmentation Techniques. SpringerBriefs in Electrical and Computer Engineering.* Springer Singapore, 17–62, 2013.

[8] Xian, M.; Zhang, Y.; Cheng, H.-D.; Xu, F.; Ding, J. Neutro-connectedness cut. *IEEE Transactions on Image Processing* Vol. 25, No. 10, 4691–4703, 2016.

[9] Chen, D. J.; Chen, H. T.; Chang, L. W. Interactive segmentation from 1-bit feedback. In: *Computer Vision–ACCV 2016. Lecture Notes in Computer Science, Vol. 10111.* Lai, S. H.; Lepetit, V.; Nishino, K.; Sato, Y. Eds. Springer Cham, 261–274, 2017.

[10] Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431–3440, 2015.

[11] Yao, R.; Lin, G.; Xia, S.; Zhao, J.; Zhou, Y. Video object segmentation and tracking: A survey *arXiv preprint* arXiv:1904.09172, 2019.

[12] Mortensen, E. N.; Barrett, W. A. Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing* Vol. 60, No. 5, 349–384, 1998.

[13] Falcão, A. X.; Udupa, J. K.; Samarasekera, S.; Sharma, S.; Hirsch, B. E.; de A Lotufo, R. User-steered image segmentation paradigms: Live wire and live lane. *Graphical Models and Image Processing* Vol. 60, No. 4, 233–260, 1998.

[14] Falcao, A. X.; Udupa, J. K.; Miyazawa, F. K. An ultra-fast user-steered image segmentation paradigm: Live wire on the fly. *IEEE Transactions on Medical Imaging* Vol. 19, No. 1, 55–62, 2000.

[15] Miranda, P. A. V.; Falcao, A. X.; Spina, T. V. Riverbed: A novel user-steered image segmentation method based on optimum boundary tracking. *IEEE Transactions on Image Processing* Vol. 21, No. 6, 3042–3052, 2012.

[16] Adams, R.; Bischof, L. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 16, No. 6, 641–647, 1994.

[17] Maninis, K.-K.; Caelles, S.; Pont-Tuset, J.; Van Gool, L. Deep extreme cut: From extreme points to object segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 616–625, 2018.

[18] Vezhnevets, V.; Konouchine, V. GrowCut: Interactive multi-label ND image segmentation by cellular automata. In: Proceedings of Graphicon, 150–156, 2005.

[19] Xian, M.; Xu, F.; Cheng, H. D.; Zhang, Y.; Ding, J. EISeg: Effective interactive segmentation. In: Proceedings of the 23rd International Conference on Pattern Recognition, 1982–1987, 2016.

[20] Meena, S.; Palaniappan, K.; Seetharaman, G. User driven sparse point-based image segmentation. In: Proceedings of the IEEE International Conference on Image Processing, 844–848, 2016.

[21] Song, G.; Myeong, H.; Lee, K. M. SeedNet: Automatic seed generation with deep reinforcement learning for robust interactive segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1760–1768, 2018.

[22] Mahadevan, S.; Voigtlaender, P.; Leibe, B. Iteratively trained interactive segmentation. In: Proceedings of the British Machine Vision Conference, 212, 2018.

[23] Xu, N.; Price, B.; Cohen, S.; Yang, J.; Huang, T. S. Deep interactive object selection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 373–381, 2016.

[24] Li, Z.; Chen, Q.; Koltun, V. Interactive image segmentation with latent diversity. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 577–585, 2018.

[25] Fan, M.; Lee, T. C. M. Variants of seeded region growing. *IET Image Process* Vol. 9, No. 6, 478–485, 2014.

[26] Xu, J.; Collins, M. D.; Singh, V. Incorporating topological constraints within interactive segmentation and contour completion via discrete calculus. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013.

[27] Friedland, G.; Jantz, K.; Rojas, R. Siox: Simple interactive object extraction in still images. In: Proceedings of the 7th IEEE International Symposium on Multimedia, 253–260, 2005.

[28] Nieuwenhuis, C.; Cremers, D. Spatially varying color distributions for interactive multilabel segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 35, No. 5, 1234–1247, 2013.

[29] Stuhmer, J.; Schroder, P.; Cremers, D. Tree shape priors with connectivity constraints using convex relaxation on general graphs. In: Proceedings of the IEEE International Conference on Computer Vision, 2336–2343, 2013.

[30] Xiang, S. M.; Nie, F. P.; Zhang, C. X.; Zhang, C. S. Interactive natural image segmentation via spline regression. *IEEE Transactions on Image Processing* Vol. 18, No. 7, 1623–1632, 2009.

[31] Long, J. W.; Feng, X.; Zhu, X. F.; Zhang, J. X.; Gou, G. L. Efficient superpixel-guided interactive image segmentation based on graph theory. *Symmetry* Vol. 10, No. 5, 169, 2018.

[32] Duchenne, O.; Audibert, J.-Y.; Keriven, R.; Ponce, J.; Ségonne, F. Segmentation by transduction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1–8, 2008.

[33] Wang, T.; Yang, J.; Sun, Q.; Ji, Z.; Fu, P.; Ge, Q. Global graph diffusion for interactive object extraction. *Information Sciences* Vols. 460–461, 103–114, 2018.

[34] Xiang, S. M.; Pan, C. H.; Nie, F. P.; Zhang, C. S. Interactive image segmentation with multiple linear reconstructions in windows. *IEEE Transactions on Multimedia* Vol. 13, No. 2, 342–352, 2011.

[35] Meshry, M.; Taha, A.; Torki, M. Multi-modality feature transform: An interactive image segmentation approach. In: Proceedings of the British Machine Vision Conference, 2015.

[36] Ren, Y.; Chua, C. S.; Ho, Y. K. Statistical background modeling for non-stationary camera. *Pattern Recognition Letters* Vol. 24, Nos. 1–3, 183–196, 2003.

[37] Kim, T. H.; Lee, K. M.; Lee, S. U. Nonparametric higher-order learning for interactive segmentation. In: Proceedings of the Computer Vision and Pattern Recognition, 3201–3208, 2010.

[38] Bai, J.; Wu, X. Error-tolerant scribbles based interactive image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 392–399, 2014.

[39] Zhang, J.; Tang, Z. H.; Gui, W. H.; Chen, Q.; Liu, J. P. Interactive image segmentation with a regression based ensemble learning paradigm. *Frontiers of Information Technology & Electronic Engineering* Vol. 18, No. 7, 1002–1020, 2017.

[40] Subr, K.; Paris, S.; Soler, C.; Kautz, J. Accurate binary image selection from inaccurate user input. *Computer Graphics Forum* Vol. 32, No. 2pt1, 41–50, 2013.

[41] Kohli P.; Ladicky, L.; Torr P. H. S. Robust higher order potentials for enforcing label consistency. *International Journal of Computer Vision* Vol. 82, 302–324, 2009.

[42] Salembier, P.; Garrido, L. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Transactions on Image Processing* Vol. 9, No. 4, 561–576, 2000.

[43] Jian, M.; Jung, C. Interactive image segmentation using adaptive constraint propagation. *IEEE Transactions on Image Processing* Vol. 25, No. 3, 1301–1311, 2016.

[44] Li, W.; Shi, Y.; Yang, W.; Wang, H.; Gao, Y. Interactive image segmentation via cascaded metric learning. In: Proceedings of the IEEE International Conference on Image Processing, 2900–2904, 2015.

[45] Cheng, D. S.; Murino, V.; Figueiredo, M. Clustering under prior knowledge with application to image segmentation. In: Proceedings of the 19th International Conference on Neural Information Processing Systems, 401–408, 2006.

[46] Luo, L.; Wang, X.; Hu, S.; Hu, X.; Chen, L. Interactive image segmentation based on samples reconstruction and FLDA. *Journal of Visual Communication and Image Representation* Vol. 43, 138–151, 2017.

[47] Mansilla, L. A. C.; Miranda, P. A. V. Oriented image foresting transform segmentation: Connectivity constraints with adjustable width. In: Proceedings of the 29th SIBGRAPI Conference on Graphics, Patterns and Images, 289–296, 2016.

[48] Taha, A.; Torki, M. Seeded laplaican: An eigenfunction solution for scribble based interactive image segmentation. *arXiv preprint* arXiv:1702.00882, 2017.

[49] Boykov, Y. Y.; Jolly, M.-P. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In: Proceedings of the 8th IEEE

International Conference on Computer Vision, Vol. 1, 105–112, 2001.

[50] Grady, L. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 11, 1768–1783, 2006.

[51] Wang, T.; Ji, Z. X.; Sun, Q. S.; Chen, Q.; Han, S. D. Image segmentation based on weighting boundary information via graph cut. *Journal of Visual Communication and Image Representation* Vol. 33, 10–19, 2015.

[52] Tang, M.; Marin, D.; Ayed, I. B.; Boykov, Y. Kernel cuts: MRF meets kernel & spectral clustering. *arXiv preprint* arXiv:1506.07439, 2015.

[53] Bai, X.; Sapiro, G. A geodesic framework for fast interactive image and video segmentation and matting. In: Proceedings of the IEEE 11th International Conference on Computer Vision, 1–8, 2007.

[54] Price, B. L.; Morse, B.; Cohen, S. Geodesic graph cut for interactive image segmentation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3161–3168, 2010.

[55] Gulshan, V.; Rother, C.; Criminisi, A.; Blake, A.; Zisserman, A. Geodesic star convexity for interactive image segmentation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3129–3136, 2010.

[56] Gong, Y.; Xiang, S.; Wang, L.; Pan, C. Fine-structured object segmentation via edge-guided graph cut with interaction simplification. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1801–1805, 2016.

[57] Vicente, S.; Kolmogorov, V.; Rother, C. Graph cut based image segmentation with connectivity priors. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1–8, 2008.

[58] Casaca, W.; Nonato, L. G.; Taubin, G. Laplacian coordinates for seeded image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 384–391, 2014.

[59] Li, Y.; Sun, J.; Tang, C.-K.; Shum, H.-Y. Lazy snapping. *ACM Transactions on Graphics* Vol. 23, No. 3, 303–308, 2004.

[60] Sung, M.-C.; Chang, L.-W. Using multi-layer random walker for image segmentation. In: Procedings of the International Workshop on Advanced Image Technology, 1–4, 2018.

[61] Wang, J. Discriminative Gaussian mixtures for interactive image segmentation. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 601–604, 2007.

[62] Yang, W. X.; Cai, J. F.; Zheng, J. M.; Luo, J. B. User-friendly interactive image segmentation through unified combinatorial user inputs. *IEEE Transactions on Image Processing* Vol. 19, No. 9, 2470–2479, 2010.

[63] Shi, R.; Liu, Z.; Xue, Y.; Zhang, X. Interactive object segmentation using iterative adjustable graph cut. In: Proceedings of the Visual Communications and Image Processing, 1–4, 2011.

[64] Wang, T.; Ji, Z. X.; Sun, Q. S.; Chen, Q.; Ge, Q.; Yang, J. Diffusive likelihood for interactive image segmentation. *Pattern Recognition* Vol. 79, 440–451, 2018.

[65] Peng, B.; Zhang, L.; Zhang, D.; Yang, J. Image segmentation by iterated region merging with localized graph cuts. *Pattern Recognition* Vol. 44, Nos. 10–11, 2527–2538, 2011.

[66] Bampis, C. G.; Maragos, P.; Bovik, A. C. Graph-driven diffusion and random walk schemes for image segmentation. *IEEE Transactions on Image Processing* Vol. 26, No. 1, 35–50, 2017.

[67] Zhang, J.; Zheng, J.; Cai, J. A diffusion approach to seeded image segmentation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2125–2132, 2010.

[68] Ducournau, A.; Bretto, A. Random walks in directed hypergraphs and application to semi-supervised image segmentation. *Computer Vision and Image Understanding* Vol. 120, 91–102, 2014.

[69] Tang, M.; Marin, D.; Ayed, I. B.; Boykov, Y. Normalized cut meets MRF. In: *Computer Vision–ECCV 2016. Lecture Notes in Computer Science, Vol. 9906*. Leibe, B.; Matas, J.; Sebe, N.; Welling, M. Eds. Springer Cham, 748–765, 2016.

[70] Jegelka, S.; Bilmes, J. Submodularity beyond submodular energies: Coupling edges in graph cuts. In: Proceedings of the Computer Vision and Pattern Recognition, 1897–1904, 2011.

[71] Kohli, P.; Osokin, A.; Jegelka, S. A principled deep random field model for image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1971–1978, 2013.

[72] Nguyen, T. N. A.; Cai, J.; Zhang, J.; Zheng, J. Robust interactive image segmentation using convex active contours. *IEEE Transactions on Image Processing* Vol. 21, No. 8, 3734–3743, 2012.

[73] Liu, D.; Xiong, Y.; Shapiro, L.; Pulli, K. Robust interactive image segmentation with automatic boundary refinement. In: Proceedings of the 17th IEEE International Conference on Image Processing, 225–228, 2010.

[74] Li, H.; Gong, M.; Miao, Q.; Wang, B. Interactive active contour with kernel descriptor. *Information Sciences* Vol. 450, 53–72, 2018.

[75] Ning, J. F.; Zhang, L.; Zhang, D.; Wu, C. K. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition* Vol. 43, No. 2, 445–456, 2010.

[76] Zhou, C. B.; Wu, D. M.; Qin, W. H.; Liu, C. C. An efficient two-stage region merging method for interactive image segmentation. *Computers & Electrical Engineering* Vol. 54, 220–229, 2016.

[77] Vallin Spina, T.; de Miranda, P. A. V.; Xavier Falcao, A. Hybrid approaches for interactive image segmentation using the live markers paradigm. *IEEE Transactions on Image Processing* Vol. 23, No. 12, 5756–5769, 2014.

[78] Falcao, A. X.; Stolfi, J.; de Alencar Lotufo, R. The image foresting transform: Theory, algorithms, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 26, No. 1, 19–29, 2004.

[79] Rother, C.; Kolmogorov, V.; Blake, A. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics* Vol. 23, No. 23, 309–314, 2004.

[80] Tang, M.; Gorelick, L.; Veksler, O.; Boykov, Y. Grabcut in one cut. In: Proceedings of the IEEE International Conference on Computer Vision, 1769–1776, 2013.

[81] Yu, H.; Zhou, Y.; Qian, H.; Xian, M.; Wang, S. Loosecut: Interactive image segmentation with loosely bounded boxes. In: Proceedings of the IEEE International Conference on Image Processing, 3335–3339, 2017.

[82] Oh, C.; Ham, B.; Sohn, K. Point-cut: Interactive image segmentation using point supervision. In: *Computer Vision–ACCV 2016. Lecture Notes in Computer Science, Vol. 10111.* Lai, S. H.; Lepetit, V.; Nishino, K.; Sato, Y. Eds. Springer Cham, 229–244, 2017.

[83] Wu, S. Q.; Nakao, M.; Matsuda, T. SuperCut: Superpixel based foreground extraction with loose bounding boxes in one cutting. *IEEE Signal Processing Letters* Vol. 24, No. 12, 1803–1807, 2017.

[84] Cheng, M. M.; Prisacariu, V. A.; Zheng, S.; Torr, P. H. S.; Rother, C. DenseCut: Densely connected CRFs for realtime GrabCut. *Computer Graphics Forum* Vol. 34, No. 7, 193–201, 2015.

[85] Rajchl, M.; Lee, M. C. H.; Oktay, O.; Kamnitsas, K.; Passerat-Palmbach, J.; Bai, W.; Damodaram, M.; Rutherford, M. A.; Hajnal, J. V.; Kainz, B. et al. Deepcut: Object segmentation from bounding box annotations using convolutional neural networks. *IEEE Transactions on Medical Imaging* Vol. 36, No. 2, 674–683, 2017.

[86] Xu, N.; Price, B.; Cohen, S.; Yang, J.; Huang, T. Deep grabcut for object selection. In: Proeedings of the 28th British Machine Vision Conference, 2017.

[87] Chen, Y. S.; Chan, A. B.; Wang, G. Adaptive figure-ground classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 654–661, 2012.

[88] Chen, Y. S.; Chan, A. B. Enhanced figure-ground classification with background prior propagation. *IEEE Transactions on Image Processing* Vol. 24, No. 3, 873–885, 2015.

[89] Lempitsky, V. S.; Kohli, P.; Rother, C.; Sharp, T. Image segmentation with a bounding box prior. In: Proceedings of the IEEE 12th International Conference on Computer Vision, 277–284, 2009.

[90] Wu, J.; Zhao, Y.; Zhu, J.-Y.; Luo, S.; Tu, Z. Milcut: A sweeping line multiple instance learning paradigm for interactive image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 256–263, 2014.

[91] Kolotouros, N., Maragos, P. A finite element computational framework for active contours on graphs. *arXiv preprint* arXiv:171004346, 2017.

[92] Choi, J.; Choi, J. Y. User interactive segmentation with partially growing random forest. In: Proceedings of the IEEE International Conference on Image Processing, 1090–1094, 2015.

[93] Dai, L.; Ding, J.; Yang, J.; Zhang, F.; Li, J. Object extraction from bounding box prior with double sparse reconstruction. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, 903–911, 2015.

[94] Tang, M.; Ayed, I. B; Boykov, Y. Pseudo-bound optimization for binary energies. In: *Computer Vision–ECCV 2014. Lecture Notes in Computer Science, Vol. 8693.* Fleet, D.; Pajdla, T.; Schiele, B.; Tuytelaars, T. Eds. Springer Cham, 691–707, 2014.

[95] Gorelick, L.; Schmidt, F. R.; Boykov, Y. Fast trust region for segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1714–1721, 2013.

[96] Ayed, I. B.; Gorelick, L.; Boykov, Y. Auxiliary cuts for general classes of higher order functionals. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1304–1311, 2013.

[97] Li, K. Q.; Tao, W. B. Adaptive optimal shape prior for easy interactive object segmentation. *IEEE Transactions on Multimedia* Vol. 17, No. 7, 994–1005, 2015.

[98] Liu, D. D.; Pulli, K.; Shapiro, L. G.; Xiong, Y. G. Fast interactive image segmentation by discriminative clustering. In: Proceedings of the ACM Multimedia Workshop on Mobile Cloud Media Computing, 47–52, 2010.

[99] Zemene, E.; Alemu, L. T.; Pelillo, M. Dominant sets for "constrained" image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 41, No. 10, 2438–2451, 2019.

[100] Oh, C.; Ham, B.; Sohn, K. Robust interactive image segmentation using structure-aware labeling. *Expert Systems With Applications* Vol. 79, 90–100, 2017.

[101] Hernández-Vela, A.; Primo, C.; Escalera, S. Automatic user interaction correction via Multi-label Graph cuts. In: Proceedings of the IEEE International Conference on Computer Vision, 1276–1281, 2011.

[102] Wang, T. H.; Han, B.; Collomosse, J. TouchCut: Fast image and video segmentation using single-touch interaction. *Computer Vision and Image Understanding* Vol. 120, 14–30, 2014.

[103] Jain, S. D.; Grauman, K. Click carving: Interactive object segmentation in images and videos with point clicks. *International Journal of Computer Vision* Vol. 127, No. 9, 1321–1344, 2019.

[104] Chen, D.-J.; Chen, H.-T.; Chang, L.-W. Toward a unified scheme for fast interactive segmentation. *Journal of Visual Communication and Image Representation* Vol. 55, 393–403, 2018.

[105] Benenson, R.; Popov, S.; Ferrari, V. Large-scale interactive object segmentation with human annotators. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 11700–11709, 2019.

[106] Jang, W.-D.; Kim, C.-S. Interactive image segmentation via backpropagating refinement scheme. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5297–5306, 2019.

[107] Agustsson, E.; Uijlings, J. R. R.; Ferrari, V. Interactive full image segmentation by considering all regions jointly. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 11622–11631, 2019.

[108] Cerrone, L.; Zeilmann, A.; Hamprecht, F. A. End-to-end learned random walker for seeded image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 12559–12568, 2019.

[109] Zheng, H. Y.; Chen, Y. F.; Yue, X. D.; Ma, C. Deep interactive segmentation of uncertain regions with shadowed sets. In: Proceedings of the 3rd International Symposium on Image Computing and Digital Medicine, 244–248, 2019.

[110] Straehle, C.; Koethe, U.; Knott, G.; Briggman, K.; Denk, W.; Hamprecht, F. A. Seeded watershed cut uncertainty estimators for guided interactive segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Providence, 765–772, 2012.

[111] Couprie, C.; Grady, L.; Najman, L.; Talbot, H. Power watershed: A unifying graph-based optimization framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 33, No. 7, 1384–1399, 2011.

[112] Rupprecht, C.; Peter, L.; Navab, N. Image segmentation in twenty questions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3314–3322, 2015.

[113] Chen, D. J.; Chen, H. T.; Chang, L. W. Interactive 1-bit feedback segmentation using transductive inference. *Machine Vision and Applications* Vol. 29, No. 4, 617–631, 2018.

[114] Sourati, J.; Erdogmus, D.; Dy, J. G.; Brooks, D. H. Accelerated learning-based interactive image segmentation using pairwise constraints. *IEEE Transactions on Image Processing* Vol. 23, No. 7, 3057–3070, 2014.

[115] Batra, D.; Kowdle, A.; Parikh, D.; Luo, J.; Chen, T. iCoseg: Interactive co-segmentation with intelligent scribble guidance. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3169–3176, 2010.

[116] Fathi, A.; Balcan, M. F.; Ren, X.; Rehg, J. M. Combining self training and active learning for video segmentation. In: Proceedings of the British Machine Vision Conference, 2011.

[117] Kowdle, A.; Chang, Y.-J., Gallagher, A.; Chen, T. Active learning for piecewise planar 3D reconstruction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 929–936, 2011.

[118] Gimp, G. N. U. Image manipulation program. User Manual. Edge-Detect Filters, Sobel, The GIMP Documentation Team, 8(2), 8–7, 2008.

[119] Bresson, X.; Esedoḡlu, S.; Vandergheynst, P.; Thiran, J. P.; Osher, S. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision* Vol. 28, No. 2, 151–167, 2007.

[120] Goldstein, T.; Bresson, X.; Osher, S. Geometric applications of the Split Bregman method: Segmentation and surface reconstruction. *Journal of Scientific Computing* Vol. 45, Nos. 1–3, 272–293, 2010.

[121] Peng, Y.; Zhang, J.; Yuan, Y.; Zhu, S.; Fang, L. Robust interactive image segmentation via iterative refinement. In: Proceedings of the IEEE International Conference on Image Processing, 4383–4387, 2014.

[122] Ali, H.; Rada, L.; Badshah, N. Image segmentation for intensity inhomogeneity in presence of high noise. *IEEE Transactions on Image Processing* Vol. 27, No. 8, 3729–3738, 2018.

[123] Badshah, N.; Chen, K. Image selective segmentation under geometrical constraints using an active contour approach. *Communications in Computational Physics* Vol. 7, No. 4, 759–778, 2010.

[124] Chan, T. F.; Vese, L. A. Active contours without edges. *IEEE Transactions on Image Processing* Vol. 10, No. 2, 266–277, 2001.

[125] Gout, C.; Le Guyader, C.; Vese, L. Segmentation under geometrical conditions using geodesic active contours and interpolation using level set methods. *Numerical Algorithms* Vol. 39, Nos. 1–3, 155–173, 2005.

[126] Abdelsamea, M. M.; Gnecco, G.; Gaber, M. M. An efficient Self-Organizing Active Contour model for image segmentation. *Neurocomputing* Vol. 149, 820–835, 2015.

[127] Cremers, D.; Osher, S. J.; Soatto, S. Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk. In: *Pattern Recognition. Lecture Notes in Computer Science, Vol. 3175.* Rasmussen, C. E.; Bülthoff, H. H.; Schölkopf, B.; Giese, M. A. Eds. Springer Berlin Heidelberg, 36–44, 2004.

[128] Lee, C. P.; Snyder, W.; Wang, C. Supervised multispectral image segmentation using active contours. In: Proceedings of the IEEE International Conference on Robotics and Automation, 4242–4247, 2005.

[129] Mille, J.; Bougleux, S.; Cohen, L. D. Combination of paths for interactive segmentation. In: Proceedings of the British Machine Vision Conference, 133.1–133.11, 2013.

[130] Mille, J.; Bougleux, S.; Cohen, L. D. Combination of piecewise-geodesic paths for interactive segmentation. *International Journal of Computer Vision* Vol. 112, No. 1, 1–22, 2015.

[131] Chen, D.; Mirebeau, J.-M.; Cohen, L. D. A new finsler minimal path model with curvature penalization for image segmentation and closed contour detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 355–363, 2016.

[132] Chen, D.; Mirebeau, J.-M.; Cohen, L. D. Finsler geodesics evolution model for region based active contours. In: Proceedings of the British Machine Vision Conference, 22.1–22.12, 2016.

[133] Chen, D.; Mirebeau, J. M.; Cohen, L. D. Global minimum for a finsler elastica minimal path approach. *International Journal of Computer Vision* Vol. 122, No. 3, 458–483, 2017.

[134] Liu, Y.; Yu, Y. Interactive image segmentation based on level sets of probabilities. *IEEE Transactions on Visualization and Computer Graphics* Vol. 18, No. 2, 202–213, 2012.

[135] Scheuermann, B.; Rosenhahn, B. Interactive image segmentation using level sets and dempster-shafer theory of evidence. In: *Image Analysis. Lecture Notes in Computer Science, Vol. 6688.* Heyden, A.; Kahl, F. Eds. Springer Berlin Heidelberg, 656–665, 2011.

[136] Li, Y. P.; Cao, G.; Wang, T.; Cui, Q. J.; Wang, B. S. A novel local region-based active contour model for image segmentation using Bayes theorem. *Information Sciences* Vol. 506, 443–456, 2020.

[137] Mylona, E. A.; Savelonas, M. A.; Maroulis, D. Automated parameterization of active contours: A brief survey. In: Proceedings of the IEEE International Symposium on Signal Processing and Information Technology, 344–349, 2013.

[138] Boykov, Y.; Funka-Lea, G. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision* Vol. 70, No. 2, 109–131, 2006.

[139] Greig, D. M.; Porteous, B. T.; Seheult, A. H. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society: Series B (Methodological)* Vol. 51, No. 2, 271–279, 1989.

[140] Boykov, Y.; Kolmogorov, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 26, No. 9, 1124–1137, 2004.

[141] Vincent, L.; Soille, P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 13, No. 6, 583–598, 1991.

[142] Blake, A.; Rother, C.; Brown, M.; Perez, P.; Torr, P. Interactive image segmentation using an adaptive GMMRF model. In: *Computer Vision - ECCV 2004. ECCV 2004. Lecture Notes in Computer Science, Vol. 3021.* Pajdla, T.; Matas, J. Eds. Springer Berlin Heidelberg, 428–441, 2004.

[143] Besag, J. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society: Series B (Methodological)* Vol. 48, 259–279, 1986.

[144] Lobacheva, E.; Veksler, O.; Boykov, Y. Joint optimization of segmentation and color clustering. In: Proceedings of the IEEE International Conference on Computer Vision, 1626–1634, 2015.

[145] Hartigan, J. A.; Wong, M. A. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* Vol. 28, No. 1, 100–108, 1979.

[146] Boykov, Y.; Veksler, O.; Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 23, No. 11, 1222–1239, 2001.

[147] Zhou, H. L.; Zheng, J. M.; Wei, L. Texture aware image segmentation using graph cuts and active contours. *Pattern Recognition* Vol. 46, No. 6, 1719–1733, 2013.

[148] Wang, T.; Ji, Z. X.; Sun, Q. S.; Han, S. D. Combining pixel-level and patch-level information for segmentation. *Neurocomputing* Vol. 158, 13–25, 2015.

[149] Criminisi, A.; Sharp, T.; Blake, A. GeoS: Geodesic image segmentation. In: *Computer Vision–ECCV 2008. Lecture Notes in Computer Science, Vol. 5302.* Forsyth, D.; Torr, P.; Zisserman, A. Eds. Springer Berlin Heidelberg, 99–112, 2008.

[150] Peng, Z. L.; Qu, S. J.; Li, Q. L. Interactive image segmentation using geodesic appearance overlap graph cut. *Signal Processing: Image Communication* Vol. 78, 159–170, 2019.

[151] Veksler, O. Star shape prior for graph-cut image segmentation. In: *Computer Vision–ECCV 2008. Lecture Notes in Computer Science, Vol. 5304.* Forsyth, D.; Torr, P.; Zisserman, A. Eds. Springer Berlin Heidelberg, 454–467, 2008.

[152] Gorelick, L.; Veksler, O.; Boykov, Y.; Nieuwenhuis, C. Convexity shape prior for segmentation. In: *Computer Vision–ECCV 2014. Lecture Notes in Computer Science, Vol. 8693.* Fleet, D.; Pajdla, T.; Schiele, B.; Tuytelaars, T. Eds. Springer Cham, 675–690, 2014.

[153] Freedman, D.; Zhang, T. Interactive graph cut based segmentation with shape priors. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 755–762, 2005.

[154] Das, P.; Veksler, O.; Zavadsky, V.; Boykov, Y. Semiautomatic segmentation with compact shape prior. *Image and Vision Computing* Vol. 27, Nos. 1–2, 206–219, 2009.

[155] Zeng, Y.; Samaras, D.; Chen, W.; Peng, Q. S. Topology cuts: A novel min-cut/max-flow algorithm for topology preserving segmentation in N–D images. *Computer Vision and Image Understanding* Vol. 112, No. 1, 81–90, 2008.

[156] Chen, L.; Cheng, H. D.; Zhang, J. P. Fuzzy subfiber and its application to seismic lithology classification. *Information Sciences–Applications* Vol. 1, No. 2, 77–95, 1994.

[157] Ciesielski, K. C.; Miranda, P. A. V.; Falcão, A. X.; Udupa, J. K. Joint graph cut and relative fuzzy connectedness image segmentation algorithm. *Medical Image Analysis* Vol. 17, No. 8, 1046–1057, 2013.

[158] Xian, M.; Cheng, H. D.; Zhang, Y. A fully automatic breast ultrasound image segmentation approach based on neutro-connectedness. In: Proceedings of the 22nd International Conference on Pattern Recognition, 2495–2500, 2014.

[159] He, K.; Wang, D.; Tong, M.; Zhang, X. Interactive image segmentation on multiscale appearances. *IEEE Access* Vol. 6, 67732–67741, 2018.

[160] Kim, T. H.; Lee, K. M.; Lee, S. U. Generative image segmentation using random walks with restart. In: *Computer Vision–ECCV 2008. Lecture Notes in Computer Science, Vol. 5304.* Forsyth, D.; Torr, P.; Zisserman, A. Eds. Springer Berlin Heidelberg, 264–275, 2008.

[161] Dong, X. P.; Shen, J. B.; Shao, L.; van Gool, L. Sub-Markov random walk for image segmentation. *IEEE Transactions on Image Processing* Vol. 25, No. 2, 516–527, 2016.

[162] Bampis, C. G.; Maragos, P. Unifying the random walker algorithm and the SIR model for graph clustering and image segmentation. In: Proceedings of the IEEE International Conference on Image Processing, 2265–2269, 2015.

[163] Ham, B.; Min, D. B.; Sohn, K. A generalized random walk with restart and its application in depth up-sampling and interactive segmentation. *IEEE Transactions on Image Processing* Vol. 22, No. 7, 2574–2588, 2013.

[164] Shen, J.; Du, Y.; Li, X. Interactive segmentation using constrained laplacian optimization. *IEEE Transactions on Circuits and Systems for Video Technology* Vol. 24, No. 7, 1088–1100, 2014.

[165] Xie, X.; Yu, Z.; Gu, Z.; Li, Y. An iterative boundary random walks algorithm for interactive image segmentation. *arXiv preprint* arXiv:1808.03002, 2018.

[166] Sener, O.; Ugur, K.; Alatan, A. A. Error-tolerant interactive image segmentation using dynamic and iterated graph-cuts. In: Proceedings of the 2nd ACM International Workshop on Interactive Multimedia on Mobile and Portable Devices, 9–16, 2012.

[167] Sinop, A. K.; Grady, L. A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. In: Proceedings of the IEEE 11th International Conference on Computer Vision, 1–8, 2007.

[168] Mehnert, A.; Jackway, P. An improved seeded region growing algorithm. *Pattern Recognition Letters* Vol. 18, No. 10, 1065–1071, 1997.

[169] Beare, R. Regularized seeded region growing. In: Proceedings of the 6th International Symposium on Mathematical Morphology, 91–99, 2002.

[170] Fan, J. P.; Zeng, G. H.; Body, M.; Hacid, M. S. Seeded region growing: An extensive and comparative study. *Pattern Recognition Letters* Vol. 26, No. 8, 1139–1156, 2005.

[171] Beare, R. A locally constrained watershed transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 28, No. 7, 1063–1074, 2006.

[172] Heimann, T.; Thorn, M.; Kunert, T.; Meinzer, H.-P. New methods for leak detection and contour correction in seeded region growing segmentation. In: Proceedings of the 20th ISPRS Congress Technical Commission V, 317–322, 2004.

[173] Comaniciu, D.; Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 24, No. 5, 603–619, 2002.

[174] Zhou, C.; Liu, C. Interactive image segmentation based on region merging using hierarchical match mechanism. In: Proceedings of the International Conference on Computer Science and Service System, 1781–1784, 2012.

[175] Dong, R.; Wang, B.; Li, S.; Zhou, Z.; Li, S.; Wang, Z. Interactive image segmentation with color and texture information by region merging. In: Proceedings of the Chinese Control and Decision Conference, 777–783, 2016.

[176] Minaee, S.; Boykov, Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image segmentation using deep learning: A survey. *arXiv preprint* arXiv:2001.05566, 2020.

[177] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778, 2016.

[178] Boroujerdi, A. S.; Khanian, M.; Breuß, M. Deep interactive region segmentation and captioning. In: Proceedings of the 13th International Conference on Signal-Image Technology & Internet-Based Systems, 103–110, 2017.

[179] Liew, J.; Wei, Y.; Xiong, W.; Ong, S.-H.; Feng, J. Regional interactive image segmentation networks. In: Proceedings of the IEEE International Conference on Computer Vision, 2746–2754, 2017.

[180] Hu, Y.; Soltoggio, A.; Lock, R.; Carter, S. A fully convolutional two-stream fusion network for interactive image segmentation. *Neural Networks* Vol. 109, 31–42, 2019.

[181] Majumder, S.; Yao, A. Content-aware multi-level guidance for interactive instance segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 11602–11611, 2019.

[182] Castrejon, L.; Kundu, K.; Urtasun, R.; Fidler, S. Annotating object instances with a polygon-RNN. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5230–5238, 2017.

[183] Acuna, D.; Ling, H.; Kar, A.; Fidler, S. Efficient interactive annotation of segmentation datasets with polygon-RNN++. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 859–868, 2018.

[184] Ling, H.; Gao, J.; Kar, A.; Chen, W.; Fidler, S. Fast interactive object annotation with curve-GCN. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5257–5266, 2019.

[185] Wang, Z.; Acuna, D.; Ling, H.; Kar, A.; Fidler, S. Object instance annotation with deep extreme level set evolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7500–7508, 2019.

[186] Sofiiuk, K.; Petrov, I.; Barinova, O.; Konushin, A. f-BRS: Rethinking Backpropagating Refinement for Interactive Segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8623–8632, 2020.

[187] He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, 2961–2969, 2017.

[188] Liew, J. H.; Cohen, S.; Price, B.; Mai, L.; Ong, S.-H.; Feng, J. MultiSeg: Semantically meaningful, scale-diverse segmentations from minimal user input. In: Proceedings of the IEEE International Conference on Computer Vision, 662–670, 2019.

[189] Lin, Z.; Zhang, Z.; Chen, L.-Z.; Cheng, M.-M.; Lu, S.-P. Interactive image segmentation with first click attention. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 13339–13348, 2020.

[190] Wang, T.; Yao, Y.; Chen, Y.; Zhang, M.; Tao, F.; Snoussi, H. Auto-sorting system toward smart factory based on deep learning for image segmentation. *IEEE Sensors Journal* Vol. 18, No. 20, 8493–8501, 2018.

[191] Noma, A.; Graciano, A. B. V.; Cesar, R. M.; Consularo, L. A.; Bloch, I. Interactive image segmentation by matching attributed relational graphs. *Pattern Recognition* Vol. 45, No. 3, 1159–1179, 2012.

[192] Noma, A.; Pardo, A.; Cesar Jr., R. M. Structural matching of 2D electrophoresis gels using deformed graphs. *Pattern Recognition Letters* Vol. 32, No. 1, 3–11, 2011.

[193] Jung, C.; Jian, M.; Liu, J.; Jiao, L. C.; Shen, Y. B. Interactive image segmentation via kernel propagation. *Pattern Recognition* Vol. 47, No. 8, 2745–2755, 2014.

[194] Hu, E. L.; Chen, S. C.; Zhang, D. Q.; Yin, X. S. Semisupervised kernel matrix learning by kernel propagation. *IEEE Transactions on Neural Networks* Vol. 21, No. 11, 1831–1841, 2010.

[195] Li, H.; Wu, W.; Wu, E. H. Robust interactive image segmentation via graph-based manifold ranking. *Computational Visual Media* Vol. 1, No. 3, 183–195, 2015.

[196] Wang, B.; Tu, Z. Affinity learning via self-diffusion for image segmentation and clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2312–2319, 2012.

[197] Likas, A.; Vlassis, N.; Verbeek, J. J. The global k-means clustering algorithm. *Pattern Recognition* Vol. 36, No. 2, 451–461, 2003.

[198] Belhumeur, P. N.; Hespanha, J. P.; Kriegman, D. J. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 19, No. 7, 711–720, 1997.

[199] Zemene, E.; Pelillo, M. Interactive image segmentation using constrained dominant sets. In: *Computer Vision–ECCV 2016. Lecture Notes in Computer Science, Vol. 9912.* Leibe, B.; Matas, J.; Sebe, N.; Welling, M. Eds. Springer Cham, 278–294, 2016.

[200] Bulò, S. R.; Pelillo, M. Dominant-set clustering: A review. *European Journal of Operational Research* Vol. 262, No. 1, 1–13, 2017.

[201] Breve, F. Interactive image segmentation using label propagation through complex networks. *Expert Systems With Applications* Vol. 123, 18–33, 2019.

[202] Wang, T.; Sun, Q. S.; Ji, Z. X.; Chen, Q.; Fu, P. Multi-layer graph constraints for interactive image segmentation via game theory. *Pattern Recognition* Vol. 55, 28–44, 2016.

[203] Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on*

[204] Mathieu, B.; Crouzil, A.; Puel, J. B. Interactive segmentation: A scalable superpixel-based method. *Journal of Electronic Imaging* Vol. 26, No. 6, 061606, 2017.

[205] Borovec, J.; Švihlík, J.; Kybic, J.; Habart, D. Supervised and unsupervised segmentation using superpixels, model estimation, and graph cut. *Journal of Electronic Imaging* Vol. 26, No. 6, 061610, 2017.

[206] Borovec, J.; Kybic, J.; Sugimoto, A. Region growing using superpixels with learned shape prior. *Journal of Electronic Imaging* Vol. 26, No. 6, 061611, 2017.

[207] Zhou, Y.; Ju, L.; Wang, S. Multiscale superpixels and supervoxels based on hierarchical edge-weighted centroidal voronoi tessellation. *IEEE Transactions on Image Processing* Vol. 24, No. 11, 3834–3845, 2015.

[208] Arbeláez, P.; Maire, M.; Fowlkes, C.; Malik, J. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 33, No. 5, 898–916, 2011.

[209] Luo, L. K.; Wang, X. F.; Hu, S. Q.; Hu, X.; Zhang, H. L.; Liu, Y. H.; Zhang, J. A unified framework for interactive image segmentation via Fisher rules. *The Visual Computer* Vol. 35, No. 12, 1869–1882, 2018.

[210] Wang, X. F.; Tang, Y. X.; Masnou, S.; Chen, L. M. A global/local affinity graph for image segmentation. *IEEE Transactions on Image Processing* Vol. 24, No. 4, 1399–1411, 2015.

[211] Shi, R.; Ngan, K. N.; Li, S. N.; Li, H. L. Interactive object segmentation in two phases. *Signal Processing: Image Communication* Vol. 65, 107–114, 2018.

[212] Wang, T.; Ji, Z. X.; Sun, Q. S.; Chen, Q.; Jing, X. Y. Interactive multilabel image segmentation via robust multilayer graph constraints. *IEEE Transactions on Multimedia* Vol. 18, No. 12, 2358–2371, 2016.

[213] Van den Bergh, M.; Boix, X.; Roig, G.; de Capitani, B.; van Gool, L. SEEDS: Superpixels extracted via energy-driven sampling. In: *Computer Vision–ECCV 2012. Lecture Notes in Computer Science, Vol. 7578.* Fitzgibbon, A.; Lazebnik, S.; Perona, P.; Sato, Y.; Schmid, C. Eds. Springer Berlin Heidelberg, 13–26, 2012.

[214] Stutz, D.; Hermans, A.; Leibe, B. Superpixels: An evaluation of the state-of-the-art. *Computer Vision and Image Understanding* Vol. 166, 1–27, 2018.

[215] Martin, D.; Fowlkes, C.; Tal, D.; Malik, J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of

the 8th IEEE International Conference on Computer Vision, 416–423, 2010.

[216] Cheng, M.-M.; Mitra, N. J.; Huang, X.; Torr, P. H. S.; Hu, S.-M. Global contrast based salient region detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 37, No. 3, 569–582, 2015.

[217] Everingham, M.; Eslami, S. M. A.; van Gool, L.; Williams, C. K. I.; Winn, J.; Zisserman, A. The pascal visual object classes challenge: A retrospective. *International Journal of Computer Vision* Vol. 111, No. 1, 98–136, 2015.

[218] Rhemann, C.; Rother, C.; Wang, J.; Gelautz, M.; Kohli, P.; Rott, P. A perceptually motivated online benchmark for image matting. In: Proceedings of theIEEE Conference on Computer Vision and Pattern Recognition, 1826–1833, 2009.

[219] Alpert, S.; Galun, M.; Brandt, A.; Basri, R. Image segmentation by probabilistic bottom-up aggregation and cue integration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 34, No. 2, 315–327, 2012.

[220] Lin, T. Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C. L. Microsoft COCO: Common objects in context. In: *Computer Vision–ECCV 2014. Lecture Notes in Computer Science, Vol. 8693*. Fleet, D.; Pajdla, T.; Schiele, B.; Tuytelaars, T. Eds. Springer Cham, 740–755, 2014.

[221] Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The cityscapes dataset for semantic urban scene understanding. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3213–3223, 2016.

[222] Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research* Vol. 32, No. 11, 1231–1237, 2013.

[223] Chen, L.-C.; Fidler, S.; Yuille, A. L.; Urtasun, R. Beat the mturkers: Automatic image labeling from weak 3D supervision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3198–3205, 2014.

[224] Unnikrishnan, R.; Pantofaru, C.; Hebert, M. Toward objective evaluation of image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 29, No. 6, 929–944, 2007.

[225] Freixenet, J.; Muñoz, X.; Raba, D.; Martí, J.; Cufí, X. Yet another survey on image segmentation: Region and boundary information integration. In: *Computer Vision–ECCV 2002. Lecture Notes in Computer Science, Vol. 2352*. Heyden, A.; Sparr, G.; Nielsen, M.; Johansen, P. Eds. Springer Berlin Heidelberg, 408–422, 2002.

[226] Meila, M. Comparing clusterings: An axiomatic view. In: Proceedings of the 22nd International Conference on Machine Learning, 577–584, 2005.

[227] Dubuisson, M.-P.; Jain, A. K. A modified Hausdorff distance for object matching. In: Proceedings of the 12th International Conference on Pattern Recognition, 566–568, 1994.

[228] Perazzi, F.; Pont-Tuset, J.; McWilliams, B.; van Gool, L.; Gross, M.; Sorkine-Hornung, A. A benchmark dataset and evaluation methodology for video object segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 724–732, 2016.

[229] Peng, B.; Zhang, L.; Zhang, D. A survey of graph theoretical approaches to image segmentation. *Pattern Recognition* Vol. 46, No. 3, 1020–1038, 2013.

[230] Shen, J.; Du, Y.; Wang, W.; Li, X. Lazy random walks for superpixel segmentation. *IEEE Transactions on Image Processing* Vol. 23, No. 4, 1451–1462, 2014.

[231] Wang, T.; Yang, J.; Ji, Z. X.; Sun, Q. S. Probabilistic diffusion for interactive image segmentation. *IEEE Transactions on Image Processing* Vol. 28, No. 1, 330–342, 2019.

[232] Tang, M.; Ben Ayed, I.; Marin, D.; Boykov, Y. Secrets of grabcut and kernel k-means. In: Proceedings of the IEEE International Conference on Computer Vision, 1555–1563, 2015.

Hiba Ramadan received her Ph.D. degree in 2018 from the University Sidi Mohamed Ben Abdellah (USMBA), Fez, Morocco. Her research interests are in video analysis, medical imaging, visual information retrieval, and pattern recognition.

Chaymae Lachqar is a medical student in Faculty of Medecine and Pharmacy, USMBA, Fez. She is interested in biotechnology and bioinformatics.

Hamid Tairi received his Ph.D. degree in 2001 from USMBA, Morocco. In 2002 he has been a postdoc in the Image Processing Group of the Laboratory in France. Since 2003, he has been an associate professor at USMBA, where he obtained his HDR in 2009. His research interests are in 3D reconstruction of artificial vision, medical imaging, visual information retrieval, and pattern recognition.