Biomedical Signal Processing and Control 59 (2020) 101924

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Dermoscopic skin lesion image segmentation based on Local Binary Pattern Clustering: Comparative study

Pedro M.M. Pereira [a,b,*], Rui Fonseca-Pinto [b,c], Rui Pedro Paiva [a], Pedro A.A. Assuncao [b,c], Luis M.N. Tavora [c], Lucas A. Thomaz [b,c], Sergio M.M. Faria [b,c]

[a] DEI – FCTUC, University of Coimbra, Portugal
[b] Instituto de Telecomunicações, Portugal
[c] Polytechnic of Leiria, Portugal

ARTICLE INFO

Article history:
Received 13 September 2019
Received in revised form 7 February 2020
Accepted 29 February 2020

Keywords:
Skin lesion
Segmentation
Medical image analysis
Dermoscopy

ABSTRACT

Accurate skin lesion segmentation is important for identification and classification through computational methods. However, when performed by dermatologists, the results of clinical segmentation are affected by a certain margin of inaccuracy (which exists since dermatologist do not delineate lesions for segmentation but for extraction) and also significant inter- and intra-individual variability, such segmentation is not sufficiently accurate for segmentation studies. This work addresses these limitations to enable detailed analysis of lesions' geometry along with extraction of non-linear characteristics of region-of-interest border lines. A comprehensive review of 39 segmentation methods is carried out and a contribution to improve dermoscopic image segmentation is presented to determine the regions-of-interest of skin lesions, through accurate border lines with fine geometric details. This approach resorts to Local Binary Patterns and k-means clustering for precise identification of lesions boundaries, particularly the melanocytic. A comparative evaluation study is carried out using three different datasets and reviewed algorithms are grouped according to their approach. Results show that algorithms from the same group tend to perform similarly. Nevertheless, their performance does not depend uniquely on the algorithm itself but also on the underlying dataset characteristics. Throughout several evaluations, the proposed Local Binary Patterns method presents, consistently, better average performance than the current state-of-the-art techniques across the three different datasets without the need of training or supervised learning steps. Overall, apart from presenting a new segmentation method capable of outperforming the current state-of-the-art, this paper provides insightful information about the behaviour and performance of different image segmentation algorithms.

1. Introduction

Medical image processing has been evolving for decades with increasingly powerful tools to assist dermatologists in the identification and classification of skin lesions [1]. Among the various types of skin lesions the melanoma is the most aggressive type of skin cancer [2], which causes the major number of death cases [3]. However, it presents high cure rates when detected and treated at an early stage it presents high cure rates [4]. Currently, dermatoscopy is the most common imaging technology used to help dermatologists on their initial assessments, making use of hand-held devices and computer vision techniques for providing useful insights to clinicians. Nevertheless, fully automated and reliable diagnosis of skin cancers still remains as one of the most challenging problems in medical image processing.

The research and development of skin lesion segmentation methods aim to improve the accuracy of the Region of Interest (ROI) delineation in dermatoscopic images. Since in general, human beings are not very good at visually discriminating subtle variations in contrast or blur, dermatologists usually do not delineate lesion borders in detail [5]. Therefore, expert drawn ground-truth border lines lack uniformity and precision. In addition, the morphological structure of a lesion (e.g. low lesion-to-skin gradient or multiple lesion regions) may induce further complexity to both manual and automated segmentation of ROIs. To this end, a large number of classical image segmentation methods, spanning across different algorithmic approaches, have been proposed over the years, but generalisation and global optimum performance are still far from ideal levels.

* Corresponding author at: Instituto de Telecomunicações, Leiria, Portugal.
E-mail address: pedrommpereira@co.it.pt (P.M.M. Pereira).

https://doi.org/10.1016/j.bspc.2020.101924

Table 1
(Implemented) conventional segmentation algorithms.

ID	Name	Ref	ID	Name	Ref
Thresholding			*Clustering*		
HT	Huang Threshold	[7]	KC	K-Means Colour	[8]
IT	Iterative Threshold	[9]	KS	K-Means Colour and Spatial	[10]
KT	Kapur Threshold	[11]	MC	Mean Shift Colour	[12]
LT	Li Threshold	[13]	MS	Mean Shift Colour and Spatial	[14]
MT	Moment Threshold	[15]	*Quantization*		
OT	Otsu Threshold	[16]	NQ	Neural Quantization	[17]
ST	Shanbhag Threshold	[18]	AQ	Wan Quantifier	[19]
YT	Yen Threshold	[20]	UQ	Wu Quantifier	[21]
BT	Bradley Threshold	[22]	RM	RGB Median Cut	[23]
RT	Renyi Threshold	[24]	PM	PCT Median Cut	[25]
Fuzzy Methods			*Active Contours*		
FT	Fuzzy Differential Evolution Entropy	[26]	CV	Chan-Vese	[27]
FL	Fuzzy Clustering LevelSet	[28]	VV	Chan-Vese Vector	[29]
FC	Fuzzy C-Means	[30]	LM	Lankton Mean Separation	[31]
Merging Threshold			*Pattern Clustering*		
SR	Statistical Region Merging	[32]	**LC**	Local Binary Patterns Clustering	**(Prop.)**

Common approaches of computational methods for skin lesion identification and classification comprise the following four stages: Preprocessing, Segmentation, Feature extraction, and Classification. Once the digital images of the skin surface are captured, their quality is improved in the Preprocessing stage, by means of image enhancement, restoration techniques, and artefact removal (e.g. hair and air bubbles). In the subsequent stage, the image is segmented in order to separate the ROI from less relevant regions that are not useful for the remaining stages of the pipeline.

This paper deals with the image segmentation stage, which is one of the most important processing stages in dermoscopic image processing pipeline due to its impact in the subsequent stages. The methods devised in this work are particularly tailored for segmentation with the specific objective of determining accurate border lines with high geometric detail. Improved segmentation results are obtained, well beyond conventional clinical-based segmentation, normally provided as ground-truth in skin lesion datasets. The robustness and validation of the proposed method (LC), which was partially presented in [6] are evaluated through a comprehensive study by direct comparison with 26 conventional segmentation algorithms using the ground-truth images of skin lesions. The proposed method outperforms all the previous works under comparison, using commonly employed metrics. Furthermore, the performance of the proposed method is also compared with 13 other recent algorithms based on diverse approaches, such as machine learning and mixture models, via the Jaccard Index (JI) metric. Our experiments show that the proposed method's performance is also superior to that of these recent methods, except for one whose performance is comparable.

The main contributions of this work are: (i) the comparison of 40 segmentation methods; and (ii) a detailed description of our method [6] which is based on Local Binary Patterns (LBP). In this work we modified the previous version by altering the colour transformation step. These algorithms are based on different segmentation paradigms that naturally have different impact on their ability to detect the transition between a lesion and normal skin. Additionally, the research study presented in this paper also provides useful insights about the comparative performance of such algorithms.

The remainder of the paper is organised as follows: Section 2 provides a summary of relevant work in the field of classical automated bi-modal/skin lesion segmentation; afterwards Section 3 presents the proposed method; then Section 4 provides details about the used datasets and evaluation metrics, and Section 5 presents results obtained from the reviewed methods and Local Binary Patterns Clustering (LC) as well as their analysis. Finally, Section 6 draws some conclusions and suggestions for future work.

Table 2
(Non-implemented) machine learning-based algorithms.

ID	Name	Ref
FCN	Fully Convolutional Networks	[33]
SSLS	Saliency-based Skin Lesion Segmentation	[34]
SCDRR	Sparse Coding with Dynamic Rule-based Refinement	[35]
MSCA	Multi-scale Superpixel with Cellular Automata	[36]
mFCN	Multi-stage Fully Convolutional Networks	[37]
JCLMM	Joint Circular-Linear Mixture Model	[38]
CDNN	Convolutional-Deconvolutional Neural Network	[39]
CDNNE	Enhanced CDNN	[40]
KL-LS	Kullback-Leibler based Level Sets	[41]
DermoNet	Densely Linked Convolution Neural Network	[42]
PSO-DEN	Particle Swarm Optimization on Deep Ensemble Network	[43]
SWSDB	Semi- and Weakly Supervised Directional Bootstrapping	[44]
DCL-PSI	Deep Class-specific Learning with Prob. based Step-wise Int.	[45]

2. Skin lesion segmentation methods

Skin lesion segmentation methods can be grouped based on their underlying algorithmic approach. In this work, a first group of 27 segmentation methods is divided into the following 7 categories: Threshold, Clustering, Fuzzy Methods, Quantization, Active Contours, Merging Threshold, and Pattern Clustering. This categorisation is shown in Table 1, where the first column (ID) is the acronym used to identify each method and the last column corresponds to the associated bibliographic reference. Most of these references are the publications where the original methods were proposed. Then, a second group of algorithms recently proposed in the literature, based on diverse approaches that do not feature in Table 1, are also considered and listed in Table 2. These are presented at the end of this section.

2.1. Thresholding methods

Segmentation algorithms based on thresholding approaches are further divided into Global and Local Thresholding, according to whether the entire input image or several smaller partitions are used to optimise the thresholding process. The presented thresholding methods perform image segmentation by generating a modified version of the image, whose grey-level values are separated using a threshold. The algorithms based on thresholding herein considered are the following:

- *Huang Threshold* (HT) [7] is a method based on the minimisation of the fuzziness of an image using a membership function to obtain the threshold. This membership function denotes the relationship between a pixel and the region where it belongs (either ROI or background);
- *Iterative Threshold* (IT) [9] chooses an optimum threshold after successive iterations, providing increasingly cleaner extractions of the ROI. It does so by transforming a smooth grey-level picture into a bi-modal picture (in the greyscale colour space) while maintaining, as close as possible, the average luminance of the picture;
- *Kapur Threshold* (KT) [11] determines the threshold by maximising the entropy on the grey-level histogram;
- *Li Threshold* (LT) [13] addresses the threshold selection problem by minimising the cross entropy between the image and its segmented version, with no assumptions about the distribution;
- *Moment Threshold* (MT) [15] uses an image moment-preserving principle where the threshold values are deterministically computed in such a way that the grey-level moments of an input picture are preserved in the output;
- *Otsu Threshold* (OT) [16] is a non-parametric and unsupervised image segmentation method for automatic threshold selection that uses a discriminant criterion. It only utilises the zeroth- and first-order cumulative moments of the grey-level histogram to calculate the optimum threshold that separates two classes in such a way that their combined intra-class variance is minimal.
- *Shanbhag Threshold* (ST) [18] is a modified version of the Kapur Threshold [11] method, where a more pertinent information measure of the image is obtained, consisting of viewing it as being composed by two fuzzy sets corresponding to two different classes, with membership coefficients associated with their grey-level;
- *Yen Threshold* (YT) [20] is a method based on the use of both the discrepancy between the thresholded and the original image, as well as the number of bits required to represent the thresholded image. A maximum correlation criterion for bi-level thresholding is considered by minimising a cost function;
- *Bradley Threshold* (BT) [22] is a local threshold method that sets each pixel to black if its brightness is a given percentage lower than the average of the surrounding pixels within a window of a specified size, otherwise each pixel is set to white;
- *Renyi Threshold* (RT) [24] is a technique based on Renyi's entropy. Similar to the maximum entropy sum method of [11] and the entropic correlation method of [20], it proposes a thresholding technique using two probability distributions (object and background) derived from the original grey-level distribution of an image, and includes the maximum entropy sum method and the entropic correlation method.

HT, IT, OT, YT, MT, and ST methods were introduced as general-purpose (bi- or multi-level) segmentation algorithms for greyscale images and visually evaluated or compared with several classical images. KT, LT, and RT are build upon two other methods, taking advantage of the entropy concept in different ways. Finally, BT was proposed as a real-time solution for live stream videos and augmented reality solution, by marking pixels as dark or light based on the spatial variation in illumination.

2.2. Clustering methods

Another type of segmentation algorithms is based on clustering, where a set of pixels is grouped according to some metric of similarity. In other words, pixels in the same group are more similar to each other (in a given sense) than those in other groups. This type of algorithms is very common in the literature for various applications and in this work these methods are used to segment skin lesions. A brief description of each one is presented below:

- *K-Means Colour* (KC) [8] is a method of vector quantization which aims to partition observations into clusters, in which every single observation belongs to the cluster whose mean value is closer to its own;
- *K-Means Colour and Spatial* (KS) [10] is an adaptive technique for colour-texture segmentation that is a generalisation of the standard KC. It adds two more dimensions to the problem that sample the local colour smoothness and the local texture complexity. In addition, it also selects the dominant colours from the input image using information from colour histograms, so that proper cluster centres may be selected;
- *Mean-shift Colour* (MC) [12] is a non-parametric density gradient estimation method of generalised kernel, in which a mean-shift estimate kernel is presented for gradient estimation;
- *Mean-shift Colour and Spatial* (MS) [14] is a general non-parametric technique for the analysis of a complex multi-modal feature space and to delineate arbitrarily shaped clusters in it, where the basic computational module is the MC. This method provides discontinuous clusters and preserves smoothing and image segmentation by augmenting the feature space with additional (spatial) parameters from the input domain.

As previously stated, both KS and MS are extensions of the more traditional KC and MC algorithms, respectively, which evolved to take advantage of information connecting similar data-points, i.e., spatial information.

The KC method was proposed with the aim of enabling a process for partitioning N-dimensional populations into k sets on the basis of a sample. The KS method was designed for image segmentation, aiming to complement KC with spatial information during the space partitioning process. The starting point for the KC approach was related to problems in optimal classification where results are theoretically justified, while in KS results are visually displayed for 6 images and visually compared with Mean-Shift algorithm. In particular, the KS algorithm is employed by the same authors for segmentation of a set of 6 skin cancer images [46]. The KS results are evaluated using mean, standard deviation and root-mean-square of the Euclidean distance between the pixels of the ground-truth image and the proposed results.

While both MC and MS algorithms were proposed to provide further understanding of dense information, they aim for different applications. MC was proposed as a non-parametric density gradient estimation method of generalised kernel that is applied to several pattern recognition problems (Gradient Clustering, Clustering and Data Filtering). Each kernel is derived, guaranteeing it is asymptotically unbiased, consistent and its estimate is uniformly consistent (evaluation is made visually). On the other hand, MS was proposed for analysis of a complex multi-modal feature space and to delineate arbitrarily shaped clusters in it, therefore the method has been successfully applied in several tasks and problems like filtering and segmentation.

2.3. Fuzzy Methods

Methods using a mixture of thresholding and clustering concepts aided by fuzzy logic or representation are grouped into this third type of segmentation methods. The following algorithms are considered:

- *Fuzzy Differential Evolution Entropy* (FT) [26] creates fuzzy partitions from the image histogram based on entropy theory. Then, the entropy measure is optimised to obtain the thresholds of the

image using a differential evolution meta-heuristic, which leads to a fast and accurate convergence;
- *Fuzzy Clustering LevelSet* (FL) [28] is another example of a clustering algorithm based on an hybrid level model, alternating between global and local region competitions. The algorithm directly evolves from the initial segmentation by using spatial information in the fuzzy clustering technique where the controlling parameters are automated through estimation from the results of fuzzy clustering;
- *Fuzzy C-Means* (FC) [30] is a clustering variant, where fuzzy partitions and prototypes are firstly generated for any set of numerical data. Then a generalised least-squares objective function is used as the clustering criterion to aggregate the subset.

Both FT and FL are presented in pursue of the same target application. FT was evaluated resorting to eight classical greyscale images in terms of computational time, mean objective value, and standard deviation of objective values. FL was proposed for medical image segmentation, and it was visually evaluated for different modalities (including carotid artery ultrasound images, liver tumours CT scans, and cerebral tissue MRI slices). Finally, the FC algorithm was originally implemented in Fortran-IV [30] that is linked to an original publication [47] (through [48]), which validates the method capability using two classes that have some degree of overlap, and compares it against other solutions.

2.4. Quantization

Segmentation approaches based on Quantization reduce the number of distinct colour levels. For skin lesion segmentation, quantization is performed until the image is segmented in two levels, i.e., the ROI and the background. In the scope of this work, five quantization methods were considered:

- *Neural Quantization* (NQ) [17] is an algorithm that uses a self-organising Kohonen neural network to quantize the colour image;
- *Wan Quantifier* (AQ) [19] is an variance-based algorithm used for multidimensional data clustering, that uses a sum-of-squared error minimisation criterion between the quantized and original image;
- *Wu Quantifier* (UQ) [21] is based on variance minimisation through linear search. The colour space cube is divided into two, in each of its axes and the division plane that minimises the sum of variances at both sides of the colour space is selected, thus creating two boxes. Then the process is repeated for the box with the largest variance. The process stops when a predefined number of boxes is found, and the boxes' centre of gravity are selected as the representative colours.
- *RGB Median Cut* (RM) [23] maps colours to their nearest neighbours in a colormap, effectively quantizing and redrawing the original image;
- *PCT Median Cut* (PM) [25] algorithm is based on an optimal transform using the Principal Components Transform (PCT), also known as Karhunen-Loeve or Hotelling transform.

All these methods were developed with the aim to advance the capabilities of representing images with less colours. NQ, AQ, and UQ were introduced as improvements over other state-of-the-art algorithms. NQ is a 24bit-to-8bit image converter for both greyscale and coloured images, which uses half the memory of other algorithms. UQ was proposed in the same scope as AQ and it was able to achieve 1/3 of the mean-square error attained by other algorithms. Finally, RM and PM are intended to display high-quality reproductions of colour images with small frame buffers. The PM method is an alternative to RGB Median Cut for two-colour-image segmentation on skin tumour images for extraction of features like tumour border, crust, hair, scale shiny areas and ulcer, with the underlying objective of developing an automated visual feature identification program. In the original paper, a success measure is defined with three metrics and the results presented when the image is passed to several different colour spaces.

2.5. Active contour

A fifth type of segmentation approach is called Active Contour algorithms, whose main advantage is the ability to delineate objects in potentially noisy 2D images. In such methods the contour is defined by a constrained spline, which is obtained by minimising a cost function. In the context of this work the following methods are used:

- *Chan-Vese* (CV) [27] is a model based on active contours proposed to detect objects resorting to techniques of curve evolution, Mumford–Shah segmentation, and level sets. It minimises an energy which can be seen as a particular case of a minimal partition problem. The level set formulation evolves the active contour until the desired boundary which does not depend on the gradient since it relates to a particular segment of the image;
- *Chan-Vese Vector* (VV) [29] is an extension of the previous method that minimises a Mumford-Shah function over the length of the contour, plus the sum of the fitting error over each component of the vector-valued image. Like the original *Chan-Vese* model, the vector-valued model also detects edges both with or without gradient;
- *Lankton Mean Separation* (LM) [31] considers local image statistics and evolves a contour based on local information. It essentially derives the curve evolution that separates two or more values of a pre-determined set of statistics computed over geometrically determined subsets of the image.

Both CV and VV are very similar as they have the same base method, but the second is specially tailored for object detection in vector-valued images (RGB). They are visually evaluated with several artificial images and a satellite image. In opposition, LM was proposed as an active contour algorithm that used local rather than global statistics, and it was visually evaluated using 10 natural and artificial images.

2.6. Merging threshold

In addition to the threshold and quantization algorithms, an intermediate type of techniques that merges several thresholds together has emerged. In this work, we explore the use of the following method:

- *Statistical Region Merging* (SR) [32] is presented as an unsupervised inference approach to border detection in dermoscopy images of pigmented skin lesions based on the statistical region merging algorithm. The inference problem models the image as an observed instance of an unknown theoretical image, whose statistical (true) regions are to be reconstructed. The optimal statistical regions share a homogeneity property, such that inside any statistical region and given any colour channel, the pixels have the same expectation, whereas the expectations of adjacent statistical regions differ in at least one colour channel. This algorithm also contains a preprocessing stage which first removes circular black frames (by verifying the lightness component of the *CIE L*a*b** colour space), and then applies a median filter (image smoothing) in order to mitigate the detrimental effects of possible artefacts (hairs and bubbles and skin lines).

SR was proposed to be a fast and accurate way of producing skin lesion segmentation, which is compared with four state-of-the-art automated methods. This is done by resorting to a border error metric and a dataset of images that lacks ground-truth images, where three expert dermatologists are said to have produced the target ground-truth segmentation.

2.7. Pattern Clustering

The last type of segmentation method (Table 1) is that of the proposed method, further detailed in Section 3, which comprises a combination of image binary patterns and clustering.

2.8. Other approaches

So far, the discussed algorithms are based on traditional approaches, which have been progressively evolving towards a new generation of methods. Taking advantage of recent developments in the field of machine learning, a new type of algorithms has emerged. Among such recent machine learning methods, a few were specifically designed to segment skin lesions. These are used in this work and their performance compared to that of the proposed method. The 13 chosen algorithms considered for this comparison are the following:

- *Fully Convolutional Networks* (FCN) [33] is a method based on Convolutional Neural Networks that is presented as a proof-of-concept showing that convolutional networks by themselves, trained end-to-end, pixels-to-pixels, are able to capture feature representations that contain a high-level of semantic information through several convolutional layers. This is done by resorting to a skip architecture that combines semantic information (from deep and coarse layers) with appearance information (from shallow and fine layers) to produce accurate and detailed segmentation;
- *Saliency-based Skin Lesion Segmentation* (SSLS) [34] presents a hair removal technique prior to the actual segmentation, consisting of a pixel level saliency map and a lesion biased Gaussian model;
- *Sparse Coding with Dynamic Rule-based Refinement* (SCDRR) [35] is an unsupervised skin lesion segmentation method for dermoscopic images that exploits the contextual information of skin image at the superpixel level with a Laplacian multi-task sparse representation;
- *Multi-scale Superpixel with Cellular Automata* (MSCA) [36] uses image-wise supervised learning to derive a probabilistic map for automated seed selection. It also enables the inclusion of additional structural information in conjunction with a Multi-scale Superpixel-based Cellular Automata;
- *multi-stage FCN* (mFCN) [37] is a recent approach that learns and refines the skin lesion segmentation results across multiple stages. The algorithm then integrates these complementary multi-stage segmentation results in an ensemble-like fashion;
- *Joining Circular-Linear Mixture Model* (JCLMM) [38] models hue and chroma information assuming the multi-modal characteristics of skin lesions and deals with heterogeneous margins for different mixture components;
- *Convolutional-Deconvolutional Neural Network* (CDNN) [39] is a fully automatic method for skin lesion segmentation by leveraging a 19-layer deep convolutional neural network. The method includes a set of strategies to ensure effective and efficient learning with limited training data, as well as a novel loss function based on Jaccard distance to eliminate the need of sample re-weighting (when using cross entropy due to the strong imbalance between the number of foreground and background pixels);
- *Enhanced CDNN* (ECDNN) [40] is an extended version of the previous work [39] that develops a deeper network architecture with smaller kernels to enhance its discriminant capacity. The potential of using a deeper network architecture with smaller convolutional kernels is investigated such that the new model has increased discriminative capacity to handle a larger variety of image acquisition conditions. It is also investigated the use of channels in other colour spaces, such as Hue-Saturation-Value (*HSV*) and *CIE L*a*b**, as additional inputs to the network that aim for a more efficient training while controlling over-fitting;
- *Kullback-Leibler based Level Sets* (KL-LS) [41] is a novel active contour based method that takes an initial lesion contour followed by the usage of Kullback-Leibler divergence between the lesion and skin to fit a curve to the lesion boundaries. These can be defined using the gradient flow that minimises an appropriate cost function (that is embedded as an external energy term in the existing distance regularised level sets evolution);
- *Densely Linked Convolution Neural Network* (DermoNet) [42] is a deep neural network that encompasses techniques from several other segmentation-aimed networks like autoencoder network funnelling and residual propagation aided with dense convolution blocks;
- *Particle Swarm Optimization on Deep Ensemble Network* (PSO-DEN) [43] is an evolving ensemble of deep networks and hybrid clustering models where the learning hyper-parameters are optimised with a cascading particle swarm optimisation algorithm and a majority voting strategy to combine the prediction results of each base model to produce the final pixelwise classification outcome;
- *Semi- and Weakly Supervised Directional Bootstrapping* (SWSDB) [44] is a combination of three deep networks: a coarse segmentation network (adapted from other works and pre-trained on other datasets), a dilated classification network (adapted and pre-trained from another work based on dense object location using dilation convolutions), and an enhanced segmentation network (consisting of an autoenconder where the encoder and decoder are separated by a enhancement layer); in which the later is trained with a hybrid loss function comprising dice loss and rank loss metrics;
- *Deep Class-specific Learning with Probability based Step-wise Integration* (DCL-PSI) [45] is a deep convolutional neural network that is refined using a step-wise integration approach that, using the image label classification probability, iteratively maximises the pixel agreement between different network-models.

3. Proposed method

The method herein proposed (LC) comprises a sequential processing flow, as depicted in Fig. 1, which accepts generic dermoscopic images as input. The three main functional blocks are the LBP processing, space transformation, and clustering – all of which was implemented and evaluated resorting to MATLAB®.

The luminance (*Y*) channel is extracted from a given input image and its pixel-wise LBPs computed. Then, a set of LBPs are selected and filtered with a Gaussian kernel to expand the region towards the remaining ones, generating an image (*L*) comprising several homogeneous regions. These regions provide rich information for the clustering algorithm, discriminating the various skin regions with different intensity levels. This is due to the fact that the extracted LBPs mark smooth surface areas on/of the object/image (a characteristic that is not usual in the lesion area, given their chaotic appearance).

The subsequent step consists of a space transformation where the *Y* and *L* images are firstly combined to ease visual separation. To this end, first a new three channel image is formed by stacking the *L* image with the luminance (*Y*) and *L* again. Then, this image stack is converted using the well-known *RGB* to *CIE L*a*b** transform, as a means of obtaining a pre-clustering function for the next step. As explained in detail in Section 3.2, the *Y* and *L* dimensions have

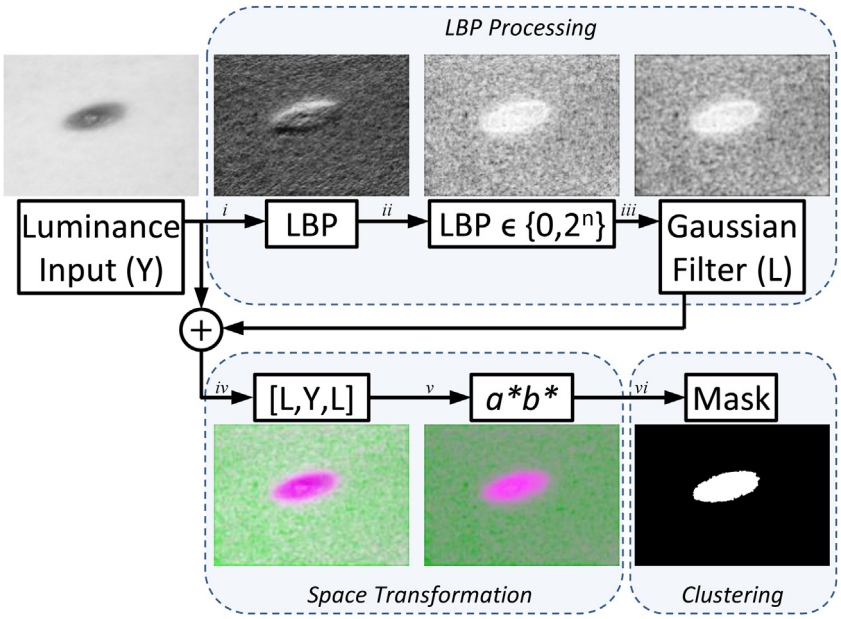


Fig. 1. Local Binary Pattern Clustering method workflow: given a luminance input image *Y*, pixel-wise LBP information is obtained (i) so that a specific set of LBPs is extracted (ii) and smoothed with a Gaussian filter (iii); then, this information, named *L*, is combined with the input *Y* in a particular fashion (iv) so that, after a conversion to the *CIE* $L^*a^*b^*$ colour space, the a^*b^* channels present the rearranged data into an optimised form (v) which, when fed to a clustering algorithm, groups the information into the two desired regions (vi).

some overlapping information where this transformation leads to a better discrimination of the image data into two groups of pixels (lesion vs. non-lesion), thus easing the final clustering process. The remainder of this section further details the processing pipeline steps, as depicted in Fig. 1. Note that the segmentation method works in the same way even if the lesion not round or if it presents blurry edges.

3.1. LBP processing

This subsection describes in detail the LBP processing module, including its three steps, as shown in Fig. 1, namely LBP computation, LBP subset binarisation and Gaussian filtering.

3.1.1. Luminance to LBP

Given an RGB image as input, the corresponding luminance image is obtained by means of a weighted sum of R, G, and B channels, as proposed by Recommendation ITU-R BT.601-7 [49] with four decimal places. The luminance information *Y* is then converted into a Local Binary Pattern (LBP) [50] code for each image using the definition in Eq. (1), where I_p and I_c are, respectively, the intensity of the peripheral and central pixels, and p is the number of neighbouring points. As the neighbourhood of each pixel consists of 8 other pixels, a total of $2^8 = 256$ different labels can be obtained, depending on the relative values of the central pixel and its neighbours.

$$\mathrm{LBP} = \sum_{p=0}^{7} s(I_p - I_c)2^p, \quad s(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{1}$$

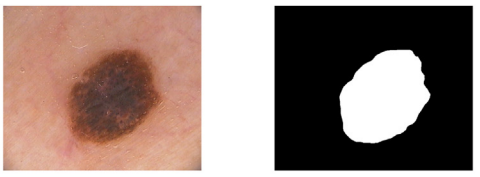
3.1.2. LBP subset

In this step it is important to understand some LBP characteristics. A given LBP can be rotated up to seven times (in the 2D plane) to generate seven other LBPs. This means that many LBPs are merely rotations of some other LBP. For example, LBP of value 1 (represented by 00000001) can be (binary) rotated to generated LBPs 2, 4, 8, 16, 32, 64 and 128 (represented by 00000010, 00000100, 00001000, 00010000, 00100000, 01000000 and 10000000, respectively). When LBPs are grouped in this fashion they are said to be Rotation Invariant [51]. This grouping turns the 256 LBPs into 36 groups of patterns.

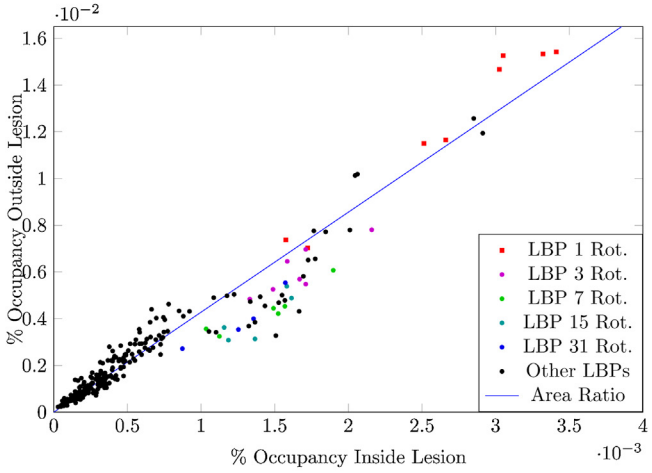
By convention, the LBP group number is defined by the lower LBP value present in the group. For example, LBP-group-3 is comprised by LBPs 3 and its rotations (which are: 6, 12, 24, 48, 96, 192 and 129). Following the previous description, it is now possible to interpret that these 36 groups have specific and comprehensible patterns. Namely, LBP 0 can represent a flat surface, since it can be constructed when the central pixel is surrounded by pixels of equal or inferior value. This is relevant because in skin lesion images the skin area is generally smooth and the lesion area is not. Therefore LBPs representing smooth patterns will be largely present in the skin area, while more noisy patterns will dominate the lesion area.

Resorting to the image presented in Fig. 2a and its ground-truth mask in Fig. 2b, plot Fig. 2c is constructed by extracting all LBP values generated from the image and checking whether they were originated from either the skin or the lesion region of the ground-truth mask. The plot displays the relationship between the frequency of occurrence (in percentage) of each LBP for each region. The ratio between the two region areas is depicted by the blue line. When an LBP is located below the blue line it means that it occurs more often inside the lesion region. Contrarily, when an LBP is located above the blue line, it means that it occurs more often outside the lesion region (i.e., in the skin area). Finally, if an LBP is on the blue line it means that it occurs by the same proportion in both regions of the image. With this information it is possible to visualise that some LBPs (and special groups) are particularly suited to discriminate either the lesion area or the skin area, which is the case for the LBPs rotated from 1 ("*LBP 1 Rot.*" in red).

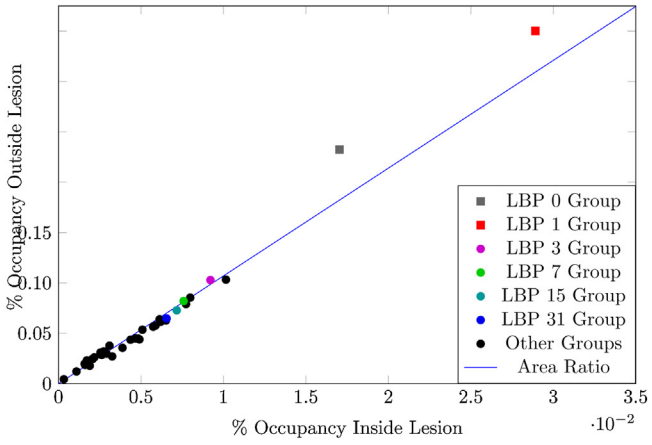
Grouping the rotation invariant LBPs from Fig. 2c allows for a simpler visualisation, hence Fig. 2d is presented. Here, apart from the LBP 0 and 1 group, it is clear that almost all LBP groups tend to be present by the same amounts in either regions. For example, it is clear that LBP group 3 is less dominant than group 0, being also of small importance since it is located much closer to the blue line. Therefore this means that only LBP 0 and 1 groups (which can also be expressed as LBPs of power of 2) are possible candidates to provide information about the type of region the current LBP



(a) *IMD021* image (b) *IMD021* mask



(c) LBPs Presence Ratio



(d) Invariant LBPs Presence Ratio

Fig. 2. Analysis of image (a) against 256 LBPs (c) and later grouped by binary rotation invariance (d). Blue-line represents the ratio between the number of pixels outside and inside the lesion given its ground-truth mask from the dataset. Images (a) and (b) were taken from [52].

represents (in this case, the normal skin where the aforementioned property regarding flat textures exists).

Hence, in this step, the obtained LBP image is filtered so that only the pixels with an LBP corresponding to one of the selected patterns remain, namely LBP values corresponding to powers of two or zero (LBP $\in \{0, 2^n\}$, $n \geq 0$), thus limiting the points to only 9 out of 256 possible LBP patterns. These selected values are represented by zeros in the image matrix and the remainder of the image is represented by ones, which are concentrated in the lesion area [6]. Only these patterns have been chosen, as they are able to represent the situation where the algorithm is more sensitive to region transitions, i.e., a luminance difference between a pixel value and its neighbours.

3.1.3. Gaussian filter

A Gaussian filter [53] (as used in [6]), defined by Eq. (2), is applied to the binary LBP image [50], thus creating the new matrix L with values ranging from zero to one. The 2D-Gaussian kernel with a standard deviation $\sigma = 3$ and size of 13 pixels, generates homogeneous regions in the LBP-image, namely in the lesion, thus removing most pixel-level noise. This regions' representation provides rich information to the clustering algorithm during the segmentation process, since they highlight only one of the two image segments. Note that these values were empirically attributed based on some of the early algorithms results. Thus, 13 pixels and $\sigma = 3$, which act as a low passing filter, may not be optimal values for all images.

$$F(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

3.2. Space transformation

As mentioned above, the space transformation is a pre-clustering operation that aims to produce an image with improved discrimination between lesion/non-lesion image regions (steps (iv) and (v) in Fig. 1), which favours the clustering algorithm ahead. The underlying idea is to use the specific characteristics of the *RGB* to *CIE L*a*b** colour space conversion [54] to obtain such discrimination as, in the case, it inherently contributes to the normalisation of the data and dimensionality reduction (as in [6]). In this colour space conversion, the a^* dimension discriminates between R and G by representing R and G in the positive and negative ranges of the a^* axis, respectively. The b^* dimension discriminates between blue and yellow and, finally, the L^* dimension represents lightness.

Taking these characteristics into account, a correspondence between RGB and the image stack *LYL* previously referred is established, such that $L \rightarrow R$; $Y \rightarrow G$, and $L \rightarrow B$. When the *LYL* image stack is observed in the *RGB* space, it is viewed as an image containing pink colouring in the lesion area, due to the low concentration of LBP values in R and B channels[1] (i.e., the L channels), and a darker colour in the Y image (due to the lower luminance values of the lesion area). On the other hand, in the non-lesion region, the presence of LBP information (i.e. low L values) flattens the R and B channels, while higher values in Y provide a green colouring, as can be seen in the images of Fig. 1. This *LYL*→*RGB* correspondence is of great importance since it places the lesion and non-lesion information, respectively, in the red (R) and green (G) channels of the *RGB* space, which are two opposite sides in the *CIE L*a*b**s' a^* colour space channel. Moreover, the b^* channel provides information to the clustering algorithm whenever the separation between red and green is not evident (i.e., when the magnitude of lesion-to-skin gradient is low). The lightness channel L^* is discarded since it provides little information to the segmentation problem.

Small blobs of pink colouring may still appear over the non-lesion region since any LBP outside the lesion has the same colouring effect as inside the lesion. However, since the luminance values are lower in this region (where such blobs are located), this pink colouring effect will have a more decolourised appearance, which does not present a problem for the algorithm.

3.3. Clustering

After the previously described colour space transformation, the resulting channels a^* and b^* are resized into a vector that is fed into a *k-means++* clustering algorithm [55] (step (vi) in Fig. 1). This algorithm was parameterised with the Euclidean distance metric, with

[1] Note: red+blue=pink and low concentration of LBPs mean high L values.

a maximum of 100 iterations and set to find 2 clusters. Additionally, the algorithm execution is replicated 3 times to reduce the chances of a bad initialisation (which are already mitigated by the *++* variant of k-means adopted). Afterwards, the best solution is selected by its lowest sums (within each cluster) of point-to-centroid (Euclidean) distances.

After the clustering step, the information is reshaped back to the number of rows and columns of the input image. This translates into a pair of binary images were each pixel is assigned to either the lesion area or the skin area, thus creating two distinct and mutually exclusive image segments/masks.

Due to the nature of the clustering algorithm, the outputted masks are not always in the same order, i.e., it is not safe to assume that the 1st segment is always the lesion area. Therefore, an automated process was also conceived to correctly select the mask that depicts the lesion area and the one related to the skin area. Since the segments match the pink and green areas of the stacked image, the variation of the mean pixel values in the image (within each segment) effectively measures the amount of pink in proportion to green that the segment contains, from the a^* to the b^* channel. That is:

1. Because the pink points present more red than blue information, when merging the values from the b^* negative channel (blue) and the values from the a^* positive channel (red), the difference between the mean values of those two layers is always positive (or zero).
2. The green is directly acquired by using the negative values in the a^* channel, thus its mean value is always negative and distant from 0.

This means that the image mask with the largest average value is always the image lesion area. Note that positive b^* values (mapped as yellow) do not exists, but are discarded nonetheless.

At this point, some morphological operations might be necessary to ensure that no small artefacts remain within the lesion perimeter or in the surrounding skin (in this case incorrectly masking the area with holes). These operations are optional since the artefacts provide insight to some structures that might be present in the lesion area, like regression structures that resemble normal skin but are part of the lesion, which are beyond the intent of this work and are therefore excluded from the mask.

After the described process, the final segmentation mask is outputted by the algorithm.

4. Datasets and evaluation metrics

In this work, three datasets were used on the experimental assessment, by performing comparisons with the proposed method plus 26 conventional segmentation algorithms; the algorithms are categorised to 7 classes, as presented in Table 1. The use of three datasets provides robustness to the study by enlarging the representativeness of the dermoscopic images and thus improving the generalisation and validation of the results. All available images where used as test data. This is an important methodological aspect to take into account in performance evaluation studies using skin lesion image datasets, because, in general, any single dataset is not well-balanced in terms of the different types of images, different technical characteristics of the acquisition setup conditions, such as, for example, the type of dermoscope, illumination, and skin surface geometry in the lesion region [32]. Another non-uniform characteristic of the chosen image datasets is the identification of ROI (lesion region) and background skin. In some datasets the ROI segmentation is absent, while others include segmentation masks that have been used for the purpose of delineating surgical excision or clinical follow-up. The datasets used in this work are the following: EDRA-Interactive Atlas of Dermoscopy [56], referred to as *Atlas*; the Pedro Hispano Hospital dataset, known as PH^2 [52], and the Dermofit dataset [57].

The remaining 13 machine-learning segmentation algorithms are later compared with the proposed method using the results reported in their original publication.

4.1. Datasets

The EDRA-Interactive Atlas of Dermoscopy (Atlas) [56] dataset is a representative set of images allowing comparisons among clinical, dermoscopic and histopathological samples. Images have 700×447 8-bit RGB pixels, and where acquired with a Dermaphot/Optotechnik dermoscope. This dataset does not include segmentation masks. However, it is possible to find in the literature some studies using images from this dataset to obtain a segmentation border with the collaboration of dermatologists. This is the case of the work by [58] (SR), whose segmentation results are used as ground-truth in its evaluation study. Since there are no ground-truth available in the original dataset, there is no guaranty that the selected ones are the optimal and validated.

Another set of dermoscopic images is the Pedro Hispano Hospital dataset (PH^2) [52]. It comprises 200 images (with sizes ranging from 761×570 to 769×577 8-bit RGB pixels) acquired with a Tuebinger Mole Analyzer system using a magnification of $20\times$ and the corresponding ground-truth segmentation masks (manually annotated by an expert dermatologist). The PH^2 has been used in different literature reviews for several purposes in dermoscopy as in [59], whose comparison methodology is followed in this work.

The *Dermofit* [57] dataset is a collection of 1300 skin lesion images (with sizes ranging from 177×189 to 3055×1630 8-bit RGB pixels), all collected under the same conditions with internal colour standards using a Canon EOS 350D SLR camera and a ring flash for controlled lighting, jointly with the segmentation ground-truth (generated with a method which uses a statistical model based on the level-set framework).

4.2. Evaluation metrics

Considering the various proposals that can be found in the literature regarding the metrics to evaluate the precision of segmentation methodologies [39,59,60], the objective of comparing segmentation metrics is central to find a global performance measure acting as a first guide for comparison purposes. Still, to assess the segmentation accuracy among the algorithms considered, the heterogeneity of the datasets and some complementary metrics must be considered (particularly when the results present high variability). Hence, the performance results are measured by resorting to five metrics (implemented in MATLAB®): *Border Error* (BE), *Hausdorff Distance* (HD), *True Detection Rate* (TDR), *False Positive Rate* (FPR), and *Jaccard Index* (JI). Only TDR and JI increase performance with higher values, the remaining metrics are preferred with lower values. To be noted that we do not resort to the Sørensen-Dice metric since it can be directly calculated from the Jaccard Index as follows:

$$\text{Dice} = 2 * \text{JI}/(1 + \text{JI}) \tag{3}$$

The *BE* is a broadly used measure to asses segmentation algorithms, as employed by [32], can be described by Eq. (4) – where the percentage area of the non-overlapping segmentation parts (exclusive-OR, $\oplus$) between the segmentation method being eval-

uated (SM) and the ground-truth segmentation (GT) is calculated in percentage.

$$\mathrm{BE}(\mathrm{SM}, \mathrm{GT}) = \frac{\mathrm{Area}(\mathrm{SM} \oplus \mathrm{GT})}{\mathrm{Area}(\mathrm{GT})} \tag{4}$$

The *HD* metric measures how far from each other are two non-empty subsets of a metric space as shown in Eq. (5), where # represents the number of pixels that a segmented lesion region contains. By definition, two sets are close in terms of the *HD* if every point of either set is close to some point of the other set.

$$\mathrm{HD}(\mathrm{SM}, \mathrm{GT}) = \frac{\#(\mathrm{SM} \cup \mathrm{GT}) - \#(\mathrm{SM} \cap \mathrm{GT}))}{\#(\mathrm{SM} \cup \mathrm{GT})} \tag{5}$$

The *TDR*, in (6), measures the ratio of pixels that are correctly classified as lesion, and the *FPR*, in (7), measures the ratio of pixels that are incorrectly classified as lesion.

$$\mathrm{TDR}(\mathrm{SM}, \mathrm{GT}) = \frac{\#(\mathrm{SM} \cap \mathrm{GT})}{\#(\mathrm{GT})} \tag{6}$$

$$\mathrm{FPR}(\mathrm{SM}, \mathrm{GT}) = \frac{\#(\mathrm{SM} \cap \bar{\mathrm{GT}})}{\#(\mathrm{GT})} \tag{7}$$

The *JI*, defined in (8), measures the similarity between two sets and it is computed as the size of the intersection divided by the size of the union of the segmentation masks.

$$\mathrm{JI}(\mathrm{SM}, \mathrm{GT}) = \frac{\left|\mathrm{SM} \cap \mathrm{GT}\right|}{\left|\mathrm{SM}\right| + \left|\mathrm{GT}\right| - \left|\mathrm{SM} \cap \mathrm{GT}\right|} \tag{8}$$

The *JI* metric is also known by intersection over union index, as it provides the ratio of intersection over union areas between two regions, enhancing the common region over the total ROI, which is seen as a central feature to distinguish segmentation outputs. Nonetheless, the remaining metrics are also important since they provide extra detail, specially convenient when comparing similar *JI*.

5. Results and discussion

The achieved results of each algorithm (identified in Table 1), using the previously mentioned metrics, are shown in Table 3 for the Atlas dataset, Table 4 for the PH^2 dataset, and Table 5 for the Dermofit dataset. Additionally, Table 6 shows results, in terms of *JI* metric, for other approaches using the PH^2 dataset. The method's results in Table 6 where gathered from different sources and compiled in reference to the most common used dataset and metric (and thus without the variety of metrics presented in the other tables).

The performance was measured for each algorithm, in percentage, using the five metrics described in the previous section. Lower *BE*, *HD* and *FPR* indicate better performance, while for *TDR* and *JI* better performances are associated to higher percentages. For each performance metric (columns), the segmentation algorithm (rows) that achieves the best performance for each metric is identified by the boldface figure.

Considering the type of images and the nature of the ground-truths (manual or automatic) the *JI* was used as the first (and most relevant) element under analysis, establishing a similarity metric between the ground-truth and the obtained segmentation for each of the 27 methods.

5.1. Atlas dataset

In the absence of segmentation masks in the Atlas dataset, the SR segmentation algorithm proposed by [32] was applied in all images to act as segmentation ground-truth. Most segmentations (generated for Table 3) tend to be slightly inside the SR's ground-truth. As it can be seen, LC, VV and BT have the best *JI* performance.

Table 3
Segmentation results for Atlas dataset (%).

ID	BE	HD	TDR	FPR	JI
Threshold					
HT	27.3 ± 63.5	519.0 ± 163.9	**88.0** ± 7.7	7.6 ± 7.8	82.2 ± 13.1
IT	21.3 ± 13.8	490.7 ± 175.8	83.0 ± 8.0	7.6 ± 4.3	80.3 ± 9.1
KT	31.1 ± 27.7	486.8 ± 165.5	72.2 ± 28.9	12.1 ± 12.4	69.7 ± 27.5
LT	24.7 ± 12.7	502.3 ± 154.3	78.5 ± 9.5	9.0 ± 4.8	76.6 ± 9.8
MT	25.0 ± 15.1	483.4 ± 172.1	80.1 ± 10.9	9.0 ± 5.2	77.0 ± 11.0
OT	21.4 ± 12.0	489.4 ± 175.6	82.3 ± 8.4	7.8 ± 4.4	79.9 ± 9.1
ST	41.6 ± 19.9	**444.9** ± 186.0	61.4 ± 20.3	14.5 ± 6.8	59.6 ± 19.0
YT	28.6 ± 25.5	525.6 ± 145.4	78.8 ± 27.7	11.2 ± 12.2	73.1 ± 25.1
BT	**17.3** ± 10.6	543.4 ± 139.8	87.7 ± 8.2	**6.3** ± 4.3	**84.0** ± 8.8
RT	27.1 ± 24.3	500.3 ± 160.4	77.1 ± 25.4	10.9 ± 12.0	73.9 ± 24.0
Clustering					
KC	**18.9** ± 10.6	485.6 ± 174.5	**84.9** ± 6.9	**6.9** ± 4.1	**82.3** ± 8.1
KS	74.8 ± 83.4	**400.1** ± 180.3	84.5 ± 11.6	21.9 ± 15.7	63.0 ± 22.4
MC	51.3 ± 96.3	443.1 ± 179.5	73.4 ± 28.9	15.0 ± 18.4	67.3 ± 29.4
MS	78.4 ± 29.8	430.2 ± 159.3	22.4 ± 29.5	30.0 ± 15.6	22.2 ± 29.3
Fuzzy Methods					
FT	58.0 ± 25.6	433.2 ± 166.8	42.9 ± 26.0	21.9 ± 12.7	42.5 ± 25.6
FL	**21.7** ± 9.3	**381.6** ± 291.1	80.7 ± 8.7	**8.1** ± 4.4	78.9 ± 8.4
FC	22.4 ± 12.9	490.7 ± 176.0	**81.4** ± 8.5	8.1 ± 4.5	**79.0** ± 9.2
Quantization					
NQ	30.3 ± 56.5	493.8 ± 176.8	87.6 ± 9.4	9.2 ± 10.1	80.0 ± 14.1
AQ	52.1 ± 65.1	580.2 ± 109.5	95.0 ± 5.8	13.8 ± 9.1	71.9 ± 17.3
UQ	**18.3** ± 9.4	**490.1** ± 174.4	85.2 ± 6.9	**6.8** ± 4.1	**82.7** ± 7.7
RM	61.5 ± 64.1	532.1 ± 151.8	95.6 ± 6.6	16.5 ± 8.5	67.9 ± 17.4
PM	61.5 ± 64.2	549.4 ± 137.5	***95.6*** ± ***6.5***	16.4 ± 8.5	67.9 ± 17.4
Active Contours					
CV	52.7 ± 34.8	404.5 ± 203.4	52.4 ± 25.2	19.0 ± 11.7	51.7 ± 25.3
VV	**16.4** ± 8.8	**340.4** ± 245.8	87.1 ± 6.7	**6.0** ± 3.8	**84.5** ± 7.3
LM	104.5 ± 80.1	624.7 ± 91.6	**93.2** ± 6.2	30.8 ± 9.2	52.8 ± 15.1
Pattern Clustering (Proposed)					
LC	***16.2*** ± ***8.7***	***71.6*** ± ***87.2***	**87.1** ± 8.4	***6.2*** ± ***4.5***	***84.5*** ± ***8.0***

In the Thresholding group both BT and HT algorithms perform well. Between them, BT appears to be more reliable as it provides better performance for all metrics, except *HD* and *TDR* (but only by 0.3% that comes with a 1.3% decrease in *FPR*).

In both Clustering and Active Contours group, the best performing algorithms are KC and VV. They provide over 80% *JI*, while the other metrics only go up to 67%.

Apart from a direct *JI* comparison, other metrics help to characterise what happened during the segmentation process and provide better means to compare similarly-performing algorithms. In the Fuzzy group, both FL and FC provide similar *JI* results with the *BE*, *TDR*, and *FPR* metrics differing only 0.68%, 0.74%, and 0.01%, respectively, with FL being better in the *BE* and *FPR* metrics. When taking into account the *HD* metric, it is clear that FL might be preferable instead of the FC since, overall, its errors are less distant from the ground-truth masks.

In the Quantization group, UQ is the best for all metrics except *TDR*. But if a small *FPR* and *HD* compromise is acceptable to attain higher *TDR*, then AQ could be the next best option. However, AQ might not be feasible since the *BE* metric would go from 18.39% to 52.12%, which means that, on average, a lesion segmentation would be always off by half the lesion ground-truth area.

Finally, the proposed LC algorithm provides the best results, in terms of *JI* and *BE*, and is second only to VV, in terms of *FPR* by 0.2%.

5.2. PH^2 dataset

Result trends for the PH^2 dataset (Table 4) are similar to those achieved with the Atlas Dataset, but now they are obtained in respect to a ground-truth provided by the dataset.

Table 4
Segmentation results for PH^2 dataset (%).

ID	BE	HD	TDR	FPR	JI
Threshold					
HT	59.8 ± 147.4	**468.5** ± 150.7	85.9 ± 8.0	10.5 ± 13.4	77.1 ± 19.1
IT	25.1 ± 48.9	495.7 ± 141.2	87.1 ± 6.5	6.5 ± 4.6	82.1 ± 10.8
KT	29.6 ± 30.5	485.0 ± 132.3	86.8 ± 19.8	11.7 ± 16.3	76.7 ± 19.2
LT	20.5 ± 9.1	477.8 ± 146.6	82.7 ± 7.7	6.9 ± 4.3	80.6 ± 7.4
MT	26.2 ± 19.1	503.5 ± 122.8	89.2 ± 8.4	7.6 ± 4.3	79.2 ± 9.3
OT	**18.3** ± 7.2	497.8 ± 138.3	86.0 ± 6.9	**6.1** ± 3.6	**82.9** ± 6.1
ST	33.8 ± 40.0	475.7 ± 143.3	77.6 ± 17.1	9.4 ± 6.9	72.8 ± 17.2
YT	25.0 ± 21.2	488.1 ± 147.1	**89.8** ± 14.6	9.4 ± 12.6	79.7 ± 14.4
BT	32.0 ± 16.6	495.7 ± 125.6	73.4 ± 17.8	11.9 ± 10.2	69.9 ± 15.1
RT	25.5 ± 23.6	489.0 ± 142.2	88.5 ± 15.6	9.8 ± 13.3	79.3 ± 15.4
Clustering					
KC	**17.9** ± 17.9	476.1 ± 155.5	**87.9** ± 6.2	**5.5** ± 3.3	**84.4** ± 7.4
KS	131.3 ± 135.1	**294.8** ± 147.1	87.6 ± 13.7	29.0 ± 14.5	51.9 ± 23.4
MC	70.2 ± 79.8	410.8 ± 148.6	44.3 ± 43.1	23.1 ± 22.3	40.0 ± 39.9
MS	91.5 ± 25.6	430.5 ± 114.5	8.6 ± 26.3	33.4 ± 22.1	8.5 ± 25.7
Fuzzy Methods					
FT	45.7 ± 27.0	414.6 ± 145.9	55.8 ± 27.6	14.5 ± 10.6	55.1 ± 26.9
FL	34.8 ± 118.0	**285.8** ± 251.7	83.4 ± 14.1	7.2 ± 8.8	81.0 ± 13.9
FC	**19.8** ± 11.7	496.4 ± 131.5	**85.5** ± 7.2	**6.4** ± 3.8	**82.0** ± 7.3
Quantization					
NQ	73.5 ± 157.1	490.5 ± 132.9	88.7 ± 16.9	14.5 ± 17.4	72.9 ± 23.5
AQ	109.3 ± 135.7	396.4 ± 176.2	***94.5*** ± ***7.4***	21.3 ± 12.4	59.8 ± 22.7
UQ	**16.3** ± 6.4	493.9 ± 139.8	87.5 ± 6.6	**5.4** ± 3.4	**84.7** ± 5.8
RM	115.4 ± 125.7	**395.5** ± 168.5	93.4 ± 10.9	23.6 ± 10.7	56.6 ± 21.1
PM	114.1 ± 125.4	397.5 ± 172.3	94.1 ± 10.7	23.3 ± 10.6	57.1 ± 21.1
Active Contours					
CV	43.8 ± 60.8	**417.4** ± 197.4	73.9 ± 22.1	12.3 ± 13.8	68.5 ± 21.2
VV	**26.9** ± 28.8	459.0 ± 183.4	82.6 ± 16.2	**7.5** ± 4.2	**77.2** ± 16.4
LM	168.0 ± 161.0	558.5 ± 67.6	**94.3** ± 6.7	36.6 ± 12.2	46.5 ± 19.8
Merging Threshold					
SR	**123.7** ± 173.4	**244.8** ± 173.7	**68.6** ± 36.2	**28.3** ± 22.9	**48.6** ± 33.8
Pattern Clustering (Proposed)					
LC	***14.1*** ± ***4.7***	***58.2*** ± ***39.3***	**88.4** ± 6.1	***5.2*** ± ***3.7***	***86.3*** ± ***4.6***

Table 5
Segmentation results for Dermofit dataset (%).

ID	BE	HD	TDR	FPR	JI
Threshold					
HT	83.9 ± 187.9	366.3 ± 306.6	**81.9** ± 11.8	12.0 ± 13.7	70.2 ± 22.4
IT	65.3 ± 144.7	343.3 ± 305.4	79.2 ± 11.7	11.0 ± 11.1	69.8 ± 19.6
KT	48.3 ± 118.7	394.1 ± 290.8	75.6 ± 22.2	11.8 ± 16.1	69.1 ± 22.8
LT	57.8 ± 126.8	**327.7** ± 297.0	75.3 ± 12.2	10.8 ± 10.1	68.5 ± 17.7
MT	53.1 ± 96.0	368.5 ± 307.8	80.5 ± 12.0	10.6 ± 9.5	69.5 ± 17.9
OT	48.9 ± 103.5	331.4 ± 297.3	76.9 ± 12.0	10.0 ± 9.7	70.6 ± 17.1
ST	65.1 ± 83.1	350.6 ± 267.2	57.0 ± 21.4	15.0 ± 8.8	51.3 ± 19.6
YT	42.4 ± 116.0	443.4 ± 291.2	80.2 ± 21.3	9.8 ± 14.7	**73.2** ± 22.1
BT	**35.0** ± 30.1	370.8 ± 307.5	72.1 ± 18.6	**9.4** ± 7.0	68.6 ± 18.4
RT	48.5 ± 131.6	386.5 ± 292.8	78.8 ± 19.9	11.0 ± 16.0	71.7 ± 21.6
Clustering					
KC	64.2 ± 136.7	349.8 ± 297.0	78.6 ± 12.1	**11.1** ± 11.8	**69.5** ± 19.8
KS	212.2 ± 188.6	348.5 ± 109.0	66.7 ± 15.3	43.4 ± 9.2	30.7 ± 15.5
MC	232.4 ± 355.8	290.1 ± 212.6	**89.7** ± 15.3	40.6 ± 34.7	51.7 ± 29.3
MS	**54.5** ± 35.2	**286.6** ± 212.8	49.7 ± 27.4	16.2 ± 11.5	49.0 ± 27.2
Fuzzy Methods					
FT	115.7 ± 133.4	372.2 ± 189.4	41.6 ± 37.3	29.8 ± 21.7	27.6 ± 24.0
FL	238.9 ± 308.8	**285.3** ± 204.2	50.4 ± 31.3	49.9 ± 41.6	38.9 ± 35.1
FC	**60.7** ± 127.8	343.7 ± 299.1	**77.2** ± 12.2	**10.8** ± 10.1	**69.0** ± 18.8
Quantization					
NQ	177.6 ± 290.3	439.8 ± 296.2	86.5 ± 17.3	25.4 ± 24.1	57.5 ± 29.4
AQ	154.8 ± 243.9	627.0 ± 222.9	91.3 ± 13.6	23.5 ± 18.3	56.2 ± 24.7
UQ	**52.6** ± 108.4	**321.9** ± 298.0	77.9 ± 12.7	**10.7** ± 12.7	**70.6** ± 18.5
RM	146.5 ± 191.2	665.4 ± 215.5	95.1 ± 9.0	23.5 ± 11.7	54.2 ± 22.5
PM	145.7 ± 190.4	664.1 ± 222.4	***95.4*** ± ***8.2***	23.4 ± 11.5	54.4 ± 22.3
Active Contours					
CV	244.6 ± 365.5	339.3 ± 236.5	69.5 ± 30.1	39.9 ± 33.1	44.7 ± 30.8
VV	240.6 ± 398.1	**295.4** ± 256.5	**77.7** ± 24.5	**33.9** ± 33.0	**52.1** ± 31.0
LM	**213.0** ± 86.8	369.8 ± 88.0	0.1 ± 0.5	53.5 ± 12.4	0.0 ± 0.3
Merging Threshold					
SR	**55.0** ± 65.9	**167.1** ± 150.8	**52.0** ± 34.2	**15.0** ± 12.6	**50.2** ± 33.1
Pattern Clustering (Proposed)					
LC	***29.6*** ± ***41.1***	***110.4*** ± ***105.1***	**78.7** ± 21.6	***7.7*** ± ***7.5***	***74.8*** ± ***21.4***

In the Thresholding group, looking at *JI*, both OT and IT perform similarly. However, IT is preferable if a gain of 1.1% in *TDR* is outperformed by the (error) gain of 0.4% *FPR*.

In the Clustering, Quantization and Active Contours groups, KC, UQ, and VV outperform all algorithms in their groups, presenting higher *JI*.

In the Fuzzy group the best algorithm in terms of detection rate is the FC with 85.54% *TDR* and 6.44% *FPR*, however it might not be the best option if its higher erratic behaviour (*HD*), in comparison to FL, is undesirable. The FL method provides almost 40% less *HD* than the FC algorithm, thus yielding smoother segmentation that better mimic the ground-truth at the cost of 2.13% *TDR* and 0.82% *FPR*.

Finally, LC provides even better results than before, while the SR provides almost the worst results in Table 4.

5.3. Dermofit dataset

The Dermofit dataset results (Table 5) are consistent with those obtained for Atlas and PH^2.

In the Thresholding group no algorithm stands out as optimal. Most results for the *JI* metric are within a close range. YT emerges with the highest value and best balance between losses and gains in the remaining metrics in comparison the remaining peers.

In the Clustering group, KC leads with the highest *JI* value. While not having the best *TDR* value, by 11.1%, it provides the best balance with a 29.5% less *FPR* in relation to MC.

Likewise, for the Fuzzy group, FC shows the best values for all metrics except *HD*.

Similarly, in the Quantization and Active Contours groups, both UQ and VV stand out. They are only outperformed by other algorithms by their lower *TDR* and *BE*, respectively. But other algorithms have worse values for all the other metrics, which are disproportional to a the small gain their provide.

Finally, the SR algorithm provided average results and the LC algorithm has the best performance balance in comparison to all others.

5.4. Across datasets

Overall, the top 3 algorithms across the datasets (sorted by an average *JI* performance weighted by the number of images in each dataset) are: LC, UQ, and KC; with 81.7%, 76.3%, and 75.5%, respectively. Generally speaking, Thresholding methods tend to provide results above average, while the remaining groups have a less linear

Table 6
Recent segmentation results for PH^2 dataset (%).

ID	FCN	SSLS	SCDRR	MSCA	mFCN	JCLMM	CDNN	CDNNE	KL-LS	DermoNet	PSO-DEN	SWSDB	DCL-PSI	**LC**
JI	82.15	68.16	76.00	72.33	83.99	70.72[a]	84.33[2]	76.5	71.54[2]	85.3	77.75[2]	***89. 04***[2]	85.90	86.34

[a] Jaccard index value generated from published SørensenDice metric.

behaviour. For dermoscopic datasets, Clustering, Fuzzy, and Quantization algorithms provide average results, while Active Contours methods tend to perform below average (except VV which is above average). For the macro images dataset (Dermofit), only Thresholding and a short number of methods provide results above average. Most algorithms tend to discard small portions of the ground-truth masks and include surrounding skin. Finally, as previously described, the proposed LC provides the best results across the different datasets.

5.5. Recent works

As discussed in the previous section, for the considered datasets, the LC algorithm was found to be the one with the best *JI* metric results in comparison to the revised classical algorithm implementations. The study carried out went beyond those approaches and, as shown in Table 6, the performance of LC was also compared with 13 more recent algorithms based on machine learning techniques (ordered by year of publication) for the PH^2 dataset (as in the 13 published algorithm works). The presented *JI* metric values were gathered from the published papers that describe these methods and from [37]; where now, the proposed method (LC) is only second to SWSDB, and even in this case for a margin smaller than 3%. This can be explained by the fact that LC algorithm is more directed and guided towards the target domain, whereas machine learning algorithms, that depend on ground-truths to learn, need large amounts of data and still might not converge to the desired results.

Although other methods such as [61] and [62] exist and could be compared with our proposal, their results are based on a different set of ground-truths that are incompatible with those of Table 6, thus these comparisons were considered out of the scope of this paper.

5.6. Summary

Globally, excluding the LC algorithm, Thresholding methods perform the best. Nevertheless, having defined *JI* as the most discriminant performance metric, the LC algorithm provides proper results across the datasets. In the presence of more recent and complex algorithms, LC comes second to SWSDB.

6. Conclusions and future work

Automated melanoma identification is crucial to help both dermatologists and computer expert systems. An accurate skin lesion segmentation can improve the initial assessment and help computer vision techniques to provide insights and enhance the users abilities. Therefore, image segmentation of skin lesion imagery plays a major role for forthcoming algorithms. This work provides a unified comparison between several segmentation methods and datasets in order to present a better understanding of some of the currently available techniques. The study compares a specific set of 27 conventional segmentation algorithms, grouped by their type of methodology, tested for 3 different datasets.

The comparison results showed that different segmentation methods of a same type tend to behave similarly. Even so, their behaviour highly depends on the datasets' image characteristics. In spite of that, some conclusions can still be drawn from these studies. The Thresholding type of algorithms seems to be the most constant in terms of expected results, however the results achieved are not that accurate in the general context. In contrast, Active Contours algorithms provide on average the worst FPR measurements, since they tend to stick close to the lesion borders, thus typically presenting more inner segmentations. Quantization, as a group, is on average one of the worst in any scenario, however the quantization UQ algorithm actually performs above average. The sole Merging Threshold algorithm, which is reported to have a good performance for the Atlas dataset, does not stand out across the selected metrics on the remaining two datasets. Finally, the proposed Local Binary Pattern based algorithm appears as the leading method when compared to the existing ground-truths. As future research, adaptions of the proposed method will be investigated on other datasets, as well as, potential clinical applications.

Across the observed results, the proposed LC method is always superior in terms of *BE*, *HD*, and *JI*. It is also the best performer in the *FPR* metric for both the PH^2 and Dermofit datasets, missing only by little this lead position on the *TDR* metric. The best performing algorithms for this metric usually create a segmentation that is too large, covering not only the lesion area but also marking other areas as lesion. Despite its higher *TDR* they perform worse in the remaining metrics.

The proposed unsupervised method is tested with three certified datasets and comprises three main phases: LBP image enhancement, space transformation, and binary clustering. It has performed better than the 27 classic methods of segmentation. Additionally, it was also compared against 13 recent skin lesion segmentation approaches, published between 2015 and 2019, for the PH^2 dataset. The proposed method outperforms most of the other algorithms while is computationally less expensive. It behaves independently from the dataset (given its' unsupervised setting) but is sensitive to hairs and possible artefacts present in the image.

With this work's insight, further research can be done in the field of skin lesion image segmentation to either improve existing segmentation methods that are lacking in performance or refine the existing top performers. This work allows one to detect which segmentation algorithm is more suitable for a given application, by inspecting the strengths and weaknesses of each of the listed algorithms and even deciding whether a given algorithm can be further improved. This is still an expanding field were other comparisons with upcoming algorithms is necessary so that.

CRediT authorship contribution statement

Pedro M.M. Pereira: Methodology, Software, Formal analysis, Investigation, Writing - original draft. **Rui Fonseca-Pinto:** Conceptualization, Writing - review & editing, Supervision, Funding acquisition. **Rui Pedro Paiva:** Writing - review & editing, Supervision, Funding acquisition. **Pedro A.A. Assuncao:** Validation, Writing - review & editing. **Luis M.N. Tavora:** Methodology, Validation, Writing - review & editing. **Lucas A. Thomaz:** Methodology, Formal analysis, Writing - review & editing. **Sergio M.M. Faria:** Conceptualization, Validation, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

This work was supported by Programa Operacional Regional do Centro, project PlenoISLA POCI-01-0145-FEDER-28325, and funded by FCT/MCTES, under PhD Grant SFRH/BD/128669/2017, through national funds and when applicable co-funded by EU funds under the Instituto de Telecomunicacoes project UIDB/EEA/50008/2020.

Conflicts of interest

None declared.

References

[1] R.J. Friedman, D.S. Rigel, A.W. Kopf, Early detection of malignant melanoma: the role of physician examination and self-examination of the skin, CA: Cancer J. Clin. 35 (3) (1985) 130–151.
[2] H.L. Kaufman, The Melanoma Book: a Complete Guide to Prevention and Treatment, vol. 1, Gotham Books, New York, USA, 2005.

[3] D.E. Elder, Skin cancer. Melanoma and other specific nonmelanoma skin cancers, Cancer 75 (S1) (1995) 245–256.
[4] L.A. Goldsmith, F.B. Askin, A.E. Chang, C. Cohen, J.P. Dutcher, R.S. Gilgor, S. Green, E.L. Harris, S. Havas, J.K. Robinson, et al., Diagnosis and treatment of early melanoma: Nih consensus development panel on early melanoma, JAMA 268 (10) (1992) 1314–1319.
[5] G. Day, R. Barbour, Automated skin lesion screening – a new approach, Melanoma Res. 11 (1) (2001) 31–35.
[6] P.M.M. Pereira, R. Fonseca-Pinto, R.P. Paiva, L.M.N. Tavora, P.A.A. Assuncao, S.M.M. Faria, Accurate segmentation of desmoscopic images based on local binary pattern clustering, in: International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia, 2019 arxiv:1902.06347.
[7] L.-K. Huang, M.-J.J. Wang, Image thresholding by minimizing the measures of fuzziness, Pattern Recognit. 28 (1) (1995) 41–51.
[8] J. MacQueen, et al., Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, Oakland, California, United States, 1967, pp. 281–297.
[9] H.J. Trussell, Comments on picture thresholding using an iterative selection method, IEEE Trans. Syst. Man Cybern. 9 (5) (1979) 311.
[10] D.E. Ilea, P.F. Whelan, Color image segmentation using a spatial k-means clustering algorithm, in: International Machine Vision and Image Processing Conference, Dublin, Ireland, 2006.
[11] J.N. Kapur, P.K. Sahoo, A.K. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, Comput. Vis. Graphics Image process. 29 (3) (1985) 273–285.
[12] K. Fukunaga, L. Hostetler, The estimation of the gradient of a density function, with applications in pattern recognition, IEEE Trans. Inf. Theory 21 (1) (1975) 32–40.
[13] C.H. Li, C. Lee, Minimum cross entropy thresholding, Pattern Recognit. 26 (4) (1993) 617–625.
[14] D. Comaniciu, P. Meer, Mean shift: A robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002) 603–619.
[15] W.-H. Tsai, Moment-preserving thresolding: a new approach, Comput. Vis. Graphics Image Process. 29 (3) (1985) 377–393.
[16] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66.
[17] A.H. Dekker, Kohonen neural networks for optimal colour quantization, Netw.: Comput. Neural Syst. 5 (3) (1994) 351–367.
[18] A.G. Shanbhag, Utilization of information measure as a means of image thresholding, CVGIP: Graph. Model. Image Process. 56 (5) (1994) 414–419.
[19] S. Wan, P. Prusinkiewicz, S. Wong, Variance-based color image quantization for frame buffer display, Color Res. Appl. 15 (1) (1990) 52–58.
[20] J.-C. Yen, F.-J. Chang, S. Chang, A new criterion for automatic multilevel thresholding, IEEE Trans. Image Process. 4 (3) (1995) 370–378.
[21] X. Wu, Efficient statistical computations for optimal color quantization, in: Graphics Gems II, Elsevier, 1991, pp. 126–133.
[22] D. Bradley, G. Roth, Adaptive thresholding using the integral image, J. Graphics Tools 12 (2) (2007) 13–21.
[23] P. Heckbert, Color Image Quantization for Frame Buffer Display, vol. 16, ACM, 1982.
[24] P. Sahoo, C. Wilkins, J. Yeager, Threshold selection using Renyi's entropy, Pattern Recognit. 30 (1) (1997) 71–84.
[25] S.E. Umbaugh, R.H. Moss, W.V. Stoecker, G.A. Hance, Automatic color segmentation algorithms-with application to skin tumor feature identification, IEEE Eng. Med. Biol. Mag. 12 (3) (1993) 75–82.
[26] S. Sarkar, S. Paul, R. Burman, S. Das, S.S. Chaudhuri, A fuzzy entropy based multi-level image thresholding using differential evolution, in: International Conference on Swarm, Evolutionary, and Memetic Computing, Springer, Bhubaneswar, India, 2014, pp. 386–395.
[27] T.F. Chan, L.A. Vese, et al., Active contours without edges, IEEE Trans. Image process. 10 (2) (2001) 266–277.
[28] B.N. Li, C.K. Chui, S. Chang, S.H. Ong, Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation, Comput. Biol. Med. 41 (1) (2011) 1–10.
[29] T.F. Chan, B.Y. Sandberg, L.A. Vese, Active contours without edges for vector-valued images, J. Visual Commun. Image Represent. 11 (2) (2000) 130–141.
[30] J.C. Bezdek, R. Ehrlich, W. Full, Fcm: the fuzzy c-means clustering algorithm, Comput. Geosci. 10 (2–3) (1984) 191–203.
[31] S. Lankton, A. Tannenbaum, Localizing region-based active contours, IEEE Trans. Image Process. 17 (11) (2008) 2029–2039.
[32] M.E. Celebi, H.A. Kingravi, H. Iyatomi, Y. Alp Aslandogan, W.V. Stoecker, R.H. Moss, J.M. Malters, J.M. Grichnik, A.A. Marghoob, H.S. Rabinovitz, et al., Border detection in dermoscopy images using statistical region merging, Skin Res. Technol. 14 (3) (2008) 347–353.
[33] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, United States, 2015, pp. 3431–3440.
[34] E. Ahn, L. Bi, Y.H. Jung, J. Kim, C. Li, M. Fulham, D.D. Feng, Automated saliency-based lesion segmentation in dermoscopic images, in: International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy, 2015, pp. 3009–3012.
[35] B. Bozorgtabar, M. Abedini, R. Garnavi, Sparse coding based skin lesion segmentation using dynamic rule-based refinement, in: International Workshop on Machine Learning in Medical Imaging, Athens, Greece, 2016, pp. 254–261.
[36] L. Bi, J. Kim, E. Ahn, D. Feng, M. Fulham, Automated skin lesion segmentation via image-wise supervised learning and multi-scale superpixel based cellular automata, in: IEEE International Symposium on Biomedical Imaging, Prague, Czech Republic, 2016, pp. 1059–1062.
[37] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, D. Feng, Dermoscopic image segmentation via multi-stage fully convolutional networks, IEEE Trans. Biomed. Eng. 64 (9) (2017) 2065–2074.
[38] A. Roy, A. Pal, U. Garain, Jclmm: a finite mixture model for clustering of circular-linear data and its application to psoriatic plaque segmentation, Pattern Recognit. 66 (2017) 160–173.
[39] Y. Yuan, M. Chao, Y. Lo, Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance, IEEE Trans. Med. Imaging 36 (9) (2017) 1876–1886.
[40] Y. Yuan, Y. Lo, Improving dermoscopic image segmentation with enhanced convolutional-deconvolutional networks, IEEE J. Biomed. Health Inform. 23 (2) (2019) 519–526.
[41] F. Riaz, S. Naeem, R. Nawaz, M. Coimbra, Active contours based segmentation and lesion periphery analysis for characterization of skin lesions in dermoscopy images, IEEE J. Biomed. Health Inform. 23 (2) (2019) 489–500.
[42] S. Baghersalimi, B. Bozorgtabar, P. Schmid-Saugeon, H.K. Ekenel, J.-P. Thiran, Dermonet: densely linked convolutional neural network for efficient skin lesion segmentation, EURASIP J. Image Video Process. 2019 (1) (2019) 71.
[43] T.Y. Tan, L. Zhang, C.P. Lim, B. Fielding, Y. Yu, E. Anderson, Evolving ensemble models for image segmentation using enhanced particle swarm optimization, IEEE Access 7 (2019) 34004–34019.
[44] Y. Xie, J. Zhang, Y. Xia, C. Shen, Semi-and Weakly Supervised Directional Bootstrapping Model for Automated Skin Lesion Segmentation, 2019 (arXiv preprint) arXiv:1903.03313.
[45] L. Bi, J. Kim, E. Ahn, A. Kumar, D. Feng, M. Fulham, Step-wise integration of deep class-specific learning for dermoscopic image segmentation, Pattern Recognit. 85 (2019) 78–89.
[46] D.E. Ilea, P.F. Whelan, Automatic segmentation of skin cancer images using adaptive color clustering, in: China-Ireland International Conference on Information and Communications Technologies, Hangzhou, China, 2006.
[47] D.E. Gustafson, W.C. Kessel, Fuzzy clustering with a fuzzy covariance matrix, in: 1978 IEEE Conference on Decision and Control Including the 17th Symposium on Adaptive Processes, IEEE San Diego, 1979, pp. 761–766.
[48] J. Bezdek, Pattern Recognition With Fuzzy Objective Function Algorithms, 1981, New York.
[49] ITU-R, Studio encoding parameters of digital television for standard 4: 3 and wide-screen 16: 9 aspect ratios, recommendation BT.601-7, in: International Telecommunication Union, Geneva, CH (March), 2011.
[50] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on featured distributions, Pattern Recognit. 29 (1) (1996) 51–59.
[51] T. Ojala, M. Pietikäinen, T. Mäenpää, Gray scale and rotation invariant texture classification with local binary patterns, in: European Conference on Computer Vision, Springer, Heidelberg, Germany, 2000, pp. 404–420.
[52] T. Mendonça, P.M. Ferreira, J.S. Marques, A.R. Marcal, J. Rozeira, Ph 2-a dermoscopic image database for research and benchmarking, in: International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan, 2013, pp. 5437–5440.
[53] R.A. Haddad, A.N. Akansu, A class of fast gaussian binomial filters for speech and image processing, IEEE Trans. Signal Process. 39 (3) (1991) 723–727.
[54] Joint ISO/CIE Standard, ISO 11664-4:2008(E)/CIE S 014-4/E:2007, Colorimetry-Part 4: CIE 1976 L* a* b* Colour Space, 2007.
[55] D. Arthur, S. Vassilvitskii, k-means++: the advantages of careful seeding, in: Proceedings of the ACM-SIAM symposium on Discrete algorithms, New Orleans, USA, 2007, pp. 1027–1035.
[56] G. Argenziano, H. Soyer, V. De Giorgi, D. Piccolo, P. Carli, M. Delfino, Interactive Atlas of Dermoscopy (Book and CD-ROM), EDRA Medical Publishing & New Media, Milan, 2000.
[57] L. Ballerini, R.B. Fisher, B. Aldridge, J. Rees, A color and texture based hierarchical k-nn approach to the classification of non-melanoma skin lesions Color Medical Image Analysis, vol. 6, 2013, pp. 63–86.
[58] M.E. Celebi, S. Hwang, H. Iyatomi, G. Schaefer, Robust border detection in dermoscopy images using threshold fusion, in: IEEE International Conference on Image Processing, Hong Kong, Hong Kong, 2010, pp. 2541–2544.
[59] M. Silveira, J.C. Nascimento, J.S. Marques, A.R. Marçal, T. Mendonça, S. Yamauchi, J. Maeda, J. Rozeira, Comparison of segmentation methods for melanoma diagnosis in dermoscopy images, IEEE J. Sel. Top. Signal Process. 3 (1) (2009) 35–45.
[60] M.E. Celebi, H.A. Kingravi, H. Iyatomi, J. Lee, Y.A. Aslandogan, W. Van Stoecker, R. Moss, J.M. Malters, A.A. Marghoob, Fast and accurate border detection in dermoscopy images using statistical region merging Medical Imaging: Image Processing, vol. 6512, SPIE, San Diego, California, United States, 2007.
[61] M. Bayraktar, S. Kockara, T. Halic, M. Mete, H.K. Wong, K. Iqbal, Local edge-enhanced active contour for accurate skin lesion border detection, BMC Bioinform. 20 (2) (2019) 91.
[62] R. Kéchichian, H. Gong, M. Revenu, O. Lézoray, M. Desvignes, New data model for graph-cut segmentation: application to automatic melanoma delineation, in: 2014 IEEE International Conference on Image Processing (ICIP), IEEE, 2014, pp. 892–896.