sciendo

Int. J. Appl. Math. Comput. Sci., 2022, Vol. 32, No. 4, 683–699
DOI: 10.34768/amcs-2022-0047

amcs

SEGMENTATION OF THE MELANOMA LESION AND ITS BORDER

GRZEGORZ SURÓWKA [a,*], MACIEJ OGORZAŁEK [a]

[a]Department of Physics, Astronomy and Applied Computer Science
Jagiellonian University
ul. prof. Stanisława Łojasiewicza 11, 30-348 Kraków, Poland
e-mail: grzegorz.surowka@uj.edu.pl

Segmentation of the border of the human pigmented lesions has a direct impact on the diagnosis of malignant melanoma. In this work, we examine performance of (i) morphological segmentation of a pigmented lesion by region growing with the adaptive threshold and density-based DBSCAN clustering algorithm, and (ii) morphological segmentation of the pigmented lesion border by region growing of the lesion and the background skin. Research tasks (i) and (ii) are evaluated by a human expert and tested on two data sets, A and B, of different origins, resolution, and image quality. The preprocessing step consists of removing the black frame around the lesion and reducing noise and artifacts. The halo is removed by cutting out the dark circular region and filling it with an average skin color. Noise is reduced by a family of Gaussian filters $3\times3-7\times7$ to improve the contrast and smooth out possible distortions. Some other filters are also tested. Artifacts like dark thick hair or ruler/ink markers are removed from the images by using the DullRazor closing images for all RGB colors for a hair brightness threshold below a value of 25 or, alternatively, by the BTH transform. For the segmentation, JFIF luminance representation is used. In the analysis (i), out of each dermoscopy image, a lesion segmentation mask is produced. For the region growing we get a sensitivity of 0.92/0.85, a precision of 0.98/0.91, and a border error of 0.08/0.15 for data sets A/B, respectively. For the density-based DBSCAN algorithm, we get a sensitivity of 0.91/0.89, a precision of 0.95/0.93, and a border error of 0.09/0.12 for data sets A/B, respectively. In the analysis (ii), out of each dermoscopy image, a series of lesion, background, and border segmentation images are derived. We get a sensitivity of about 0.89, a specificity of 0.94 and an accuracy of 0.91 for data set A, and a sensitivity of about 0.85, specificity of 0.91 and an accuracy of 0.89 for data set B. Our analyses show that the improved methods of region growing and density-based clustering performed after proper preprocessing may be good tools for the computer-aided melanoma diagnosis.

Keywords: computer-aided diagnosis, DBSCAN, melanoma, region growing, segmentation.

1. Introduction

Melanoma is a tumor of the pigment cells in the epidermis. It is usually acquired during the lifetime *de novo* or from malicious transformation of the benign forms of pigmented lesions called displastic (atypical) nevi. Although non-melanocytic types of skin cancer are more common, melanoma is the most malignant human cancer (it may affect young people), and its mortality is increasing year by year by 4–6% (ACS, 2020).

The diagnosis of melanoma is based on observations, first with bare eyes (patient self-examination, family doctors), then by a dermoscope (dermatologists) (Celebi *et al.*, 2019). When the lesion expands horizontally and reveals some distinctive geometric/color features, it should be soon excised before it starts vertical expansion to the dermis. It is essential to resect the malignant lesion at an early, non-metastatic stage. Unfortunately, visual examinations at the early stages of the melanoma process may often lead to a false diagnosis, even by experienced specialists. Since excision of the lesion is the ultimate procedure, non-invasive, computer-based systems supporting medical diagnosis are of key importance (Vestergaard *et al.*, 2008).

Computer acquisition systems integrated with dermoscopes can store and compare, exchange (telemedicine), and attempt to (roughly) analyze dermoscopic images. Offline computer systems are often used to help medical doctors to take a decision as to a biopsy, if clinical (dermoscopy) diagnosis is inefficient (no visual melanoma features, so-called featureless or

*Corresponding author

small melanoma, or lack of experience). The latter is an example of computer aided diagnosis (CAD) and is supposed to increase the accuracy of a diagnosis, plus it may reduce the time and cost of treatment (even if formally no CAD system is standardized and approved for use in clinical settings).

Medical doctors diagnose pigmented skin lesions by means of some descriptive metrics (Celebi *et al.*, 2007). They include the ABCD(E) rule, the 7-Point Checklist, Menzies, 7FFM, CASH, etc. (Jensen and Elewski, 2015), to mention the most common. The ABCD(E) rule is a semi-quantitative analysis of five criteria: (A)symmetry, irregular (B)order, variety of (C)olors and (D)ifferential Structural Components found in the lesion (dots, globules, pseudopods, a white veil, featureless areas). Sometimes an extra factor is taken into account, which is the (E)volutionary change of the lesion. Each criterion is weighted to contribute to the total dermoscopic score (TDS) classifying the lesion into benign, suspicious, and melanoma classes.

The standard melanoma CAD (Masood and Al-Jumaily, 2013; Oliveira *et al.*, 2018; Mishra and Celebi, 2016) is based on segmentation methods to reveal and/or enhance those above-mentioned visual features of the neoplasm.

In particular, border irregularity is the most significant factor to discriminate melanoma from benign pigmented lesions (Keefe *et al.*, 1990). This is due to the uneven spread of melanocytes and/or the immune system regression pattern on the lesion surface. The lesion segmentation itself is a fundamental step before feature extraction and classification of melanoma. The development of segmentation methods directly influences the performance that a CAD system for the automatic diagnosis of pigmented skin lesions can achieve. It is relevant for the following reasons:

- it helps directly identify some of the criteria of the clinical metrics,
- it makes possible to extract some derivative information on lesion characteristics, e.g., diameter, skewness, etc. and,
- is the determinant part of the subsequent steps of the CAD pipeline.

There exist different taxonomies for skin lesion segmentation methods. According to Celebi *et al.* (2015), the segmentation methods can be divided into the three main groups:

- unsupervised (computer vision) methods that use no training data,
- supervised (shallow) learning methods that train and use classifiers,
- supervised deep learning methods.

The unsupervised segmentation methods use properties of the color space to cluster pixels into homogenous regions (Møllersen *et al.*, 2010). This is a broad field divided according to the topological (similarity) information derived from the pixels. Pixel-based (local) segmentation methods (e.g., thresholding) are based on the intensity of separate pixels. A threshold value is usually calculated by analyzing predefined image features, e.g., intensity histograms. Such methods are simple, robust, and not computationally intensive, but they only work if there is enough contrast between the region of interest (ROI) and the background (BG). They are further limited by the inhomogenous intensity distribution of the lesion and may fail if the distribution contains multiple peaks (Celebi *et al.*, 2013).

Edge-based segmentation methods are aimed at detecting discontinuities in the pixel intensity. They make use of differential (gradient) operators (e.g., Sobel, Prewitt, Canne, Roberts, etc.) to delineate the edge between ROI and BG (Sadeghi *et al.*, 2011). Pixel-, and edge-based segmentations are sensitive enough only if there is a distinct border between ROI and BG. They are not efficient in the soft/fuzzy transition between ROI and BG. Image characteristics such as edges, smoothness, or other statistical distributions may be used in some energy/cost functions which, when minimized, determine the boundaries of ROIs.

Contour lines are primarily used in digital elevation models (DEMs) for determining slopes and generating a surface. There exists many contour tracing algorithms: square tracing, Moore neighbor, radial sweep, etc. (Pradhan *et al.*, 2010). Other examples are active contour models that may utilize metaheuristics and genetic algorithms (Zhou *et al.*, 2011). Active contour models are based on deformable splines (so-called snakes), which are selected according to the desired contour shape or learned from a training set. The snake is then matched to the contour of the object by some 'energy minimization' operator. Such a procedure may strongly depend on interaction with a user, other examples (adjacent in space or time), or a higher-level knowledge (Mete and Sirakov, 2010).

Region-based (global) segmentation looks for regions satisfying a given homogeneity criterion (color, texture, brightness). One can merge small regions into larger ones (region merging or region growing), or split the image (step-by-step) into areas with similar pixel properties (region splitting). The split&merge methods first split the image into a number of dissimilar areas, and then merge the most similar ones. Watershed segmentation methods, on the other hand, interpret the brightness of pixels as an elevation above the sea level,

and analyze the slopes to determine the closed 'water reservoirs', i.e., similar regions (Wang *et al.*, 2011). There are also iterative/statistical region-merging methods recursively merging pixels or regions in a hierarchical manner (Celebi *et al.*, 2008).

Recently other unsupervised methods have become popular: saliency-based segmentation, Delaunay triangulation (Pennisi *et al.*, 2016), sparse coding methods (Bozorgtabar *et al.*, 2016), or cellular automata (Bi *et al.*, 2016), etc. Saliency segmentation (Ahn *et al.*, 2017) computes the most informative region in an image (e.g., by color features) based on human perception. For that task, pixel quality maps are produced. Such representations are more meaningful and easier to analyze, but they require high resolution and computational efficiency. The supervised segmentation methods segment the skin lesion by training classifiers (e.g., shallow artificial neural networks, support vector machines, decision trees, etc.) that separate the skin lesions from the surrounding healthy skin. For that purpose, those methods extract features on different hierarchy levels (pixel, regions, colors and textures) (Ashour *et al.*, 2018).

Quite new, but promising for skin lesion segmentation is the paradigm of deep learning (DL) (Esteva *et al.*, 2017). Current DL-based segmentation methods (Goyal *et al.*, 2019; Codella *et al.*, 2019) mainly use different architectures of convolutional neural networks (CNNs), especially fully convolutional networks (FCNs). They are able to derive image-wide semantic information from a full hierarchy of features. Such an approach is robust and successful in segmentation and/or classification of dermoscopy images (Bi *et al.*, 2017). More detailed information on these methods can be found in the current reviews of segmentation techniques used in the skin lesions (e.g., Pathan *et al.*, 2018b).

Despite advances in the field and ample literature, the segmentation step of melanoma lesions in dermoscopy images is still unsatisfactory and not fully tested. Here efficiency, complexity and stability (regarding the data sets and within a given data set) come into play. Efficiency of segmentation is still subject to noise and artifacts, conditions for reception of dermoscopy images, and regression patterns. In this paper we propose a method for segmentation of the lesion border based on region growing (RG) with an automatic detection of the seed points/threshold and a comparative study of segmentation of the pigmented lesion (location, shape, size) by RG and a density clustering algorithm (DBSCAN). We quantitatively evaluate the method on two different data sets, one publicly available database included.

In the following sections, we present all the steps of our experimental procedure, show the results, and discuss their importance in terms of the available literature on the subject.

2. Materials and methods

In this section, we describe the process of preparation (with two different procedures for hair removal), segmentation (by two methods, i.e., a 'classic' RG and the DBSCAN algorithm) and, finally, diagnostic assessment of the pigmented lesions from dermoscopy images. The pipeline can be summarized as follows:

1. Preprocessing:
   (a) black frame: crop/floodfill,
   (b) filtering: median/Gaussian/ADF/Sigma/Bilateral/Non-Local Means,
   (c) hair removal: Black Top Hat/DullRazor.
2. Pigmented lesion segmentation (Task 1):
   (a) RG,
   (b) DBSCAN.
3. Pigmented lesion border segmentation (Task 2):
   (a) RG.
4. Assessment.

2.1. Preprocessing. Prior to the segmentation procedure, the data are preprocessed in three steps: (i) removing the black frame (halo), (ii) noise reduction, (iii) removing artifacts in the image.

Not all, but some dermoscopy pictures are taken with dark corners, where the illumination drops to an arbitrarily low value. This may produce a significant gradient that can affect the segmentation results. To automatically improve such images, two strategies can be applied. The first, and more common strategy is to remove those rows and columns of the image whose content exceeds a certain fraction of the dark pixels. The luminance of the 'dark' pixel is determined over the full data set (normalized luminance is usually less than 20%). Another strategy is to find a circle that surrounds the bright center of the image and fill all pixels outside with an 'average' skin color. The latter method is safer for the segmentation result.

In order to improve the image quality and thus the expected segmentation result, one should reduce the noise. Noise reduction is performed through image filtering. This is the key step for the detection of borders in dermoscopy images, since: (i) it improves the contrast, (ii) smooths out noise/distortions, and finally, (iii) it helps remove artifacts like thin or light hairs, droplets of immersion fluid/sweat/water, glare of light, etc. in the last stage of preprocessing. Filters can be applied to each RGB component separately, or on the luminance representation (see further).

One disadvantage of filtering is the blurring of local structures, which is especially prominent with a simple

low-pass filter (such a filter is not used). The most important aspect of filters in terms of segmentation is the ability to preserve the edges. Kernel-like, median-like, anisotropic diffusion filters (ADFs), Sigma, Bilateral and Non-Local-Means filters may set a trade-off between the image contrast and the image blur. After some experiments with the filter masks in the range from 3×3 to 9×9, we took for the reference data set A a Gaussian filter ($\sigma = 1$) of size 7×7 (for data set B the mask was smaller, 5×5 or 3×3, proportional to the image size). The median filter was of the same quality but introduced distortions in fine-grained objects. The ADFs were almost of the same quality (for data set A), but more computationally intensive. The other filters gave worse results (see Tables 2 and 3).

The last preprocessing step involves removing big artifacts. Big artifacts that can be found in the dermoscopy images are hairs and ruler markers (less often ink markers). Dark, thick hairs (or ruler ticks) covering a part of each lesion were removed by using either the DullRazor algorithm (Lee *et al.*, 1997) or the Black Top Hat method (Jaworek-Korjakowska and Tadeusiewicz, 2013).

The DullRazor algorithm (Lee *et al.*, 1997) derives a hair mask which is a union of individual masks for three color bands R, G, B, where each such a color mask is a generalized (over 0, 45, and 90 degrees) closing image for a hair brightness threshold (arbitrarily) below a value of 25 (in the range of 0–255).

The alternative algorithm of Jaworek-Korjakowska and Tadeusiewicz (2013) derives hairmasks (or masks of other dark elongated objects) with the help of the Black Top Hat (BTH) transform, $\mathrm{BTH}(I) = I \bullet b - I$, where I stands for an image, subject to the transform (luminance), b stands for a structuring element, $\bullet$ denotes the morphological closing operation (erosion(dilation)).

One can impose different thresholds analyzing the image in a finer or a coarser scale. Pixels of the detected hairs were replaced with a mean value of the surrounding pixels.

Color features play a crucial role in the analysis of dermoscopy images for discrimination between benign and malicious skin lesions (Stanley *et al.*, 2007). For the segmentation process, we tested luminance-based color models (YC_rC_b, HSL, HSI) and finally used the normalized $Y'C_rC_b$ representation from the JFIF standard: RGB-elementary colors $\in [0, \ldots, 255]$, where

$$Y', Cb, Cr \in [0, \ldots, 255],$$
$$Y' = 0.299 \times R + 0.587 \times G + 0.114 \times B,$$
$$Cr = 128 + 0.5 \times R - 0.418688 \times G - 0.081312 \times B,$$
$$Cb = 128 - 0.168736 \times R - 0.331264 \times G + 0.5 \times B.$$

2.2. Region growing. As we pointed out in Introduction, segmentation of the pigmented skin lesion is the first, essential step in the pipeline of (computer-aided) clinical diagnosis of melanoma. The border detection problem is related to the lesion segmentation but goes even further. It attempts to derive the lesion border itself as the feature B of the ABCD rule. Discrimination of irregular, fuzzy borders that can be a sign of malignancy or (at least) a sign of suspicious atypical growth, from regular, 'sharp' borders, which are common for benign lesions, plays a crucial role as a diagnostic factor for (expected) malignancy. Therefore in this work, we use (selectively) our segmentation methods both for the quantitative determination of the lesion in the dermoscopy image and the lesion border.

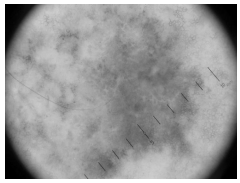

Fig. 1. Preprocessing step: elimination of the black halo. Strategy: draw a circle that cuts off the black region(s).

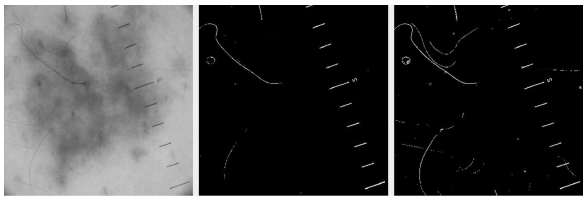

Fig. 2. Preprocessing step: elimination of artifacts like hairs, droplets of fluids, rulers, etc. by filtering and the DullRazor/BTH algorithm. Left: input dermoscopy image, middle: DullRazor, right: BTH mask of the artifacts.

The latter is a two-step segmentation: first we segment the lesion and then the skin (background). As a result, we get two binary masks. The difference between the two segmentation masks determines the border and thus indicates the type of lesion (benign or malicious).

As for the tools, we take advantage of a simple region growing method and one density-based method. The segmentation procedure by region growing (RG) can be described in the following way. Each particular region is grown from a single starting seed point, which must definitely belong to the proper class. The region is iteratively grown by adding unallocated adjacent pixels that resemble the initial seed point. The procedure is based on similarity/continuity criteria between the pixels, and the extent of the changes is controlled by a structuring element (a set of pixels showing the extent of the neighborhood).

Only pixels that differ from the region average intensity less than a specified threshold are allocated to the respective region. The growth continues until the full

Algorithm 1. Triangle method.

Require: H_L {Histogram of the image luminance}
1: draw a line $l1$ between the peak and the first nonzero channel of H_L
2: draw a line $l2$ perpendicularly crossing $l1$ at a point S, where the distance from $l1$ to H_L is maximum
3: **return** S {the cut-off level}

Algorithm 2. RG-A.

Require: I {Image}
1: determine the cut-off level S of I {from the triangle method}
2: determine the luminance L of I
3: $R = 3 \times 3$ square {structuring element}
4: **for** $R \in L$ **do**
5: calculate $Avg(R) = \sum_{p \in R} L_p / \#(p \in R)$
6: $diff = 0$
7: **while** $(diff \leq |Avg(R) - S|)$ **do**
8: {region accumulation}
9: **for** $p \in R$ **do**
10: $diff\,(p, R) = |L(p, R) - Avg(R)|$ {p: neighbors in the region}
11: **end for**
12: $seed = \arg\min(diff\,(p))$
13: $R := R \setminus seed$
14: **end while**
15: **end for**

homogenous region is determined (no adjacent points can be attached to the region).

There are three hyperparamaters of the method:

- the seed point that starts the expansion of a region,
- the structuring element,
- the threshold, necessary for the partition of the image into homogeneous regions.

For regular, well-centered lesions, one can use a 'naive' way to fix the seed points for the lesion and the background region. For the lesion seed point, the center of the analyzed image can be selected.

If the black frame of the dermoscopy image has already been removed, the background skin color can be estimated from the corners of the image. For the moles, which are completely inside the image, the seed point can be sets $x = 1$, $y = 1$ (or any other corner). If a lesion is displaced from the center, or is too big (it is indicated by the fact that there can be no full bounding circle around it), the background segmentation procedure can be attempted from any bright corner and then the resulting 'background' region usually consists of several smaller regions. In such cases, the separate background masks should be merged together to form a homogenous region for the final assessment. The 'naive' procedure described above for the lesion and the background seed point is problematic when dealing with nonstandard (too big or misplaced) skin lesions in the dermoscopy images. An automatic and more robust procedure is required.

The choice of the seed points is adapted from Smaoui and Bessassi (2013). A region of lesion is assumed to be darker than the surrounding skin (background). Potential pixels are ordered into a list according to a minimum value of the gray level. For each pixel from this list, an average value is calculated in a 3×3 window. For the seed point, this pixel is selected, which has the minimum difference with its neighbors.

For the background, the same algorithm is applied, except that we now search for the brightest pixels in their homogeneous neighborhoods.

Our structuring element (neighborhood) is a 3×3 square (Smaoui and Bessassi, 2013). As an alternative, this can also be a 3×3 cross of distance of 1 (Kroon, 2004) with a neigborhood matrix to loop for x and y $[-1, 0; 1, 0; 0, -1; 0, 1]$ (the results from both assumptions are statistically comparable). Larger structuring elements are less accurate.

Determination of the threshold and the accumulation of regions is explained by an algorithm presented herein, after the idea found in the work of Smaoui and Bessassi (2013)—we call it RG-A, and alternatively in that of Kroon (2004) (implemented in Matlab)—we call it RG-B.

At the end, a morphological closing (with the diameter of 5 pixels) is produced to remove some isolated pixels.

2.3. Density-based segmentation. Density-based clustering algorithms respond to the general problem of segmentation that can be summarized by the four requirements that we usually have about the data: (i) no *a priori* knowledge about the number of clusters, (ii) clusters with arbitrary shapes, (iii) the presence of outliers and noise in the data and, (iv) density variation and scarcity of data. Some popular density-based algorithms on the market are DBSCAN (Ester *et al.*, 1996), DENCLUE (Hinneburg and Keim, 1998), OPTICS (Ankerst *et al.*, 1999), and MDCUT (Louhichi *et al.*, 2017).

The main idea of pixels in the DBSCAN algorithm is explained below. The pixels can be classified into three types (Louhichi *et al.*, 2017):

- a core point is a point whose neighborhood ($radius < eps$) includes more than $minPts$ points,
- a border point is a point whose neighborhood ($radius < eps$) includes less than $minPts$ points, but is reachable from a core point,

Algorithm 3. RG-B.

Require: I {Image}
1: $threshold = 0.20$
2: determine the luminance L of I
3: $R = 3 \times 3$ cross {structuring element}
4: **for** $R \in L$ **do**
5: calculate $Avg(R) = \sum_{p \in R} L_p / \#(p \in R)$
6: $diff = 0$
7: **while** ($diff \leq |Avg(R) - threshold|$) **do**
8: {region accumulation}
9: **for** $p \in R$ **do**
10: $diff\,(p, R) = |L(p, R) - Avg(R)|$
{p: neighbors in the region}
11: **end for**
12: $seed = \arg\min(diff\,(p))$
13: $R := R \setminus seed$
14: **end while**
15: **end for**

Algorithm 4. DBSCAN.

Require: I {Image}
1: hyperparameters: eps, $minPts$
2: **for** (all not yet assigned points $p \in I$) **do**
3: select a point p
4: identify $N \in neighborhood_{eps}(p)$
5: **if** ($size(N) < minPts$) **then**
6: $C(N) \leftarrow p$ {form a cluster}
7: **else**
8: $N' : C'(N') \leftarrow p$
{form and identify a new cluster}
9: **end if**
10: **end for**

- a noise point (outlier) is neither the core, nor the border point.

Based on these definitions, two points can be directly density reachable (when they are within eps, and the other one is a core point), density reachable (when there exists a sequence of points between the two points and each pair of points in the sequence is directly density reachable), and finally, density connected (when there exists a point such that the two points are directly density reachable to it). With the two input parameters, eps and $minPts$, DBSCAN can start clustering from an arbitrary point and keep looking for other points that are directly density reachable (Suer *et al.*, 2011). A sparse neighbourhood (with respect to eps and $minPts$) cannot develop a cluster (the points are then labelled as noise points), otherwise a cluster is initiated and iteratively built up. If the cluster cannot be expanded any further, another unlabelled point is taken and is subject to the procedure.

The above mentioned concepts are shown graphically by, e.g., Kockara *et al.* (2010) and Suer *et al.* (2011). The pseudocode of DBSCAN is presented by Kockara *et al.* (2010) and in more detail by Ester *et al.* (1996) and Kockara *et al.* (2010); an efficient implementation based on kd-trees was first developed by Bentley (1975).

DBSCAN starts by mapping the data into a kd-tree data structure and identifying the (Euclidean) distance between each point and its k-th nearest neighbor (k-dist). Such queries are stored in a matrix of sorted distances (the first column contains values of the nearest neighbors, second is for the second nearest neighbors etc.).

The main steps of DBSCAN (Ester *et al.*, 1996) can be shown in the following simplified scheme: A data point can get to a cluster (when below a certain k-distance), start a new one, or remain as a noise/not classified point. In this work, we test the DBSCAN algorithm as an extention to the original region growing segmentation method.

All numerical calculations, i.e., preprocessing (filtering, BTH, DullRazer), segmentation (RG-A/RG-B, DBSCAN), and postprocessing (metrics) were performed in the R environment (Hornik, 2021). Our motivation for choosing the environment was the choice of a ready-made and community-recognized implementation of the density-based algorithms, that is "Density Based Clustering of Applications with Noise (DBSCAN) and Related Algorithms—R package" on the GitHub (Hahsler, 2021). This R package provides a fast, community-recognized, C++-based implementation of several density-based algorithms for clustering data which, last but not least, also includes the fast nearest neighbor search implementation using kd-trees (Hahsler, 2021).

2.4. Postprocessing. The third step is a quantitative estimation of compliance between the performed automatic segmentation methods RG and DBSCAN and the ground truth.

For data set A, manual borders were obtained by an expert dermatologist in Matlab/ Image Segmenter (data set A) and provided by the public repository (data set B).

For the lesion segmentation task, in the case of a fuzzy, smeared border, the edge was painted to the maximum lesion extent. For data set B, segmentation masks were part of the data set.

We divided our research into two tasks discussed below.

Task 1: Lesion segmentation. To assess the lesion segmentation process, we should take into account two factors characterizing the binary mask: location in the image and extent and orientation.

There exist many metrics characterizing these aspects in the literature. Before we describe the main approach, i.e., metrics based on the information retrieval

Table 1. DBSCAN parameters.

source	$minPts$	eps	remarks
Ester *et al.*, 1996	$\leq$ (data dim+1)	5	–
Ester *et al.*, 1996	10	'chosen best'	to compare DBSCAN and MDCUT
Hahsler, 2021	5	–	–
Louhichi *et al.*, 2017	10	0.05–17	eps to fit different data sets
Kockara *et al.*, 2010	60	5	–

terminology, we report on metrics found in the work of Agarwal *et al.* (2017).

Agarwal *et al.* (2017) compare different binary masks by geometrical features like solidity and extent. Solidity is defined as (the ratio of the area to the convex hull area) of the binary mask. The extent, on the other hand, is defined as (the ratio of the area to the bounding box area). As the final (best) binary mask to the one with the largest solidity and extent. The overlap of such a binary mask with the ground truth marked by a medical expert is calculated by the so-called overlapping score (%) as a fraction of the intersection and the union of the two compared binary masks. They also define a more precise correlation coefficient taking into account the exact location of the segmented lesion with respect to the ground truth lesion. These metrics quantify the first and second factors in the above list.

We do not use such metrics for two reasons: (i) they are not popular and therefore not easily comparable with the results of other segmentation works in the field, (ii) rather than making a production system that picks the best method for each image, we want to test the methods in general.

The metrics based on the information retrieval terminology quantify how much the automatic border (AB) is different from the manual border (MB), the ground truth. They define the following basic categories (Kockara *et al.*, 2010), where all definitions are meant for the number of pixels in a particular region:

- TP: indicates the intersection between AB and MB, so a correct lesion region found by AB,
- TN: shows healthy region (background), both AB and MB agree on,
- FN: shows a region of the missed lesion, i.e., MB-AB,
- FP: shows erroneous positive regions, i.e., AB-MB.

Having these definitions, one can build common metrics:

- $Sensitivity = TP/(TP + FN)$,
- $Precision = TP/(TP + FP)$,
- $BorderError = Area(AB \oplus MB)/Area(MB) = (FP+FN)/(TP+FN)$, where $\oplus$ = XOR; currently the most important metric for assessing the quality of any automatic border detection algorithm (Pennisi *et al.*, 2016; Suer *et al.*, 2011)

Specificity and accuracy are more arbitrary due to the TN category, which is not well normalized (the area of the background skin can be unlimited).

- $Specificity = TN/(TN + FP)$,
- $Accuracy = (TP + TN)/(TP + FN + FP + TN)$.

Using the dermatologist-determined MBs, and AMs obtained from RG and DBSCAN for each segmented lesion, we calculate the basic numbers (TP, FN, TN, FP), from which information-retrieval-based definitions can be calculated, e.g., those listed above: sensitivity, precision and border error.

Task 2: Lesion border segmentation. For the lesion border segmentation, the segmented border (AB) is assessed by an expert dermatologist (MB) to comply with (a) the dermoscopy image, (b) with the hist-pat label (those are different factors as some lesions can visually fake the hist-pat examination, e.g., a benign lesion can look like the malignant one and vice versa). In this task we want to check how efficiently AM can recognize the fuzziness/regularity of the lesion border, and this to contribute to the irregular border criterion of the ABCD(E)'s diagnostic system.

As metrics we use here sensitivity (meaning the ability of the method to detect the melanoma), specificity (the ability of the method to detect the benign lesion) and (overall) accuracy.

3. Results and a discussion

3.1. Results. We did our experiments on two data sets from different sources. Data set A (Melanoma, ML 2018) consists of 185 images: 101 malignant melanoma (M) and 84 dysplastic nevus (D) cases. Those are 2272×1704 (JPEG) images collected using a Minolta Dimage Z5 digital camera equipped with an epiluminescence lens with white halogen lighting. This data set is available

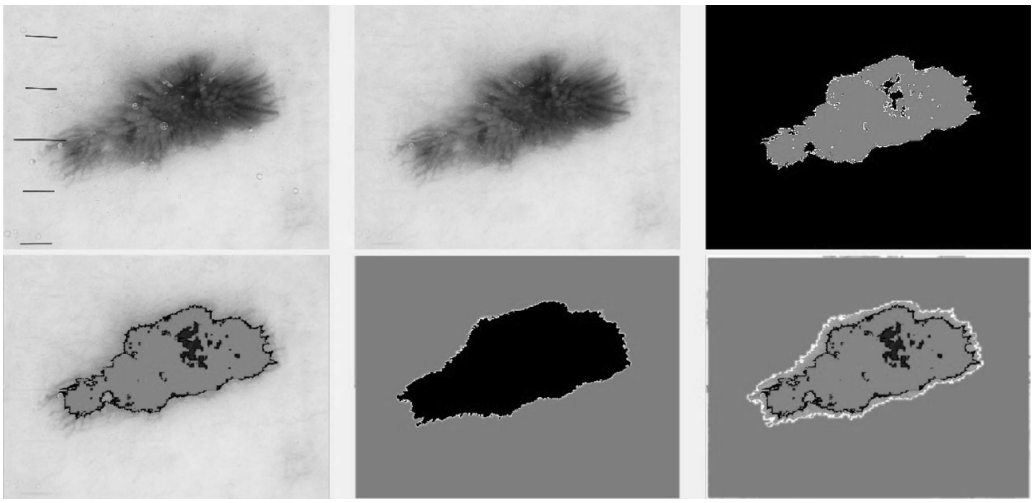

Fig. 3. RG-A determination of the lesion border: the case of a benign skin lesion. Top (from left to right): raw lesion, removal of hairs/tics and bubbles, lesion segmentation mask (gray). Bottom (from left to right): lesion segmentation superimposed on the image, background segmentation mask (gray), lesion border as the difference between the background and the lesion mask (the area between the black and white outline).

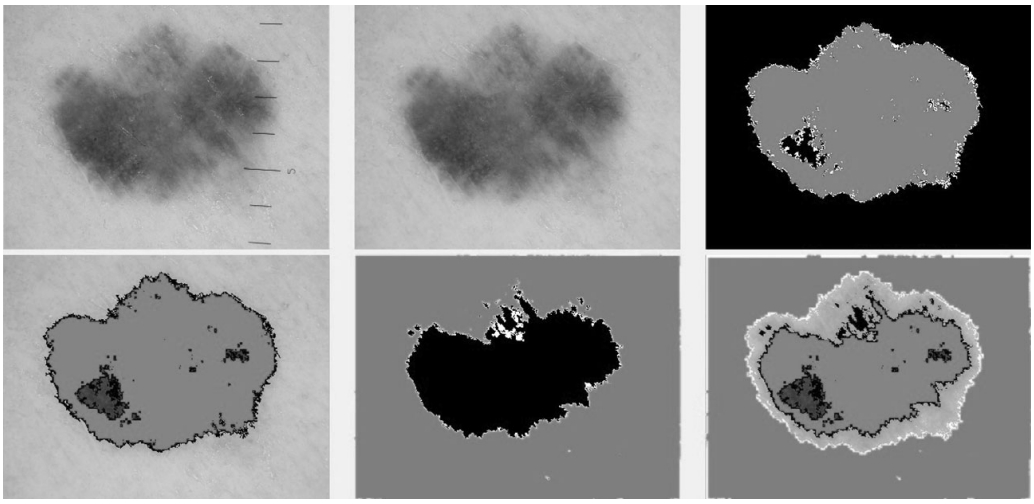

Fig. 4. RG-A determination of the lesion border: the case of a malignant skin lesion. Top (from left to right): raw lesion, removal of hairs/tics and bubbles, lesion segmentation mask (gray). Bottom (from left to right): lesion segmentation superimposed on the image, background segmentation mask (gray), lesion border as the difference between the background and the lesion mask (the area between the black and white outline).

online after registration (except for the labels, neither expert details, nor segmentation masks are available).

Data set B is a public data set PH^2 (Mendonça *et al.*, 2013) developed for research and benchmarking purposes, which contains 40M and 77D cases (common nevi were removed from the data set since our discrimination task only applies to melanoma and displastic moles). Images B (BMP) were collected using Tuebinger Mole Analyzer and have a resolution of 768×560 and a magnification of about $20\times$. There is auxiliary information in this database provided by expert dermatologists: binary masks for the lesion and the background of the analyzed dermoscopy images, plus the class label (from the hist-pat examination), and other important dermoscopic structures.

Databases A and B include images of only white Caucasian skin types. In our research the ground truth about the class of the lesion is based on the declared hist-pat examination. For the quality of the lesion segmentation in data set A, we used our own clinical expert dermatologist.

The preprocessing step is rather straightforward. In Fig. 1 we can see a typical black halo around the skin lesion. In this case (data set B), the lesion is centered

in the image and covers a large amount of area of the image. The first strategy to remove the halo described in Section 2.1 would not be economical in this case. It could remove too much of the lesion area. Thus, here we draw a circle surrounding the bright center of the image and fill all pixels outside with the 'average' skin color.

An example of filtering and the DullRazor/BTH procedure is shown in Fig. 2 for an image from data set A. In this case we took a Gaussian filter ($\sigma = 1$) with a mask of 7×7 to smooth the image (left). The BTH transform (far right) is apparently better than DullRazor (middle), hence the BTH masks of artifacts to be removed were further used for segmentation.

After the preparation of the dermatoscopic images, we performed segmentation according to the previously described algorithms. As for the seed points, we did not follow the 'naïve', but adapted the 'minimum difference' (Smaoui and Bessassi, 2013) method.

Let us start from the more straightforward RG-B method. RG-B is rather fast, but it works definitely good (for numerical results are presented in the remainder of this section) only for (i) homogenous lesions with a well-defined border separating the lesion content from the background healthy skin (i.e., mainly benign lesions) and, at the same time, (ii) well localized lesions, i.e., centered in the dermoscopy image and with a regular, evenly illuminated background. In this case the growth of either region does not depend strongly on the choice of the seed point (because it is always well chosen), nor on the 'stability' of the threshold(s).

For the majority of cases (in both tested data sets), this approach fails unless we carefully adjust the procedure to each individual image. However then the analysis is no longer 'automatic', algorithmwise, but manual.

RG-A is much better (for numerical results, see the remainder of this section) mainly due to the adaptive threshold embedded in the algorithm. An example of how algorithm RG-A works is shown in Figs. 3 and 4 below. What we focus on is the lesion border, which is the difference between the background and the lesion mask. In Fig. 4 we can see a case of Malignant Melanoma with a fuzzy, smeared border, the indicator (B in the ABCD metric) to diagnose it as a tumor.

Experiments on the density-based algorithm DBSCAN were performed in the R package. We took advantage of the full 'infrastructure' found in (Hahsler, 2021), where DBSCAN plus some important auxiliary procedures (density calculation, kd-tree structure, k-NN search, etc.) are already implemented.

The local density at each data point was defined as the number of points within a given neighborhood of radius eps (including the query point itself). For a fast, distance-based nearest neighbor search, the kd-tree structure was used. We took the default values of

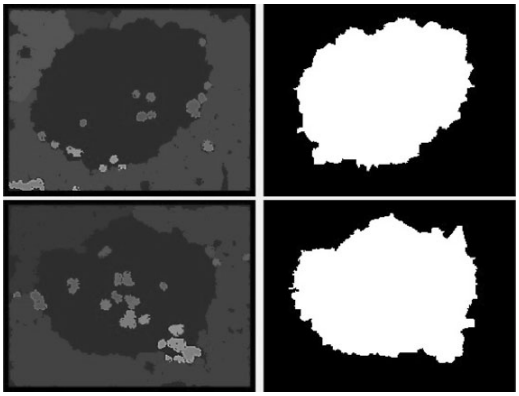

Fig. 5. DBSCAN segmentation (left) and its extra binarization (right). Top: case of a benign lesion, bottom: case of a malignant lesion.

the parameters: the maximum size of the kd-tree leafs (bucketSize) as 10, and the split rule (splitTRule) as 'best guess'.

DBSCAN, implemented following Hahsler (2021), Mount and Arya (2010), Hahsler *et al.* (2019) and Campello *et al.* (2013), estimates the density around each data point in an eps neighborhood and applies a $minPts$ threshold (the minimum density for nonnoisy areas) to identify core, border and noisy points. The core points are joined into a cluster if they are density-reachable (i.e., there is a chain of core points where one falls inside the eps-neighborhood of the next). Finally, border points are assigned to clusters. Table 1 shows values for the user-defined parameters $minPts$ and eps of DBSCAN found in the literature.

Increasing $minPts$ in principle suppresses noise, as it requires more points to form a cluster. After some experiments, we fixed $minPts = 60$ and $eps = 5$ (Kockara *et al.*, 2010), to be good for both data sets A and B, where data set A was subject (only for DBSCAN) to a two-fold reduction in resolution from 2272×1704 to 568×426. The resolution reduction was also beneficial for the DBSCAN segmentation in data set A.

DBSCAN is very time demanding for region queries (for comparison, a time of 102 s for a single full-resolution image of data set A is required). A major drawback of DBSCAN is that no desired number of clusters can be set up in the algorithm. As a result, although the general shape and size of the lesion are detected (the main cluster), some extra small clusters may, and usually are, present in the image (as expected, more for smeared, malignant-like lesions). This is due to the nonuniformity of color within the lesion and the normal skin. Although, for a human assessment, the DBSCAN mask for a lesion is fairly comparable with the ground truth, the small clusters should be removed for a fully automatic determination of the lesion mask. This can be

Table 2. Evaluated lesion segmentation methods on data set A. Presented metrics are the average metrics calculated for individual images.

filter: Gaussian ($\sigma = 1$) 7×7			
method	sensitivity	precision	border error
RG-A	0.92	0.98	0.08
Bin/DBSCAN	0.91	0.95	0.09
filter: median 7×7			
method	sensitivity	precision	border error
RG-A	0.92	0.96	0.09
Bin/DBSCAN	0.91	0.96	0.09
filter: AFD exp. minimal 5 iter.			
method	sensitivity	precision	border error
RG-A	0.90	0.95	0.11
Bin/DBSCAN	0.89	0.91	0.14
filter: bilateral 2× variance			
method	sensitivity	precision	border error
RG-A	0.90	0.89	0.13
Bin/DBSCAN	0.85	0.89	0.15

Table 3. Evaluated lesion segmentation methods on data set B. Presented metrics are the average metrics calculated for individual images.

filter: Gaussian ($\sigma = 1$) 3×3			
method	sensitivity	precision	border error
RG-A	0.85	0.91	0.15
Bin/DBSCAN	0.89	0.93	0.12
filter: median 5×5			
method	sensitivity	precision	border error
RG-A	0.83	0.90	0.16
Bin/DBSCAN	0.88	0.91	0.12
filter: AFD exp. minimal 5 iter.			
method	sensitivity	precision	border error
RG-A	0.84	0.90	0.15
Bin/DBSCAN	0.85	0.88	0.16
filter: bilateral 2× variance			
method	sensitivity	precision	border error
RG-A	0.79	0.80	0.19
Bin/DBSCAN	0.86	0.90	0.14

done (morphological closing + thresholding, see Fig. 5) with less effort (on a presegmented image) than on a raw dermoscopy image. Henceforth, we will refer to this procedure as Bin/DBSCAN. Bin/DBSCAN cannot be used for segmenting the lesion border itself.

We evaluated the lesion segmentation (RG-A, Bin/DBSCAN) and border detection (RG-A) errors by comparing our results with dermatologist-determined boundaries performed in Matlab/ImageSegmenter (data set A) and provided by the public repository (data set B). Those binary masks plus the class label (from the hist-pat examination) are our ground truth. For data set A, in the case of non-sharp borders the contour includes the most outer region of the lesion (where the healthy skin begins), for data set B no such details are available.

The quality of the methods was assessed as a joint preprocessing (BTH) and segmentation step (RG-A, Bin/DBSCAN). We evaluated data set B to make an easier reference to the literature (PH^2 database).

Numerical results are presented in Tables 2 and 3 (Task 1) and Tables 4 and 5 (Task 2). Now we comment on the results.

The overall accuracy of the pigmented lesion segmentation methods (both RG-A and Bin/DBSCAN) may be influenced by the following factors:

- ground truth: the inter-dermatologist agreement on manual borders is not perfect, especially for lesions with a broad indefinite border; most dermatologists draw the largest possible region of influence of the lesion, but in fact it may not include some tumor area 'outside' the arbitrary border;

- the above argument also applies to segmentation: the threshold between the healthy skin (background) and the tumor may be by far arbitrary since the border may be imperceptible;

- ground truth: since MB we compare with (data set A) is obtained from a human expert, we cannot quantitatively evaluate how different local structures recognized in the lesion (dark and light ones) affect the decision (is it only the border?); if it plays a role, similar dermoscopy images (in the sense of the lesion edge) may differ in the extent of the lesion due to the additional structures (dots and globules, white/blue veil, pigment mesh, pseudopods, etc.);

- the above argument also applies to segmentation: (visually) homogeneous regions may obscure structures that affect the performance of the segmentation algorithm.

A comparison between RG-A and Bin/DBSCAN depends on the available data set. RG-A is performed on the nominal resolutions in both data sets, so the higher resolution (data set A) seems to better correspond to the region growing method. This is apparent both in Table 2 (Task 1) and in Tables 4 and 5 (Task 2). Bin/DBSCAN is applied to approximately similar resolutions of data sets (resolution of the data set A is reduced twice for Bin/DBSCAN). This may explain why Bin/DBSCAN in Table 2 (data set A) and Table 3 (data set B) are relatively equal (in sensitivity and precision).

For both RG-A and Bin/DBSCAN, data set B has apparently worse metrics which we think can be due to many corelevant factors:

If we consider the segmentation methods themselves, we can see that the detected lesions by RG-A are more

similar to those delimited by the experts (Task 1). We think this is mainly due to the adaptive threshold and seed points embedded in the algorithm, which surpasses the fixed threshold and seed points. The latter is the main deficiency of the 'classic' region growing methodology. RG-A is the only one among the tested methods that allows for border detection by the lesion and background segmentation.

The efficiency of DBSCAN depends on several parameters (directly: eps, $minPts$; in the sense of the distance structure: $bucketSize$, $splitTRule$) plus no desired number of clusters can be set up in the algorithm. All those are probably the reason why this method did not live up to our expectations. Our working hypothesis was that DBSCAN, as a density-based segmentation method, should perform better than region growing. It turned out that only the use of secondary binarization could preserve the advantages of the method. Although the general shape and size of the lesion was detected (the main cluster), some extra small clusters were present in the image (more for smeared, malignant-like lesions).

In terms of Task 2 we clearly see that the well-defined lesions (well-separated from the background) have similar lesion and background masks, while the 'smeared', worse-defined (fuzzy) lesions differ much in this aspect (Figs. 1 and 2). The latter case gets complicated in terms of segmentation when the edge is too 'smeared.' Such cases may decrease the numerical performance; however, they are still evaluated as 'suspicious' by a human expert. The proper lesion mask should clearly show the shape of the lesion and allow for further inspection of the local structures (dots and globules, pigment mesh, streaks). Separate segmentation of the lesion and the background makes it possible to observe bright areas within the lesion (intensity close to the background). Those structures show regression patterns and/or the blue veil, but artifacts are also possible. Slightly different scores in reference to the image or the hist-pat label appear because there are some malignant melanoma images that look similar to benign lesions and vice versa.

- Data set B is used "as is", i.e., its original resolution is relatively low (about 16 times lower image area than in Data Set A). In theory, this can have various meanings, sometimes it can improve the result of segmentation (fewer pixels/details, a more compact image), another time it may worsen it (less precise preprocessing/filtering, tuning of the segmentation method),

- Data set B ground truth, i.e., the provided segmentation masks (MB), although generally very good, are anyway 'arbitrary.' For quantitative assessment even minor changes can contribute to noticeable errors when we 'intersect' MB and AB,

Table 4. Metrics for border detection by RG-A for the best Gaussian filter regarding the visual examination of the dermatologist. Sensitivity (specificity) means the ability of the method to detect the melanoma (benign) lesion.

data set	sensitivity	specificity	accuracy
A	0.8911	0.9405	0.9135
B	0.8500	0.9091	0.8889

Table 5. Metrics for border detection by RG-A for the best Gaussian filter regarding the hist-pat labels. Sensitivity (specificity) means the ability of the method to detect the melanoma (benign) lesion.

data set	sensitivity	specificity	accuracy
A	0.8962	0.9367	0.9135
B	0.8723	0.8714	0.8718

- Examples in data set B are more demanding than in data set A, both in terms of preprocessing (a lot of hair, different illumination and magnification of the lesion) and segmentation.

As an ablation study, we tested how RG-A and Bin/DBSCAN (Tasks 1 and 2) are affected when the preprocessing steps are completely removed. For well centered, denoised and hair-free dermoscopy images it did not make any measurable difference; however, for most of the images from both data sets A and B the performance of RG-A fell down by more than 30%. Although the same dramatic decrease was observed for DBSCAN (different small clusters), Bin/DBSCAN weakened the lack of the preprocessing step. The decrease in efficiency in the latter case was estimated at 10%.

In what follows, we will compare our numerical results with the literature of the subject, especially in terms of data set B. A general comparison is not possible due to different preprocessing and segmentation methods and different sources and statistics of the input data.

3.2. Comparison with the literature. In the work of Ali *et al.* (2020), a system for classification of border irregularity is shown through a CNN and Gaussian naive Bayes ensemble. This approach achieves high results: Acc $= 93.6\%$, Sens $= 100\%$, and Spec $= 92.5\%$.

Guaragnella and Rizzi (2020) analyze the significant RGB bits and the histogram representation of the dermoscopy images to achieve Acc $= 95.6\%$. Rizzi and Guaragnella (2020) presented SVD-based image denoising plus a binary orthogonalization-like procedure for lesion segmentation. They yield Acc $= 95.37\%$ on the PH^2 database.

Segmentation of melanoma skin lesions using the Perceptual Color Difference Saliency (PCDS) algorithm with morphological analysis plus thresholding applied to

Table 6. Metrics for border detection by RG-A for the best Gaussian filter regarding the hist-pat labels. Sensitivity (specificity) means the ability of the method to detect the melanoma (benign) lesion, respectively.

source	data set	methods	sens	spec	acc
Ali *et al.*, 2020		CNN + Gaussian naïve Bayes ensemble	1.00	0.925	0.936
Guaragnella and Rizzi, 2020		RGB histogram	–	–	0.956
Rizzi and Guaragnella, 2020	PH^2	SVD denoising + binary orthogonalization	–	–	0.954
Olugbara *et al.*, 2018	PH^2	PCDS + thresholding	–	–	0.94–0.98
Pathan *et al.*, 2018a	PH^2	new hair removal + deformable model	0.876	0.953	0.934
Aljanabi *et al.*, 2018	PH^2	bee colony	–	–	0.952–0.976
Patiño *et al.*, 2018	PH^2	superpixel oversegmenting	0.865–0.921	0.687–0.964	0.752–0.952
Khan *et al.*, 2018	PH^2	entropy-based	–	–	0.975

binary segmentation of the saliency map is proposed by Olugbara *et al.* (2018). They report Acc = 94–98% for the PH^2 database.

In the work of Pathan *et al.* (2018a), a new hair detection and chroma-based geometric deformable model is used to effectively differentiate the lesion from the surrounding skin. On average, Acc = 93.4%, Sens = 87.6%, Spec = 95.3% are achieved for the PH^2 dataset.

Aljanabi *et al.* (2018) report a segmentation method based on an artificial bee colony. For the melanoma detection in PH^2, the method achieved an average accuracy and Jaccard's coefficient in the range of 95.24–97.61%.

Patiño *et al.* (2018) present a superpixel-based strategy for (over)segmenting skin lesions on dermoscopic images to be capable of dealing with problems such as hairs, oil bubbles, changes in illumination, etc. without any additional steps. The method was evaluated on data set B (PH^2) yielding Sens = 86.45-92.12%, Spec = 68.70-96.42%, and Acc = 75.19-95.24%.

A novel entropy-based method to derive features for skin lesion segmentation and classification is reported by Khan *et al.* (2018). The proposed method validated on the PH^2 data achieved Acc = 97.5%. These results are collected in Table 6.

A review of computational methods for the image segmentation of pigmented skin lesions is compiled by Oliveira *et al.* (2016) and Ferreira *et al.* (2013). Various forms of region growing and density-based segmentation (clustering) methods are reported below.

The work of Smaoui and Bessassi (2013) is the closest to what we have done regarding the segmentation method, except for only lesion segmentation here, and double segmentation in our case (lesion, background). Its preprocessing step consists of three steps. They use a median 5×5 filter on each RGB component separately and perform morphological closing to eliminate hair. Finally, they convert the image to grayscale and do a histogram adjustment to improve the contrast. The choice for seed points within the lesion is as follows. A list of points with a minimum value of the gray level is produced. For each pixel from this list, an average value is calculated in a 5×5 window of its neighbors. As the seed point, a pixel is selected, which has the minimum difference with its neighbors. The threshold is calculated using a few steps. First, the triangle method is used to find an 'optimal' threshold from the histogram of the image and for a given region with 3×3 pixels, an average value is calculated. The actual threshold is calculated by measuring the difference between the average and the 'optimal' threshold. The region growing process around a seed point can be described as the following steps:

1. collect a window of neighbors;
2. calculate the difference between the gray level of each neighbor and the region average;
3. mark a seed pixel in the region as one with the smallest difference;
4. update the average, threshold and gray level difference;
5. delete the pixel from the list of neighbors;
6. return to Step 1 each time the difference found in Step 4 is less than the threshold.

At the end, a morphological closing (with a diameter of 5 pixels) is produced to remove some isolated pixels.

The result of the segmentation is then used in the next steps for deriving the ABCD features and then calculating the TDS (total dermoscopic score) to classify into melanoma, suspicious, and benign skin lesions. They report the accuracy, sensitivity, and specificity as 92.5% (4 false diagnoses of the 40 samples), 88.88% and 92.30%, respectively.

Indraswari *et al.* (2017) analyze two setups: (i) RG with automatic seeds and a fixed threshold 0.18, and (ii)

RG with automatic seeds and threshold from analysis of the interclass variance.

First, they (Fourier) transform the images to the frequency domain to apply a Gaussian low-pass filter (smoothing) and transform back to the spatial domain. An analysis of the interclass variance of the overall intensity of the melanoma image is implemented to obtain the seed point and threshold parameter values that can provide optimal segmentation results for each image automatically. The interclass variance analysis is based on the histogram of the image (in grayscale). Using each gray level as a threshold, one can calculate the sum of the lesion class variance and the background class variance. To determine the gray level intensity of the seeds, two values must be found: the gray level with the smallest interclass variance and, the lowest gray level found in the image (melanoma is 'darker' than the background skin). Then the gray level intensity of the seed is the mean value between the two. Such pixels are searched for in the image. The threshold value for the region growing is the difference between the intensity of the seed and the level with the smallest interclass variance.

Setup (i) yields average accuracy, sensitivity, and specificity of 96.2%, 93.8%, and 97.4%, respectively. Setup (ii) yields average accuracy, sensitivity, and specificity of 97.6%, 94.8%, and 98.7%, respectively. The results are high because the threshold value used is adaptive and the method provides the seed points automatically.

After segmentation, Indraswari *et al.* (2017) also classify the lesions by an SVM according to some features (area ratio, circumference ratio, color) but this task is out of our scope of our article.

Agarwal *et al.* (2017) show a region growing approach, where for each individual image two arbitrary thresholds $t_1 = 0.1$, $t_2 = 0.2$ and two characteristics, 'solidity' and 'extent' of the lesion are used to segment the lesion border. For solidity, they use a fraction of the object area to the object convex area. The extent, on the other hand, is defined as the ratio of the object area to the object bounding box area. Finally, a threshold wins that maximizes both the solidity and extent of the lesion. To prepare the images, they employ a median filter to remove the hair edges. Two metrics are used to quantify their procedure against the ground truth: an overlapping score of 91.13% and an average correlation coefficient of 93.79%.

Kockara *et al.* (2010) compare two clustering algorithms DBSCAN and FCM (fuzzy c-means). Each approach is examined on a set of 100 dermoscopy images (probably the same as in the work of Suer *et al.* (2011)). Both methods are quantitatively analyzed over three accuracy measures: border error, precision, and recall. For DBSCAN they show on average 7%, 100%, and 76.66%, respectively, while for FCM 100% (?), 99.26% and 55%, respectively. Also visually, the DBSCAN algorithm more effectively delineates the targeted lesions, and FCM has poor performance, especially regarding the border error metric.

In the work of Suer *et al.* (2011) fast density based lesion detection (FDBLD) with a new elaborated (normalized) distance measure is shown, which is an extension to the results of Mete *et al.* (2011). The new distance measure not only considers pixel positions, but also their colors. On a data set with one hundred dermoscopy images, they show BE = 0.01–0.17 (few images show a much higher BE), precision = 0.89–1.00, and sensitivity = 0.70–0.94 (few images show a much lower sensitivity).

Louhichi *et al.* (2018) set forth algorithm MDCUT. It uses most of the concepts utilized in DBCSAN (the core, border and noise points), but makes the process of clustering more automatic by visual determination of local levels of density. MDCUT starts by mapping the data into a kd-tree data structure and identifying the (Euclidean) distance between each point and its k-th nearest neighbor (k-dist). Such queries are stored in a matrix of sorted distances (the first column contains values of the nearest neighbors, the second is for the second nearest neighbors etc.). When we plot all pairs: (order-of-the-point [=column-of-matrix], k-dist) we can see a scatterplot, which is then interpolated by an exponential spline to form a continous function F (interpolation uses a tension parameter t). This function is monotone by intervals and shows the distribution of densities among the data. Points belonging to the same cluster have a very low variation in density (more or less "horizontal" intervals) and different clusters are separated by more sparse, nearly "vertical" (discontinous) regions in the k-NN plot (the so-called change/inflection points). The change points can be calculated from the curve as those where the sign of the curvature changes (a necessary condition, $F''(x) = 0$). Determination of the change points in the curve means detection of the core (=seed) pixels for the expansion of regions. Each such point has a value (localized on the curve F) which is a local density threshold. Now the region expansion occurs (called SBRG, seed-based region growing).

The algorithm grows different clusters (each with a range of density) starting from the seed points, directly localized on the curve F. A data point can get to a cluster (when below a certain k-distance), start a new one, or remain as a noise/not classified point.

Louhichi *et al.* (2018) report the Dice similarity measure (= $F1$-score) to be 71%, the total sensitivity of 81% and specificity of 74%.

The above results are collected in Table 7.

An interesting adaptive thresholding (AT) system for automatic diagnosis of pigmented lesions is presented by De Vita *et al.* (2012). Its first stage is devoted to lesion

Table 7. Skin lesion border detection—RG and/or density based algorithms.

src	preprocessing	methods	metric 1	metric 2	metric 3
Smaoui and Bessassi, 2013	median filter, morph. closing, hist. adjust.	RG	acc=0.925	spec=0.923	sens=0.889
Indraswari *et al.*, 2017	Gaussian filter (based on Fourier)	RG: thresh=0.18	acc=0.962	spec=0.974	sens=0.938
		automatic RG	acc=0.976	spec=0.987	sens=0.948
Agarwal *et al.*, 2017	median filter	solidity, extent; threshold selection	corr. coeff. =0.940	overlap. score =0.911	–
Kockara *et al.*, 2010	unknown	DBSCAN	border err. =0.07	prec=1.00	sens=0.767
		FCM	border err. =1.00	prec=0.993	sens=0.550
Suer *et al.*, 2011	unknown	FDBLD	border err. =0.01–0.17	prec= 0.89–1.00	sens= 0.70–0.94
Louhichi *et al.*, 2018	RGB, HSV, XYZ; median filter	MDCUT	Dice=0.71	prec=0.74	sens=0.81

border detection, then feature extraction and, finally, feature classification according to the 7-point check list.

Extraction of the skin lesion border consists of three steps: (i) conversion to the separate RGB images, (ii) binarization using an adaptive threshold, and (iii) border identification based on a blob-finding algorithm. In stage (i) histograms of the R, G, and B color components are analyzed to best derive the two modes: the pigmented lesion (image foreground) and the surrounding skin (the image background). The optimum threshold for each histogram is selected by the Otsu algorithm. The image foreground threshold defines a binary mask for the next step. Finally, a modified Moore's Neighbour Contour Tracing algorithm is adopted to extract the contour of the lesion from the binary mask. As an alternative method, statistical region merging (SRM) is performed.

This border from AT or SRM is superimposed on the colour dermoscopy image for visible inspection to the human experts. They independently indicate the border points which are then connected by a second-order B-spline. A majority vote is used to select the final contour of the lesion. A comparison between the manual and automatic border is quantified by the border error (Pennisi *et al.*, 2016; Suer *et al.*, 2011) introduced in Section 2.4. They yield $BorderError(AT) = 8.7 \pm 4.8\%$ and $BorderError(SRM) = 10.8 \pm 6.9\%$. The automatic border resulting from AT is better than the result from SRM. Their results are comparable to ours for the best filter and for data set A, $BorderError(\textbf{\textit{RG-A}}) = 8.0\%$, $BorderError(BIN/DBSCAN) = 9.0\%$, and surpass ours for data set B, $BorderError(\textbf{\textit{RG-A}}) = 15.0\%$, $BorderError(BIN/DBSCAN) = 12.0\%$.

4. Conclusion

In this paper, we demonstrated the applicability of region growing and the density-based clustering algorithm in the detection of skin lesions and their borders. Our numerical results are comparable to other methods in the field (taking into account possible accuracies of the component methods), but we noticed that the cited works usually use worse statistics of images.

In fact, due to the different selection of the images (data set B = PH^2) and different methodologies among the presented research groups, the results cannot be directly compared.

Although region growing is a well-known and widely used technique, our method overcomes its limitations by the automatic search for the seed pixels and threshold values.

Our procedure is fairly robust to regular and irregular/fuzzy borders, and after initial adjustments (different requirements for image preprocessing), it may be applied to any (standard) data set.

The proposed density-based method DBSCAN shows less efficiency for the lesion segmentation and has to be supported by a secondary binarization. The latter method is capable of correctly segmenting the pigmented lesion in the dermoscopy images.

Border extraction is only one of several factors included in the metrics of melanoma (an essential one, although). The proposed region growing segmentation method is simple in terms of morphological operations deployed, but has performance comparable to other results indicated in the literature. It is also sensitive to dermoscopy images with fuzzy borders, where malignancy is more probable.

Data availability

Data set A (Melanoma ML, 2018) used to support the findings of this study may be released upon application to https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:114463, which can be contacted through DOI 10.17026/dans-zue-zz2y.

Data set B (Mendonça *et al.*, 2013) used to support the findings of this study is available via DOI 10.1109/EMBC.2013.6610779.

References

ACS (2020). Key statistics for melanoma skin cancer, American Cancer Society, Atlanta, https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html.

Agarwal, A., Issac, A. and Dutta, M. (2017). A region growing based imaging method for lesion segmentation from dermoscopic images, *4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics, Mathura, India*, pp. 632–637.

Ahn, E., Kim, J., Bi, L., Kumar, A., Li, C., Fulham, M. and Feng, D. (2017). Saliency-based lesion segmentation via background detection in dermoscopic images, *Journal of Biomedical and Health Informatics* **21**(6): 1685–1693.

Ali, A.-R., Li, J., Yang, G. and O'Shea, S. (2020). A machine learning approach to automatic detection of irregularity in skin lesion border using dermoscopic images, *PeerJ Computer Science* **6**: e268.

Aljanabi, M., Özok, Y., Rahebi, J. and Abdullah, A. (2018). Skin lesion segmentation method for dermoscopy images using artificial bee colony algorithm, *Symmetry* **10**(8): 347.

Ankerst, M., Breunig, M., Kriegel, H. and Sander, J. (1999). Optics: Ordering points to identify the clustering structure, *ACM Sigmod Record* **28**(2): 49–60.

Ashour, A., Hawas, A., Guo, Y. and Wahba, M. (2018). A novel optimized neutrosophic k-means using genetic algorithm for skin lesion detection in dermoscopy images, *Signal, Image and Video Processing* **12**(7): 1311–1318.

Bentley, J. (1975). Multidimensional binary search trees used for associative searching, *Communications of the ACM* **18**(9): 509–517.

Bi, L., Kim, J., Ahn, E., Feng, D. and Fulham, M. (2016). Automated skin lesion segmentation via image-wise supervised learning and multi-scale superpixel based cellular automata, *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic*, pp. 1059–1062.

Bi, L., Kim, J., Ahn, E., Kumar, A., Fulham, M. and Feng, D. (2017). Dermoscopic image segmentation via multistage fully convolutional networks, *IEEE Transactions on Biomedical Engineering* **64**(9): 2065–2074.

Bozorgtabar, B., Abedini, M. and Garnavi, R. (2016). Sparse coding based skin lesion segmentation using dynamic rule-based refinement, *International Workshop on Machine Learning in Medical Imaging, Athens, Greece*, pp. 254–261.

Campello, R., Moulavi, D. and Sander, J. (2013). Density-based clustering based on hierarchical density estimates, *Pacific-Asia Conference on Knowledge Discovery and Data Mining, Gold Coast, Australia*, pp. 160–172.

Celebi, E., Codella, N. and Halpern, A. (2019). Dermoscopy image analysis: Overview and future directions, *Journal of Biomedical and Health Informatics* **23**(2): 474–478.

Celebi, E., Wen, Q., Hwang, S., Iyatomi, H. and Schaefer, G. (2013). Lesion border detection in dermoscopy images using ensembles of thresholding methods, *Skin Research and Technology* **19**(1): e252–e258.

Celebi, M., Kingravi, H. and Iyatomi, H. (2008). Border detection in dermoscopy images using statistical region merging, *Skin Research and Technology* **14**(3): 347–353.

Celebi, M., Kingravi, H. and Uddin, B. (2007). A methodological approach to the classification of dermoscopy images, *Computerized Medical Imaging and Graphics* **31**(6): 362–373.

Celebi, M., Wen, Q., Iyatomi, H., Shimizu, K., Zhou, H. and Schaefer, G. (2015). A state-of-the-art survey on lesion border detection in dermoscopy images, *Dermoscopy Image Analysis* **10**: 97–129.

Codella, N., Rotemberg, V., Tschandl, P., Celebi, M., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., Kittler, H. and Halpern, A. (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the International Skin Imaging Collaboration (ISIC), *arXiv* 1902.03368.

De Vita, V., Di Leo, G., Fabbrocini, G., Liguori, C., Paolillo, A. and Sommella, P. (2012). Statistical techniques applied to the automatic diagnosis of dermoscopic images, *Acta Imeko* **1**(1): 7–18.

Ester, M., Kriegel, H., Sander, J. and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise, *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, USA*, pp. 226–231.

Esteva, A., Kuprel, B., Novoa, R., Ko, J., Swetter, S., Blau, H. and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks, *Nature* **542**(7639): 115–118.

Ferreira, P., Mendonça, T. and Rocha, P. (2013). A wide spread of algorithms for automatic segmentation of dermoscopic images, *Iberian Conference on Pattern Recognition and Image Analysis, Madeira, Portugal*, pp. 592–599.

Goyal, M., Oakley, A., Bansal, P., Dancey, D. and Yap, M. (2019). Skin lesion segmentation in dermoscopic images with ensemble deep learning methods, *IEEE Access* **8**: 4171–4181.

Guaragnella, C. and Rizzi, M. (2020). Simple and accurate border detection algorithm for melanoma computer aided diagnosis, *Diagnostics* **10**(6): 423.

Hahsler, M. (2021). Density based clustering of applications with noise (DBSCAN) and related algorithms—R package, https://github.com/mhahsler/dbscan.

Hahsler, M., Piekenbrock, M. and Doran, D. (2019). DBSCAN: Fast density-based clustering with R, *Journal of Statistical Software* **91**: 1–30.

Hinneburg, A. and Keim, D. A. (1998). An efficient approach to clustering in large multimedia databases with noise, *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining, New York, USA*, pp. 58–65.

Hornik, K. (2021). *The Comprehensive R Archive Network*, https://cran.r-project.org.

Indraswari, R., Herulambang, W. and Rokhana, R. (2017). Melanoma classification using automatic region growing for image segmentation, *Proceeding of the International Conference on Technology and Applications, Surabaya, Indonesia*, pp. 165–172.

Jaworek-Korjakowska, J. and Tadeusiewicz, R. (2013). Hair removal from dermoscopic color images, *Bio-Algorithms and Med-Systems* **9**(2): 53–58.

Jensen, D. and Elewski, B. (2015). The ABCDEF rule: Combining the 'ABCDE RULE' and the 'ugly duckling sign' in an effort to improve patient self-screening examinations, *Journal of Clinical and Aesthetic Dermatology* **8**(2): 15.

Keefe, M., Dick, D. and Wakeel, R. (1990). A study of the value of the seven-point checklist in distinguishing benign pigmented lesions from melanoma, *Clinical and Experimental Dermatology* **15**(3): 167–171.

Khan, M., Akram, T., Sharif, M., Shahzad, A., Aurangzeb, K., Alhussein, M., Haider, S. and Altamrah, A. (2018). An implementation of normal distribution based segmentation and entropy controlled features selection for skin lesion detection and classification, *BMC Cancer* **18**(1): 1–20.

Kockara, S., Mete, M., Chen, B. and Aydin, K. (2010). Analysis of density based and fuzzy c-means clustering methods on lesion border extraction in dermoscopy images, *BMC Bioinformatics* **11**(6): 1–11.

Kroon, D. (2004). Region growing, https://www.mathworks.com/matlabcentral/fileexchange/19084-region-growing.

Lee, T., Ng, V., Gallagher, R., Coldman, A. and McLean, D. (1997). DullRazor: A software approach to hair removal from images, *Computers in Biology and Medicine* **27**(6): 533–543.

Louhichi, S., Gzara, M. and Abdallah, H. (2018). Skin lesion segmentation using multiple density clustering algorithm mdcut and region growing, *IEEE/ACIS 17th International Conference on Computer and Information Science, Singapore, Singapore*, pp. 74–79.

Louhichi, S., Gzara, M. and Ben-Abdallah, H. (2017). Unsupervised varied density based clustering algorithm using spline, *Pattern Recognition Letters* **93**: 48–57.

Masood, A. and Al-Jumaily, A. (2013). Computer aided diagnostic support system for skin cancer: A review of techniques and algorithms, *International Journal of Biomedical Imaging* **7**: 323268.

Melanoma ML (2018). Data set, https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:114463, DOI: 10.17026/dans-zue-zz2y.

Mendonça, T., Ferreira, P., Marques, J., Marcal, A. and Rozeira, J. (2013). Ph2—A dermoscopic image database for research and benchmarking, *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan*, pp. 5437–5440.

Mete, M., Kockara, S. and Aydin, K. (2011). Fast density-based lesion detection in dermoscopy images, *Computerized Medical Imaging and Graphics* **35**(2): 128–136.

Mete, M. and Sirakov, N. (2010). Lesion detection in dermoscopy images with novel density-based and active contour approaches, *BMC Bioinformatics* **11**(6): 1–13.

Mishra, N. and Celebi, M. (2016). An overview of melanoma detection in dermoscopy images using image processing and machine learning, *arXiv* 1601.07843.

Møllersen, K., Kirchesch, H.M., Schopf, T.G. and Godtliebsen, F. (2010). Unsupervised segmentation for digital dermoscopic images, *Skin Research and Technology* **16**(4): 401–407.

Mount, D. and Arya, S. (2010). ANN: A library for approximate nearest neighbor searching, https://github.com/dials/annlib.

Oliveira, R., Mercedes, F., Ma, Z., Papa, J., Pereira, A. and Tavares, J. (2016). Computational methods for the image segmentation of pigmented skin lesions: A review, *Computer Methods and Programs in Biomedicine* **131**: 127–141.

Oliveira, R., Papa, J. and Pereira, A. (2018). Computational methods for pigmented skin lesion classification in images: Review and future trends, *Neural Computing and Applications* **29**(3): 613–636.

Olugbara, O., Taiwo, T. and Heukelman, D. (2018). Segmentation of melanoma skin lesion using perceptual color difference saliency with morphological analysis, *Mathematical Problems in Engineering* **2018**, Article ID: 1524286.

Pathan, S., Prabhu, K. and Siddalingaswamy, P. (2018a). Hair detection and lesion segmentation in dermoscopic images using domain knowledge, *Medical & Biological Engineering & Computing* **56**(11): 2051–2065.

Pathan, S., Prabhu, K. and Siddalingaswamy, P. (2018b). Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review, *Biomedical Signal Processing and Control* **39**: 237–262.

Patiño, D., Avendaño, J. and Branch, J. (2018). Automatic skin lesion segmentation on dermoscopic images by the means of superpixel merging, *International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain,* pp. 728–736.

Pennisi, A., Bloisi, D., Nardi, D., Giampetruzzi, A., Mondino, C. and Facchiano, A. (2016). Skin lesion image segmentation using Delaunay triangulation for melanoma detection, *Computerized Medical Imaging and Graphics* **52**: 89–103.

Pradhan, R., Kumar, S., Agarwal, R., Pradhan, M.P. and Ghose, M. (2010). Contour line tracing algorithm for digital topographic maps, *International Journal of Image Processing* **4**(2): 156–163.

Rizzi, M. and Guaragnella, C. (2020). Skin lesion segmentation using image bit-plane multilayer approach, *Applied Sciences* **10**(9): 3045.

Sadeghi, M., Razmara, M., Lee, T. and Atkins, M. (2011). A novel method for detection of pigment network in dermoscopic images using graphs, *Computerized Medical Imaging and Graphics* **35**(2): 137–143.

Smaoui, N. and Bessassi, S. (2013). Melanoma skin cancer detection based on region growing segmentation, *International Journal of Computer Vision and Signal Processing* **1**(1): 1–7.

Stanley, R., Stoecker, W. and Moss, R. (2007). A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images, *Skin Research and Technology* **13**(1): 62–72.

Suer, S., Kockara, S. and Mete, M. (2011). An improved border detection in dermoscopy images for density based clustering, *BMC Bioinformatics* **12**(10): 1–10.

Vestergaard, M., Macaskill, P., Holt, P. and Menzies, S. (2008). Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: A meta-analysis of studies performed in a clinical setting, *British Journal of Dermatology* **159**(3): 669–676.

Wang, H., Moss, R., Chen, X. and Stanley, R. (2011). Modified watershed technique and post-processing for segmentation of skin lesions in dermoscopy images, *Computerized Medical Imaging and Graphics* **35**(2): 116–120.

Zhou, H., Schaefer, G., Celebi, M., Lin, F. and Liu, T. (2011). Gradient vector flow with mean shift for skin lesion segmentation, *Computerized Medical Imaging and Graphics* **35**(2): 121–127.

Grzegorz Surówka is an assistant professor in the Department of Physics, Astronomy and Applied Computer Science at Jagiellonian University. His research interests include machine learning, deep learning and pattern recognition. He has published several papers on malanoma computer aided diagnosis.

Maciej J. Ogorzałek is a professor of electrical engineering and computer science and the head of the Department of Information Technologies, Jagiellonian University, Kraków, Poland. For one year (2000) he worked at the Microelectronic Center, Sevilla, Spain. From 2006 to 2009 he held the Chair of Biosignals and Systems, Hong Kong Polytechnic University, under the Distinguished Scholars Scheme. He is an author or a co-author of over 350 technical papers published in journals and conference proceedings, and the author of the book *Chaos and Complexity in Nonlinear Electronic Circuits* (World Scientific, 1997). He has given over 60 plenary and keynote talks at major conferences world-wide. He has served the IEEE Circuits and Systems Society in various capacities, including the 2008 Society President. He was a CAS Society Distinguished Lecturer (2004–2005) and received the 2002 Guillemin–Cauer Award, the IEEE CAS Golden Jubilee Award and the IEEE CAS Meritorious Service Award. Currently he serves on the IEEE Board of Directors as IEEE Director Division 1. He is an elected member of the Academia Europaea and the Polish of Academy of Sciences, and an IEEE fellow.

Received: 13 December 2021
Revised: 6 April 2022
Accepted: 20 May 2022