

## Article

# Malignant and Benign Mass Segmentation in Mammograms Using Active Contour Methods

Marcin Ciecholewski <sup>†</sup>

Institute of Informatics, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, 80-308 Gdańsk, Poland; marcin.ciecholewski@ug.edu.pl

<sup>†</sup> Current address: ul. Wita Stwosza 57, 80-308 Gdańsk, Poland.

Received: 1 October 2017; Accepted: 12 November 2017; Published: 16 November 2017

**Abstract:** The correct segmentation of tumours can simplify formulate the diagnostic hypothesis, particularly in cases of irregular shapes, with fuzzy margins or spicules growing into the surrounding tissue, which are more likely to be malignant. In this study, the following active contour methods were used to segment the masses: an edge-based active contour model using an inflation/deflation force with a damping coefficient (EM), a geometric active contour model (GAC) and an active contour without edges (ACWE). The preprocessing techniques presented in this publication are to reduce noise and at the same time amplify uniform areas of images in order to improve segmentation results. In addition, the use of image sampling by bicubic interpolation was tested to shorten the evolution time of active contour methods. The experiments used a test set composed of 100 cases taken from two publicly available databases: Digital Database for Screening Mammography (DDSM) and Mammographic Image Analysis Society (MIAS) database. The qualitative assessment concerned the ability to formulate an adequate diagnostic hypothesis and, for the individual methods (malignant and benign cases together), it amounted to at least: 81% (EM), 76% (GAC), and 69% (ACWE). The quantitative test consisted of measuring the following indexes: overlap value (OV) and extra fraction (EF). The OV of the segmentation for malignant and benign cases had the following average values:  $0.81 \pm 0.10$  (EM),  $0.79 \pm 0.09$  (GAC),  $0.76 \pm 0.18$  (ACWE). The average values of the EF index, in turn, amounted to:  $0.07 \pm 0.06$  (EM),  $0.07 \pm 0.05$  (GAC)  $0.34 \pm 0.32$  (ACWE). The qualitative and quantitative results obtained are the best for EM and are comparable or better than for other methods presented in the literature.

**Keywords:** active contour; edge-based active contour; region-based active contour; image processing; segmentation; masses; breast cancer; mammography

## 1. Introduction

Mammography systems for the computer-aided detection (CAD) of cancer masses perform the following steps: preprocessing [1–3], segmentation [4–7], feature extraction [8–10] and classification [11–13]. However, whether a CAD system will be successfully adopted in clinical practice depends mainly on the segmentation algorithm or algorithms used. In medical image analysis, segmentation is defined as a method allowing the precise margins of the potential lesions to be determined or the shape of the organ to be determined.

A tumor is a pathology occupying a certain area, ranging from medium-grey to white shades in the mammogram. The smallest tumors visible in mammograms are approx. 0.5 cm in diameter. The most significant features indicating whether the tumor is malignant or benign are its shape and the nature of its margins [8,14]. The shape can be: round, oval, lobulated, and irregular. The margins, in turn, can be described as: circumscribed, microlobulated, obscured, ill-defined, and spiculated. Benign tumours are usually: regular (round or oval) in shape, with smooth margins and uniform

texture. They are circumscribed well against the surrounding tissue. Tumors of irregular shapes and with fuzzy margins or spicules growing into the surrounding tissue are more likely to be malignant.

Digital mammograms frequently contain strong noise while cancerous tumours are of varied shapes and appearances. Furthermore, the contrast of suspicious-looking regions of mammograms is frequently low and heterogeneous, and the margins between masses are fuzzy and difficult to identify. All of this means that the segmentation of the lesions is an important and frequently very difficult task.

In general, algorithms used for the detection and segmentation of masses as well as their further possible classification can be divided into two approaches: supervised segmentation and unsupervised segmentation. Supervised segmentation mainly includes model-based methods [15–17]. Model-based methods use previously acquired (e.g., defined or learned) knowledge of objects and background regions that are being segmented. Previous knowledge is used to determine whether specific regions occur in the image or not. Supervised segmentation methods also include template matching approaches [15,18], in which the training set contains templates or patterns of objects that can be detected. Unfortunately, the main limitation of model—or template-based methods—is their reduced effectiveness in case of irregular masses with spiculated margins that are difficult to distinguish.

Unsupervised segmentation methods work by dividing the image into areas that are different or uniform with regard to defined features, such as grey levels, their texture or colour. Three main groups of unsupervised segmentation methods can be listed:

- Clusterization methods that group pixels with the same properties and make a division into non-connected regions possible. An example of such a method is presented in the publication of Suliga et al. [19], in which the Markov random field is used to clusterize pixels belonging to lesions. The mass detection algorithm developed in [20] works by using the area of fraction under minimum (AFUM) filter. The AFUM filter determines the degree to which the surrounding region of a point radially decreases in intensity. The last step in the algorithm is to threshold the image to identify suspicious-looking areas.
- Region-based methods allow the image to be divided into homogeneous and connected regions, e.g., according to their texture properties. In the publication by Wei et al. [21], potential lesions are first extracted by their clusterization based on the region-growing method. The morphological and spatial relationships of grey levels, i.e., the characteristics of texture features, are extracted for every suspicious-looking object. Then, rule-based and linear discriminant analysis classifiers are used to differentiate normal tissues from masses. Unfortunately, this method has a certain limitation, namely the wrong classification of the background with fragments of lesions, causing oversegmentation. This applies to images of a low contrast. In Ref. [22], the authors proposed an automated method (requiring no interaction by the user) with the use of the marker-controlled watershed segmentation [22] and presented good results for malignant and benign cases, as well as for different types of tissue densities and margins of the analysed tumours.
- Contour methods that find the boundaries of approximated areas. Active contour methods (ACMs) are among the most popular because they are able to integrate image pixels into smooth, connected borders that delineate the shape being approximated. ACMs can be further subdivided, e.g., into edge-based [23–25], or region-based ones [26–29]. The way in which the energy function is defined in the ACM determines the effectiveness and the range of applications of the active contour. The advantage of region-based methods, in turn, consists of their topological adaptability to complex shapes. Several active contour approaches using level-sets have been used to segment masses in mammograms [6,30,31]. Rahmati et al. [6] presented the maximum likelihood active contour model using level sets (MLACMLS) and demonstrated that the proposed method produced better results than earlier approaches presented in [30] and also in [31]. In [6], experiments were conducted on 100 regions of interest (ROI) with the resolution of  $256 \times 256$  pixels taken from mammograms from the DDSM database, and the mean value of

Jaccard's index [32] (Overlap Value—OV) achieved was 86.85%. The time needed to segment masses for ROIs with the resolution of  $256 \times 256$  ranged from 10 to 60 s [6].

A more detailed review of publications dealing with the classification of various methods of segmenting lesions in mammograms can be found, among others, in [5,33,34]. Articles [35–38], in turn, present approaches supporting interactive and automatic image segmentation for grey level [36] and colour [35,37,38] images.

This article is structured as follows. Section 2 presents the Material and Methods. Section 3 contains the Results and Discussion. Section 4 is the Conclusion.

## 2. Material and Methods

In this research, just as in [6,30,31], ACMs were used to segment masses in mammograms. This publication describes the application of the recently developed edge-based active contour model using an inflation/deflation force with a damping coefficient [25], which is abbreviated to EM. This model has been proven to be useful in the segmentation of various types of digital images, including medical ones: produced by ultrasonography (USG) and magnetic resonance (MRI). The EM was compared to two popular active contour methods, i.e.,

- A geometric active contour (GAC) using a morphological approach [39]. It enables the active contour to evolve very effectively because there are no floating point operations and differential operators have been replaced with mathematical morphological methods on a binary-level set [39].
- Active Contour Without Edges (ACWE). This is a region-based method from publication [26], very frequently used to segment images, including medical ones. This model enables a global segmentation—for the entire image—and also a local segmentation (e.g., of a selected object in the image after the contour is initiated inside the approximated object).

This study also contains an assessment of the ability to segment masses using two state-of-the-art region-based active contours that support global segmentation, i.e., the region-scalable fitting model (RSF) [28] and the B-spline level-set model [29].

In this study, 90 mammograms from the DDSM database [40,41] and 30 from the MIAS database [42] were used. They were selected for the experiments by two experienced breast radiologists:  $R_1$ ,  $R_2$ .  $R_1$  has 30 years of professional experience and  $R_2$  10 years.

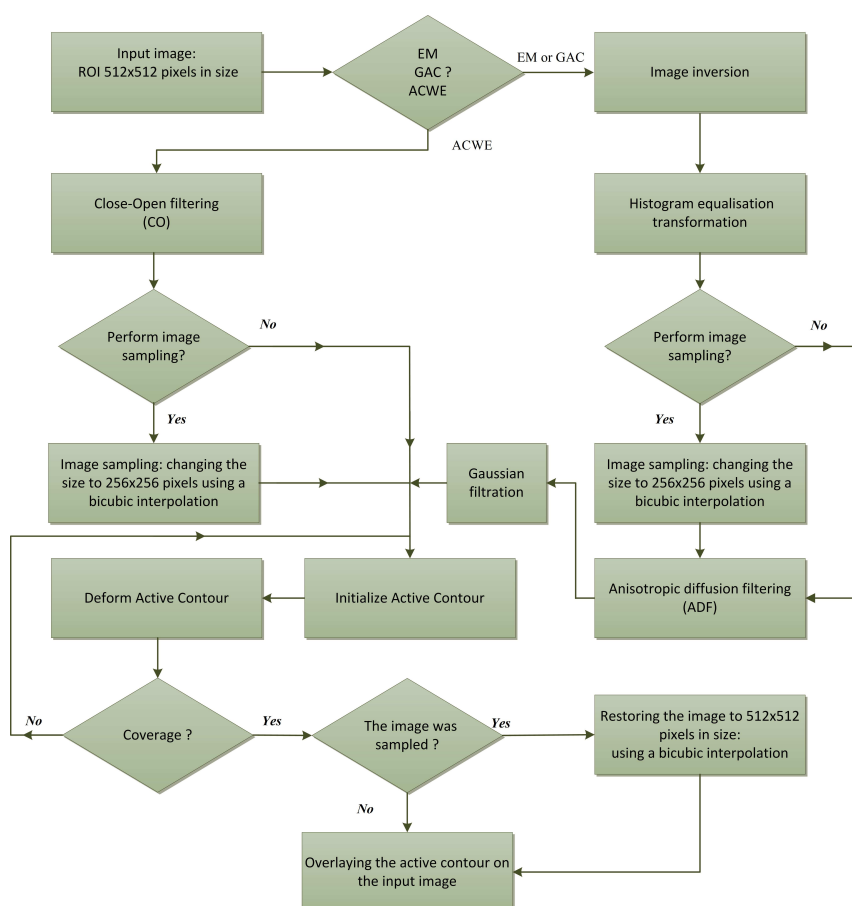
An ROI with the constant dimensions of  $512 \times 512$  pixels and 256 grey levels was sampled from each selected mammogram. Among the 120 analysed ROIs, there were 60 benign cases and 60 cancerous ones, with every ROI corresponding to a different patient. Each case in the DDSM database has four images available in the CC and MLO projections for the left and right breast. CC is the cranio-caudal projection showing the central and medial part of the mamma while MLO stands for the medio-lateral oblique projection. In the experiments, a single view was taken, namely the CC or MLO view, for each patient.

Detailed information about the analysed datasets with regards to the shape of the disorder, different margins and types of tissues is presented in Table 1. It is worth noting that, in the case of the DDSM database, there are four tissue categories given, which is consistent with the BI-RADS [43] classification. The description of cases from the MIAS mammography database contains only three categories of tissues.

It should be noted that twenty images—showing masses of various shapes, dimensions, margins and brightnesses—were used as a training set to establish the necessary parameters of the methods applied. The remaining 100 images were used to test the proposed segmentation method. A convenient functionality of the CAD program can stem from using methods that automatically mark ROIs containing suspicious-looking masses in mammograms. Example solutions are known from literature [44–46]. On the other hand, these methods may lead to identifying false positive areas, i.e., implied ROIs that actually contain no masses, so the judgement of an experienced breast radiologist is

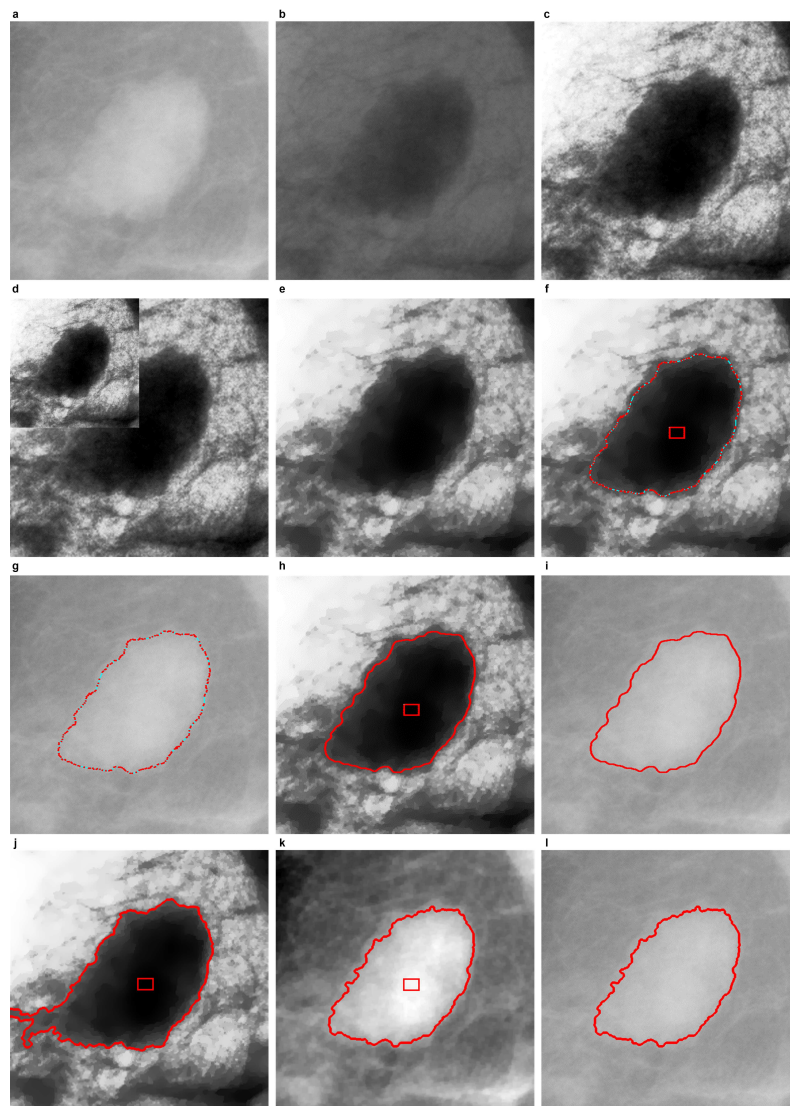
indispensable here. This publication focuses on the subject of segmentation and the methods used in it accept rectangular ROIs marked by a radiologist on an input mammogram in such a way that the suspicious-looking anomalies are located in their centre.

The proposed method of segmenting lesions is illustrated in Figure 1. The preprocessing methods used—namely input image inversion, histogram equalization [1], anisotropic diffusion filtering (ADF) [47,48], Gaussian filtration—are to reduce noise, amplify uniform areas of the processed image and thus improve the operation of the active contour models, namely: EM and GAC. In the case of ACWE, it is sufficient to use only the Close-Open filtering (CO), which helps reduce noise. Image sampling by bicubic interpolation [49], in turn, reduces the image resolution by half and thus makes it possible to shorten the evolution time of the active contour for a pre-set number of iterations in all three of the models used: ACWE, EM and GAC. If the image has been sampled, then its size should be restored and this is done using bicubic interpolation too, which is also applied to the points of the active contour from the last iteration. In the last step, the active contour nodes are overlaid on the input image.



**Figure 1.** A block diagram of the proposed method of segmenting suspicious-looking masses from mammograms using the following active contour models: an active contour without edges ACWE, an edge-based active contour model using an inflation/deflation force with a damping coefficient EM and a geometric active contour model GAC.

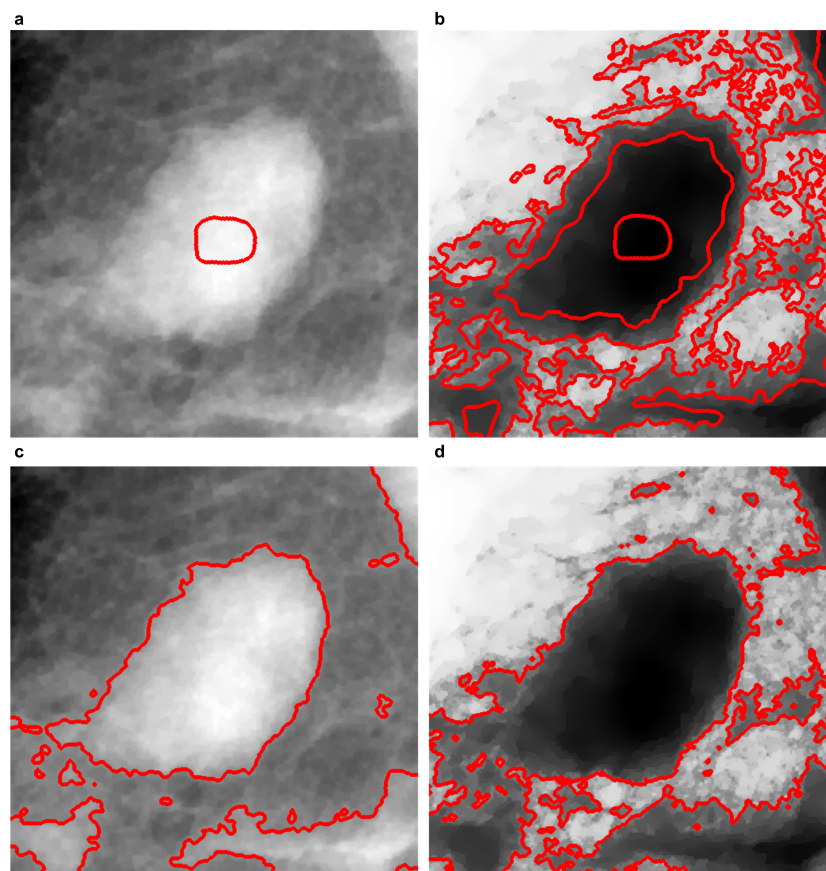




**Figure 2.** An example illustrating subsequent steps of the methods for approximating masses in the mammogram, in the order of the diagram from Figure 1. (a) ROI  $512 \times 512$  pixels in size, based on the image ‘mdb015’ from the MIAS database; (b) inverted image from letter (a); (c) executing the histogram equalisation transformation; (d) sampling the image from letter (c) and reducing its size to  $265 \times 256$  pixels. The sampled image sized  $256 \times 256$  pixels is overlaid on the processed image from letter (c); (e) the sampled image from letter (d),  $256 \times 256$  pixels in size, after anisotropic diffusion filtering [47,48]. To emphasise details better, the example is of the same height and width as the image before sampling, i.e., from letter (c); (f) EM model use: the rectangular seed contour and the final contour approximating the boundaries of the lesion; (g) EM model use: the final contour overlaid on the input image from letter (a) after the processed image has been restored to the size of  $512 \times 512$  pixels; (h) GAC model use: the rectangular seed contour and the final contour approximating the boundaries of the lesion; (i) GAC model use: The final contour overlaid on the input image from letter (a) after the processed image has been restored to the size of  $512 \times 512$  pixels; (j) the use of the ACWE model (the same preprocessing as in the EM and GAC models): a rectangular seed contour and the final contour; (k) ACWE: a rectangular seed contour and the final contour approximating the edges of the lesion in a source image after the CO filtration; (l) ACWE: the final contour overlaid on the input image from item (a).

Figure 2 contains examples illustrating the methods based on the diagram from Figure 1. If the ACWE model uses the same methods of preprocessing as the EM and GAC methods, then the active

contour will frequently produce an extra fraction, located outside the approximated area of the mass. This situation is illustrated by Figure 2j. Only the use of CO filtering makes better results possible: Figure 2k,l. Literature describes various active contour models that can be used to segment medical images and the authors of papers present very good results. For instance, the region-scalable fitting model (RSF) [28] and the B-spline level-set model [29] are region-based active contours recognised in the literature. The RSF model can produce very good segmentation results for, among others, MRI images of the brain and blood vessels. The B-spline level-set model, in turn, is very resistant to stronger noise and also allows a calcaneus bone in 3D micro-CT images to be segmented with a high accuracy. Unfortunately, the results of segmenting masses in mammograms using both models [28,29] that execute the global segmentation of the entire analysed image are unsatisfactory. Example results are illustrated in Figure 3.



**Figure 3.** The application of the following region-based active contours: the region-scalable fitting model (RSF) [28] and the B-spline level-set [29] to the example mammogram ‘mdb015’ from the MIAS database. The seed contour was put in the same place as in examples from Figure 2f,h. The first column compares results based on the source image in which the CO filtration with a disk-shaped structuring element with the radius of 3 had been applied, and the second column shows segmentation results for an image subjected to the anisotropic diffusion filtration (ADF) as in the example from Figure 2e. (a,b) RSF; (c,d) B-spline level-set.

**Table 1.** Disorders in the analysed set of 120 ROIs, depending on the shape of the mass, its margins, and tissue density types, based on the following databases: DDSM (90 ROIs) and MIAS (30 ROIs).

No. of Lesions of Various	No. of Lesions of Various	No. of Lesions of Various	No. of Lesions of Various
Mass Shape Types	Margin Types	Densities (DDSM)	Densities (MIAS)
Round 20	Circumscribed 40	Almost entirely fatty 23	Fatty 18
Oval 20	Spiculated 50	Scattered fibroglandular 43	Fatty—glandular 8
Lobulated 20	Ill-defined 10	Heterogeneously dense 16	Dense—glandular 4
Irregular 60	Microlobulated 10	Extremely dense 8	
	Obscured 10		

### 2.1. Anisotropic Diffusion Filtering (ADF)

Anisotropic diffusion reduces the noise in the image while keeping region boundaries, which is very important in the segmentation carried out later. According to [48], the anisotropic diffusion ( $I_t$ ) of the image ( $I$ ) is defined by the following equation:

$$I_t = \text{div}[c(x, y, t) \nabla I] = c(x, y, t) \Delta I + \nabla c \cdot \nabla I, \quad (1)$$

where  $\text{div}$  is the divergent operator,  $\Delta$  and  $\nabla$  are, respectively, the gradient and Laplace operators, and  $c(x, y, t)$  is the diffusion coefficient with the iteration step  $t$ . The diffusion coefficient  $c(x, y, t)$  influences the rate of diffusion, which means that the boundaries in the image are selectively equalised or amplified, depending on the function selected. It may be written as  $c(\|\nabla I\|)$ . In [47], two functions were proposed:

$$c(\|\nabla I\|) = e^{[-(\|\nabla I\|)/(K)^2]}, \quad (2)$$

$$c(\|\nabla I\|) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2}, \quad (3)$$

where  $K$  is a constant controlling the edge sensitivity. The function (2) emphasises high-contrast edges in relation to low-contrast edges, while function (3) emphasises larger regions in relation to smaller ones in the image processed.

In this study, function (3) with the value of  $K = 2$  was used. The ADF filtering was executed for 50 iterations. An example of utilising an ADF filter using function (3) is presented in Figure 2e.

### 2.2. Active Contour Models Used

The following two edge-based active contour models were used in the study: EM and GAC, as was the ACWE, which belongs to the group of region-based active contours.

#### 2.2.1. Edge-Based Active Contour Model Using an Inflation/Deflation Force with a Damping Coefficient (EM)

The (EM) model is a recent development [25,50] and its iteration equation has the following form:

$$v_{i+1}(t) = v_i(t) - \overline{w}_1 F_i^{\text{tensile}}(t) + \overline{w}_2 F_i^{\text{flexural}}(t) + w_3 F_i^{\text{external}}(t) - \tau(i) F_i^{\text{inflation/deflation}}(t), \quad (4)$$

whereas:

- $\overline{w_1}, \overline{w_2}$  and  $w_3$  are weighing coefficients.
- $F_i^{tensile}$  is the tensile force. For the node  $i$  at iteration  $t$ :  $F_i^{tensile}(t) = 2v_i(t) - v_{i-1}(t) - v_{i+1}(t)$
- $F_i^{flexural}$  is the flexural force. For the node  $i$  at iteration  $t$ :  $F_i^{flexural}(t) = 2F_i^{tensile}(t) - F_{i-1}^{tensile}(t) + F_{i+1}^{tensile}(t)$ .
- $F_i^{external}(t) = \nabla P(x_i(t), y_i(t))$  is the external force that allows the contour to be moved to regions with higher gradient values in the image. When  $P(x, y) = -c||\nabla(G_\sigma * I(x, y))||$ , the convolution  $G_\sigma * I(x, y)$  serves the purpose of removing noise from the image  $I$  using the Gaussian filter  $G_\sigma$ .
- $F_i^{inflation/deflation}(t) = F(I_s(x_i, y_i))n_i(t)$  is the force which inflates or deflates the contour in the direction of the edges being identified, whereas  $n_i(t)$  represents a unit vector for the node with the index  $i$ :

$$F(x, y) = \begin{cases} -1, & \text{if } I(x, y) < T, \\ +1, & \text{if } I(x, y) \geq T, \end{cases} \quad (5)$$

in turn, is a function that links the inflation force with the image function  $I$  for a set value of the threshold  $T$ .

- The parameter  $\tau_i$  makes it possible to dampen the inflation/deflation force for the node with the index  $i$  at iteration  $t$ . The damping is performed in the vicinity of the identified edges, according to the relationship  $\tau_i \leftarrow DF * \tau_i$ , whereas  $DF$  represents the value of the damping factor.

In addition, the following parameters are determined according to Algorithm 1 based on the publication [25]:

- $\alpha_{min}, \alpha_{max}$  i.e., the minimum and maximum angle between a pair of adjacent nodes;
- $L_{min}, L_{max}$ , which are the minimum and maximum distances between adjacent nodes;
- $inRv$  the maximum number of reversals of the node; after it is exceeded, the value of the inflation/deflation force is dampened;
- number of iterations executed.

### 2.2.2. Geometric Active Contour Using Morphological Operators (GAC)

The GAC model has been presented in publication [39]. Iteration equations of the evolving contour are as follows:

$$u^{n+\frac{1}{3}}(x) = \begin{cases} D_d u^n(x), & \text{if } I(x) \in [I_0, I_1] \text{ and } v > 0, \\ E_d u^n(x), & \text{if } I(x) \in [I_0, I_1] \text{ and } v < 0, \\ u^n(x), & \text{otherwise.} \end{cases} \quad (6)$$

$$u^{n+\frac{2}{3}}(x) = \begin{cases} 1, & \text{if } \nabla u^n(x) \nabla g(I)(x) > 0, \\ 0, & \text{if } \nabla u^n(x) \nabla g(I)(x) < 0, \\ u^{n+\frac{1}{3}}(x), & \text{if } \nabla u^n(x) \nabla g(I)(x) = 0, \end{cases} \quad (7)$$

$$u^{n+1}(x) = \begin{cases} (SI_d \circ IS_d u^n)(x), & \text{if } g(I)(x) \geq 0, \\ u^{n+\frac{2}{3}}(x), & \text{otherwise,} \end{cases} \quad (8)$$

whereas  $D_d$  and  $E_d$  represent the operators of dilation and erosion with a square structural element  $3 \times 3$  pixels in size. The parameter  $v$  represents an inflating ( $v > 0$ ) or a deflating contour ( $v < 0$ ) and is set by the user. The centre of the brightness interval  $[I_0, I_1]$  is calculated as the median value of brightnesses calculated for the contour neighbourhood region using the selected function  $g(I)$ , i.e.,

$$g(I) = \frac{1}{\sqrt{10^{-20} + |\nabla G_\sigma * I|}}, \quad (9)$$

$$g(I) = |G_\sigma * I|. \quad (10)$$

The first function (9) reaches minima where the edges occur in the image  $I$ , while the second function (10) reaches the minima where dark lines occur.  $G_\sigma$  stands for a Gaussian function removing noise from image  $I$ .  $SI_d$  and  $IS_d$  represent morphological line operators [39], and the purpose of the operation  $SI_d \circ IS_d$  is to smooth out the contour during its evolution. In the binary image, the  $SI_d$  and  $IS_d$  operators have the same effect. However, the operator  $SI_d$  removes sharp ends of the contour for the white pixels, while  $IS_d$  does the opposite.

In the computer implementation, the parameter ‘Interval Radius’ represents a certain radius of the brightness interval  $[I_0, I_1]$ , while the ‘attraction force’ represents the selection of the function  $g(I)$  (i.e., Label (9) or (10)). In addition, the number of iterations carried out is also determined.

### 2.2.3. Active Contour without Edges (ACWE)

This is a region-based method, based on publication [26], very frequently used to segment images, including medical ones. In the ACWE model, the following energy criterion has been adopted:

$$E(\phi) = \int_{\Omega} F(I(x), \phi(x)) dx + \lambda \int_{\Omega} \delta(\phi(x)) \|\nabla\| dx, \quad (11)$$

where  $\delta$  is a Dirac function and the  $F$  function is as follows:

$$F(I(x), \phi(x)) = H(\phi(x))(I(x) - v)^2 + (1 - H(\phi(x)))(I(x) - u)^2. \quad (12)$$

$H$ —is a Heaviside function, while  $u$  and  $v$  are parameters updated at each iteration and are as follows:

$$u = \frac{\int_{\Omega} (1 - H(\phi(x))) \cdot I(x)}{\int_{\Omega} (1 - H(\phi(x)))}, \quad (13)$$

$$v = \frac{\int_{\Omega} H(\phi(x)) \cdot I(x)}{\int_{\Omega} H(\phi(x))}. \quad (14)$$

The first integral in Equation (11) corresponds to the data attachment term, while the second integral is a regularization term minimising the length of the curve and thus allowing the contour to be smoothed out during its evolution. The evolution equation is as follows:

$$\frac{\partial \phi}{\partial t}(x) = \delta(\phi(x))((I(x) - v)^2 - (I(x) - u)^2) + \lambda \delta(\phi(x)) \kappa, \quad (15)$$

where  $\kappa = \text{div}\left(\frac{\nabla \phi(x)}{\|\nabla \phi(x)\|}\right)$  corresponds to the curvature of the evolving contour. In the computer implementation, two parameters are used:

- The curvature term  $\lambda$ . This parameter controls the regularization term from the Equation (15).
- The number of iterations executed  $N$ .

### 2.3. Accuracy Measurements of Segmentations Executed in Mammograms

The accuracy of segmenting masses in mammograms from the MIAS and DDSM databases was estimated by comparing the results produced by the three active contour models used—ACWE, EM and GAC—to areas traced manually by the more experienced breast radiologist ( $R_1$ : more than 30 years of radiology practice). Two indexes were used in the measurements: the overlap value  $OV = \frac{|M \cap E|}{|M \cup E|}$  and the extra fraction  $EF = \frac{|M \cap \bar{E}|}{|E|}$ , where:

- $M$  represents a segmentation by an active contour method, while  $|M|$  is the number of pixels in the extracted area.
- $E$  is the area traced by the expert  $R_1$ , while  $|E|$  is the number of pixels in the extracted area.



- $|M \cap R|$ ,  $|M \cup R|$  are respectively: the number of pixels from the common area and the number of all pixels found in the following areas:  $M$  and  $E$ .

If the  $OV$  index is close to 1 and, at the same time, the  $EF$  index is close to 0, this means that the segmentation executed by the computer method is similar to the region of the image hand-traced by the expert. The measurements of the  $OV$  and  $EF$  indexes constitute the quantitative criteria of the segmentations completed. In addition, a qualitative assessment of the segmentation results was made based on the opinion of two breast radiologists:  $R_1$  with over 30 years of professional experience and  $R_2$  with 10 years of experience. The radiologists assessed every completed segmentation with a view to the ability to formulate an adequate or an inadequate hypothesis for the diagnosis to be made. It is obvious that an assessment based only on the radiologists' opinion may be subjective and dependent on their experience. However, the same can be said about the manual tracing of the contour. Executing two assessment procedures, one quantitative and one qualitative, is to minimise the subjectivity because they complement each other.

#### 2.4. Selecting Parameters and Initialising of Active Contour Models

A training set composed of 20 ROIs (10 benign cases and 10 cases of malignant masses of various shapes, sizes, margins and brightnesses) was used to set the parameters of active contour models so that the segmentation results obtained were as accurate as possible, namely the  $OV$  index was the highest, and, at the same time, the  $EF$  index was as low as possible.

Table 2 contains the values of parameters used in the preprocessing for the EM and GAC models. The CO filtering was done with a disk-shaped structuring element with the radius of 3.

The location of the seed contour in individual mammograms is identical for three models: ACWE, EM and GAC. An example initialisation and the location of the contour in the last iteration are presented in Figure 2f,h,k. Both the EM and GAC models used an inflating contour, which increases its surface area from the initial location during subsequent iterations.

Tables 3–5 present the optimal values of parameters used in the ACWE, EM and GAC methods. The values of parameters shown in Tables 3 and 4 were selected so as to:

- Enable the effective evolution of the contour from its initial location and also the effective moving of nodes during subsequent iterations, if they were far from the boundary searched for.
- Produce the best results of the segmentations carried out.

In the GAC model (Table 3), the following three parameters—the maximum iteration number, the attraction force  $g(I)$ , the interval radius—were used to ensure the control of the moving active contour.

For the EM method, Table 4 contains 12 parameters in total, but it is worth emphasising that 11 parameters were defined so that their values remained constant for all experiments carried out. To improve segmentation results, it is enough to select the brightness threshold  $T$  from the interval of values shown in Table 4.

**Table 2.** Preprocessing parameters for the following models without image sampling: EM and GAC, and with image sampling: EM(S) and GAC(S).

Parameter name	ADF: Equation (3)	ADF: K	ADF: No. of iter.	$\sigma$
Parameter value	-	2	50	3
Image Pre-processing: EM, EM(S), GAC, GAC(S)				



**Table 3.** The optimal set of parameters for the GAC model without image sampling, and with image sampling—GAC(S)—for the segmentations completed in mammograms.

Parameter name	No. of iter.	Attraction force	Interval radius
Parameter value	245	(9)–(10)	[24, 83]
GAC			
Parameter name	No. of iter.	Attraction force	Interval radius
Parameter value	130	(9)–(10)	[24, 83]
GAC(S)			

**Table 4.** The optimal set of parameters for the EM model without image sampling, and with image sampling—EM(S)—for the segmentations completed in mammograms.

Par. name	No. of iter.	$\overline{w}_1$	$\overline{w}_2$	$w_3$	$\tau_{init}$	$T$	$L_{min}$	$L_{max}$	$\alpha_{min}$	$\alpha_{max}$	$DF$	$inRv$
Par. value	300	1	0.5	40	8	[0.1, 0.4]	4 pixels	9 pixels	21°	30°	0.8	3
EM												
Par. name	No. of iter.	$\overline{w}_1$	$\overline{w}_2$	$w_3$	$\tau_{init}$	$T$	$L_{min}$	$L_{max}$	$\alpha_{min}$	$\alpha_{max}$	$DF$	$inRv$
Par. value	135	1	0.5	40	8	[0.1, 0.4]	4 pixels	9 pixels	21°	30°	0.8	3
EM(S)												

**Table 5.** The optimal set of parameters for the ACWE model without image sampling, and with image sampling—ACWE(S)—for the segmentations completed in mammograms.

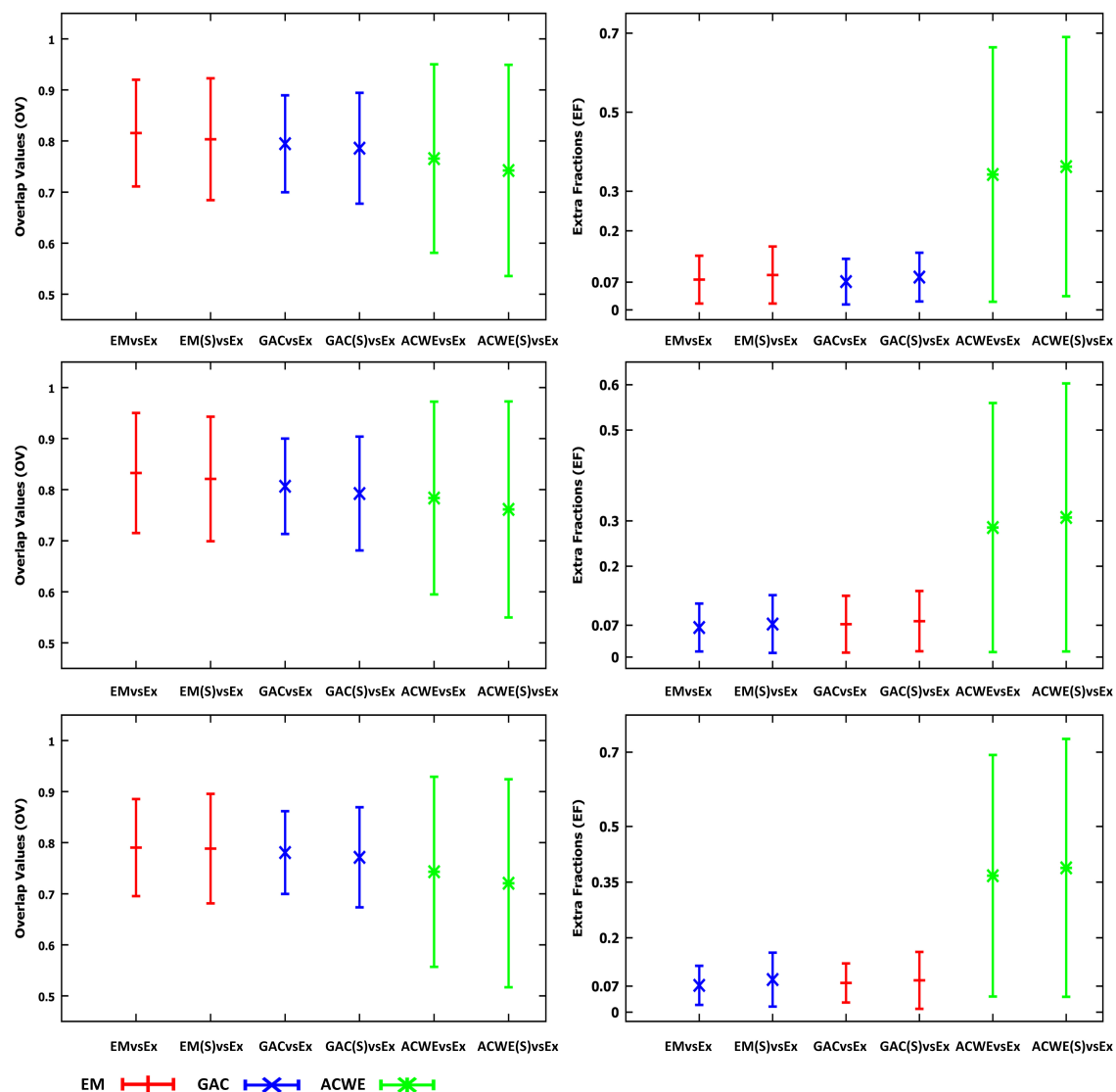
Parameter name	No. of iter.	$\lambda$
Parameter value	800	0.2
ACWE		
Parameter name	No. of iter.	Attraction force
Parameter value	400	0.2
ACWE(S)		

### 3. Results and Discussion

After the parameters enabling the segmentation to be controlled had been established, the active contour methods allowing mass segmentation were tested on the remaining 100 ROIs from mammograms, using contours traced manually by a radiologist ( $R_1$ ). The results obtained are presented in Table 6. Table 6 presents the determined statistical parameters, such as: the maximum value (max), the minimum value (min), the mean value (mean) and the standard deviation (sd) of the following calculated indexes: OV and EF. Figure 4 shows a graph of data from Table 6.

**Table 6.** Measurements of the overlap value (OV), extra fraction (EF) indexes for 100 mammograms from the MIAS and DDSM databases. The symbol (S) represents results of measurements with image sampling. The mean values in bold characters are the best among the values obtained.

	OV	EF	OV(S)	EF(S)		OV	EF	OV(S)	EF(S)
mean	<b>0.8156</b>	0.0764	0.8034	0.088	mean	0.7946	<b>0.0713</b>	0.7858	0.0828
sd	0.1045	0.0605	0.1193	0.0722	sd	0.095	0.0578	0.1086	0.0616
min	0.5467	0	0.5293	0.0031	min	0.7223	0	0.551	0.0015
max	0.9651	0.3028	0.9187	0.3234	max	0.9782	0.3072	0.9223	0.318
Benign and Malignant: EM vs. Expert					Benign and Malignant: GAC vs. Expert				
	OV	EF	OV(S)	EF(S)		OV	EF	OV(S)	EF(S)
mean	<b>0.8327</b>	<b>0.0651</b>	0.821	0.0726	mean	0.8067	0.0723	0.7925	0.079
sd	0.1175	0.0527	0.1219	0.0636	sd	0.0934	0.0626	0.1115	0.0663
min	0.5384	0	0.5293	0.0037	min	0.5677	0	0.551	0.0019
max	0.9263	0.1867	0.9187	0.1934	max	0.9347	0.2208	0.9223	0.2464
Benign: EM vs. Expert					Benign: GAC vs. Expert				
	OV	EF	OV(S)	EF(S)		OV	EF	OV(S)	EF(S)
mean	<b>0.7904</b>	0.0877	0.7884	0.0982	mean	0.7714	<b>0.0854</b>	0.7806	0.0785
sd	0.0951	0.0726	0.1073	0.0858	sd	0.098	0.0767	0.081	0.0526
min	0.5968	0	0.5882	0.0031	min	0.6039	0.0015	0.6157	0
max	0.8972	0.3159	0.8814	0.3205	max	0.8549	0.318	0.8676	0.2957
Malignant: EM vs. Expert					Malignant: GAC vs. Expert				
	OV	EF	OV(S)	EF(S)		OV	EF	OV(S)	EF(S)
mean	0.7656	0.3423	0.7423	0.3623	mean	0.7836	0.2853	0.7614	0.3076
sd	0.1847	0.3218	0.2068	0.3278	sd	0.1887	0.2745	0.2115	0.2953
min	0.4934	0	0.3132	0	min	0.4372	0	0.9004	0
max	0.9508	2.0354	0.9004	2.0598	max	0.9233	1.5237	0.3811	1.5449
Benign and Malignant: ACWE vs. Expert					Benign: ACWE vs. Expert				
	OV	EF	OV(S)	EF(S)		OV	EF	OV(S)	EF(S)
mean	0.7656	0.3423	0.7423	0.3623					
sd	0.1847	0.3218	0.2068	0.3278					
min	0.4934	0	0.3132	0					
max	0.9508	2.0354	0.9004	2.0598					
Malignant: ACWE vs. Expert									



**Figure 4.** Graphs of the mean value and the standard deviation based on measurements of two indexes—overlap value (OV), extra fraction (EF)—for the three applied methods—ACWE, EM and GAC—compared to the contours traced by the radiologist (expert) based on data from Table 6. The symbol (S) represents results of measurements with image sampling, i.e., the bicubic interpolation, applied. First row: benign and malignant cases, second row: benign cases, third row: malignant cases.

The EM model produced insignificantly better results of the OV index (81%) compared to the GAC (79%). The values of the EF index are comparable for EM and GAC and amount to 0.07. The lowest convergence with the contour traced manually by a radiologist ( $R_1$ ) occurs for the ACWE, where the average OV index amounts to 76%, while the extra fraction cannot, unfortunately, be very high as the EF index averages 36%.

After sampling the image by the bicubic interpolation and then applying the EM and GAC models, a slight drop in the OV index occurs along with a slight increase in the EF index, with the value of change for both indexes amounting to about 1%. In the case of the ACWE method, image sampling also causes a slight drop in the OV index coupled with a slight increase of the EF index, at the level of 2% for both of these indexes. The bicubic interpolation may cause an overshoot (a haloing effect) and increase the apparent sharpness [49], but considering the results of measuring the OV and EF indexes, the undesirable effects can be said to be small.

Table 7 presents quantitative segmentation results based on the OV index and the results available in the literature. Unfortunately, segmentation methods very rarely achieve OV values of 90% and higher [51]. Manual segmentations are often large enough to make it possible to place the entire tumour inside the contour traced. In addition, the accuracy of a manual segmentation may be limited, and this greatly depends on the experience of the radiologist participating in the research. According to the data from Table 7, the best results of the average OV were achieved in publication [6], where the index amounted to 87%. The second and third best results were achieved in this study using the following methods: (EM) and (EM(S)), for which the average values of the OV amounted to, respectively 81% and 80%. It should, however, be emphasised that the results compiled in Table 7 concerned different numbers of masses analysed and also different sizes of ROIs, which are not presented in some papers. In addition, the time measurement allows the computational efficiency of the methods used to be determined.

**Table 7.** A comparison of average overlap values of masses, for selected methods, with the number of analysed masses (or the number of images) as well as information whether the segmentation was carried out in a mammogram of the source size (M) or in the ROI (R). In the case of the ROI, its size is also presented if this information is available in the source paper. The table also contains the average time, in seconds, of the segmentations conducted and the database/databases used. The mean values in bold characters are the best among the values obtained.

Methods	No. of Masses or No. of Images	Mean OV	Full Mammo. (M)	ROI (R); Size of ROI in Pixels	Mean Time in Seconds	Database
Ball et al. [30], 2007	60 masses	0.57 $\pm$ 0.06	-	R	-	DDSM
Ball et al. [31], 2007	60 masses	0.74 $\pm$ 0.07	-	R; 2048 $\times$ 2048	-	DDSM
Tao et al. [52], 2010	54 masses	0.69 $\pm$ 0.16	-	R	-	priv. held database
Xu et al. [22], 2011	363 masses	0.72 $\pm$ 0.13	-	R; 125 $\times$ 125	-	DDSM
Rahmati et al. [6], 2012	100 masses	<b>0.87 <math>\pm</math> 0.05</b>	-	R; 256 $\times$ 256	From 10 to 60 s	DDSM
Abbas et al. [53], 2013	480 masses	0.73	-	R	-	DDSM and MIAS
Pereira et al. [54], 2014	160 masses (640 images)	0.79 $\pm$ 0.08	M	-	11.05	DDSM
Cordeiro et al. [55], 2016	57 masses	0.58 $\pm$ 0.24	-	R	-	MIAS
ACWE	100 masses	0.76 $\pm$ 0.18	-	R; 512 $\times$ 512	126 $\pm$ 2	DDSM and MIAS
ACWE(S)	100 masses	0.74 $\pm$ 0.20	-	R; 512 $\times$ 512	20.9 $\pm$ 2	DDSM and MIAS
EM	100 masses	0.81 $\pm$ 0.10	-	R; 512 $\times$ 512	26.6 $\pm$ 5.72	DDSM and MIAS
EM(S)	100 masses	0.80 $\pm$ 0.12	-	R; 512 $\times$ 512	8.42 $\pm$ 1.77	DDSM and MIAS
GAC	100 masses	0.79 $\pm$ 0.09	-	R; 512 $\times$ 512	16.42 $\pm$ 2.3	DDSM and MIAS
GAC(S)	100 masses	0.78 $\pm$ 0.10	-	R; 512 $\times$ 512	<b>5.86 <math>\pm</math> 0.7</b>	DDSM and MIAS

Example differences in the segmentation produced by the ACWE(S), EM(S) and GAC(S) methods and also the regions manually marked by an experienced radiologist ( $R_1$ ) are presented in Figures 5 and 6. Figure 5 shows example results for the MIAS database, Figure 6—examples for the DDSM database.

Table 8 sums up the results of the quantitative assessment of the segmentations carried out. It is worth noting that the assessments of segmentations by radiologists  $R_1$  and  $R_2$  are identical for active contour methods without image sampling and those with image sampling by bicubic interpolation.

The data from Table 8 justifies the following statements:

- The more experienced radiologist ( $R_1$ ) assessed the adequacy of the segmentations carried out as follows: 69%, 81%, and 76%. The assessments of the second radiologist ( $R_2$ ), in turn, are as follows: 74%, 84%, and 81%.
- In general, benign lesions produced higher assessments than malignant ones.

A comparison of the results obtained using image sampling by the bicubic interpolation method presented in Tables 6 and 8 reveals that, whereas image sampling leads to a slight fall of the OV and the simultaneous increase of the EF, it does not influence the qualitative assessment by both radiologists,  $R_1$  and  $R_2$ . In addition, image sampling allows the image processing time to be significantly reduced. Table 9 is a summary of the above results. For the ACWE model, sampling cuts the total average time of image processing six times, for the EM model over three times and for the GAC: 2.6 times.

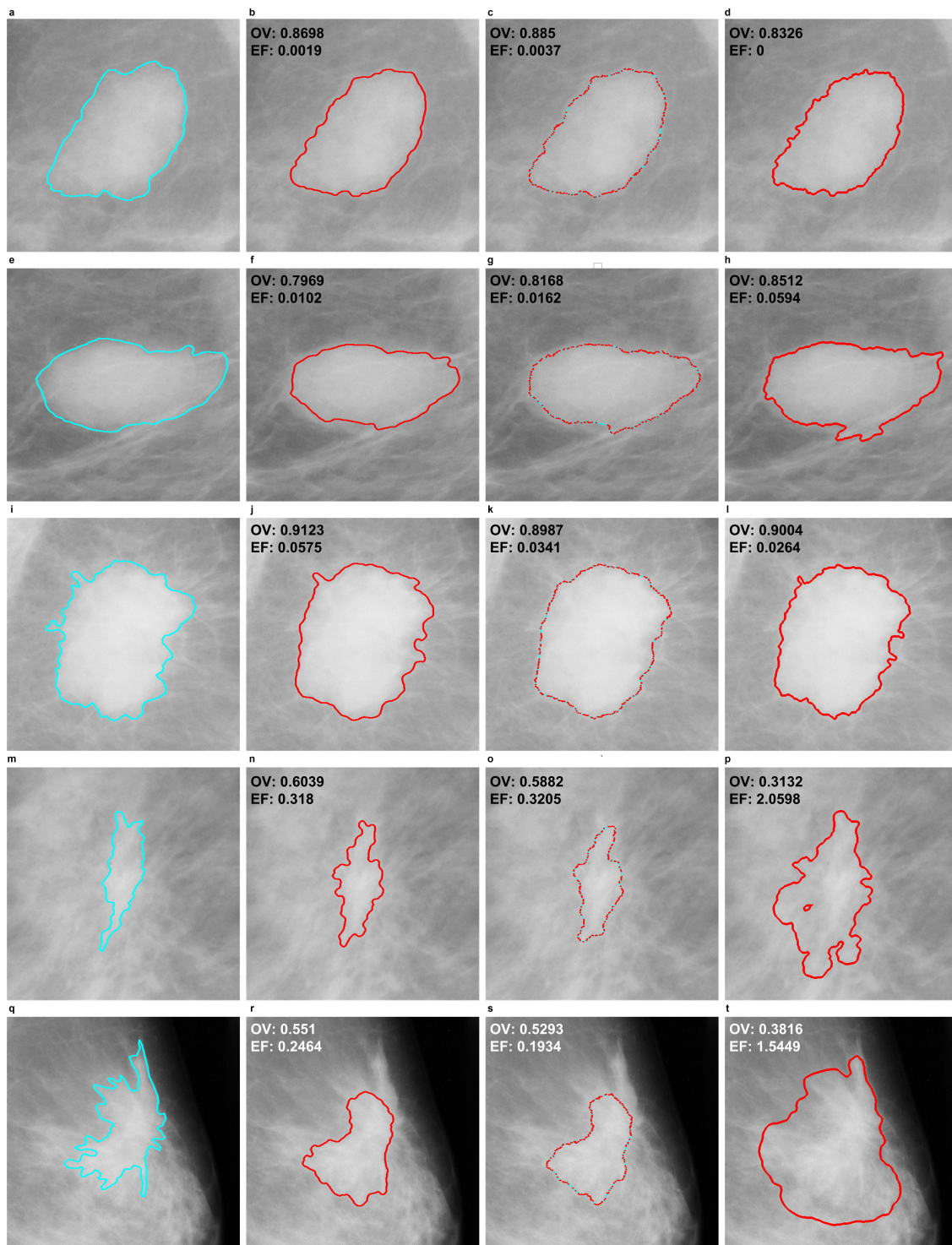
The ACWE, EM and GAC models were implemented in the MatlabR2015a (MathWorks, Inc., Natick, MA, USA) environment under Windows 8/10 and their time was measured on a PC with an Intel Core i7 CPU @ 2GHz and 16 GB of RAM.

**Table 8.** Indicators of the adequate (AD) and inadequate (INAD) segmentation of 100 masses for benign and malignant cases, based on the analysed ROIs, for the following methods: ACWE, ACWE(S), EM, EM(S), GAC, and GAC(S). The results obtained for the ACWE, ACWE(S) are identical. Similarly, the same results were obtained for the following pairs: EM, EM(S) and GAC, GAC(S).  $R_1$ —is the assessment of the segmentation by the breast cardiologist with 30 years of experience, while  $R_2$  is the assessment by the one with 10 years. The values in bold characters are the best among the values obtained.

	AD (benign)	INAD (benign)	AD (malignant)	INAD (malignant)	AD (total)	INAD (total)	Total
$R_1$	36 (72%)	14 (28%)	33 (66%)	17 (34%)	69 (69%)	31 (31%)	100
$R_2$	38 (76%)	12 (24%)	36 (72%)	14 (28%)	74 (74%)	26 (26%)	100
ACWE/ACWE(S)							
	AD (benign)	INAD (benign)	AD (malignant)	INAD (malignant)	AD (total)	INAD (total)	Total
$R_1$	41 (82%)	9 (18%)	40 (80%)	10 (20%)	81 (81%)	20 (19%)	100
$R_2$	<b>43 (86%)</b>	<b>7 (14%)</b>	<b>41 (82%)</b>	<b>9 (18%)</b>	<b>84 (84%)</b>	<b>16 (16%)</b>	100
EM/EM(S)							
	AD (benign)	INAD (benign)	AD (malignant)	INAD (malignant)	AD (total)	INAD (total)	Total
$R_1$	39 (78%)	11 (22%)	37 (74%)	13 (26%)	76 (76%)	24 (24%)	100
$R_2$	41 (82%)	9 (18%)	40 (80%)	10 (20%)	81 (81%)	20 (19%)	100
GAC/GAC(S)							

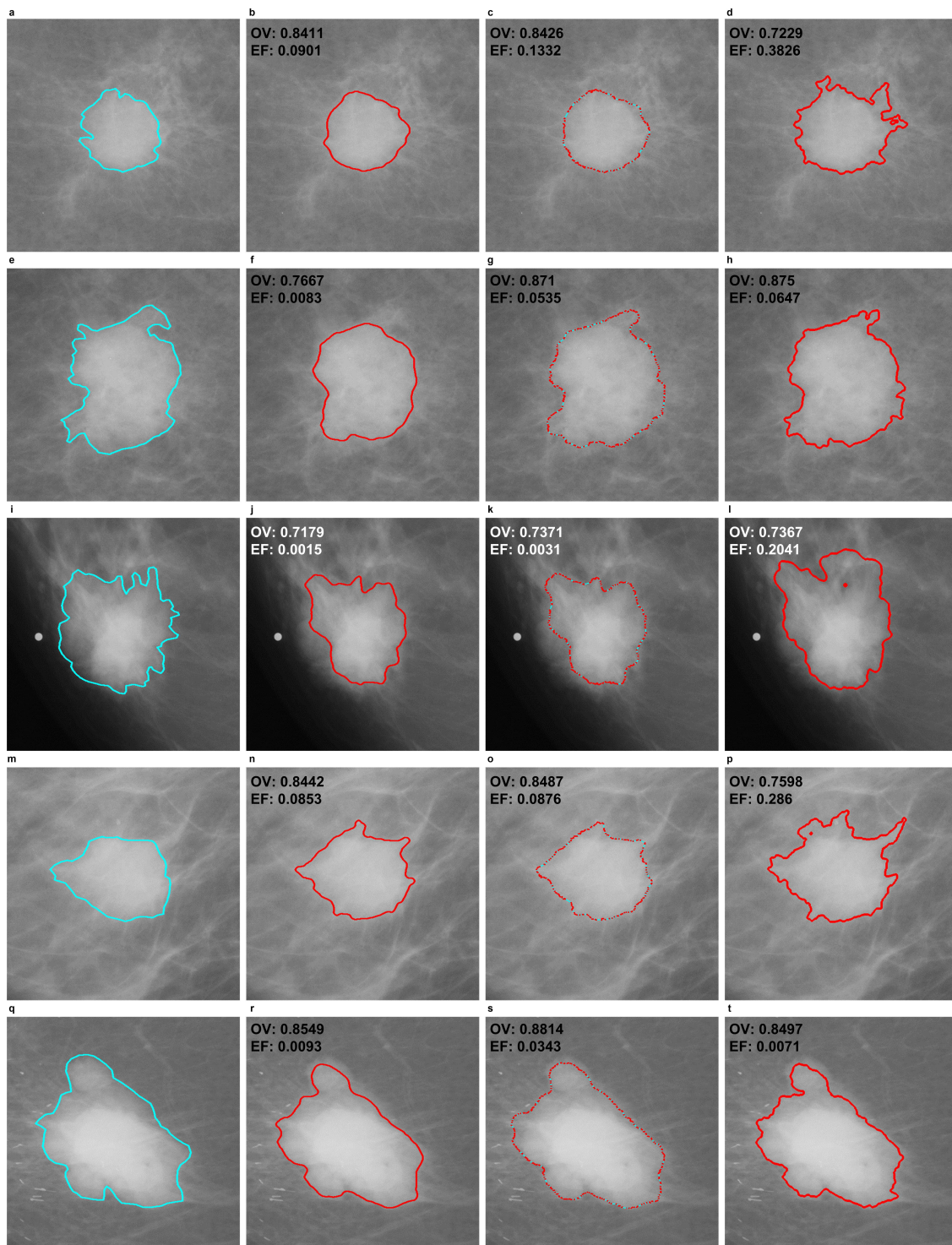
**Table 9.** Time measurements (min, max, mean, sd—standard deviation) of individual steps of the applied methods, for 100 mammograms from the DDSM and MIAS database. The symbol (S) represents mammogram processing with image sampling, the lack of an (S) means that the mammogram was processed without image sampling. Preprocessing: GAC or EM—measurements of the preprocessing time, during which the following are executed in this order: input image inversion, histogram equalisation and ADF, as shown in the diagram from Figure 1. Preprocessing(S): ACWE, represents the measurements of the preprocessing time during which the CO filtration and image sampling are executed in this order—Figure 1.  $\Sigma(1),(5),(11)$  represents the total time measurements for the three methods in the following order: (1),(5),(11). Similarly:  $\Sigma(2),(6)$ ,  $\Sigma(1),(7),(11)$ ,  $\Sigma(2),(8)$ ,  $\Sigma(3),(9),(11)$ ,  $\Sigma(4),(10)$ . The mean values in bold characters are the best of average values obtained for individual steps of the methods used.

Method	Min	Max	Mean	sd
(1) Preproc.(S): EM or GAC	2.906	3.346	3.113	0.098
(2) Preproc.: EM or GAC	6.17	6.954	6.686	0.184
(3) Preproc.(S): ACWE	0.89	0.902	0.894	0.007
(4) Preproc.: ACWE	0.775	0.843	<b>0.792</b>	0.009
(5) GAC(S)	1.538	4.106	<b>2.626</b>	0.706
(6) GAC	5.484	12.512	8.711	2.098
(7) EM(S)	2.307	9.06	5.223	1.744
(8) EM	13.047	29.027	19.823	5.477
(9) ACWE(S)	14.972	21.254	17.4	2.207
(10) ACWE	122.03	129.21	125.4	2.021
(11) Postproc.(S)	0.119	0.142	<b>0.138</b>	0.003
$\Sigma(1),(5),(11)$	4.803	7.436	<b>5.862</b>	0.737
$\Sigma(2),(6)$	11.721	19.532	15.427	2.317
$\Sigma(1),(7),(11)$	5.491	12.274	8.429	1.774
$\Sigma(2),(8)$	19.324	36.127	26.607	5.723
$\Sigma(3),(9),(11)$	18.362	25.287	20.904	2.288
$\Sigma(4),(10)$	126.3	130.6	126.321	2.032



**Figure 5.** The use of the GAC(S), EM(S) and ACWE(S) methods to segment the lesions—masses, for example mammograms from the MIAS database, in which the image was sampled (S) by bicubic interpolation. The contours in the first column have been traced by a radiologist, and columns two, three and four show contours were determined using the following methods: GAC(S), EM(S), and ACWE(S). (a–d) image ‘mdb015’ (mass shape: oval, margins circumscribed, pathology benign); (e–h) ‘mdb025’ (mass shape: oval, margins circumscribed, pathology benign); (i–l) ‘mdb184’ (mass shape: oval—irregular, margins spiculated, pathology malignant); (m–p) ‘mdb145’ (mass shape: irregular, margins spiculated, pathology benign); (q–t) ‘mdb198’ (mass shape: irregular, margins spiculated, pathology benign).





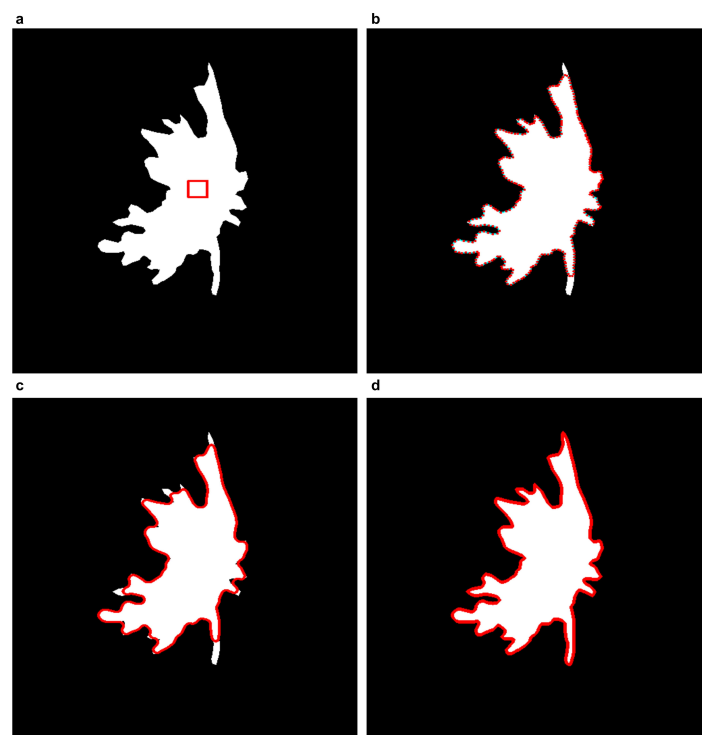
**Figure 6.** The use of the GAC(S), EM(S) and ACWE(S) methods to segment the lesions—masses, for example mammograms from the DDSM database, in which the image was sampled (S) by bicubic interpolation. The contours in the first column have been traced by a radiologist, and columns two, three and four show contours determined using the following methods: GAC(S), EM(S), and ACWE(S) (a–d) image B\_3049\_1.LEFT\_CC (mass shape: round, margins spiculated, pathology malignant); (e–h) B\_3084\_1.RIGHT\_CC (mass shape: lobulated-irregular, margins spiculated, pathology malignant); (i–l) C\_0006\_1.RIGHT\_CC (mass shape: irregular, margins spiculated, pathology malignant); (m–p) C\_0009\_1.RIGHT\_MLO. (mass shape: irregular, margins spiculated, pathology malignant); (q–t) C\_0011\_1.RIGHT\_CC (mass shape: lobulated, margins circumscribed, pathology malignant).

### 3.1. Intensity Inhomogeneity of Pixels and Fuzzy Edges in the Analysed ROIs

If the segmentation is carried out in ROIs in which the intensity of pixels is strongly inhomogeneous within the masses and/or the approximated edges are very fuzzy and merge with the image background, this may make active contour methods extract the masses inaccurately. The use of image preprocessing techniques can only partly improve the segmentation results. The problems described can be illustrated by the following examples for masses that are irregular in shape and have spiculated margins, i.e.,

- Figure 5n–p as compared to Figure 5m (mass shape: irregular, margins spiculated, pathology benign) and also Figure 5r–t (mass shape: irregular, margins spiculated, pathology benign) as compared to Figure 5q. It can be seen that the region segmented by computer methods is both under- and overestimated
- Figure 6j–l as compared to Figure 6i (mass shape: irregular, margins spiculated, pathology malignant). Here, it can be seen that ACMs underestimate the region to some extent.
- Figure 6n–p (mass shape: irregular, margins spiculated, pathology malignant) on the right-hand side, one can see a small overestimated area in the segmentations produced by ACMs. In the case of ACWE, the overestimated area is the largest.

The active contour methods—ACWE, EM and GAC—correctly extract the concavities of masses, unless the image has an inhomogeneous pixel intensity and/or the edges approximated are very fuzzy and merge with the background. Figure 7 shows an example of extracting a mass of an irregular shape with spiculated margins from the example mammogram ‘mdb198’ (Figure 5q). The results produced by ACWE in a black-and-white image are the best, but, in the real ROI, the opposite situation occurs, as shown in Figure 5t (OV = 0.38, EF = 1.54).



**Figure 7.** The use of the EM, GAC and ACWE methods to extract an irregular shape (with concavities), based on the margins traced by a radiologist, from the example mammogram ‘mdb198’, from Figure 5q. The analysed image is black-and-white. (a) initial contour; (b) EM (c) GAC (d) ACWE.

### 3.2. A Comparison of the Approach Used and the Experiment Results with Other Methods

Regardless of valuable research being completed, not all articles from the literature include detailed information significant for the full assessment of the segmentations executed. Taking into account the results obtained as part of this project and those from the four most recent articles from Table 7, it can be stated that:

- Articles [53,55] do not give the processing time of the segmentation methods used.
- Articles [53,55] do not specify the resolution of the analysed ROIs.
- The results from articles [6,54,55] are not split into benign and malignant cases.
- Articles [6,54] do not specify the type of density of tissues analysed.
- Article [54] does not describe the types of shapes and margins of the analysed masses. In addition, it is not known whether cases with protruding spicules and with a strong intensity inhomogeneity were analysed, and they are the most difficult to segment.
- The impact of the decrease or increase of the OV index on the radiologists' quality assessment of the mass segmentations executed was not researched. This is why it is worth proposing more computationally efficient approaches, as a slight drop in the OV index does not impact the qualitative assessment of the segmentation by radiologists, as presented in Table 8 of this paper.
- Segmentation results can be improved using preprocessing methods suitably matched to the approach followed.
- The properties of the EM model, namely the use of different forces, i.e., the external force, the tensile force and the inflation/deflation force with a damping coefficient made it possible to produce results better than those from the GAC model [39] and other methods, as shown by the data based on Table 7, with the exception of the MLACMLS model [6], which yielded the best segmentation results as assessed by the OV index. The EM model received a higher qualitative assessment than the GAC model, as shown by results from Table 8. The computational efficiency, in turn, of the GAC model is better than of the EM model.

Whereas the existing or new approaches can produce good results for prominent edges (including those amplified by preprocessing methods applied), the most difficult thing is to segment cases with protruding spicules and a strong intensity inhomogeneity. This is still an open problem that requires further research. Another important direction of research could be to automate the process of segmenting masses for mammograms in full resolution: promising results were described in article [54].

## 4. Conclusions

The original contributions of this paper consist of: (1) employing an edge-based active contour model using an inflation/deflation force with a damping coefficient (EM) together with the proposed preprocessing methods to segment suspicious looking masses in mammograms; (2) assessing the results of the EM segmentation, in comparison with other active contour methods implemented, namely ACWE, B-spline level set, GAC, RSF; and (3) using radiologists' knowledge at the last step of the segmentation. The EM model has been able to produce segmentation results that are better than or similar to those of other methods presented in literature. However, in the case of masses that are irregular in shape and have spiculated margins merging with the background of the mammogram, the ability to use the EM and also the other compared models—ACWE and GAC—is limited. In addition, the presence of a strong inhomogeneity of pixel intensity within the mass also makes obtaining satisfactory segmentation results difficult. To summarise the results of the qualitative assessment of the EM using the knowledge of radiologists, it can be said that the segmentation results would have been adequate for formulating the diagnostic hypothesis in at least 80% (40/50) of malignant cases and 82% (41/50) of benign ones. These results are promising and show some potential of the EM model for the computer assistance of diagnostics of mammograms. Further research should focus on improving the segmentation results for the difficult examples described, as well as the possible automation of the selection of active contour parameters (EM). Another direction of further research

could be to develop or adapt existing methods to enable classifying masses and/or microcalcifications because multifunctional CAD systems are the most desirable in practice.

**Acknowledgments:** I thank Olga Kruszelnicka from the John Paul II Specialist Hospital in Kraków in Poland for providing support and medical consultations.

**Conflicts of Interest:** The author declares no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

EM	edge-based active contour model using an inflation/deflation force with a damping coefficient
GAC	geometric active contour model
ACWE	active contour without edges
MLACMLS	maximum likelihood active contour model using level sets
OV	overlap value
EF	extra fraction
DDSM	Digital Database for Screening Mammography
MIAS	Mammographic Image Analysis Society
ACM	active contour method
ROI	region of interest
RSF	region-scalable fitting model
CC	cranio-caudal projection
MLO	medio-lateral oblique projection
CAD	computer aided diagnosis
ADF	anisotropic diffusion filtering
CO	close-open filtering

## References

1. Bankman, I. *Handbook of Medical Image Processing and Analysis*; Academic Press: Cambridge, MA, USA, 2008.
2. Bhateja, V.; Misra, M.; Urooj, S.; Lay-Ekuakille, A. A robust polynomial filtering framework for mammographic image enhancement from biomedical sensors. *IEEE J. Sens.* **2013**, *13*, 4147–4156.
3. He, W.; Hogg, P.; Juetten, A.; Denton, E.R.; Zwiggelaar, R. Breast image pre-processing for mammographic tissue segmentation. *Comput. Biol. Med.* **2015**, *67*, 61–73.
4. Kupinski, M.A.; Giger, M.L. Automated seeded lesion segmentation on digital mammograms. *IEEE Trans. Med. Imaging* **1998**, *17*, 510–517.
5. Oliver, A.; Freixenet, J.; Martí, J.; Pérez, E.; Pont, J.; Denton, E.R.; Zwiggelaar, R. A review of automatic mass detection and segmentation in mammographic images. *Med. Image Anal.* **2010**, *14*, 87–110.
6. Rahmati, P.; Adler, A.; Hamarneh, G. Mammography segmentation with maximum likelihood active contours. *Med. Image Anal.* **2012**, *16*, 1167–1186.
7. Thieu, Q.T.; Luong, M.; Rocchisani, J.M.; Sirakov, N.M.; Viennet, E. Efficient segmentation with the convex local-global fuzzy Gaussian distribution active contour for medical applications. *Ann. Math. Artif. Intell.* **2015**, *75*, 249–266.
8. Demirkazık, F.B.; Gülsün, M.; Fırat, P. Mammographic features of nonpalpable spiculated lesions. *Clin. Imaging* **2003**, *27*, 293–297.
9. Mete, M.; Sirakov, N.M. Dermoscopic diagnosis of melanoma in a 4D space constructed by active contour extracted features. *Comput. Med. Imaging Graph.* **2012**, *36*, 572–579.
10. Liu, X.; Zeng, Z. A new automatic mass detection method for breast cancer with false positive reduction. *Neurocomputing* **2015**, *152*, 388–402.
11. Cheng, H.; Shi, X.; Min, R.; Hu, L.; Cai, X.; Du, H. Approaches for automated detection and classification of masses in mammograms. *Pattern Recognit.* **2006**, *39*, 646–668.
12. Choi, J.Y.; Kim, D.H.; Plataniotis, K.N.; Ro, Y.M. Computer-aided detection (CAD) of breast masses in mammography: Combined detection and ensemble classification. *Phys. Med. Biol.* **2014**, *59*, 3697.

13. Dong, M.; Lu, X.; Ma, Y.; Guo, Y.; Ma, Y.; Wang, K. An efficient approach for automated mass segmentation and classification in mammograms. *J. Digit. Imaging* **2015**, *28*, 613–625.
14. Jeske, J.; Bernstein, J.R.; Stull, M.A. Screening and diagnostic imaging. In *American Cancer Society Atlas of Clinical Oncology*; B.C. Decker: London, UK, 2000; pp. 41–63.
15. Tourassi, G.D.; Vargas-Voracek, R.; Catarious, D.M.; Floyd, C.E. Computer-assisted detection of mammographic masses: A template matching scheme based on mutual information. *Med. Phys.* **2003**, *30*, 2123–2130.
16. Székely, N.; Tóth, N.; Pataki, B. A hybrid system for detecting masses in mammographic images. *IEEE Trans. Instrum. Meas.* **2006**, *55*, 944–952.
17. Freixenet, J.; Oliver, A.; Martí, R.; Lladó, X.; Pont, J.; Pérez, E.; Denton, E.R.; Zwigelaar, R. Eigendetection of masses considering false positive reduction and breast density information. *Med. Phys.* **2008**, *35*, 1840–1853.
18. Bator, M.; Nieniewski, M. Detection of cancerous masses in mammograms by template matching: Optimization of template brightness distribution by means of evolutionary algorithm. *J. Digit. Imaging* **2012**, *25*, 162–172.
19. Suliga, M.; Deklerck, R.; Nyssen, E. Markov random field-based clustering applied to the segmentation of masses in digital mammograms. *Comput. Med. Imaging Graph.* **2008**, *32*, 502–512.
20. Heath, M.D.; Bowyer, K.W. Mass detection by relative image intensity. In Proceedings of the 5th International Workshop on Digital Mammography, Toronto, ON, Canada, 11–14 June 2000; pp. 219–225.
21. Wei, J.; Sahiner, B.; Hadjiiski, L.M.; Chan, H.P.; Petrick, N.; Helvie, M.A.; Roubidoux, M.A.; Ge, J.; Zhou, C. Computer-aided detection of breast masses on full field digital mammograms. *Med. Phys.* **2005**, *32*, 2827–2838.
22. Xu, S.; Liu, H.; Song, E. Marker-controlled watershed for lesion segmentation in mammograms. *J. Digit. Imaging* **2011**, *24*, 754–763.
23. Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1988**, *1*, 321–331.
24. Paragios, N.; Deriche, R. Geodesic active regions and level set methods for motion estimation and tracking. *Comput. Vis. Image Underst.* **2005**, *97*, 259–282.
25. Ciecholewski, M. An edge-based active contour model using an inflation/deflation force with a damping coefficient. *Expert Syst. Appl.* **2016**, *44*, 22–36.
26. Chan, T.F.; Vese, L.A. Active contours without edges. *IEEE Trans. Image Process.* **2001**, *10*, 266–277.
27. Dubrovina-Karni, A.; Rosman, G.; Kimmel, R. Multi-region active contours with a single level set function. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1585–1601.
28. Li, C.; Kao, C.Y.; Gore, J.C.; Ding, Z. Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans. Image Process.* **2008**, *17*, 1940–1949.
29. Bernard, O.; Fréchet, D.; Thévenaz, P.; Unser, M. Variational B-spline level-set: A linear filtering approach for fast deformable model evolution. *IEEE Trans. Image Process.* **2009**, *18*, 1179–1191.
30. Ball, J.E.; Bruce, L.M. Digital mammogram spiculated mass detection and spicule segmentation using level sets. In Proceedings of the IEEE 29th Annual International Conference on Engineering in Medicine and Biology Society (EMBS), Lyon, France, 22–26 August 2007; pp. 4979–4984.
31. Ball, J.E.; Bruce, L.M. Digital mammographic computer aided diagnosis (CAD) using adaptive level set segmentation. In Proceedings of the IEEE 29th Annual International Conference on Engineering in Medicine and Biology Society (EMBS), Lyon, France, 22–26 August 2007; pp. 4973–4978.
32. Jaccard, P. Nouvelles recherches sur la distribution florale. *Bull. Soc. Vard. Sci. Nat.* **1908**, *44*, 223–270.
33. Elter, M.; Horsch, A. CADx of mammographic masses and clustered microcalcifications: A review. *Med. Phys.* **2009**, *36*, 2052–2068.
34. Rojas-Domínguez, A.; Nandi, A.K. Development of tolerant features for characterization of masses in mammograms. *Comput. Biol. Med.* **2009**, *39*, 678–688.
35. Gulshan, V.; Rother, C.; Criminisi, A.; Blake, A.; Zisserman, A. Geodesic star convexity for interactive image segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 13–18 June 2010; pp. 3129–3136.
36. Couprie, C.; Najman, L.; Talbot, H. Seeded segmentation methods for medical image analysis. In *Medical Image Processing*; Springer: Berlin, Germany, 2011; pp. 27–57.
37. Panagiotakis, C.; Grinias, I.; Tziritas, G. Natural image segmentation based on tree equipartition, bayesian flooding and region merging. *IEEE Trans. Image Process.* **2011**, *20*, 2276–2287.



38. Panagiotakis, C.; Papadakis, H.; Grinias, E.; Komodakis, N.; Fragopoulou, P.; Tziritas, G. Interactive image segmentation based on synthetic graph coordinates. *Pattern Recognit.* **2013**, *46*, 2940–2952.
39. Marquez-Neila, P.; Baumela, L.; Alvarez, L. A morphological approach to curvature-based evolution of curves and surfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *36*, 2–17.
40. Heath, M.; Bowyer, K.; Kopans, D.; Kegelmeyer, P., Jr.; Moore, R.; Chang, K.; Munishkumaran, S. Current status of the digital database for screening mammography. In *Digital Mammography*; Springer: Berlin, Germany, 1998; pp. 457–460.
41. Heath, M.; Bowyer, K.; Kopans, D.; Moore, R.; Kegelmeyer, W.P. The digital database for screening mammography. In Proceedings of the 5th International Workshop on Digital Mammography, Toronto, ON, Canada, 11–14 June 2000; pp. 212–218.
42. Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, I.; Stamatakis, E.; Cerneaz, N.; Kok, S.; et al. The mammographic image analysis society digital mammogram database. In *Excerpta Medica. International Congress Series*; Elsevier: Amsterdam, The Netherlands, 1994; Volume 1069, pp. 375–378.
43. Sickles, E.; D'Orsi, C.; Bassett, L.; Appleton, C.; Berg, W.; Burnside, E. *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*; American College of Radiology: Reston, VA, USA, 2013; pp. 39–48.
44. Bator, M.; Chmielewski, L.J. Finding regions of interest for cancerous masses enhanced by elimination of linear structures and considerations on detection correctness measures in mammography. *Pattern Anal. Appl.* **2009**, *12*, 377.
45. Agrawal, P.; Vatsa, M.; Singh, R. Saliency based mass detection from screening mammograms. *Signal Process.* **2014**, *99*, 29–47.
46. Jen, C.C.; Yu, S.S. Automatic detection of abnormal mammograms in mammographic images. *Expert Syst. Appl.* **2015**, *42*, 3048–3055.
47. Perona, P.; Malik, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 629–639.
48. Alemán-Flores, M.; Álvarez, L.; Caselles, V. Texture-oriented anisotropic filtering and geodesic active contours in breast tumor ultrasound segmentation. *J. Math. Imaging Vis.* **2007**, *28*, 81–97.
49. Keys, R. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* **1981**, *29*, 1153–1160.
50. Ciecholewski, M. An edge-based active contour model using an inflation/deflation force with a damping coefficient. Available online: [https://www.youtube.com/playlist?list=PLrwO0xQ2T\\_8SAvs8yXZnvkNIMjulGsgUF](https://www.youtube.com/playlist?list=PLrwO0xQ2T_8SAvs8yXZnvkNIMjulGsgUF) (accessed on 15 November 2016).
51. Timp, S.; Karssemeijer, N. A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography. *Med. Phys.* **2004**, *31*, 958–971.
52. Tao, Y.; Lo, S.C.B.; Freedman, M.T.; Makariou, E.; Xuan, J. Multilevel learning-based segmentation of ill-defined and spiculated masses in mammograms. *Med. Phys.* **2010**, *37*, 5993–6002.
53. Abbas, Q.; Celebi, M.E.; Garcia, I.F. Breast mass segmentation using region-based and edge-based methods in a 4-stage multiscale system. *Biomed. Signal Process. Control* **2013**, *8*, 204–214.
54. Pereira, D.C.; Ramos, R.P.; Do Nascimento, M.Z. Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. *Comput. Methods Progr. Biomed.* **2014**, *114*, 88–101.
55. Cordeiro, F.R.; Santos, W.P.; Silva-Filho, A.G. An adaptive semi-supervised Fuzzy GrowCut algorithm to segment masses of regions of interest of mammographic images. *Appl. Soft Comput.* **2016**, *46*, 613–628.

