Applied Soft Computing 46 (2016) 665–676

Contents lists available at ScienceDirect

Applied Soft Computing

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

On the performance of nature inspired algorithms for the automatic segmentation of coronary arteries using Gaussian matched filters

Ivan Cruz-Aceves [a,*], Arturo Hernandez-Aguirre [b], S. Ivvan Valdez [b]

[a] CONACYT Research Fellow – Centro de Investigación en Matemáticas (CIMAT), A.C., Jalisco S/N, Col. Valenciana, C.P. 36000 Guanajuato, Gto, Mexico
[b] Centro de Investigación en Matemáticas (CIMAT), A.C., Jalisco S/N, Col. Valenciana, C.P. 36000 Guanajuato, Gto, Mexico

ARTICLE INFO

Article history:
Available online 28 January 2016

Keywords:
Automatic segmentation
Coronary angiograms
Gaussian matched filters
Medical imaging
Nature inspired algorithms
Vessel detection

ABSTRACT

This paper presents a comparative analysis of four nature inspired algorithms to improve the training stage of a segmentation strategy based on Gaussian matched filters (GMF) for X-ray coronary angiograms. The statistical results reveal that the method of differential evolution (DE) outperforms the considered algorithms in terms of convergence to the optimal solution. From the potential solutions acquired by DE, the area (A_z) under the receiver operating characteristic curve is used as fitness function to establish the best GMF parameters. The GMF-DE method demonstrated high accuracy with $A_z = 0.9402$ with a training set of 40 angiograms. Moreover, to evaluate the performance of the coronary artery segmentation method compared to the ground-truth vessels hand-labeled by a specialist, measures of sensitivity, specificity and accuracy have been adopted. According to the experimental results, GMF-DE has obtained high coronary artery segmentation rate compared with six state-of-the-art methods provided an average accuracy of 0.9134 with a test set of 40 angiograms. Additionally, the experimental results in terms of segmentation accuracy, have also shown that the GMF-DE can be highly suitable for clinical decision support in cardiology.

1. Introduction

Coronary angiography is an X-ray imaging procedure designed to help cardiologists in diagnosing and treating coronary artery abnormalities. In recent years, the application and development of computational methods for automatic segmentation of coronary arteries has become essential for systems that perform computer-aided diagnosis (CAD) in cardiology. The two main challenges in X-ray coronary angiograms are the uneven illumination and the weak contrast between blood vessels and image background, which are illustrated in Fig. 1. Since the nonuniform illumination along vessel structures produces multimodal histograms, the vessel segmentation problem has been frequently addressed in two stages. The first stage is vessel detection, also called enhancement, which is used to enhance vessel-like structures while removing noise from the angiogram, and the second stage focuses on applying a classification technique over the enhanced angiogram to discriminate vessel and nonvessel pixels.

Due to the importance of the automatic vessel segmentation problem, a number of computational methods have been introduced for different types of medical images. In literature, a few methods in the image frequency domain have been proposed such as Gabor filters [1–3], and wavelet transform [4]. Most of the proposed methods are computed in the spatial image domain including single-scale top-hat operator [5], multiscale top-hat operator [6,7], hit-or-miss transform [8], region-growing with differential-geometry [9], Hessian matrix [10–14], and Gaussian matched filters (GMF) [15], which have been successfully applied in retinal imaging studies [16,17].

The GMF method represents a template matching technique used for detecting vessel-like structures in the spatial image domain. GMF assumes that by using a Gaussian curve as matching template, the shape of blood vessels can be detected. Generally, the Gaussian template is rotated at different orientations, and then convolved with the input image to form a filter bank. The maximum response at each pixel is preserved to acquire the enhanced image. The appropriate performance of the GMF directly depends of four parameters, which have to be tuned in order to increase the vessel detection rate. The first parameter L determines the length of the vessel segment to be detected. The second parameter σ defines the spread of the intensity profile. The third parameter T determines the position in which, the Gaussian curve trails will cut, and the

* Corresponding author. Tel.: +52 473 732 7155/735 0800x4506; fax: +52 473 732 5749.
E-mail addresses: ivan.cruz@cimat.mx (I. Cruz-Aceves), artha@cimat.mx (A. Hernandez-Aguirre), ivvan@cimat.mx (S.I. Valdez).

http://dx.doi.org/10.1016/j.asoc.2016.01.030
1568-4946/

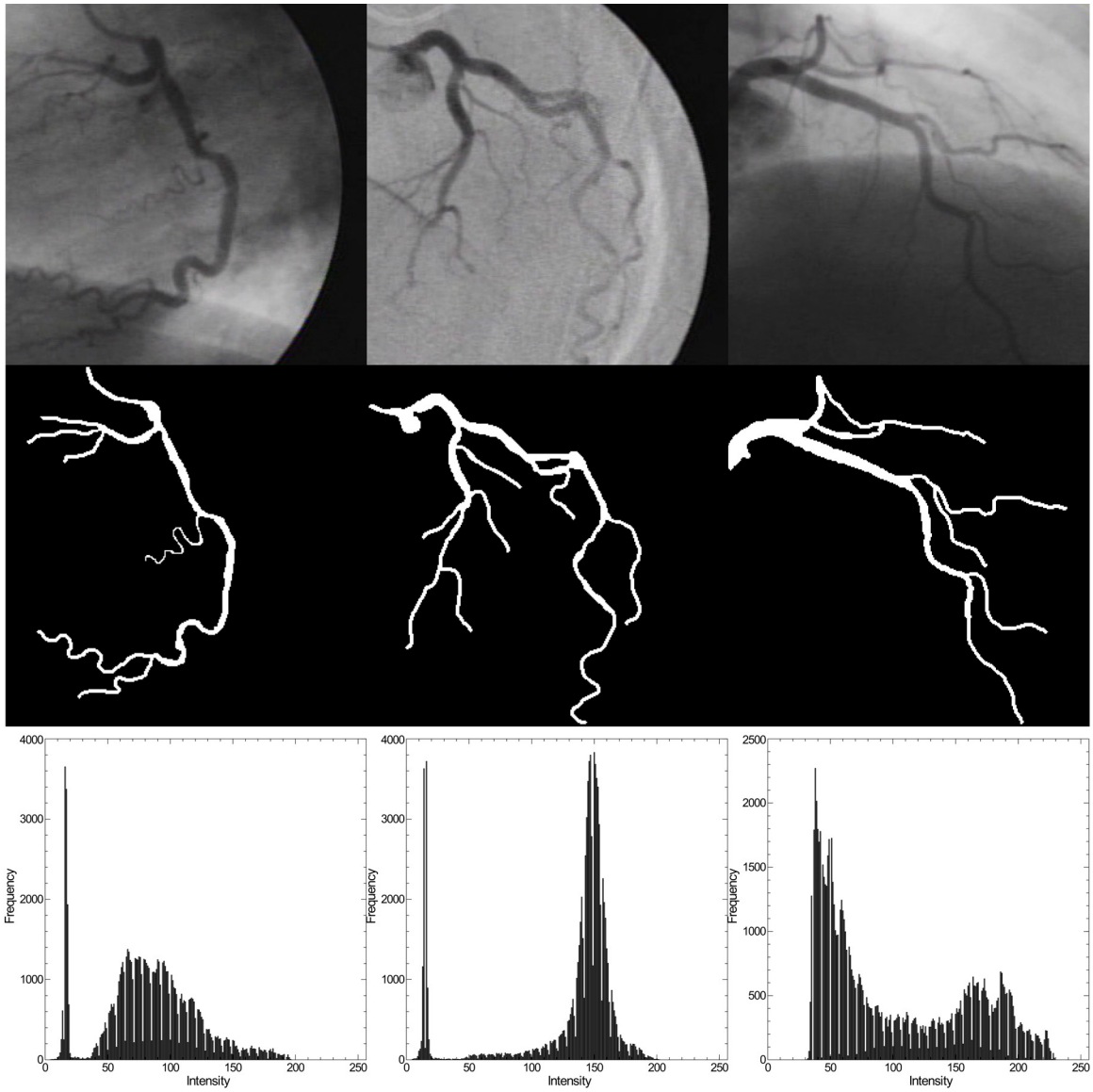

Fig. 1. First row: three coronary angiograms. Second row: manual delineations of vessels in the images of the first row, as drawn by an expert (ground-truth). Last row: histograms of the angiograms shown in the first row of the figure.

last parameter κ defines the number of evenly spaced filters with angular resolution θ.

This vessel detection technique has been successfully applied as a preprocessing stage in diverse automatic segmentation and registration methods. Chanwimaluang and Fan [16] applied the GMF as detection method followed by the local entropy thresholding technique as classification procedure to segment vessels in retinal fundus images. Subsequently, Chanwimaluang et al. [17] applied the above mentioned method, as part of a system for retinal image registration. Kang et al. [18] applied the GMF in a fusion strategy with the morphology-based top-hat operator to detect vessels in coronary angiograms. This fusion strategy was also used by Kang et al. [19,20] by applying the degree segmentation method to classify vessel pixels.

Since the GMF method was introduced by Chaudhuri et al. [15], different values for each parameter of the detection filter have been proposed. Kang et al. [18–20] proposed different values for σ and number of oriented filters κ. Cinsdikici and Aydin [21] used the original parameters of the method, just modifying the parameter of number of oriented filters κ. Al-Rawi et al. [22] proposed to apply an exhaustive search over an extended range of the variables L, T, and σ, keeping constant the number of filters. This method obtains better detection results than the aforementioned methods, taking into account the area (A_z) under the receiver operating characteristic (ROC) curve. Al-Rawi and Karajeh [23] proposed to apply the population-based technique of genetic algorithms (GAs) instead of the exhaustive search to select the optimal L, T, and σ values. According to the tests, the performance of the genetic algorithm working together with the GMF method provided better detection results than the empirically determined methods.

The population-based methods have become very popular to solve optimization problems in discrete and continuous spaces. These methods consist of a set of potential solutions, which are gradually improved by using a fitness function until a stopping criterion is satisfied. Recently, nature inspired algorithms such as particle swarm optimization (PSO), differential evolution (DE), and estimation of distribution algorithms (EDAs) have begun to attract more attention for solving optimization problems with a fast convergence including parameter estimation in computational biology [24], image segmentation [25], and cancer classification [26].

Considering the relevance of the Gaussian matched filters to remove image noise while enhancing vessel-like structures, an optimization process to avoid an exhaustive search for selecting the optimal GMF parameters plays an essential role. In this paper a comparative analysis of four nature inspired algorithms to improve the training stage of the GMF method through the optimal parameter selection is presented. In this approach, the optimization process is carried out over the L, T, and σ parameters whose values are computed by using GAs, EDAs, DE or the PSO algorithm by using the area A_z under the ROC curve. These algorithms are experimentally compared based on the statistical significance of the number of evaluations, dispersion, and computational time. Finally, the coronary artery segmentation results obtained from the GMF method are compared with those obtained with six, previously mentioned, state-of-the-art vessel segmentation methods using the measures of sensitivity, specificity, and accuracy.

The remainder of this paper is organized as follows. In Section 2, the database of X-ray coronary angiograms, and the fundamentals of the GMF and considered nature inspired methods along with the optimization problem are introduced. The experimental results are presented and discussed in Section 3, and conclusions are given in Section 4.

2. Materials and methods

Due to the suitable performance of genetic algorithms to select the optimal parameters of the Gaussian matched filters for blood vessel detection [23], a comparative analysis with estimation of distribution algorithms, differential evolution, and particle swarm optimization is of interest in the present work; these methods are described in detail in the present section.

2.1. Database of coronary angiograms

The database used in the present research is composed of 80 X-ray coronary angiograms of 27 different patients. Each angiographic image is of size 300×300 pixels. To evaluate the performance of the computational methods, the database has been divided into two subsets, 40 of the 80 angiograms are used as training set, and the remaining 40 angiograms are used as test set. The ground-truth angiograms were hand-labeled by a specialist, and ethics approval was provided by the cardiology department of the Mexican Social Security Institute, UMAE León.

2.2. Gaussian matched filters (GMF)

The Gaussian matched filters method was proposed by Chaudhuri et al. [15] for detecting vessel-like structures in different types of medical images. The GMF method is founded on the fact that the shape of blood vessels in the spatial image domain can be detected by applying a Gaussian curve as matching template, which is defined as following:

$$G(x,y) = -\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad |y| \leq L/2, \tag{1}$$

where L represents the length (in pixels) of the vessel segment to be processed, and σ is the average width of the vessel-like structures. Since the Gaussian curve has infinitely long double sided trails; the trails are commonly truncated at $u = \pm 3\sigma$. A neighborhood N is defined as following:

$$N = \{(u,v), |u| \leq T, |v| \leq L/2\} \tag{2}$$

where T is a discrete parameter (in pixels) used to define the position where the Gaussian curve trails will cut. In order to detect blood vessels at different orientations, the Gaussian kernel $G(x,y)$ is rotated at different angles (θ) using $\kappa = 180/\theta$ oriented filters as in Eq. (3), where κ is the number of evenly spaced filters in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$. These oriented kernels are convolved with the original image, and pixels with the maximum response over all orientations are conserved to generate the final filtered image.

$$\kappa = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}. \tag{3}$$

Moreover, the process of selection of the optimal parameters of the GMF method plays a vital role for each application. To select the highest vessel detection performance of the method, one continuous parameter (σ), and three discrete parameters (L,T,θ) have to be determined.

In Fig. 2, an X-ray coronary angiogram along with its ground-truth image drawn by a specialist are illustrated. Fig. 2(c) presents the Gaussian kernel originally proposed by Chaudhuri et al. [15] as $L = 9$, $T = 13$, $\kappa = 12$, and $\sigma = 2.0$. Fig. 2(d) and (e), introduce the matching templates acquired from the Gaussian kernel oriented at $\theta = 0^\circ$, and $\theta = 45^\circ$, respectively. The final enhanced image obtained from the above mentioned parameters of the GMF method is given in Fig. 2(f).

2.3. Genetic algorithms

Genetic algorithms (GAs) are from the family of evolutionary computation (EC) techniques that emulate genetic evolution. GAs are population-based methods to solve optimization problems that can be considered as stochastic global search strategies [27,28]. GAs use a set of individuals (potential solutions) called population, which is gradually improved until a stopping criterion is satisfied (e.g., number of iterations). For each iteration also called generation, the individuals are evaluated according to an objective function, and by applying the evolutionary operators of selection, crossover, and mutation. The selection operator is used to obtain the best individuals at each generation to form a base for a new improved population. The crossover operator is used to exchange information between two or more selected individuals. Finally, the mutation operator is often applied to introduce randomly information in the population.

According to the above description, GAs can be implemented by using the following procedure:

1. Establish number of generations G, population size Np, crossover rate CR, and mutation rate MR.
2. Initialize the individuals within the search space.
3. Evaluate population in the objective function.
4. For each generation $g = \{1, \ldots, G\}$:
   (a) Select a subset of the best individuals;
   (b) Generate new individuals by applying the crossover and mutation operators;
   (c) Evaluate new individuals in the objective function;
   (d) Update population by replacing worst individuals with the best new individuals.
5. Stop if the convergence criterion is satisfied (e.g., stability or number of generations).

2.4. Estimation of distribution algorithms (EDAs)

EDAs represent a family of population-based methods that incorporate statistical information of potential solutions to solve optimization problems in discrete and continuous domains [29,30]. Similar to EC techniques, EDAs use selection operators, binary or real encoding, and a population of potential solutions to perform the optimization task. The key feature of EDAs resides in the fact that the crossover and mutation operators are not required, since

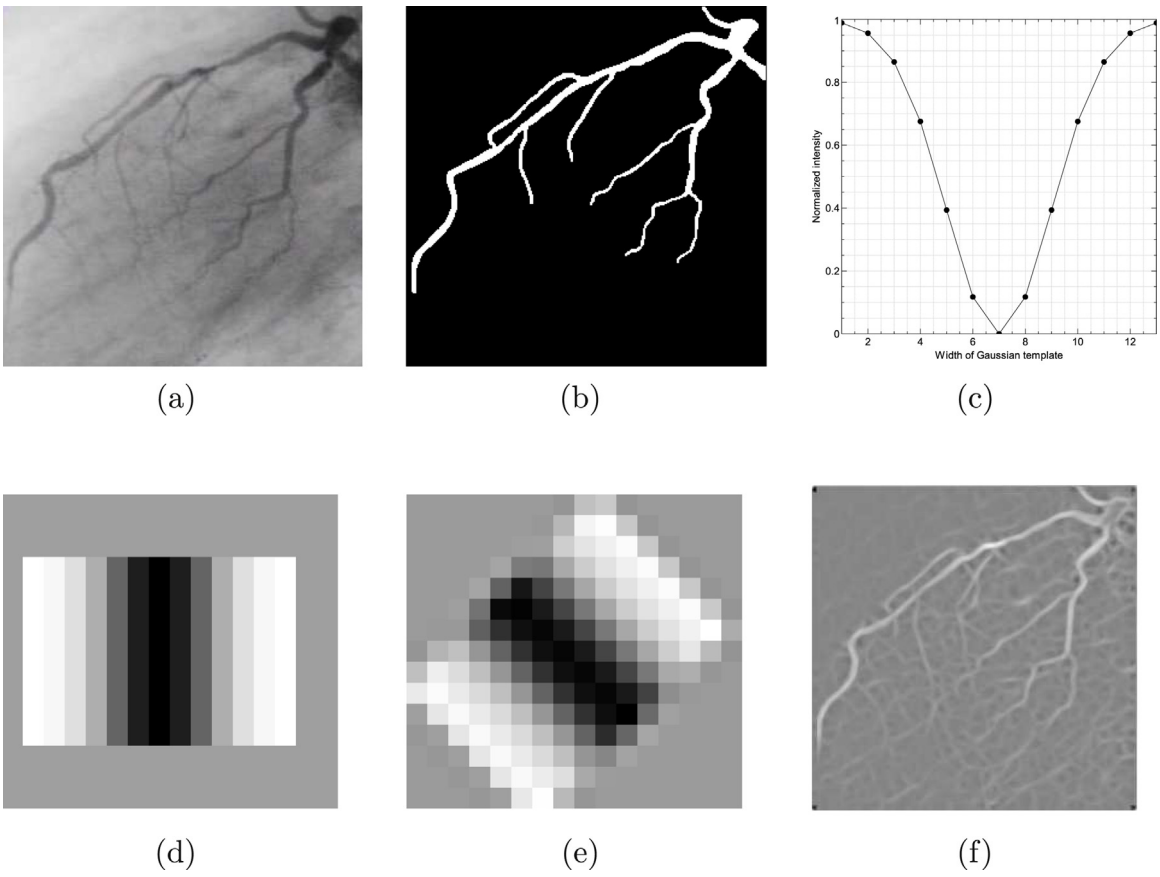

Fig. 2. (a) Original X-ray coronary angiogram. (b) Ground-truth of angiogram in (a). (c) Gaussian kernel proposed by Chaudhuri et al. [15]. (d) Gaussian template with $\theta = 0^{\circ}$. (e) Gaussian template with $\theta = 45^{\circ}$. (f) Resulting enhanced image using the angiogram in (a) and the Gaussian kernel in (c).

the new individuals are created by building a probabilistic model based on global statistical information of promising solutions at each generation. In the present work, the univariate marginal distribution algorithm (UMDA) has been adopted because it works ideally for linear problems with not many significant dependencies [31,32].

UMDA employs a binary codification of the problem, and it generates a probability vector $\mathbf{p} = (p_1, p_2, \ldots, p_n)^T$, where p_i is the marginal probability for the i variable empirically computed, and it is used for simulating new potential solutions for each variable independently. The fundamental idea of UMDA is to approximate the probability distribution of the individuals in $\mathbb{P}_t$ as the product of the univariate frequencies calculated from selected solutions assuming that all the variables are independent [33]. The evolutionary process of UMDA can be summarized in the steps of selection of promising individuals, estimation of probability distribution, and generation of new potential solutions. In the selection step, the individuals within the search space Ω are arranged according to fitness value, then a proportional selection s is computed as follows:

$$\mathbb{P}^s(x) = \frac{\mathbb{P}(x) f(x)}{\sum_{\tilde{x} \in \Omega} \mathbb{P}(\tilde{x}) f(\tilde{x})}. \tag{4}$$

The second step consists in the estimation of the univariate marginal probabilities $\mathbb{P}$. The probability model for independent variables can be stated as follows:

$$\mathbb{P}(x) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i), \tag{5}$$

where $x = (x_1, x_2, \ldots, x_n)^T$ is the binary value of ith bit in the individual, and X_i is the ith random value of the vector X. In the last step of UMDA, new individuals are generated through the estimated probability distribution and evaluated by the fitness function at each generation. Finally, these three steps are performed until a convergence criterion is satisfied.

According to the above description, UMDA can be implemented by the following procedure:

1. Establish number of n individuals and number of generations t.
2. Initialize the individuals within the search space.
3. Select a subset of individuals S of $m \leq n$ according to a selection method.
4. Calculate the univariate marginal probabilities $p_i^s(x_i, t)$ of S.
5. Generate n new individuals by applying $p(x, t+1) = \prod_{i=1}^{n} p_i^s(x_i, t)$.
6. Stop if convergence criterion is satisfied (e.g., stability or number of generations), otherwise, repeat steps (3)–(5).

2.5. Differential evolution

Differential evolution (DE) is a stochastic real-parameter technique introduced by Storn and Price [34,35] for numerical global optimization problems. DE uses a set of Np randomly initialized individuals $X = \{x_1, x_2, \ldots, x_{Np}\}$, which are gradually improved according to an objective function and the evolutionary principles of mutation, crossover and selection. The mutation operator generates a mutant vector $V_{i,g+1}$ at each generation g based on the distribution of the current population $\{X_{i,g} | i = 1, 2, \ldots, Np\}$ by applying the following mutation strategy:

$$V_{i,g+1} = X_{r1,g} + F(X_{r2,g} - X_{r3,g}), \quad r1 \neq r2 \neq r3 \neq i, \tag{6}$$

where $r1, r2$ and $r3$ are the indexes of three individuals from the current population mutually different and uniformly selected from the set {1,. . .,Np}, g indicates the current generation of the evolutionary process, and F is the differentiation factor also known as mutation parameter. The second step corresponding to the crossover operator is used to create a trial vector $U_{i,g+1}$ as follows:

$$U_{i,g+1} = \begin{cases} V_{i,g+1}, & \text{if } r \leq CR \\ X_{i,g}, & \text{if } r > CR \end{cases} \quad (7)$$

where r represents a uniform random value on the interval [0, 1], which is compared with the crossover rate, $CR \in [0, 1]$. The CR is a constant value that controls the rate of parameters values that are copied from the mutant vector, and it also controls the diversity of solutions in the population. If the value of r is bigger than CR, the current individual $X_{i,g}$ is preserved, otherwise, the mutant vector $V_{i,g+1}$ is copied to the trial vector $U_{i,g+1}$. Commonly, a low CR value leads to a slow convergence to the optimal solution and it poses challenges for multi-modal problems. On the other hand, a high CR value provides high diversity to the pool of candidate solutions, which is useful for the present work. The CR value was determined statistically in the range [0, 1] in steps of 0.5, using the training set. Finally, the selection operator uses the best one between the individual $X_{i,g}$ and the trial vector $U_{i,g+1}$ to replace the current individual in the next generation.

$$X_{i,g+1} = \begin{cases} U_{i,g+1}, & \text{if } f(U_{i,g+1}) < f(X_{i,g}) \\ X_{i,g}, & \text{otherwise} \end{cases} \quad (8)$$

According to the above description, the DE algorithm can be implemented by the following procedure:

1. Establish number of generations G, population size Np, differentiation factor F and crossover rate CR.
2. Initialize each individual X_i by generating random solutions within the search space.
3. For each individual $X_{i,g}$, where $g = \{1, \ldots, G\}$:
   (a) Compute $V_{i,g+1}$ using the mutation operator (6);
   (b) Assign $U_{i,g+1}$ applying the crossover operator (7);
   (c) Update $X_{i,g+1}$, if $U_{i,g+1}$ is better than $X_{i,g}$ employing the selection operator (8).
4. Stop if the convergence criterion is satisfied (e.g., stability or number of generations).

2.6. Particle swarm optimization (PSO)

PSO is a computational intelligence technique proposed by Eberhart and Kennedy [36], and improved by Shi and Eberhart [37] to solve numerical optimization problems. PSO consists of a set of potential solutions known as swarm, where each solution also called particle represents a point in an N-dimensional space $X_i = \{x_{i1}, x_{i2}, \ldots, x_{iN}\}$. At each iteration, the particles move through hyperspace to a new position using the following velocity equation:

$$v_i(t+1) = \varphi v_i(t) + \kappa r_1(P_{best} - x_i(t)) + \kappa r_2(P_{gbest} - x_i(t)), \quad (9)$$

where x_i is the particle in the time (t) and v_i is its current velocity, φ represents the inertia weight, κ is the learning factor, and r_1, $r_2 \sim U(0, 1)$ are random values from a uniform distribution. P_{best} represents the best solution found by the current particle, and P_{gbest} is the best solution of the whole swarm. The last step of the PSO model consists in updating the current position of the particle within the search space according to the position equation as follows:

$$x_i(t+1) = x_i(t) + v_i(t+1). \quad (10)$$

Moreover, the PSO algorithm described above can be implemented through the following procedure:

1. Establish number of particles of the swarm, inertia weight and learning factor.
2. Initialize position and velocity of each particle within the search space.
3. Evaluate each particle in objective function to update its P_{best}, if the new fitness is better.
4. Find the best particle in the whole swarm to update the P_{gbest}, if the fitness value is better.
5. Stop if the convergence criterion is satisfied (e.g., stability or number of iterations).
6. Update velocity and position of the whole swarm by using (9) and (10), respectively, then repeat steps 3–5.

2.7. Optimization of Gaussian matched filters

Since the detection performance of the GMF method directly depends of four parameters, many researchers have proposed different values for each parameter. According to the work of Chaudhuri et al. [15], these parameters can be set as $L = 9$, $T = 13$, $\sigma = 2.0$, and $\theta = 15$ acquiring $\kappa = 12$ evenly spaced filters. Kang et al. [18–20] modified the angular resolution of the original work to $\theta = 30^\circ$, obtaining 6 oriented templates, and the average vessel width to $\sigma = 1.5$, in order to be applied on coronary angiograms. Cinsdikici and Aydin [21] modified the angular resolution to $\theta = 10^\circ$ obtaining $\kappa = 18$ oriented templates. This method was applied in a hybrid model with the ant colony algorithm on ophthalmoscope images. Al-Rawi et al. [22] modified the GMF method by extending the range of the variables to $L = \{7, 7.1, \ldots, 11\}$, $T = \{2, 2.25, \ldots, 10\}$, $\sigma = \{1.5, 1.6, \ldots, 3\}$, and keeping constant the number of evenly spaced filters to $\kappa = 12$. In this method, an exhaustive global search was performed over all possible combinations of the extended parameters, and the best detection rate was determined through the area (A_z) under the receiver operating characteristic curve. To avoid the exhaustive global search, Al-Rawi and Karajeh [23] applied the population-based technique of genetic algorithms (GAs) to select the optimal L, T, and σ values in retinal images. The GMF based on GAs obtained superior performance than the previous empirically methods in terms of blood vessel detection, and also it reduced the number of evaluations regarding the exhaustive search method.

Based on the appropriate performance of the evolutionary method, in the present work a comparative analysis of four population-based methods has been performed to select the optimal parameters of the GMF method while reducing the number of evaluations of the objective function. The search space of the GMF method was defined taking into account the above described methods and the dataset of X-ray coronary angiograms. The search space of the parameters was established as $L = \{8, 9, \ldots, 15\}$, $T = \{8, 9, \ldots, 15\}$, and $\sigma = [1, 5]$ with an step size of $\Delta = 0.01$ (25,664 possible solutions). The parameter of number of oriented filters was established as $\kappa = 12$, since similar results were obtained with $\kappa = 15$, 20, 30, 45, and 60 oriented filters. For further analysis, $\kappa = 12$ with angular resolution of $\theta = 15^\circ$ was applied.

To perform the optimization task, the area A_z under the ROC curve has been used as the objective function to be maximized. The receiver operating characteristic (ROC) curve is a plot that represents the true-positive fraction (TPF) against false-positive fraction (FPF). TPF is the rate of vessel pixels correctly detected by the computational method regarding the ground-truth image. FPF is the rate of nonvessel pixels incorrectly classified as vessel pixels by the method. The A_z value is in the range [0, 1], where 1 is perfect detection, and 0 otherwise.

On the other hand, to assess the vessel and nonvessel pixel classification, the measures of sensitivity, specificity, and accuracy have been adopted.

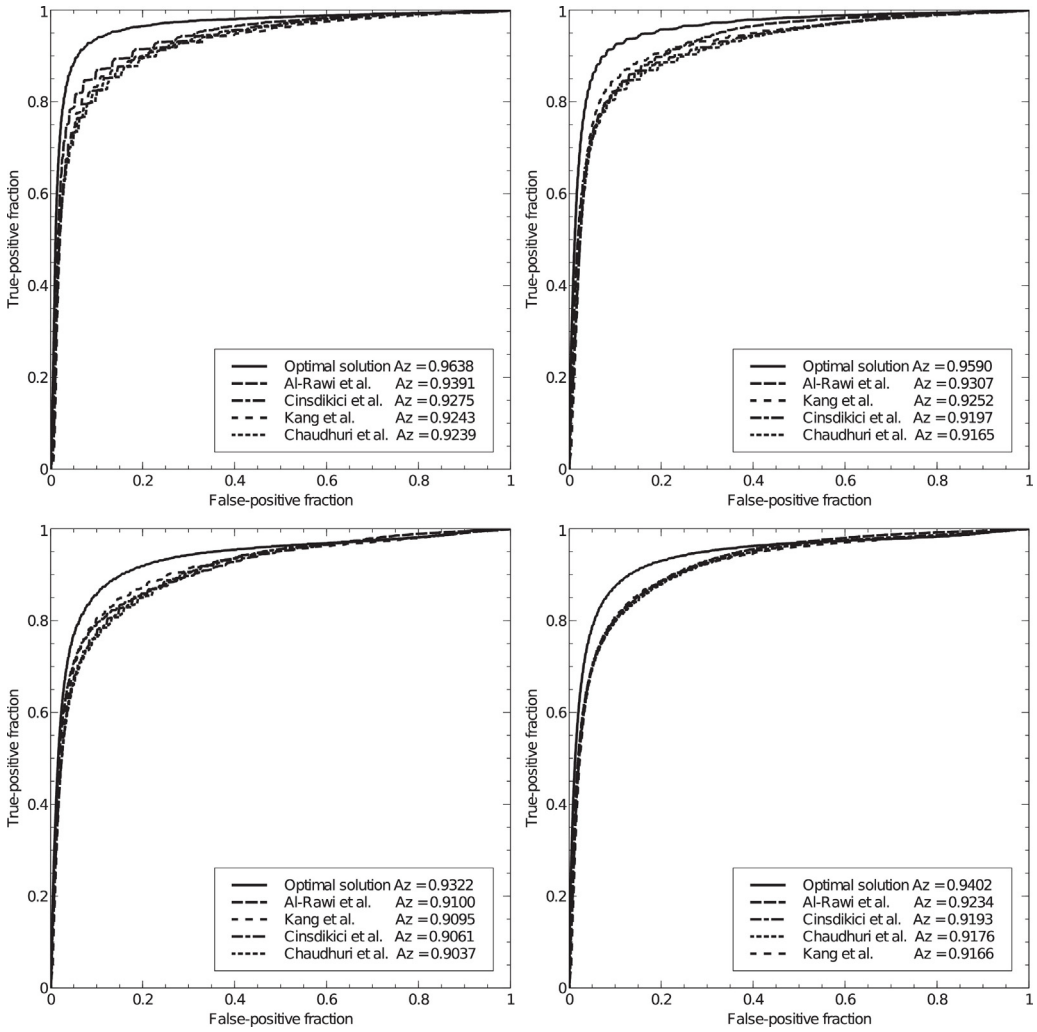

Fig. 3. Comparison of ROC curves for blood vessel detection: first row, from left to right, subsets of 4, and 8 angiograms respectively. Second row, from left to right, subsets of 20, and 40 angiograms respectively.

The sensitivity measure, also called (TPF), is defined as the fraction of vessel pixels hand-labeled by the specialist that are correctly detected by the method as follows:

$$Sensitivity = \frac{TP}{TP + FN}, \tag{11}$$

where TP is the fraction of vessel pixels correctly classified as such by the method, and FN represents the fraction of vessel pixels incorrectly classified by the method.

The specificity measure is defined as the fraction of nonvessel (background) pixels that are correctly classified as such by the method, and can be computed as follows:

$$Specificity = \frac{TN}{TN + FP}, \tag{12}$$

where TN is defined as the fraction of nonvessel pixels correctly classified as such by the method, and FP represents the fraction of nonvessel pixels that are incorrectly classified.

The accuracy measure indicates the fraction of correctly classified pixels, i.e., the sum of TP and TN divided by the number of pixels in the image, as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}. \tag{13}$$

The measures of sensitivity, specificity, and accuracy are in the range [0, 1], where 1 is perfect classification, and 0 otherwise. These measures are used to evaluate the vessel classification performance according to the vessels hand-labeled by the specialist.

In Section 3, the segmentation results obtained from the proposed method using X-ray angiographic images are presented and analyzed by the evaluation metrics.

3. Results and discussion

In this section, we evaluate the performance of the proposed segmentation method, and it is compared with different state-of-the-art vessel detection and segmentation methods. The computational experiments are performed on a computer with an AMD A10-5800B, 8GB of RAM, 3.8 GHz processor, using the Matlab software version 2012a.

As part of the analysis of vessel detection techniques via GMF, in Fig. 3 the results of four state-of-the-art methods are compared with the optimal solution (the highest detection rate obtained by applying the GMF method) using the area (A_z) under the ROC curve. The comparative study is carried out over four different subsets with 4, 8, 20, and 40 images from the training set of angiograms, respectively. The optimal solution for each subset was acquired by varying the values of the GMF parameters in an exhaustive search. In this performance analysis, the method of Chaudhuri et al. [15] presents, in general, the lowest vessel detection rate. The methods of Kang et al. [20], and Cinsdikici and Aydin [21], present similar

Table 1
Statistical analysis of the vessel detection performance using the area A_z over 30 runs of the GA, UMDA, DE, and PSO methods. (a) Subset of 4 angiograms, (b) subset of 8 angiograms, (c) subset of 20 angiograms, and (d) whole training set of 40 angiograms.

Training set	Measure	GA	UMDA	DE	PSO
(a)	Maximum	0.9637	0.9638	0.9638	0.9638
	Minimum	0.9110	0.9221	0.9391	0.9275
	Median	0.9424	0.9391	0.9586	0.9499
	Mean	0.9429	0.9396	0.9567	0.9488
	Std.Dev.	0.0155	0.0141	0.0080	0.0116
(b)	Maximum	0.9590	0.9512	0.9590	0.9590
	Minimum	0.9032	0.9148	0.9357	0.9165
	Median	0.9280	0.9252	0.9482	0.9308
	Mean	0.9314	0.9280	0.9489	0.9361
	Std.Dev.	0.0142	0.0120	0.0084	0.0156
(c)	Maximum	0.9322	0.9322	0.9322	0.9322
	Minimum	0.9012	0.9037	0.9100	0.9012
	Median	0.9095	0.9098	0.9275	0.9202
	Mean	0.9119	0.9119	0.9269	0.9186
	Std.Dev.	0.0097	0.0088	0.0060	0.0113
(d)	Maximum	0.9402	0.9387	0.9402	0.9402
	Minimum	0.9112	0.9142	0.9193	0.9112
	Median	0.9234	0.9218	0.9352	0.9201
	Mean	0.9248	0.9264	0.9314	0.9275
	Std.Dev.	0.0092	0.0097	0.0090	0.0112

results for detecting vessel-like structures overcoming the detection performance of the original GMF method. In addition, the results of these two methods have also shown that the parameter of number of oriented filters can be predefined to $\kappa = 12$, preserving relevant information and reducing the search space of the GMF parameters. The exhaustive method of Al-Rawi et al. [22] obtains the best detection rate in the comparative tests, which is achieved by extending the range of the variables L, T, and σ.

In order to illustrate the vessel detection results of the comparative analysis, in Fig. 4, the Gaussian filter response of the aforementioned four methods over a subset of training angiograms is introduced. By visual inspection, it can be observed that the exhaustive method and the optimal solution present a higher contrast between vessel pixels and background image, and a better discrimination of false-positive pixels than the previous methods.

Since the GMF-based methods described above present low detection rate compared with the optimal solution, the computational intelligence techniques of GA, UMDA, DE, and PSO, are analyzed to avoid an exhaustive search and to maximize the vessel detection performance of the GMF method. From the EC techniques, UMDA is applied with a selection rate = 0.6, and the genetic algorithm uses a crossover fraction = 0.7, mutation rate = 0.3, elite = 1, with an heuristic multi-point as crossover method, as in the work of Al-Rawi et al. [23]. In these two techniques, a 15 bits binary encoding for each individual has been applied, where 3 bits are used for L parameter, 3 bits for T parameter, and the remaining 9 bits for the σ parameter. The real-parameter DE method was tuned as differentiation factor = 0.9 and crossover rate = 0.9. The parameters of the PSO method were established as inertia weight = 0.5, and learning factor = 0.9. In the experiments, the four EC techniques have been applied with a population size = 30, and two stopping criteria; convergence to optimal solution, otherwise, number of iterations = 50.

In Table 1, a statistical analysis of the vessel detection performance obtained by the EC techniques using the area A_z under the ROC curve is introduced. The experiment was performed with 30 runs over four training subsets with 4, 8, 20, and 40 angiograms, respectively. In this analysis, the measures of mean and standard deviation show that DE is more stable and robust to work with the GMF method than the comparative techniques.

Table 2
Comparison of DE algorithm with considered methods using Student's t-test for number of evaluations. (a) Subset of 4 angiograms, (b) subset of 8 angiograms, (c) subset of 20 angiograms, and (d) whole training set of 40 angiograms.

Training set	GA	UMDA	PSO
(a)	4.129+	3.282+	1.586+
(b)	2.331+	1.856+	1.073+
(c)	5.833+	3.883+	1.146+
(d)	6.820+	6.346+	0.627+

In addition, Fig. 5 presents a comparative analysis in terms of number of evaluations over the four different subsets from the training set of angiograms. The illustrated results suggest that the EC technique of Differential Evolution presents the best performance in terms of stability and dispersion, obtaining in most of the tests, the lowest mean and number of required iterations to converge to the optimal solution. Moreover, the best running time per angiogram in the training set was acquired by differential evolution with 2.01 min, followed by PSO, UMDA, and GA with 2.42 min, 4.23 min, and 4.68 min, respectively. The executing time of the exhaustive search was of 25.6 min. Additionally, since all the methods are implemented in Matlab and no optimization process is performed, the computational time could be significantly improved.

Based on the results of Fig. 5, and with the aim of ensuring a significant difference between the DE algorithm and the considered methods, in Table 2, a statistical comparison is performed for the number of evaluations using Student's t-test. In this experiment, a confidence level of 0.95 (t-critical = 0.063) has been applied to prove the null hypothesis that there is no difference in the number of evaluations for 30 runs of the methods. In the table, t-test values are presented, and the symbol "+" indicates that there are significant differences between the DE method and the comparative technique, i.e., null hypothesis is rejected.

In order to segment vessel structures from the background image, different automatic thresholding strategies can be applied. In the present work, the vessel segmentation obtained by the proposed method is a soft classification, where a single threshold value is used to classify vessel-like structures and produce a binary segmentation for each image. This threshold value (t) is computed by using the ROC curve of the training set of angiograms. Since the procedure to compute the ROC curve is performed by applying a sliding threshold to the gray-scale Gaussian filter response, the true-positive fraction (TPF) and false-positive fraction (FPF) are acquired by comparing the different binary results with the corresponding ground-truth angiogram. The threshold $t = 0.54$ is the value where the best trade-off between TPF and FPF was achieved in the training stage. Hence, the threshold value $t = 0.54$ is used to segment all the angiograms in the test set.

Table 3
Comparative analysis of the proposed GMF-DE method with six state-of-the-art vessel segmentation methods using the test set of 40 angiograms.

Segmentation method	Sensitivity	Specificity	Accuracy	Average
Kang et al. [18]	0.5113	0.9863	0.9565	0.8180
Qian et al. [6]	0.8355	0.8521	0.8511	0.8462
Wang et al. [12]	0.6217	0.9753	0.9532	0.8501
Chanwimaluang and Fan [17]	0.8647	0.8721	0.8717	0.8695
Al-Rawi et al. [22]	0.8492	0.8818	0.8798	0.8703
Kang et al. [20]	0.8359	0.9110	0.9063	0.8844
Proposed GMF-DE	0.8247	0.9620	0.9534	0.9134

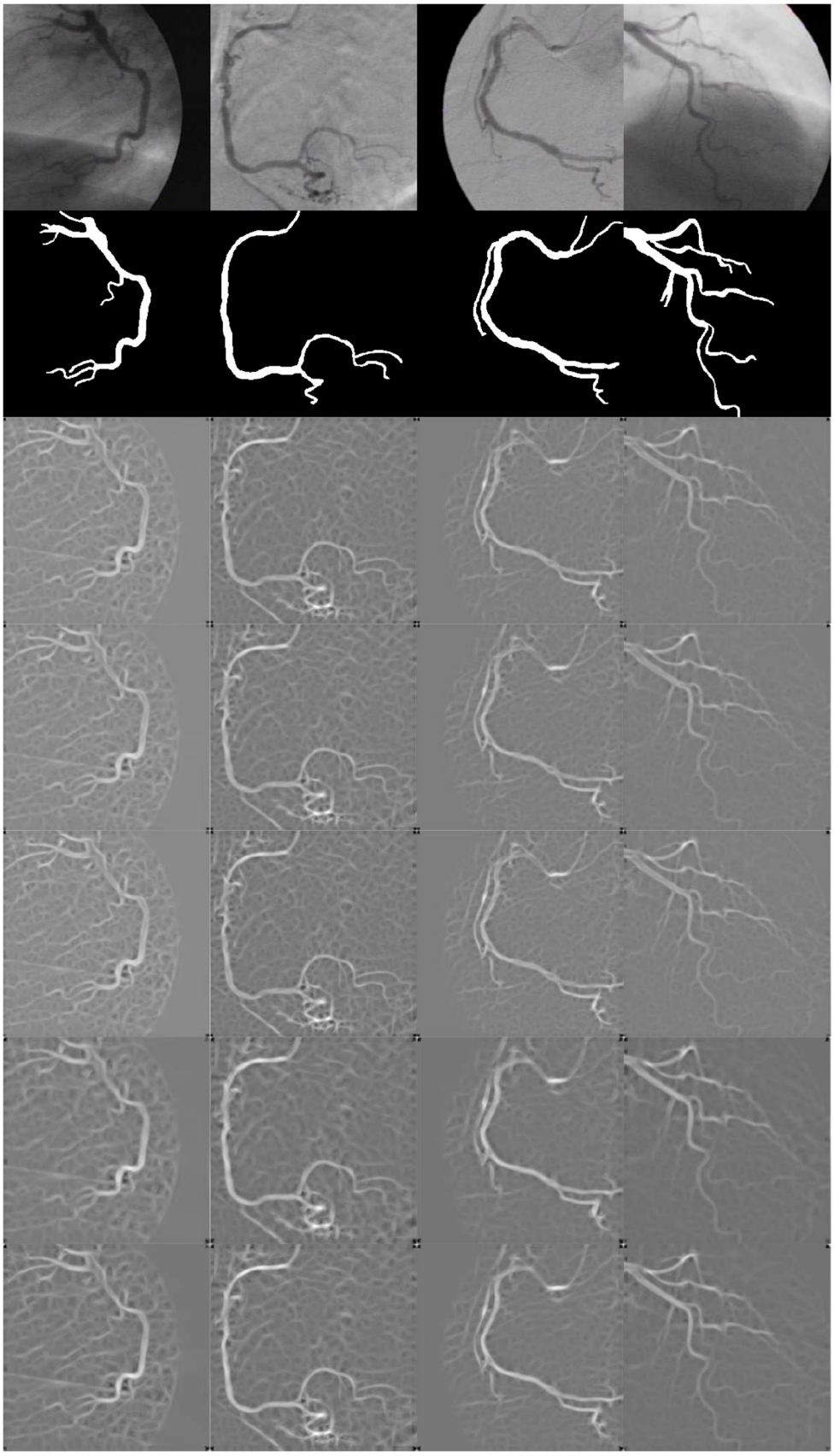

Fig. 4. First row: subset of angiograms from the training set. Second row: ground-truth of the images in first row. The remaining five rows present the Gaussian filter response of the methods of Chaudhuri et al. [15], Cinsdikici and Aydin [21], Kang et al. [20], Al-Rawi et al. [22], and the optimal solution obtained from the set of 40 angiograms, respectively.

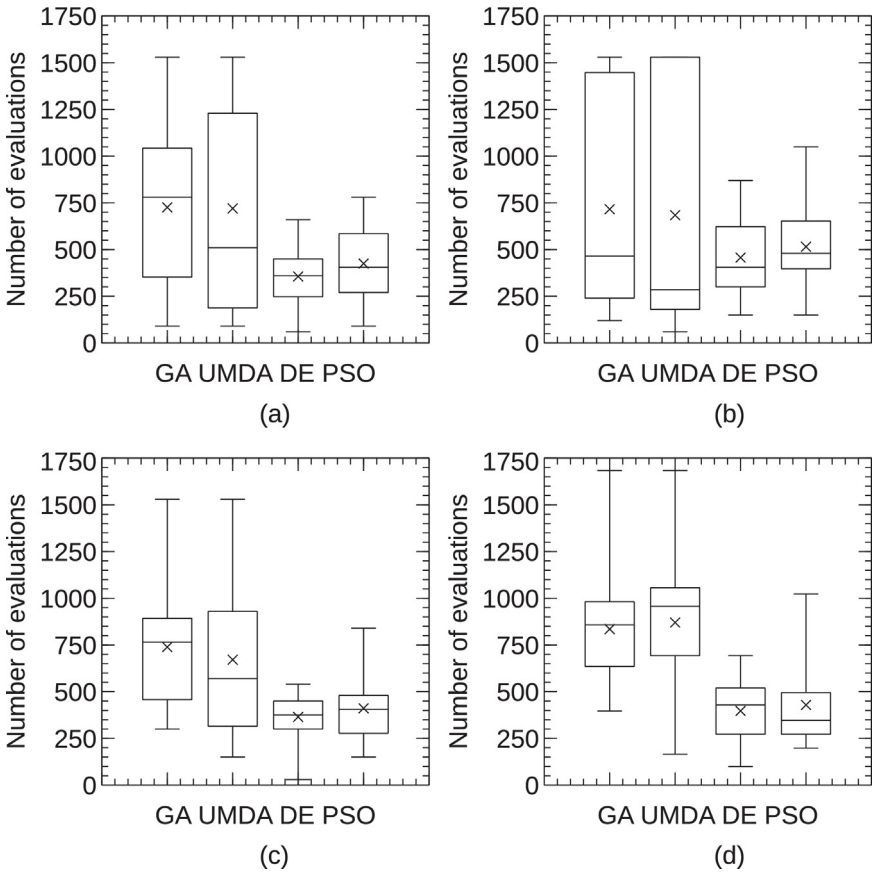

Fig. 5. Dispersion and symmetry of the number of evaluations over 30 runs of the GA, UMDA, DE, and PSO methods. (a) Subset of 4 angiograms, (b) subset of 8 angiograms, (c) subset of 20 angiograms, and (d) whole training set of 40 angiograms.

In Table 3, the results of the GMF based on DE algorithm are compared with six vessel segmentation methods of the state-of-the-art using the test set of 40 angiograms. The best GMF parameters found by DE over the whole training set of angiograms were established as $L = 13$, $T = 15$, and $\sigma = 2.82$. In these segmentation results, the average of the sensitivity, specificity, and accuracy measures is used as the main metric for evaluation, since the specificity and accuracy values are always high due to the vessel pixels often occupy less than 14% of an angiogram. The values of these measures were acquired by concatenating the test set of angiograms to form just one large binary image. The obtained results show that the proposed GMF-DE method provides the highest trade-off in terms of vessel segmentation using the test set.

Additionally, the training and execution times of the proposed and comparative methods are introduced in Tables 4 and 5, respectively. In the training stage, the GMF-based methods of Kang et al. [18,20] and Chanwimaluang and Fan [17] use empirically or statistically determined values for the four parameters (L, T, σ, κ) of the method. The method of Al-Rawi et al. [22] is computationally more expensive as compared to the other methods, because it is based on an exhaustive global search of the parameters in a predefined search space. The proposed method avoids empirically determined values and also an exhaustive global search by introducing the nature inspired algorithm of differential evolution and the optimization process discussed above. On the other hand, after the training stage, the segmentation step is carried out by using the obtained values for the GMF parameters. The proposed method shows competitive results in terms of computational time in both training stage and execution time over the test set of coronary angiograms. Although the proposed method is computationally more expensive as compared to other methods, it provides the best trade-off in segmentation accuracy. It is important to point out that the method of Chanwimaluang and Fan [17] obtains the best computational time because the GMF parameters are based on the statistically determined values of Chaudhuri et al. [15]. Due

Table 4
Performance of the segmentation methods in terms of training time (in seconds) using the training set of 40 angiograms. The symbol "–" represents that the method does not use a training stage applying GMF or other method, and the symbol "*" represents that the method uses empirically or statistically determined values.

Segmentation method	Training time (GMF)	Training time (other)
Kang et al. [18]	*	62.23 s
Qian et al. [6]	–	1016.27 s
Wang et al. [12]	–	63.78 s
Chanwimaluang and Fan [17]	*	–
Al-Rawi et al. [22]	1124.4 s	–
Kang et al. [20]	*	62.23 s
Proposed GMF-DE	120.6 s	–

Table 5
Performance of the segmentation methods in terms of executing time per image in the test set of angiograms.

Segmentation method	Execution time (s)
Kang et al. [18]	4.70
Qian et al. [6]	2.36
Wang et al. [12]	6.25
Chanwimaluang and Fan [17]	1.88
Al-Rawi et al. [22]	3.03
Kang et al. [20]	4.68
Proposed GMF-DE	3.23

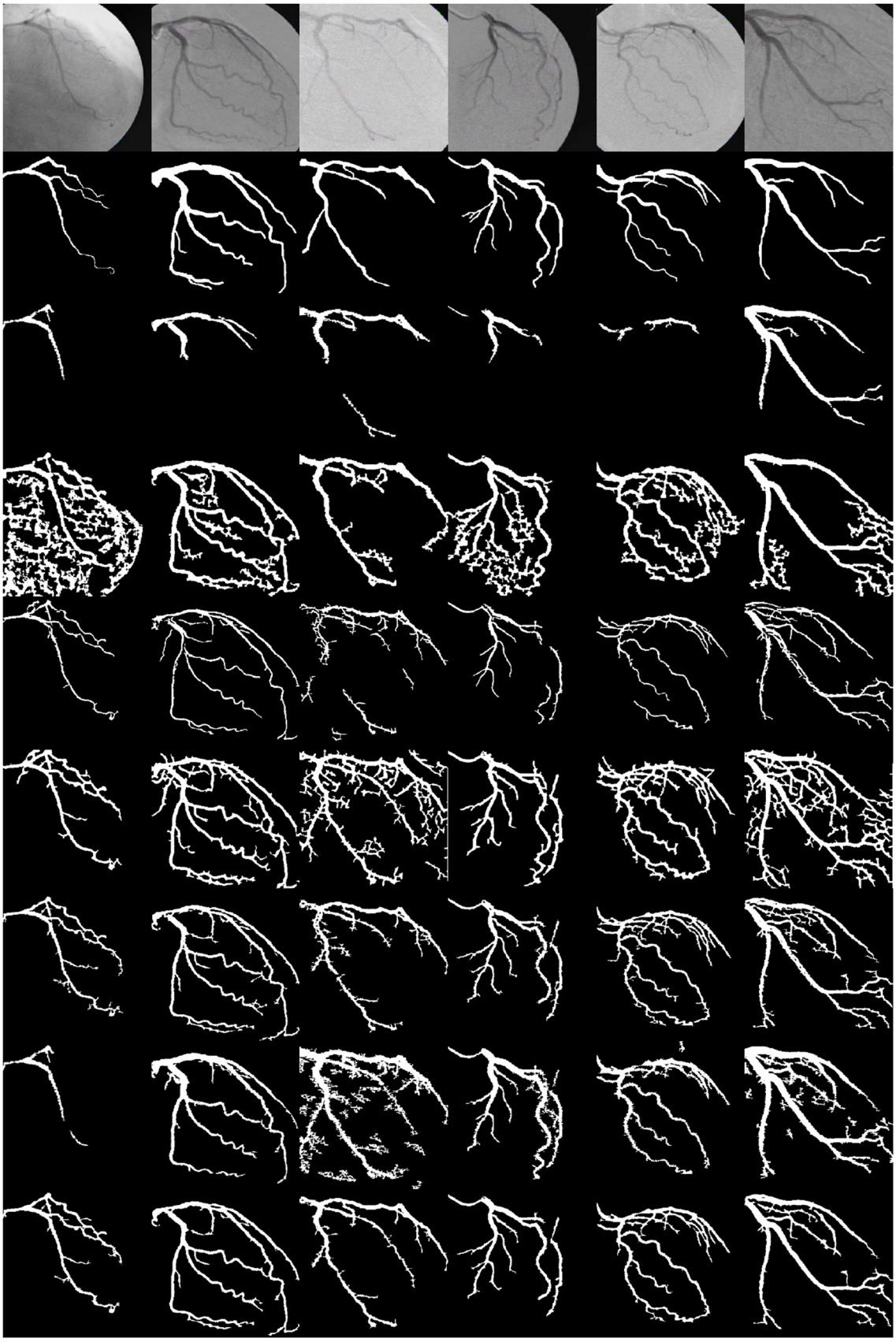

Fig. 6. First row: subset of original angiographic images from the test set. Second row: ground-truth of the images of the first row. The remaining seven rows present the segmentation results of the methods of Kang et al. [18], Qian et al. [6], Wang et al. [12], Chanwimaluang and Fan [17], Al-Rawi et al. [22], Kang et al. [20], and the proposed GMF-DE method, respectively, applied on the angiograms of the first row.

to this, the method obtains a high rate of false-positive pixels, which decreases the specificity and accuracy obtaining a low trade-off between the evaluation measures. In the methods of Kang et al. [18] and Kang et al. [20] a training stage to compute the GMF parameters is not required since they were empirically determined. In Kang et al. [18] the problem of low sensitivity is acquired where a high number of vessel pixels were not detected. The method of Kang et al. [20] also presents a high number of false-positive pixels which poses problems to obtain an appropriate trade-off in segmentation accuracy. Consequently, in order to obtain the best performance of these GMF-based methods, the L, T, σ, and κ parameters have to be tuned properly to be used on specific medical applications such as detection of coronary stenosis.

Fig. 6 illustrates a subset of X-ray coronary angiograms from the test set along with the corresponding vessel ground-truth image. The segmentation results of the method of Kang et al. [18] present broken vessel-like structures and low rate of true-positive pixels, which leads to low values in terms of the sensitivity measure. The method of Qian et al. [6] shows a high rate of false-positive pixels, which leads to a poor performance in terms of specificity and accuracy measures. Although the method of Wang et al. [12] illustrates vascular structures at different calibers, many broken vessels are acquired, which decrease the segmentation results in terms of the sensitivity measure. The method of Chanwimaluang and Fan [16] presents a high rate of true-positive and false-positive pixels, which increases the sensitivity while decreasing the specificity and accuracy measures. The method of Al-Rawi et al. [22] based on an exhaustive search shows an appropriate results detecting vessels of different diameters and high rate of true-positive pixels. The method of Kang et al. [20] presents low performance for most of the angiographic images having nonuniform illumination and high rate of false-positive pixels, leading to a poor results in terms of specificity and accuracy. Furthermore, the segmentation results obtained from the GMF-DE proposed method show the highest average value over the test set of 40 angiograms. The method presents a suitable rate of true-positive pixels and high values in terms of specificity and accuracy involving vessels with different diameters while obtaining a low rate of broken vessels and false-positive pixels.

The considered state-of-the-art vessel segmentation methods provide suitable results based on the evaluation metrics over the test set of angiograms. However, the comparative analysis reveals that the performance of the proposed GMF-DE method is robust in images with nonuniform illumination acquiring an appropriate vessel segmentation rate and providing the highest average performance with the analyzed metrics.

4. Conclusion

In this paper a comparative analysis of four state-of-the-art nature inspired algorithms to improve the training stage of a segmentation strategy based on GMF for X-ray angiograms has been introduced. The statistical results in terms of number of evaluations show a superior performance of differential evolution than the considered algorithms. The process of parameter selection through DE, in general achieved the best detection rate of the GMF method according to the area under ROC curve with the training set of 40 angiograms. According to the experimental results, the proposed GMF-DE strategy has obtained the highest average performance rate of 0.9134 using the test set of 40 angiograms compared with six state-of-the-art vessel segmentation methods. In addition, considering the hand-labeled angiograms by specialist and the vessel segmentation results obtained from the GMF-DE strategy, it can be highly appropriate for X-ray image applications in cardiology.

Acknowledgments

This work has been supported by the Mexican National Council on Science and Technology (Cátedras-CONACYT No. 3150-3097). The authors thank Dr. F.J. Solórzano-Zepeda with the cardiology department of the Mexican Social Security Institute, UMAE León, for his valuable clinical advice and collaboration providing us the sources of the coronary angiograms.

References

[1] N. Sang, Q. Tang, X. Liu, W. Weng, Multiscale centerline extraction of angiogram vessels using Gabor filters, in: Computational and Information Science, Springer LNCS 3314, 2004, pp. 570–575.
[2] Z. Shoujun, Y. Jian, W. Yongtian, C. Wufan, Automatic segmentation of coronary angiograms based on fuzzy inferring and probabilistic tracking, Biomed. Eng. Online 9 (2010) 21.
[3] S.W. Franklin, S.E. Rajan, Retinal vessel segmentation employing ANN technique by Gabor and moment invariants-based features, Appl. Soft Comput. 22 (2014) 94–100.
[4] A. Fathi, A.R. Naghsh-Nilchi, Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation, Biomed. Signal Process. Control 8 (2013) 71–80.
[5] S. Eiho, Y. Qian, Detection of coronary artery tree using morphological operator, Comput. Cardiol. 24 (1997) 525–528.
[6] Y. Qian, S. Eiho, N. Sugimoto, M. Fujita, Automatic extraction of coronary artery tree on coronary angiograms by morphological operators, Comput. Cardiol. 25 (1998) 765–768.
[7] K. Sun, N. Sang, Morphological enhancement of vascular angiogram with multiscale detected by Gabor filters, Electron. Lett. 44 (2008).
[8] B. Bouraoui, C. Ronse, J. Baruthio, N. Passat, P. Germain, Fully automatic 3D segmentation of coronary arteries based on mathematical morphology, in: 5th IEEE International Symposium on Biomedical Imaging (ISBI): From Nano to Macro, 2008, pp. 1059–1062.
[9] D. Lara, A. Faria, A. Araujo, D. Menotti, A semi-automatic method for segmentation of the coronary artery tree from angiography, in: XXII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI), 2009, pp. 194–201.
[10] A. Frangi, W. Niessen, K. Vincken, M. Viergever, Multiscale vessel enhancement filtering, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI'98), Springer LNCS 1496, 1998, pp. 130–137.
[11] N. Salem, A. Nandi, Unsupervised segmentation of retinal blood vessels using a single parameter vesselness measure, in: Sixth Indian Conference on Computer Vision, Graphics and Image Processing, vol. 34, IEEE, 2008, pp. 528–534.
[12] S. Wang, B. Li, S. Zhou, A segmentation method of coronary angiograms based on multi-scale filtering and region-growing, in: International Conference on Biomedical Engineering and Biotechnology 2012, 2012, pp. 678–681.
[13] T. Tsai, H. Lee, M. Chen, Adaptive segmentation of vessels from coronary angiograms using multi-scale filtering, in: International Conference on Signal-Image Technology and Internet-Based Systems, 2013, pp. 143–147.
[14] F. M'hiri, L. Duong, C. Desrosiers, M. Cheriet, Vesselwalker: coronary arteries segmentation using random walks and Hessian-based vesselness filter, in: IEEE 10th International Symposium on Biomedical Imaging (ISBI): From Nano to Macro, 2013, pp. 918–921.
[15] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, M. Goldbaum, Detection of blood vessels in retinal images using two-dimensional matched filters, IEEE Trans. Med. Imaging 8 (1989) 263–269.
[16] T. Chanwimaluang, G. Fan, An efficient blood vessel detection algorithm for retinal images using local entropy thresholding, in: Proc. IEEE International Symposium on Circuits and Systems, vol. 5, 2003, pp. 21–24.
[17] T. Chanwimaluang, G. Fan, S. Fransen, Hybrid retinal image registration, IEEE Trans. Inf. Technol. Biomed. 10 (2006) 129–142.
[18] W. Kang, K. Wang, W. Chen, W. Kang, Segmentation method based on fusion algorithm for coronary angiograms, in: 2nd International Congress on Image and Signal Processing (CISP), 2009, pp. 1–4.
[19] W. Kang, W. Kang, W. Chen, B. Liu, W. Wu, Segmentation method of degree-based transition region extraction for coronary angiograms, in: 2nd International Conference on Advanced Computer Control, 2010, pp. 466–470.
[20] W. Kang, W. Kang, Y. Li, Q. Wang, The segmentation method of degree-based fusion algorithm for coronary angiograms, in: 2nd International Conference on Measurement, Information and Control, 2013, pp. 696–699.
[21] M. Cinsdikici, D. Aydin, Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm, Comput. Methods Progr. Biomed. 96 (2009) 85–95.
[22] M. Al-Rawi, M. Qutaishat, M. Arrar, An improved matched filter for blood vessel detection of digital retinal images, Comput. Biol. Med. 37 (2007) 262–267.
[23] M. Al-Rawi, H. Karajeh, Genetic algorithm matched filter optimization for automated detection of blood vessels from digital retinal images, Comput. Methods Progr. Biomed. 87 (2007) 248–253.
[24] D. Penas, J. Banga, P. Gonzalez, R. Doallo, Enhanced parallel differential evolution algorithm for problems in computational systems biology, Appl. Soft Comput. 33 (2015) 86–99.

[25] I. Cruz-Aceves, J. Avina-Cervantes, J. Lopez-Hernandez, et al., Automatic image segmentation using active contours with univariate marginal distribution, Math. Probl. Eng. 419018 (2013) 9.
[26] K. Chen, K. Wang, K. Wang, M. Angelia, Applying particle swarm optimization-based decision tree classifier for cancer classification on gene expression data, Appl. Soft Comput. 24 (2014) 773–780.
[27] D. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison Wesley, New York, 1989.
[28] M. Mitchell, An Introduction to Genetic Algorithms, The MIT Press, Cambridge, MA, 1997.
[29] P. Larrañaga, J. Lozano, Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation, Kluwer Academic, Boston, MA, USA, 2002.
[30] M. Hauschild, M. Pelikan, An introduction and survey of estimation of distribution algorithms, Swarm Evol. Comput. 1 (2011) 111–128.
[31] S. Bashir, M. Naeem, S. Shah, A comparative study of heuristic algorithms: GA and UMDA in spatially multiplexed communication systems, Eng. Appl. Artif. Intell. 23 (2010) 95–101.
[32] H. Muehlenbein, The equation for response to selection and its use of prediction, Evol. Comput. 5 (1997) 303–346.
[33] L. Lozada-Chang, R. Santana, Univariate marginal distribution algorithm dynamics for a class of parametric functions with unitation constraints, Inf. Sci. 181 (2011) 2340–2355.
[34] R. Storn, K. Price, Differential Evolution – A simple and efficient adaptive scheme for global optimization over continuous spaces, Technical Report TR-95-012, International Computer Sciences Institute, Berkeley, CA, USA, 1995.
[35] R. Storn, K. Price, Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces, J. Glob. Optim. 11 (1997) 341–359.
[36] R. Eberhart, J. Kennedy, New optimizer using particle swarm theory, in: Proceedings of the 6th International Symposium on Micro Machine and Human Science, 1995, pp. 39–43.
[37] Y. Shi, R. Eberhart, Modified particle swarm optimizer, in: Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC'98), 1998, pp. 69–73.