



# CSA-DE/EDA: a Novel Bio-inspired Algorithm for Function Optimization and Segmentation of Brain MR Images

Zhe Li<sup>1</sup> · Yong Xia<sup>1,2</sup> · Hichem Sahli<sup>2,3,4</sup>

Received: 2 October 2018 / Accepted: 3 July 2019 / Published online: 7 August 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

The clonal selection algorithm (CSA), which describes the basic features of an immune response to an antigenic stimulus, has drawn a lot of attention in the biologically inspired computing community, due to its highly adaptive and easy-to-implement nature. Despite many successful applications, CSA still suffers from limited ability to explore the solution space. In this paper, we incorporate the differential evolution (DE) algorithm and the estimation of distribution algorithm (EDA) into CSA, and thus propose a novel bio-inspired algorithm referred to as CSA-DE/EDA. In the proposed algorithm, the hypermutation and receptor editing processes are implemented based on DE and EDA, which provide improved local and global search ability, respectively. We have applied the proposed algorithm to five commonly used benchmark functions for optimization and brain magnetic resonance (MR) image segmentation. Our comparative experimental results show that the proposed CSA-DE/EDA algorithm outperforms several bio-inspired computing techniques. CSA-DE/EDA is a compelling bio-inspired algorithm for optimization tasks.

**Keywords** Bio-inspired computing · Clonal selection algorithm (CSA) · Differential evolution (DE) · Estimation of distribution algorithm (EDA) · Image segmentation

## Introduction

Optimization problems arise in almost every field of engineering [1]. In spite of their wide-spread applications, numerical optimization algorithms, such as gradient descent, pose many constraints on objective functions, including convexity, continuity, derivability, and unimodality, which, unfortunately, are not always satisfied in most real-world problems [1]. In recent decades, biologically inspired computing techniques [2] have

witnessed a lot of attention, from the information science community, for their capacity to accommodate the way nature uses powerful computational techniques, such as natural selection [3] and clonal selection [4], to obtain highly suitable solutions. Biologically inspired computing techniques are used to solve a variety of single and multiobjective problems, are highly adaptive and easy-to-implement, and pose less constraints on the objective functions [5]. Among such techniques, evolutionary algorithms and swarm intelligence [6–8] have attracted tremendous research attention.

The genetic algorithm (GA) [9], as one of the most prevalent evolutionary algorithms (EAs), uses heuristics-guided search that simulates the process of natural selection and survival of the fittest, and generates the next population of solutions by performing a combination of genetic operators, including selection, crossover, and mutation, on the current population, which enables GA to adapt to environmental changes. Although GA has the potential to search a global optimal solution, it often falls into local optima due to the limited runtime [5]. To overcome such problem, researchers devised many improved EAs. Among others, the estimation of distribution algorithms (EDAs) [10] and the differential evolution (DE) [5] are two of the most well-known solutions.

✉ Yong Xia  
yxia@nwpu.edu.cn

<sup>1</sup> National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China

<sup>2</sup> Research & Development Institute, Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China

<sup>3</sup> Audio Visual Signal Processing (AVSP), Department of Electronics & Informatics (ETRO), Vrije Universiteit Brussel (VUB), VUB-ETRO, Pleinlaan, 2, B-1050 Brussels, Belgium

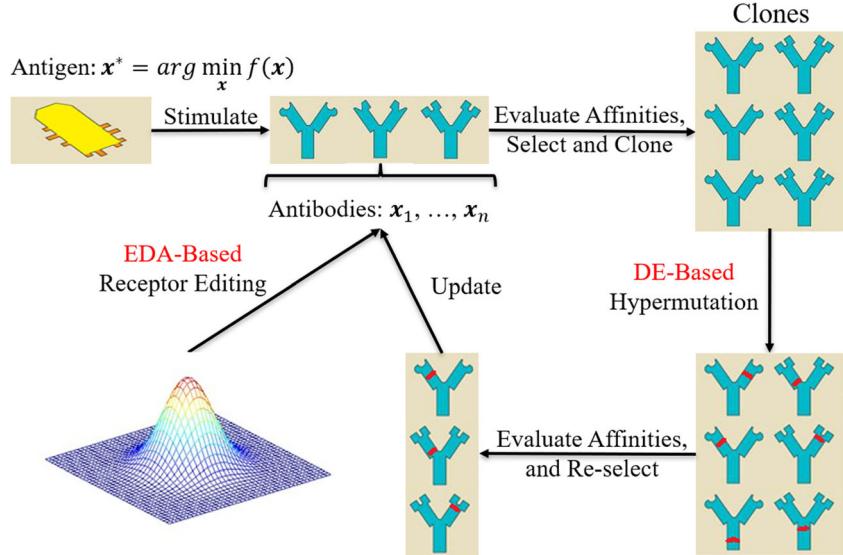
<sup>4</sup> Interuniversity Microelectronics Center, Kapeldreef 75, 3001 Leuven, Belgium

**Table 1** Five DE mutation strategies, denoted as DE/x/y, with “x” the mutated vector, “y” denotes the number of difference vectors,  $r_1, r_2, r_3, r_4, r_5$  are mutually exclusive and random values in  $[1, N_p]$ ,  $X_{best}^G$  the best individual in the current population, and  $F$  is a user specified positive scaling factor that controls the scaling of the difference vectors

Strategy notation	Operation
DE/rand/1	$V_i^G = X_{r_1}^G + F \cdot (X_{r_2}^G - X_{r_3}^G)$
DE/best/1	$V_i^G = X_{best}^G + F \cdot (X_{r_1}^G - X_{r_2}^G)$
DE/rand-to-best/1	$V_i^G = X_i^G + F \cdot (X_{best}^G - X_i^G) + F \cdot (X_{r_1}^G - X_{r_2}^G)$
DE/best/2	$V_i^G = X_{best}^G + F \cdot (X_{r_1}^G - X_{r_2}^G) + F \cdot (X_{r_3}^G - X_{r_4}^G)$
DE/rand/2	$V_i^G = X_{r_1}^G + F \cdot (X_{r_2}^G - X_{r_3}^G) + F \cdot (X_{r_4}^G - X_{r_5}^G)$

EDAs, also known as probabilistic model-building genetic algorithms, replace the crossover and mutation operators with learning and sampling from the solution distribution to generate offspring. The advantages of EDAs include (i) the expressiveness and transparency of the probabilistic model that guide the search process (i.e., global search), (ii) the absence of multiple parameters to be tuned, (iii) a compact representation, and (iv) the ability to avoid premature convergence [11]. EDAs have been proven to be better suited for some applications than GAs, while achieving competitive and robust results in the majority of the tackled problems [12]. DE generates new individuals called trial individuals by perturbing the current individuals with a scaled difference between two randomly selected individuals. DE employs differential information for efficient local search [13]. DE has been proven, in benchmarks [14] and a variety of real applications, to produce more accurate results than several other optimization methods including GA, simulated annealing, and evolutionary programming [15].

**Fig. 1** The proposed CSA-DE/EDA algorithm



The artificial immune system (AIS) [16–20], a well-known paradigm in swarm intelligence [8], has also drawn considerable research attention. Many AIS algorithms have been designed to solve multimodal functional optimization problems by mimicking the behavior of living organisms in protecting themselves against antigens. Inspired by the clonal selection principle of acquired immunity that explains how B and T lymphocytes improve their response to antigens over time [21], De Castro and Von Zuben [21, 22] developed the clonal selection algorithm (CSA). CSA has shown superior performance compared to several other bionic algorithms [4] and traditional optimization mechanisms in a variety of applications [21, 23]. CSA has been designed to simulate the affinity maturation process based on the clonal selection theory, which claims that only those cells that recognize the antigens will be selected to proliferate, and the proliferated cells will improve their affinity to the antigens through an affinity maturation process [4, 21, 24]. Despite the success and prevalence of CSA, it can be further improved [25, 26]. The Baldwinian clonal selection algorithm (BCSA) proposed by Gong et al. [25] is a famous version of hybrid CSA, which incorporates the Baldwinian learning operator into CSA to improve its global search ability.

The problem of most optimization algorithms is that they usually get stuck in local minima (or maxima) [13, 25, 26]. Hence, they should use both local information around the current solutions and the global information about the search space [13]. The former is of great importance for local search (exploitation of the surrounding space), and the latter can guide the search for new promising areas (broader exploration of the search space) [21]. Although hypermutation and receptor editing in traditional CSA perform local and global search, respectively, their search efficiency is not satisfactory [22]. Considering the excellent local search performance of DE algorithms and the unique global search strategy of the

**Table 2** The CSA-DE/EDA algorithm

Line	Pseudo code of CSA-DE/EDA
1	<b>Initialization:</b> Randomly generate N antibodies $\mathbf{Ab}^0$ ( $k = 0$ )
2	<b>while</b> the stopping criterion is not met
3	Evaluate the vector $\mathbf{f}^k$
4	Select $\mathbf{Ab}_{\{n\}}^k$ ( <b>Section 3.1</b> )
5	Select $\mathbf{Ab}_{\{M\}}^k$ ( <b>Section 3.4</b> )
6	Clone $\mathbf{Ab}_{\{n\}}^k \rightarrow \mathbf{C}^k$ ( <b>Section 3.1</b> )
7	DE-based hypermutation $\mathbf{C}^k \rightarrow \mathbf{C}^{*k}$ ( <b>Section 3.2</b> )
8	Reselect from $\mathbf{C}^{*k}$ to obtain $\mathbf{Ab}^{k+1}$ ( <b>Section 3.3</b> )
9	EDA-based receptor editing to replace the $r * N$ lowest affinity from $\mathbf{Ab}^{k+1}$ ( <b>Section 3.4</b> )
10	<b>end</b>

EDAs algorithms [12, 27], we suggest using them to replace the hypermutation and receptor editing in traditional CSA or other local and global search operator in improved CSAs. In this work, we combine the strength of two famous and important bio-inspired algorithms, namely, evolutionary algorithms and swarm intelligence. We use the DE and EDA principles to perform hypermutation and receptor editing in CSA, respectively, and propose a novel bio-inspired algorithm, referred to as CSA-DE/EDA, for global optimization problems. We evaluate the proposed bio-inspired algorithm on five commonly used test functions for optimization, and compare its performance against four bio-inspired computing techniques, including DE, EDA, simple combination of the DE and EDA (DE/EDA) [13], and CSA. Furthermore, the proposed CSA-DE/EDA is also employed for brain magnetic resonance (MR) image segmentation of gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF).

**Table 3** Parameter settings of CSA-D E/EDA

Function	$D$	Parameter settings				
		$N$	$F$	$CR$	$\alpha$	$r$
G. Ros.	5	30	0.6	1.0	1	0.1
	10	40	0.6	0.9	1	0.1
M. Lan.	5	100	0.2	0.3	5	0.1
	10	300	0.4	0.8	10	0.1
S. Fox.	5	150	0.3	0.3	2	0.1
	10	300	-0.5	0.9	5	0.1
E. Mic.	5	40	0.5	0.3	4	0.1
	10	40	0.5	0.1	4	0.1
T. Pol.	9	90	0.5	1.0	2	0.1
	17	170	0.5	1.0	2	0.05

$D$  function dimension,  $N$  population size,  $F$  scaling factor,  $CR$  crossover rate,  $\alpha$  clonal multiplying factor, and  $r$  replacement ratio

## Related Work

### Differential Evolution Algorithm

DE [28] maintains a population of  $N_p$  individuals in every generation, each is a potential solution to the optimization problem. DE evolves and improves the population iteratively. At each iteration, a new population is generated by mutation, crossover, and combination operations. The mutation operation extracts the distance and direction information from the current individuals [13] and adds random deviation for diversity. The five most frequently used DE mutation strategies [29] are listed in Table 1, where  $\{\mathbf{X}_i^G : i = 1, 2, \dots, N_p\}$  denote the  $G$ th generated population and  $\mathbf{V}_i^G$  the  $i$ th mutant vector.

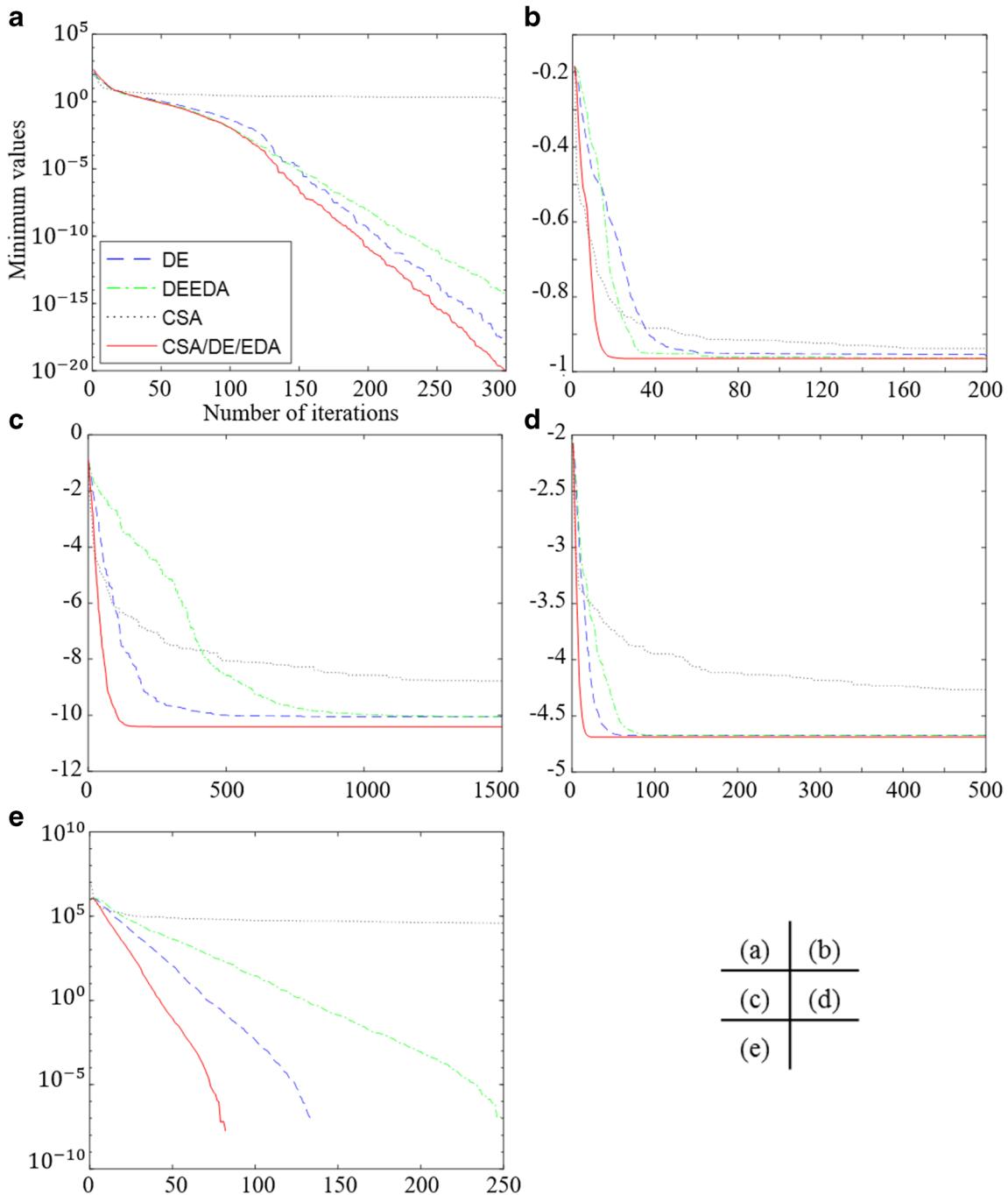
After mutation, the following crossover operation is applied to each obtained mutant vector  $\mathbf{V}_i^G$  and its parent  $\mathbf{X}_i^G$  to generate a trial vector

$$U_{i,j}^G = \begin{cases} \mathbf{V}_{i,j}^G, & \text{if } (\text{rand} < CR) \text{ or } (j = j_{\text{rand}}) \\ \mathbf{X}_{i,j}^G, & \text{otherwise} \end{cases}, j = 1, \dots, D \quad (1)$$

where  $\text{rand}$  is a uniform distributed random value within the range  $[0, 1]$ ,  $j_{\text{rand}}$  is a random integer within the range  $[1, D]$ , and  $CR$  is a user specified crossover rate. In the combination operation, if a trial vector  $\mathbf{U}_i^G$  has an improved fitness, usually defined as the corresponding objective function value, its parent  $\mathbf{X}_i^G$  will be replaced in the new generation. This evolution process continues until a stopping criterion is met (e.g., the value of the objective function is smaller than a given threshold, or after a certain number of generations). It results in a faster convergence speed [14].

### Estimation of Distribution Algorithms

EDAs also search for the optimal solution through iteratively evolving a population of solutions. In contrast to DE, EDAs use global information instead of local information to produce the new generation. They select a set of parents to build a distribution model of promising solutions. To form the next population, they replace fully or partly the current population with new solutions sampled from the probabilistic model. There are three methods used to select parent individuals, namely the proportional selection, truncation selection, and tournament selection. For global optimization problems, the probabilistic model  $p(\mathbf{X}^G)$  can be a Gaussian distribution, a Gaussian mixture, a histogram, or a Gaussian model with diagonal covariance matrix (GM/DCM). The probability density parameters are estimated from parent individuals using maximum likelihood estimation. Depending on the used probabilistic model, several EDAs have been proposed in the literature [13, 30–32].



**Fig. 2** Evolution procedure for the consider benchmark functions in their five-dimensional form. **a** G. Ros, **b** M. Lan, **c** S. Fox, **d** E. Mic, **e** T. Pol

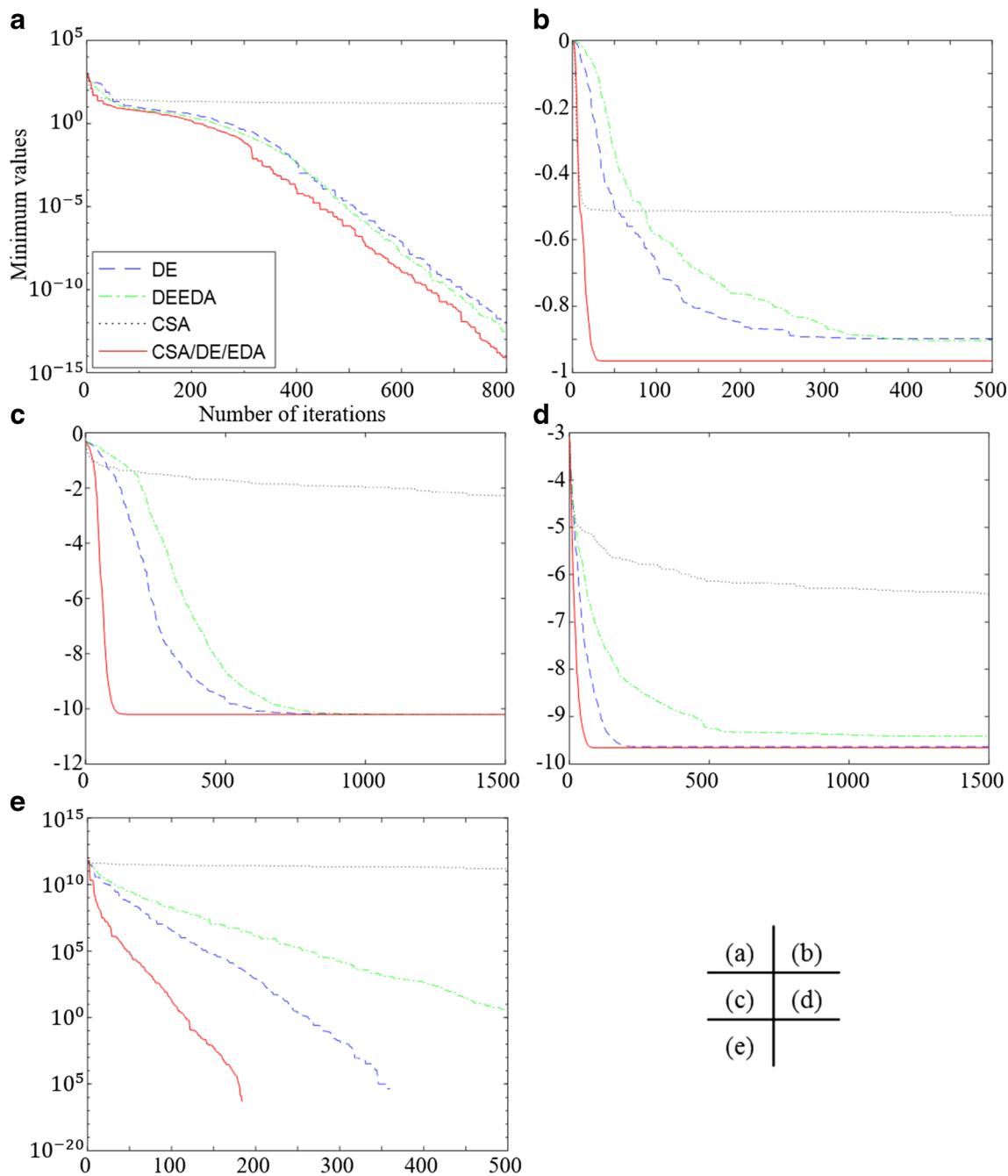
## Proposed CSA-DE/EDA Algorithm

Like the CSA, the proposed CSA-DE/EDA algorithm uses the clonal selection principle to solve the following optimization problem

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in R^D} f(\mathbf{x}). \quad (2)$$

where each admissible solution  $\mathbf{x}$  is encoded as an antibody  $\mathbf{Ab}_i^k = [Ab_{i,1}^k, Ab_{i,2}^k, \dots, Ab_{i,D}^k] \in R^D$ , and the objective function  $f(\cdot)$

is defined as the adaptive immune response, namely the affinity, to the corresponding antigen. Solving this optimization problem is equivalent to searching the antibody that has the maximum affinity [21]. To this end, CSA-DE/EDA, see Fig. 1, evolves (at  $k$ th generation) a population of  $N$  antibodies, denoted as  $\mathbf{Ab}^k = \{\mathbf{Ab}_1^k, \mathbf{Ab}_2^k, \dots, \mathbf{Ab}_N^k\}$ . The population can be randomly initialized and updated on a generation-by-generation basis using five operators: selection, clone, DE-based hypermutation, reselection, and EDA-based receptor editing [21, 33].



**Fig. 3** Evolution procedure for the considered benchmark functions in their 10-dimensional form. **a** G. Ros, **b** M. Lan, **c** S. Fox, **d** E. Mic, **e** T. Pol

### Selection and Clone

For the  $k$ th generation of antibodies  $\mathbf{Ab}^k = \{\mathbf{Ab}_1^k, \mathbf{Ab}_2^k, \dots, \mathbf{Ab}_N^k\}$ , the corresponding affinities,  $f^k = \{f(\mathbf{Ab}_1^k), f(\mathbf{Ab}_2^k), \dots, f(\mathbf{Ab}_N^k)\}$ , are first evaluated, then we select  $n$  highest-affinity antibodies, denoted as  $\{\mathbf{Ab}_{\{1\}}^k, \mathbf{Ab}_{\{2\}}^k, \dots, \mathbf{Ab}_{\{n\}}^k\}$ , and clone each selected antibody  $n_i$  times. The number of antibodies in the clone set  $\mathbf{C}^k$  is

$$N_c = \sum_{i=1}^n n_i = \alpha * n, \quad (3)$$

with  $\alpha$  the clonal multiplying factor that takes a positive integer and controls the cloning number [21]. In this paper,  $n_i$  is the same for all selected antibodies and equals to  $\alpha$ , to locate multiple optima and reach the global optimum [21].

### Hypermutation

The DE-based hypermutation is applied to each selected antibody  $\mathbf{Ab}_i^k$  and its clones in  $\mathbf{C}^k$  in three steps. First, a temporary antibody is generated as follows

**Table 4** Performance of five optimization algorithms for the considered benchmark functions in their five-dimensional form. Italic indicates the best performance

		G. Ros. 5	M. Lan. 5	S. Fox. 5	E. Mic. 5	T. Pol. 5
CSA	VTR	1.0e–6	N/A	N/A	–4.687	1.0e–7
	<i>f</i>	1.173 <sup>+</sup>	–0.953 <sup>+</sup>	–8.683 <sup>+</sup>	–4.364 <sup>+</sup>	2.30e+4 <sup>+</sup>
	BVAT	0.387	–0.962	–10.17	–4.560	3.20e+3
DE	ENES	N/A	N/A	N/A	N/A	N/A
	<i>f</i>	VTR $\approx$	–0.963 $\approx$	–10.20 <sup>+</sup>	–4.67 <sup>+</sup>	VTR $\approx$
	BVAT	0.0	–0.965	–10.40	–4.826	4.3e–29
EDA	ENES	4561	N/A	N/A	5339	17325
	<i>f</i>	2.442 <sup>+</sup>	–0.899 <sup>+</sup>	–6.342 <sup>+</sup>	–3.745 <sup>+</sup>	5.24e+2 <sup>+</sup>
	BVAT	1.392	–0.947	–8.439	–4.628	2.69e+2
DE/EDA	ENES	N/A	N/A	N/A	N/A	N/A
	<i>f</i>	VTR $\approx$	–0.964 $\approx$	–10.40 $\approx$	–4.680 $\approx$	VTR $\approx$
	BVAT	0.0	–0.965	–10.40	–4.687	0.0
CSA-DE/EDA	ENES	4554	N/A	N/A	2894	14,009
	<i>f</i>	VTR	–0.964	–10.40	–4.687	VTR
	BVAT	0.0	–0.965	–10.40	–4.687	0.0
ENES		3985	N/A	N/A	4132	13,456

$$\mathbf{Z} = \frac{(\mathbf{Ab}_d^k + \mathbf{Ab}_i^k)}{2} + F \cdot [(\mathbf{Ab}_d^k - \mathbf{Ab}_i^k) + (\mathbf{Ab}_b^k - \mathbf{Ab}_c^k)], \quad (4)$$

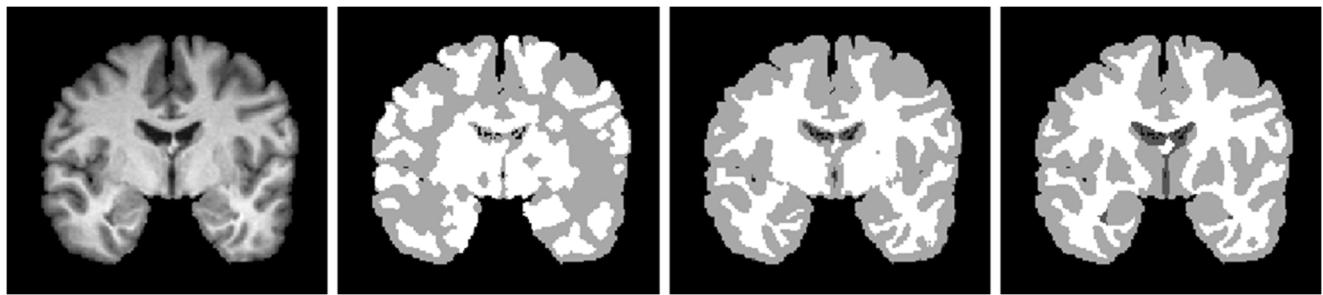
with  $\mathbf{Ab}_d^k$  randomly selected from  $\mathcal{C}^k$  such that  $f(\mathbf{Ab}_d^k) \leq f(\mathbf{Ab}_i^k)$ ,  $\mathbf{Ab}_b^k$  and  $\mathbf{Ab}_c^k$  are also randomly selected from  $\mathcal{C}^k$ , and  $F$  a user specified scaling factor. Second, the

following crossover is applied to each temporary antibody  $\mathbf{Z}$  and its parent  $\mathbf{Ab}_i^k$  to generate a trial antibody  $\mathbf{Ab}_i^{*k}$

$$\mathbf{Ab}_{i,j}^{*k} = \begin{cases} \mathbf{Z}_j & \text{if } (\text{rand} < CR) \\ \mathbf{Ab}_{i,j}^k & \text{otherwise} \end{cases}, j = 1, \dots, D, \quad (5)$$

**Table 5** Performance of five optimization algorithms for the considered benchmark functions in their 10-dimensional form

		G. Ros. 10	M. Lan. 10	S. Fox. 10	E. Mic. 10	T. Pol. 10
CSA	VTR	1.0e–6	N/A	N/A	–9.660	1.0e–7
	<i>f</i>	15.590 <sup>+</sup>	–0.587 <sup>+</sup>	–2.218 <sup>+</sup>	–6.876 <sup>+</sup>	1.69e+11 <sup>+</sup>
	BVAT	9.139	–0.784	–2.655	–7.645	1.32e+10
DE	ENES	N/A	N/A	N/A	N/A	N/A
	<i>f</i>	VTR $\approx$	–0.834 <sup>+</sup>	–10.20 <sup>+</sup>	–9.440 <sup>+</sup>	VTR $\approx$
	BVAT	0.0	–0.965	–10.28	–9.660	5.7e–29
EDA	ENES	27,169	N/A	N/A	77,365	78937
	<i>f</i>	8.168 <sup>+</sup>	–0.180 <sup>+</sup>	–0.900 <sup>+</sup>	–7.166 <sup>+</sup>	3.39e+6 <sup>+</sup>
	BVAT	7.755	–0.617	–1.446	–9.184	1.69e+6
DE/EDA	ENES	N/A	N/A	N/A	N/A	N/A
	<i>f</i>	VTR $\approx$	–0.902 <sup>+</sup>	–10.26 $\approx$	–9.450 <sup>+</sup>	VTR $\approx$
	BVAT	0.0	–0.965	–10.28	–9.660	0.0
CSA-DE/EDA	ENES	22,709	N/A	N/A	57,892	71,293
	<i>f</i>	VTR	–0.965	–10.28	–9.659	VTR
	BVAT	0.0	–0.965	–10.28	–9.660	0.0
ENES		20,121	N/A	N/A	16,133	70,638



**Fig. 4** An example coronal slice from the image “IBSR\_09” (left), and the corresponding results of region-based segmentation (middle left), region- and pixel-based segmentation (middle right), and ground truth (right)

with  $rand$  a uniformly distributed random value within  $[0, 1]$  and  $CR$  a user-specified crossover rate. At the combination step, the obtained trial antibody  $\mathbf{Ab}_i^{*k}$  replaces  $\mathbf{Ab}_i^k$  in  $\mathbf{C}^k$  if it has a higher affinity than the original antibody  $\mathbf{Ab}_i^k$ ; otherwise,  $\mathbf{Ab}_i^k$  is kept. After the hypermutation,  $\mathbf{C}^k$  is termed  $\mathbf{C}^{*k}$ . Since CSA-DE/EDA uses the arithmetic combination, it can capture the local information in the current population for more efficient exploitation [5].

### Reselection

For each antibody  $\mathbf{Ab}_i^k$ , there are  $n_i$  cloned and hypermutated copies, which form a subset of antibodies. To keep the size of antibody population unchanged, we reselect the best antibody in each subset and thus form a trial population [33].

### Receptor Editing

The EDA-based receptor editing is applied to the trial population. To explore the global information, we assume that the probabilistic distribution of the elite antibody is Gaussian and each component of the antibody is mutually independent [32].

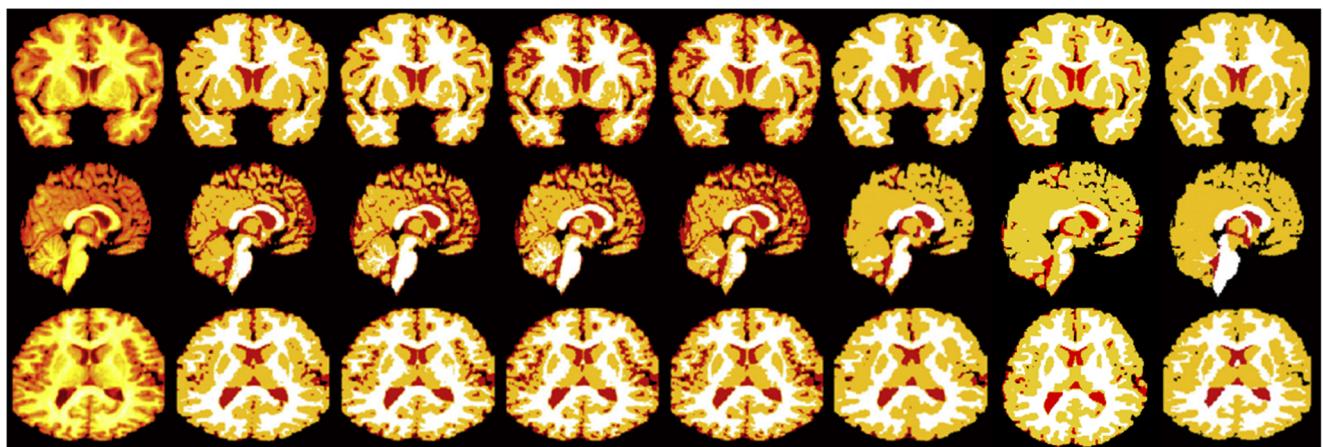
Thus, we have

$$p(\mathbf{Ab}^*) = \prod_{d=1}^D p(Ab_d^*) = \prod_{d=1}^D N(Ab_d^*; \mu_d, \sigma_d), \quad (6)$$

with  $\mathbf{Ab}^* = [Ab_1^*, Ab_2^*, \dots, Ab_D^*]$  the globally optimal antibody, and  $N(\cdot; \mu, \sigma)$  a univariate Gaussian distribution with mean  $\mu$  and standard deviation  $\sigma$ . To estimate the distribution  $p(\mathbf{Ab}^*)$ , we select  $M$  highest-affinity antibodies,  $\mathbf{Ab}_{\{M\}}^k$ , from the current generation using the truncation selection and apply the maximum likelihood estimation [13] to obtain the estimated Gaussian parameters  $\{\mu_d^k, \sigma_d^k; d = 1, 2, \dots, D\}$ . Then, we sample  $r \cdot N$  antibodies from the distribution  $\prod_{d=1}^D N(Ab_d^*; \mu_d^k, \sigma_d^k)$  and use them to replace  $r \cdot N$  lowest-affinity antibodies in the trial population. Usually, the number of selected antibodies  $M$  is set to  $N/2$  [13].

### CSA-DE/EDA Algorithm Overview

The major steps of the proposed CSA-DE/EDA algorithm are summarized in Table 2.



**Fig. 5** Segmentation results. Coronal, sagittal, and transverse slices of image “IBSR\_14” (1st column), and segmentation results obtained by SPM (2nd column), FSL (3rd column), D-C (4th column), GA-GMM (5th column), HMRF-CSA (6th column), HMRF-CSA/DE/EDA (7th column), and the ground truth (8th column) (modified from [51])

## Application to MR Image Segmentation

Due to its excellent optimization capability, the proposed CSA-DE/EDA algorithm can be applied to many image analysis problems [34–37], such as the segmentation of brain MR images to delineate the GM, WM, and CSF.

Let an observed brain MR image be an instance of the image random field  $\mathbf{Y}$ , and the corresponding segmentation result be an instance of the label random field  $\mathbf{X}$ . According to the hidden Markov random field (HMRF) model [38], this segmentation task can be converted into the following optimization problem.

$$\mathbf{X}^* = \arg \max_{\mathbf{X}} p(\mathbf{X}|\mathbf{Y}, \mathbf{b}, \boldsymbol{\theta}^*), \quad (7)$$

$$\boldsymbol{\theta}^* = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{b}, \mathbf{X}^*), \quad (8)$$

where  $\mathbf{b}$  is the bias field,  $\boldsymbol{\theta} = \{\theta_k | k=1, 2, \dots, K\}$  the ensemble of model parameters,  $K$  the number of classes ( $K=3$  in our case), and

$$\begin{aligned} p(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{b}, \mathbf{X}^*) &\propto p(\mathbf{Y}|\mathbf{X}^*, \mathbf{b}, \boldsymbol{\theta})p(\boldsymbol{\theta}) \\ &= \prod_j \sum_k [N(Y_j - b_j | X_j = k; \theta_k) p(\theta_k)], \end{aligned} \quad (9)$$

with  $N(x; \theta_k)$  a Gaussian distribution with parameter  $\theta_k = (u_k, \sigma_k)$ . Starting from a random initialization, this problem can be solved using a three-step iterative process: (1) apply the method in [39] to update the bias field  $\mathbf{b}$ , (2) then use the CSA-DE/EDA algorithm to solve the sub-problem given in Eq. (10) and update the model parameters (i.e.,  $\boldsymbol{\theta}$  being the antibody), and finally (3) use the iterated conditional mode (ICM) approach [40] to update the segmentation result.

In this work we jointly use region- and pixel-based HMRF methods. Indeed, region-based HMRF method [41] is more robust to noise and artifacts than pixel-based HMRF; however, it may produce a too smooth and relatively holistic view of the image to be segmented, hence we combine it with pixel-based HMRF. Considering an input brain MR image, we first apply the super pixel algorithm-TurboPixel [42] to generate an image segmentation into  $m$  regions  $\mathbf{R} = \{R_i | i=1, 2, \dots, m\}$ . Then we define a region adjacency graph (RAG) [43] composed of nodes, representing the regions  $R_i$  of  $\mathbf{R}$ , and edges representing their adjacency. To each node we associate the mean gray value of the pixels within the region. This coarse segmentation allows estimating the model parameters, which will be then used for a final pixel-based segmentation. The overall process is summarized in Algorithm 1.

**Algorithm 1:** HMRF-CSA/DE/EDA MR segmentation algorithm

**Input:** observed brain MR image  $\mathbf{Y}$   
**Output:** optimal voxel class labels  $\mathbf{X}$  and model parameters  $\boldsymbol{\theta}$ .

**Initialization:** segmentation by k-means, random model parameters and initial bias field  $\mathbf{b}$

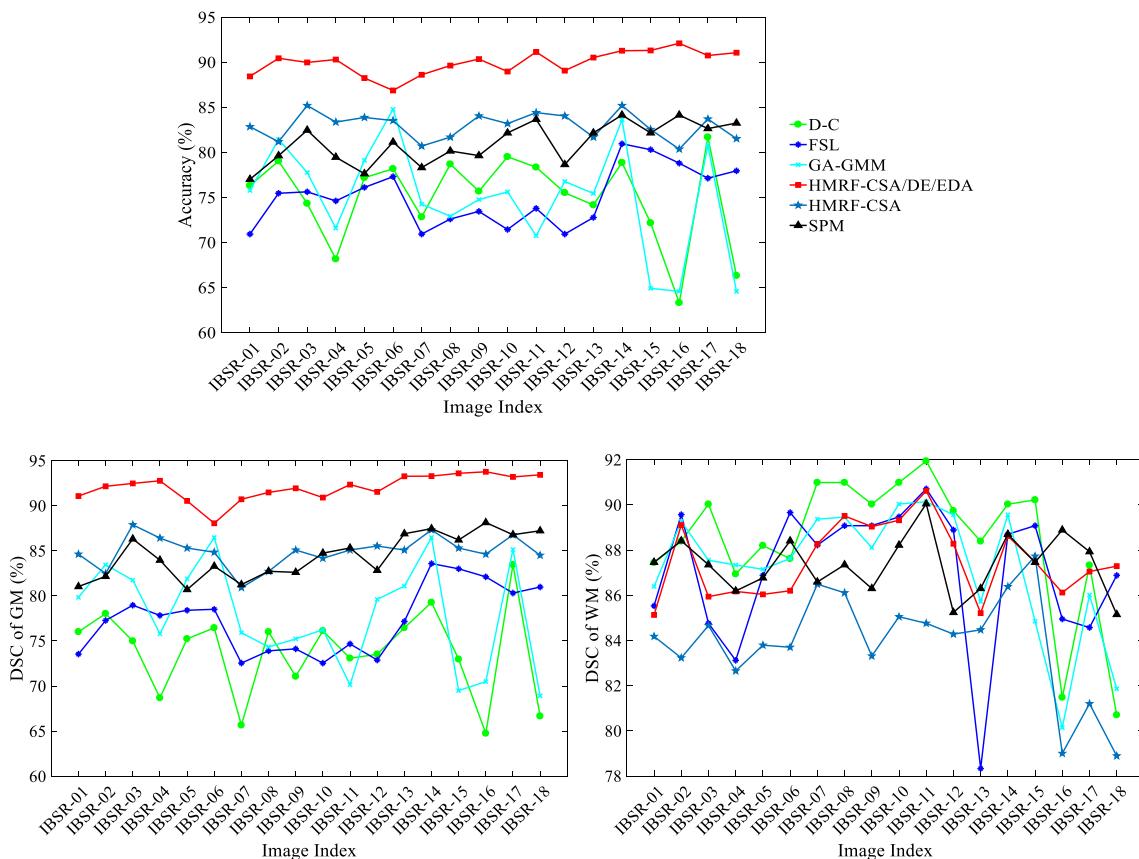
- 1: **while** the stopping criterion is not met (region-based)
- 2:     Evaluate the vector  $\mathbf{f}^k$  through Eq. (11)
- 3:     Select  $\mathbf{Ab}_{(n)}^k$
- 4:     Select  $\mathbf{Ab}_{(M)}^k$
- 5:     Clone  $\mathbf{Ab}_{(n)}^k \rightarrow \mathbf{C}^k$
- 6:     Hypermutation using DE  $\mathbf{C}^k \rightarrow \mathbf{C}^{*k}$
- 7:     Reselect from  $\mathbf{C}^{*k}$  to partly replace  $\mathbf{Ab}^k$
- 8:      $\mathbf{Ab}_{(d)}^k$  replace the d lowest affinity from  $\mathbf{Ab}^k$  using EDA
- 9:     Update class labels in Eq. (9) using the ICM algorithm [40]
- 10:    Update the bias field using the method in [39]
- 11:    Update the bias field using the method in [39]
- 12: **end**
- 13: Execute step 2 to 12 again with the ranges of  $\boldsymbol{\theta}$  (pixel-based)

## Experiments and Results

### Benchmark Functions

In this section we compare the proposed CSA-DE/EDA algorithm to four bio-inspired computing techniques, including DE [14], CSA [21], EDA, and DE/EDA [13]. We make use of five benchmark functions for optimization, namely, the Generalized Rosenbrock (G. Ros), Modified

Langerman (M. Lan.), Shekel's Foxholes (S. Fox.), Epistatic Michalewicz (E. Mic.), and Chebychev Polynomial (T. Pol.) [13]. We set the dimensions of these functions to 5 and 10, respectively, and ran our algorithm 50 times on each function and each dimension value. To quantitatively assess the performance of different methods, we make use of three performance metrics, including the best value found (BVAT), the expected number of evaluations per success (ENES), and the average of the best



**Fig. 6** Segmentation accuracy for gray matter (bottom left), white matter (bottom right), and overall brain volume (top) on each image of the IBSR\_V2.0 dataset

objective function ( $f$ ) [13]. BVAT is the best value found during the 50 trials or runs. In our implementation, we stop the iterations when a given value-to-reach (VTR) is attained or after a certain number of iterations has been reached. ENES is estimated as the total number of function evaluations over 50 runs, divided by the number of runs that successfully attained the VTR. It measures the speed of the algorithm. The smaller ENES, the more efficient the algorithm is. The third metric  $f$  is the average of the best objective function values obtained in 50 runs. For a fair comparison, the parameters of the algorithm are set as DE [28]. Table 3 lists the parameters of the proposed algorithm. We implemented CSA with the default parameters used in [21] and tested it in the same way. As for DE, EDA, and DE/EDA, we adopted the performance reported in [13].

The plots of  $f$  versus the generations of populations in each algorithm are displayed in Fig. 2 (function dimension = 5) and

Fig. 3 (function dimension = 10). The figures reveal that the proposed algorithm converges faster than the other four algorithms.

The performances of the five optimization algorithms are presented in Tables 4 and 5. Moreover, to evaluate the difference between any two algorithms effectively, the paired Wilcoxon signed-rank test at 0.05 significance level for  $f$  was conducted on the obtained results. The symbols “+”, “~”, and “-” in Tables 4 and 5 express that the  $f$  obtained by CSA-DE/EDA is significantly better than (“+”), worse than (“-”), or almost similar to (“~”) the other algorithms. One can notice the low performance of CSA and EDA, since they failed to reach the VTRs in most of test cases, whereas the proposed CSA-DE/EDA algorithm achieved the second best BVAT and ENES on the five-dimensional E. Mic. function and the best performance in all other cases.

**Table 6** Average segmentation accuracy on the IBSR\_V2.0 dataset. Italic indicates the best performance

	Accuracy	FSL (%)	D-C (%)	SPM (%)	GA-GMM (%)	HMRF-CSA (%)	HMRF-CSA/DE/EDA (%)
Overall	75.06	75.02	81.02	74.97	82.95	89.90	
GM	77.35	73.80	84.42	77.90	84.92	92.02	
WM	87.08	88.41	87.38	87.23	83.88	87.52	

**Table 7** Average accuracy and computation time of PHMRF-CSA/DE/EDA and HMRF-CSA/DE/EDA on one 3D MR image of IBSR\_V2.0. Italic indicates the best performance

Methods	Overall accuracy (%)	DSC for GM (%)	DSC for WM (%)	Time cost (s)
PHMRF-CSA/DE/EDA	85.09	86.85	86.66	2164
HMRF-CSA/DE/EDA	<i>89.90</i>	92.02	87.52	<i>1510</i>

## Brain MR Image Segmentation

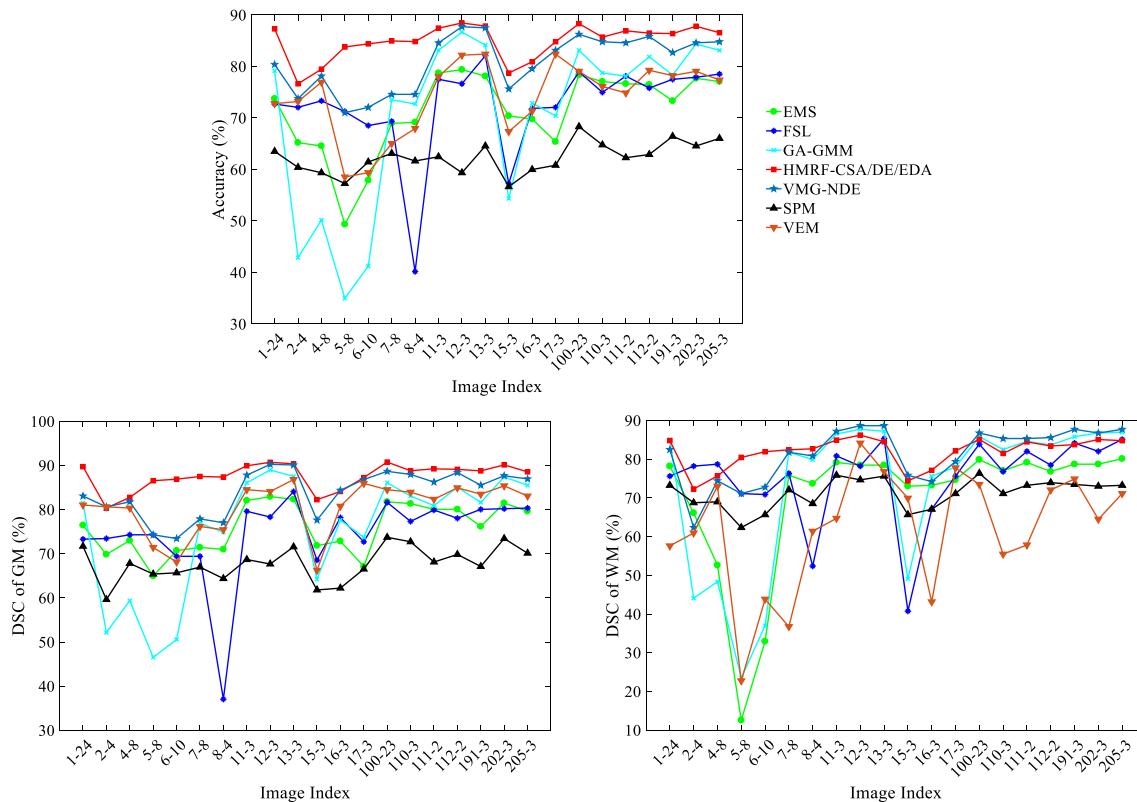
In the following we evaluate the performance of the CSA-DE/EDA algorithm for brain MR images segmentation using the proposed HMRF-CSA/DE/EDA MR segmentation algorithm of Section 4. For our experiments, we made use of MR images provided by the brain segmentation repository (IBSR) [44].

We empirically set the parameters of the CSA-DE/EDA algorithm as follows: population size  $N=50$ , scaling factor  $F=0.9$ , crossover rate  $CR=0.1$ , clonal multiplying factor  $\alpha=2$ , replacement ratio  $r=0.1$ , and 15 as maximum iterations, including 5 iterations for region-based segmentation and 10 iterations for pixel-based segmentation. The accuracy of delineating gray matter and white matter is quantitatively assessed via the Dice similarity coefficient (DSC) [45]:

$$D(V_s(k), V_g(k)) = 2 * \frac{|V_s(k) \cap V_g(k)|}{|V_s(k)| + |V_g(k)|} \quad (10)$$

with  $V_s(k)$  the volume of brain tissue class  $k$  in the segmentation result,  $V_g(k)$  the corresponding volume in the ground truth, and  $|V|$  represents the number of voxels in volume  $V$ . DSC values are in  $[0, 1]$ ; a high value indicates more accurate brain tissue delineation. The overall segmentation accuracy of a given MR image is measured as the percentage of correctly classified voxels.

We first performed segmentation experiments on 18 T1-weighted clinical brain MR images with expert segmentations (IBSR\_V2.0) [46]. Each image was first spatially normalized into the Talairach orientation, and then resliced into  $256 \times 256 \times 128$  voxels with a voxel size of  $1.0 \times 1.0 \times 1.5 \text{ mm}^3$ . Figure 4 shows the 68th coronal slice of image ‘IBSR\_09’ from IBSR\_V2.0, the intermediate region segmentation and final segmentation results of the proposed algorithm, along with the ground truth brain tissue map. As it can be noticed, the combined region- and pixel-based segmentation provides better delineation compared to the region-based segmentation. Next,



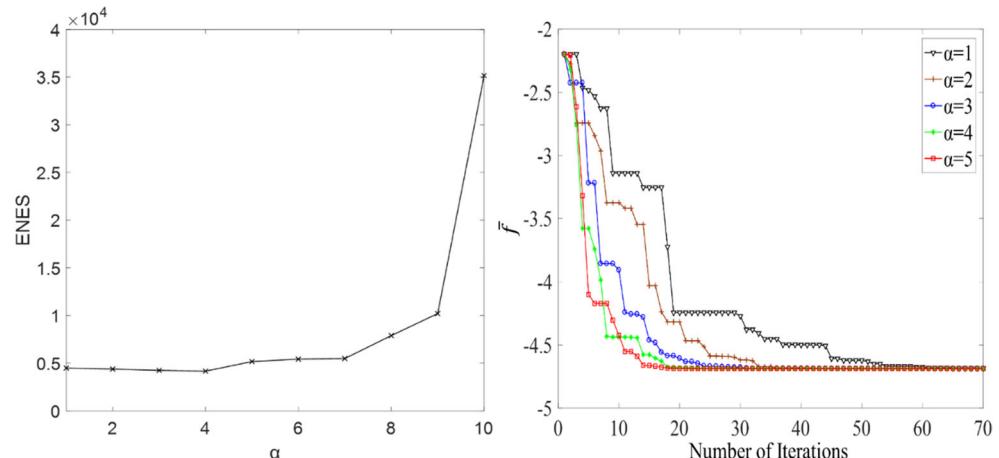
**Fig. 7** Segmentation accuracy on each image of the IBSR\_20Normals dataset; gray matter (bottom left), white matter (bottom right), and overall brain volume (top)

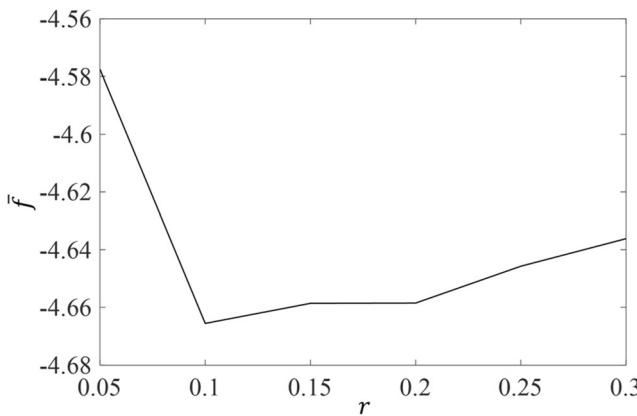
**Table 8** Average segmentation accuracy on the IBSR\_20Normals dataset

Accuracy	FSL (%)	EMS (%)	SPM (%)	GA-GMM (%)	VEM (%)	VMG-NDE (%)	HMRF-CSA/DE/EDA (%)
Overall	72.34	71.37	62.30	70.71	74.08	80.82	84.89
GM	74.51	75.86	67.76	75.49	80.47	83.83	87.57
WM	75.18	69.97	71.20	73.00	62.11	81.25	81.89

we compared the proposed HMRF-CSA/DE/EDA algorithm to the segmentation routines of the widely used statistical parametric mapping (SPM) toolbox [47] and FMRIB Software Library (FSL) [48] and three existing brain MR image segmentation algorithms, including the GA with Gaussian mixture model (GA-GMM) algorithm proposed in [49], the deformable co-segmentation (D-C) algorithm of [50], and the HMRF-CSA algorithm of [51]. The SPM segmentation routine is a unified registration-segmentation algorithm that is based on the Gaussian mixture model (GMM), and the FSL segmentation routine is based on the HMRF model solved using the expectation-maximization (EM) algorithm. Figure 5 depicts the coronal, sagittal, and transverse slices of image “IBSR\_14” from the IBSR\_V2.0 data set, the segmentation results of each of the considered methods, along with the ground truth. Figure 6 depicts the segmentation accuracy of each of the considered methods, and Table 6 lists their average segmentation accuracy. It shows that the proposed algorithm outperforms the other five algorithms on IBSR\_V2.0. Compared to the D-C algorithm, the proposed algorithm has a slightly lower accuracy in segmenting the white matter (WM) (87.52% vs. 88.41%), but substantially improved accuracy in segmenting the gray matter (GM) (92.02% vs. 73.80%) and the entire brain volume (89.90% vs. 75.02%). It seems that the higher performance of the D-C algorithm in WM segmentation is due to an over-segmentation of WM and under-segmentation of GM.

Using the IBSR\_V2.0 dataset, we conducted ablation experiments, in which the HMRF-CSA/DE/EDA algorithm was compared to its counterpart without using the region-based HMRF part (denoted by PHMRF-CSA/DE/EDA). Table 7 gives the obtained average accuracy and processing time using an Intel Core i7-4710HQ CPU, 16GB memory and Matlab R2015a. As it can be noticed, the HMRF-CSA/DE/EDA algorithm achieves a higher average segmentation accuracy and lower computational time compared to the pixel-based PHMRF-CSA/DE/EDA segmentation algorithm. The results also indicate that the region-based HMRF step does effectively contribute to the improvement of the image segmentation performance. However, although more efficient than the PHMRF-CSA/DE/EDA algorithm, our HMRF-CSA/DE/EDA algorithm still has a relatively high computational complexity due to the time-consuming nature of bio-inspired algorithms. Our final experiments were carried out on 20 T1-weighted clinical brain MR images of normal subjects with expert segmentations (IBSR\_20Normals). All images were reduced to 8-bit from 16-bit by thresholding intensities above or below 99.99% of the total number of different intensities scaled to the range [0, 255]. Each image was first spatially normalized into a standard 3D brain coordinate system by using rotation and non-deformable transformation [52, 53], and then resliced into  $256 \times 256 \times 63$  voxels of size  $1.0 \times 1.0 \times 3.1 \text{ mm}^3$ . We compared the proposed HMRF-CSA/DE/EDA algorithm to the segmentation routines in the

**Fig. 8** Effect of clonal multiplying factor  $\alpha$  on the ENES (left) and  $f$  (right) for five-dimensional E. Mic function



**Fig. 9** Effect of replacement ratio  $r$  on  $f$  for five-dimensional E. Mic function

SPM toolbox [47] and FSL package [48] and four other state-of-art brain MR image segmentation algorithms, including the GA-GMM algorithm [49], expectation-maximization segmentation (EMS) package [54], variational EM (VEM)-based Bayesian inference [55] and variational mixture of Gaussians model-based niche differential evolution algorithm (VMG-NDE) [56]. Figure 7 illustrates, for each image of the data set, the obtained accuracy for segmenting the GM, WM, and overall brain volume. The average segmentation accuracies of the different algorithms are given in Table 8. One can notice that the obtained results are consistent with the results of the previous experiments, demonstrating the ability of the proposed algorithm in producing more accurate segmentation of gray matter and overall brain volume compared to state-of-art algorithms. Moreover, it seems that our algorithm also outperforms the other algorithms in delineating the white matter.

## Discussion on the Parameters Setting

By incorporating the DE and EDA strategies into the CSA process, the proposed CSA-DE/EDA algorithm has more user-defined parameters than the original CSA. Among these parameters, the population size  $N$ , scaling factor  $F$ , and the DE-based crossover rate  $CR$  have been thoroughly discussed in the literature [13]. In our experiments, we mainly focused on settings the clonal multiplying factor  $\alpha$  for the DE-based hypermutation and the replacement ratio  $r$  for the EDA-based receptor editing.

To investigate the impact of the clonal multiplying factor  $\alpha$  on the performance of the proposed algorithm, we used the five-dimensional E. Mic function as a case study. We fixed other parameters as  $N=40$ ,  $F=0.5$ ,  $CR=0.3$ , and  $r=0.1$ , and evaluated the ENES and  $f$  when setting  $\alpha$  to different values. Figure 8 plots the ENES versus  $\alpha$  and the variation of  $f$  over the number of iterations when  $\alpha$  ranges from 1 to 5.

Since increasing the clonal multiplying factor  $\alpha$  leads to an increased number of evaluations of the fitness function, a large  $\alpha$  implies high computational cost in one iteration of the proposed algorithm. However, Fig. 8 shows that, with the increase of  $\alpha$ , ENES does not change much until  $\alpha$  becomes larger than 7. The reason is that for a larger number of  $\alpha$  the proposed algorithm uses the differential information or local information more adequately in the current population when the algorithm executes the DE-based hypermutation, and hence speeds up the convergence. Figure 8 depicts the convergence curves of the proposed algorithm with different values of  $\alpha$ , being consistent with our analysis. Therefore, we suggest setting  $\alpha$  to a value smaller than 7. Nevertheless, since increasing  $\alpha$  can improve somewhat the accuracy of the proposed algorithm, we can set  $\alpha$  to a larger value when the computational complexity is not a concern. This was the case for the 10-dimensional M. Lan. function for which we set  $\alpha$  to 10 in Table 3.

We also investigated the impact of replacement ratio  $r$  on the performance. We fixed the other parameters as  $N=40$ ,  $F=0.5$ ,  $CR=0.3$ ,  $\alpha=4$ , and set the maximum generation number to 15. The plot of  $f$  versus  $r$  is shown in Fig. 9. It reveals that  $f$  has smaller values when  $r$  takes a value in [0.1, 0.2]. Therefore, we suggest taking the value of  $r$  from this range.

## Conclusion

In this paper, we proposed a novel bio-inspired algorithm-CSA-DE/EDA for function optimization and brain image segmentation. This algorithm incorporates DE and EDA into the CSA process, and thus generates offspring by jointly using both local and global information from the current generation. Our experimental results, using benchmark functions, suggest that the proposed algorithm is superior to DE, EDA, CSA, and DE/EDA in function optimization. We also showed that it outperforms the GA-GMM, VEM, EMS, HMRF-CSA, D-C algorithms and the brain image segmentation routine in the commonly used SPM and FSL packages for brain MR image segmentation. Our future work will mainly focus on reducing the computation time of the algorithm by using parallel computing techniques.

**Acknowledgements** We appreciate the efforts devoted by the Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC) to collect and share the clinical MR brain data sets and their manual segmentations for comparing interactive and (semi)-automatic segmentation algorithms for MRI of major brain tissues.

**Funding** This work was supported in part by the National Natural Science Foundation of China under Grant 61771397, in part by the Science and Technology Innovation Committee of Shenzhen

Municipality, China, under Grant JCYJ20180306171334997, in part by the Seed Foundation of Innovation and Creation for Graduate Students in Northwestern Polytechnical University (NPU) under Grant ZZ2019029, in part by Synergy Innovation Foundation of the University and Enterprise for Graduate Students in NPU under Grant XQ201911, and in part by the Project for Graduate Innovation team of NPU.

## Compliance with Ethical Standards

**Conflict of Interest** We would like to submit a manuscript entitled “*CSA-DE/EDA: A Novel Bio-inspired Algorithm for Function Optimization and Segmentation of Brain MR Images*” for possible publication in *Cognitive Computation*. There are no potential conflicts of interest to report.

**Ethical Approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Molina D, LaTorre A, Herrera F. An insight into bio-inspired and evolutionary algorithms for global optimization: review, analysis, and lessons learnt over a decade of competitions. *Cogn Comput*. 2018;10(4):517–44. <https://doi.org/10.1007/s12559-018-9554-0>.
- Siddique N, Adeli H. Nature inspired computing: an overview and some future directions. *Cogn Comput*. 2015;7(6):706–14. <https://doi.org/10.1007/s12559-015-9370-8>.
- Srinivas M, Patnaik LM. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Trans Syst Man Cybern*. 1994;24(4):656–67. <https://doi.org/10.1109/21.286385>.
- Haktanirlar Ulutas B, Kulturel-Konak S. A review of clonal selection algorithm and its applications. *Artif Intell Rev*. 2011;36(2):117–38. <https://doi.org/10.1007/s10462-011-9206-1>.
- Binitha S, Sathy SS. A survey of bio inspired optimization algorithms. *Int J Soft Comput Eng*. 2012;2(2):137–51.
- Ma X, Li X, Zhang Q, Tang K, Liang Z, Xie W, et al. A survey on cooperative co-evolutionary algorithms. *IEEE Trans Evol Comput*. 2018;23:421–41. <https://doi.org/10.1109/TEVC.2018.2868770>.
- Gupta A, Ong Y-S. Back to the roots: multi-X evolutionary computation. *Cogn Comput*. 2019;11:1–17. <https://doi.org/10.1007/s12559-018-9620-7>.
- Tang Q, Shen Y, Hu C, Zeng J, Gong W. Swarm intelligence: based cooperation optimization of multi-modal functions. *Cogn Comput*. 2013;5(1):48–55. <https://doi.org/10.1007/s12559-012-9144-5>.
- Horn J, Nafpliotis N, Goldberg DE, editors. A niched Pareto genetic algorithm for multiobjective optimization. In: Proceedings of the first IEEE conference on evolutionary computation, IEEE World Congress on Computational Intelligence; 1994 Jun. Vol. 1, p. 82–87.
- Zhang Q, Sun J, Tsang E, Ford J. Hybrid estimation of distribution algorithm for global optimization. *Eng Comput*. 2004;21(1):91–107.
- Wang SY, Wang L. An estimation of distribution algorithm-based memetic algorithm for the distributed assembly permutation flowshop scheduling problem. *IEEE Trans Syst Man Cybern Syst Hum*. 2015;46(1):139–49.
- Wu CG, Wang L, Zheng XL. An effective estimation of distribution algorithm for solving uniform parallel machine scheduling problem with precedence constraints. *IEEE Congress on Evolutionary Computation*; 2016 Jul 24. p. 2626–32.
- Sun JY, Zhang QF, Tsang EPK. DE/EDA: a new evolutionary algorithm for global optimization. *Inf Sci*. 2005;169(3–4):249–62. <https://doi.org/10.1016/j.ins.2004.06.009>.
- Price KV. Differential evolution vs. the functions of the 2nd ICEO. In: Proceedings of 1997 IEEE International Conference on Evolutionary Computation; 1997 Apr 13. p. 153–157.
- Tang L, Dong Y, Liu J. Differential evolution with an individual-dependent mechanism. *IEEE Trans Evol Comput*. 2014;19(4):560–74.
- Kim SS, Mcloone S, Byeon JH, Lee S, Liu H. Cognitively inspired artificial bee colony clustering for cognitive wireless sensor networks. *Cogn Comput*. 2017;9(2):207–24.
- Xiaoyang F, Shuqing Z. An improved artificial immune recognition system based on the average scatter matrix trace criterion. *Adv Swarm Intell*. 2012;14(1):284–90.
- Vidal JM, Orozco ALS, Villalba LJJG. Adaptive artificial immune networks for mitigating DoS flooding attacks. *Swarm Evol Comput*. 2018;38:94–108.
- Zhang X, Zhang E, Li R. Optimized feature extraction by immune clonal selection algorithm. *IEEE Congress on Evolutionary Computation*; 2012 Jun 10. p. 1–6.
- Banerjee S, Moses M. Scale invariance of immune system response rates and times: perspectives on immune system architecture and implications for artificial immune systems. *Swarm Intell*. 2010;4(4):301–18.
- Lind C, Zuben FJV. Learning and optimization using the clonal selection principle. *IEEE Trans Evol Comput*. 2002;6(3):239–51.
- Zhang L, Gong M, Jiao L, Yang J. Optimal approximation of linear systems by an improved clonal selection algorithm. 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence); 2008 Jun 1. p. 527–534.
- Batista L, Guimaraes FG, Ramirez JA. A distributed clonal selection algorithm for optimization in electromagnetics. *IEEE Trans Magn*. 2009;45(3):1598–601. <https://doi.org/10.1109/tmag.2009.2012752>.
- Ulutas BH, Kulturel-Konak S. A review of clonal selection algorithm and its applications. *Artif Intell Rev*. 2011;36(2):117–38.
- Gong M, Jiao L, Zhang L. Baldwinian learning in clonal selection algorithm for optimization. *Inf Sci*. 2010;180(8):1218–36.
- Xu N, Ding Y, Ren L, Hao K. Degeneration recognizing clonal selection algorithm for multimodal optimization. *IEEE Trans Cybern*. 2018;48(3):848–61. <https://doi.org/10.1109/TCYB.2017.2657797>.
- Das S, Mullick SS, Suganthan PN. Recent advances in differential evolution—an updated survey. *Swarm Evol Comput*. 2016;27:1–30.
- Bersini H, Dorigo M, Langerman S, Seront G, Gambardella L. Results of the first international contest on evolutionary optimisation (1st ICEO). In: Proceedings of IEEE International Conference on Evolutionary Computation; 1996 May 20. p. 611–615.
- Efron MM, Coello CAC. A comparative study of differential evolution variants for global optimization. *Genetic and Evolutionary Computation Conference*; 2006. p. 485–92.
- Li G, Chan TM, Leung KS, Lee KH. An estimation of distribution algorithm for motif discovery. *IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*; 2008 Jun 1. p. 2411–8.
- Chen T, Tang K, Chen G, Yao X. On the analysis of average time complexity of estimation of distribution algorithms. *IEEE Congress on Evolutionary Computation*; 2007 Sep 25. p. 453–460.
- Larrañaga P, Lozano JA. Estimation of distribution algorithms: a new tool for evolutionary computation: Springer Science & Business Media; 2001 Oct 31.
- Campelo F, Guimaraes FG, Igarashi H, Ramirez JA. A clonal selection algorithm for optimization in electromagnetics. *IEEE Trans Magn*. 2005;41(5):1736–9. <https://doi.org/10.1109/TMAG.2005.846043>.
- Xie Y, Xia Y, Zhang J, Song Y, Feng D, Fulham M, et al. Knowledge-based collaborative deep learning for benign-

- malignant lung nodule classification on chest ct. *IEEE Trans Med Imaging*. 2018;38:991–1004. <https://doi.org/10.1109/TMI.2018.2876510>.
- 35. Zhang J, Xie Y, Xia Y, Shen C. Attention residual learning for skin lesion classification. *IEEE Trans Med Imaging*. 2019 Jan 21. <https://doi.org/10.1109/TMI.2019.2893944>.
  - 36. Song Y, Tan EL, Jiang X, Cheng JZ, Ni D, Chen S, et al. Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Trans Med Imaging*. 2016;36(1):288–300.
  - 37. Lei B, Yang P, Wang T, Chen S, Ni D. Relational-regularized discriminative sparse learning for Alzheimer’s disease diagnosis. *IEEE Trans Cybern*. 2017;47(4):1102–13.
  - 38. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation–maximization algorithm. *IEEE Trans Med Imaging*. 2001;20(1):45–57. <https://doi.org/10.1109/42.906424>.
  - 39. Li C, Gatenby C, Li W, Gore JC. A robust parametric method for bias field estimation and segmentation of MR images. *IEEE Conference on Computer Vision and Pattern Recognition*. 2009 Jun 20: 218–223.
  - 40. Besag J. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1986 Jul; 48(3): 259–279.
  - 41. Zheng C, Hu Y, Wang L, Qin Q. Region-based MRF model with optimized initial regions for image segmentation. *International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE)*; 2011 Jun 24, p. 3354–7.
  - 42. Levinstein A, Stere A, Kutulakos KN, Fleet DJ, Dickinson SJ, Siddiqi K. TurboPixels: fast superpixels using geometric flows. *IEEE Trans Pattern Anal Mach Intell*. 2009;31(12):2290–7. <https://doi.org/10.1109/TPAMI.2009.96>.
  - 43. Modestino JW, Zhang J, editors. A Markov random field model-based approach to image interpretation. *Computer Vision and Pattern Recognition*; 1989. p. 4–8.
  - 44. Iglesias JE, Liu CY, Thompson PM, Tu Z. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE transactions on medical imaging*. 2011 Apr 5;30(9):1617–34.
  - 45. Bharatha A, Hirose M, Hata N, Warfield SK, Ferrant M, Zou KH, et al. Evaluation of three-dimensional finite element-based deformable registration of pre- and intraoperative prostate imaging. *Med Phys*. 2001;28(12):2551–60.
  - 46. Rohlfing T. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. *IEEE Trans Med Imaging*. 2012;31(2):153–63. <https://doi.org/10.1109/TMI.2011.2163944>.
  - 47. Flandin G, Friston KJ. Statistical parametric mapping (SPM). *Scholarpedia*. 2008;3(4):6232.
  - 48. Library FS. in <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>. Accessed Jun 2016.
  - 49. Tohka J, Krestyannikov E, Dinov ID, Graham A, Shattuck DW, Ruotsalainen U, et al. Genetic algorithms for finite mixture model based voxel classification in neuroimaging. *IEEE Trans Med Imaging*. 2007;26(5):696–711.
  - 50. Zhang T, Xia Y, Feng DD. A deformable cosegmentation algorithm for brain MR images. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012 Aug 28. p. 3215–8.
  - 51. Zhang T, Xia Y, Feng DD. Hidden Markov random field model based brain MR image segmentation using clonal selection algorithm and Markov chain Monte Carlo method. *Biomed Signal Process Control*. 2014;12:10–8.
  - 52. Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. New York: Thieme; 1988.
  - 53. Filipek PA, Kennedy DN, Caviness VS. Volumetric analysis of central nervous system neoplasm based on MRI. *Pediatr Neurol*. 1991;7(5):347–51.
  - 54. Van Leemput K, Maes F, Vandermeulen D, Suetens P. A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans Med Imaging*. 2003;22(1):105–19.
  - 55. Bishop CM. *Pattern recognition and machine learning*. New York: Springer Stat Sci. 2006.
  - 56. Li Z, Xia Y, Ji Z, Zhang Y. Brain voxel classification in magnetic resonance images using niche differential evolution based Bayesian inference of variational mixture of Gaussians. *Neurocomputing*. 2017;269:47–57.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.