sciendo
JAISCR, 2020, Vol. 10, No. 1, pp. 57 – 69
10.2478/jaiscr-2020-0005

EFFICIENT IMAGE RETRIEVAL BY FUZZY RULES FROM BOOSTING AND METAHEURISTIC

Marcin Korytkowski[1,*], Roman Senkerik[2],
Magdalena M. Scherer[3], Rafal A. Angryk[4], Miroslaw Kordos[5], Agnieszka Siwocha[6]

[1]*Department of Computer Engineering, Częstochowa University of Technology, Częstochowa, Poland*

[2]*Tomas Bata University in Zlín, 760 05 Zlín, Czech Republic*

[3]*Faculty of Management, Częstochowa University of Technology, Częstochowa, Poland*

[4]*Department of Computer Science, Georgia State University, Atlanta, GA, USA*

[5]*Department of Computer Science and Automatics, University of Bielsko-Biała, Poland*

[6]*Information Technology Institute, University of Social Science, Łódź, Poland*
Clark University, Worcester, MA 01610, USA

[*]*E-mail: marcin.korytkowski@pcz.pl*

Submitted: 14th September 2019; Accepted: 16th November 2019

Abstract

Fast content-based image retrieval is still a challenge for computer systems. We present a novel method aimed at classifying images by fuzzy rules and local image features. The fuzzy rule base is generated in the first stage by a boosting procedure. Boosting meta-learning is used to find the most representative local features. We briefly explore the utilization of metaheuristic algorithms for the various tasks of fuzzy systems optimization. We also provide a comprehensive description of the current best-performing DISH algorithm, which represents a powerful version of the differential evolution algorithm with effective embedded mechanisms for stronger exploration and preservation of the population diversity, designed for higher dimensional and complex optimization tasks. The algorithm is used to fine-tune the fuzzy rule base. The fuzzy rules can also be used to create a database index to retrieve images similar to the query image fast. The proposed approach is tested on a state-of-the-art image dataset and compared with the bag-of-features image representation model combined with the Support Vector Machine classification. The novel method gives a better classification accuracy, and the time of the training and testing process is significantly shorter.

Keywords: image retrieval, fuzzy rules, local image features

1 Introduction

Content-based image retrieval (CBIR) systems allow for browsing, searching and retrieving images relevant to the query. With the popularity of imaging devices, the demand for such systems is growing. To design a CBIR system, we need to develop image representation in the form of relevant features. Then, we have to compare the features efficiently as we usually deal with collections of thousands to billions of images. The similarity between images is traditionally reflected by the similarity between their features. At the beginning of the century, methods based on local image features [38, 41, 42, 45, 64] started to gain popularity with the most popular SURF [5], SIFT [38] or ORB [57] invariant local features. Despite the popularity of the image keypoints, image retrieval of images similar to the query image is not easy. Objects can be recorded under various angles and perspectives. Images and objects can be described by many kinds of features such as color [29, 33, 47], textures [14, 23, 31, 65], shape [30, 32, 73] or edge detectors [85]. The output of these methods is one or more descriptor vectors which have to be compared in order to search for images. Such comparison is enormously time consuming, and there is ongoing worldwide research aimed at speeding up the process. The next partial breakthrough, after local invariant features, was adopting the bag of words model from information retrieval to computer vision. Yet, the current state of the art in the case of high dimensional computer vision applications is not fully satisfactory. The literature presents countless methods and variants utilizing e.g. a voting scheme or histograms of clustered keypoints. They are mostly based on some form of approximate search. One of the solutions to the problem can be descriptor vector hashing. In [19] the authors proposed a locality-sensitive hashing method for the approximate nearest neighbour algorithm. In [45] the authors built a hierarchical quantizer in the form of a tree. Such a tree is a kind of an approximate nearest neighbour algorithm and constitutes a visual dictionary.

As aforementioned, the bag-of-features (BoF) approach [27, 48, 64, 78, 82] is a popular image retrieval and classification method. In BoF, histograms of descriptors are computed and the method can be modified, e.g. by applying the earth mover's distance, presented in [27]. The main problem with this family of methods is that vector comparison is very time consuming, and if we add new classes, the set of histograms needs to be rebuilt. Our approach is a fast index of descriptors based on fuzzy rules. The words used in BoF are a somewhat similar concept to the fuzzy rules from our approach.

In this paper we present a method for classifying and fast retrieving images (partially inspired by [70, 77, 84]) which uses boosting metalearning to search for the most salient image features. In [70, 77] certain feature values become weak classifiers for detecting faces. In our approach AdaBoost is used to select the salient image descriptors to generate fuzzy rules which use fuzzy sets to describe information [61, 62]. We draw randomly one descriptor from the positive set to make a base for a new fuzzy rule (new classifier). The parameters of this rule are changed to better accommodate the rule to its class. Then, the differential evolution SHADE algorithm described in Section 2 is used to optimise the fuzzy rule base. We chose SHADE as it proved to work well in high-dimensional search spaces. The presented approach can use various image local features, hand-crafted (e.g. SIFT or SURF) and learned ones. The remainder of the paper is organised as follows. In Section 2 we briefly explore the utilization of metaheuristic algorithms for the various tasks of fuzzy systems optimization and the SHADE algorithm. In Section 3 we present the fuzzy rule generation algorithm. In Section 4 we provide a description of a new, query image classification and the retrieval of similar images. Section 5 compares the algorithm with an established image retrieval algorithm and Section 6 concludes the paper.

2 Metaheuristics for Fuzzy Rules

Currently, there exist many types of fuzzy systems (FS) and Fuzzy rule–based systems (FRBS) with different structures, features, and requirements for robust and effective optimization and learning. The need for effective optimization and learning of highly–accurate FSs/FRMSs is in most cases motivated by the requirement to efficiently process high–dimensional and high–volume data for which a manual (apriori knowledge) design by experts is not feasible [4]. The metaheuristic algorithms are then a clear choice.

A wide portfolio of metaheuristic techniques, including all classic evolutionary and swarm-based algorithms, altogether with the symbolic nature of some evolutionary methods (e.g., genetic programming), have been introduced for the design, learning, and optimization of FSs/FRBSs [16, 15, 22], thus creating evolutionary (genetic) fuzzy systems [16]. Algorithms such as ant colony optimization (ACO) [13], genetic algorithms (GA) [12], genetic programming (GP) [25, 24, 7], multi–gene genetic programming [34], artificial bee colony (ABC) optimization [6], differential evolution (DE) [36, 53, 71, 40], especially its state-of-the-art version SHADE/L-SHADE [72], have all been used to address research tasks. Also, multi-objective optimization has become an important approach for optimal design and learning of FSs/FRBSs [58, 56, 2].

All above-mentioned metaheuristic methods have addressed various aspects of FSs/FRBSs design including rule base learning [13, 7, 16, 4, 34, 6, 1] and optimization [16, 25, 44, 37], membership function scaling and tuning [16, 12], evolutionary synthesis of fuzzy rules [35], and so on.

There is a broad application field of evolutionary optimized (synthesized) FSs such as fuzzy intrusion detection system [20], classification of healthcare data [44], decision support in healthcare [1], medical diagnostics [26], advanced driver assistance systems [36], linguistic modelling [15, 16] or deep evolving fuzzy neural networks [52].

Finally, FS/FRBS optimization and design can be seen as a challenging high–dimensional optimization problem, mostly with many objectives. Since the modern DE versions have been used in the most recent frameworks (like DECO3RUM) [72, 71] and single/multi-objective applications [36], the following text is focused on this powerful algorithm.

Among the existing implementations of metaheuristic algorithms for fuzzy rules optimization, DE [66] modifications such as DISH [75] and others [18, 50, 51] can be considered. DE has been thoroughly investigated with an emphasis on the theoretical insight and insights into inner population dynamics [81, 76, 68, 46]. The DE algorithm-based family is often represented in contests at the Congress on Evolutionary Computation (CEC) [54, 17, 43, 39, 55, 79]. For this reason, we expect these advanced versions of DE to be effective for the fuzzy rules optimization problem, especially in high dimensional applications. One of the newest DE algorithms is the Success-History based Adaptive Differential Evolution (SHADE) [67], which has a line of recent improvements following JADE [83] that is based on jDE [9], upgraded as L-SHADE [69], SPS-L-SHADE-EIG [28], LSHADE-cnEpSin [3], jSO [11], aL-SHADE [49], and most recently, DISH [75]. To make the paper self-contained we describe the canonical DE followed by necessary improvements leading to the most recent DISH version.

The canonical 1995 DE computes the parameters via evolution of a set of solutions of population $\boldsymbol{P}$ of size NP and is based on parameter estimation through evolution from a randomly generated set of solutions using population $\boldsymbol{P}$, which has a preset size of NP. Individuals are vectors $\boldsymbol{x}$ of length D. Objective function $f(\boldsymbol{x})$ expresses the quality of the solution. First, all individuals in the initial population $\boldsymbol{P}$ are uniformly generated at random with constraints $[x_{\mathrm{lower},j}, x_{\mathrm{upper},j}]$, $\forall j = 1, \ldots, D$

$$x_i = \{\mathcal{U}[x_{\mathrm{lower},j},\ x_{\mathrm{upper},j}]\}; \\ \forall j = 1, \ldots, D; \forall i = 1, \ldots, NP, \tag{1}$$

then, three indices r_1, r_2, and r_3, are used to compute a differential vector (hence the name DE for the algorithm) and combine it

$$v_i = x_{r1} + F(x_{r2} - x_{r3}), \tag{2}$$

which is then taken into crossover with the current vector at index i

$$u_{j,i} = \begin{cases} v_{j,i} & \text{if } \mathcal{U}[0,1] \le CR \text{ or } j = j_{\mathrm{rand}} \\ x_{j,i} & \text{otherwise} \end{cases}, \tag{3}$$

and then a selection operator yields a new vector $x_{i,G+1}$ at this location i for next generation $G+1$

$$x_{i,G+1} = \begin{cases} u_{i,G} & \text{if } f(u_{i,G}) \le f(x_{i,G}) \\ x_{i,G} & \text{otherwise} \end{cases}. \tag{4}$$

The SHADE algorithm has a self-adaptive mechanism of control parameters i.e. scaling factor F and crossover rate CR and is inspired slightly by JADE [83]. The main difference is in the historical memories M_{F} and M_{CR} for successful scaling factor and crossover rate values with their update mechanism.

The mutation scheme "current-to-$p$best/1" combines four index-wise mutually different vectors in computation of the mutated vector $\boldsymbol{v}$, with the index of $\boldsymbol{x}_{p\mathrm{best}}$ being different from $r1$, $r2$, and i, as

$$v_i = x_i + F_i(x_{p\mathrm{best}} - x_i) + F_i(x_{r1} - x_{r2}), \tag{5}$$

where $\boldsymbol{x}_{p\mathrm{best}}$ is a randomly selected individual from the best $NP \times p$ individuals in the current population. The p value is generated randomly for each mutation by PRNG, with the uniform distribution from the range $[p_{\min}, 0.2]$ and $p_{\min} = 2/NP$. Vector $\boldsymbol{x}_{r1}$ is selected randomly from the current population $\boldsymbol{P}$. Vector $\boldsymbol{x}_{r2}$ is obtained randomly from the union of the current population $\boldsymbol{P}$ and external archive $\boldsymbol{A}$. The scaling factor value F_i is given by

$$F_i = \mathcal{C}\left[M_{\mathrm{F},r}, 0.1\right], \tag{6}$$

where $M_{\mathrm{F},r}$ is a randomly selected value (index r is generated by PRNG from the range 1 to H) from $\mathbf{M}_\mathrm{F}$ memory, and $\mathcal{C}$ stands for the Cauchy distribution. Therefore the F_i value is generated from the Cauchy distribution with location parameter value $\mathbf{M}_\mathrm{r}$ and scale parameter value of 0.1. If the generated value F_i is higher than 1, it is truncated to 1, and if F_i is less or equal to 0, it is generated again by (6).

The crossover uses the very same scheme as in (3), only with the difference, that CR value is not static, but it is generated from a Gaussian distribution with a mean parameter value $M_{\mathrm{CR},r}$ selected from the crossover rate historical memory $\mathbf{M}_\mathrm{CR}$ by the same index r as in the scaling factor case and Standard Deviation value of 0.1:

$$CR_i = \mathcal{N}\left[M_{CR,r}, 0.1\right]. \tag{7}$$

When the generated CR_i value is less than 0, it is replaced by 0, and when it is greater than 1, it is replaced by 1. Selection process is again the same as for the canonical DE (4).

Historical memories $\mathbf{M}_\mathrm{F}$ and $\mathbf{M}_\mathrm{CR}$ serve to store successful values of F and CR used in the mutation and crossover steps. During every single generation, these successful values are stored in their corresponding arrays $\mathbf{S}_F$ and $\mathbf{S}_{CR}$. After each generation, one cell of $\mathbf{M}_\mathrm{F}$ and $\mathbf{M}_\mathrm{CR}$ memories is updated. This cell is given by index k, which starts at 1 and increases by 1 after each generation. When it overflows the memory size H, it is reset to 1. The new values of the k-th cell for $\mathbf{M}_\mathrm{F}$ and $\mathbf{M}_\mathrm{CR}$ are calculated, respectively

$$M_{F,k} = \begin{cases} \mathrm{mean}_{WL}\left(\mathbf{S}_F\right) & \text{if } \mathbf{S}_F \neq \emptyset \\ M_{F,k} & \text{otherwise} \end{cases}, \tag{8}$$

$$M_{CR,k} = \begin{cases} \mathrm{mean}_{WL}\left(\mathbf{S}_{CR}\right) & \text{if } \mathbf{S}_{CR} \neq \emptyset \\ M_{CR,k} & \text{otherwise} \end{cases}, \tag{9}$$

where $\mathrm{mean}_{WL}()$ stands for the weighted Lehmer mean

$$\mathrm{mean}_{WL}\left(\mathbf{S}\right) = \frac{\sum_{k=1}^{|\mathbf{S}|} w_k \bullet S_k^2}{\sum_{k=1}^{|\mathbf{S}|} w_k \bullet S_k} \tag{10}$$

and the weight vector $\boldsymbol{w}$ is based on the improvement in the objective function value between the trial and the original individuals in current generation G, as follows

$$w_k = \frac{\mathrm{abs}\left(f\left(u_{k,G}\right) - f\left(x_{k,G}\right)\right)}{\sum_{m=1}^{|\mathbf{S}_{CR}|} \mathrm{abs}\left(f\left(u_{m,G}\right) - f\left(x_{m,G}\right)\right)}. \tag{11}$$

Because both arrays $\mathbf{S}_F$ and $\mathbf{S}_{CR}$ have the same size, it is arbitrary which size will be used for the upper boundary for m in Equation (11).

Another operation that distinguishes L-SHADE from SHADE algorithm is the linear population decrease. The basic idea is to reduce the population size to promote exploitation in later phases of the evolution. Therefore, a new population size is calculated after each generation (12). Whenever the new population size NP_new is smaller than the current population size NP, the population is sorted according to the objective function value, and the worst $NP - NP_\mathrm{new}$ individuals are discarded. The size of the external archive is reduced as well, using the formula

$$\begin{aligned} NP_\mathrm{new} = \mathrm{round}(NP_\mathrm{init} - \\ - \frac{FES}{MAXFES} * \left(NP_\mathrm{init} - NP_\mathrm{f}\right)), \end{aligned} \tag{12}$$

where NP_{init} is the initial population size and NP_f is the final population size. FES and $MAXFES$ are objective function number evaluations and a maximum number of objective function evaluations, respectively.

Later, the iL-SHADE [10] extends L-SHADE by initializing all values in $\mathbf{M}_\mathrm{F}$ and $\mathbf{M}_\mathrm{CR}$ at 0.8, additional historical memory entry $M_{\mathrm{F},H} = M_{\mathrm{CR},H} = 0.9$, limiting F and CR values in the early stages, and a new formula for computing p for pBest mutation strategy.

Finally, the next algorithmic evolution was represented by the jSO algorithm [11] as the announced winner at the CEC 2017 Competition, and introduces mainly a new weighted version of the mutation strategy and further parameter setup improvements (for details, please, refer to [11]).

The jSO algorithm was used later as a final stage of the evolving of the DISH algorithm which uses distance based parameter adaptation. The original adaptation mechanism for the scaling factor and the crossover rate values uses weighted forms of the means (11), where weights are based on the improvement in the objective function value (10). This approach promotes exploitation over exploration and, therefore, might lead to premature convergence.

The distance approach is based on the Euclidean distance between the trial and the original individual (11). In this modification, the scaling factor and crossover rate values connected with the individual that have moved the furthest will have the highest weight

$$w_k = \frac{\sqrt{\sum_{j=1}^{D} \left(u_{k,j,G} - x_{k,j,G}\right)^2}}{\sum_{m=1}^{|S_{CR}|} \sqrt{\sum_{j=1}^{D} \left(u_{m,j,G} - x_{m,j,G}\right)^2}}. \tag{13}$$

Therefore, the exploration ability and higher population diversity are rewarded, and this should lead to avoidance of the premature convergence in higher dimensional objective spaces. This distance based approach can be easily implemented to any variant of SHADE/L-SHADE family of algorithms [74].

3 Boosting-Generated Simple Fuzzy Classifiers

In this Section, we find the most representative fuzzy rules for visual class ω_c, $c = 1, \ldots, V$ which we use to retrieve similar images or to just index images in a large repository. As we use the SIFT descriptors, classifiers have $N = 128$ features. The fuzzy rules have the following form

$$\begin{aligned} R_t^c:\ &\text{IF } x_1 \text{ is } G_{1,t}^c \text{ AND } x_2 \text{ is } G_{2,t}^c \text{ AND} \ldots \\ &\ldots \text{ AND } x_{128} \text{ is } G_{128,t}^c \text{ THEN image } i \in \ \omega_c(\beta_t^c) \end{aligned}, \tag{14}$$

where $t = 1, \ldots, T^c$ is the rule number in the current run of boosting, T^c is the number of rules for the class ω_c and β_t^c is the weak hypothesis significance. In the paper we use the Gaussian membership functions

$$G_{n,t}^c(x) = e^{-\left(\frac{x - m_{n,t}^c}{\sigma_{n,t}^c}\right)^2}, \tag{15}$$

where $m_{n,t}^c$ is the center of the Gaussian function (15) and $\sigma_{n,t}^c$ is its width. For the clarity of presentation this Section describes generating the ensemble of weak classifiers for a class ω_c, thus the class index c will be omitted.

The training dataset has I images (I_{pos} positive ones and I_{neg} negative ones). Initially, descriptors have the same boosting weights

$$D_1^l = \frac{1}{L} \text{ for } l = 1, \ldots, L\,, \tag{16}$$

where L is the number of descriptors. Two matrices are the training dataset of image descriptors

$$\mathbf{P}_t = \begin{bmatrix} \mathbf{p}^1 & D_t^1 \\ \vdots & \vdots \\ \mathbf{p}^{L_{pos}} & D_t^{L_{pos}} \end{bmatrix} = \begin{bmatrix} p_1^1, \ldots, p_N^1 & D_t^1 \\ \vdots & \vdots \\ p_1^{L_{pos}}, \ldots, p_N^{L_{pos}} & D_t^{L_{pos}} \end{bmatrix}, \tag{17}$$

$$\mathbf{N}_t = \begin{bmatrix} \mathbf{n}^1 & D^1 \\ \vdots & \vdots \\ \mathbf{n}^{L_{neg}} & D_t^{L_{neg}} \end{bmatrix} = \begin{bmatrix} n_1^1, \ldots, p_N^1 & D^1 \\ \vdots & \vdots \\ n_1^{L_{neg}}, \ldots, p_N^{L_{neg}} & D_t^{L_{neg}} \end{bmatrix}. \tag{18}$$

The learning process consists in creating T simple classifiers (weak learners in the boosting terminology) in the form of fuzzy rules (14). After each run t, $t = 1, \ldots, T$, of the algorithm, we create rule R_t and the algorithm is presented below.

1. Randomly choose one vector $\mathbf{p}^r$, $1 \leq r \leq L_{pos}$ from the positive samples using normalized distribution of elements $D_t^1, \ldots, D_t^{L_{pos}}$ in matrix (17). This drawn vector becomes a basis for generating a new classifier and the learning set weights contribute to the probability of choosing a keypoint.

2. For each image from the positive image set find the feature vector which is nearest to $\mathbf{p}^r$ (for example according to the Euclidean distance) and store this vector in matrix $\mathbf{M}_t$ of the size $I_p \times N$. Every row represents one feature from a different image v_i, $i = 1, \ldots, I_{pos}$, and no image occurs more than once

$$\mathbf{M}_t = \begin{bmatrix} \tilde{p}_{t,1}^1 & \cdots & \tilde{p}_{t,N}^1 \\ \vdots & \cdots & \vdots \\ \tilde{p}_{t,1}^j & \ddots & \tilde{j}_{t,N}^j \\ \vdots & \cdots & \vdots \\ \tilde{p}_{t,1}^{I_{pos}} & \cdots & \tilde{p}_{t,N}^{I_{pos}} \end{bmatrix}, \tag{19}$$

Each vector $\begin{bmatrix} \tilde{p}_{t,1}^j & \cdots & \tilde{p}_{t,N}^j \end{bmatrix}$, $j = 1, \ldots, I_{pos}$, in matrix (19) contains one visual descriptor from set $\{\mathbf{p}^i; i = 1, .., L_{pos}\}$.

3. In this step a weak classifier is built, i.e. we find centres and widths of Gaussian functions which are membership functions of fuzzy sets in a fuzzy rule (14).

   (a) Compute absolute value $d_{t,n}$ as the difference between the smallest and the highest values in each column of the matrix (19)

   $$d_{t,n} = |\min_{i=1,\dots,I_p} p_n^i - \max_{i=1,\dots,I_p} p_n^i,| \quad (20)$$

   where $n = 1,\dots,N$. Compute the center of fuzzy Gaussian membership function (15) $m_{t,n}$ in the following way

   $$m_{t,n} = \max_{i=1,\dots,I_p} p_n^i - \frac{d_{t,n}}{2} . \quad (21)$$

   Now we have to find the widths of these fuzzy set membership functions. We have to assume that for all real arguments in the range of $\left[m_{t,n} - \frac{d_{t,n}}{2}; m_{t,n} + \frac{d_{t,n}}{2}\right]$, the Gaussian function (fuzzy set membership function) values should satisfy $G_{n,t}(x) \geq 0.5$. Only in this situation do we activate the fuzzy rule. As we assume that $G_{n,t}(x)$ is at least 0.5 to activate a fuzzy rule, using simple substitution $x = m_{t,n} - \frac{d_{t,n}}{2}$, we obtain the relationship for $\sigma_{t,n}$

   $$\sigma_{t,n} = \frac{d_{t,n}}{2\sqrt{-\ln(0.5)}}. \quad (22)$$

   Finally, we have to calculate values $m_{t,n}$ and $\sigma_{n,t}$ for every element of the nth column of matrix (19), thus we have to repeat the above steps for all N dimensions. In this way, we obtain N Gaussian membership functions of N fuzzy sets. Of course, we can label them using fuzzy linguistic expressions such as 'small', 'large' etc., but for the time being we mark them only in a mathematical sense by $G_{n,t}$, where n, $n = 1,..,N$, is the index associated with feature vector elements and t means the fuzzy rule number.

   (b) Using values obtained in point a) we can construct a fuzzy rule which creates a fuzzy classifier (14).

4. Now we have to evaluate the quality of the classifier obtained in step 3. We do this using the standard AdaBoost algorithm [60]. Let us determine the activation level of the rule R_t which is computed by a t-norm of all fuzzy sets membership function values

   $$f_t(\bar{\mathbf{x}}) = \mathop{T}_{n=1}^{N} G_{n,t}(\bar{x}_n) , \quad (23)$$

   where $\bar{\mathbf{x}} = [\bar{x}_1,\dots,\bar{x}_N]$ is a vector of the values of linguistic variables $x_1,\dots,x_N$. In the case of the minimum t-norm, formula (23) becomes

   $$f_t(\bar{\mathbf{x}}) = \min_{n=1}^{N} G_{n,t}(\bar{x}_n) . \quad (24)$$

   As a current run of the AdaBoost is for a given class ω_c, we can treat the problem as a binary classification (dichotomy) i.e. $y^l = 1$ for descriptors of positive images and $y^l = 0$ for descriptors of negative images. Then the fuzzy classifier decision is computed by

   $$h_t(\bar{\mathbf{x}}^l) = \begin{cases} 1 & \text{if} \quad f_t(\bar{\mathbf{x}}^l) \geq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} . \quad (25)$$

   For all the keypoints stored in matrices $\mathbf{P}_t$ and $\mathbf{N}_t$ we calculate new weights D_t^l. To this end, we compute the error of classifier (25) for all $L = L_{pos} + L_{neg}$ descriptors of all positive and negative images

   $$\varepsilon_t = \sum_{l=1}^{L} D_t^l I(h_t(\bar{\mathbf{x}}^l) \neq y^l) , \quad (26)$$

   where I is the indicator function

   $$I(a \neq b) = \begin{cases} 1 & \text{if} \quad a \neq b \\ 0 & \text{if} \quad a = b \end{cases} . \quad (27)$$

   If $\varepsilon_t = 0$ or $\varepsilon_t > 0.5$, we finish the training stage. If not, we compute new weights

   $$\alpha_t = 0.5 \ln \frac{1-\varepsilon_t}{\varepsilon_t} . \quad (28)$$

   $$D_{t+1}^l = \frac{D_t^l \exp\{-\alpha_t I(h_t(\bar{\mathbf{x}}^l) = y^l)\}}{C} , \quad (29)$$

   where C is a constant such that $\sum_{l=1}^{L} D_{t+1}^l = 1$. Finally, classifier importance is determined by

   $$\beta_t = \frac{\alpha_t}{\sum_{t=1}^{T} \alpha_t} . \quad (30)$$

The obtained set of rules R is then fine-tuned by the SHADE algorithm described in Section 2 against achieving the best image classification accuracy. Later, we present image indexation and retrieval.

4 Classification of a Query Image

Each group of images ω_c, $c = 1, \ldots, V$ requires generation of fuzzy rules, thus after the training procedure and the evolutionary optimization, we obtain a set of V strong classifiers. A new query image has its u descriptors in $\mathbf{Q}$

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}^1 \\ \mathbf{q}^2 \\ \vdots \\ \mathbf{q}^u \end{bmatrix} = \begin{bmatrix} q_1^1 \cdots q_N^1 \\ q_1^2 \cdots q_N^2 \\ \vdots \\ q_1^u \cdots q_N^u \end{bmatrix} . \tag{31}$$

Let us determine the value of

$$F_t(\mathbf{Q}) = \mathop{S}_{j=1}^{u} \left(\mathop{T}_{n=1}^{N} G_{n,t}(q_n^j) \right) , \tag{32}$$

where S and T are t-norm and t-conorm, respectively (see [59]). To compute the overall output of the ensemble of classifiers designed in Section 3, for each class ω_c we sum weak classifiers outputs (32) taking into consideration their importance (30), i.e.

$$H^c(\mathbf{Q}) = \sum_{t=1}^{T^c} \beta_t F_t(\mathbf{Q}) . \tag{33}$$

Eventually, we assign a class label to the query image in the following way

$$f(\mathbf{Q}) = \arg \max_{c=1,\ldots,V} H^{*c}(\mathbf{Q}) . \tag{34}$$

In formulas (33) and (34) we restored class label index c, which had been removed at the beginning of Section 3. In formula (32) t-norm and t-conorm can be chosen as min and max operators, respectively.

The fuzzy rules created during the boosting learning and tuned by the metaheuristic are used to fast retrieve images similar to the query image, which we show in Figure 1. We create a database index by identifying the ranges in the Gaussion functions having values greater than 0.5. The database index determines which image feature values fall into the ranges in which the fuzzy sets which constitute the predecessor of the rule have values greater than 0.5.

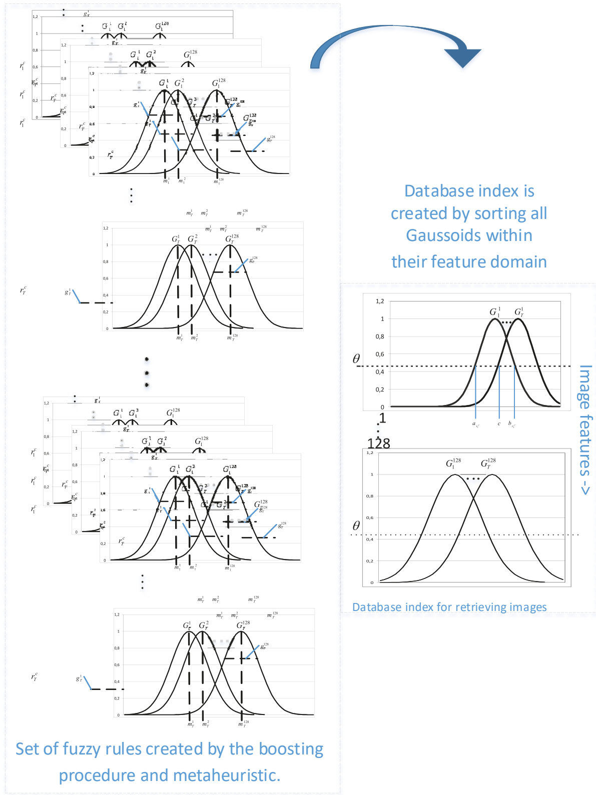

Figure 1. Set of fuzzy rules for image classification created by boosting and differential evolution (left part). Right part: database index from the rules allowing fast retrieval of similar images.

5 Experiments

We evaluated the presented approach on images taken from the PASCAL Visual Object Classes (VOC) dataset [21] by checking the speed and accuracy. We present some examples in Fig. 2. We divided each class of objects into training and testing examples (15 %). We generated local keypoint descriptors with the SIFT algorithm; for complex images there would be even thousands of descriptors.

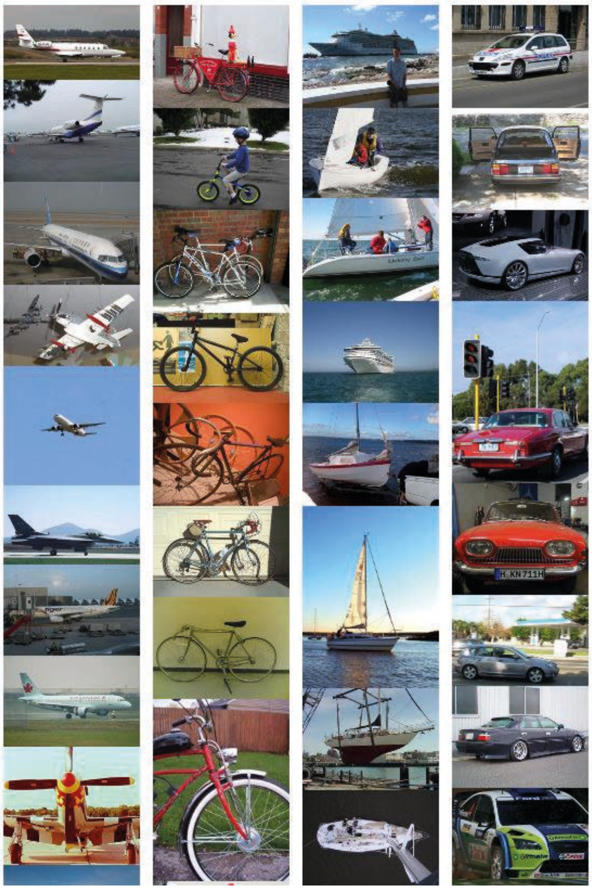

Figure 2. Examples of images from the PASCAL Visual Object Classes (VOC) dataset, namely aeroplanes, bicycles, boats and cars.

We used negative images from a different kind of images from the dataset. We checked the proposed method performance against the Support Vector Machine (SVM) with the Chi-Square kernel. The training procedure described in Section 3 requires a set of negative examples for each considered class of objects. We picked randomly negative examples from other classes. We ran it with a dictionary of the size of 400 words. We created dictionaries for BoF in C++ language, based on the OpenCV Library [8]. Both methods were evaluated with the same images (Table 1). In the BoF algorithm the column "Training time" is empty as the training is performed for the whole dataset. As we can see, the algorithm presented in the paper is faster and more accurate than the BoF approach.

6 Conclusions

We proposed a new method of creating fuzzy rules from image local features by boosting and differential evolution. We briefly described content-based image retrieval and utilization of metaheuristic algorithms for various tasks in fuzzy system optimization. Further, we gave a comprehensive description of the current best-performing DISH algorithm, which represents a powerful version of a differential evolution algorithm with effective embedded mechanisms for stronger exploration and preservation of the population diversity, designed for higher dimensional and complex optimization tasks. We used the DISH algorithm to fine-tune the fuzzy rules obtained by the boosting procedure. The proposed approach outperformed the state-of-the-art method in image retrieval, which is a combination of the bag of features method with SVM. Our approach is faster and more accurate. Moreover, contrary to the bag-of-features approach, it is relatively simple to train the system to recognize new image classes. We used the SIFT image features, but the proposed method can use other image keypoint detectors and descriptors, hand-crafted as SURF or ORB and learned ones as LIFT [80] or that proposed in [63].

Acknowledgements

This work was supported by the project financed under the program of the Polish Minister of Science and Higher Education under the name "Regional Initiative of Excellence" in the years 2019–2022 project number 020/RID/2018/19, the amount of financing 12,000,000.00 PLN.

References

[1] Alharbi, A., Tchier, F.: Using a genetic-fuzzy algorithm as a computer aided diagnosis tool on saudi arabian breast cancer database. Mathematical Biosciences 286, 39 – 48 (2017)

[2] Antonelli, M., Ducange, P., Marcelloni, F.: A fast and efficient multi-objective evolutionary learning scheme for fuzzy rule-based classifiers. Information Sciences 283, 36 – 54 (2014). New Trend of Computational Intelligence in Human-Robot Interaction

[3] Awad, N.H., Ali, M.Z., Suganthan, P.N., Reynolds,

Table 1. Comparison of the proposed method with the bag of words combined with the support vector machines.

	Proposed approach			Bag of features and SVM		
	Classification accuracy on testing set	Training time [s]	Testing time [s]	Classification accuracy on testing set	Training time [s]	Testing time [s]
bicycle	87.45%		2.236	69.54%		7.141
boat	77.55%		2.435	66.84%		6.274
bus	86.38%		3.023	70.89%		5.241
car	78.43%		3.274	88.45%		7.274
cat	78.74%		3.137	88.72%		5.134
plane	86.89%		3.272	80.45%		6.233
train	73.53%		3.458	54.34%		5.381
Total	81.28%	287.381	20.925	74.17%	544.323	42.678

R.G.: An ensemble sinusoidal parameter adaptation incorporated with L-SHADE for solving CEC2014 benchmark problems. In: 2016 IEEE Congress on Evolutionary Computation (CEC), pp. 2958–2965. IEEE (2016)

[4] Aydogan, E.K., Karaoglan, I., Pardalos, P.M.: hga: Hybrid genetic algorithm in fuzzy rule-based classification systems for high-dimensional problems. Applied Soft Computing 12(2), 800 – 806 (2012)

[5] Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). Comput. Vis. Image Underst. 110(3), 346–359 (2008)

[6] Beloufa, F., Chikh, M.: Design of fuzzy classifier for diabetes disease using modified artificial bee colony algorithm. Computer Methods and Programs in Biomedicine 112(1), 92 – 103 (2013)

[7] Berlanga, F., Rivera, A., del Jesus, M., Herrera, F.: Gp-coach: Genetic programming-based learning of {COmpact} and {ACcurate} fuzzy rule-based classification systems for high-dimensional problems. Information Sciences 180(8), 1183 – 1200 (2010)

[8] Bradski, G.: The opencv library. Doctor Dobbs Journal 25(11), 120–126 (2000)

[9] Brest, J., Greiner, S., Bošković, B., Mernik, M., Bošković, V.: Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems. IEEE Transactions on Evolutionary Computation 10(6), 646–657 (2006)

[10] Brest, J., Maučec, M.S., Bošković, B.: il-shade: Improved l-shade algorithm for single objective real-parameter optimization. In: 2016 IEEE Congress on Evolutionary Computation (CEC), pp. 1188–1195. IEEE (2016)

[11] Brest, J., Maučec, M.S., Bošković, B.: Single objective real-parameter optimization: Algorithm jSO. In: 2017 IEEE Congress on Evolutionary Computation (CEC), pp. 1311–1318. IEEE (2017)

[12] Casillas, J., Cordon, O., del Jesus, M.J., Herrera, F.: Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. IEEE Transactions on Fuzzy Systems 13(1), 13–29 (2005)

[13] Casillas, J., Cordón, O., Herrera, F.: Learning fuzzy rules using ant colony optimization algorithms. In: Abstract proceedings of ANTS2000 From Ant Colonies to Arti Ants: A Series of International Workshops on Ant Algorithms, pp. 13–21 (2000)

[14] Chang, T., Kuo, C.C.: Texture analysis and classification with tree-structured wavelet transform. Image Processing, IEEE Transactions on 2(4), 429–441 (1993). DOI 10.1109/83.242353

[15] Cordón, O.: A historical review of evolutionary learning methods for mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems. International Journal of Approximate Reasoning 52(6), 894 – 913 (2011)

[16] Cordón, O., Gomide, F., Herrera, F., Hoffmann, F., Magdalena, L.: Ten years of genetic fuzzy systems: current framework and new trends. Fuzzy Sets and Systems 141(1), 5 – 31 (2004)

[17] Das, S., Abraham, A., Chakraborty, U.K., Konar, A.: Differential Evolution Using a Neighborhood-Based Mutation Operator. IEEE Transactions on Evolutionary Computation 13(3), 526–553 (2009)

[18] Das, S., Mullick, S.S., Suganthan, P.N.: Recent advances in differential evolution – An updated survey. Swarm and Evolutionary Computation 27, 1–30 (2016)

[19] Datar, M., Immorlica, N., Indyk, P., Mirrokni, V.S.: Locality-sensitive hashing scheme based on p-stable distributions. In: Proceedings of the Twentieth Annual Symposium on Computational Geometry, SCG ’04, pp. 253–262. ACM, New York, NY, USA (2004)

[20] Elhag, S., Fernández, A., Bawakid, A., Alshomrani, S., Herrera, F.: On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on intrusion detection systems. Expert Systems with Applications 42(1), 193 – 202 (2015)

[21] Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International Journal of Computer Vision 88(2), 303–338 (2010)

[22] Fernández, A., Herrera, F.: Evolutionary Fuzzy Systems: A Case Study in Imbalanced Classification, pp. 169–200. Springer International Publishing, Cham (2016)

[23] Francos, J., Meiri, A., Porat, B.: A unified texture model based on a 2-d wold-like decomposition. Signal Processing, IEEE Transactions on 41(8), 2665–2678 (1993). DOI 10.1109/78.229897

[24] Freischlad, M., Schnellenbach-Held, M.: A machine learning approach for the support of preliminary structural design. Advanced Engineering Informatics 19(4), 281 – 287 (2005)

[25] Freischlad, M., Schnellenbach-Held, M., Pullmann, T.: Evolutionary generation of implicative fuzzy rules for design knowledge representation. In: I. Smith (ed.) Intelligent Computing in Engineering and Architecture, Lecture Notes in Computer Science, vol. 4200, pp. 222–229. Springer Berlin Heidelberg (2006)

[26] Gorzałczany, M.B., Rudziński, F.: Interpretable and accurate medical data classification – a multi-objective genetic-fuzzy optimization approach. Expert Systems with Applications 71, 26 – 39 (2017)

[27] Grauman, K., Darrell, T.: Efficient image matching with distributions of local invariant features. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 2, pp. 627–634 vol. 2 (2005). DOI 10.1109/CVPR.2005.138

[28] Guo, S.M., Tsai, J.S.H., Yang, C.C., Hsu, P.H.: A self-optimization approach for L-SHADE incorporated with eigenvector-based crossover and successful-parent-selecting framework on CEC 2015 benchmark set. In: 2015 IEEE Congress on Evolutionary Computation (CEC), pp. 1003–1010. IEEE (2015)

[29] Huang, J., Kumar, S., Mitra, M., Zhu, W.J., Zabih, R.: Image indexing using color correlograms. In: Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on, pp. 762–768 (1997). DOI 10.1109/CVPR.1997.609412

[30] Jagadish, H.V.: A retrieval technique for similar shapes. SIGMOD Rec. 20(2), 208–217 (1991)

[31] Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using gabor filters. Pattern Recognition 24(12), 1167 – 1186 (1991)

[32] Kauppinen, H., Seppanen, T., Pietikainen, M.: An experimental comparison of autoregressive and fourier-based descriptors in 2d shape classification. Pattern Analysis and Machine Intelligence, IEEE Transactions on 17(2), 201–207 (1995). DOI 10.1109/34.368168

[33] Kiranyaz, S., Birinci, M., Gabbouj, M.: Perceptual color descriptor based on spatial distribution: A top-down approach. Image Vision Comput. 28(8), 1309–1326 (2010)

[34] Koshiyama, A., Escovedo, T., Dias, D., Vellasco, M., Tanscheit, R.: Gpf-class: A genetic fuzzy model for classification. In: Evolutionary Computation (CEC), 2013 IEEE Congress on, pp. 3275–3282 (2013). DOI 10.1109/CEC.2013.6557971

[35] Krömer, P., Platos, J.: Simultaneous prediction of wind speed and direction by evolutionary fuzzy rule forest. In: International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland, pp. 295–304 (2017). DOI 10.1016/j.procs.2017.05.195. URL https://doi.org/10.1016/j.procs.2017.05.195

[36] Krömer, P., Prauzek, M., Stankuš, M., Konečnỳ, J.: Adaptive fuzzy video compression control for advanced driver assistance systems. In: 2018 26th International Conference on Systems Engineering (ICSEng), pp. 1–9. IEEE (2018)

[37] Liang, S., Kuo, C., Shaw, F., Chen, Y., Hsu, C., Chen, J.: Combination of expert knowledge and a genetic fuzzy inference system for automatic sleep staging. IEEE Transactions on Biomedical Engineering 63(10), 2108–2118 (2016). DOI 10.1109/TBME.2015.2510365

[38] Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)

[39] Mallipeddi, R., Suganthan, P.N., Pan, Q.K., Tasgetiren, M.F.: Differential evolution algorithm with ensemble of parameters and mutation strategies. Applied Soft Computing 11(2), 1679–1696 (2011)

[40] Marinaki, M., Marinakis, Y., Stavroulakis, G.E.: Fuzzy control optimized by a multi-objective differential evolution algorithm for vibration suppression of smart structures. Computers & Structures 147, 126–137 (2015)

[41] Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing 22(10), 761 – 767 (2004). British Machine Vision Computing 2002

[42] Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. International Journal of Computer Vision 60(1), 63–86 (2004)

[43] Mininno, E., Neri, F., Cupertino, F., Naso, D.: Compact Differential Evolution. IEEE Transactions on Evolutionary Computation 15(1), 32–54 (2011)

[44] Nguyen, T., Khosravi, A., Creighton, D., Nahavandi, S.: Classification of healthcare data using genetic fuzzy logic system and wavelets. Expert Systems with Applications 42(4), 2184 – 2197 (2015)

[45] Nister, D., Stewenius, H.: Scalable recognition with a vocabulary tree. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2, CVPR ’06, pp. 2161–2168. IEEE Computer Society, Washington, DC, USA (2006)

[46] Opara, K.R., Arabas, J.: Differential Evolution: A survey of theoretical analyses. Swarm and Evolutionary Computation (2018). . DOI: https://doi.org/10.1016/j.swevo.2018.06.010

[47] Pass, G., Zabih, R.: Histogram refinement for content-based image retrieval. In: Applications of Computer Vision, 1996. WACV ’96., Proceedings 3rd IEEE Workshop on, pp. 96–102 (1996). DOI 10.1109/ACV.1996.572008

[48] Philbin, J., Chum, O., Isard, M., Sivic, J., Zisserman, A.: Object retrieval with large vocabularies and fast spatial matching. In: Computer Vision and Pattern Recognition, 2007. CVPR ’07. IEEE Conference on, pp. 1–8 (2007)

[49] Piotrowski, A.P.: aL-SHADE optimization algorithms with population-wide inertia. Information Sciences (2018)

[50] Piotrowski, A.P., Napiorkowski, J.J.: Some metaheuristics should be simplified. Information Sciences 427, 32–62 (2018)

[51] Piotrowski, A.P., Napiorkowski, J.J.: Step-by-step improvement of JADE and SHADE-based algorithms: Success or failure? Swarm and Evolutionary Computation (2018). . DOI: https://doi.org/10.1016/j.swevo.2018.03.007

[52] Pratama, M., Pedrycz, W., Webb, G.I.: An incremental construction of deep neuro fuzzy system for continual learning of non-stationary data streams. CoRR abs/1808.08517 (2018). URL http://arxiv.org/abs/1808.08517

[53] Prauzek, M., Krömer, P., Rodway, J., Musilek, P.: Differential evolution of fuzzy controller for environmentally-powered wireless sensors. Applied Soft Computing 48, 193–206 (2016)

[54] Qin, A.K., Huang, V.L., Suganthan, P.N.: Differential Evolution Algorithm With Strategy Adaptation for Global Numerical Optimization. IEEE Transactions on Evolutionary Computation 13(2), 398–417 (2009)

[55] Qu, B.Y., Suganthan, P.N., Liang, J.J.: Differential Evolution With Neighborhood Mutation for Multimodal Optimization. IEEE Transactions on Evolutionary Computation 16(5), 601–614 (2012)

[56] Rey, M., Galende, M., Fuente, M., Sainz-Palmero, G.: Multi-objective based fuzzy rule based systems (frbss) for trade-off improvement in accuracy and interpretability: A rule relevance point of view. Knowledge-Based Systems 127, 67 – 84 (2017)

[57] Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: Orb: An efficient alternative to sift or surf. In: Computer Vision (ICCV), 2011 IEEE International Conference on, pp. 2564–2571 (2011). DOI 10.1109/ICCV.2011.6126544

[58] Rudziński, F.: A multi-objective genetic optimization of interpretability-oriented fuzzy rule-based classifiers. Applied Soft Computing 38, 118 – 133 (2016)

[59] Rutkowski, L.: Computational Intelligence Methods and Techniques. Springer Berlin Heidelberg (2008)

[60] Schapire, R.E.: A brief introduction to boosting. In: Proceedings of the 16th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI’99, pp. 1401–1406. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1999)

[61] Scherer, R.: Designing boosting ensemble of relational fuzzy systems. International Journal of Neural Systems 20(5), 381–388 (2010)

[62] Scherer, R.: Multiple Fuzzy Classification Systems. Springer Publishing Company, Incorporated (2014)

[63] Simo-Serra, E., Trulls, E., Ferraz, L., Kokkinos, I., Fua, P., Moreno-Noguer, F.: Discriminative learning of deep convolutional feature point descriptors.

In: Proceedings of the IEEE International Conference on Computer Vision, pp. 118–126 (2015)

[64] Sivic, J., Zisserman, A.: Video google: a text retrieval approach to object matching in videos. In: Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on, pp. 1470–1477 vol.2 (2003)

[65] Śmietański, J., Tadeusiewicz, R., Łuczyńska, E.: Texture analysis in perfusion images of prostate cancer—a case study. International Journal of Applied Mathematics and Computer Science 20(1), 149–156 (2010)

[66] Storn, R., Price, K.: Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. Journal of Global Optimization 11, 341–359 (1997)

[67] Tanabe, R., Fukunaga, A.: Success-history based parameter adaptation for Differential Evolution. In: 2013 IEEE Congress on Evolutionary Computation (CEC), pp. 71–78. IEEE (2013)

[68] Tanabe, R., Fukunaga, A.: How far are we from an optimal, adaptive de? In: J. Handl, E. Hart, P.R. Lewis, M. López-Ibáñez, G. Ochoa, B. Paechter (eds.) Parallel Problem Solving from Nature – PPSN XIV, pp. 145–155. Springer International Publishing, Cham (2016)

[69] Tanabe, R., Fukunaga, A.S.: Improving the search performance of SHADE using linear population size reduction. In: 2014 IEEE Congress on Evolutionary Computation (CEC), pp. 1658–1665. IEEE (2014)

[70] Tieu, K., Viola, P.: Boosting image retrieval. Int. J. Comput. Vision 56(1-2), 17–36 (2004)

[71] Tsakiridis, N.L., Theocharis, J.B., Zalidis, G.C.: Deco3r: A differential evolution-based algorithm for generating compact fuzzy rule-based classification systems. Knowledge-Based Systems 105, 160–174 (2016)

[72] Tsakiridis, N.L., Theocharis, J.B., Zalidis, G.C.: Deco3rum: A differential evolution learning approach for generating compact mamdani fuzzy rule-based models. Expert Systems with Applications 83, 257–272 (2017)

[73] Veltkamp, R.C., Hagedoorn, M.: State of the art in shape matching. In: M.S. Lew (ed.) Principles of Visual Information Retrieval, pp. 87–119. Springer-Verlag, London, UK, UK (2001)

[74] Viktorin, A., Senkerik, R., Pluhacek, M., Kadavy, T., Zamuda, A.: Distance based parameter adaptation for differential evolution. In: 2017 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 1–7. IEEE (2017)

[75] Viktorin, A., Senkerik, R., Pluhacek, M., Kadavy, T., Zamuda, A.: Distance Based Parameter Adaptation for Success-History based Differential Evolution. Swarm and Evolutionary Computation (Available online 12 November 2018). DOI 10.1016/j.swevo.2018.10.013

[76] Viktorin, A., Senkerik, R., Pluhacek, M., Zamuda, A.: Steady success clusters in Differential Evolution. In: 2016 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 1–8. IEEE (2016)

[77] Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. I–511–I–518 vol.1 (2001)

[78] Voloshynovskiy, S., Diephuis, M., Kostadinov, D., Farhadzadeh, F., Holotyak, T.: On accuracy, robustness, and security of bag-of-word search systems. In: IS&T/SPIE Electronic Imaging, pp. 902,807–902,807. International Society for Optics and Photonics (2014)

[79] Wu, G., Shen, X., Li, H., Chen, H., Lin, A., Suganthan, P.N.: Ensemble of differential evolution variants. Information Sciences 423, 172–186 (2018)

[80] Yi, K.M., Trulls, E., Lepetit, V., Fua, P.: Lift: Learned invariant feature transform. In: European Conference on Computer Vision, pp. 467–483. Springer (2016)

[81] Zamuda, A., Brest, J.: Self-adaptive control parameters' randomization frequency and propagations in differential evolution. Swarm and Evolutionary Computation 25, 72–99 (2015)

[82] Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: A comprehensive study. In: Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on, pp. 13–13 (2006). DOI 10.1109/CVPRW.2006.121

[83] Zhang, J., Sanderson, A.C.: JADE: Adaptive Differential Evolution With Optional External Archive. IEEE Transactions on Evolutionary Computation 13(5), 945–958 (2009)

[84] Zhang, W., Yu, B., Zelinsky, G., Samaras, D.: Object class recognition using multiple layer boosting with heterogeneous features. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 2, pp. 323–330 vol. 2 (2005). DOI 10.1109/CVPR.2005.251

[85] Zitnick, C., Dollar, P.: Edge boxes: Locating object proposals from edges. In: D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (eds.) Computer Vision – ECCV 2014, Lecture Notes in Computer

Marcin Korytkowski received the Ph.D. and degree in computer science from Częstochowa University of Technology, Poland, in 2007. Currently he is an associate professor at Częstochowa University of Technology. He has published over 70 technical papers. His present research interests include deep learning architectures and their applications in computer security, databases and image retrieval. He is involved in organization of various scientific events and conferences.

Roman Senkerik was born in Zlin, the Czech Republic, in 1981. He received his M.Sc. degree in technical cybernetics from the Tomas Bata University in Zlin, Faculty of applied informatics in 2004, his Ph.D. degree also in technical Cybernetics, in 2008, from the same university, and the Assoc. prof. Degree in Informatics from VSB - Technical University of Ostrava, in 2013. From 2008 to 2013 he was a Research Assistant and Lecturer with the Tomas Bata University in Zlin, Faculty of Applied Informatics. Since 2014 he has been an Associate Professor and Head of the A.I.Lab with the Department of Informatics and Artificial Intelligence, Tomas Bata University in Zlin. He is an author of more than 40 journal papers, 250 conference papers, and several book chapters as well as editorial notes. His research interests are soft computing methods and their interdisciplinary applications in optimization and cyber-security, development of evolutionary algorithms, machine learning, data science, the theory of chaos, and complex systems. He is a recognized reviewer for many leading journals in computer science/ computational intelligence. He was a part of the organizing teams for special sessions/symposiums at IEEE WCCI, CEC, or SSCI events. He was a guest editor of several special issues in journals, and editor of proceedings for several conferences.

Mirosław Kordos obtained his M.Sc. in electrical engineering from the Technical University of Łódź, Poland in 1994 and his Ph.D. in computer science from Silesian University of Technology, Poland in 2005. In years 1994-2005 he worked in the IT industry as a software developer and systems engineer. In 2006-2007 he was a research fellow 18 at the Division of Biomedical Informatics in the Cincinnati Children's Hospital Research Center, USA. In 2008-2009 he was an assistant professor at the Silesian University of Technology, Poland and since 2010 he has been an assistant professorat the University of Bielsko-Biała, Poland. His research interest comprise various aspects of artificial intelligence, especially neural networks and industrial applications of machine learning.

Magdalena Scherer received her M.Sc. degree in computer science from the Częstochowa University of Technology, Poland, in 2008 and her Ph.D. in management in 2016 from the same university. Currently, she is an assistant professor at Częstochowa University of Technology. Her present research interests include machine learning for prediction and classification.

Rafał A. Angryk received his M.S. and Ph.D. degrees in Computer Science in 2004 from Tulane University. Dr. Angryk also holds two other master's degrees, which he obtained from universities in Poland. In 1999, Dr. Angryk received a M.A. in Business Management from the University of Szczecin and three years later a M.Eng. degree in Computer Science from the Technical University of Szczecin. Dr. Angryk is a 2CI Professor of Computer Science at Georgia State University (GSU). He also holds two professor by courtesy appointments at the Physics and Astronomy Department and in the Institute for Insight (in J. Mack Robinson College of Business). Before joining GSU in August 2013, he spent almost a decade as a faculty in Montana State University, where his interdisciplinary and collaborative research on mining massive repositories of solar data was initiated. Dr. Angryk is the founding director of MSU/GSU Data Mining Laboratory, and GSU's M.Sc. Program in Data Science and Analytics. Dr. Angryk's research and teaching interests lie in the areas of very large databases (image and spatio-temporal databases, and kNN indexing of high dimensional data), data mining (frequent patterns discovery, clustering and classification of real-life large-scale data), and information retrieval (text and image data). He has published over 160 journal articles, book chapters and peer-reviewed conference papers in these areas. His research has been sponsored by the federal agencies: NASA, NSF, NGA, and industry: Intergraph Corporation, RightNow Technologies (now Oracle), with the successful grant history exceeding $13M. (http://grid.cs.gsu.edu/~rangryk/)

Agnieszka Siwocha received M.Sc. the degree from Łódź University of Technology, Faculty of Technical Physics, Computer Science and Applied Mathematics, and her a Ph.D. in 2015 in the field of computer science, computer graphics at the University of Social Sciences, Łódź, Poland. She is currently a assistant professor at the Social Academy of Sciences in Łódź. Author of over 20 publications related to various problems of computer science, computer graphics and IT applications.