Yang ZHANG[1], Weiming LIU[1]

South China University of Technology, School of Civil Engineering and Transportation (1)

A novel pedestrian detection method based on Cost-Sensitive Support Vector Machine and Chaotic Particle Swarm Optimization with T mutation

***Abstract**. This paper presents a novel pedestrian detection method based on chaotic particle swarm optimization with T mutation (CTPSO) and cost-sensitive support vector machine (CS-SVM). In order to solve the problem of class-imbalanced in pedestrian detection, a new improve SVM named CS-SVM is proposed, which is based on the idea of assigning different weights to the errors of the two classes when the numbers of data samples from each class are imbalanced. In addition, a new type of PSO called CTPSO is used to select suitable parameters of CS-SVM, which could improve the classification ability of CS-SVM prominently. CTPSO is a novel optimization algorithm, which not only has strong global search capability but also helps to find the optimum quickly by using chaos queues and T mutation. The experiment carried out on videos from INRIA, MIT and Daimler datasets, result indicates that the effectiveness and efficiency of the proposed method, which can achieve higher accuracy than other three state of the art algorithms.*

***Streszczenie.** Przedstawiono nową metode detekcji pieszych bazującą na algorytmie mrówkowym z mutacją T oraz mechanizmie SVM. Zaproponowano nowy algorytm CS-SVM polegający na przyporządkowaniu różnych wag błędów w dwóch klasach kiedy liczba próbek w każdej klasie jest nierówna. Optimum znajdowane jest szybko przy wykorzystaniu mutacji T. Przeprowadzono eksperymenty bazujące na różnych bazach danych. (**Nowa metoda detekcji pieszych bazująca na mechanizmie SVM i algorytmie mrówkowym z mutacją T**)*

Keywords: Cost-Sensitive SVM, Chaotic PSO, T mutation, Pedestrian detection.
Słowa kluczowe: SVM, algorytm mrówkowy, wykrywanie pieszych

Introduction

Pedestrian detection has very important applications in video surveillance, robotics and intelligent vehicles. However, because of variable appearance and the wide range of poses, pedestrian detection is still a challenging task. Recently, more and more machine learning methods are used to solve pedestrian detection problem [1].Support vector machine (SVM) is a method based on machine learning, which can find global optimum solutions for problems with small training samples, high dimensions, non-linear. SVM is applied in pedestrian detection field has achieved certain success. However, the largest problem encountered in constructing the SVM model is how to select the training parameter values of SVM, and inappropriate parameter settings lead to poor classification results [2].

In order to solve the above problem, particle swarm optimization (PSO) has been applied to select the proper parameters of SVM, but the method is easy to trap into local optimum. Chaos particle swarm optimization (CPSO) is a kind of improved particle swarm optimization, which can not only avoid the search being trapped in local optimum and but also help to search the optimum quickly by using chaos queues. In addition, the simultaneous perturbation method is a kind of stochastic gradient method. The scheme can obtain the local information without direct calculation of the gradient. Combination of PSO and the simultaneous perturbation optimization also avail to search the global optimum [3].

In this paper, we combine the simultaneous perturbation strategy based on T mutation with Chaos particle swarm optimization. This new type of PSO called CTPSO. Then, we improve standard SVM to the CS-SVM, which is more suitable to solve pedestrian detection problem, and CTPSO is applied to the parameter optimization of CS-SVM. The rest of this paper is organized as follows. Standard PSO is described in Section 2. PSO based on chaos queues and T mutation is derived in Section 3. The CS-SVM is introduced in Section 4. The steps of the optimal parameters selection of CS-SVM by CTPSO are arranged in Section 5. An experiment in pedestrian detection is given in Section 6. Section 7 draws the conclusions.

Standard Particle Swarm Optimization

Particle swarm optimization (PSO) is inspired by social behavior among individuals like the birds blocking or the fish grouping [3]. It uses a set of particles search the optimal position through coexist and collaborate among the individuals. Suppose that a particle swarm consists of m particles and the search space is n-dimensional, therefore, the position of each particle can be represented by an n-dimensional vector, $x_i = (x_{i1}, x_{i2},\ldots, x_{in})$ and the velocity of every particle can be represented as $v_i = (v_{i1}, v_{i2},\ldots, v_{in})$. In PSO system, every particle remembers its previous best position and knows the best position of the whole swarm. Then, the velocity of every particle is constantly adjusted according to the above information. The velocity of particle and its new position can be assigned by using the following equations:

$$v_i^{t+1} = \omega v_i^t + c_1 r_1 (pbest_i^t - x_i^t) + c_2 r_2 (gbest_i^t - x_i^t) \quad (1)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (2)$$

where t is the iteration counter, ω is called inertia weight, c_1 and c_2 are two positive constant parameters called acceleration coefficients, r_1 and r_2 are random variables with range [0, 1], $pbest_i^t$ is the individual previous best position and $gbest_i^t$ is the whole swarm global best position.

Chaotic T Mutation Particle Swarm Optimization

T Mutation Particle Swarm Optimization

PSO has gained much attention and widespread applications in more and more fields. However, the performance of the standard PSO greatly depends on its parameters, and unsatisfactory parameters often lead particle to be trapped in local optimum [2]. In order to overcome this shortage, Cauchy mutation PSO is applied to regulate the inertia weight of velocity on the fitness value of object function [4]. The new method can improve the capacity of the global search of the PSO algorithm and obtain highly efficient on the conditions of real number code. However, this method also has some inherent defects. Because too small central part of Cauchy distribution, the fine-tuning performance of Cauchy mutation PSO is very poor. Therefore, in this method, particle local exploitation ability is relatively low, which hinder particle to find the optimization result. T mutation is another famous evolutionary programming, which has a larger hill around the center [10]. For this reason, its fine-tuning ability is better than Cauchy mutation. Moreover, by adjusting the free degree in the distribution, T mutation can tune the probability density. In this way, T distribution can get broad level direction and large central part simultaneously. Hence, the individuals not only could search better position in biggish scope and discard local best solution easily but also have better fine-tuning ability, which could improve their ability to find optimization result. From this point of view, in this paper, we try to apply T operator to improve PSO algorithm.

The probability density function of the T distribution can be defined as follows:

(3) $$f(x) = \frac{1}{B(0.5, 0.5n)\sqrt{n}}(1 + \frac{x^2}{n})^{-\frac{n+1}{2}}$$

Where n is a free degree parameter, which can change the shape of T distribution. Through adjusting the value of the free degree, T mutation can generate offspring have both good global search ability and well fine-tuning ability synchronously. B is B function that is described as below:

(4) $$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

Then, we make T mutation for the previous velocity of the particle, and the T mutation PSO algorithm can be represent as follows:

(5) $$v_{ij}^{t+1} = (1-\mu)\omega_{ij}^t v_{ij}^t + \mu(\eta_i(j) \cdot T(n_i^t)) + c_1 r_1 (pbest_{ij}^t - x_{ij}^t) + c_2 r_2 (gbest_{ij}^t - x_{ij}^t)$$

(6) $$x_{ij}^{t+1} = x_{ij}^t + v_{ij}^{t+1}$$

(7) $$n_i^{t+1} = n_i^t \exp(\tau_1 \cdot \varepsilon + \tau_2 \cdot N(0, \Delta n))$$

where $T(n)$ denotes a T distributed with the n free degree number, $\eta_i(j)$ is standard deviations of T mutation, τ_1 and τ_2 denote the parameter of operator set, ε is T distribution number.

The inertia weight is one of the critical for the performance of PSO, which balances global exploration and local exploitation abilities of the swarm. A big inertia weight facilitates global exploration and a small inertia weight is beneficial to local exploitation. Therefore, when particle is far away from the optimal value, a biggish inertia weight should be given. In contrast, a smaller inertia weight should be endowed since particle is nearby the optimization result. In this paper, feedback takes from how far the distance between particle and the estimated optimal value is used for one of the factor to decide how to adjust the inertia weight value. The feedback value (fv) is defined as follows:

(8) $$fv_i^t = \frac{f(x_i^1) - f_{KN}}{f(x_i^t) - f_{KN} + \varphi}$$

Where $f(x_i^t)$ is the fitness of the ith particle and f_{KN} denotes the known estimated optimal solution value, $\varphi \in (0,1)$ is the coefficient of smoothing, which is used for preventing the denominator to zero.

The value of inertia weight can calculate with the following formula:

(9) $$\omega_{ij}^t = \frac{\lambda}{1 + \exp[-(\beta \cdot fv_i^t)^{-1}]} + (1-\lambda)\omega_{ij}^0 \exp(-\varepsilon t^2)$$

The first item 1 / $\{1+exp[-(\beta \times fv_i^t)^{-1}]\}$ of the above formula represents adjust the value of inertia weight depends on the feedback *fv* and the value of parameter β. A small *fv* means particle is far away from the optimal value and a biggish inertia weight can get from the first item of formula (9). On the other hand, when *fv* is big, particle needs a strong fine-tuning local exploitation, and inertia weight can change small too. The value of parameter β also can control the decreasing speed of inertia weight similar to *fv*. The second item $exp(-\varepsilon t^2)$ of formula (9) denotes the mutation based on the iterative variable *t*. In smaller iterative variable, the particles mutate in big scope and a big inertia weight value obtains to enhance the ability of global exploration. Contrary to above, when iterative variable is bigger, inertia weight value changes smaller to strengthen the performance of local exploitation.

Improve Chaotic T mutation PSO

Although PSO algorithm is simple and easy to implement, it often suffers the problem of trapping into local optimum [2,14]. In order to overcome the above drawback, Cai et al. [5] introduced a chaotic PSO method, which make use of chaotic local search to avoid arithmetic trapping into local minimum and achieved certain success. In this paper, an algorithm named improves chaotic T mutation PSO (CTPSO) is proposed to enhance the ability of search the global optimum.

1. Chaotic initialization

In standard PSO, the particles positions are random, and some of the particles are far away from the optimum position, which is detrimental to fast convergence. In our proposed method, chaotic sequences are used for initializing positions of the particles and make them near the optimum position [15]. The Logistic map is employed to obtain chaos queues, which is described as follow:

(10) $$z_{n+1} = \mu z_n (1 - z_n), \quad n = 0,1,2,\ldots,N$$

Where μ is the control parameter, z_n is the nth chaotic number and n is the iteration number, $0 < z_0 < 1$. $\mu = 4$ is adopted in the experiments.

The steps of chaotic initialization are expounded as follows:

Step 1: An n-dimensional vector $z_1 = (z_{11}, z_{12}, \ldots, z_{1n})$ is generated randomly, and $z_{i+1} = \mu z_i(1-z_i)$, $i = 1,2, \ldots N-1$.

Step 2: Make use of chaos queues $z = (z_1, z_2, \ldots, z_N)$ initialize positions of the particles, map each component of z_i into corresponding optimized range:

(11) $$x_{ij} = p_{\min} + (p_{\max} - p_{\min})z_{ij}$$

In above equation, P_{min} and P_{max} are the minimum and maximum of the optimized range, respectively, $i = 1,2,\ldots,N$, $j = 1,2,\ldots,n$, and x_i is the initial position of particle.

Step 3: Compute fitness value of the particles x_i and choose the top s particles that are sorted in an ascending order of fitness value. The final initialize particles can denote as $x_i = (x_1, x_2, \ldots, x_s)$.

Step 4: generate s velocities $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})$ Randomly ($i = 1,2, \ldots,s$).

2. Chaotic disturbance

Chaotic disturbance can help particles get out of the local optimum [11]. In this paper, chaotic disturbance is used for helping particles search global optimal solution and avoid local optimum. The procedure of chaotic disturbance is illustrated as below:

Step 1: Randomly generate an n-dimension vector $\tau_0 = (\tau_{01}, \tau_{02}, \ldots, \tau_{0n})$, $\tau_{0i} \in [0,1]$ and $i = 1,2,\ldots,n$. The Logistic map also used to engender chaos queues. Hence, $\tau_{1j} = 4\tau_{0j}(1-\tau_{0j})$, $j = 1,2, \ldots,n$.

Step 2: Suppose chaotic disturbance range is $[-\delta, \delta]$, the disturbance range of particle can be obtained by following equation:

(12) $$\Delta x_j = -\delta + 2\delta\tau_{1j}, \ j = 1,2,\ldots,n$$

(13) $$\Delta x = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)$$

Step 3: compute the positions x_i^{k+1} and Δx_i^{k+1} as follow two expressions respectively. $x_i^{k+1} = x_i^k + v_i^{k+1}$; $\Delta x_i^{k+1} = x_i^k + v_i^{k+1} + \Delta x$.

Step 4: compute the fitting values of the positions x_i^{k+1} and Δx_i^{k+1} respectively. If the fitting values of the position x_i^{k+1} is bigger than the position Δx_i^{k+1}, use Δx_i^{k+1} replace x_i^{k+1}.

Cost-sensitive Support Vector Machine

SVM is a crucial approach to solve classification problems. However, when the numbers of data samples from each class are not balanced, the standard SVM is prone to generating a classifier that has a strong estimation bias toward the majority class, resulting in a large number of misclassification [6,12]. In addition, the optimal parameters of the SVM are difficult to confirm [12,13]. In order to overcome the above shortages, the cost-sensitive SVM (CS-SVM) is used to deal with class-imbalanced problem and the chaotic T mutation PSO algorithm is applied to the parameter optimization of CS-SVM.

CS-SVM is based on the idea of assigning different weights to the errors of the two classes when the numbers of data samples from each class are imbalanced. The cost-sensitive SVM handles this issue by controlling the cost asymmetry between false positives and false negatives [7].

Consider a set of training sample data $\{x_i, y_i\}_i^n$, where $x_i \in R^d$ is a input feature vector, $y_i \in \{-1,1\}$ is a class label. The objective of CS-SVM is to find an optimal separating hyperplane and this hyperplane can be found by solving the following constrained optimization problem:

(14) $$\min_{w,\xi,p} \frac{1}{2}\|w\|^2 - vp + \frac{\gamma}{N}\sum_{i\in I_+}\xi_i + \frac{1-\gamma}{N}\sum_{i\in I_-}\xi_i$$

(15) $$s.t.\quad y_i(w\cdot\varphi(x_i)+b) \ge p-\xi_i \quad i=1,2,\cdots N$$

(16) $$\xi_i \ge 0,\, p \ge 0, \quad i=1,2,\cdots N$$

Where w and x_i are d dimensional column vectors, $\gamma\in[0,1]$ is the cost asymmetry controlling the tradeoff between false negatives and false positives, $v\in[0,1]$ is a parameter that controls the support vector numbers, and ζ_i is slack variables, $I_+=\{i{:}y_i=+1\}$ and $I_-=\{i{:}y_i=-1\}$ denote the numbers of training samples that belong to the positive and negative classes, respectively.

The above optimization problem can solve by its dual, which depends on a set of Lagrange multipliers.

(17) $$\min_{\alpha} \quad \frac{1}{2}\sum_i\sum_j \alpha_i\alpha_j y_i y_j k(x_i,x_j) - \sum_i \alpha_i$$

(18) $$s.t.\quad 0 \le \alpha_i \le \frac{1_{\{y_i<0\}} + y_i\gamma}{N}, \quad \forall i$$

(19) $$\sum_i y_i\alpha_i = 0, \qquad \sum_i \alpha_i \ge v$$

Where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_N)$, the indicator function $1_{\{A\}}$ returns 1 if the condition A is true and 0 otherwise.

We can solve the optimal problem of formula (17) by selecting the optimal solution $\alpha^*(\gamma)= (\alpha_1^*(\gamma), \alpha_2^*(\gamma), \cdots, \alpha_N^*(\gamma))$, and the sign of the decision function of CS-SVM can be expressed as follow:

(20) $$f(x) = sign(\sum_i \alpha_i^*(\gamma) y_i k(x,x_i) + b)$$

The optimal parameters selection of CS-SVM by chaotic T mutation PSO

The parameters of CS-SVM have a great effect on the generalization performance of classify. In order to improve performance of CS-SVM, the intelligent algorithm named chaotic T mutation PSO (CTPSO) is used to search the optimal parameters of the CS-SVM model. CTPSO is an excellent technique to solve the optimization problems.

For valuating classification ability of the proposed algorithm, the fitness function of CTPSO is designed as follow:

(21) $$fitness = \frac{1}{N}\sum_{i=1}^{N}\frac{\gamma f_+ + (1-\gamma)f_-}{f+t}$$

Where N is the number of the sample, f_+ denotes the false positive classification and f_- denotes the false negative classification, t is the correct classification. The steps of the proposed method are described as below:

Step 1: Set up parameters of CTPOS. Initialize positions and velocity of particles and set up the learning parameters v, γ and α, the inertia weight, the maximum number of iterations and T distribution.

Step 2: Compute the fitness values of each chaos particle, the particle with minimal fitness value is the global extreme point.

Step 3: Adopt T mutation operator by Eqs. (7) and (9) to control particle velocity, and update the particle position by formula (5) and (6).

Step 4: The global best gbest is optimized by Chaos initialization and Chaos disturbance, and get the best global extreme point.

Step 5: The procedure proceeds until stopping condition are met, then the global optimum values of v, γ and α are obtained. Otherwise, loop to step 2.

Step 6: End the training procedure, the CS-SVM model is established based on the optimal v, γ and α.

Experiments for pedestrian detection based on CTPSO and CS-SVM algorithm

In this section, we describe the completely pedestrian detection system based on CTPSO and CS-SVM model. Firstly, we extract the pedestrian features from image database. In this study, LABH feature is used to extract the feature vector [8]. Then, CTPSO and CS-SVM model is applied to classify pedestrian features. Finally, Full Binary Tree structure classifier is used to improve system's real-time performance. The process of the pedestrian detection based on CTPSO and CS-SVM is shown in Fig.1.

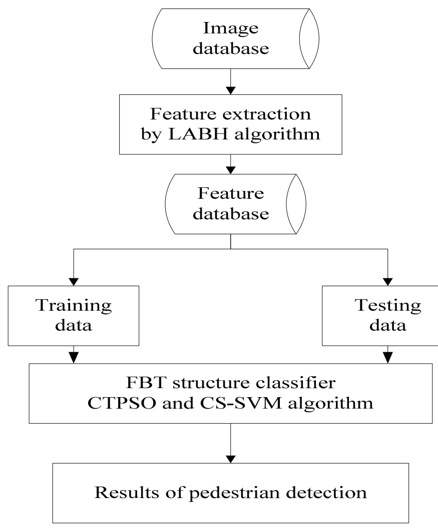

Fig.1 The process of the proposed pedestrian detection method

To evaluate the proposed pedestrian detection method, we use three challenging datasets (the INRIA dataset, the MIT dataset and the Daimler pedestrian dataset [1]) in our experiments. These datasets contains of 16702 pedestrian training samples as well as a test sequence comprising 21822 images (320×240pixels). The training and testing sets in these datasets are well designed. In our work, the training sets contain 9641 pedestrian images and 13523 pedestrian-free images. The testing sets contain 8708 pedestrian images and 10236 pedestrian-free images. The pedestrian-free and pedestrian images in the testing set are used to evaluate the false-positive and true-detection rates, respectively.

In our experiments, Full Binary Tree (FBT) structure classifier is applied to pedestrian detection. The FBT structure has advantages of both series connection structure and parallel connection structure, and brings in a principle of “Early-rejection” to improve system’s real-time performance [9]. We used CTPSO and CS-SVM algorithm in FBT structure for improving accuracy performance of pedestrian detection system. The structure of the FBT is shown in Fig.2.

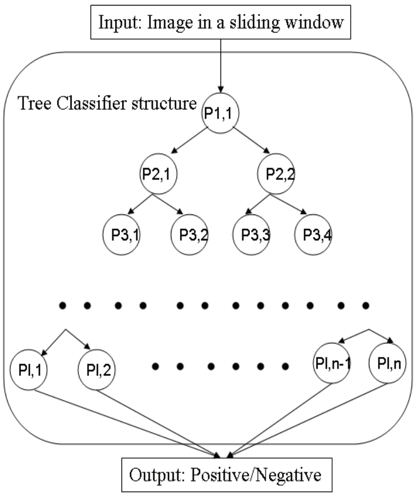

Fig 2. Full Binary Tree structure Classifier

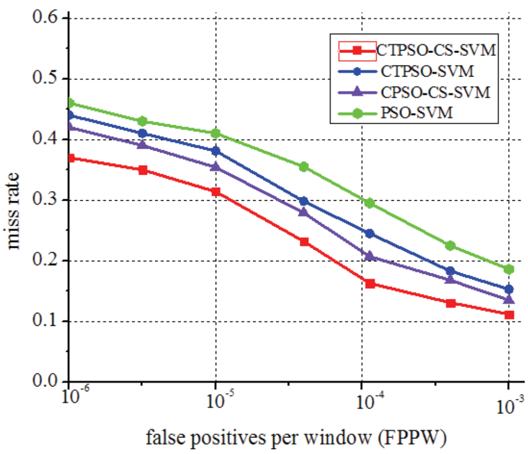

Fig 3. Performance comparison between proposed algorithm and the other three state of the art algorithms

In CTPSO and CS-SVM algorithm, CTPSO is used to choose the optimal parameters of the CS-SVM model. The initial parameters of CTPSO are given as follows: the maximal iterative number k_{max} is 150, inertia w is 0.85, positive acceleration constants c_1 and c_2 are 2, T distribution free degree parameter n is 1.5, increment coefficient μ is 0.1, the coefficient of controlling particle velocity attenuation ε is 2.

In this section, our proposed CTPSO and CS-SVM algorithm is compared with several state of the art algorithms to show the significance of exploring CTPSO and CS-SVM algorithm applied to pedestrian detection. The detection miss rate vs. FPPW is employed to compare propose algorithm with other three state of the art algorithms (CTPSO-SVM, CPSO-CS-SVM and PSO –SVM [2]) on above dataset. Where FPPW is defined as:

$$FPPW = \frac{Number \quad of \quad false \quad positives}{Total \quad number \quad of \quad testing \quad negative \quad windows}$$

Results are shown in Fig.3, performance of the proposed algorithm is the best in pedestrian detection. CS-SVM is better than standard SVM in pedestrian detection problem, and the parameters optimized by CTPSO are of better choice to construct CS-SVM model for the design of pedestrian detection system than the ones by PSO and CPSO.

Conclusion

In this paper, a novel pedestrian detection method based on cost-sensitive SVM and chaotic PSO with T mutation algorithm is presented. Pedestrian detection is a class-imbalanced problem and cost-sensitive SVM is based on the idea of assigning different weights to the errors of the two classes when the numbers of data samples from each class are imbalanced. Then, CTPSO is used to select suitable parameters of CS-SVM, which could improve the classification ability of CS-SVM prominently. CTPSO is a novel optimization algorithm, which not only has strong global search capability, but also helps to find the optimum quickly. Therefore, it is very suitable for parameters selection of CS-SVM. Three challenging datasets is adopted in our experimental, result indicates that the proposed method can achieve higher accuracy than other three state of the art algorithms (CTPSO-SVM, CPSO-CS-SVM and PSO -SVM).

REFERENCES

[1] M. Enzweiler, D. M. Gavrila. Monocular pedestrian detection: Survey and experiments. Journal of IEEE Trans. on Pattern Analysis and Machine Intelligence. 31(12):2179–2195, 2009.

[2] Ren Gang, Zhou Zhuping. Traffic safety forecasting method by particle swarm optimization and support vector machine. Expert Systems with Applications. 38(2011): 10420-10424,

[3] Zhao Chenglin, Sun Vuebin, Sun Songlin, Jiang Ting. Fault diagnosis of sensor by chaos particle swarm optimization algorithm and support vector machine. Expert Systems with Applications. 38(2011): 9908-9912, 2011.

[4] Qi Wu, Rob Law. Cauchy mutation based on objective variable of Gaussian particle swarm optimization for parameters selection of SVM. Expert Systems with Applications. 38(2011): 6405-6411,

[5] Cai Jiejin, Ma Xiaoqian, Li Lixiang, Peng Haipeng. Chaotic particle swarm optimization for economic dispatch considering the generator constraints. Energy Conversion and Management 48 (2007): 645–653, 2007.

[6] Yuchun Tang, Nitesh V. Chawla. SVMs Modeling for Highly Imbalanced Classification. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS. 39(1): 281-288, 2009.

[7] Gyemin Lee, Clayton Scott. Nested Support Vector Machines. IEEE Trans. ON Signal Processing58(3): 1648-1660.

[8] S. Yan, S. Shan, X. Chen, and W. Gao. Locally Assembled Binary (LAB) feature with feature-centric cascade for fast and accurate face detection. In Proceedings of IEEE conference on Computer Vision and Pattern Recognition, pages 1–7, 2008.

[9] Xu Y.W., Cao X.B., Qiao H. An efficient tree classifier ensemble-based approach for pedestrian detection, J. of IEEE Trans. on Systems, Man, Cybernetics. 41(1):107-117, 2011.

[10] Rob W. W. Hooft, Leo H. Straver, Anthony L. Spek. Probability plots based on Student’s t-distribution. Journal of Acta Cryst. A65, 319-321, 2009.

[11] Liu Bo, Wang Ling, Jin Yi Hui, Tang Fang, Huang De Xian. Improved particle swarm optimization combined with chaos. Chaos, Solitons Fractals 25(5):1261–1271, 2005.

[12] Chan-Yun Yang, Jr-Syu Yang, Jian-Jun Wang. Margin calibration in SVM class-imbalanced learning. Neurocomputing. 73 (2009): 397–411, 2009.

[13] Gedikpinar, M. The speed control of DC motors with Support Vector Machine. Przeglad Elektrotechniczny. 87(5):269-271, 2011.

[14] Dems, M, lachecinski, S, Wiak, S. Influence of the constraints on the optimal problem solution using particle swarm algorithm. Przeglad Elektrotechniczny. 86(8):127-130, 2010.

[15] Luo, YF, Li JH, Huang P, Ji SC, Li YM. Chaotic characteristics of time series of partial discharges in oil-paper insulation and their applications in pattern recognition. Przeglad Elektrotechniczny. 87(7): 219-224, 2011.

***Authors**: dr. Yang Zhang, School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510640, China, E-mail: zhang_yang1983@163.com; prof. dr Weiming Liu, School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510640, China, E-mail: zhang_yang1983@163.com.*