

RBF KERNEL OPTIMIZATION METHOD WITH PARTICLE SWARM OPTIMIZATION ON SVM USING THE ANALYSIS OF INPUT DATA'S MOVEMENT

Rarasmaya Indraswari, Agus Zainal Arifin, Darlis Herumurti

Department of Informatics, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember (ITS), Kampus ITS Sukolilo, Surabaya, 60111, Indonesia

E-mail: rarasmaya15@mhs.if.its.ac.id

Abstract

SVM (Support Vector Machine) with RBF (Radial Basis Function) kernel is a frequently used classification method because usually it provides accurate results. The focus of most SVM optimization research is the optimization of the input data, whereas the parameter of the kernel function (RBF), the sigma, which is used in SVM also has the potential to improve the performance of SVM when optimized. In this research, we proposed a new method of RBF kernel optimization with Particle Swarm Optimization (PSO) on SVM using the analysis of input data's movement. This method performed the optimization of the weight of the input data and RBF kernel's parameter at once based on the analysis of the movement of the input data which was separated from the process of determining the margin on SVM. The steps of this method were the parameter initialization, optimal particle search, kernel's parameter computation, and classification with SVM. In the optimal particle's search, the cost of each particle was computed using RBF function. The value of kernel's parameter was computed based on the particle's movement in PSO. Experimental result on Breast Cancer Wisconsin (Original) dataset showed that this RBF kernel optimization method could improve the accuracy of SVM significantly. This method of RBF kernel optimization had a lower complexity compared to another SVM optimization methods that resulted in a faster running time.

Keywords: *parameter, Particle Swarm Optimization, RBF kernel, sigma, Support Vector Machine*

Abstrak

Metode klasifikasi SVM (*Support Vector Machine*) dengan RBF (*Radial Basis Function*) kernel merupakan metode yang sering digunakan karena memberikan hasil klasifikasi yang cukup akurat. Penelitian mengenai optimasi pada SVM sementara ini masih banyak berfokus pada optimasi dari nilai data masukan padahal parameter fungsi kernel (RBF), yaitu parameter *sigma*, yang digunakan pada SVM juga memiliki potensi untuk meningkatkan performa dari SVM apabila dioptimasi. Pada penelitian ini diajukan metode baru optimasi RBF kernel dengan *Particle Swarm Optimization* (PSO) pada SVM berdasar analisis persebaran data masukan. Metode ini melakukan optimasi terhadap bobot data masukan sekaligus parameter RBF kernel berdasarkan analisis persebaran data masukan sehingga terpisah dari proses penentuan margin pada SVM. Tahapan dari metode ini adalah inisialisasi parameter, pencarian partikel optimal, perhitungan nilai parameter kernel, dan klasifikasi dengan SVM. Pada proses pencarian partikel optimal, nilai *cost* dari tiap partikel dihitung berdasar fungsi RBF. Nilai parameter kernel dihitung berdasar pergerakan partikel data masukan pada PSO. Hasil uji coba pada dataset *Breast Cancer Wisconsin (Original)* menunjukkan bahwa metode optimasi RBF kernel mampu meningkatkan akurasi klasifikasi SVM secara cukup signifikan. Metode optimasi parameter RBF kernel ini memiliki kompleksitas yang lebih rendah dibandingkan dengan metode optimasi SVM lainnya sehingga menghasilkan *running time* yang lebih cepat.

Kata Kunci: *parameter, Particle Swarm Optimization, RBF kernel, sigma, Support Vector Machine*

1. Introduction

SVM (Support Vector Machine) classification is a method proposed by Boser, Guyon, and Vapnik in 1992 and often used in various fields such as pattern recognition, bioinformatics, and text categorization [1]. One of the SVM method that is commonly used is SVM with RBF kernel. A lot of

research about SVM with RBF kernel optimization used an optimization methods such as Particle Swarm Optimization (PSO) and Genetic Algorithm (GA). In general, the focus of SVM optimization that was performed is divided in two types, namely the data optimization and parameter optimization.

On the data optimization type, optimization methods are used to determine the weight of the

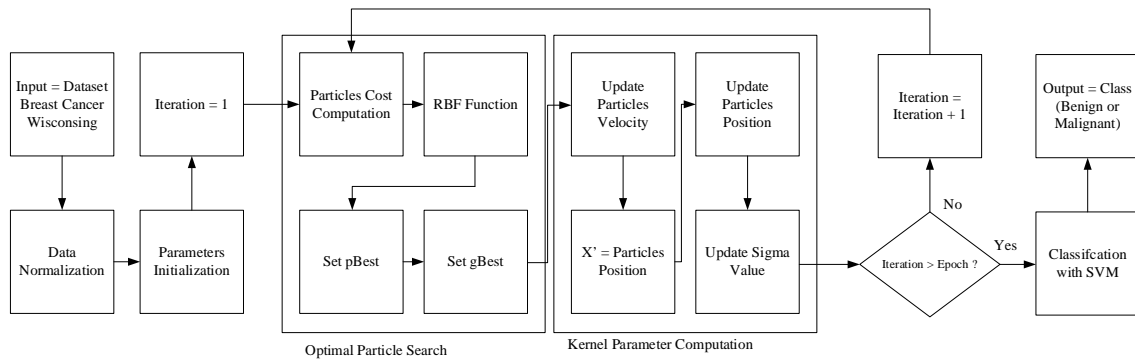


Figure 1. The steps of RBF kernel optimization method with Particle Swarm Optimization on SVM using the analysis of input data's movement.

input data to provide a more optimal result of the SVM classification. In 2014, Devos et. al was using a Genetic Algorithm to perform the optimization of input data which was performed on olive oils data [2]. Fitness value that was used is the error of the k-fold cross validation applied to the SVM.

On the parameter optimization type, researches that were conducted usually about selecting the combination of SVM margin parameter (C) with a kernel parameter (σ) that can provide the most optimal result of SVM classification. So that on the usual optimization process, the input used the combination of SVM margin parameter and kernel parameter, and the output is the accuracy of SVM classification with k-fold cross validation method. Grid algorithm is an alternative method to find the optimal combination of parameter C and σ on SVM with RBF kernel. However, this method required a long computational time and sometimes it did not go well [3] [4]. On of the research about the optimization of parameter C and σ on SVM with RBF kernel was done by Ding and Li in 2009 which showed that the SVM parameter optimization with PSO can improve the accuracy of SVM classification result [5].

Those two types of SVM optimization should have been combined so that in one optimization process, the data optimization and the SVM parameter optimization could be conducted simultaneously. In 2006, Cheng-Lung Huang and Chieh-Jen Wang proposed an SVM parameter C and σ optimization method using Genetic Algorithm. This method also performed the input data optimization, which is the selection of the features of the dataset using GA [1]. Features subset from the dataset became the part of the GA chromosom together with the values of parameter C and σ . In 2008, Lin et al. optimized the parameters of SVM and did feature selection using PSO [6]. However, the method proposed in both of those researchs called the SVM classification

process in many times that resulted on a high complexity.

In this study we proposed a new method of RBF kernel optimization with Particle Swarm Optimization (PSO) on SVM using the analysis of input data's movement. This method performed the optimization of the input data values and the RBF kernel parameter (σ) at once based on the analysis of the input data's movement so that it was separated from the process of determining the margin on SVM. Classification process with SVM was called once after the optimal value of the data and the kernel's parameter was obtained so that the complexity of this RBF kernel optimization method is not increased too much compared with the usual SVM with RBF kernel method. PSO was selected to do the optimization process because PSO method is represent a swarm of data group, while the RBF kernel shape is circular so the PSO is suitable to be applied to the RBF kernel optimization.

Particle Swarm Optimization

PSO was introduced by Kennedy and Eberhart to imitate social behavior of animals such as birds flocking in searching for food [7]. Each particle flies in hyperspace searching for the best solution by adjusting position and velocity based on its own flying experience ($pBest$) and its companions' experience ($gBest$). Each particle has a fitness value or cost which was evaluated using the fitness function to be optimal, position, and velocity that controls the movement of the particles. The inertia weight w was later introduced to improve the PSO optimizer [8]. The steps of PSO are as follow :

- 1) Initialization of PSO parameters.
- 2) Compute the cost of each particle using the fitness function.
- 3) Search the $pBest$ dan $gBest$ values.
- 4) Update the particle velocity v_i using equation(1)

$$v_i(t) = wv_i(t-1) + c_1r_1(x_{pBest_i}(t) - x_i(t)) + c_2r_2(x_{gBest}(t) - x_i(t)) \quad (1)$$

where the c_1 and c_2 is a variable value namely *correction factor*, r_1 and r_2 is a random variable which value is between 0 and 1.

- 5) Update the particle position x_i using equation (2).

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

RBF Kernel

Kernel function is used to change the manufacturing process SVM models that are linear to non-linear computing without overly complicated. RBF kernel is the kernel that can generally be used for all types of data. It uses a Gaussian kernel function RBF to get the inner product of x and x' using equation (3)

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (3)$$

where $\|x - x'\|$ is the Euclidean Distance of the data values in two different feature space and σ (*sigma*) is a free parameter on RBF kernel which determine the weight of the kernel. In SVM, parameter σ need to be adjusted to provide a more accurate classification result. The default value of σ is 1. In RBF function, we can also used parameter *gamma* which value is $\gamma = \frac{1}{2\sigma^2}$.

2. Methods

RBF kernel optimization methods with Particle Swarm Optimization (PSO) on SVM using the analysis of input data's movement has four main steps, which are parameters initialization, optimal particle search, kernel's parameter calculation, and classification with SVM. The details of the stages of this method is shown in Figure 1.

Parameters Initialization

In this step, we initialize the PSO parameters which are inertia w , correction factor c , and the maximal number of iteration (*epoch*). In this research, the value of those parametes are $w=0.4$, $c=0.7$, dan *epoch*=100 which are the optimal value based on the experiment that had been conducted. We also set the initial velocity $v=0$.

In this step, we also set the initial value of RBF kernel parameter that will be optimized, σ , as 0.01. The value of data matrix x and the value of

the data in the new feature space x' is initialized to be equal with the value of the input data. The value of x' is the particles position in PSO. We also performed the *pBest* value initialization of each particle, where the initial *pBest* should be a big number because the initial *pBest* value must be substituted by the cost that obtained in the first iteration.

Optimal Particle Search

After the parameters were initialized, we calculate the cost of each particle. Fitness function that is used is the RBF kernel function in equation (3) with the value of x is always fixed, the value of x' is the particle position in PSO, and the σ value is the value of the parameter σ at the current iteration.

After obtained the cost of each particle, we search the new local optimal value (*pBest*) from each particle i . The new *pBest* value ($pBest(t)$) was obtained by comparing the value of the previous *pBest* ($pBest(t-1)$) with the cost of the particle at iteration (t). To get the value of the new *pBest* we use equation (4).

$$pBest(t)_i = \min(pBest(t-1)_i, cost_i) \quad (4)$$

From the all *pBest* values that have been obtained, we search the minimum value of *pBest*. This minimum *pBest* value is the value of *gBest*

TABLE 1
EXPERIMENTAL RESULT ON INERTIA VALUE

Epoch	Optimized Kernel Parameter σ	Accuracy (%)
0.1	0.475	94.3
0.2	0.558	92.9
0.3	0.669	92.9
0.4	0.809	95.0
0.5	0.992	94.3

TABLE 2
EXPERIMENTAL RESULT ON CORRECTION FACTOR VALUE

Correction Factor	Optimized Kernel Parameter σ	Accuracy (%)
0.1	0.074	64.3
0.3	0.256	85.0
0.5	0.520	94.3
0.7	0.809	95.0
0.9	1.073	94.3

TABLE 3
EXPERIMENTAL RESULT ON EPOCH VALUE

Epoch	Optimized Kernel Parameter σ	Accuracy (%)
50	0.311	90.0
100	0.809	95.0
150	1.250	94.3
200	1.596	93.6
250	1.785	94.3

TABLE 4
COMPARISON WITH OTHER METHODS

Fold	RBF Kernel Optimization Method		SVM (sigma = 1)		SVM Parameters Optimization [5]	
	Accuracy (%)	Time (s)	Accuracy (%)	Time (s)	Accuracy (%)	Time (s)
2	71.75	0.970	65.35	0.110	69.40	10.490
5	71.74	0.796	67.66	0.076	71.70	14.542
10	71.16	0.720	65.18	0.063	73.27	15.513
Average	71.55	0.830	66.06	0.080	71.46	13.520

TABLE 5
ANALYSIS OF THE EFFECT OF DATA OPTIMIZATION PROCESS ON NUMBER OF FOLD = 2

K-	Optimized σ Value	Accuracy of RBF Kernel Optimization Method (%)	Accuracy of SVM with Optimized σ Value (%)
1	0.470	71.40	71.40
2	1.755	70.50	68.90
Average		70.95	70.15

(global optimal value). In this RBF kernel optimization method, the fitness function that used is the RBF function so that the cost of each particle that obtained in accordance with the data position to the kernel. The minimum value of the RBF kernel has been chosen as $gBest$ because the smaller value of $K(x, x')$, the greater the distance between the initial value of the data and the value of the data in the new feature space so that the swarm of the data in the new dimension is moving by following the particle which move furthest.

Kernel's Parameter Computation

Having obtained the value of $pBest$ and $gBest$, we calculate the velocity of each particle. In PSO, we use equation (1) to calculate the velocity. However, based on the experiment that had been conducted, if equation (1) is applied directly to this RBF kernel optimization method, then the given accuracy of the SVM classification results will become unstable because of the influence of the random variables r_1 and r_2 . Therefore, the RBF kernel optimization method is using equation (5), which is a modification of equation (1), to calculate the value of velocity at which the value of variable c_1 and c_2 in equation (1) are combined into a single variable called the correction factor c . Besides, the random variable r_1 and r_2 in equation (1) is set to be the same value, namely r .

$$v_i(t) = wv_i(t-1) + cr \left((x_{pBest_i}(t) - x_i(t)) + (x_{gBest}(t) - x_i(t)) \right) \quad (5)$$

After we get the velocity value of each particle, the position of each particle is updated using equation (2). The new position of the particle is the value of data x' in the new feature space. The position value of the particle is updated based on the velocity of the particle. Therefore, changes in the σ value is also updated based on the average

velocity of the particles using equation (6) where n is the number of the particles.

$$\sigma(t) = \sigma(t-1) + \frac{1}{n} \sum_{i=1}^n v_i(t) \quad (6)$$

After the position of each particle is updated and the new value of parameter σ is obtained, we do optimal particle search using the new value of the particle's position and σ . This process is done repeatedly until the number of iteration reaches the epoch value.

Classification With SVM

After the epoch value is reached, the matrix of the mapping result from RBF kernel $K(x, x')$ on the last iteration is used to perform the classification process using SVM as described by Cortes and Vapnik in their research in 1995 [9]. Because the kernel used is RBF, then the *hyperplane* that used is circular. The optimal margin value $\left(\frac{1}{\|w\|}\right)$ calculation using equation (7) is done to determine the support vectors, where the parameter C is the parameter of soft margin SVM and ξ is the margin error. Output of this process is a SVM model that will be used to classify the data test.

$$\text{minimize} \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \right) \quad (7)$$

3. Results and Analysis

The dataset used in this study is the Breast Cancer Wisconsin (Original) dataset from UCI machine learning repository [10]. This dataset consists of 699 data with 9 numeric attributes which are divided into two classes : benign and malignant.

TABLE 6
ANALYSIS OF THE EFFECT OF DATA OPTIMIZATION PROCESS ON NUMBER OF FOLD = 5

K-	Optimized σ Value	Accuracy of RBF Kernel Optimization Method (%)	Accuracy of SVM with Optimized σ Value (%)
1	0.539	72.00	72.00
2	0.617	72.00	72.00
3	0.675	72.00	72.00
4	0.579	72.00	72.00
5	0.646	70.80	64.60
Average		71.76	70.52

The experiment conducted is the experiment to determine the optimal parameters of PSO and the comparison of the accuracy between the RBF kernel optimization method with the existing SVM optimization method. At the experiment on determining the optimal parameters, there are three parameters were tested, which are the inertia parameter, correction factor, and the epoch value.

In the comparison of the accuracy of the method, there are two experiments that were conducted. The first experiment compared the classification accuracy of RBF kernel optimization method proposed, the SVM method with default σ value (which is 1), and the methods of parameters C and σ on SVM optimization as proposed in [5]. The second experiment comparing the results of the classification accuracy between RBF kernel optimization method with SVM methods without optimization using σ value which has been optimized with PSO to analyze the influence of the movement of the input data to increase the accuracy of RBF kernel optimization method that was proposed. In these two experiments, we used k-fold cross validation with many fold is 2, 5, and 10 pieces to perform the separation between the training and the test data.

Inertia

Inertia parameter is the percentage of the velocity of the particle at the previous iteration that will be used in the current iteration. The greater the inertia, the greater the velocity of the particles in a iteration, likely. Table 1 shows the results of the experiment on determining the optimal value of inertia that was measured from the accuracy that is achieved. It appears from Table 1 that the optimal accuracy is obtained by the inertia value = 0.4.

Correction Factor

Correction factor is the parameter that determines the magnitude of the effects of optimal swarm particle ($pBest$ and $gBest$) value on the particle's velocity. Table 2 shows the results of the experiment on determining the optimal value of correction factor that was measured from the accuracy that is achieved. It appears from Table 2

that the optimal accuracy is obtained by the correction factor value = 0.7.

Epoch

Epoch is the limit of iterations performed while running the PSO algorithm. The greater the value of the epoch, could be that the position of the particles in the swarm is getting closer and centralized, but could also further apart when the central point was missed because the value of epoch is too big. The greater the value of another parameters (inertia and correction factor) the less epoch value needed to reach the central point. Table 3 shows the results of the experiment on determining the optimal value of epoch that was measured from the accuracy that is achieved. It appears from Table 3 that the optimal accuracy is obtained by the epoch value = 100.

Comparison With Another Methods

We conducted an experiment to determine whether the RBF kernel optimization method can provide the optimal value of parameter σ . The optimality of the parameter σ is measured from the given accuracy of the classification. The accuracy is compared with the usual SVM with RBF kernel using the default σ value, which is 1. We also made comparison between RBF kernel optimization method with SVM parameter optimization method as in [5]. This experiment is done using k-fold cross validation with the number of folds are 2, 5, and 10 as shown in Table 4.

In Table 4, it appears that the method proposed in this research, RBF kernel optimization method, have a higher average accuracy than the usual SVM with RBF kernel method and SVM parameter optimization method. Seen in Table 4 that the bigger the value of the fold, which means the amount of training data is also bigger, the accuracy of RBF kernel optimization method is decreased. This could have happened if the input data lying scattered so that the value of the σ parameter obtained is not optimal, and not good as the value obtained by SVM parameter optimization method. Seen in Table 4 that the accuracy obtained RBF kernel optimization method is quite stable and

TABLE 7
ANALYSIS OF THE EFFECT OF DATA OPTIMIZATION PROCESS ON NUMBER OF FOLD = 10

K-	Optimized σ Value	Accuracy of RBF Kernel Optimization Method (%)	Accuracy of SVM with Optimized σ Value (%)
1	0.160	69.20	69.20
2	0.227	69.20	69.20
3	0.224	69.20	69.20
4	0.239	69.20	69.20
5	0.233	69.20	69.20
6	0.216	75.00	75.00
7	0.246	75.00	75.00
8	0.210	75.00	75.00
9	0.219	75.00	75.00
10	0.213	72.70	72.70
Average		71.90	71.90

better than the SVM method without optimization.

In terms of complexity, RBF kernel optimization method has a lower complexity than SVM parameter optimization method. In general, the complexity of SVM with RBF kernel is $O(nd)$ [11]. In SVM parameter optimization with PSO method, SVM is executed repeatedly according to the value epoch (i) and the number of particles PSO (p) so that the complexity of this SVM parameter optimization method is $O(ipnd)$. In the RBF kernel optimization method, although the function of RBF also executed many times according to the number of PSO epoch, the process of finding the optimal margin value on SVM is only run once so that the complexity of RBF kernel optimization method is lower than the complexity of SVM parameter optimization method. This is proven by the value of the running time of both methods that shown in Table 4.

Analysis of the Effect of Data Optimization Process

We conduct an experiment to analyze the effect of the PSO on the movement of the input data to increase the accuracy of the RBF kernel optimization methods. We do a comparison between RBF kernel optimization method with the usual SVM with RBF kernel method in which the σ values used in the usual SVM method is the optimal σ parameters obtained from RBF kernel optimization method. The experiment is done using the number of folds are 2, 5, and 10 that is shown in Table 5, Table 6 and Table 7 respectively. Average accuracy obtained RBF kernel optimization methods and SVM with optimal parameters σ at this experiment is 71.54% and 70.86% respectively.

From Table 5, Table 6 and Table 7 it appears that the accuracy obtained with the RBF kernel optimization method and the usual SVM method that uses optimized σ value does not vary much. Therefore it can be concluded that the data optimization process on RBF kernel optimization method can improve the accuracy of RBF kernel

optimization method, but the effect is not significant.

In the RBF kernel optimization method, the process of determining the optimal σ value have a more significant influence on the improvement of SVM classification accuracy than the input data optimization process. In addition, the value of the optimal parameter σ obtained from the analysis of the movement of the input data can also improve the classification accuracy of usual SVM significantly as shown in Table 4 in which the average of the results of the classification accuracy of usual SVM without optimization of only 66.06%, while the average accuracy results of usual SVM classification which uses optimal parameter σ is 70.86%.

4. Conclusion

RBF Kernel optimization with PSO (Particle Swarm Optimization) can improve the accuracy of SVM (Support Vector Machine) classification method quite significantly. According to the experimental results using the Breast Cancer Wisconsin (Original) dataset, RBF kernel optimization methods gives the average accuracy of 71.55% while the SVM method without optimization gives the average accuracy of 66.06%. In addition, the RBF kernel optimization methods also provide accuracy results which were not much different from the SVM parameter optimization method with PSO. The advantages of RBF kernel optimization method lies in its lower complexity so that the running time of RBF kernel optimization method is faster than SVM parameter optimization methods.

Particle Swarm Optimization (PSO) which is used in the optimization method RBF kernel affects two things, which are the movement of the input data and the movement of the parameter σ in RBF kernel. According to experiment, it can be concluded that the movement of the input data increases the accuracy of RBF kernel optimization method, but the effect is not significant. While the optimal parameter σ values obtained from the

analysis of the movement of the input data has a significant influence on the accuracy of the classification results when applied to either RBF kernel optimization method and the usual SVM with RBF kernel.

Determination of the proper parameters of PSO reasonably affect the accuracy of the classification of RBF kernel optimization method. So that a further research about the methods for determining the value of PSO parameters automatically based on the analysis of the input data can be done.

References

- [1] C.-L. Huang and C.-J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert System with Applications*, vol. 31, pp. 231-240, 2006.
- [2] O. Devos, G. Downey and L. Duponchel, "Simultaneous data pre-processing and SVM classification model selection based on a parallel genetic algorithm applied to spectroscopic data of olive oils," *Food Chemistry*, vol. 148, pp. 124-130, 2014.
- [3] C. W. Hsu and C. J. Lin, "A simple decomposition method for support vector machine," *Machine Learning*, vol. 46, pp. 219-314, 2002.
- [4] S. M. LaVelle and M. S. Branicky, "On the relationship between classical grid search and probabilistic roadmaps," *International Journal of Robotics Research*, vol. 97, pp. 673-692, 2002.
- [5] S. Ding and S. Li, "PSO Parameters Optimization Based Support Vector Machines for Hyperspectral Classification," in *International Conference on Information Science and Engineering (ICISE)*, 2009.
- [6] S.-W. Lin, K.-C. Ying, S.-C. Chen and Z.-J. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," *Expert Systems with Application*, vol. 35, pp. 1817-1824, 2008.
- [7] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proceedings of IEEE International Conference on Neural Networks*, vol. 4, pp. 1942-1948, 1995.
- [8] Y. Shi and R. Eberhart, "A modified particle swarm optimization," *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 69-73, 1998.
- [9] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.
- [10] C. Blake and C. Merz, "UCI repository of machine learning databases," 1998. [Online]. Available: <http://www.ics.uci.edu/~mllearn/MLRepository.html>.
- [11] H. Cao, T. Naito and Y. Ninomiya, "Approximate RBF Kernel SVM and Its Applications in Pedestrian Classification," in *The 1st International Workshop on Machine Learning for Vision-based Motion Analysis - MLVMA*, Marseille, France, Oct 2008.