Accelerometer-based human activity classification using water wave optimization approach

Moataz Kilany[1,*], Aboul Ella Hassanien[2,3,*], Amr Badr[2]

[1]Minia University, Faculty of Computers and Information, El-Minia, Egypt
[2]Faculty of Computers and Information, Cairo University, Egypt.
[3]Faculty of Computers and Information, Beni Suef University, Egypt
[*]Scientific Research Group in Egypt (SRGE)
http://www.egyptscience.net

***Abstract*—This paper introduces an approach of classifying accelerometer data for simple human activities using Support Vector Machine (SVM). Classifier results are being optimized by (WWO) Water Wave Optimization Algorithm. Human activity classification has been a hot research point and has been an aid for smart human activity recognition-based systems and for analysis purposes. Numerous classification methods are introduced by researchers including decision trees, bagging of trees, boosting of trees and random forests [1] as well as the SVM that is employed in this work. Accelerometers are very effective source of data for human activity recognition purposes and have been used in many research efforts as well as this paper. This work shows the process of optimizing SVM parameters in order to get better classification results for a set of human activity accelerometer-based data. The applied optimization algorithm is (WWO) [2], a meta-heuristic evolutionary algorithm. 15-Fold cross validation is accomplished with different data preparation using a leave-one-subject-out approach. Classification accuracy on the applied dataset ranges from 81% to 97% for a first run, and approached 100% accuracy for a second run. validation being applied on 15 different folds on 15 separate dataset files each represents a different participant. We employ four different data preparation configurations.**

***Index Terms*—Human Activity Classification, Accelerometer Data, SVM, WWO, Parameters Optimization, Data Preparation**

I. INTRODUCTION

Work introduced here covers a very rich topic that makes use of numerous computer science advancements related to optimization, classification algorithms as well as pervasive computing. The following sections provide an introduction to such approaches being applied and related works.

A. Human Activity classification

Human activity classification has been a focus of research efforts for its importance in our daily life. There are many sources of data for classification such as video feeds, Images, voice and accelerometer data analysis and for numerous purposes such as security, medical and entertainment. Video feeds and Images are popular for being handy and useful sources of data such as surveillance systems, these are called vision-based systems, and those also appear every where nowadays [3], [4], for instance airports and subway stations in order to detect and identify abnormal behaviors and take proper precautions accordingly.

B. Accelerometer-based Classification

One main advantage of this source of information is its availability and simplicity. They are cheap, easy to use, does not corrupt the natural behavior of participants and many other advantages. Currently almost every one is carrying a smart phone or PDA and most such devices are currently providing an accelerometer installed on it, which is a facility and a lead to solve many problems such as human activity recognition from walking, moving or any other situations that encompass motions with such devices accompanied. This, as well as computer vision systems, is a big aid for surveillance and medical applications [5], life improvement applications, and for entertainment purposes. Research introduced numerous efforts employing accelerometer-based wearable devices. [6] showed a study that uses a single accelerometer mounted to the waist of participants performing natural activities in real life environments. In addition to other publications employ similar approaches [7], [8], [9], [10], [11] with some variations. They change the wearable sensor systems, locations on the body such as wrist, arm, chest, waist, head depending on the desired classification application. employing different sampling frequencies, and other different configurations. Casale et.al showed a study on accelerometer-based activity recognition in [12] where the random forests algorithm is applied for classification.

C. Metaheuristic algorithms

Lately there are a variety of nature-inspired meta-heuristic optimization algorithms that can solve a range of hard search and optimization problems [13] such as ant colony, bat swarm [14], social spider [15], cuckoo search [16], monkey search algorithm [17] and counting still. Such algorithms have been proving their efficiency for solving optimization problems. Research efforts performed optimization for SVM parameters using PSO optimization algorithm with considerable improvement in classification accuracies with different parameter selection configurations [18]. Such algorithms have common procedures and phases, the mission is to search a space of solutions in a non-exhaustive heuristic paradigm which differs from one another. They involve an exploration stage where candidate best solutions are selected and an exploitation phase where recent best solutions found are exploited to find better solutions at similar locations. Nature-inspired algorithms generally assign each solution to an object that mimics a natural object that performs a search-like procedure. This object can represent a bat in swarm, fish in swarm or others. Important to note that such optimization problem cannot be solved by conventional optimization algorithms because of high dimensionality and time cost.

D. Water Wave Optimization Algorithm (WWO)

WWO is a novel optimization algorithm published by Yu-Jun Zheng, October 2014 [2]. The search object employed by the algorithm is a water wave that passes through a number of exploration and exploitation phases that changes the wave location for possible better solutions. Wave passes through propagation, refraction and breaking processes during the optimization. WWO is featured by its simplicity and fast search. It is simple for the algorithm itself and for the small number of parameters applied. Being compared against a number of meta-heuristic algorithm using a set of uni-modal and multi-modal benchmark functions and showed great comparative results.

E. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a well-known classifier and successful in many machine learning applications and has proved excellent generalization capability in classification problems. It was introduced in 1990's and employed for pattern recognition problems. The classifier theory aims for maintaining the largest margin between different classes of data using a kernel function [19]. SVMs are being applied in classification problems for many purposes and fields, [20] is a recent research aims to locate and evaluate lung diseases from image feeds. Other text-mining application appears in [21] where SVM is employed to classify sexual predators patterns. This as well as a number of generalized applications on data sets in [22], this work shows an approach for data set noise and skewness reduction by optimizing SVM performance using PSO algorithm.

F. Classifiers Optimization

Optimizing classifiers is a known approach and there are a number of research efforts in this area for different applications. In [23] an optimization algorithm is introduce for the sake of optimizing convex hull classifier results and for binary classification in general. Genetic Algorithm is employed for optimizing k-nearest neighbors classifiers on the weight and offset parameters [24]. GA optimized SVM and was applied as well on colon cancer, leukemia cancer, and lung cancer datasets, this appeared in a study in In [25]. In [26], gradient descent optimization algorithm is used to optimize the AUC metric (Area Under The Curve) for different classifiers. Also PSO optimization algorithm has been applied in many research efforts to optimize classification results. [27] uses PSO to optimize generalization of SVM classification results of EMG signals where the classifier is applied for each PSO particle separately on different folds of validation, focusing on the kernel parameter settings. [28] also shows a study for SVM parameters optimization applying simulated annealing as the optimizer in that case, with the focus on optimizing four features of classifier that are Kernel selection, adjusting kernel parameters, penalty parameter, and feature selection, and hence, the optimization algorithm tries to find the best combination of the four classification features. This paper follows a similar approach with a different optimizer and dataset, employing WWO for optimization on human activity datasets. [29] Showed similar approaches for predicting a successful growth cycle of the Spirulina platensis from raceway experiments data.

II. Human Activity Classification

This paper introduces an optimization method for SVM accuracy to obtain better classification results using a meta-heuristic optimization algorithm (WWO). The focus of optimization here is the gamma and penalty (C) parameters of SVM Kernel function.

A. Data Sets

The employed data set for this research is provided by Casale et al. [30]. Its data source is a tri-axial accelerometer device fixed to the chest of a human participant performing a number of simple activities. Samples are recorded on 52 HZ frequency. Whole data set is assembled of 15 files for different participants, each recorded 100 to 150 thousand samples of different 7 simple activities. Accelerometer readings are recorded in the form of X, Y and Z accelerations for each sample with a label picked from 1 through 7 for the following simple activities :-

1) Working at Computer
2) Standing Up, Walking and Going up / down stairs
3) Standing
4) Walking
5) Going Up / Down Stairs
6) Walking and Talking with Someone
7) Talking while Standing

The overall samples provided for 7 activities by 15 participants are enumerated as 1926896 samples are all being fed to the classifier in different folds. Which is a magnitude of 2 million samples. This number should give support for the validity of the results of this work.

B. Data Preparation

We put the dataset under a very simple preparation stage. Here preparation is done for the purpose of reduction of dataset size as well as canceling variations in samples and resulting more generalized classification. Preparation is done by obtaining the mean value of a window of data series for each accelerometer reading axis X, Y and Z as shown in Eq. 1.

Algorithm 1 WWO-SVM Optimization Procedure

```
1: Input: HA Data Set (DS) , FPD (Files Per Fold) , FD (Folds Count)
2: Output: Overall Average Accuracy
3: for Each Participant DataSet P in DS do
4:    Perform preparation of P
5:    Add P to prepared dataset PDS
6:    TSI ← 1 to FD  (Training Set Indices)
7:    VSI ← ϕ  (Validation Set Indices)
8:    for FoldID ← 1 to FD do
9:       TSI ← TSI ∩ VSI (Training Set Indices)
10:      TSI ← ϕ
11:      VSI ← TSI{0 to FPD} (Validation Set Indices)
12:      TSI ← TSI − TSI{0 to FPD}
13:      for Each Ind ∈ TSI  (Build Training Set) do
14:         TDS ← TDS ∪ PDS[Ind]
15:      end for
16:      for Each Ind ∈ VSI  (Build Validation Set) do
17:         VDS ← VDS ∪ PDS[Ind]
18:      end for
19:      Acc ← Train_Optimize_SVMTDS,VDS)
20:   end for
21: end for
22: Exit
```

$$S_k = \frac{1}{w} \sum_{i}^{i+w} X_i \tag{1}$$

Where, S_k represents the new averaged sample of index k after averaging a window of w samples, w is the window of samples to be averaged. This work will emphasize that different windows of data preparation will affect the accuracy of classification results.

C. SVM Optimization

The optimization process is done by a simple modification for WWO algorithm by redefining the objective function being applied. In each WWO wave propagation step, it relocates by updating the SVM parameters and recalculates its fitness by performing training and validation on the given dataset and returning classification accuracy. And hence, the optimizer task is to optimize classification accuracy based on systematic updates on the given SVM kernel parameters.

1) WWO Configuration: In this work, the focus is on two parameters for SVM to be optimized, those are; the penalty parameter (C) together with the gamma parameter. Penalty parameter controls the generalization ability of SVM, also said to represent the tradeoff between generalization of the algorithm (correct classification of new data - maximum margin) and how much training is accurate (obtain minimum margin). Gamma parameter affects how far the algorithm can correctly reach (classify) new data.

The algorithm has performed two runs with two different gamma parameter configurations, both parameters are assigned to two dimensions of the water wave optimizer with dimension bounds set to [1,100] and [0.001 , 100] for penalty and gamma parameters respectively on a first run.

For the second run the algorithm is assigned a greater margin for the gamma parameter which showed exceptionally better optimization results where gamma parameter is set in $[10^{-5}, 10^5]$

2) SVM Configuration: SVM is continuously reconfigured as WWO wave suggests for both penalty and gamma parameters, while the kernel function applied is RBF (Radial Basis Function).

3) The optimization procedure: The new hybrid algorithm as stated before assigns SVM classification accuracy metric as the objective function of WWO wave. Optimization and classification processes runs according to the configurations discussed in sections II-C2 and II-C1. Algorithm 1 shows a high level algorithmic framework for data preparation and fold data selection from 15 participant files.

Step 19 represents the actual optimized classification that runs separately at each fold. The run is a normal WWO run with the objective function $f(X)$ represents an SVM Training and validation upon the selected training and validation sets TDS, VDS respectively.

A presentation for the overall architecture of the system is presented in Fig. 1 in terms of an activity diagram.

D. Validation

Validation process is done through 15 Folds. Each fold puts 14 data set for 14 participant as training dataset against the remaining single participant dataset for testing. Each fold picks a different participant from the provided 15 to be tested on and the others are preserved for training so that testing data are always new to the classifier in order to prevent the problem of over-fitting [31].

In this paper, four different configurations for data preparations are selected with different windows of samples as discussed in section II-B. Testing process is accomplished using a personal computer PC with the following configuration; CPU Core i7 2.4 GHZ with memory of 8 GB.

III. Results And Discussion

A. First Run

The results of four optimization runs each on 15 folds are shown in table II. It shows accuracies for all 15 folds for each run representing four selected different configurations W1, W-2, W3 and W4 for windows of 4, 8, 12 and 32 seconds of original samples respectively.

B. Second Run

This run undergone same configuration with greater gamma parameter range in $[10^{-5}, 10^5]$. Results of this run appears in Table III-A. Here we obtained the results only for two different windows of data preparation (averaging) that are 4-seconds and 16 seconds.

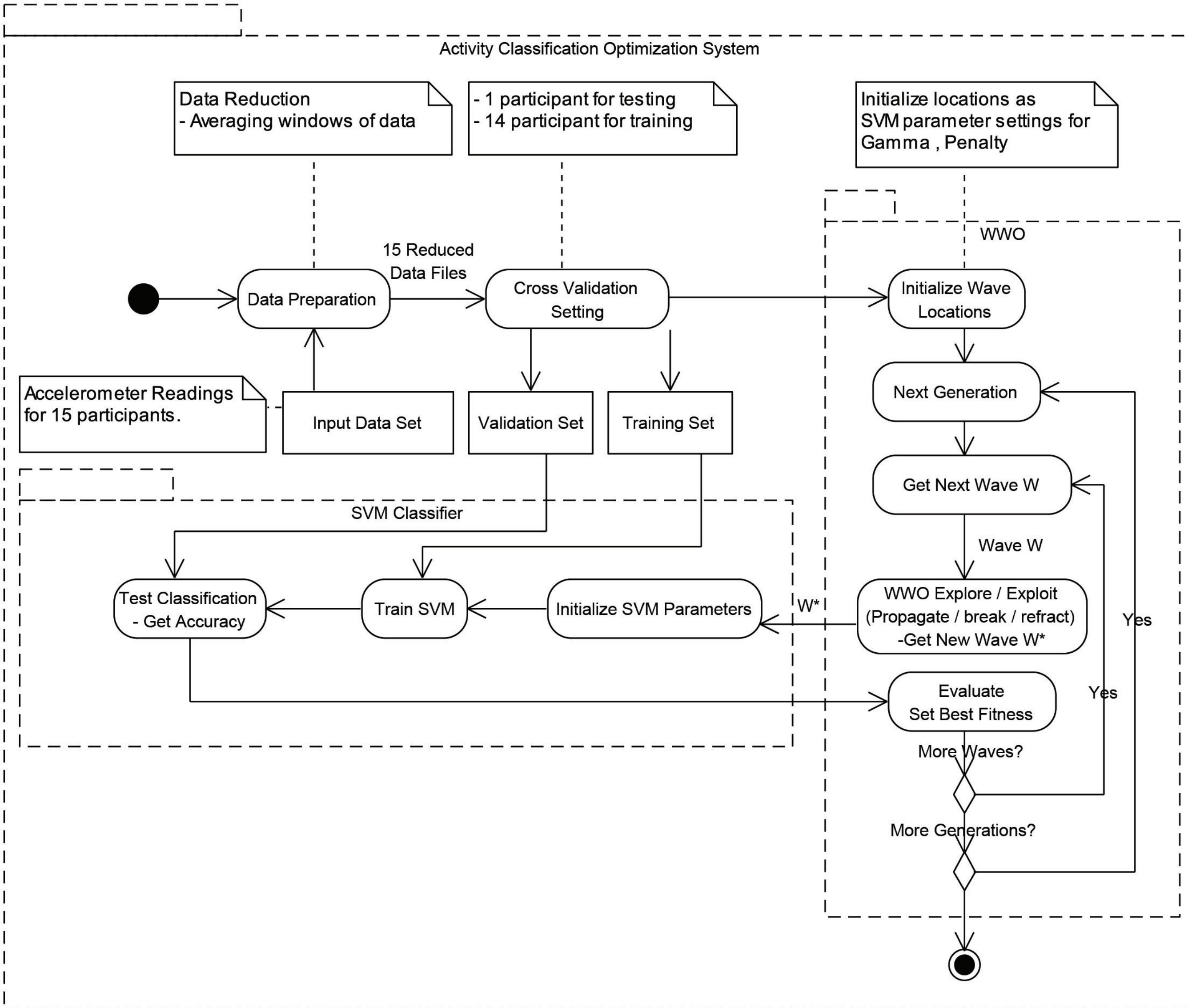

Fig. 1. Activity Diagram - Overall System Architecture

TABLE I
SVM OPTIMIZATION RUN-1 ON HUMAN ACTIVITY DATASETS

Fold	W-1(4-sec)	W-2(8-sec)	W-3(12 sec)	W-4(32sec)
1	84.97%	88.32%	82.56%	89.22%
2	91.89%	94.07%	89.50%	97.75%
3	83.29%	84.86%	96.51%	95.45%
4	89.61%	89.19%	89.15%	97.40%
5	83.33%	85.60%	93.33%	90.00%
6	92.58%	93.90%	88.38%	93.48%
7	91.44%	91.16%	95.43%	93.20%
8	86.85%	88.10%	91.58%	88.89%
9	85.80%	86.91%	87.72%	95.24%
10	84.68%	87.06%	90.73%	93.83%
11	81.71%	84.31%	91.03%	92.54%
12	84.63%	87.05%	93.85%	93.24%
13	81.32%	82.04%	88.73%	86.67%
14	84.58%	87.32%	93.75%	97.33%
15	81.25%	82.94%	86.05%	90.91%
Avg	85.86%	87.52%	90.55%	93.01%

The results show overall average classification accuracy from 85% to 90% for the first run, and approaches 100%for the second run with different data windowing configurations. Variation in accuracy also is related to the selected participants for training and validation with results varying between 81% to 97% along the 15 folds. It is obvious that larger windows of samples to be averaged before classification process results in greater classification accuracy. It is important to note that the windowing strategy is limited. The limitation depends on the observation technique employed and the sampling frequency. A window of 32 seconds will force the observer to record human movements not less than 32 seconds in order to produce a successful classification. This is not available in all cases and depends on the nature of space being observed

TABLE II
SVM OPTIMIZATION RUN-2 ON HUMAN ACTIVITY DATASETS

Fold	W-1(4-sec)	W-2(16-sec)
1	99.69%	99.49%
2	99.64%	99.70%
3	98.79%	99.60%
4	99.39%	100.00%
5	98.91%	99.74%
6	100.00%	100.00%
7	98.78%	99.49%
8	98.56%	98.81%
9	98.80%	100.00%
10	99.80%	99.03%
11	99.76%	100.00%
12	100.00%	98.80%
13	100.00%	100.00%
14	99.57%	100.00%
15	99.76%	100.00%
Avg	99.43%	99.62%

and the speed of human activities being recorded.

IV. CONCLUSION

This paper emphasizes the effect of optimization techniques on classification accuracy, as well as the effect of data preparation. The application of the meta-heuristic water optimization algorithm yielded classification accuracy ranges from 81% to 94% on gamma parameter being optimized in the window of $[0.001, 100]$ being applied on dataset of accelerometer readings for 15 different persons each acting 7 simple activities. The algorithm showed much greater classification accuracies approached 100% on gamma parameter optimization in a greater window of $[10^{-5}, 10^{5}]$.

The final conclusion of results show that optimizing SVM gamma on a very large window in a very small to a very large number results in high classification accuracy. This as well as the effect of data preparation and resampling in greater windows which results in more accurate results too. Greater windows of data results in more accuracy as stated; However this is limited by the ability of observer to capture greater windows of time intervals for human activities.

V. ACKNOWLEDGMENTS

This work is partially supported by UCI - Machine learning repository [32] for the provided dataset authored by Casale et. al.

REFERENCES

[1] S. Pirttikangas, K. Fujinami, and T. Nakajima, “Feature selection and activity recognition from wearable sensors,” in *Ubiquitous Computing Systems*. Springer, 2006, pp. 516–527.

[2] Y.-J. Zheng, “Water wave optimization: a new nature-inspired metaheuristic,” *Computers & Operations Research*, vol. 55, pp. 1–11, 2015.

[3] J. K. Aggarwal and M. S. Ryoo, “Human activity analysis: A review,” *ACM Computing Surveys (CSUR)*, vol. 43, no. 3, p. 16, 2011.

[4] P. C. Ribeiro and J. Santos-Victor, “Human activity recognition from video: modeling, feature selection and classification architecture,” in *Proceedings of International Workshop on Human Activity Recognition and Modelling*. Citeseer, 2005, pp. 61–78.

[5] O. D. Lara and M. A. Labrador, “A survey on human activity recognition using wearable sensors,” *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 3, pp. 1192–1209, 2013.

[6] X. Long, B. Yin, and R. M. Aarts, “Single-accelerometer-based daily physical activity classification,” in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 6107–6110.

[7] C. Randell and H. Muller, “Context awareness by analysing accelerometer data,” in *Wearable Computers, The Fourth International Symposium on*. IEEE, 2000, pp. 175–176.

[8] L. Bao and S. S. Intille, “Activity recognition from user-annotated acceleration data,” in *Pervasive computing*. Springer, 2004, pp. 1–17.

[9] S.-W. Lee and K. Mase, “Activity and location recognition using wearable sensors,” *IEEE pervasive computing*, vol. 1, no. 3, pp. 24–32, 2002.

[10] N. Kern, B. Schiele, and A. Schmidt, “Multi-sensor activity context detection for wearable computing,” in *Ambient Intelligence*. Springer, 2003, pp. 220–232.

[11] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, “Activity classification using realistic data from wearable sensors,” *Information Technology in Biomedicine, IEEE Transactions on*, vol. 10, no. 1, pp. 119–128, 2006.

[12] P. Casale, O. Pujol, and P. Radeva, “Human activity recognition from accelerometer data using a wearable device,” in *Pattern Recognition and Image Analysis*. Springer, 2011, pp. 289–296.

[13] I. BoussaïD, J. Lepagnot, and P. Siarry, “A survey on optimization metaheuristics,” *Information Sciences*, vol. 237, pp. 82–117, 2013.

[14] A. H. Gandomi, X.-S. Yang, A. H. Alavi, and S. Talatahari, “Bat algorithm for constrained optimization tasks,” *Neural Computing and Applications*, vol. 22, no. 6, pp. 1239–1255, 2013.

[15] E. Cuevas, M. Cienfuegos, D. Zaldívar, and M. Pérez-Cisneros, “A swarm optimization algorithm inspired in the behavior of the social-spider,” *Expert Systems with Applications*, vol. 40, no. 16, pp. 6374–6384, 2013.

[16] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, “Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems,” *Engineering with computers*, vol. 29, no. 1, pp. 17–35, 2013.

[17] A. Mucherino and O. Seref, “Monkey search: a novel metaheuristic search for global optimization,” in *Data Mining, Systems Analysis and Optimization in Biomedicine*, vol. 953, no. 1. AIP Publishing, 2007, pp. 162–173.

[18] Y. Bao, Z. Hu, and T. Xiong, “A pso and pattern search based memetic algorithm for svms parameters optimization,” *Neurocomputing*, vol. 117, pp. 98–106, 2013.

[19] J. A. Suykens and J. Vandewalle, “Least squares support vector machine classifiers,” *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.

[20] J. Xi, W. Zhao, J. E. Yuan, J. Kim, X. Si, and X. Xu, “Detecting lung diseases from exhaled aerosols: Non-invasive lung diagnosis using fractal analysis and svm classification,” *PloS one*, vol. 10, no. 9, p. e0139511, 2015.

[21] C. Morris and G. Hirst, “Identifying sexual predators by svm classification with lexical and behavioral features.” in *CLEF (Online Working Notes/Labs/Workshop)*, 2012.

[22] J. Cervantes, F. García-Lamont, A. López, L. Rodriguez, J. S. R. Castilla, and A. Trueba, “Pso-based method for svm classification on skewed data-sets,” in *Advanced Intelligent Computing Theories and Applications*. Springer, 2015, pp. 79–86.

[23] M. Cococcioni, P. Ducange, B. Lazzerini, and F. Marcelloni, “A new multi-objective evolutionary algorithm based on convex hull for binary classifier optimization,” in *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*. IEEE, 2007, pp. 3150–3156.

[24] M. R. Peterson, T. E. Doom, and M. L. Raymer, “Ga-facilitated knn classifier optimization with varying similarity measures,” in *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, vol. 3. IEEE, 2005, pp. 2514–2521.

[25] H.-N. Nguyen, S.-Y. Ohn, S.-H. Chae, D. H. Song, and I. Lee, “Optimizing weighted kernel function for support vector machine by genetic algorithm,” in *MICAI 2006: Advances in Artificial Intelligence*. Springer, 2006, pp. 583–592.

[26] S. Gao and Q. Sun, “Classifier optimization for multimedia semantic concept detection,” in *Multimedia and Expo, 2006 IEEE International Conference on*. IEEE, 2006, pp. 1489–1492.

[27] A. Subasi, “Classification of emg signals using pso optimized svm for diagnosis of neuromuscular disorders,” *Computers in biology and medicine*, vol. 43, no. 5, pp. 576–586, 2013.

[28] P. Gaspar, J. Carbonell, and J. L. Oliveira, “On the parameter optimization of support vector machines for binary classification,” *J Integr Bioinform*, vol. 9, no. 3, p. 201, 2012.

[29] P. G. Nieto, E. García-Gonzalo, J. A. Fernández, and C. D. Muñiz, "A hybrid pso optimized svm-based model for predicting a successful growth cycle of the spirulina platensis from raceway experiments data," *Journal of Computational and Applied Mathematics*, 2015.

[30] P. Casale, O. Pujol, and P. Radeva, "Personalization and user verification in wearable systems using biometric walking patterns," *Personal and Ubiquitous Computing*, vol. 16, no. 5, pp. 563–580, 2012.

[31] D. M. Hawkins, "The problem of overfitting," *Journal of chemical information and computer sciences*, vol. 44, no. 1, pp. 1–12, 2004.

[32] M. Lichman, "UCI machine learning repository," 2013. [Online]. Available: http://archive.ics.uci.edu/ml