applied sciences

MDPI

Article

A Novel MOGA-SVM Multinomial Classification for Organ Inflammation Detection

Kwok Tai Chui [1,*] and Miltiadis D. Lytras [2,3]

1 Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China
2 School of Business & Economics, Deree College—The American College of Greece, 6 Gravias Street GR-153 42, Aghia Paraskevi, 15342 Athens, Greece; mlytras@acg.edu
3 Effat College of Engineering, Effat University, Jeddah P.O. Box 34689, Saudi Arabia
* Correspondence: ktchui3-c@my.cityu.edu.hk

Received: 6 April 2019; Accepted: 29 May 2019; Published: 3 June 2019

Featured Application: In this paper, a novel multi-objective genetic algorithm based support vector machine (MOGA-SVM) has been proposed. A customized similarity kernel has been optimally designed for the multinomial classification of the inflammations of appendix, pancreas, and duodenum. Practically, this methodology can be applied to other classification problems as the concept of the methodology is to customize the kernel to specific application. In order to achieve a better performance using kernel based algorithm, it is highly recommended to use customize kernel instead of traditional kernels.

Abstract: Wrist pulse signal (WPS) contains crucial information of humans' health condition. It can serve as an alternative method for diagnosing of organ inflammation instead of traditional clinical measurement. In this paper, a novel multi-objective genetic algorithm based support vector machine (MOGA-SVM) has been proposed for the multinomial classification of the inflammations of appendix, pancreas, and duodenum. A customized similarity kernel (K_{CS}) has been optimally designed. The performance of multinomial classification using K_{CS} is compared with five types of kernels, linear, radial basis function (RBF), polynomial and sigmoid kernel, as well as mixtures of polynomial and RBF, to verify the effectiveness of K_{CS}. The sensitivity, specificity and accuracy (Acc) of the proposed method are 92%, 91.2%, and 91.6% respectively. The results have demonstrated that K_{CS} improves the accuracy of classification from 8.9% to 59.6%. When compared to related work, the proposed method increases the performance by more than 10%. It is believed that WPS can serve as alternative measures to diagnose organ inflammations.

Keywords: bioinformations; genetic algorithm; multiobjective optimization; organ inflammation; support vector machine; wrist pulse signal

1. Introduction

Health is crucial element in today's life. Researchers have devoted vast efforts in proposing new policies, algorithms, systems, and architectures for healthcare. According to the World Health Organization (WHO), in 2013, the global requirement and the actual number of health workforce were 60.4 million and 43 million, respectively [1]. These figures will be increased to 81.8 million and 67.3 million, respectively, by 2030. Hence, it is believed that the shortage of medical personnel is unsolved and remained serious in the coming decade. Automatic decision making via machine learning is believed to be the only way out to solve the shortage of medical personnel [2,3]. Medical workers may argue that the automatic system has a conflict of interest with them; nevertheless, it is not the truth. First, the current workload of medical workers (ratio of workers to patients) is heavy

and will become normal. Second, an automatic system focuses on routine works, so that medical workers can devote more time to professional consultation and surgery activities. Third, the increase in quality of medical services will lead to higher acceptance and satisfaction by the public. Thus, medical workers will earn a higher social status and better job satisfaction.

Many diseases and abnormal human conditions can be examined by digital imaging diagnostic, like X-ray, Computed Tomography (CT) scan, Magnetic Resonance Imaging (MRI), Ultrasonography, Electrocardiogram, and Biopsy. In this paper, Wrist pulse signal (WPS) of human is considered which provides key information regarding health conditions. In the literature, WPS can be utilized for various applications, for instance, pre-meal and post-meal classification [4], physical exercise [5], diabetes classification [6], hypertension association [7,8], lung cancer recognition [9], and inflammation classification [10,11]. Various signal processing techniques on WPS can be found in [12–14], for instance, dynamic time warping, wavelet analysis, periodic decomposition, principal component analysis, and linear discriminant analysis.

In this paper, four common types of organ inflammation are considered, namely, appendicitis, acute appendicitis, duodenitis, and pancreatitis. According to the WHO, the annual deaths hat are attributable to appendicitis, duodentitis and pancreatitis in 2000, 2005, 2010, and 2015 are shown in Table 1 [15]. It is noted that acute appendicitis is embedded into Global Health Expand (GHE) code 1240. From Table 1, the number of deaths in each category is increasing by an increment of29%, 24%, and 60% for appendicitis, duodenitis, and pancreatitis, respectively, from 2000 to 2015. Among three types of organ inflammations, pancreatitis is the leading cause, which is followed by duodenitis and appendicitis. To conclude, the issues of deaths in these organ inflammations remain unsolved.

Table 1. Number of deaths due to Appendicitis, Pancreatitis, and Duodenitis in 2000, 2005, 2010, and 2015.

GHE Code	GHE Cause	Number of Deaths (Annual)			
		2000	2005	2010	2015
1240	Appendicitis	34,800	39,400	43,300	45,000
1241	Duodenitis	37,900	40,400	43,800	47,000
1248	Pancreatitis	64,400	77,800	93,900	103,500

There have been more than million of sufferers and thus it is necessary to have a reliable and accurate method for the diagnosis of organ inflammations. Based on literature finding, there are a few publications working on binary classification of healthy, appendicitis, acute appendicitis, duodenitis, and pancreatitis sufferers [10,11]. In [10], the features extraction process, an auto-regression (AR) based model was proposed. Two features, the standard deviation and mean of the prediction error from AR model, were chosen to represent the information of the WPS, and for further analysis. With regard to the classification, the support vector machine (SVM) with linear kernel was adopted for the binary classification, which yields an accuracy of 77.8–91.2%. For further improvement, a radial basis kernel (RBF) has been utilized to replace the linear kernel [11]. The idea is that most of the classification problems are not linearly separable. The enhanced method achieved an accuracy of 88.6–98.4%.

Nevertheless, as a pragmatic application, it is deemed to be formulated as classifying instances into one of the more than two classes, and multinomial classification is desired. A novel multi-objective genetic algorithm (MOGA) based SVM, abbreviated as MOGA-SVM, has been proposed for the multinomial classification of the organ inflammations of appendicitis, acute appendicitis, duodenitis, and pancreatitis. MOGA is a heuristic approach that has been widely adopted to obtain tradeoff solutions between two or more conflicting objectives [16–18]. SVM receives a lot of attention as a supervised learning algorithm for classification problems [19–21]. In this paper, a customized similarity kernel (K_{CS}) has been optimally designed for specific application, the classification of appendicitis, acute appendicitis, duodenitis, and pancreatitis. It is worth mentioning that traditional kernels, like linear, RBF, quadratic, and polynomial kernels are not designed for any particular application. It is

recommended that the customized kernel should be utilized for organ inflammations classification instead of traditional kernels in order to improve the classification accuracy.

This paper is organized, as follows. Section 2 provides the background of organ inflammations and an overview of MOGA-SVM. The methodology of the proposed algorithm is explained in Section 3. Performance evaluation and comparison are given in Section 4. Finally, a conclusion is made in Section 5.

2. Dataset and Overview of MOGA-SVM

The background symptoms of each organ inflammation, appendicitis, acute appendicitis, duodenitis, and pancreatitis will firstly describe. Only a summary is provided in each topic. Readers who are interested in the details of the inflammations are suggested to refer to appendicitis [22], acute appendicitis [23], duodenitis [24], and pancreatitis [25]. Subsequently, the overview of the MOGA-SVM is briefly discussed, in which the details will be explained in the next section.

2.1. Background of Organ Inflammation

2.1.1. Appendicitis

Appendicitis is an inflammation of the appendix. It is not uncommon abdominal emergency at any age. The causes are due to the blocking of appendix by stool, cancer, or foreign body, or from infection. Typical symptoms are abdominal pain, loss of appetite, diarrhea, and inability to pass gas. The clinical approaches for diagnosis include abdominal exam, urine test, rectal exam, blood test, CT scans, and ultrasound. The lifetime risk of suffering from appendicitis is about 7%, with different severity levels [22]. The occurrence of this inflammation is approximately 11 persons per 10,000 populations in each year.

2.1.2. Acute Appendicitis

Acute appendicitis is more severe than appendicitis, which has annual incidence of 90 to 140 per 10,000 populations [23]. Although this inflammation has been documented for more than 500 years, its etiology is not well known. It is usually results from injury of its mucosa and spread from that injury via its wall. The symptoms and examinations of acute appendicitis are similar to that in appendicitis.

2.1.3. Duodenitis

Duodenitis is inflammation of the duodenum. The known causes include helicobacter pylori infection, bacterial infection, Nonsteroidal anti-inflammatory drug, viral infection, coeliac disease, and idiopathic [24]. Abdominal pain, nausea, vomiting, and discomfort in stomach are the four known symptoms. The most common examination is an Oesophago-Gastro-Duodenoscopy. For the global annual years of healthy life lost, the estimation is about 58 persons per 100,000.

2.1.4. Pancreatitis

Pancreatitis is inflammation of the pancreas. It is more important than the aforementioned three organ inflammations, because it often characterized by irreversible change, permanent loss of function [25]. The clinical features of pancreatitis include fibrosis, chronic and recurrent inflammation, duct distortion, atrophy, and the risk of pancreatic cancer. The estimated incidence of pancreatitis is 42 persons per 100,000 population [26].

2.2. Overview of MOGA-SVM

Figure 1 shows the flow chart of MOGA-SVM for organ inflammations classification. The typical waveform of the wrist pulse signal is shown in Figure 2, which is characterized by a percussion wave, tidal wave, dicrotic wave, peak systolic velocity, reverse velocity, peak diastolic velocity, and end diastolic velocity.

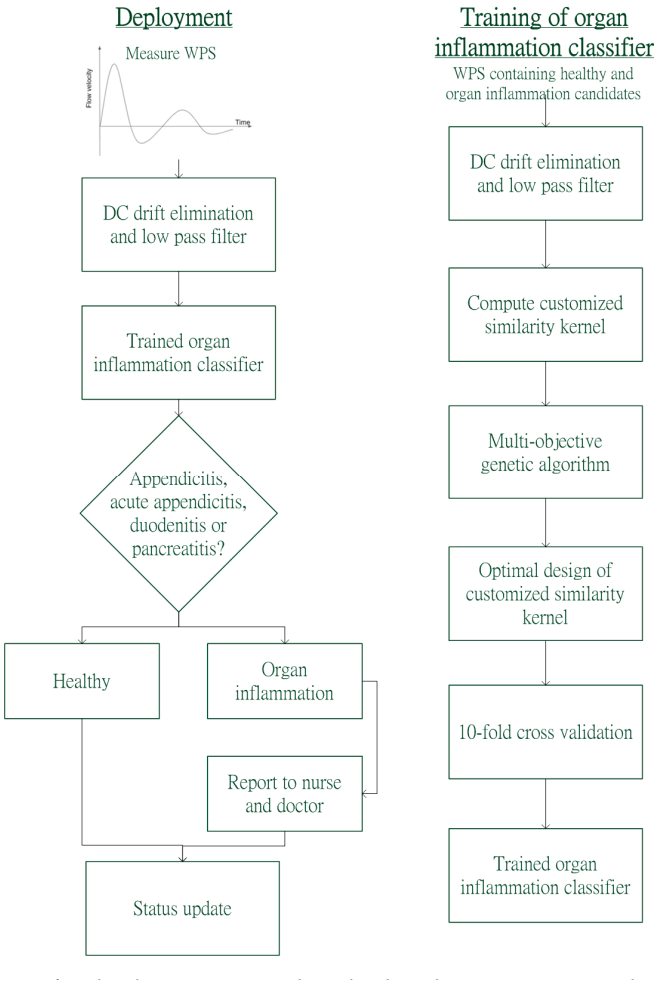

Figure 1. Overview of multi-objective genetic algorithm based support vector machine (MOGA-SVM) for organ inflammations classification.

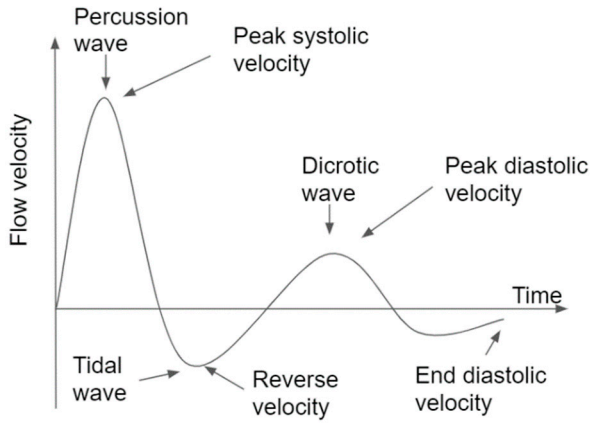

Figure 2. Typical waveform of wrist pulse signal.

The training of organ inflammations classifier, the datasets contain WPS of healthy, appendicitis, acute appendicitis, duodenitis, and pancreatitis candidates were considered [10]. Each record of WPS is carried out DC drift elimination and low-pass filter following the approach, as in [11]. Afterwards, the local maxima and minima points of the WPS are located. The detail is not being discussed in this paper, as the authors would like to mainly focus on the proposed MOGA-SVM.

The similarity coefficients of every pair of WPS are computed, which form the customized similarity kernel. After MOGA, the optimal kernel K_{CS} is designed. A classifier for organ inflammations classification is constructed. Section 3 discusses the details (Figure 3 is drawn to summarize the key steps of the MOGA-SVM). In this paper, the 10-fold cross-validation is adopted to evaluate the classifier, as it is a practical order in literature [27,28].

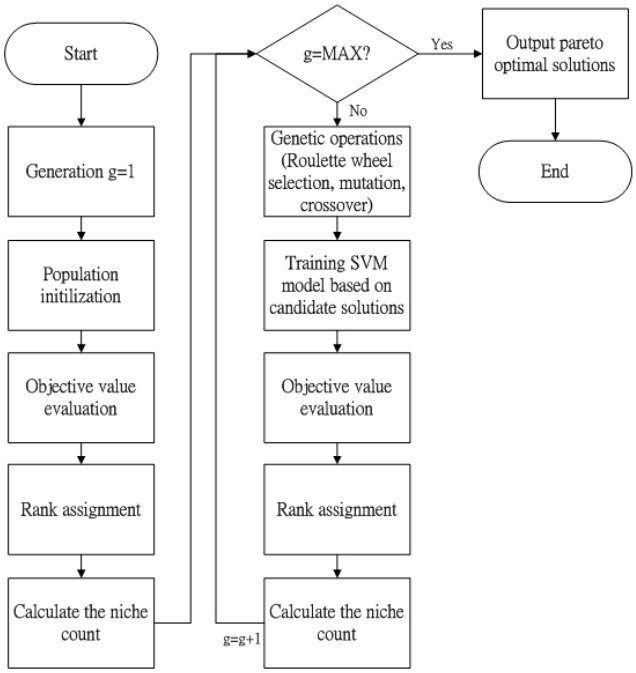

Figure 3. Optimal design of K_{CS} and classifiers using MOGA-SVM.

When it comes to practical application, the WPS of the candidate is measured and it served as the input of the trained organ inflammations classifier. The outputs maintain five possibilities, healthy, appendicitis, acute appendicitis, duodenitis, and pancreatitis. If the status is one of the four organ inflammations, a report will be sent to a nurse and doctor for further examination and treatment.

3. Methodology

This section is composed of three parts. First, the datasets of healthy, appendicitis, acute appendicitis, duodenitis, and pancreatitis candidates are illustrated in Section 3.1. Next, the data preprocessing of the datasets is explained in Section 3.2. At last, Section 3.3 formulates the optimal design of the K_{CS}.

3.1. Datasets of Organ Inflammations Classifier

Gratitude is expressed to researchers in [10] for sharing the datasets. The WPSs were measured while using Doppler ultrasonic blood analyzer module. In each measurement, three steps were followed. Firstly, an approximated position was located where the fluctuation of signal was larger than the other positions. Subsequently, a fine tuning of position with slight variation of angle and position until the largest signal was observed. Finally, WPS was recorded under the setting of largest signal amplitude.

Table 2 summarizes the details of the datasets. Assign the class label to each of the category, Class 0: healthy, Class 1: appendicitis, Class 2: acute appendicitis, Class 3: duodenitis, and Class 4: pancreatitis. The datasets are formed by four age groups, $[0, 20)$, $[20, 40)$, $[40, 60)$ and $[60, 100)$. The total number of samples is 248 and the corresponding samples in Class 0–4 are 100, 22, 38, 42, and 46, respectively.

Table 2. Sample distribution of the datasets.

Class	Name	Age				Total
		[0,20)	[20,40)	[40,60)	[60,100)	
0	Healthy	8	26	30	16	100
1	Appendicitis	0	22	0	0	22
2	Acute Appendicitis	20	8	10	0	38
3	Duodenitis	4	26	6	6	42
4	Pancreatitis	16	26	4	0	46

3.2. Data Preprocessing

The data preprocessing of the aforementioned samples is following the related work [11]. It includes DC drift elimination, six-order Butterworth low-pass filter, and the detection of local maxima and minima points. In this analysis, the WPS has a cycle less than 120 samples. The individual sample is formed by the portion between the two largest maxima points.

There are 1800, 630, 972, 1386, and 828 samples for healthy, appendicitis, acute appendicitis, duodenitis, and pancreatitis candidates, respectively. For equal division using 10-fold cross validation, two, six, and eight samples have been removed for acute appendicitis, duodenitis, and pancreatitis candidates. Overall, there are 5600 samples.

3.3. Formulation of Optimal K_{CS} and MOGA-SVM Classifier

Kernel is essential in SVM classification and it has to obey Mercer's theorem. That is, the kernel is positive semi-definite. A common interpretation of kernel is that it captures the correlation between pairs of data. Thus, the proposed K_{CS} is optimally designed using convolution and cross-correlation. The K_{CS} is formulated as multi-objective optimization problem and is solved by MOGA [29].

Let $X_{i,j}(n)$ of length 120 (zero padding for length<120) be the WPS sample. The subscript i refers to the class label from 0 to 4 and that of j refers to the sample number. Therefore, the sets in Class 0 to Class 4 are $\{X_{0,1}(n), \ldots, X_{0,1800}(n)\}$, $\{X_{1,1}(n), \ldots, X_{1,630}(n)\}$,$\{X_{2,1}(n), \ldots, X_{2,970}(n)\}$,$\{X_{3,1}(n), \ldots, X_{3,1380}(n)\}$,and $\{X_{4,1}(n), \ldots, X_{4,820}(n)\}$, respectively.

The convolution between two WPSs $X_{a,b}(n)$ and $X_{c,d}(n)$ is given by

$$C_{c,d}^{a,b}(n) = X_{a,b}(n) * X_{c,d}(n) = \sum_{k=0}^{N-1} X_{a,b}(k) X_{c,d}(n-k) \tag{1}$$

where N = 120 is the length of the WPS sample.

The cross-correlation between two WPSs $X_{a,b}(n)$ and $X_{c,d}(n)$ can be expressed as

$$R_{c,d}^{a,b}(k) = \begin{cases} \sum_{n=k}^{N-1} X_{a,b}(n) X_{c,d}(n-k), & k \geq 0 \\ \sum_{n=0}^{N-|k|-1} X_{a,b}(n) X_{c,d}(n-k), & k < 0 \end{cases} \tag{2}$$

The customized similarity kernel K_{CS} is formulated by customized convolution kernel K_c and customized cross-correlation kernel K_{cc}. K_c, and K_{cc} are defined as

$$K_c = \begin{bmatrix} X_{c,1,1} & \cdots & X_{c,1,N_t} \\ \vdots & \ddots & \vdots \\ X_{c,N_t,1} & \cdots & X_{c,N_t,N_t} \end{bmatrix} \tag{3}$$

$$K_{cc} = \begin{bmatrix} X_{cc,1,1} & \cdots & X_{cc,1,N_t} \\ \vdots & \ddots & \vdots \\ X_{cc,N_t,1} & \cdots & X_{cc,N_t,N_t} \end{bmatrix} \tag{4}$$

where N_t = 5040 is the 90% of the training samples in Class 0 to Class 4. Here, $X_{c,i,j}$ refers to the weighting sum of convolution coefficients between ith and jth WPS sample. The 1st to 1620th samples come from Class 0. The 1621th to 2187th samples come from Class 1. The 2188th to 3060th samples come from Class 2. The 3061th to 4302th samples come from Class 3. The 4303th to 5040th samples come from Class 4.

$X_{c,i,j}$ and $X_{cc,i,j}$ are computed by

$$X_{c,i,j} = \sum_{m=1}^{2N-1} w_{c,m} C_{c,d}^{a,b}(m) \tag{5}$$

$$X_{cc,i,j} = \sum_{m=1}^{2N-1} w_{cc,m} R_{c,d}^{a,b}(m) \tag{6}$$

where $w_{c,m}$ and $w_{cc,m}$ are the weightings of convolution coefficients and cross-correlation coefficients, respectively. These weightings are optimally designed while using MOGA.

The kernels K_c and K_{cc} predominantly affect the maximum margin and the accuracy (Acc) of the organ inflammations classifier. From (3) and (4), the maximum margin is directly related to both $X_{c,i,j}$ and $X_{cc,i,j}$. Based on (5) and (6), an optimal design of both $X_{x,i,j}$ and $X_{cc,i,j}$ can be obtained by an optimal design of the weighting s$w_{c,m}$ and $w_{cc,m}$ for$m$ = 1, ... ,2N−1, for the given sequences of $C_{c,d}^{a,b}(m)$ and $R_{c,d}^{a,b}(m)$, respectively. In general, an optimally designed kernel will speed up the convergence of the training algorithm MOGA-SVM. However, varying the combinations of $w_{c,m}$ and $w_{cc,m}$ will deduce the different kernels. Searching the optimal weightings require a large computational power. As a result, there is a tradeoff between the accuracy and computational power. As it is difficult to find the optimal values of $w_{c,m}$ and $w_{cc,m}$ that are attributable to the complexity of the objective function, a good trial of $w_{c,m}$ and $w_{cc,m}$ are primarily important, which determines the accuracy.

In this paper, an multi-objective optimization approach, MOGA, is employed to determine the weighting s$w_{c,m}$ and $w_{cc,m}$. However, exhaustive search algorithms may not be the appropriate choices for searching the solution of the optimization problem. The reason is that it deals with a huge range of combinations. Indeed, heuristic search algorithms efficiently and effectively perform the searching of the optimal solutions. In particular, the GA is a robust searching heuristic algorithm that imitates the process of the natural evolution for searching the solution of the optimization problem by the operations in selection, inheritance, crossover and mutation.

Aforementioned, to be a proper kernel for SVM classification, the fulfillment of Mercer's theorem is essential [30]. The evaluation of eigenvalues helps to determine the positive semi-definite of K_c and K_{cc}.

$$\begin{cases} K_c V_c = D_c V_c \\ K_{cc} V_{cc} = D_{cc} V_{cc} \end{cases} \tag{7}$$

where V_c and V_{cc} are non-zero eigenvectors for K_c and K_{cc}, respectively. D_c and D_{cc} are the corresponding eigenvalues. All of the eigenvalues must be positive in order to ensure K_c and K_{cc} are positive semi-definite.

Define K_{cs} as the sum of K_c and K_{cc}. It takes the advantageous from K_c and K_{cc}. It is proved below that the sum of Mercer's kernels is also a Mercer's kernel. If K_c and K_{cc} are positive semi-definite, then for any $c \in \Re^n$, $c^T K_c c \geq 0$ and $c^T K_{cc} c \geq 0$. Hence,

$$c^T K_{CS} c = c^T (K_c + K_{cc}) c \tag{8}$$

$$c^T K_{CS} c = c^T K_c c + c^T K_{cc} c \geq 0 \tag{9}$$

Therefore, the K_{CS} is positive semi-definite. Thus, it is a Mercer's kernel. The multi objective optimization problem is formulated with two objective functions:

$$\begin{array}{ll} Max & F_1 = M(\alpha, w) \\ Max & F_2 = OA = 0.5(S_e + S_p) \end{array} \tag{10}$$

$$s.t. \begin{cases} \alpha_i \geq 0, \sum\limits_{i=1}^{N} \alpha_i y_i = 0, i = 1, \ldots, N \\ \sum\limits_{n=1}^{2N-1} w_{c,n} = 1, \sum\limits_{n=1}^{2N-1} w_{cc,n} = 1 \\ D_{c,i} \geq 0, D_{cc,i} \geq 0, \forall i \end{cases} \tag{11}$$

where $M(\alpha, w)$ is the margin function of the classifier, α_i is the Lagrange multiplier, S_e is the sensitivity, S_p is the specificity, $y_i \in \{-1, +1\}$ is the output of the classifier, and $D_{c,i}$ and $D_{cc,i}$ are the entries of D_c and D_{cc}, respectively. The margin function is defined as

$$M(\alpha, w) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K_{CS} \tag{12}$$

Figure 3 shows the MOGA-SVM for the optimal design of the K_{CS}. The procedures are as follows: (i) The population size and values of objective function are initialized; (ii) The values of objective function of individuals in the population are computed while using the values of objective function defined in (i); (iii) Ranking the individuals according to the values of objective function; (iv) The population convergence is dependent on small group of pareto optimal solutions, but not all optimal solutions are attributable to the nature of the stochastic selection errors, given a limited population size; (v) Niche count is introduced to enhance the population diversity by lengthening the distance between two optimal solutions along the axis of objective functions. The convergence to small group solutions will be avoided; (vi) A new offspring is generated and the values of the objective functions are evaluated; (vii) Ranks assignment and niche count calculation are carried out repeatedly in the new offspring; and, (viii) The algorithm is terminated if it attains the maximum number of generations or if the output reaches the pareto front. To facilitate readers' understanding, Algorithm 1 and Algorithm 2 are presented as the pseudo code of MOGA-SVM.

A pseudo code for the MOGA-SVM is given for better illustration and understanding.

Algorithm 1 *Segmentation*(X_m)

Data: Organ inflammations of appendicitis, acute appendicitis, duodenitis and pancreatitis retrieved from 248 candidates [10], X_m
Output: WPS samples $X_{i,j}$
Step 1: dc drift elimination
Step 2: Filter X_m using low pass filter H_{low}
Step 3: Locate local maxima and minima points of the X_m;
Step 4: Locate two maxima points with interval of 120 sampling points;
$X_{i,j}$ (i = 1:4 = class label, j = length(Class))←Portion of signal between two maxima points with interval of 120 sampling points

Algorithm 2 *TrainClassifier*(*Classlabel*, K_c, K_{cc})

Data: Classlabel, K_c, K_{cc}
Output: Model
Step 1: generations = 1
Step 2: initialization (population)
Step 3: Evaluate the individuals with the fitness function (F1 and F2)
Step 4: rank the individuals by their fitness values by step 3
Step 5: do the Niche count calculation
while generations <= max_generation **do**
Step 6: Select two parents from the population
Step 7: Create the offspring using Roulette wheel selection,crossover and mutation
Step 8: Train SVM model for each individual
Step 9: Evaluate the offspring with the fitness function (F1and F2)
Step 10: rank the individuals by their fitness values by step 3
Step 11: do the Niche count calculation
Step 12: Decide the new population based on the offspring
Step 13: generations = generations + 1
End while
Model←Pareto solutions

4. Performance Evaluation and Comparison

Section 4 is divided into three sub-sections. Firstly, the performance of the proposed K_{CS} is analyzed. Afterwards, it is compared with five other kernels using the feature extraction approach. Finally, performance comparison between proposed and related work is discussed.

4.1. Performance of Proposed MOGA-SVM Using K_{CS}

The performance evaluation of the proposed MOGA-SVM using K_{CS} adopts 10-fold cross validation. Randomly divide 5600WPS samples into 10 equal-sized subsets; with each set containing 560 (10%) samples with Class 0: 180 samples, Class 1: 63 samples, Class 2: 97 samples, Class 3: 138 samples, and Class 4: 82 samples. In each fold of validation, 90% of datasets (nine subsets) from each class serves as training dataset and 10% of the remaining subset serves as the testing datasets. This process completes one-fold of operations. Subsequently, another set is chosen for validation and the remaining nine subsets are used for training. It is noted that this chosen validation set must be different from the validation sets that were selected in the previous folds of operations. The process is repeated until all of the 10 subsets have been validation.

Applying 10-fold cross validation, the proposed MOGA-SVM using KCS achieves average S_e, S_p, and Acc of 92%, 91.2%, and 91.6%, respectively.

4.2. Evaluation of Other Kernels Using Feature Extraction Approach

In this subsection, feature extraction using convolution coefficients and cross-correlation coefficients as features is adopted. The following five kernels, linear, RBF polynomial and sigmoid kernel, and mixtures of polynomial and RBF kernels [31] are applied. They can be expressed by:

$$\text{Linear kernel}: \ k_1(x_i, x_j) = \left\langle x_i, x_j \right\rangle \tag{13}$$

$$\text{RBF kernel}: \ k_2(x_i, x_j) = \exp(\left\| x_i - x_j \right\|_2 / 2\sigma) \tag{14}$$

$$\text{Polynomial kernel}: k_3(x_i, x_j) = \left(\left\langle x_i, x_j \right\rangle + c\right)^p \tag{15}$$

$$\text{Sigmoid kernel}: \ k_4(x_i, x_j) = \tanh\left(\left\langle x_i, x_j \right\rangle + c\right) \tag{16}$$

$$\text{Mixtures of polynomial and RBF kernels } k_5(x_i, x_j) = \rho k_3(x_i, x_j) + (1-\rho) k_4(x_i, x_j) \tag{17}$$

Three scenarios are considered: (i) Only convolution coefficients serve as features (1–199 coefficients); (ii) Only cross-correlation coefficients serve as features (1–199 coefficients); and, (iii) Both convolution and cross-correlation coefficients serve as features (1–398 coefficients).

Table 3 summarizes the performance of kernels K_1–K_5 in three scenarios. Only the best scenario is given. The results reveal that scenario (iii) achieves highest performance, because it takes the advantages from both the convolution and cross-correlation coefficients. Compared K_{CS} with K_1–K_5, the ranking (from highest to lowest) is $K_{CS} > K_5 > K_3 > K_2 > K_4 > K_1$. When compared to scenarios (i), (ii), and (iii), K_{CS} improves the Acc by 14.4–58.2%, 12.4–59.6%, and 8.9–53.7%, respectively.

Table 3. Analysis of traditional kernels in organ inflammation classifications.

Kernel	Performance		
	Scenario (i) (S_e,S_p,**Acc)%**	**Scenario (ii)** (S_e,S_p,**Acc)%**	**Scenario (iii)** (S_e,S_p,**Acc)%**
$k_1(x_i,x_j)$	(57.6, 58.2, 57.9)	(57.7, 57.1, 57.4)	(58.8, 60.4, 59.6)
$k_2(x_i,x_j)$	(76.7, 77.5, 77.1)	(76.8, 76.6, 76.7)	(77.3, 78.3, 77.8)
$k_3(x_i,x_j)$	(77.6, 78.2, 77.9)	(78.3,78.9, 78.6)	(78.7, 80.1, 79.4)
$k_4(x_i,x_j)$	(73.8, 74.6, 74.2)	(73.2, 73.0, 73.1)	(74.8, 75.8, 75.3)
$k_5(x_i,x_j)$	(79.9, 80.3, 80.1)	(82.0, 81.0, 81.5)	(83.8, 84.4, 84.1)

4.3. Comparison between Proposed and Related Work

Based on our finding, the multinomial classification of appendicitis, acute appendicitis, duodenitis, and pancreatitis is the first of its kind. Previous works [10,11] have considered the problem as binary classification. To compare the performance between the proposed and related work [10,11], it is analyzed in two directions. (i) Table 4 gives the raw comparison between the works. (ii) Table 5 gives the matched comparison between the works. The forms of the datasets, application, and cross-validation in [10,11] will be changed into those in this paper. Thus, every work considers 5600 samples for multinomial classification and evaluates using 10-fold cross validation.

From the raw comparison, it can be seen that the performance, S_e, S_p, and Acc of classification between [10] and [11] are similar for binary classification between healthy and appendicitis candidates, and between healthy and duodenitis candidates. For that between healthy and acute appendicitis, and between healthy and pancreatitis, the improvements are 8% and 9%, respectively. If the proposed work is taken into account, it outperforms [10] in the classification of all inflammations. By averaging the Acc in [11], it is approximately equal to the proposed work. Therefore, it can be interpreted that multinomial classification can be achieved without deteriorating the performance in inflammations classification.

A matched comparison environment is setup to compare the performance between algorithms in organ inflammations classification. Repeated simulation is carried out for [10,11,14,32] while using the identical datasets and 10-fold cross validation. It is concluded that the proposed MOGA-SVM improves the Acc from 6.9% to 13.4%.

Table 4. Raw comparison between proposed and related work [10,11].

Work	Method	Feature Extraction	Dataset (Samples)	Cross Validation	Class Labels	S_e (%)	S_p (%)	Acc (%)
[10]	Binary Classification using modified auto-regressive model and linear kernel SVM	Mean and standard deviation of prediction error	Healthy (100), appendicitis (22), acute appendicitis (38), duodenitis (42) and pancreatitis (46)	No	Class 0: healthy; Class 1: appendicitis	81.8	93.3	91.2
					Class 0: healthy; Class 1: acute appendicitis	76.5	82.4	80.8
					Class 0: healthy; Class 1: duodenitis	80.0	91.4	88.0
					Class 0: healthy; Class 1: pancreatitics	83.3	94.4	90.9
					Class 0: healthy; Class 1: All inflammations	80.4	89.7	87.3
[11]	Binary Classification using RBF SVM	peak systolic velocity; reverse velocity; peak diastolic velocity; end diastolic velocity; duration of systole; and duration of diastole	Healthy (100), appendicitis (100), acute appendicitis (100), duodenitis (100) and pancreatitis (100)	10-fold	Class 0: healthy; Class 1: appendicitis	N/A	N/A	92.8
					Class 0: healthy; Class 1: acute appendicitis	N/A	N/A	88.1
					Class 0: healthy; Class 1: duodenitis	N/A	N/A	88.6
					Class 0: healthy; Class 1: pancreatitics	N/A	N/A	98.4
Our work	Multinomial Classification using customized kernel	Cross-correlation and convolution coefficients	Healthy (1800), Appendicitis (630), Acute Appendicitis (970), Duodenitis (1380) and Pancreatitis (820)	10-fold	Class 0: health; Class 1: appendicitis; Class 2: acute appendicitis; Class 3: duodenitis; Class 4: pancreatitis	92.0	91.2	91.6

Table 5. Matched comparison between proposed and related work [10,11,32].

Work	Method	Feature Extraction	Dataset (Samples)	Cross Validation	Class Labels	S_e (%)	S_p(%)	Acc (%)
[10]	Binary Classification using modified auto-regressive model and linear kernel SVM	Mean and standard deviation of prediction error	Healthy (1800), appendicitis (630), acute appendicitis (970), duodenitis (1380) and pancreatitis (820)	10-fold	Class 0: healthy; Class 1: appendicitis; Class 2: acute appendicitis; Class 3: duodenitis; Class 4: pancreatitis	81.3	80.3	80.8
[11]	Binary Classification using RBF SVM	peak systolic velocity; reverse velocity; peak diastolic velocity; end diastolic velocity; duration of systole; and duration of diastole	Healthy (1800), appendicitis (630), acute appendicitis (970), duodenitis (1380) and pancreatitis (820)	10-fold	Class 0: healthy; Class 1: appendicitis; Class 2: acute appendicitis; Class 3: duodenitis; Class 4: pancreatitis	81.7	82.9	82.3
[32]	A recursive cluster elimination based SVM	spatial features obtained from a bi-modal Gaussian model	Healthy (1800), appendicitis (630), acute appendicitis (970), duodenitis (1380) and pancreatitis (820)	10-fold	Class 0: healthy; Class 1: appendicitis; Class 2: acute appendicitis; Class 3: duodenitis; Class 4: pancreatitis	84.7	84.1	84.4
[14]	RBF SVM	Periodic and non-periodic feature extension	Healthy (1800), appendicitis (630), acute appendicitis (970), duodenitis (1380) and pancreatitis (820)	10-fold	Class 0: healthy; Class 1: appendicitis; Class 2: acute appendicitis; Class 3: duodenitis; Class 4: pancreatitis	85.3	86.1	85.7
Our work	Multinomial Classification using customized kernel	Cross-correlation and convolution coefficients	Healthy (1800), appendicitis (630), acute appendicitis (970), duodenitis (1380) and pancreatitis (820)	10-fold	Class 0: healthy; Class 1: appendicitis; Class 2: acute appendicitis; Class 3: duodenitis; Class 4: pancreatitis	92.0	91.2	91.6

5. Conclusions

In this paper, a novel MOGA-SVM has been proposed for the multinomial classification of four common organ inflammations, appendicitis, acute appendicitis, duodenitis, and pancreatitis. A customized similarity kernel K_{CS} is optimally designed using MOGA. K_{CS} captures the characteristics of the inflammations, which is an ideal approach in the kernel selection perspective. Typical kernel functions are generally built-in package as the analytic tool that does not aim at yielding best performance for all applications, and it is thus highly recommended that the customized kernel should be utilized for organ inflammations classification. The results show that the proposed algorithm achieves sensitivity, specificity, and accuracy of 92%, 91.2%, and 91.6%, respectively. It achieves a significant improvement using traditional kernels and related works by 60% and 10%, respectively. It is believed that WPS can be utilized as alternative, reliable and accurate method to determine whether a candidate is suffering from organ inflammation. Besides accuracy, the proposed method is a timely and inexpensive approach. Bringing machine learning into real-world healthcare application is always a good solution to relieve the workload of medical personnel, as everybody needs regular body check and timely examination.

Author Contributions: K.T.C. proposed and implemented the methodology. K.T.C. and M.D.L. have evaluated the performance of proposed work and drafted the paper.

Funding: Authors would like to thank Effat University in Jeddah, Saudi Arabia, for funding the research reported in this paper through the Research and Consultancy Institute.

Acknowledgments: Gratitude is expressed to Y.C., L.Z., D.Z. and D.Z., Department of Computing, Biometrics Research Center, The Hong Kong Polytechnic University, for sharing the datasets.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. WHO. *Health Workforce Requirements for Universal Health Coverage and the Sustainable Development Goals*; World Health Organization: Geneva, Switzerland, 2016.
2. Chui, K.T.; Alhalabi, W.; Pang, S.S.H.; Pablos, P.O.D.; Liu, R.W.; Zhao, M. Disease Diagnosis in Smart Healthcare: Innovation, Technologies and Applications. *Sustainability* **2017**, *9*, 2309. [CrossRef]
3. Spruit, M.; Lytras, M. Applied Data Science in Patient-centric Healthcare. *Telemat. Inform.* **2018**, *35*, 643–653. [CrossRef]
4. Khaire, N.N.; Joshi, Y.V. Disgnosis of disease using wrist pulse signal for classification of pre-meal and post-meal samples. In Proceedings of the 2015 International Conference on Industrial Instrumentation and Control, Maharashtra, India, 28–30 May 2015; IEEE: Piscataway, NJ, USA.
5. Reddy, R.K.; Pooni, R.; Zaharieva, D.P.; Senf, B.; El Youssef, J.; Dassau, E.; Castle, J.R. Accuracy of Wrist-Worn Activity Monitors during Common Daily Physical Activities and Types of Structured Exercise: Evaluation Study. *JMIR mHealth uHealth* **2018**, *6*, e10338. [CrossRef] [PubMed]
6. Li, J.; Zhang, B.; Lu, G.; You, J.; Zhang, D. Body surface feature-based multi-modal Learning for Diabetes Mellitus detection. *Inf. Sci.* **2019**, *472*, 1–14. [CrossRef]
7. He, D.; Wang, L.; Fan, X.; Yao, Y.; Geng, N.; Sun, Y.; Xu, L.; Qian, W. A new mathematical model of wrist pulse waveforms characterizes patients with cardiovascular disease—A pilot study. *Med. Eng. Phys.* **2017**, *48*, 142–149. [CrossRef] [PubMed]
8. Qiao, L.J.; Qi, Z.; Tu, L.P.; Zhang, Y.H.; Zhu, L.P.; Xu, J.T.; Zhang, Z.F. The Association of Radial Artery Pulse Wave Variables with the Pulse Wave Velocity and Echocardiographic Parameters in Hypertension. *Evid. Based Complement. Altern. Med.* **2018**, *2018*, 5291759. [CrossRef]
9. Zhang, Z.; Zhang, Y.; Yao, L.; Song, H.; Kos, A. A sensor-based wrist pulse signal processing and lung cancer recognition. *J. Biomed. Inform.* **2018**, *79*, 107–116. [CrossRef] [PubMed]
10. Chen, Y.; Zhang, L.; Zhang, D.; Zhang, D. Computerized wrist pulse signal diagnosis using modified auto-regressive models. *J. Med. Syst.* **2011**, *35*, 321–328. [CrossRef]

11. Chow, W.H.; Wu, C.K.; Tsang, K.F.; Li, B.Y.S.; Chui, K.T. Wrist pulse signal classification for inflammation of appendix, pancreas, and duodenum. In Proceedings of the 40th Annual Conference of the IEEE Industrial Electronics Society, Dallas, TX, USA, 30 October–1 November 2014; IEEE: Piscataway, NJ, USA.
12. Garg, N.; Bisht, A.; Ryait, H.S.; Kumar, A. Identification of motion outliers in wrist pulse signal. *Comput. Electr. Eng.* **2018**, *67*, 776–790. [CrossRef]
13. Liu, X.; Ji, Z.; Tang, Y. Recognition of pulse wave feature points and non-invasive blood pressure measurement. *J. Signal Process. Syst.* **2017**, *87*, 241–248. [CrossRef]
14. Wang, D.; Zhang, D.; Lu, G. Generalized Feature Extraction for Wrist Pulse Analysis: From 1-D Time Series to 2-D Matrix. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 978–985. [CrossRef] [PubMed]
15. WHO. *Global Health Estimates 2015: Deaths by Cause, Age, Sex, by Country and by Region, 2000–2015*; World Health Organization: Geneva, Switzerland, 2016.
16. Chui, K.T.; Tsang, K.F.; Chi, H.R.; Ling, B.W.K.; Wu, C.K. An accurate ECG based transportation safety drowsiness detection scheme. *IEEE Trans. Ind. Inform.* **2016**, *12*, 1438–1452. [CrossRef]
17. Montazeri, A.; West, C.; Monk, S.D.; Taylor, C.J. Dynamic modelling and parameter estimation of a hydraulic robot manipulator using a multi-objective genetic algorithm. *Int. J. Control* **2017**, *90*, 661–683. [CrossRef]
18. Tseng, F.H.; Wang, X.; Chou, L.D.; Chao, H.C.; Leung, V.C. Dynamic Resource Prediction and Allocation for Cloud Data Center Using the Multiobjective Genetic Algorithm. *IEEE Syst. J.* **2018**, *12*, 1688–1699. [CrossRef]
19. Chui, K.T.; Tsang, K.F.; Wu, C.K.; Hung, F.H.; Chi, H.R.; Chung, H.S.H.; Man, K.F.; Ko, K.T. Cardiovascular diseases identification using electrocardiogram health identifier based on multiple criteria decision making. *Expert Syst. Appl.* **2015**, *42*, 5684–5695. [CrossRef]
20. Fuchida, M.; Pathmakumar, T.; Mohan, R.E.; Tan, N.; Nakamura, A. Vision-based perception and classification of mosquitoes using support vector machine. *Appl. Sci.* **2017**, *7*, 51. [CrossRef]
21. Wu, J.L.; Chang, P.C.; Tsao, C.C.; Fan, C.Y. A patent quality analysis and classification system using self-organizing maps with support vector machine. *Appl. Soft Comput.* **2016**, *41*, 305–316. [CrossRef]
22. Ryan, W.L. *Digestive Diseases—Research and Clinical Developments: Appendicitis: Symptoms, Diagnosis, and Treatments*; Nova Science: New York, NY, USA, 2011.
23. Keyzer, C.; Gevenois, P.A. *Imaging of Acute Appendicitis in Adults and Children*; Springer: Berlin, Germany, 2011.
24. Serra, S.; Jani, P.A. An approach to duodenal biopsies. *J. Clin. Pathol.* **2006**, *59*, 1133–1150. [CrossRef]
25. Adams, D.B.; Cotton, P.B.; Zyromski, N.J.; Windsor, J. *Pancreatitis: Medical and Surgical Management*; Wiley Blackwell: Chichester, UK, 2017.
26. Yadav, D.; Timmons, L.; Benson, J.T.; Dierkhising, R.A.; Chari, S.T. Incidence, prevalence, and survival of chronic pancreatitis: A population-based study. *Am. J. Gastroenterol.* **2011**, *106*, 2192. [CrossRef]
27. De Haan, R.R.; Visser, J.B.; Pons, E.; Feelders, R.A.; Kaymak, U.; Hunink, M.M.; Visser, J.J. Patient-specific workup of adrenal incidentalomas. *Eur. J. Radiol. Open* **2017**, *4*, 108–114. [CrossRef]
28. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schroder, B.S.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [CrossRef]
29. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley & Sons, Inc.: New York, NY, USA, 2001.
30. Herbrich, R. *Learning Kernel Classifiers Theory and Algorithms*; The MIT Press: London, UK, 2002.
31. Smits, G.F.; Jordan, E.M. Improved SVM regression using mixtures of kernels. In Proceedings of the 2002 International Joint Conference on Neural Networks, Honolulu, HI, USA, 12–17 May 2002; IEEE: Piscataway, NJ, USA.
32. Rangaprakash, D.; Dutt, D.N. Study of wrist pulse signals using time domain spatial features. *Comput. Electr. Eng.* **2015**, *45*, 100–107. [CrossRef]