Computers in Biology and Medicine 81 (2017) 167–175

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

An expert system for selecting wart treatment method

Fahime Khozeimeh[a], Roohallah Alizadehsani[b], Mohamad Roshanzamir[c], Abbas Khosravi[d,*], Pouran Layegh[e], Saeid Nahavandi[d]

[a] Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
[b] Department of Computer Engineering, Sharif University of Technology, Azadi Ave, Tehran, Iran
[c] Department of Electrical and Computer Engineering, Isfahan University of Technolgy, Isfahan, Iran
[d] Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Victoria 3217, Australia
[e] Cutaneous Leishmaniasis Research Center, Qaem Hospital, Mashhad University of Medical Sciences, Mashhad, Iran

ARTICLE INFO

Keywords:
Warts
Immunotherapy
Cryotherapy
Fuzzy rule-based system

ABSTRACT

As benign tumors, warts are made through the mediation of Human Papillomavirus (HPV) and may grow on all parts of body, especially hands and feet. There are several treatment methods for this illness. However, none of them can heal all patients. Consequently, physicians are looking for more effective and customized treatments for each patient. They are endeavoring to discover which treatments have better impacts on a particular patient. The aim of this study is to identify the appropriate treatment for two common types of warts (plantar and common) and to predict the responses of two of the best methods (immunotherapy and cryotherapy) to the treatment. As an original work, the study was conducted on 180 patients, with plantar and common warts, who had referred to the dermatology clinic of Ghaem Hospital, Mashhad, Iran. In this study, 90 patients were treated by cryotherapy method with liquid nitrogen and 90 patients with immunotherapy method. The selection of the treatment method was made randomly. A fuzzy logic rule-based system was proposed and implemented to predict the responses to the treatment method. It was observed that the prediction accuracy of immunotherapy and cryotherapy methods was 83.33% and 80.7%, respectively. According to the results obtained, the benefits of this expert system are multifold: assisting physicians in selecting the best treatment method, saving time for patients, reducing the treatment cost, and improving the quality of treatment.

1. Introduction

Machine learning and data mining algorithms are utilized to analyze large datasets and discover and extract knowledge from them. They are also employed as a tool in medical sciences [1–18], crime detection, risk assessment, and sales of products. These algorithms can analyze data in order to discover the unknown patterns in large databases. Industries such as banking, insurance, health, and marketing commonly apply them in order to reduce costs, improve the quality of research, and increase the amount of sales.

Classification is one of the important tasks in machine learning and data mining. Fuzzy rule-based systems have recently been employed for classification [19–23] to handle the concept of partial truth. Truth values may range between *completely false* and *completely true* [24].

Fuzzy rule-based systems are applied in many different fields, including artificial intelligence, control theory [25] and medical fields. In the medical field, they are utilized for the early diagnosis of diseases and important factors influencing them [12,13].

In medical research, one of the most important fields is skin disease, and among the skin diseases, researchers generally apply machine-learning methods to Melanoma treatment [26–40]. Melanoma is a type of skin cancer developing from melanocytes which is a type of pigment-containing cells [41].

A number of studies have been performed on other skin diseases, using machine-learning algorithms [42–45]. However, as far as we know, there has been no machine-learning research conducted in the field of wart treatment thus far. Although there are different wart treatment methods [46], physicians have not recognized which one is more effective for each patient. They are obliged to test each method individually.

In this research, we investigated immunotherapy with candida antigen and cryotherapy with liquid nitrogen on 180 patients with plantar and common warts [46] who had referred to the dermatology clinic of Ghaem Hospital, Mashhad, Iran. These two treatment methods were selected as they are two of the best wart treatment methods. Cryotherapy is the most common wart treatment method.

* Corresponding author.
E-mail address: abbas.khosravi@deakin.edu.au (A. Khosravi).

http://dx.doi.org/10.1016/j.compbiomed.2017.01.001
Received 10 August 2016; Received in revised form 31 December 2016; Accepted 3 January 2017

However, a number of difficulties arise when applying this method. The first problem is that it has side effects. Second, it is painful, and all the warts must be treated together. Third, many treatment sessions are required. Accordingly, experts are looking for novel ways to treat this issue. Immunotherapy is a new treatment method which has lately been employed. Fortunately, it lacks the majority of the deficiencies cryotherapy has encountered. In the present study, we propose a fuzzy rule-based algorithm to detect which one of these two treatment methods has better results for each patient. Not only do we aim to find a good classifier, but we also recommend some useful and interpretable rules to physicians so as to assist them in treating their patients. This diagnosis would help these patients spend less time and money. To the best of our knowledge, our study is the first one conducted in the domain of wart treatment.

2. Datasets

The two datasets utilized in this study were collected in the dermatology clinic of Ghaem Hospital in Mashhad from January 2013 to February 2015. The datasets were collected from patients, with plantar and common warts, who had referred to the dermatology clinic. These two types of warts are two of the most common wart types.

The first dataset consists of seven features gathered when the cryotherapy method was applied. These features are demonstrated in Table 1. The second dataset has eight features collected when the immunotherapy method was employed. Table 2 presents these features. The class attribute in these datasets is the *Response to Treatment* feature.

T-test and chi-square test were used for statistical analysis, and the difference was considered significant when p-value < 0.05. Based on our results, there were no significant differences between age, time elapsed before treatment, surface area of warts, and the number of warts in these two datasets. Their p-values are 0.20, 0.37, 0.62, and 0.28, respectively.

These two types of treatment were selected since they are common wart treatment methods. These features, as important factors for treatment, were selected based on the physicians' opinion. They guessed that in employing these two methods for the treatment, these factors are probably more important than others. However, no research has been performed so far to confirm this hypothesis. A number of medical studies have been done to compare these two treatment methods [47–55]. However, none of them have been aimed at investigating this issue from the perspective of machine learning.

3. Method

The patients were randomly divided into two groups. Each group contained 90 patients. Once a patient came to clinic, he/she was sent to the first group. Then the next one was sent to the second group. This procedure continued until each group had 90 patients. The first group was treated by immunotherapy with candida antigen method. Treating these patients lasted up to three sessions by intralesional injection of vaccine. There was a three-week lapse between the sessions. The vaccination would halt in case a complete treatment was achieved before the third session. In the second group, the cryotherapy method with liquid nitrogen was applied. The treatment continued for a maximum of 10 sessions or until the complete treatment of warts occurred before the 10th session. In this method, there was a one-week time interval between the sessions. After three sessions of immunotherapy and ten sessions of cryotherapy, if the treatment was not achieved, we had to change the method because it was not effective [46].

Table 1
Features utilized in the cryotherapy method.

Feature name	Values	Mean ± SD[a]
Response to treatment	Yes or No	
Gender	47 Man 43 Woman	
Age (year)	15–67	28.6 ± 13.36
Time elapsed before treatment (month)	0–12	7.66 ± 3.4
The number of warts	1–12	5.51 ± 3.57
Types of wart (Count)	1- Common (54), 2- Plantar (9), 3- Both (27)[b]	
Surface area of the warts[c] (mm^2)	4–750	85.83 ± 131.73

[a] Standard deviation.
[b] Patients have both types of common and plantar warts.
[c] Surface area of biggest wart.

Table 2
Features employed in the immunotherapy method.

Feature name	Values	Mean ± SD
Response to treatment	Yes or No	
Gender	41 Man 49 Woman	
Age (year)	15–56	31.04 ± 12.23
Time elapsed before treatment (month)	0–12	7.23 ± 3.10
The number of warts	1–19	6.14 ± 4.2
Types of wart (Count)	1– Common (47), 2– Plantar (22), 3– Both (21)	
Surface area of the warts[a] (mm^2)	6–900	95.7 ± 136.61
Induration diameter of initial test(mm)	5–70	14.33 ± 17.22

[a] Surface area of biggest wart.

3.1. Feature selection

One of the methods employed for feature selection is information gain. This method measures the reduction in the entropy of data. If a feature has the highest information gain, it can separate two classes completely. For the purpose of classification, this type of features is the best one. In other words, the higher the values of the information gain, the better candidate the feature is in the classification task [56]. In Eq. (1), H(T) indicates the entropy of the training data T.

$$H(T) = -\sum_{c \in classes} P(c)\log\,(P(c)) \tag{1}$$

where, c is the class value. The class value is "Response to treatment". The probability that a record is in class c is denoted by P(c).

Eq. (2) indicates the information gain of the a^{th} feature of the training data T.

$$IG(T,\,a)=H(T)-\sum_{v \in values(a)} \frac{|\{x \in T \mid x_a = v\}|}{|T|}.\,H(\{x \in T \mid x_a = v\}) \tag{2}$$

where, x_a represents the value of the a^{th} feature of example x.

3.2. Association rule learning

Association rule learning is used to discover the relationship between features in databases [57]. For each rule, support and confidence are defined. The support of a rule shows the proportion of data, which includes the antecedent and consequence of that rule. The confidence of a rule demonstrates the probability of finding the consequence of the rule in those item sets which include the rule antecedent [56].

Apriori is an algorithm for rule extraction. It is an algorithm for association rule learning over databases [58]. For rule extraction, at first, the high-probability conditions are extracted from the dataset. Next, these conditions are broken in all possible ways. Extracting the rules with the highest confidence is the final step [56]. In this research, only the rules having "Response to treatment" as their consequence were considered.

3.3. Fuzzy rule-based system

In 1965, fuzzy logic was introduced by Lotfi A. Zadeh [59]. Since then, fuzzy logic has been applied as a powerful method for diagnosis of diseases and different decision support systems. In comparison with classical logical systems, fuzzy logic is much closer to human thinking. In point of fact, it can handle the uncertain and inexact nature of the real world. In some cases, the classic rules can decrease the performance of decision support systems. For example, when we have a rule like "if age < 20, then the patient responds to the treatment.", this means that if a patient is 21 years old, he/she does not at all respond to the treatment. However, in fuzzy rule-based systems, this means that our confidence in the treatment of a 19-year-old patient is higher than that of a 21-year-old patient. In fuzzy logic, this problem has been solved by linguistic terms. Employing this method eases the representation of human knowledge. For instance, in the above-mentioned case, the rule changes to something like "if the patient is young, then he/she responds to the treatment.". Meanwhile, in the fuzzy rule-based systems, one error of a rule is not fatal to the whole system, as there is more than one control rule.

In fuzzy rule-based systems, fuzzy rules have the following common form: "IF A, THEN B.", where, A and B are linguistic variables. A and B are called the antecedent and consequence of the rule, respectively. The input variables in a fuzzy system are defined by membership function sets. In the input space, each point is mapped to a number between 0 and 1. This number shows the membership degree. The membership degree for all input points is called membership function and is commonly shown as a membership curve [24]. In this research, we applied the bell membership function. It is defined according to Eq. (3) as follows:

$$bell(x; a, b, c) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}} \tag{3}$$

The three parameters of {a, b, c} specify a generalized bell membership function. These parameters show a physical meaning: 'c' is the center of the membership function, 'a' is the half width, and the slopes of crossover points are controlled by 'a' and 'b'. Fig. 1 illustrates these concepts.

Training fuzzy rules is performed by using a set of classified samples. The trained system can predict the class label of input variables.

There are two types of fuzzy inference systems: Mamdani-type and Sugeno-type. They are selected according to the way in which outputs are determined. The output membership functions in the Mamdani-type must be fuzzy sets. However, in the Sugeno-type systems, they should be either linear or constant [60].

3.4. Adaptive network-based fuzzy inference system (ANFIS)

Based on the Sugeno fuzzy inference system, Jang defined a type of neural network called ANFIS [24]. It was developed in the early 1990s. It captures the benefits of two famous methods of neural networks and fuzzy logic principles as it integrates both of them. Its inference system utilizes a set of fuzzy IF–THEN rules.

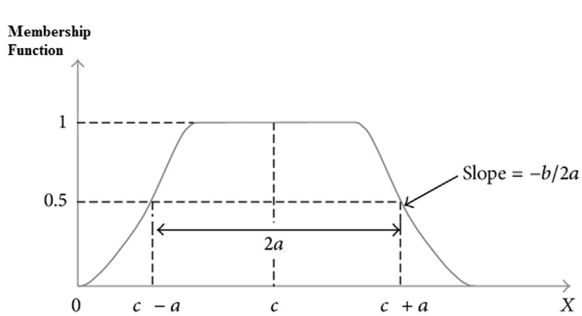

Fig. 1. Bell-shaped membership function [24].

Developing a fuzzy system requires suitable rules and membership functions. To produce correct rules, it is necessary to extract expert knowledge. However, it is a time-consuming process. This is where ANFIS can play its role. It can find suitable parameters for membership functions. Takagi-Sugeno reasoning mechanism is illustrated in Fig. 2(a). Fig. 2(b) indicates the equivalent ANFIS architecture. In this figure, nodes of the same layer have similar functions. The output of the i^{th} node in layer l is represented by $O_{l,i}$.

As Fig. 2(a) demonstrates, suppose that a Sugeno-type fuzzy system has the following rules:

1. If x is A_1 and y is B_1, then $f_1 = p_1x + q_1y + k_1$.
2. If x is A_2 and y is B_2, then $f_2 = p_2x + q_2y + k_2$.

The following two rules were eliminated for the purpose of simplification. However, we can simply extend Eq. (4) to Eq. (13) for these two rules.

1. If x is A_1 and y is B_2, then $f_3 = p_3x + q_3y + k_3$.
2. If x is A_2 and y is B_1, then $f_4 = p_4x + q_4y + k_4$.

Let μ_{Ai} and μ_{Bi} show the membership functions of fuzzy sets A_i *and* B_i, i=1, 2, respectively. To evaluate the rules, "logical and" is estimated by production.

The premises of the rules are calculated by Eq. (4):

$$w_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad i = 1, 2 \tag{4}$$

The consequence of the rules is calculated by Eq. (5):

$$f(x, y) = \frac{w_1(x, y)f_1(x, y) + w_2(x, y)f_2(x, y)}{w_1(x, y) + w_2(x, y)} \tag{5}$$

If we leave the arguments out of Eq. (5), Eq. (6) is obtained.

$$f = \frac{w_1f_1 + w_2f_2}{w_1 + w_2} \tag{6}$$

If Eq. (7) is defined as:

$$\overline{w_i} = \frac{w_i}{w_1 + w_2}, \tag{7}$$

then Eq. (6) can be written as Eq. (8):

$$f = \overline{w_1}f_1 + \overline{w_2}f_2 \tag{8}$$

In Fig. 2(b), the function of every node i in layer 1 is as Eq. (9):

$$O_{1,i} = \mu_{A_i}(x) for i = 1, 2, \quad or$$

$$O_{1,i} = \mu_{B_{i-2}}(y) for i = 3, 4 \tag{9}$$

where, x and y are node i inputs. The linguistic labels corresponding with this node are A_i and B_{i-2}. This equation specifies how much the quantifiers of A and B are satisfied with the given inputs of x and y.

In layer 2, as it is given in Eq. (10), the output of every node is the production of its inputs.

$$O_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad i = 1, 2 \tag{10}$$

In layer 3, the outputs of the nodes are calculated by Eq. (11):

$$O_{3,i} = \overline{w_i} = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \tag{11}$$

Node i in layer 4 is an adaptive node. Its function is shown in Eq. (12):

$$O_{4,i} = \overline{w_i}f_i = \overline{w_i}(p_ix + q_iy + r_i) \tag{12}$$

where, $\{p_i, q_i, r_i\}$ is the parameter set of node i. Parameters in this layer are consequence parameters.

As it is given in Eq. (13), the sum of all inputs computes the overall output in layer 5.

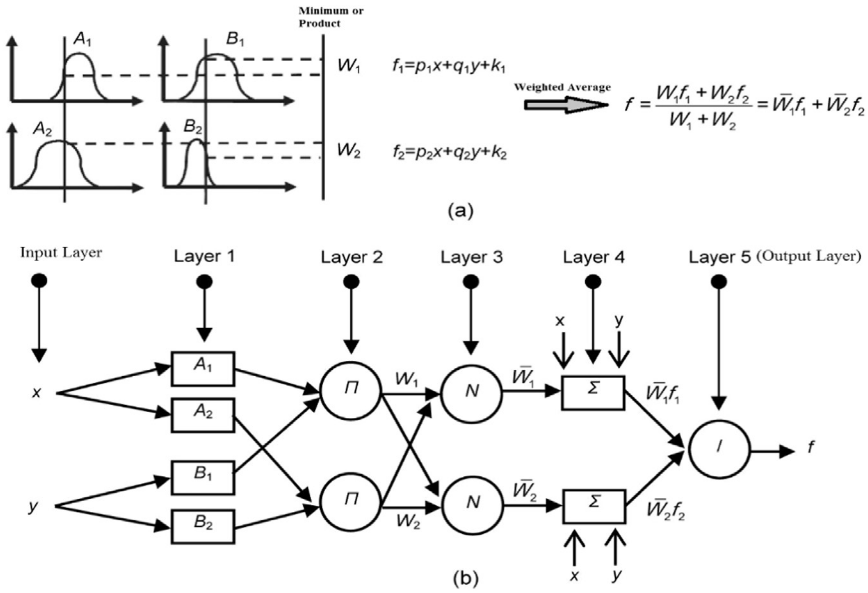

Fig. 2. (a) First-order Takagi-Sugeno fuzzy model with two inputs and two rules; (b) The equivalent ANFIS structure [24].

$$O_5 = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (13)$$

Now, an adaptive network is constructed that is equivalent to the Sugeno fuzzy model. To learn the parameters, ANFIS can be trained by a hybrid algorithm [24].

3.5. The proposed method

In this research, feature effectiveness in classification was first determined using information gain measure and is demonstrated in Tables 3 and 4. Some other measurement methods such as Gini index, filter method, and wrapper method were investigated for feature selection. However, the information gain was selected as it has the best performance on the training data. Following that, k features with higher information gain were selected as Apriori algorithm arguments. We changed k between 1 and 8, and the best accuracy rate on training data was for k=5 in both immunotherapy and cryotherapy methods. For accuracy measurement, 10-fold cross validation was applied. In each fold, the Apriori algorithm extracted the rules. Then, the rules with support and confidence values more than 0.2 and 0.6, respectively, were selected for test data classification. These thresholds for support and confidence were selected in order to prevent overfitting. Afterwards, the extracted rules were converted into Sugeno-Type fuzzy rules as our system outputs had constant values. The initial values of membership functions were determined in consultation with the physicians.

Finally, ANFIS was employed to optimize membership function parameters. These rules were applied to test data in order to determine the accuracy. The flowchart of our proposed method is illustrated in Fig. 3.

Table 3
Feature effectiveness in treatment with immunotherapy using the information gain.

Features	Weight
Time elapsed before treatment	1
Initial induration diameter test	0.769
Age	0.596
Types of wart	0.579
Surface area of the warts	0.462
Gender	0.052
The number of warts	0.032

Table 4
Feature effectiveness in treatment by cryotherapy, using the information gain.

Features	Weight
Age	1
Time elapsed before treatment	0.917
Types of wart	0.841
Surface area of the warts	0.653
The number of warts	0.107
Gender	0.03

4. Results

In this section, the effective features on wart treatment by immunotherapy and cryotherapy methods are introduced first. Then, some rules, which were obtained by association rules, are described. Finally, the results of classification algorithms are discussed.

4.1. Feature selection

Information gain was applied to determine the most important features in wart treatment, using immunotherapy. Table 3 indicates the results of using information gain in immunotherapy dataset. The higher the weight, the more effective the features.

As this table depicts, the effective treatment factors are in a descending order, the time elapsed before treatment, initial induration diameter test, age, and types of wart while the immunotherapy method was used.

The same procedure was followed to determine the most important features in the treatment of warts by cryotherapy. Table 4 illustrates the results of using information gain in cryotherapy dataset.

This table indicates that while cryotherapy was utilized, the important treatment factors are in a descending order age, the time elapsed before treatment, types of wart, and surface area of the warts.

In Tables 3 and 4, age and the time elapsed before treatment are among the most important features, while gender and the number of warts have less significant impacts.

4.2. Rules extracted from dataset

A number of rules related to immunotherapy and cryotherapy methods were extracted from the training data, using the association

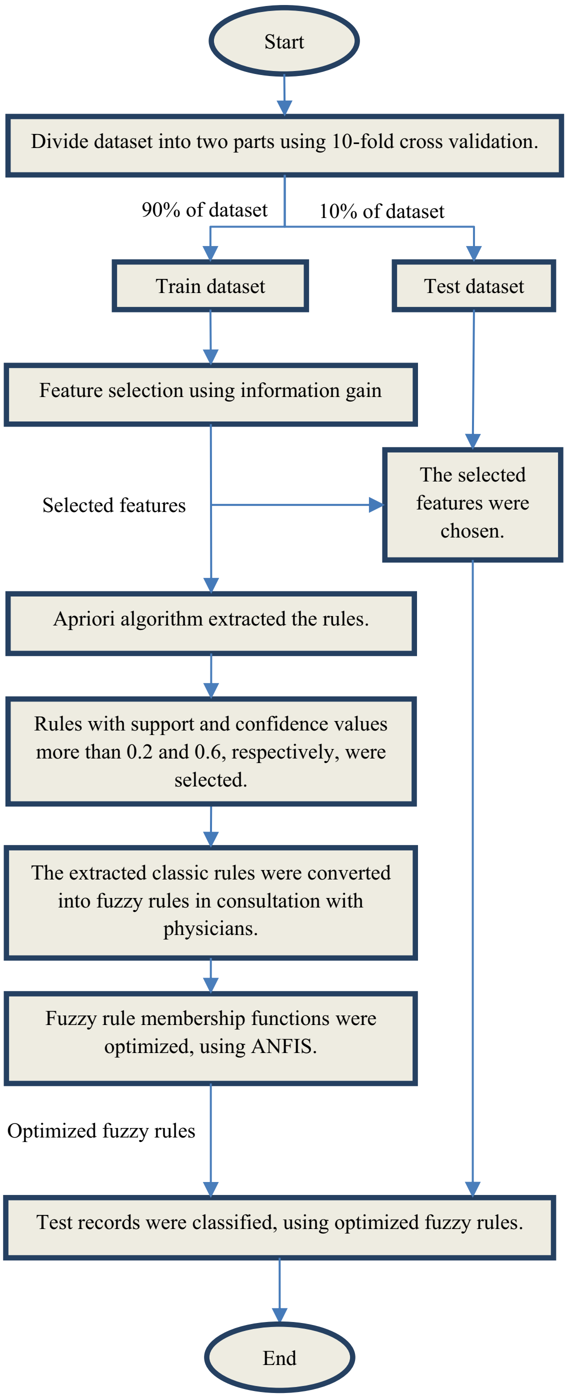

Fig. 3. The proposed method.

rules. These are some common rules in all folds. They were expressed in this article for the first time. In these rules, S and C represent the support and the confidence of the extracted rules, respectively. They are average of ten values extracted in each fold.

For cryotherapy method, seven important rules are presented as follows:

1. [Types of wart is both] = > Unresponsiveness, S=0. 3, C=0.83
2. [Age < 50 and Time elapsed before treatment < 6 months] = > Response to treatment, S=0.33, C=1
3. [Time elapsed before treatment < 6 months] = > Response to treatment, S=0.35, C=0.85
4. [Age≤21] = > Response to treatment, S=0.4, C=0.87
5. [Types of wart is common and Surface area of the warts > 63] = > Response to treatment, S =0.26, C=0.9
6. [Age≤31] = > Response to treatment, S=0.633, C=0.789
7. [Age > 31] = > Unresponsiveness, S=0.366, C=0.818

Rule 1 indicates that the patients with both types of plantar and common warts do not respond to cryotherapy treatment well. Rule 2 states that the patients who are 50 or younger and for whom the time elapsed before treatment is less than 6 months are treated by cryotherapy. Rule 4 declares that the patients younger than 21 respond to the treatment by cryotherapy very well.

The extracted rules show that both young people and the patients who referred for treatment in less than 6 months after the occurrence of warts (time elapsed before treatment < 6 months) have better chance of being treated by the cryotherapy method.

For the immunotherapy method, three important rules are given below:

1. [Time elapsed before treatment < 6 months] = > Response to treatment, S=0.35, C=0.85
2. [Induration diameter of initial test ≤7 mm] = > Response to treatment, S=0.63, C=0.789
3. [Types of wart is plantar and Time elapsed before treatment ≥6 month] = > Response to treatment, S=0.20, C=0.833

Rule 1 states that if the time elapsed before the treatment is less than 6 months, the immunotherapy treatment results will be very effective. These rules show that in this method, the time elapsed before the treatment is also important.

4.3. Conversion of extracted rules into fuzzy rules

The rules extracted by the Apriori algorithm were converted into fuzzy rules. Their input variable membership functions were determined in consultation with the physicians. Subsequently, they were optimized, using ANFIS. Fuzzy variable membership functions for cryotherapy and immunotherapy methods after the optimization are presented in Figs. 4 and . 5, respectively. Fig. 4 and Table 5 indicate the property of fuzzy system for cryotherapy dataset. As it is clear in Fig. 4, we have four variables. Each variable has its own membership function. For example, the variable of age has three membership functions. They are young, middle age, and old. The type of each membership function and their parameters are shown in Table 5. They are depicted according to Eq. (3) and Fig. 1. All types of membership function are bell-shaped functions. Parameters in Table 5 are a, b, and c in Eq. (3). Fig. 5 and Table 6 show the property of fuzzy system for immunotherapy dataset. Meanwhile, in Tables 5 and 6, parameters before and after the optimization are demonstrated for cryotherapy and immunotherapy methods, respectively. We selected the weighted average method for defuzzification and product for “logical and” [24].

In the cryotherapy group, the rules converted into fuzzy are as follows:

1. If (types_of_wart is Both) then (response_to_treatment is No) (0.83).
2. If (age is not Old) and (time_elapsed_before_treatment is not VeryLate) then (respose_to_treatment is Yes) (1).
3. If (time_elapsed_before_treatment is not VeryLate) then (respose_to_treatment is Yes) (0.85)
4. If (age is Young) then (respose_to_treatment is Yes) (0.87).
5. If (types_of_wart is Common) and (surface_area_of_warts is not Low) then (respose_to_treatment is Yes) (0.9).
6. If (age is not Old) then (respose_to_treatment is Yes) (0.789).
7. If (age is not Young) then (respose_to_treatment is No) (0.818).

In addition, in the immunotherapy group, the fuzzy rules are as follows:

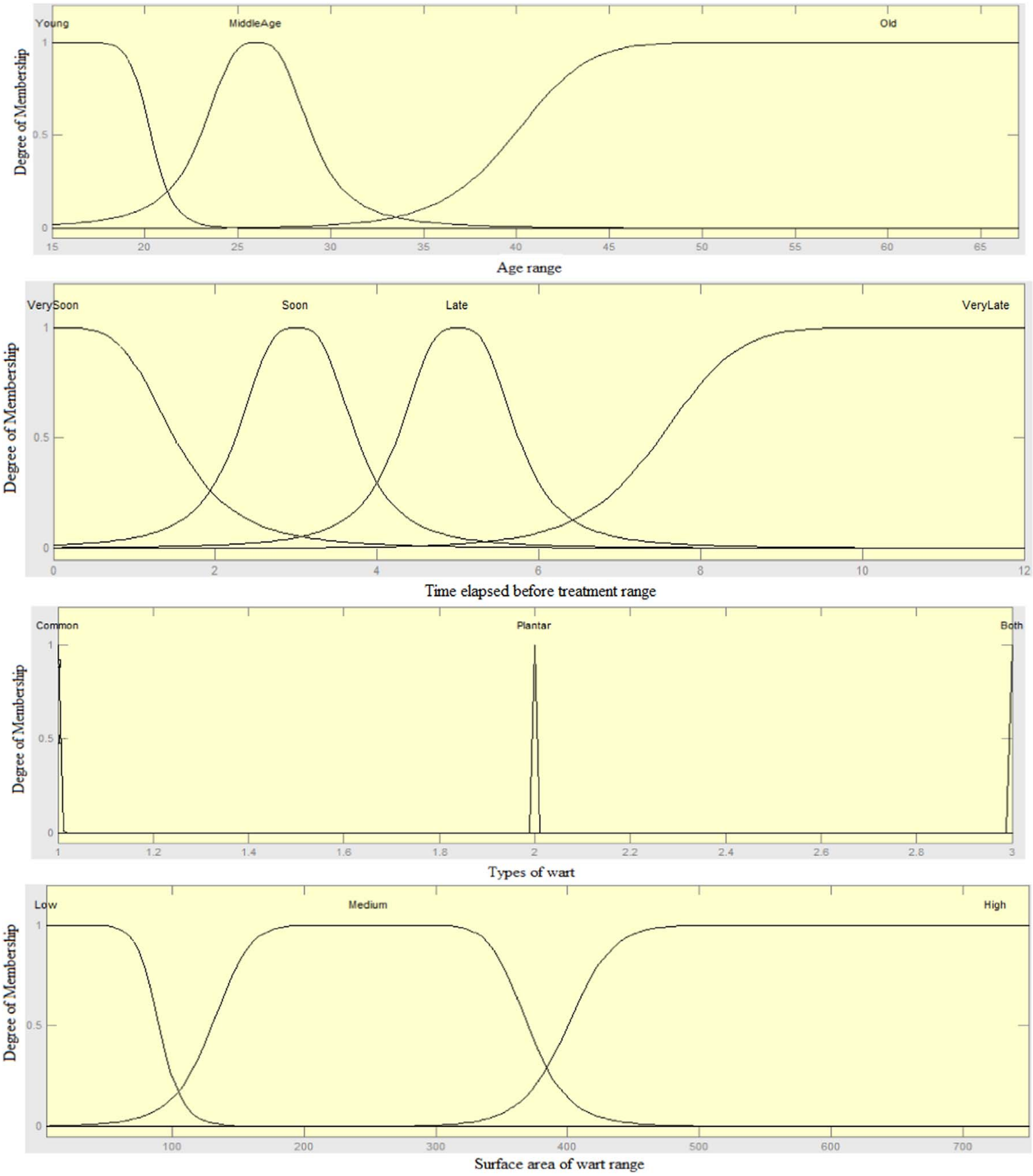

Fig. 4. Input variable membership functions for cryotherapy rules.

1. If (time_elapsed_before_treatment is not VeryLate) then (response_to_treatment is Yes) (0.85).
2. If (induration_diameter_of_initial_test is Low) then (response_to_treatment is Yes) (0.79).
3. If (types_of_wart is Plantar) and (time_elapsed_before_treatment is VeryLate) then (response_to_treatment is Yes) (0.83).

It is clear that the rules converted into fuzzy are more understandable. The numbers in parentheses represent the rule weights. They are selected according to their confidence in the Apriori algorithm.

According to Section 3.4, if all possible rules were added, we would have 108 rules for cryotherapy and 24 rules for immunotherapy. This will have some disadvantages. There will be many rules which are difficult for physicians to understand and use. Meanwhile, most of them have low support and confidence. Overfitting is another big obstacle that we will face if we use a lot of rules. So, we used the simplified version of rules.

4.4. Classification results

Table 7 compares the performance of classic rule-based method and our fuzzy rule-based method before and after training, using 10-fold cross-validation on both cryotherapy and immunotherapy datasets. According to the accuracy rate, after training, fuzzification could increase the performance of our method for approximately 10% with respect to the classic rule-based method.

The Receiver Operating Characteristic (ROC) Curves of our proposed method on cryotherapy and immunotherapy datasets are shown in Fig. 6. The Area Under Curve (AUC) of cryotherapy and immunotherapy datasets is 0.902 and 0.813, respectively. Using this method can assist physicians in predicting, with high accuracy, the better

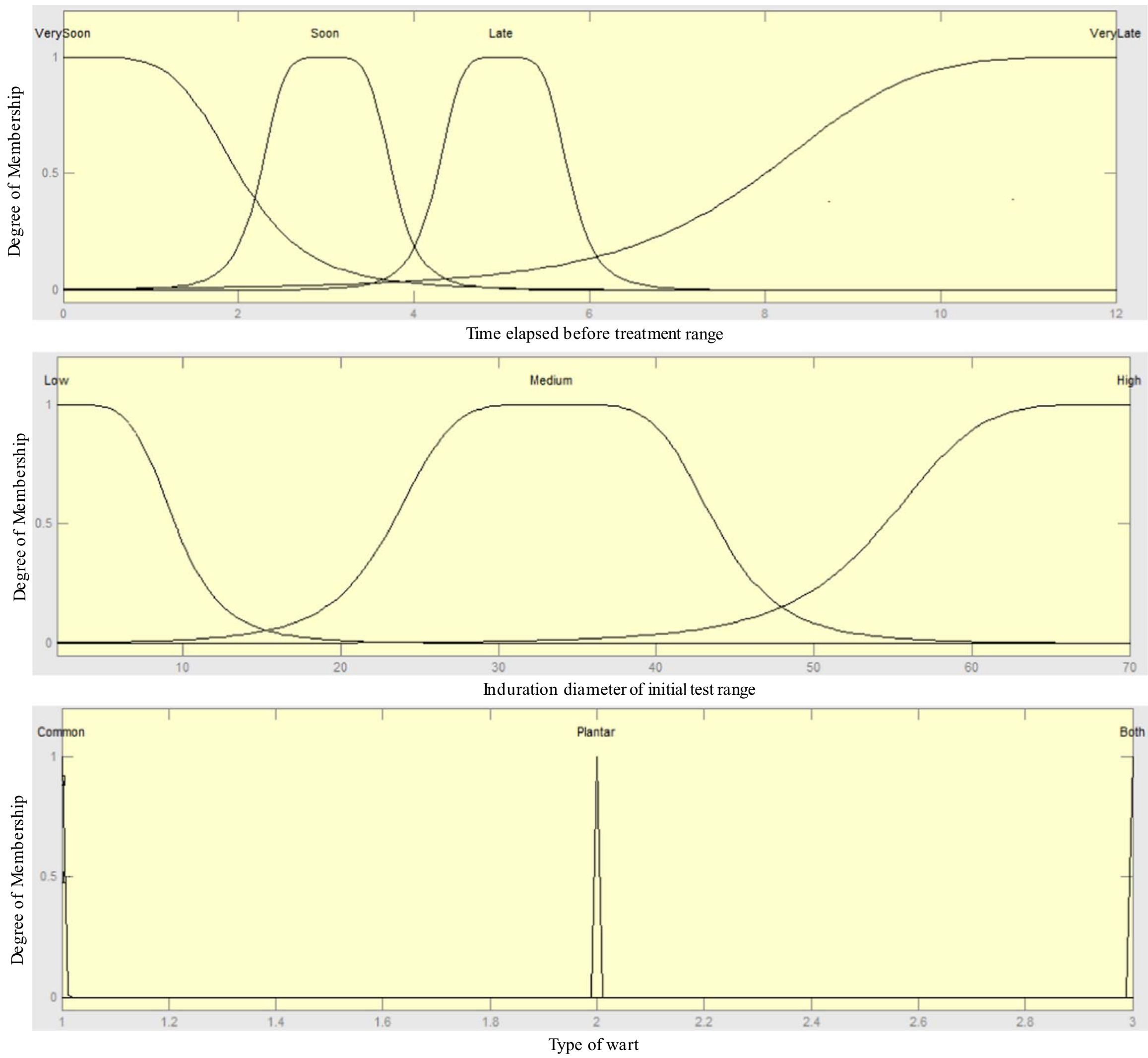

Fig. 5. Input variable membership functions for immunotherapy rules.

Table 5
Fuzzy variables and their properties for cryotherapy rules.

Variables	Range	Membership Function	Type	Parameters before optimization	Parameters after optimization
Age	[15 67]	Young	gbellmf	[6.5 5 15]	[5.93 4.96 14.4]
		Middle Age	gbellmf	[3 1.5 26]	[3 1.5 26]
		Old	gbellmf	[5,20,60]	[20.08 4.941 59.94]
Time_elapsed_before_treatment	[0 12]	Very Soon	gbellmf	[1.5 2 0]	[1.5 2 1.781]
		Soon	gbellmf	[0.75 1.5 3]	[0.75 1.5 3.751]
		Late	gbellmf	[0.75 1.5 5]	[0.75 1.5 5]
		Very Late	gbellmf	[4 4 11.5]	[5.008 4.221 10.29]
Types_of_wart	{1, 2 ,3}	Common	gbellmf	[0.00542 3.13 1]	[0.00542 3.13 1]
		Plantar	gbellmf	[0.00433 3.13 2]	[0.00433 3.13 2]
		Both	gbellmf	[0.00325 3.13 3]	[0.00325 3.13 3]
Surface_area_of_warts	[4 750]	Low	gbellmf	[140 8 -50]	[139.9 8.042 -50.05]
		Medium	gbellmf	[4,120,250]	[4,120,250]
		High	gbellmf	[10,350,750]	[10,350,750]

method according to patients' features. This method also helps patients not only by saving time and reducing the costs, but also by decreasing the side effects of ineffective treatment methods.

5. Discussion

In this novel study, we reviewed two important methods of treatment for plantar and common warts, using fuzzy rule-based system. These two methods were cryotherapy and immunotherapy with candida antigen. First, we expressed the important features for each treatment method. Then, a number of rules related to both types of treatment were extracted. Finally, by converting the rules into fuzzy rules and optimizing the fuzzy variables by ANFIS, we achieved the accuracy rates of 83.33% in predicting the treatment of plantar and common warts by immunotherapy and 80% in predicting the treatment of plantar and common warts by cryotherapy with liquid nitrogen.

Table 6
Fuzzy variables and their properties for immunotherapy rules.

Variables	Range	Membership function	Type	Parameters before optimization	Parameters after optimization
time_elapsed_before_treatment	[0 12]	Very Soon	gbellmf	[2 2.5 0]	[2 2.5 0]
		Soon	gbellmf	[0.75 2.5 3.5]	[0.75 2.5 3.028]
		Late	gbellmf	[0.75 2.5 4.5]	[0.75 2.5 5.027]
		Very Late	gbellmf	[4.5 2.5 12.5]	[3.977 3.387 12.88]
induration_diameter_of_initial_test	[2 70]	Low	gbellmf	[7.5 2.5 2.07]	[6.02 2.072 0.2132]
		Medium	gbellmf	[10.25 2.5 33.5]	[10.25 2.5 33.5]
		High	gbellmf	[16 2.5 70.5]	[16 2.5 70.5]
types_of_wart	{1 , 2, 3}	Common	gbellmf	[0.00542 3.13 1]	[0.00542 3.13 1]
		Plantar	gbellmf	[0.00433 3.13 2]	[0.00433 3.13 2]
		Both	gbellmf	[0.00325 3.13 3]	[0.00325 3.13 3]

Table 7
The performance of the classic rule-based method and fuzzy rule-based method before and after triaing using 10-fold cross-validation.

	Accuracy % (Sensitivity; Specificity)		
	Classic Rule based method	Fuzzy Rule based Method	
		Before Training	After Training
Cryotherapy	70.00 ± 7.45 (0.70; 0.69)	75.56 ± 4.33 (0.79; 0.71)	80.00 ± 5.23 (0.82; 0.77)
Immunotherapy	73.33 ± 5.98 (0.78; 0.57)	77.78 ± 6.82 (0.81; 0.63)	83.33 ± 6.02 (0.87; 0.71)

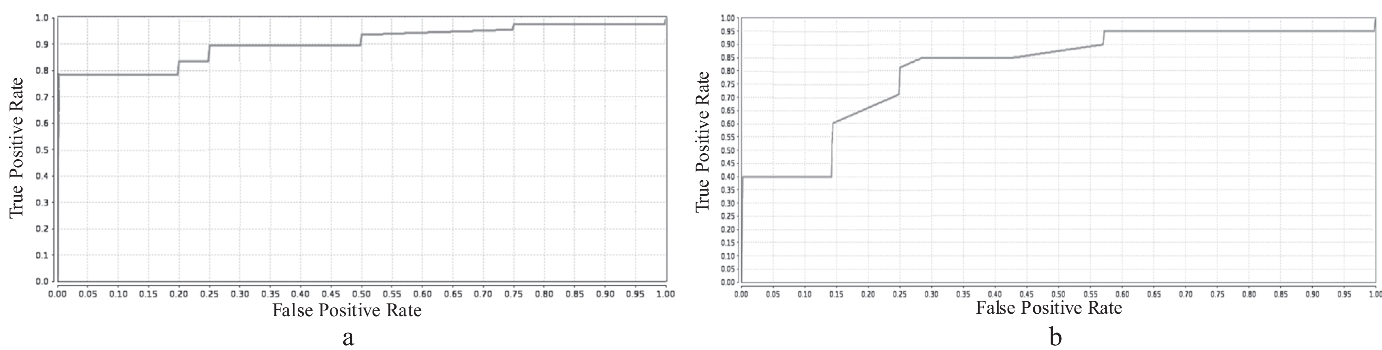

Fig. 6. ROC of proposed method on a) Cryotherapy dataset and b) Immunotherapy dataset.

6. Conclusion and future works

A computational intelligence-based expert system was developed in this paper to select the best methods for wart treatment. In the core of the proposed expert system exists a fuzzy logic system which analyzes patients' information and generates recommendation. It was successfully applied for selecting the treatment methods using the rules generated from the real data. It was observed that this system can greatly and effectively reduce both the time and cost of treatment for patients. It was found that faster treatment with lower complications is achieved through applying this system.

In a near future work, the team is going to increase the number of patients in order to obtain more precise and accurate results. Furthermore, other methods of wart treatment will be investigated for the sake of comparison. In addition, alternative machine-learning and data-mining algorithms will be applied to achieve a higher accuracy for wart treatment prediction by different methods.

Conflicts of interest

The authors have no conflicts of interest to declare.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.compbiomed.2017.01.001.

References

[1] D.E. Jones, H. Ghandehari, J.C. Facelli, A review of the applications of data mining and machine learning for the prediction of biomedical properties of nanoparticles, Comput. Methods Prog. Biomed. 132 (2016) 93–103.
[2] I. Kononenko, Machine learning for medical diagnosis: history, state of the art and perspective, Artif. Intell. Med. 23 (2001) 89–109.
[3] A. Kalinli, F. Sarikoc, H. Akgun, F. Ozturk, Performance comparison of machine learning methods for prognosis of hormone receptor status in breast cancer tissue samples, Comput. Methods Prog. Biomed. 110 (2013) 298–307.
[4] A. Majid, S. Ali, M. Iqbal, N. Kausar, Prediction of human breast and colon cancers from imbalanced data using nearest neighbor and support vector machines, Comput. Methods Prog. Biomed. 113 (2014) 792–808.
[5] J. Arevalo, F.A. González, R. Ramos-Pollán, J.L. Oliveira, M.A.G. Lopez, Representation learning for mammography mass lesion classification with convolutional neural networks, Comput. Methods Prog. Biomed. 127 (2016) 248–257.
[6] K.J. Cios, G.W. Moore, Uniqueness of medical data mining, Artif. Intell. Med. 26 (2002) 1–24.
[7] L. Li, H. Tang, Z. Wu, J. Gong, M. Gruidl, J. Zou, M. Tockman, R.A. Clark, Data mining techniques for cancer detection using serum proteomic profiling, Artif. Intell. Med. 32 (2004) 71–83.
[8] D. Delen, G. Walker, A. Kadam, Predicting breast cancer survivability: a comparison of three data mining methods, Artif. Intell. Med. 34 (2005) 113–127.
[9] R. Bellazzi, B. Zupan, Predictive data mining in clinical medicine: current issues and guidelines, Int. J. Med. Inform. 77 (2008) 81–97.
[10] A. Kupusinac, E. Stokic, I. Kovacevic, Hybrid EANN-EA system for the primary estimation of cardiometabolic risk, J. Med. Syst. 40 (2016) 1–9.
[11] A.J. Masino, R.W. Grundmeier, J.W. Pennington, J.A. Germiller, E.B. Crenshaw, Temporal bone radiology report classification using open source machine learning and natural langue processing libraries, BMC Med. Inform. Decis. Mak. 16 (2016) 1–10.
[12] L. Verma, S. Srivastava, P.C. Negi, A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data, J. Med. Syst. 40 (2016) 1–7.
[13] T. Manikandan, N. Bharathi, Lung cancer detection using fuzzy, J. Med. Syst. 40

(2016) 1–9.
[14] C. Barbieri, F. Mari, A. Stopper, E. Gatti, P. Escandell-Montero, J.M. Martínez-Martínez, J.D. Martín-Guerrero, A new machine learning approach for predicting the response to anemia treatment in a large cohort of End Stage renal disease patients undergoing dialysis, Comput. Biol. Med. 61 (2015) 56–61.
[15] C.S. Tucker, I. Behoora, H.B. Nembhard, M. Lewis, N.W. Sterling, X. Huang, Machine learning classification of medication adherence in patients with movement disorders using non-wearable sensors, Comput. Biol. Med. 66 (2015) 120–134.
[16] G. Wang, K.-M. Lam, Z. Deng, K.-S. Choi, Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques, Comput. Biol. Med. 63 (2015) 124–132.
[17] N. Memarian, S. Kim, S. Dewar, J. Engel Jr., R.J. Staba, Multimodal data and machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy, Comput. Biol. Med. 64 (2015) 67–78.
[18] N. Habibi, A. Norouzi, S.Z.M. Hashim, M.S. Shamsir, R. Samian, Prediction of recombinant protein overexpression in Escherichia coli using a machine learning based model (RPOLP), Comput. Biol. Med. 66 (2015) 330–336.
[19] L.S. Riza, C.N. Bergmeir, F. Herrera, J.M. Benítez Sánchez, frbs: fuzzy rule-based systems for classification and regression in R, J. Stat. Softw. 56 (2015) 1–30.
[20] R.A. Mohammadpour, S.M. Abedi, S. Bagheri, A. Ghaemian, Fuzzy rule-based classification system for assessing coronary artery disease, Comput. Math. Methods Med. 2015 (2015) 1–8.
[21] A. Fernández, M.J. del Jesus, F. Herrera, Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets, Int. J. Approx. Reason. 50 (2009) 561–577.
[22] H. Ishibuchi, T. Yamamoto, Rule weight specification in fuzzy Rule-based classification systems, IEEE Trans. Fuzzy Syst. 13 (2005) 428–435.
[23] S.M. Fakhrahmad, M.Z. Jahromi, Constructing accurate fuzzy classification systems: a new approach using weighted fuzzy rules, computer graphics, Imaging Vis. (2007) 408–413.
[24] J.-S.R. Jang, C.-T. Sun, Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence, Prentice-Hall, Inc, 1997.
[25] M.G. Yunusoglu, H. Selim, A fuzzy rule based expert system for stock evaluation and portfolio construction: an application to Istanbul Stock exchange, Expert Syst. Appl. 40 (2013) 908–920.
[26] A.G. Isasi, B.G. Zapirain, A.M. Zorrilla, Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms, Comput. Biol. Med. 41 (2011) 742–755.
[27] P.G. Cavalcanti, J. Scharcanski, Automated prescreening of pigmented skin lesions using standard cameras, Comput. Med. Imaging Graph. 35 (2010) 481–491.
[28] K. Korotkov, R. Garcia, Methodological review: computerized analysis of pigmented skin lesions: a review, Artif. Intell. Med. 56 (2012) 69–90.
[29] D. Ruiz, V. Berenguer, A. Soriano, B. Sánchez, A decision support system for the diagnosis of melanoma: a comparative approach, Expert Syst. Appl. 38 (2011) 15217–15223.
[30] P.G. Cavalcanti, J. Scharcanski, G.V.G. Baranoski, A two-stage approach for discriminating melanocytic skin lesions using standard cameras, Expert Syst. Appl. 40 (2013) 4054–4064.
[31] I. Giotis, N. Molders, S. Land, M. Biehl, M.F. Jonkman, N. Petkov, MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images, Expert Syst. Appl. 42 (2015) 6578–6585.
[32] H. Mirzaalian, T.K. Lee, G. Hamarneh, Skin lesion tracking using structured graphical models, Med. Image Anal. 27 (2016) 84–92.
[33] H. Vasudevan, A.R. Joshi, N.M. Shekokar, R. Sumithra, M. Suhil, D.S. Guru, Segmentation and classification of skin lesions for disease diagnosis, Procedia Comput. Sci. 45 (2015) 76–85.
[34] Q. Abbas, I. Fondón, M. Rashid, Unsupervised skin lesions border detection via two-dimensional image analysis, Comput. Methods Prog. Biomed. 104 (2011) e1–e15.
[35] R.J. Stanley, R.H. Moss, W. Van Stoecker, C. Aggarwal, A fuzzy-based histogram analysis technique for skin lesion discrimination in dermatology clinical images, Comput. Med. Imaging Graph. 27 (2003) 387–396.
[36] R.J. Stanley, W.V. Stoecker, R.H. Moss, H.S. Rabinovitz, A.B. Cognetta, G. Argenziano, H.P. Soyer, A basis function feature-based approach for skin lesion discrimination in dermatology dermoscopy images, Ski. Res. Technol. 14 (2008) 425–435.
[37] M.E. Celebi, Q. Wen, S. Hwang, H. Iyatomi, G. Schaefer, Lesion border detection in dermoscopy images using ensembles of thresholding methods, Ski. Res. Technol. 19 (2013) e252–e258.
[38] S. Jain, V. jagtap, N. Pise, Computer aided melanoma skin cancer detection using image processing, Procedia Comput. Sci. 48 (2015) 735–740.
[39] R.B. Oliveira, N. Marranghello, A.S. Pereira, J.M.R.S. Tavares, A computational approach for detecting pigmented skin lesions in macroscopic images, Expert Syst. Appl. 61 (2016) 53–63.
[40] E. Flores, J. Scharcanski, Segmentation of melanocytic skin lesions using feature learning and dictionaries, Expert Syst. Appl. 56 (2016) 300–309.
[41] C.M. Balch, J.E. Gershenwald, S.J. Soong, J.F. Thompson, M.B. Atkins, D.R. Byrd, A.C. Buzaid, A.J. Cochran, D.G. Coit, S. Ding, A.M. Eggermont, K.T. Flaherty, P.A. Gimotty, J.M. Kirkwood, K.M. McMasters, M.C. Mihm, D.L. Morton, M.I. Ross, A.J. Sober, V.K. Sondak, Final version of 2009 AJCC melanoma staging and classification, J. Clin. Oncol. 27 (2009) 6199–6206.
[42] H. Lamminen, K. Ruohonen, H. Uusitalo, Visual tests for measuring the picture quality of teleconsultations for medical purposes, Comput. Methods Prog. Biomed. 65 (2001) 95–110.
[43] S. Ribarič, L. Todorovski, J. Dimec, T. Lunder, Presentation of dermatological images on the Internet, Comput. Methods Prog. Biomed. 65 (2001) 111–121.
[44] K. Bunte, M. Biehl, M.F. Jonkman, N. Petkov, Learning effective color features for content based image retrieval in dermatology, Pattern Recognit. 44 (2011) 1892–1902.
[45] V.K. Shrivastava, N.D. Londhe, R.S. Sonawane, J.S. Suri, Exploring the color feature power for psoriasis risk stratification and classification: a data mining paradigm, Comput. Biol. Med. 65 (2015) 54–68.
[46] D. McGibbon, Rook's textbook of dermatology (7th edition)Clin. Exp. Dermatol. 31 (2006) 178–179.
[47] D.W. Russell, W.A. McCann, H.E. Maroncelli, Candida and subsequent cell-mediated-panel driven intralesional immunotherapy of common warts in children and adults, J. Allergy Clin. Immunol. 125 (2010) AB204.
[48] M.M. Clifton, S.M. Johnson, P.K. Roberson, J. Kincannon, T.D. Horn, immunotherapy for recalcitrant warts in children using intralesional mumps or candida antigens, Pediatr. Dermatol. 20 (2003) 268–271.
[49] A. Nofal, E. Nofal, Intralesional immunotherapy of common warts: successful treatment with mumps, measles and rubella vaccine, J. Eur. Acad. Dermatol. Venereol. 24 (2010) 1166–1170.
[50] T.D. Horn, S.M. Johnson, R.M. Helm, P.K. Roberson, Intralesional immunotherapy of warts with mumps, Candida, and Trichophyton skin test antigens: a single-blinded, randomized, and controlled trial, Arch. Dermatol. 141 (2005) 589–594.
[51] S.M.JOHNSON, P.K.ROBERSON, T.D.HORN, Intralesional injection of mumps or Candida skin test antigens. A Novel Immunotherapy for Warts, American Medical Association, Chicago, IL, ETATS-UNIS, 2001.
[52] M. Maronn, C. Salm, V. Lyon, S. Galbraith, One-year experience with candida antigen immunotherapy for warts and molluscum, Pediatr. Dermatol. 25 (2008) 189–192.
[53] H. Gamil, I. Elgharib, A. Nofal, T. Abd-Elaziz, Intralesional immunotherapy of plantar warts: report of a new antigen combination, J. Am. Acad. Dermatol. 63 (2010) 40–43.
[54] K. Khurshid, S.S. Pal, Role of candida antigen in treatment of viral warts: a placebo-controlled study, J. Pak. Assoc. Dermatol. 19 (2009) 146–150.
[55] N.B. Silverberg, J.K. Lim, A.S. Paller, A.J. Mancini, Squaric acid immunotherapy for warts in children, J. Am. Acad. Dermatol. 42 (2000) 803–808.
[56] R. Alizadehsani, J. Habibi, M.J. Hosseini, H. Mashayekhi, R. Boghrati, A. Ghandeharioun, B. Bahadorian, Z.A. Sani, A data mining approach for diagnosis of coronary artery disease, Comput. Methods Prog. Biomed. 111 (2013) 52–61.
[57] G. Piatetsky-Shapiro, Discovery, Analysis, and Presentation of Strong Rules, Knowl. Discov. Databases (1991) 229–238.
[58] R.Agrawal, R.Srikant, Fast Algorithms for Mining Association Rules in Large Databases, in: Proceedings of the 20th International Conference on Very Large Data Bases, Morgan Kaufmann Publishers Inc., 1994 pp. 487–499.
[59] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[60] A. Hamam, N.D. Georganas, A comparison of Mamdani and Sugeno fuzzy inference systems for evaluating the quality of experience of Hapto-Audio-Visual applications, in: Proceedings of the IEEE International Workshop on Haptic Audio visual Environments and Games, pp. 87–92, 2008.

Fahime Khozeimeh is currently a MS student in Mashhad University of Medical science. Fahime's research interests include in the areas of cardiovascular disease, bioinformatics, eye disease, skin disease and cardiovascular imaging. She has published 6 journal papers and one US patent.

Roohallah Alizadehsani obtained a Bachelor of Science degree in computer engineering-software from Sharif University of Technology, and then received a Master in Science in computer engineering-software from Sharif University of Technology in 2012. He has worked in Software engineering Lab under supervision of an Associate Professor, Jafar Habibi since 2011. Roohallah’s research interests include mainly in the areas of data mining, machine learning, bioinformatics, and evaluation of computer systems performance. He has published 8 journal papers and one US patent. His approach for these topics includes new algorithms in data mining and mathematically prove them. He is a reviewer of Computer Methods and Programs in Biomedicine.

Mohamad Roshanzamir received BSc in computer engineering-software from Isfahan University of Technology, Iran 2003, MSc in computer engineering-artificial intelligence From University of Isfahan, Iran 2006. Currently he is a PhD candidate in Isfahan University of Technology, Iran. His primary research interests include data mining, machine learning and evolutionary optimization.

Abbas Khosravi received BSc in Elec. Eng. from Sharif University of Technology, Iran 2002, MSc in Elec. Eng. From Amir kabir University of Technology, Iran 2005, and PhD from Deakin University, Australia 2010. Currently he is a senior research fellow in the Center for Intelligent Systems Research (CISR) at Deakin University. His primary research interests include artificial intelligence and evolutionary optimization and their applications for modeling, pattern recognition, and decision-making. Mr Khosravi is recipient of Alfred Deakin Postdoctoral Research Fellowship in 2011.

Pouran Layegh is Associate Professor of Dermatology. She has published more than 50 journal and 10 conference papers and 3 books.

Saeid Nahavandi received a Ph.D. from Durham University, U.K. He is an Alfred Deakin Professor, Chair of Engineering, and the Director of the Center for Intelligent Systems Research at Deakin University, Australia. He has published over 450 papers in various international journals and conferences. His research interests include modeling of complex systems, robotics and haptics. He is the Co-Editor-in-Chief of the IEEE Systems Journal, an Editor (South Pacific Region) of the International Journal of Intelligent Automation and Soft Computing. He is a Fellow of Engineers Australia (FIEAust), the Institution of Engineering and Technology (FIET) and Senior member of IEEE (SMIEEE).