Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

Article

A Quantum Hybrid PSO Combined with Fuzzy K-NN Approach to Feature Selection and Cell Classification in Cervical Cancer Detection

Abdullah M. Iliyasu [1, 3, 4, *] **and Chastine Fatichah** [2]

[1] Electrical Engineering Department, College of Engineering, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, KSA
[2] Informatics Department, Institut Nepuluh Nopember, ITS Campus, Surabaya 60111, Indonesia
[3] School of Computing, Tokyo Institute of Technology, Yokohama 226-8502, Japan
[4] School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China
* Correspondence: a.iliyasu@psau.edu.sa; +966-115-888-8259

Abstract: A quantum hybrid (QH) intelligent approach that blends the adaptive search capability of the quantum-behaved particle swarm optimisation (QPSO) method with the intuitionistic rationality of traditional fuzzy *k*-nearest neighbours (Fuzzy *k*-NN) algorithm (known simply as the Q-Fuzzy approach) is proposed for efficient feature selection and classification of cells in cervical smeared (CS) images. From an initial multitude of seventeen (17) features describing the geometry, colour, and texture of the CS images, the QPSO stage of our proposed technique is used to select the best subset features (i.e. global best particles) that represent a pruned down collection of seven (7) features. Using a dataset of almost 1000 images, performance evaluation of our proposed Q-Fuzzy approach assesses the impact of our feature selection on classification accuracy by way of three experimental scenarios that are compared alongside two other approaches: The All-features (i.e. classification without prior feature selection) and another hybrid technique combining the standard PSO algorithm with the Fuzzy *k*-NN technique (P-Fuzzy approach). In the first and second scenarios, we further divided the assessment criteria in terms of classification accuracy based on the choice of best features and those in terms of the different categories of the cervical cells. In the third scenario, we introduced new QH hybrid techniques, i.e. QPSO combined with other supervised learning methods, and compared the classification accuracy alongside our proposed Q-Fuzzy approach. Furthermore, we employed statistical approaches to establish qualitative agreement with regards to the feature selection in scenarios 1 and 3. The synergy between the QPSO and Fuzzy *k*-NN in the proposed Q-Fuzzy approach marginally improves classification accuracy as manifest in the reduction in number cell features, which is crucial for effective cervical cancer detection and diagnosis.

Keywords: computational intelligence; quantum hybrid intelligent systems; quantum machine learning; medical image processing; disease diagnosis; Fuzzy *k*-NN; Quantum-behaved PSO; cervical smear images; cancer detection

1. Introduction

Hybrid intelligent systems (HIS) simultaneously integrate (or combine) two or more intelligent approaches, such as fuzzy techniques, genetic algorithms, neural networks, agent-based techniques, case-based reasoning and other computationally (or artificially) intelligent approaches, conducive to overcome individual limitations and achieve synergetic outcomes. Such hybridisation offers the capability to handle real world complex problems involving imprecision, uncertainty, vagueness, high-dimensionality, etc. [1]. HIS systems are applied in almost every area of life, but notable applications can be found in science, technology, business, commerce, and medicine.

© 2017 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

Cervical cancer is one of the most common lethal malignant diseases among women. Thankfully, however, with improvement in medical technologies, over the past few years, it is easier to detect such disease at an early stage by doing Pap smear image tests. These tests typically involve filtering out abnormal cervical cells [2] and use of the results to detect precancerous changes in cervical cells based on colour and shape properties of their nuclei and cytoplasm [3]. Screening for cervical cancer by Pap smear image tests provide an inference regarding the presence of the Papilloma virus that is responsible for cervical cancer [4]. However, performing the Pap smear test manually is known to be a time consuming and error-prone exercise that is further exacerbated by the lack of adequate pathology expertise [5]. In addition, due to subjective disparity from different cytologists, the results of the screening often show a lot of inconsistencies [6]; thereby further compromising the screening process [7]. Moreover, since hundreds of images need to be analysed daily, the manual screening process to classify the cells is a challenging pursuit that is susceptible to error [8]. A Single cell from the Pap smear test can be classified into one of seven classes [9-12], which are Superficial squamous, Intermediate Squamous, Columnar, Mild dysplasia, Moderate dysplasia, Severe dysplasia, and Carcinoma in situ.

Meanwhile, in furtherance of improving the accuracy of automated cervical cell detection systems, dynamic segmentation techniques are required to delineate the contours of the cytoplasm and nucleus in the cell images from Pap smear tests. This has led to the proposal of numerous approaches aimed at improving the assessment of Pap smear image test results.

Some studies related to the cervical cancer detection classification of various cervical cell features are presented in [8-14]. The study in [8] uses fourteen (14) features and five (5) classifiers to validate their classification results with focus on digital imaging colposcopy. In [11], a neuro-fuzzy method was used to classify the twenty (20) features in cervical cell types; while in [13] an automatic cervical cell segmentation and classification method was applied on three datasets of Pap smeared images. Therein, nine (9) features were used and the results obtained were compared alongside five other classification methods. A support vector machine technique based on recursive feature elimination (SVM-RFE) was used to select the features and classify the cervical cell types in [14]. In that study, eleven (11) nuclei features and nine (9) cytoplasm features were used to differentiate the cervical cell types. The study in [14] combined four (4) feature selection approaches with the traditional support vector machine (SVM) algorithm in order to classify the cell types. It was reported therein that the accuracy of the classification results depended on good choice of features, which further indicates the importance of choosing the best subset of features enhances to accuracy of classification results.

In terms of feature selection, three (3) methods: filter, wrapper, and hybrid methods, are widely used in the literature. The filter method is considered computationally fast, easy to interpret, and is scalable for high-dimensional data [15]. Nonetheless, equipped with advanced machine learning algorithms to select the best feature subsets, the wrapper and hybrid methods are known to demonstrate better performances than the filter methods. Besides, hybrid methods commonly use supervised learning techniques and swarm-based intelligent methods as integral components of their feature selection. Many previous studies utilised the swarm intelligence algorithms for their feature selection stages, among which is the particle swarm optimisation (PSO) that is widely used to solve optimisation problems [17].

Recently, quantum machine learning, which is an approach integrating quantum mechanics into traditional machine learning approaches, has been used to further understand and enhance the learning process. One of such techniques is the quantum-behaved particle swarm optimisation (QPSO), which is a variant of the standard PSO algorithm, that was proposed in [18] by exploiting some proven properties of quantum mechanics. Among other superlative properties, QPSO eliminates the velocity term and control parameters that are used in the traditional PSO approach [17]. This ensures that, in comparison with the original PSO algorithm [18], the QPSO offers improved performance in terms of its search capability.

Since that effort, numerous approaches have been suggested to further enhance the QPSO. Among them, a new QPSO (NQPSO) algorithm was proposed in [19] to further improve on the QPSO by employing and balancing the choice of local and one global neighbourhood search strategies.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

Similarly, in [20] an improved QPSO (IQPSO) algorithm was proposed and utilised for visual features selection (VFS). Overall, the standard QPSO approach was proposed to deal premature convergence, and it is simple and easy to understand [19]. The use of a global optimal to determine the best subset features makes the QPSO a veritable choice to accelerate convergence in feature selection tasks.

Therefore, the study presented in this work exploits the proven versatility of the QPSO algorithm as the main component of a new hybrid approach for selecting features of cervical cells in Pap smeared images by blending the QPSO with the intuitionistic descriptiveness of Fuzzy *k*-Nearest Neighbours (Fuzzy *k*-NN) algorithms. Specifically, the best choice of subset features is enhanced by combining the QPSO with the Fuzzy *k*-NN, which itself is an extension of the k-nearest neighbours (*k*-NN) algorithms but with fuzzy intuition integrated into it [21-23]. As envisaged, this hybridisation leads to modest improvements in cervical cell classification accuracy.

To validate the expected potency of the proposed technique, the Herlev dataset [24], which contains original and segmented images collated using the CHAMP software [10-12], is used. To establish the cogency of the outcomes from our proposed technique, we utilised seventeen (17) geometric, chromatic, and textural features that are employed in some available literature [3, 11, 12] to describe our cervical cell images. The acuteness of the QPSO unit of the proposed technique ensures that the best features are selected and then pruned down to a collection of seven (7) features. This feature selection step, combined with the rationality from the Fuzzy *k*-NN guarantees improvement in cell classification accuracy, which is crucial in cervical cancer prediction.

The remainder of the paper is organised as follows: advances in cervical cell classification for smeared images are highlighted in the next section as well as a succinct overview of the QPSO and Fuzzy *k*-NN algorithms including arguments supporting their adoption as the core units of our proposed approach to feature selection and cell classification for cervical cancer detection. These two units and how they coalesce into our quantum hybrid (QH) system are discussed in Section 3, while experimental results to demonstrate the utility of our proposed technique are reported in Section 4.

2. **Overview of cell classification in smeared cervical images, quantum-behaved PSO and fuzzy K-NN**

As mentioned earlier, in this section we will highlight the advances made in cell classification for smeared cervical images. Additionally, we will present a succinct overview of the QPSO and Fuzzy *k*-NN algorithms and conclude with a few arguments supporting their adoption as the core units of our proposed approach to feature selection and cell classification in cervical cancer detection.

2.1. Cell classification in smeared cervical images

Cervical cancer is a malignant cancer that forms in cervical tissues (i.e. the organ that connects the uterus to the vagina) [14]. The Pap smear image test is one of the first procedures used to extract medical inference regarding the presence of the Papilloma virus, which is known to be responsible for causing cervical cancer [3]. Pap Smear image tests also provide a window for early detection and treatment before the condition deteriorates [6]. A single cell pap smear image can be classified into one of the seven classes presented in Table 1 [9-12].

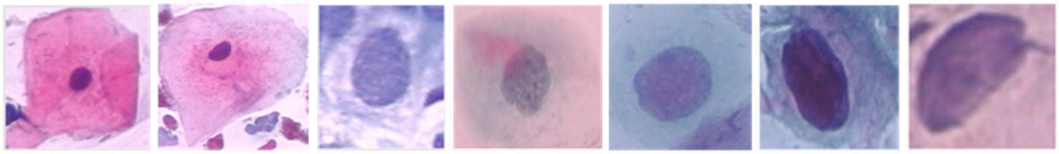
As seen from the table, abnormal cervical cells that have undergone precancerous changes (called a dysplastic cell) are further divided into four main phases. The first phase, called mild dysplasia, occurs when the nucleus grows larger and brighter than normal. In the second phase, called moderate dysplasia, the nucleus becomes darker and larger in size. The third phase, known as severe dysplasia, occurs when the size and texture of both the nucleus and cytoplasm change: the nucleus grows (in size) while becoming darker with strange shapes; and the cytoplasm also becomes darker but smaller in size. The last phase, known as Carcinoma in situ, is characterised by a very large nucleus and it occurs at a point when cytology strongly suggests malignancy.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

Table 1. Classification of Pap smear single image cell types

1	Normal cells	Superficial squamous
		Intermediate squamous
		Columnar
2	Abnormal cells	Mild dysplasia
		Moderate dysplasia
		Severe dysplasia
		Carcinoma in situ

Knowledge of these cell attributes allow us to describe cells in cancerous and precancerous stages since they are marked by many changes in morphology and architecture, including geometry (in terms of both shape and size) of the cytoplasm and nucleus, changes in nuclear-cytoplasm ratios, and others [15-17]. Some sample results for Pap smear image tests are shown in Figure 1.



a b c d e F g

Figure 1. A single cell cervical smear image: (a) Superficial squamous, (b) Intermediate Squamous, (c) Columnar, (d) Mild dysplasia, (e) Moderate dysplasia, (f) Severe dysplasia, (g) Carcinoma in situ [9-11].

2.2. Quantum-behaved Particle Swarm Optimisation (QPSO)

Quantum computing is a new and exciting computing paradigm that is tailored towards exploiting the physical attributes of quantum mechanics to harness information processing [25]. Many new applications have been proposed either for use with quantum computing hardware or just to exploit some of the confounding properties of quantum mechanics to enhance traditional computing protocols, notably those in the areas of image processing, quantum machine learning (QML), and general areas of computational intelligence [25-28].

QML, which incorporates integration of 'quantumness' into traditional machine learning algorithms, and its application in different domains is an emerging sub-discipline that seems to receiving increased attention [34]. Among others, a major objective of this study is to explore the integration of QML techniques (specifically, QPSO) in image-based cervical cancer detection [34, 35].

We begin our discussion of the QPSO with a pithy highlight of the standard particle swarm optimisation (PSO) algorithm. PSO was proposed (by Eberheart and Kennedy in 1995 [21]) to mimic the social behaviour of a swarm of birds searching for food in a predefined space with each of them behaving in accordance with the expected intelligence of the swarm population in such an environment [20]. Furthermore, the scenario imposes the additional restriction that there is only one source of food and none of the birds has prior knowledge of this location. Although rather inefficacious, based on the foregoing scenario, the easiest solution would be to trail any bird that perchance stumbles into this location. Consequently, all the birds in the population would traverse (albeit, randomly) the same path to the food source irrespective of their own proximity to the source and the remaining birds in the population. In PSO parlance, each single solution in the search space is called a particle, like a bird. PSO is initialised with random particles (solutions) and an optimal solution is determined within the search space by updating each generation. All particles have fitness values and velocities with which they fly over the search space as they follow paths covered by particles that are perceived as better solutions [19].

The d dimension of the i^{th} particle is represented as $X_i = (x_{i1}, x_{i2},\ldots,x_{id})$ and for each generation each particle is updated using two 'best' values [15]. The first of these solutions, called P_{i_best} (personal best), is the best solution for a specific particle, while the second one, called P_{global} (global best) is the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

best value of any particle in the population. The fitness function of particle X (denoted as $F(X)$), is defined using the *F1 score* (which will be presented much later in Eq. (29)).

The quantum-behaved version of the PSO (i.e. the QPSO) was proposed to improve on the capabilities of the PSO algorithm. In it, the probability of the particle appearing in position X may be obtained from the quantum mechanical interpretation of the wave function of the particle at current position (t) as described in Eq. (1) [20].

$$\psi\big(x(t)\big) = \frac{1}{\sqrt{Q}} e^{-\frac{|m_{best}(t) - X(t)|}{Q}} \tag{1}$$

where the parameter Q depends on the mean of best and current positions of the particles and it helps to specify the search scope for a particle; and m_{best} is calculated as the mean of the best positions of all particles (S) in the population, such that:

$$m_{best} = \frac{1}{S} \sum P_{i\_best}\,. \tag{2}$$

The parameter Q and the position X are updated according to the constraints in Eqs. (3) to (5).

$$Q(t) = 2. \propto . |m_{best}(t) - X(t)|, \tag{3}$$

$$X(t+1) = p(t) - \propto . |m_{best}(t) - X(t)| . ln \frac{1}{u(t)}, \qquad \text{if } s(t) \geq 0.5 \tag{4}$$

and

$$X(t+1) = p(t) - \propto . |m_{best}(t) - X(t)| . ln \frac{1}{u(t)}, \qquad \text{if } s(t) < 0.5, \tag{5}$$

where u and s are uniformly distributed random numbers in the interval (0, 1); the parameter α is the contraction-expansion coefficient [18]; and $p(t)$ takes the form defined in Eq. (6).

$$p(t) = \phi(t) . P_{i\_best}(t) + \big(1 - \phi(t)\big) . P_{global}(t), \tag{6}$$

where ϕ is a uniformly distributed random number in the interval (0,1), and to update the new best position of particle I ($P_{\text{i\_best}}$ (t+1)), Eq. (7) is used.

$$P_{i\_best}(t+1) = \begin{cases} X(t), & if\ F\big(X(t)\big) > F\left(P_{i\_best}(t)\right) \\ P_{i\_best}(t), & if\ F\big(X(t)\big) \leq F\left(P_{i\_best}(t)\right) \end{cases} \tag{7}$$

The Pseudocode in Table 2 outlines the steps to execute the QPSO algorithm (as discussed in [18, 19]).

Table 2. Pseudocode for executing the QPSO algorithm [18, 19].

Initialise the current positions and the P_{i_best} positions of all the particles
Do
Calculate m_{best} in Eq. (2)
Select a suitable value for α
For particles i = 1 to S
1. Calculate the fitness value of particle i according to classification accuracy
2. Update P_i_best and P_{global} in Eq. (7)
3. For dimension 1 to d
φ = rand (0,1)
u = rand (0,1)
If s = rand (0,1) ≥ 0.5
update particle positions in Eq. (4)
else
update particle positions in Eq. (3)
Until terminal condition is satisfied.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

As outlined in the pseudocode, each particle is encoded into a binary string whose length is equal to the size of feature vector. A value '1' indicates that a feature is selected whereas a value '0' indicates otherwise. In this manner, each particle is encoded into a binary string as presented in Eq. (8) and further depicted in Figure 2 [19].

$$x = \begin{cases} 1, & \text{if } Sigmoid(x) > U(0,1) \\ 0, & \text{otherwise} \end{cases}, \quad (8)$$

where $Sigmoid(x) = 1/(1+e^{-x})$ and $U(0,1)$ is a uniformly distributed random number in the interval (0,1).

The 100101 six-bit string depicting a particle (shown in Figure 2), shows that the first, fourth and sixth features are selected, whereas the second, third and fifth features are not selected. The fitness value of each particle is calculated based on the accuracy of the classification results.

Particle

1	0	0	1	0	1

Figure 2. Illustration of particle encoded as a binary string where the bit value '1' denotes a selected feature and '0' denotes a non-selected feature.

2.3. Fuzzy k-nearest neighbours (Fuzzy k-NN)

Fuzzy *k*-nearest neighbours (Fuzzy *k*-NN) is an extension of the standard *k*-nearest neighbours (*k*-NN) algorithm but with fuzzy intuition is integrated into it [22]. More precisely, fuzzy theory is used to generalise definitions of the *k*-NN membership values of the data in each class as defined in Eq. (9).

$$u(x, c_i) = \frac{\sum_{k=1}^{K} u(x_k, c_i).\, d(x, x_k)^{\frac{-2}{m-1}}}{\sum_{k=1}^{K} d(x, x_k)^{\frac{-2}{m-1}}}, \quad (9)$$

where $u(x,\ c_i)$ is membership values of data x in the class c_i; k value is the number of nearest neighbours; $u(x_k,\ c_i)$ is membership value of k nearest neighbours' data x in the class c_i; $d(x,\ x_k)$ is distance between data x and k nearest neighbours; and m is weight exponent, which should be greater than 1.

The pseudocode in Table 3 outlines the Fuzzy *k*-NN algorithm (as discussed in [22, 23]).

Table 3. Pseudocode for executing the Fuzzy *k*-NN algorithm [22, 23].

Normalise the data
Find the *k*-nearest neighbours
Calculate the membership value $u(x,\ c_i)$ in Eq. (9)
Select the maximum value of c from $u(x,\ c_i)$
Assign class c to the data

3. Methodology for quantum hybrid approach to cervical cancer feature selection and classification

Pap cervical smear (CS) images are rich in various features like colour, shape, and texture. The process of accurate extraction of these unique visual features from the images is very crucial in developing an automated cancer screening process [8].

The layout of our proposed quantum hybrid technique (QHT) that combines the quantum-behaved PSO with the Fuzzy *k*-NN (henceforth referred to as the Q-Fuzzy approach) to select appropriate features for accurate classification of cervical cells in smeared images is presented in Figure 3.

As seen in that figure, the inputs are the original and segmented versions of the cervical smear images. In the first stage of the proposed technique, features relevant to the colour, shape or texture of the input smeared images are extracted. Following this, in the second stage, the proposed QH (i.e. the combination of QPSO and Fuzzy *k*-NN) algorithm is used for feature selection and subsequent cell classification. The output of the system provides an inference about the predicted class from data testing of cervical smear images.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

The flowchart in Figure 4 further highlights the steps of our proposed feature selection and classification using proposed the QH or Q-Fuzzy (i.e. QPSO blended with Fuzzy *k*-NN) approach.

To ensure effective cell classification, the feature extraction stage is designed to target seventeen (17) features related to the geometry, colour, and texture of the input CS images. These features have been widely cited in previous studies, notably [3, 11, 13]. However, for integrality, we further define these features in Eqs. (10) to (26).

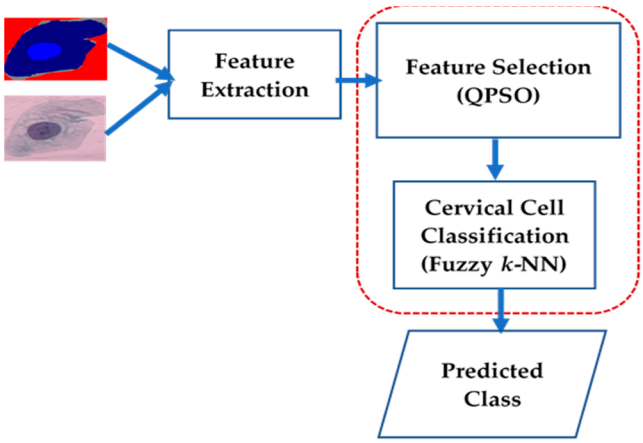


Figure 3. Layout of proposed quantum hybrid (Q-Fuzzy) technique to select and classify cells in smeared cervical images

a. Area of nucleus, A_n

A_n = number of pixels in the nucleus region (10)

b. Major axis of nucleus, L_n

L_n = the length of the major axis of an ellipse that completely encloses the nucleus region. (11)

c. Minor axis of nucleus, D_n

D_n = the length of the minor axis of an ellipse that completely encloses the nucleus region. (12)

d. Aspect ratio of nucleus, R_n

$$R_n = \frac{W_n}{H_n}, \quad (13)$$

where W_n is the width of the nucleus and H_n is the height of the nucleus region.

e. Perimeter of Nucleus, P_n

P_n = the perimeter of the nucleus region, (14)

f. Roundness of nucleus, N_{circle}

$$\mathrm{N_{circle}} = \frac{\pi}{4} . {L_n}^2 \rightarrow N_{\mathrm{roundness}} = \frac{A_n}{\mathrm{N_{circle}}}. \quad (15)$$

g. Maxima of nucleus, Max_n

Max_n = number of local maximum value from eight pixels in the neighbourhood of the nucleus region. (16)

h. Minima of nucleus, Min_n

Min_n = number of local minimum value from eight pixels in the neighbourhood of the nucleus region. (17)

i. Homogenity of nucleus, H_n

$$H_n = \sum_{i=1}^{m}\sum_{i=1}^{m}\frac{p(i,j)}{1+\lceil i-j \rceil}, \quad (18)$$

where $p(i,j)$ is the probability pixel pairs in the nucleus region calculated by gray level co-occurrence matrix and m is the number of grey level in image.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

j. Brightness of nucleus, B_n

B_n = the average pixel intensity of nucleus region. (19)

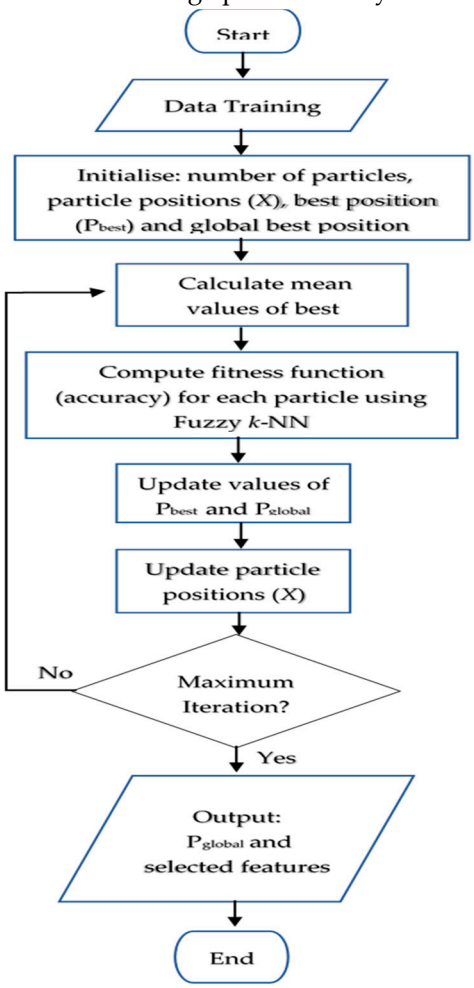


Figure 4. Flowchart depicting process of feature selection using the proposed Q-Fuzzy approach

k. Maxima of cytoplasm, Max_c

Max_c = number of local maximum value from eight (8) pixels in the neighbourhood of the cytoplasm region. (20)

l. Minima of cytoplasm, Min_c

Min_c = number of local minimum value from eight (8) pixels in the neighbourhood of the cytoplasm region. (21)

m. Brightness of cytoplasm, Bc

Bc = the average pixel intensity of cytoplasm region. (22)

n. Area of entire cell, A_{cell}

A_{cell} = number of pixels in the cell region. (23)

o. Compactness of the entire cell, C_{cell}

$$C_{\text{cell}} = \frac{\text{P}^2_{\text{cell}}}{\text{A}_{\text{cell}}}. \quad (24)$$

p. Ratio of nucleus and cell, R_{cell}

$$R_{cell} = \frac{A_{nu}}{A_{cy}}, \quad (25)$$

where A_{cell} is area of the cell region

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

q. LBP_{cell} = the local binary pattern of cell region. (26)

To extract textural features from the CS images, we use Local Binary Pattern Histogram Fourier (LBP-HF) protocol; wherein we start by applying operator LBP to obtain the pattern of data and create its histogram using Uniform LBP approach. Second, the LBP-HF feature is computed by applying Discrete Fourier Transform (DFT) from *n* histogram Uniform LBPs. Finally, the feature vector of LBP-HF is obtained by combining histogram values of all zeros, all ones, non-uniform, and Fourier spectrum values.

Based on the LBP-HF and relevant texture features in Eqs. (10) to (26), we obtain a texture feature vector of 38 entries, while the total feature vector has 54 entries. As mentioned earlier, further details about these features can be found in [3, 10, 13].

As outlined (in the red short-dashed rectangle) in Figure 3, the feature selection stage of our proposed QHT technique comprises of two units that utilise the quantum-behaved PSO technique and the Fuzzy *k*-NN algorithm so that both feature selection and classification accuracy in smeared cervical images are enhanced.

We further clarify that (in Figure 4) the fuzzy *k*-NN is applied prior to computing the fitness function, i.e. using the F*1* score (defined later in Eq. (29)). In other words, the QPSO approach is used to find the variation of subset features by generating particles that are each evaluated by calculating fitness values from the accuracy of classification result. Following this, the Fuzzy *k*-NN method is applied to classify the categories of smear images. Based on their fitness values, the particles with best fitness values are assigned as local best position and global best position. The particle that is assigned as global best position represents the best subset features.

In this study, we shall use Fuzzy *k*-NN is used to enhance the accuracy of our classification of cells in smeared cervical images and the best position P_{i_best} of particle i is determined using Eq. (18) [18].

4. Experiments on smeared cervical images using Q-Fuzzy approach

Using the framework outlined in Figures 3 and 4, which were discussed in latter parts of the preceding section, in this section, we present an experimental validation regarding the utility of our proposed QH approach for feature selection in cervical cell classification.

4.1. Description of dataset

For our experimental validation, the unit cell microscopic cervical smear (CS) images from the Herlev dataset [24] will be used as input dataset. This dataset consists of 917 CS images that were collated by cytology experts using a microscope connected to a digital camera. Each image was taken with a resolution of 0.201μm/pixel [10-12].

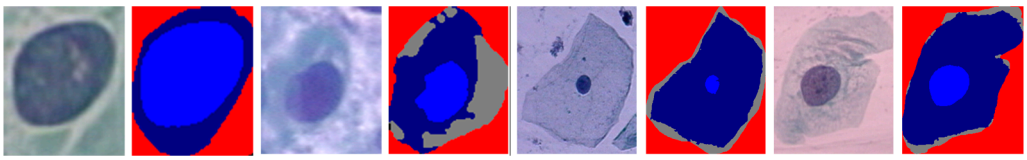


Figure 5. Sample cervical smear images and manual segmentation (ground truth) images [15-17]

After capturing the images, cytology experts were tasked with manually classifying the CS images into the seven classes discussed in Section 2. Each cell is assessed by two experts while a medical doctor was asked to further examine the samples that are deemed hard to deal with. All the images in the dataset are segmented into the cytoplasm, nucleus, and background regions using CHAMP software [12]. Results of the segmentation were further examined by cytology experts to ensure accuracy. A few sample images from the Herlev dataset (with their respective ground truth images) are shown in Figure 5. Furthermore, Table 3 presents the description of the Herlev smeared cervical dataset as well as their distribution and categories.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

Table 3. Classification of Pap smear single image cell types in the Herlev dataset [24]

Cell	Class Name	Cell count	Sub-total
Normal	Normal superficial squamous	74	242
	Normal intermediate squamous	70	
	Normal columnar	98	
Abnormal	Carcinoma in situ	150	675
	Light dysplastic	182	
	Moderate dysplastic	146	
	Severe dysplastic	197	
	Total	**917**	**917**

4.2. Evaluation method

Many approaches are utilised to evaluate the performance of classification algorithms. Here, we use the Precision, Recall, and F*1* Score analysis (or PRS analysis). Precision is used to ascertain the accuracy of classification using the number of correctly classified positive examples divided by the number of examples labelled by the system as positive (Eq. (27)). Recall is the number of correctly classified positive examples divided by the number of positive examples in the data (Eq. (28)) and F*1* score is combination of Precision and Recall [29] as defined in Eq. (29).

$$Precision\ (P) = \frac{tp}{tp+fp} \tag{27}$$

$$Recall\ (R) = \frac{tp}{tp+fn} \tag{28}$$

$$F1\ score = 2.\frac{P \times R}{P+R} \tag{29}$$

where *tp* is true positive, *fp* is false positive, and *fn* is false negative.

The *r*-fold cross validation method is employed for validating the experimental results, which considering the size of our data, 5-fold cross-validation is used.

To establish the relationship between the different feature classification approaches, we employ Cohen's kappa statistical measure, which is widely used to quantify agreement between two raters (i.e. mechanism to assess or observe a variable or system), that each classify N items into C mutually exclusive categories [30] as defined in Eq. (30).

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{30}$$

where $\boldsymbol{p_o}$ is the relative observed agreement between raters (akin to accuracy) and $\boldsymbol{p_e}$ is the hypothetical probability of chance agreement. If raters are in complete agreement $\boldsymbol{\kappa = 1}$, whereas when there is no agreement among the raters other than what would be expected by chance (i.e. as given by $\boldsymbol{p_e}$), then $\boldsymbol{\kappa \leq 0}$.

4.3. Experimental Result for Cervical Smear Image Classification

Earlier in Section 3, the mechanism via which the best out of an assemblage of seventeen (17) features with size of feature vector of 54 entries for the CS images were pruned down to seven (7) features with a feature vector of 32 entries.

The seven (7) best features utilised for the remainder of our performance assessment are area of nucleus (i.e. Eq. (10), roundness of nucleus i.e. Eq. (15)), brightness of nucleus (i.e. Eq. (19)), brightness of cytoplasm (i.e. Eq. (22)), area of entire cell (i.e. Eq. (23)), ratio of nucleus and cytoplasm (i.e. Eq. (25)), and local binary pattern with histogram Fourier (LBP-HF) (i.e. Eq. (26)). Furthermore, for the LBP-HF, 25 of 38 feature vectors were selected. Meanwhile, in the Fuzzy *k*-NN unit of the proposed Q-Fuzzy approach (Figure 3), we computed the accuracy for varying values of *k* (i.e. the number of nearest neighbours as outlined earlier in Section 2 and discussed in standard literature including [22,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

23]) starting with $k = 2$, which implies that at least 2 nearest neighbours are considered in determining the prediction result. In addition, it was observed that beyond $k = 4$ the accuracy of prediction results decreased, and so the upper bound was limited at $k = 7$.

We clarify here that our experiments are mainly focused on establishing the impact of the feature selection on the classification accuracy and how well our proposed QH approach (which combines the QPSO algorithm with the Fuzzy *k*-NN - or simply the Q-Fuzzy approach) contributes towards improving the performance of classification results, which is deemed crucial for cervical cancer detection.

Consequently, we considered three experimental scenarios whereby, in the first two experiments we compared the classification results obtainable with and without prior feature selection. This encompasses comparisons between our proposed Q-Fuzzy approach and other Fuzzy *k*-NN hybrid methods including one that blends the standard PSO with the Fuzzy *k*-NN technique (which we refer to as the P-Fuzzy approach) alongside the All-Features approach, i.e. classification without prior feature selection imposing the use of all the image features. Based on this setup, in the first experimental scenario, we present an assessment of the classification results for the All-features, P-Fuzzy and Q-Fuzzy approaches.

For a more objective assessment, we maintained the same parameter values of PSO and QPSO as well as the number of particles and number of iterations (maintained as 20 and 200, respectively) while varying the number of nearest neighbours, *k*. The outcomes (in Table 4) show that (for all values of *k*) applying prior feature selection (i.e. in both the P-Fuzzy or Q-Fuzzy approaches) improved the classification results in comparison to instances using all features (i.e. All-Features approach). The highest performance obtained without feature selection (i.e. using all features) is realised at $k = 5$, whereas the best performance for the P-Fuzzy and Q-Fuzzy approaches was obtained at $k = 4$.

Table 4. Comparison of classification results for approaches with (i.e. P-Fuzzy and Q-Fuzzy) and those without prior feature selection (i.e. All-features)

k	All-Features				P-Fuzzy				Q-Fuzzy			
	Precision	**Recall**	**F1 score**	**κ**	**Precision**	**Recall**	**F1 score**	**κ**	**Precision**	**Recall**	**F1 score**	**κ**
2	0.67	0.74	0.70	0.66	0.73	0.80	0.76	0.72	0.73	0.79	0.76	0.71
3	0.68	0.69	0.68	0.64	0.76	0.77	0.76	0.72	0.77	0.77	0.77	0.73
4	0.74	0.74	0.74	0.70	**0.83**	**0.84**	**0.83**	**0.81**	**0.85**	**0.86**	**0.85**	**0.83**
5	**0.76**	**0.76**	**0.76**	**0.72**	0.80	0.81	0.80	0.76	0.80	0.81	0.80	0.76
6	0.73	0.73	0.73	0.68	0.74	0.76	0.75	0.70	0.74	0.75	0.74	0.69
7	0.69	0.70	0.69	0.64	0.71	0.72	0.71	0.67	0.73	0.74	0.73	0.69

Furthermore, using the same approaches (i.e. All-features, P-Fuzzy and Q-Fuzzy), in the second experimental scenario, we assessed the impact of prior feature selection on the classification accuracy for all the seven cervical cell categories. Our results, in Table 5, indicate that our proposed QH feature selection approach that blends the QPSO algorithm with the Fuzzy *k*-NN (i.e. the Q-Fuzzy approach) yielded better classification accuracy than both the All-features and P-Fuzzy approaches in terms of outcomes of the PRS analysis. Further investigation of the outcomes indicates an improvement in the classification accuracy for all cell categories when prior feature selection (P-Fuzzy or Q-Fuzzy) was utilised. Moreover, our proposed Q-Fuzzy approach outperformed the P-Fuzzy hybrid approach for all cervical cell categories (except the Normal Columnar category) with its best classification results in the Normal superficial cell category.

To further establish the utility of the proposed Q-Fuzzy QH approach, in the third experimental scenario, we compared it (i.e. the proposed Q-Fuzzy approach) with other hybrid QH approaches formed by blending the QPSO algorithm with other supervised learning methods, namely, the Naïve Bayes (NB) and the support vector machines (SVM). Additionally, we examined the performance of the proposed approach alongside other feature selection methods, such as the filtering approach i.e.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

Relief [29] (by selecting the 25 top-ranked features from a feature vector of 54 entries (i.e. based on the Relief score), the sequential feature selection (SFS) method, which has an in-built capability to add or remove features sequentially [31] and the random forest (RF) method, which is an ensemble learning based technique that is widely employed in classification, regression, etc. [35].

Table 5. Comparison of classification results with and without prior feature selection using the all-features, P-Fuzzy and proposed Q-Fuzzy approaches

Cell category	All features (k = 5)			P-Fuzzy (k = 4)			Q-Fuzzy (k = 4)		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 Score
Normal superficial	0.83	0.86	0.84	0.95	0.91	0.93	0.95	0.95	0.95
Normal intermediate	0.82	0.74	0.78	0.89	0.84	0.86	0.89	0.89	0.89
Normal columnar	0.63	0.67	0.65	0.65	0.72	0.68	0.61	0.74	0.67
Carcinoma in situ	0.74	0.83	0.78	0.81	0.87	0.84	0.84	0.90	0.87
Light dysplastic	0.83	0.78	0.81	0.88	0.94	0.91	0.89	0.97	0.93
Moderate dysplastic	0.69	0.85	0.76	0.83	0.96	0.89	0.89	0.96	0.93
Severe dysplastic	0.79	0.62	0.70	0.86	0.65	0.74	0.88	0.61	0.72

Table 6. Comparison of classification outcomes between hybrid QH approaches that combine QPSO with Naïve Bayes, Support Vector Machines (SVM), Random Forest (RF) and Fuzzy k-NN

Classifier Methods		All Features	Feature Selection Methods			
			Non-quantum HIS			QHT
			Relief	SFS	PSO	QPSO
Naïve Bayes (NB)	Precision	0.72	0.75	0.76	0.73	0.76
	Recall	0.72	0.76	0.77	0.74	0.77
	F1 score	0.72	0.75	0.76	0.73	0.76
	κ	0.67	0.71	0.72	0.68	0.72
SVM	Precision	0.79	0.83	0.85	0.84	0.84
	Recall	0.79	0.83	0.86	0.84	0.85
	F1 score	0.79	0.82	0.84	0.84	0.84
	κ	0.74	0.80	0.82	0.81	0.81
Random Forest (RF)	Precision	0.79	0.83	0.85	0.84	0.85
	Recall	0.80	0.83	0.86	0.84	0.86
	F1 score	0.79	0.82	0.85	0.84	0.85
	κ	0.74	0.80	0.83	0.81	0.82
Fuzzy k-NN (k = 4 and k = 5)	Precision	0.76	0.74	0.80	0.83	**0.85**
	Recall	0.76	0.79	0.85	0.84	**0.86**
	F1 score	0.76	0.76	0.83	0.83	**0.85**
	κ	0.72	0.72	0.80	0.81	**0.83**

To clearly assess the potential impacts of integrating 'quantumness' into cell classification of cervical smear images for the purpose of cervical cancer detection, our discussion of the results in Table 6 will be predicated in two directions.

First, we analyse the results in terms of the resulting HIS models that are realised by blending the classifiers (in the extreme left of Table 6) with the feature selection methods (Naïve Bayes (NB) support vector machines (SVM) and random forest (RF)) that produce the NB, RF, SVM and Fuzzy k-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

NN (F*k*NN) HIS models. To simplify our discussion, we further divided the models into two categories. The first category consists of non-quantum HIS models wherein neither the classifier nor the feature selection units has any 'quantumness' in it. The second category comprises of quantum HIS techniques (or simply quantum hybrid techniques) that are realised by blending the QPSO feature selection method with each of the four classifiers (NB, SVM, RF and F*k*NN). The resulting NB, SVM, RF and F*k*NN QHTs are shown highlighted in Table 6. Notably, we clarify that the F*k*NN QHT (whose results are shaded in darker background) is also our proposed Q-Fuzzy approach that blends the QPSO into the Fuzzy *k*-NN classifier method.

Our analysis in terms of the HIS models pitches the non-quantum models against the QH techniques for each classifier method. As reported in Table 6, the QHT technique of each classification method (i.e. QPSO blended with NB, QPSO blended with SVM, QPSO blended with RF and QPSO blended with F*k*NN) present between marginal to modest improvements across all the parameters (i.e. Precision, Recall, F1 score and Cohen's Kappa measure) used to assess the classification accuracy than most of the non-quantum HIS models for that same classifier. Similarly, the results obtained when All-Features (i.e. without prior feature selection) methods is used are improved along all parameters in comparison with using the QHT of each corresponding classification method.

To focus the assessment in terms of our proposed Q-Fuzzy approach, the second perspective of our discussion of the results in Table 6 is confined to comparisons between the four QHTs. The results (shaded area) show that, with the exception of the RF-based QHT, our proposed Q-Fuzzy approach (or F*k*NN-based QHT) performs better than the NB and SVM-based QHTs in terms of all four parameters that are reported (in Table 6). Even though these improvements are modest (ranging between 1 to 9 percent for Precision, Recall and F1 score and 2 to 11 percent for the Cohen's Kappa statistical measure) their impact in terms of better detection of cancerous cells in cervical smear images are significant. Moreover, as reported earlier in Table 5, our proposed Q-Fuzzy approach outperformed its closest competition (the P-Fuzzy technique or PSO-based non-quantum HIS) in terms of accuracy of classifying most of the cell categories in the cervical smear images. Similarly, as reported in Table 4, the proposed Q-Fuzzy approach presented better classification accuracy in terms of abnormal cell categories than the P-Fuzzy method.

In concluding, we reiterate that all along our objective has been to explore the impact of integrating 'quantumness' into HIS models. As reported earlier in this section, the quantum hybridised QHT models enhanced the classification accuracy across most of classifier methods (i.e. NB, SVM and Fuzzy *k*-NN) that were evaluated. Furthermore, as noted in earlier Section 2, the improvements recorded via the proposed Q-Fuzzy approach could be attributed to the manner that the intuitionistic rationality of the Fuzzy *k*-NN complements the adaptive search capability of the quantum-behaved QPSO algorithm. Consequently, the foregoing narrative enunciates potential applications for quantum machine learning algorithms in image-based approaches to disease diagnosis and treatment.

5. Concluding remarks

Accuracy is an important aspect of cell detection and delineation, especially in many image-based applications for disease diagnosis. To enhance the accuracy of classification of cells in smeared cervical images, our study proposes a hybrid feature selection technique that combines the potency of the quantum-behaved particle swarm optimisation (QPSO) algorithm with the versatility of the Fuzzy *k*-nearest neighbours (Fuzzy *k*-NN) algorithm (i.e. the proposed quantum hybrid (QH) or Q-Fuzzy technique).

From an initial multitude of seventeen (17) features that describe the geometry, colour, and texture of the cervical images, the QPSO stage of our proposed technique is used to select the best subset features (i.e. global best particles) that represent a pruned down collection of seven (7) features.

Using the Herlev dataset [24]) consisting of almost 1000 images, our verification of the cogency of the proposed QH approach in cervical cancer detection was predicated upon its use in three experimental scenarios. The purpose of the first two scenarios is to establish the impact of prior feature selection on classification accuracy. In these experiments, we compared the All-Features

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

approach (i.e. when no prior feature selection is used in the classification) alongside HIS models including the P-Fuzzy approach (which is a hybrid approach that combines the standard PSO algorithm with the Fuzzy *k*-NN technique) on one side and alongside our proposed Q-Fuzzy approach that blends the quantum-behaved PSO algorithm with the Fuzzy *k*-NN technique. Outcomes from these tests confirm that integrating 'quantumness' into the feature selection units led to enhanced classification accuracy. Moreover, the efficacy of our proposed Q-Fuzzy approach manifests in the manner that it rivals the P-Fuzzy approach in terms of classification accuracy for six out of the seven cervical cell categories. In the third experimental scenario, we established the synergy between the quantum-behaved PSO and the Fuzzy *k*-NN technique that form our proposed Q-Fuzzy approach by comparing its classification accuracy alongside that from a trio of hybrid QH approaches that each blends the QPSO with other classification methods, namely the Naïve Bayes (NB) networks, the random forest (RF) ensemble method and the support vector machine (SVM). Outcomes establish that the adaptive search capability of the quantum-behaved QPSO algorithm is best complemented by the intuitionistic rationality inherent to the Fuzzy *k*-NN technique. Overall, our proposed QPSO blended Fuzzy *k*-NN QH approach (i.e. Q-Fuzzy approach) outperformed the other approaches that QPSO blended with the NB and SVM with modest average increases ranging from 2 to 11 percent in terms of the classification parameters that were evaluated.

Buoyed by the foregoing outcomes, in ongoing work, additional 'quantumness' will be exploited and integrated into the QPSO (as well as in QH techniques generally) so that the accuracy and performance of the feature selection could be further enhanced. These efforts would improve the effectiveness of image-based approaches used in cervical cancer diagnosis and treatment. Later, we hope to exploit these strategies (and others in [33]) in image-based detection of other carcinogens.

Acknowledgments: This work is sponsored, in full, by the Prince Sattam Bin Abdulaziz University, Saudi Arabia via the Deanship for Scientific Research funding granted to the CIIS Research group project number 2016/01/6441.

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

References

1. A. Abraham, "Hybrid artificial intelligent systems", in Innovations in Hybrid Intelligent Systems Series on Advanced Soft Computing, 2007, 44(2007). xvi-xvi.
2. K. A. Abuhasel, C. Fatichah, A. M. Iliyasu, "A Bi-Stage Technique for segmenting Cervical Smear Images Using Possibilistic Fuzzy C-Means and Mathematical Morphology", J. of Medical Imaging and Health Info., 2016, 6(7): 1663–1669.
3. WHO, "Comprehensive Cervical Cancer Control: A Guide to Essential Practice," WHO Press, 2006
4. A. Gençtava, S. Aksoy, and S. Onder, Unsupervised segmentation and classification of cervical cell images, Pattern Recognition, 2012, 45(12): 4151–4168.
5. Parkway Cancer Centre, 2014. [Online]. Available: http://www.parkwaycancercentre.com/. [Accessed October 2017].
6. C. M. Wang, H.T. Chen, S. F. Yang-Mao, Y. K. Can, S. F. Lin, New Methods for Image De-noising and Edge Enhancement in Cervical Smear Images Segmentation, Int. J. of Computer, Consumer and Control (IJ3C), 2013, 2(1): 2013.
7. R.M. Demay, Common problems in Papanicolaou smear interpretation, Archives of Pathology and Laboratory Medicine, 1997, 121(3): 229–238.
8. B. L. Craine, E. R. Craine, J. R. Engel, N. T. Wemple, A clinical system for digital imaging Colposcopy, Medical Imaging II, Calif, USA, Newport Beach press 1998, pp. 505–511.
9. E. J. Mariarputham and A. Stephen, Nominated Texture Based Cervical Cancer Classification, Computational and Mathematical Methods in Medicine, 2015, 2015(Article ID 586928): 2015.
10. M. H. Tsai, Y. K. Chan, Z. Z. Lin, S. F. Yang-Mao, P. C. Huan, Nucleus and cytoplast contour detector of cervical smear image, Pattern Recognition Letter, 2008, 29(9): 1441-1453.
11. E. Martin, Pap-smear classification, Master's Thesis, Technical University of Denmark: Oersted-DTU, Automation, 2003.
12. J. Norup, Classification of pap-smear data by transductive neuro-fuzzy methods, Master's thesis, Technical University of Denmark: Oersted-DTU, Automation, 2005.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2017 doi:10.20944/preprints201711.0193.v1

13. J. Jantzen, J. Norup, G. Dounias and B. Bjerregaard, Pap-smear Benchmark Data for Pattern Classification, 2005.
14. T. Chankong, N. Theera-Umpan and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in pap smear," *Computer Methods and Program in Biomedicine*, 2014, 2(113): 539-556.
15. A. H. Mbaga, P. ZhiJun, "Pap Smear Images Classification for Early Detection of Cervical Cancer," *Int. J. of Comp. Appl*., 2015, 118(7): 10-16.
16. B. Ashok, P. Aruna, Comparison of Feature selection methods for diagnosis of cervical cancer using SVM classifier," *Int. J. of Eng. Research & Appl*., 2016, 6(1): 94-99.
17. K. A. Abuhasel, A. M. Iliyasu, C. Fatichah, "A Hybrid Particle Swarm Optimisation and Neural Network with Fuzzy Membership Function Technique for Epileptic Seizure Classification," *J. of Adv. Comput. Intelligence and Intelligent Informatics*, 2015, 19(3): 447-455.
18. J. Sun, B. Feng, and W. B. Xu, Particle swarm optimization with particles having quantum behaviour, Proceedings of the IEEE Congress on Evolutionary Computation (CEC '04), 2004 June 2-4; pp. 325–331.
19. X. Fu, W.S. Liu, B. Zhang, H. Deng, "Quantum behaved particle swarm optimization with neighbourhood search for numerical optimization," *Math. Problems Eng*. 201(Article ID 469723): 2013.
20. C. Jin, S.W. Jin, Automatic image annotation using feature selection based on improving quantum particle swarm optimization, *Signal Processing*, 2015, 109: 172–181.
21. J. Kennedy, R.C. Eberhart, Y. Shi, Swarm Intelligence, Morgan Kaufmann, San Francisco, 2001.
22. D. Li, J.S. Deogun, and K. Wang, Gene function classification using fuzzy k-nearest neighbour approach, Proceedings of IEEE Int. Conference on Granular Computing (GRC), pp. 644-644, 2007.
23. S. Y. Kim, J. Sim, and J. Lee, "Fuzzy k-Nearest Neighbour Method for Protein Secondary Structure Prediction and Its Parallel Implementation," Computational Intelligence and Bioinformatics, 2006, 4115, Lecture Notes in Computer Science: 444-453.
24. The Management and Decision Engineering Laboratory, Herlev Databases http://mde-lab.aegean.gr/downloads
25. A. M. Iliyasu, Towards the Realisation of Secure and Efficient Image and Video Processing Applications on Quantum Computers, *Entropy*, 2013, 15(8): 2874-2974
26. A. M. Iliyasu, P. Q. Le, F. Dong, K. Hirota, Restricted geometric transformations and their applications for quantum image watermarking and authentication, In Proc. of the 10th Asian Conf. on Quantum Info. Science (AQIS 2010), 18-19 August 2010; Tokyo, Japan, pp. 212–214.
27. A. M. Iliyasu, F. Yan, S. E. Venegas-Andraca, A. S. Salama, "Hybrid Quantum-Classical Protocol for Storage and Retrieval of Discrete-Valued Information," *Entropy*, 2014, 16(6): 3537-3551.
28. F. Yan, A. M. Iliyasu, A survey of quantum image representation, Quant. Info. Proc. 2016, 15 (1): 1–35.
29. M. Sokolova, G. Lapalme, "A systematic analysis of performance measures," *Info. Proc. & Managemen*t, 2009, 45: 427–437.
30. Cohen, Jacob, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, 1960, 20 (1): 37–46. doi:10.1177/001316446002000104
31. R. Thomas, C. Osendorfer, P. Van der Smagt, "Sequential Feature Selection for Classification," *Advances in Artificial Intelligence*, 2011, 132-141, doi:10.1007/978-3-642-25832-9_14
32. A. Sakthi, M. Rajaram, "Density based Multiclass Support Vector Machine using IoT driven service oriented architecture for predicting cervical cancer", *Int. J. of u- and e-Service, Science and Technology*, 2016, 9(11):195-216
33. A. M. Iliyasu, C. Fatichah, K. A. Abuhasel, "Evidence Accumulation Clustering with Possibilitic Fuzzy C-Means base clustering approach to disease diagnosis," *Automatika: Journal for Control, Measurement, Electronics, Computing and Communications*, 2016, 57(3): 822-835
34. M. Schuld, I. Sinayskiy, F. Petruccione, "An introduction to quantum machine learning," *Contemporary Physics*, 2014, 56 (2): *172.* arXiv:1409.3097
35. A. M. Iliyasu, C. Fatichah, A. S. Salama, J. She "Quantum hybrid technique for feature selection in classification of cervical cells," 7th Int. Sym. on Comp. Intelligence & Industrial Appl. (ISCIIA 2016), 3-6 Nov., 2016; Beijing-China: 822-835