Journal of Computers Vol. 30 No. 2, 2019, pp. 103-115
doi:10.3966/199115992019043002009

Feature Selection Using Fuzzy-based Firefly Algorithm for Glistenings Detection on Intraocular Lenses

Parisut Jitpakdee[1*], Bunyarit Uyyanonvara[2], Chris Hull[3]

[1] Sirindhorn International Institute of Technology, Thammasat University, Thailand
parisut@gmail.com

[2] Sirindhorn International Institute of Technology, Thammasat University, Thailand
bunyarit@siit.tu.ac.th

[3] School of Health Sciences, City University London, London, United Kingdom
c.c.hull@city.ac.uk

Received 15 December 2018; Revised 15 February 2019; Accepted 15 February 2019

Abstract. Glistenings are liquid-filled microvacuoles in intraocular lenses (IOLs) appear when the IOL is in an aquatic environment that affect the quality of vision. In our glistenings Detection method, the candidate glistenings are automatically detected by mathematic morphology methodology. Machine learning approaches, feature selection and classification are used in this paper. The 68 features are extracted and used as training data for fine segmented using the classifiers. The detected glistenings are validated by object-based with ophthalmologist's hand-drawn ground-truth. Our proposed method, Feature Selection using Fuzzy-based Firefly Algorithm (FS-FFA) applied the concept of fuzzy entropy to calculating the membership of features data for in order to select good sets of the relevant features that maximize the classification performance in glistenings Detection. The proposed FS-FFA is compared with feature selection methods the standard firefly algorithm (FS-FA) and without feature selection using basic classifier k-nearest neighbor. The results have shown that the Matthews correlation coefficient (MCC) and the diagnostic odds ratio (DOR) value increase after feature selection using firefly algorithm and fuzzy entropy. Small size of features set also decreased that classification time in testing phase.

Keywords: classification, edical image processing, feature extraction, feature selection, firefly algorithm, fuzzy entropy, machine learning, object detection, optimization algorithm

1 Introduction

In irregular conditions of visibility, for example, cataract, the regular lens of a human eye is being supplanted by a counterfeit Intraocular Lens (IOL) embedded in the capsular bag [1]. Glistenings is an appearance of fluid-filled microvacuoles in embedded lenses and decrease the clearness of Lens affect optical capacity [2]. Many researches, [3-5] and [6] proposed the result that glistenings were found after implantation in spite of the fact that the times of onset were distinctive. The visual impacts of glistenings due to forward light disperse have been assessed in various studies [7-9]. Many researches consider demonstrate the glistenings arrangement in various ways [2-4, 6, 10-11]. The work of Werner [2] portrayed how to grade the glistenings by utilizing the density.

Computer vision and image processing methods are applied in medical application, for example, Cancer diagnosis [12], exudate and micro aneurysms detection in retinal images [13]. Blobs detection, the imperative issue in medical image processing utilizes systems like those executed for glistenings detection. Glistenings detection with simple image processing methodology were presented in our

* Corresponding Author

previous work [14], we used 20 features and basic classifier, k-Nearest Classifier and Naive Bayes to increase the accuracy of glistenings detection. In this study we want to improve the performance of glistenings detection by extract more features and used feature selection to find the suitable features.

The principle thought of feature selection is to pick a subset of information factors by wiping out features with practically no prescient data. Utilizing feature subset selection procedures, it could be overlooked repetitive and insignificant features to lessen the run time in classification [15]. Feature selection for classification was reviewed in [16] and many feature extraction methods for Image Registration are introduced in [17-18] introduced a technique for classification problems by using fuzzy entropy measure for select subset of feature. While choosing features, it considers limit samples and selects relevant features to get higher average classification accuracy.

Many optimization algorithms are used for selected subset of features. [19] proposed feature selection algorithm that combines Firefly algorithm with Rough Set Theory. The experiment with four different medical datasets obtained from UCI showed that their approach better that other methods in terms of time and optimality. [20] introduced a Feature Selection method using Forest Optimization Algorithm for search the best feature to improving the accuracy of classification. Genetic algorithm is applied for selecting good subsets of features in vehicle detection and face detection application [21]. They considered PCA for feature extraction and support vector machines (SVMs) for classification. In [22] presented a filter feature selection algorithm based on PSO and the mutual information for each pair of features, which is utilized to assess the pertinence and excess in the chose feature subset. In [23] proposed an Opposition based firefly algorithm (OBFA) to extract the relevant features that maximize the classification performance of SVM for prediction fatal risk. Cuckoo search algorithm is applied for selection features in [24]. They tested with four public datasets and compared with Bat Algorithm, Firefly Algorithm and Particle Swarm Optimization.

Feature selection based on firefly algorithm (FFA) optimization is proposed in [25]. They used fitness function calculated from classification accuracy and size of selected features, so can obtain the minimum size and maximum accuracy. [26] introduced an optimized feature selection algorithm, utilizing fuzzy entropy and firefly concepts. They selected initial subset of features using fuzzy entropy threshold and then used firefly algorithm selected the features from this initial subset. In our proposed algorithm we used fuzzy entropy not just pre-selection, but we use fuzzy entropy in the learning process of firefly algorithm, too.

In this study, we proposed a new approach of classification-based feature selection using firefly algorithm and fuzzy entropy, in order to increase the accuracy and performance of glistenings detection. We compared the performance of our method with basic k-Nearest Classifier using all features and compared the result with feature selection method using normal firefly algorithm by evaluating average accuracy, precision and F-measure, against ophthalmologist's hand-drawn ground-truth. The paper is organized as follows. Section 2 presented our methodology candidate glistenings detection, feature extraction and feature selection. The experimental results are showed in Section 3. Finally, Section 4 is a conclusion of this study.

2 Methodology

2.1 Overall Process

The overall process is shown in Fig. 1. To start with the Color IOL Image is changed over to the Grayscale picture by utilizing a Red band. Then candidate glistenings are detected with blob detection method described in Section 2.2. Each blob is labeled in to type "glistening" and "non-glistening" by comparing with the ground-truth image. Features data are extracted from all "glistening" and "non-glistening" objects.

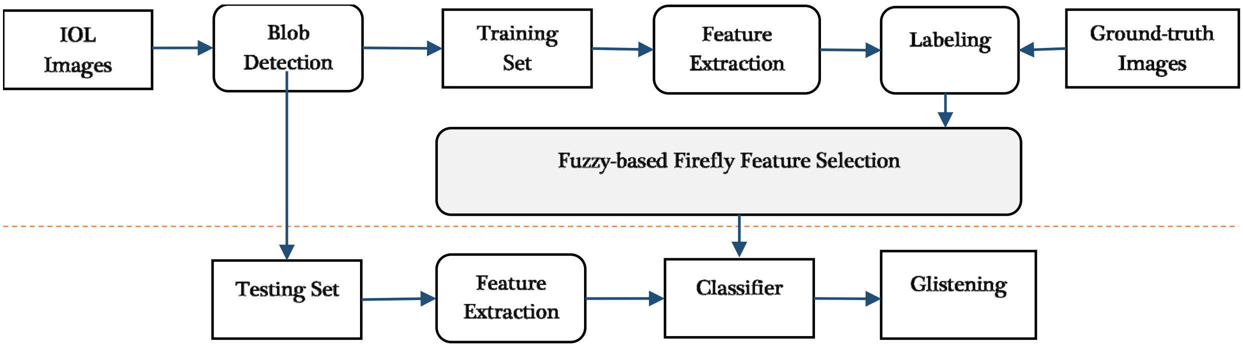
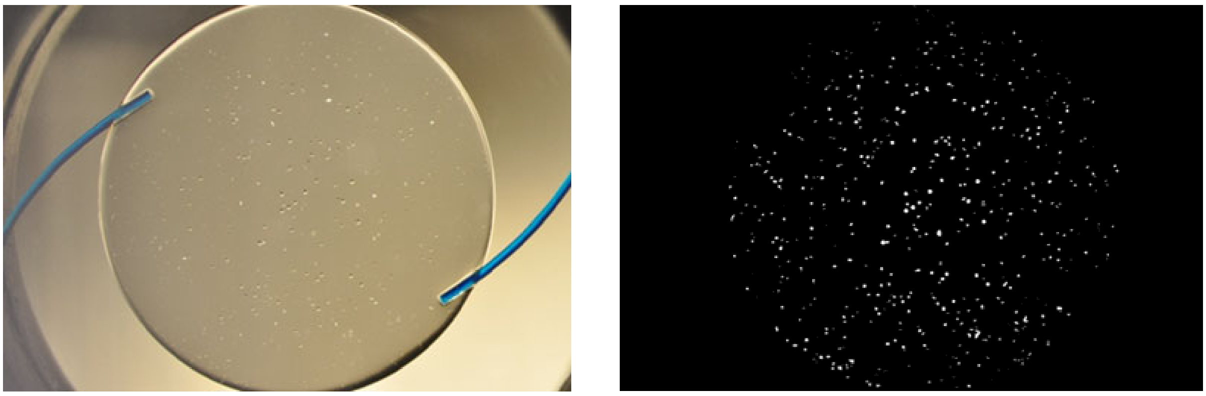


Fig. 1. Overall process of glistening detection using fuzzy-based firefly for feature selection

2.2 Glistenings Detection

In collaboration with the Applied Vision Research Centre, School of Health Sciences, City University London, we obtained 27 digital Intraocular Lens Glistenings images captured with a digital Nikon FS-2 photo slit-lamp. The images taken for in vitro experiment downsized from 3216 x 2136 to 750 x 499 with the ground-truth images drew by the ophthalmologists. Examples of IOL images and ground-truth image are shown in Fig. 2.

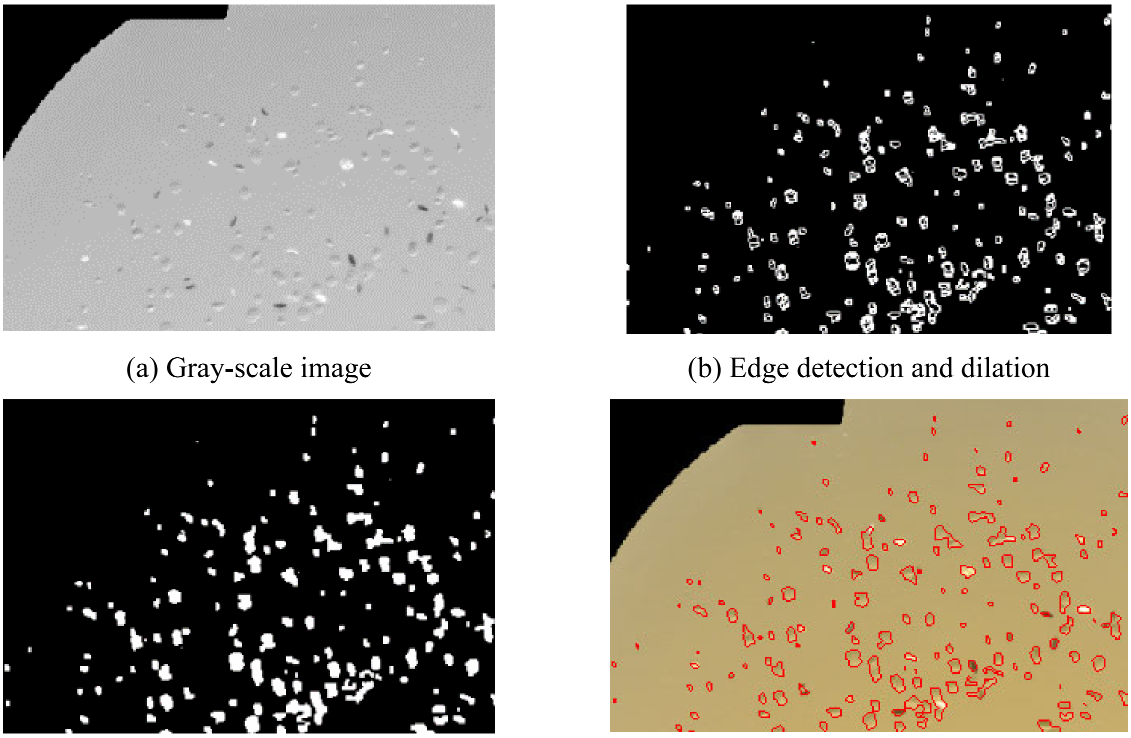


(a) IOL Image (b) ground-truth image

Fig. 2. Sample of (a) IOL Image and (b) ground-truth image

A set of mathematical morphology method are used to detection the glistening candidates. The blobs' edges are detected by Sobel operator [27] both horizontal and vertical direction and used the sensitivity threshold 0.02 to ignore all weak edges. Then the sub-sequence lines are grouped by a dilation operation. Dilation will fill every single close locale with white shading. The example images of close-up blobs detection method are shown in Fig. 3.

From 27 IOL Images, 9,858 blobs are detected. All blobs are labelled by comparing with the ground-truth images, as "glistenings" and "non-glistenings". Next step, all blobs are extracted the information in the feature extraction process described in section 2.3.



(a) Gray-scale image (b) Edge detection and dilation

(c) Hole filling (d) Candidate glistenings edges

Fig. 3. Blobs detection

2.3 Feature Extraction

In this stage, we proposed 68 features to recognize glistening and non-glistening objects. All features are extracted from each blob images and can divided to 6 main types: statistical features, geometric features, filter responses features, differential features, spatial-domain features and colour features [17]. The detail of feature in each group described in Table 1.

Table 1. Detail of feature extraction for glistening detection

No.	Feature Name	Detail
Type 1: Statistical Features		
1	Most dominant intensity	The intensity that have maximum member in histogram in gray-scale image
2	Mean Intensity	Average of intensity gray-scale image
3	Maximum Intensity	Maximum of intensity in gray-scale image
4	Minimum Intensity	Minimum of intensity in gray-scale image
5	Standard deviation intensity	Standard deviation of intensity in gray-scale image
6	Skewness	Measure of the asymmetry of the data around the sample mean.
7	Kurtosis	Measure of how outlier-prone a distribution is.
8-11	Energy	$\Sigma_{i_1}\Sigma_{i2}[p(i_1,i_2 \mid \theta)]^2$, θ = 0°, 45°, 90°, 135° **(4)**
12-15	Contrast	$\Sigma_{i_1}\Sigma_{i2}(i_1-i_2)^2 p(i_1,i_2 \mid \theta)$, θ = 0°, 45°, 90°, 135° **(5)**
16-19	Correlation	Measures a varying pattern $\Sigma_{i_1}\Sigma_{i_2}\frac{(i_1-\mu_{i_1})(i_2-\mu_{i_2})}{\sigma_{i_1}\sigma_{i2}}p(i_1,i_2 \mid \theta)$ **(6)** where μ_{i_1} and σ_{i_1} are the mean and the standard deviation of $\Sigma_{i_2}h(i_1,i_2 \mid \theta)$ and μ_{i_2} and σ_{i_2} are the mean and the standard deviation of $\Sigma_{i_1}h(i_1,i_2 \mid \theta)$, θ = 0°, 45°, 90°, 135°
20-23	Entropy	shows the probability of two intensities appearing together $\Sigma_{i_1}\Sigma_{i2}p(i_1,i_2 \mid \theta)\log p(i_1,i_2 \mid \theta)$, θ = 0°, 45°, 90°, 135° **(7)**

Table 1. Detail of feature extraction for glistening detection (continue)

No.	Feature Name	Detail	
24-27	Homogeneity	measures the smoothly of image $\Sigma_{i_1}\Sigma_{i_2}\frac{p(i_1,i_2\mid\theta)}{1+(i_1-i_2)^2}$, θ = 0°, 45°, 90°, 135°	**(8)**
28-31	Short-runs emphasis	measures the noisy of image $\Sigma_i\Sigma_l\frac{p(i,l\mid\theta)}{l^2}$, θ = 0°, 45°, 90°, 135°	**(9)**
32-35	Long-runs emphasis	$\Sigma_i\Sigma_l l^2 p(i,l\mid\theta)$, θ = 0°, 45°, 90°, 135°	**(10)**
36-39	Gray-level non-uniformity	measures the varying of intensities $\frac{\Sigma_i(\Sigma_l h(i,l\mid\theta))^2}{\Sigma_i\Sigma_l h(i,l\mid\theta)}$, θ = 0°, 45°, 90°, 135°	(11)
40-43	Long-runs non-uniformity	measures the varying of patterns $\frac{\Sigma_l(\Sigma_i h(i,l\mid\theta))^2}{\Sigma_i\Sigma_l h(i,l\mid\theta)}$, θ = 0°, 45°, 90°, 135°	**(12)**
44-47	Run percentage	measures the smoothly varying of image $\frac{100}{MN}\Sigma_i\Sigma_l h(i,l\mid\theta)$, θ = 0°, 45°, 90°, 135°	**(13)**
Type 2: Geometric Features			
48	Normalized Area	the actual number of pixels in the region	
49	Eccentricity	the ratio of the distance between the foci of the ellipse and its major axis length	
50	EquivDiameter	the diameter of a circle with the same area as the region	
51	Perimeter	the distance around the boundary of the region	
52	Compactness	the variance of the radial distance of the object's pixels from the centroid divided by the area.	
53	Solidity	the proportion of the pixels in the convex hull that are also in the region. Computed as Area/Convex Area	
Type 3: Filter Responses			
54	Gradient Magnitude	Mean of intensity of red band image	
55-60	Gabor Filter	Mean of intensity of red band image, θ = 0, π/6, 2π/6, 3π/6, 4π/6, and 5π/6.	
Type 4: Differential Features			
61	Gaussian Filter	Mean of intensity of red band image smoothed by Gaussian filter, filter size = 3 with standard deviation sigma = 0.25	
62	Laplacian Filter	Mean of intensity of red band image smoothed by Laplacian filter, filter size = 3 with standard deviation sigma = 0.25	
63	LoG (Laplacian of Gaussian) filter	Mean of intensity of red band image smoothed by Laplacian of Gaussian filter, filter size = 3 with standard deviation sigma = 0.25	
Type 5: Spatial-Domain Features			
64	Deviation from mean	The average absolute difference between the mean intensity and other intensities in a red band image	
65	Absolute centre contrast	The average absolute difference between the centre intensity and other intensities in a red band image	
Type 6: Colour Features			
66	Average colour	Average colour at the pixel using 3x3 neighbourhood in red band image $\frac{1}{9}\sum_{k=-1}^{1}\sum_{l=-1}^{1}R(x+k,\ y+l)$	**(14)**
67	Density of dominant colour	The ratio of number of pixels with dominant colour to the total number of pixels in the red band images $\frac{Nc_i}{\Sigma_k Nc_k}$, where Nc_i is the number of pixels with the dominant colour.	
68	Standard deviation of Saturation	colourfulness of an area judged in proportion to its brightness $1-\frac{3\min(R,G,B)}{(R+G+B)}$	**(15)**

Feature No. 8-27 are gray level spatial-dependence features calculated from the joint conditional probability density (JCDP) to describe texture in an image. The joint conditional probability density $p(i_1, i_2 | \theta)$ is calculated by divided a co-occurrence matrix $h(i_1, i_2 | \theta)$ with the sum of its entries. Run-length features are created in feature No. 28 to No. 47. The run-length matrix $h(i, l | \theta)$ used to describe the intensities variations in the image in particular direction. A joint conditional probability density $p(i, l | \theta)$ is calculated by divided a run-length matrix with the sum of its entries. Feature No. 55-60 are the features from Gabor filter. A two-dimensional Gabor filter is a Gaussian kernel function modulated by a complex sinusoidal plane wave, defined as follows [27-28]:

$$G(x, y) = \frac{f^2}{\pi\gamma\mu}\exp(-\frac{x'^2 + y^2 y'^2}{2\sigma^2})\exp(j2\pi fx' + \phi) \quad \textbf{(1)}$$

$$x' = x\cos\theta + y\sin\theta \quad \textbf{(2)}$$

$$y' = -x\sin\theta + y\cos\theta \quad \textbf{(3)}$$

Where f is the frequency of the sinusoidal factor, θ is the orientation of the normal to the parallel stripes of a Gabor function, ϕ is the phase offset, σ is the standard deviation and γ is the spatial aspect ratio. In this paper, we used θ = 0, π/6, 2π/6, 3π/6, 4π/6, and 5π/6.

2.4 Firefly Algorithm (FA)

A Firefly Algorithm (FA) first proposed by Yang [29], is an optimization algorithm that emulates characteristics of fireflies and their shine pattern. The Firefly Algorithm is a nature-inspired algorithm, population-based algorithm, based on swarm intelligence aims to find the global optima of objective functions. Each firefly is pulled in by the brighter gleam of other neighboring fireflies. At the point when more far between the couples of fireflies, the appeal is diminishing.

In Firefly algorithm, there are three idealized rules defined by Yang [29]: (1) All fireflies are unisex so that one firefly will be pulled in to different fireflies not mind to their sex; (2) Attractiveness is corresponding to their shine. Thus, for any two fireflies, the less splendid one will move towards the brighter one. On the off chance that there is no brighter one than a specific firefly, it will move arbitrarily; (3) the shine of a firefly is from the objective function. For a maximization problem, the brilliance can basically be relative to the estimation of the objective function [29-30]. The pseudo code of Firefly algorithm can be shown as Fig. 4.

Firefly Algorithm: FA

Objective function $f(x), x = (x_1, \dots, x_d)^T$
Define the parameters
Generate initial population of fireflies x_i $(i = 1,2, \dots, n)$
Calculate Light intensity I_i *at* x_i *is determined by* $f(x_i)$
Repeat:
  for $i = 1{:}n$ *all n fireflies*
    for $j = 1{:}n$ *all n fireflies*
      if $(I_j > I_i)$, *Move firefly i towards j in d-dimension Attractiveness varies with distance r via* $exp[-\gamma r]$
        Evaluate new solutions and update light intensity
    if no one of firefly brighter than I_i, I_i *move randomly*
  Rank the fireflies and find the current best
Until: Maximum iteration or minimum change of objective function

Fig. 4. The pseudo code of firefly algorithm

The attractiveness function β calculate from the distance $r_{i,j}$ of the firefly is determined by:

$$\beta(r_{i,j}) = \beta_0 e^{-\gamma r_{i,j} r^2}. \quad \textbf{(16)}$$

where β_0 is the attractiveness at $r=0$ and γ is the light absorbtion coefficient at the source. It should be noted that the $r_{i,j}$ which is the euclidence distance between any two fireflies i and j at x_i and x_j, where x_i and x_j are the spatial coordinate of the fireflies i and j, respectively.

The movement of a firefly i, which is attracted to another more attractive firefly j is determined by:

$$x_i = x_i + \beta_0 e^{-\gamma r_{i,j}^2}(x_i - x_j) + \alpha(rand - \frac{1}{2}) \quad \textbf{(17)}$$

where the second term is the attraction while the third term is randomization parameter with α being the randomization parameter. For most cases in implementation, Yang [30] introduced $\beta_0 = 1, \alpha \in [0,1]$ and $\gamma \in [0.01, 100]$.

2.5 Feature Selection Using Firefly Algorithm (FS-FA)

Feature selection using firefly algorithm, first step, each firefly are randomly generated as weight of feature,

$$x_i = [w_{i_1}, w_{i_2}, w_{i_3}, \ldots, w_{i_D}] \quad \textbf{(18)}$$

where w_{ij} is weight of feature j^{th}, and D is a number of all features. Feature j^{th} is selected if w_{ij} > thresholdτ.

The fitness function is the Matthews correlation coefficient (MCC) of classification using k-nearest neighbours (k-NN) algorithm [2]. The k -NN algorithm is one of familiar machine learning method. This classification algorithm finds closest neighbours of a test object [31]. In this study, we used k = 9 and measure the distance with the Euclidean distance [32]. The Matthews correlation coefficient (MCC) is a balanced measure, which can be utilized regardless of the possibility that the classes are of altogether different sizes [34]. MCC value is in a range between −1 and +1. A coefficient of +1 indicates a plenary prediction, 0 no better than random prediction and −1 indicates total depreciation between prediction and observation. MCC is calculated by:

$$MCC = \frac{(TPxTN) - (FPxFN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad \textbf{(19)}$$

True Positive (TP) is the number of true glistenings detected as glistenings, False Positive (FP) is the number of false glistenings detected as glistenings and False Negative (FN) is the number of true glistenings that were not detected. True Negative (FN) is the number of false glistenings that detected as false glistenings.

The pseudo code of feature selection using firefly algorithm (FS-FA) is shown in Fig. 5

Feature Selection using Firefly Algorithm (FS-FA)

```
Objective function f(x), x = (x_1, ..., x_d)^T,
Define the parameters
Generate initial population of fireflies x_i , x_i = [w_i1, w_i2, w_i3, ..., w_id] , i = 1,2, ..., n
Calculate Light intensity I_i at x_i is determined by f(x_i)
Repeat:
    for i = 1:n all n fireflies
        for j = 1:n all n fireflies
            if (I_j > I_i), Move firefly i towards j in d-dimension Attractiveness varies with
                    distance r via exp[-γr]
                Select subset of features where w_ij > τ
                Calculate accuracy with classifier
                update light intensity
        if no one of firefly brighter than I_i, I_i move randomly
    Rank the fireflies and find the current best
Until: Maximum iteration or minimum change of objective function
```

Fig. 5. The pseudo code of feature selection using firefly algorithm (FS-FA)

2.6 Feature Selection Using Fuzzy-based Firefly Algorithm (FS-FFA)

The proposed algorithm enhanced version of the FS-FA algorithm to find the small size subset of features that have high discriminant information and maximizes the fitness function by using fuzzy entropy. The entropy of fuzzy sets is a measure of fuzziness between fuzzy sets [33]. Fuzzy entropy (EC) [26] calculation from fuzzy membership in Fuzzy matrix (U).

Fuzzy matrix $U = [\mu_{ij}]$ with c rows and f columns in which n is the number of class and f is the number of features:

$$\mu_{ij} \in [0,1] \qquad \forall i = 1, 2, \ldots, c; \qquad \forall j = 1, 2, \ldots, f \tag{20}$$

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{f} \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}}} \tag{21}$$

where N= size of data, m (m>1) is a scalar termed the weighting exponent and controls the fuzziness of the resulting clusters, in this study we used m=2, d_{ij} is the Euclidian distance from data x_i to the centroid of feature data j^{th}, z_j

$$d_{ij} = \| x_i - z_i \| \tag{22}$$

The fuzzy entropy of the feature of class c FE_C, is given by

$$FE_C = -D_c \log D_c \tag{23}$$

Where D_c is a match degree of class c calculated by:

$$D_c = \frac{\Sigma_{x_{i \in c}} U_c(x_i)}{\Sigma_{x_{i \in c}} U_c(x_i)} \tag{24}$$

Fuzzy entropy of feature used for selected feature calculated by

$$EC = \Sigma_{c=1}^{C} FE_c \tag{25}$$

From FS-FA described in previous section, feature j^{th} is selected if $wij >$ threshold τ. In this approach, to select the related features, the subset of features is filtered again with fuzzy entropy, using threshold ϕ. The pseudo code of feature selection using firefly algorithm (FS-FA) is shown in Fig. 6.

```
Feature Selection using Fuzzy-based Firefly Algorithm (FS-FFA)
Objective function f(x), x = (x_1, ..., x_d)^T,
Define the parameters
Generate initial population of fireflies x_i , x_i = [w_i1, w_i2, w_i3, ..., w_id] , i = 1,2, ..., n
Calculate Light intensity I_i at x_i is determined by f(x_i)
Repeat:
    for i = 1:n all n fireflies
        for j = 1:n all n fireflies
            if (I_j > I_i), Move firefly i towards j in d-dimension Attractiveness varies with
                    distance r via exp[-γr]
                Select subset of features F' where w_ij > τ
                Calculate Fuzzy Entropy of feature in F'
                Select subset of features from F' where EC_j > ϕ
                Calculate MCC
                Update light intensity
        if no one of firefly brighter than I_i, I_i move randomly
    Rank the fireflies and find the current best
Until: Maximum iteration or minimum change of objective function
```

Fig. 6. The pseudo code of feature selection using fuzzy-based firefly algorithm (FS-FFA)

The experiments of glistenings detection compared FS-FA and FS-FFA method are described in section 3.

3 Experiments and Results

In this experiment, we implemented FS-FA and FS-FFA algorithm and compared the result with the simple detection method, no classification, to test that our approach can improve the performance of glistenings detection.

From section 2.3, 9,858 blobs are detected from 27 IOL images and 68 features are extracted. The classification model is evaluated using leave-one-out cross-validation. Leave-one-out cross-validation (LOOCV) used data of 1 image as a testing set and the rest data from 26 images as a training set. In learning process of firefly algorithm, 1/3 of the training set are divided as a validation set. Parameters used in the experiment shown in Table 2.

Table 2. Parameter setting

Parameter	FS-FA	FS-FFA
Number of fireflies	5	5
Neighbors (k in k-NN)	15	15
Maximum iteration	50	50
Randomness	0.2	0.2
Attractiveness	1	1
Light absorption coefficient	0.9	0.9
Threshold τ	0.4	0.4
Threshold ϕ	-	0.2

The different performance measures precision, recall, f-measure, MCC, and DOR are used for validation. Precision, Recall, and F-Measure are chosen as the measurements of the accuracy of the algorithms. Precision measures the exactness of the classifiers while recall measures the completeness of the classifiers. The F-measure score shows the balance between the precision and the recall. The precision, recall and F-measure are computed by:

$$Precision = \frac{TP}{TP + FP} \quad \textbf{(26)}$$

$$Recall = \frac{TP}{TP + FN} \quad \textbf{(27)}$$

$$F\_Measure = 2\frac{Precision \times Recall}{Precision + Recall} \quad \textbf{(28)}$$

The diagnostic odds ratio (DOR) is a measure of the effectiveness of a diagnostic test [35]. DOR is calculated by:

$$DOR = \frac{(TP / FP)}{(FN / TN)} \quad \textbf{(29)}$$

The result of classification is shown in Table 3. The result shows that all measure values from FS-FFA better than FS-FA and without feature selection. Graphical comparisons of the glistenings detection result with mathematical morphology method and fine segmentation with feature selection and classifier, according to precision, recall and f-measure of the glistenings detection are shown in Fig. 7, while the comparison of MCC and DOR are shown in Fig. 8 and Fig. 9.

Table 3. Object-based classification validation result

	Precision	Recall	F1	MCC	DOR
morphology	0.7045	0.8338	0.7452	-0.1454	0.3116
morphology+k-NN,	0.8319	0.8479	0.8263	0.4100	14.2768
morphology+Naive Bayes	0.8739	0.7165	0.7704	0.3932	15.6563
morphology+FS-FA	0.8289	0.8492	0.8250	0.4007	13.6224

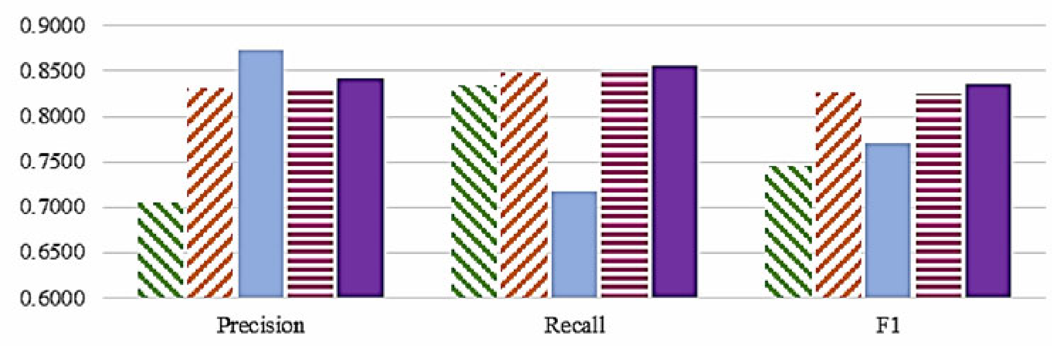


Fig. 7. Graphical comparisons according to precision, recall and f-measure

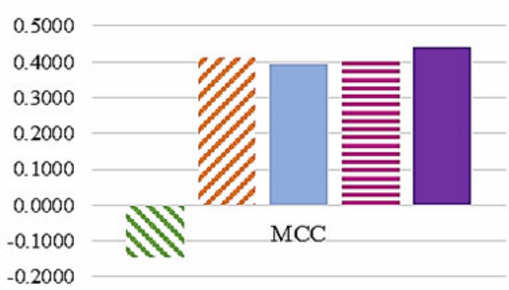


Fig. 8. Graphical comparisons according to Matthews's correlation coefficient (MCC)

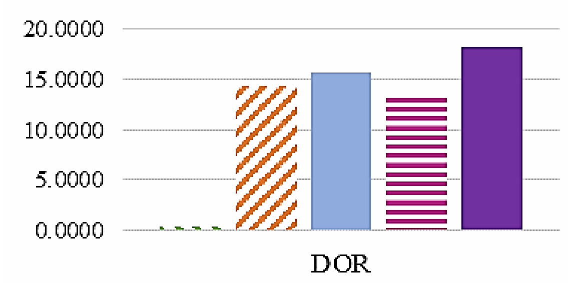


Fig. 9. Graphical comparisons according to diagnostic odds ratio (DOR)

Table 4 show the average number of selected features and the execution time of training data and testing data. Glistenings detection time in real application in testing time because the model of classifier is already learning. The result shows that feature selection can reduce the running time of detection.

Table 4. Comparing the number of features and running time of algorithm

	Precision	Recall	F1
morphology	0.0	0.0000	0.3093
morphology+k-NN	68.0	0.0000	0.3093
morphology+Naive Bayes	68.0	0.0000	0.0434
morphology+FS-FA	47.8	348.5801	0.1530
morphology+FS-FFA	26.4	112.1230	0.1144

From using FS-FFA, we can reduce the number of false positive (FP) and false negative (FN) in glistenings classification, Fig. 10 shows the classification result on IOL image from k-NN classifier compared with FS-FFA. The true glistenings (TP) are shown in the green edge. Blue edge objects are the candidate glistenings the classified correctly to not glistenings class (TN), while the red edge objects are wrong classified as glistenings (FP). Missing glistenings (FN) are on the yellow edge.

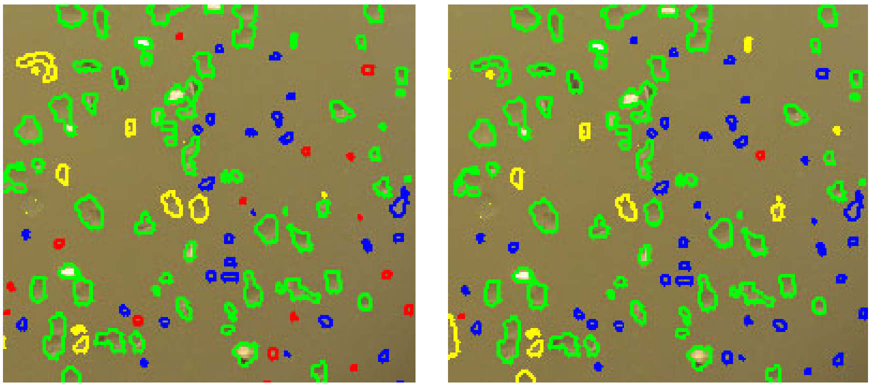


Fig. 10. Classification results, left is the results of fine segmentation using k-NN, right is the results of glistenings detection with FS-FFA

4 Conclusion

Firefly algorithm and fuzzy entropy are proposed for selection feature set to improve the performance of glistenings detection. The outcomes have shown that the MCC and DOR value increase after feature selection utilizing firefly algorithm and fuzzy entropy. A small size of features set also decreased that classification time in a testing phase. Note that, in Fig. 10 while FP and FN are expanded, some TP and TN are diminished as well. For the future study, the detection algorithm needs to enhance to keep up the exactness of classification. From the result, there are many false positive (FP), one reason is the ground-truth pictures are hand-drawn by an ophthalmologist, subsequently, there is an arbitrary mistake gotten from hand dependability and subjective judgment of what is and isn't a glistening. For future study, the mathematical morphological for blobs detection needs to improve, in some image the proposed method cannot segmented glistenings from the images, it causes of a false negative (FN). The problem of overlap of glistenings needs to fix too. The proposed method, Fuzzy-based firefly algorithm for glistenings detection, also can study to apply to other blobs detection problem.

Acknowledgements

We would like to thank the Applied Vision Research Centre, School of Health Sciences, City University London, for the IOL images and ground truth data.

References

[1] R. Bellucci, An introduction to intraocular lenses: material, optics, haptics, design and aberration, Cataract, ESASO Course Series 3(2013) 38-55.

[2] M. Malley, AcrySof® "glistenings" and questions of haze, Ophthalmology Times 20(18)(1995) 1.

[3] L. Werner, Glistenings and surface light scattering in intraocular lenses, Cataract & Refractive Surgery 36(8)(2010) 1398-1420.

[4] K. Hayashi, A. Hirata, M. Yoshida, K. Yoshimura, H. Hayashi, Long-term effect of surface light scattering and glistenings of intraocular lenses on visual function, American Journal of Ophthalmology 154(2)(2012) 240-251.

[5] D. Tognetto, L. Toto, G. Sanguinetti, G. Ravalico, Glistenings in foldable intraocular lenses, Cataract & Refractive Surgery 28(7)(2001) 1211-1216.

[6] N. Mamalis, Intraocular lens glistenings, Cataract & Refractive Surgery 38(7)(2012) 1119-1120.

[7] E. Mönestam, A. Behndig, Impact on visual function from light scattering and glistenings in intraocular lenses, a long-term study, Acta Ophthalmol 89(8)(2011) 724-728.

[8] K. Hayashi, A. Hirata, M. Yoshida, K. Yoshimura, H. Hayashi, Long-term effect of surface light scattering and glistenings of intraocular lenses on visual function, American Journal of Ophthalmology 154(2)(2012) 240-251.

[9] M. van der Mooren, L. Franssen, P. Piers, Effects of glistenings in intraocular lenses, Biomedical optics express 4(8)(2013) 1294-1304.

[10] K. Kato, M. Nishida, H. Yamane, K. Nakamae, Y. Tagami, K. Tetsumoto, Glistening formation in an AcrySof lens initiated by spinodal decompositionof the polymer network bytemperature change, Cataract & Refractive Surgery 27(9)(2001) 1493-1498.

[11] N.Z. Gregori, T.S. Spencer, N. Mamalis, R.J. Olson, In vitro comparison of glistening formation among hydrophobic acrylic intraocular lenses, Cataract & Refractive Surgery 28(77)(2002) 1262-1268.

[12] A. Abarghouei, A. Ghanizadeh, S. Sinaie, S. Shamsuddin, A survey of pattern recognition applications in cancer diagnosis, in: Proc. 2009 International Conference of Soft Computing and Pattern Recognition, 2009.

[13] A. Sopharak, M. Dailey, B. Uyyanonvara, S. Barman, T.H. Williamson, K.T. New, Y.A. Moe, Machine learning approach to automatic exudate detection in retinal images from diabetic patients, Modern Optics 57(2)(2010) 124 -135.

[14] P. Jitpakdee, B. Uyyanonvara, Computer-aided detection and quantification in glistenings on intra-ocular lenses, Multimedia Tools and Applications 76(18)(2017) 1-14.

[15] M. Dash, H. Liu, Feature Selection for Classification, Intelligent Data Analysis 1(1)(1997) 131-156.

[16] J. Tang, S. Alelyani, H. Liu, Feature selection for classification: a review, in: C. C. Aggarwal (Ed.), Data Classification: Algorithms and Applications, CRC Press, New York, 2014, pp. 37-64.

[17] S.B. Kang, Advances in Computer Vision and Pattern Recognition, Springer, London, 2012.

[18] J.-D Shie, S.-M. Chen, Feature subset selection based on fuzzy entropy measures for handling classification problems, Appl Intell 28(2008) 69-82.

[19] H. Banati, M. Bajaj, Fire-fly based feature selection approach, Computer Science Issues 8(4)(2011) 473-480.

[20] M. Ghaemi, M. Derakhshi, Feature selection using forest optimization algorithm, Pattern Recognition 60(2016) 121-129.

[21] Z. Sun, G. Bebis, R. Miller, Object detection using feature subset selection, Pattern Recognition 37(11)(2004) 2165-2176.

[22] L. Cervante, B. Xue, M. Zhang, L. Shang, Binary particle swarm optimization for feature selection: a filter based approach, in: Proc. 2012 IEEE Congress on Evolutionary Computation, 2012.

[23] V. Subha, D. Murugan, Opposition-based firefly algorithm optimized feature subset selection approach for fetal risk anticipation, Machine Learning and Applications 3(2)(2016) 55-64.

[24] L.A.M.Pereira, D.Rodrigues, T.N.S.Almeida, C.O.Ramos, A.N.Souza, X.S.Yang, J.P.Papa, A Binary Cuckoo Search and Its Application for Feature Selection, Springer International Publishing, Switzerland, 2014.

[25] E. Emary, H. Zawbaa, K. Ghany, A. Hassanien, B. Parv, Firefly optimization algorithm for feature selection, in: Proc. the 7th Balkan Conference on Informatics Conference, 2015.

[26] D. Sheela, A. Pethalakshmi, Optimized feature selection algorithm for high dimensional data, Indian Journal of Science and Technology 9(31)(2016). doi: 10.17485/ijst/2016/v9i3/79656

[27] A. Rhodes, L. Bai, Circle detection using a gabor annulus, in: Proc. 2011 British Machine Vision Conference, 2011.

[28] J.G. Zhang, T. Tan, L. Ma, Invariant texture segmentation via circular gabor filters, in: Proc. 2002 International Conference on Pattern Recognition, 2002.

[29] X. S. Yang, Nature-Inspired Metaheuristic Algorithms, Luniver Press, 2008.

[30] X. S. Yang, Firefly algorithm stochastic test functions and design optimization, International Journal of Bio-inspired

Computation 2(2010) 78-84.

[31] N.S. Altman, An introduction to kernel and nearest-neighbor nonparametric regression, The American Statistician 46-3(1992) 175-185.

[32] E. Deza, M. Marie, Encyclopedia of Distances, Springer, London, 2009.

[33] M. Dhar, R. Chutia, S. Mahanta, A note on the existing definition of fuzzy entropy, Energy, Information and Communications 3(1)(2012) 17-22.

[34] B.W. Matthews, Comparison of the predicted and observed secondary structure of T4 phage lysozyme, Biochimica et Biophysica Acta (BBA) - Protein Structure 405(2)(1975) 442-451.

[35] S.G. Afina, G.L. Jeroen, H.P. Martin, J.B. Gouke, M.M. Patrick, The diagnostic odds ratio: a single indicator of test performance, Clinical Epidemiology 56(11)(2003) 1129-1135.