

PAPER • OPEN ACCESS

## Corneal arcus classification for hyperlipidemia detection using gray level co-occurrence matrix features

To cite this article: R A Ramlee *et al* 2020 *J. Phys.: Conf. Ser.* **1432** 012084

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Corneal arcus classification for hyperlipidemia detection using gray level co-occurrence matrix features

Ramlee R A<sup>1\*</sup>, Subramaniam S K<sup>1</sup>, Yaakob S B<sup>2</sup>, Rahman A S F<sup>2</sup> and Saad N M<sup>1</sup>

<sup>1</sup>Advanced Sensors & Embedded Control System Research Group(ASECs), Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100, Durian Tunggal, Melaka, Malaysia.

<sup>2</sup>Pusat Pengajian Kejuruteraan Sistem Elektrik, Universiti Malaysia Perlis, Kampus Pauh Putra, Perlis

Email: ridza@utem.edu.my

**Abstract:** The arcus cornea is an eye problem that is often encountered among older people, but the situation is risky to teenagers and young people, in which it gave a sign the presence of lipid abnormalities in their blood and the risk of disease problems such as diabetes and heart disease. This paper presents the classification of the arcus cornea, using the extraction of texture features of the gray level co-occurrence matrix, along with several models of the classifiers, namely as scale conjugate gradient, Bayesian regulation, and Levenberg-Marquardt. Percentage fractions for training, testing and validation for classifier are 70%, 15%, and 15% respectively. The comparison of the classifiers used by the past researchers for classification the eye abnormalities, also were analyzed and studied in this work. In this experiment, a total of 125 image eyes were used, consisting of two classes of the eye image, which is normal and abnormal. The best result demonstrated in this proposed framework using Bayesian regulation classifier is, a sensitivity of 96%, and a specificity of 100%. However, this classifier did not achieve perfectly classification or an accuracy of 100%. Nevertheless, it is able and evident that the system is effective by the output of 98.4% accuracy.

## 1. Introduction

Corneal arcus (CA) is the symptom of hyperlipidemia. The forming of lipids around the iris is due to excessive levels of fats in the blood, will cause the development of CA. Figure 1 (a) shows, the examples of eye abnormalities, with the presence of lipid as a white ring, at the sclera-iris area. While, the normal eye in Figure 1(b), show it is clear from the presence of CA.



**Figure 1.** Examples of images used in this experiment, (a) the situation seemed to be like the white ring is clearly visible as a result of lipids that exist, (b) the eyes are free from the condition of the white ring (lipid).

The presence of CA normally happen to the elderly group, aged 50 years and above[1]. However, this situation is unusual if it happens to individuals aged 40 years and below as specified by [2]. This lipid thickness according to [1], normally occur between 0.3 to 1 mm width.

The researchers [3]–[11] have studied about CA, some of them agreed, the presence of CA are associated with the abnormal lipids in the blood. While others say, it is due to other factors such as smoking and alcohol intake [4].

In [12], the fuzzy k-means was used in order to extract the features from the eye images. These features are fed to the radial basis function network (RBFN) for training and testing. About 150 samples including normal and abnormal (i.e. suffer from CA, glaucoma, and cataract) are used in their experiment. Their proposed system shows an excellent performance for classification, with a specificity of 100%, a sensitivity of 90%, and the overall accuracy is 95%.

In another work Acharya [13] et al., have made a comparison between the three types of classification techniques for classify the abnormalities of the eye. A total of 135 subjects used for the classification of eye diseases. Their results produce more than 85% for sensitivity, and specificity was 100%.

Mahesh [14] et al. used the support vector machine (SVM) for diagnosing the CA. Using five attributes of statistical features such as mean, standard deviation, entropy, skewness, and kurtosis. These features are swapped to SVM classifier for training and testing. The result demonstrated by their proposed system as following: sensitivity (94%), sensitivity (98%) and accuracy (96%).

Another research study on CA, such as work done by Nasution et al. [15]–[17]. The proposed system developed based on a digital image acquisition system, using low cost components. In their work, the hybrid N-feature neural network (HNFNN), for recognition is used. The system they developed, providing an accuracy of 93.6% in determining the length of the arc CA [15].

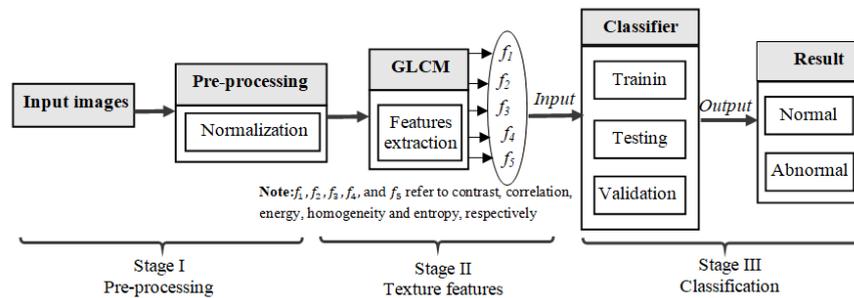
The proposed method using OTSU threshold has been explained by the authors in previous studies for detecting the CA [18]. It has been identified the drawback related to the image intensity, because the work depends on the threshold value of the histogram for image classification to distinguish between normal and abnormal eye's image. The problem will appear when various intensity values of images are used. To overcome this problem, in this paper, a different approach to classifying the data-set classes is proposed.

The novelty of this work, the used of the frame-work as proposed. Where, the classification data-set using the classifiers (BR, LM, and SCG) which never been tried for this problem (classification CA). This is for filling the gap of the studies for trying another method to solve the problem of CA's classification. This work conducted to improve the findings of previous studies conducted by other researchers. The CA's presence normally occurs most frequently in the older peoples. However, in the case of CA happen upon the younger, it is reported significant to show, there is the abnormality of lipid in blood.

## 2. Methodology

The automated system for CA classification uses the eye image as an input in order to find the lipid deposition (like a white ring) presence in the eye, specifically in the peripheral cornea. This paper presents a method for the screening of CA appearance, using eye images. Figure 2, shows the proposed framework applied in this paper. This proposed system divided into three stages, i.e. image extraction, features vectors and image classification. In the first stage (stage I), the image input in the form of RGB images will be converted to the grayscale intensity image. The extraction of the image features is acquired from the properties of the mage using the gray level co-occurrence matrix (GLCM). In stage II, the features statistical values are calculated from this GLCM's matrix. In final

stage, the classifier is trained, tested and validated using the statistic's features calculated from the GLCM's output.

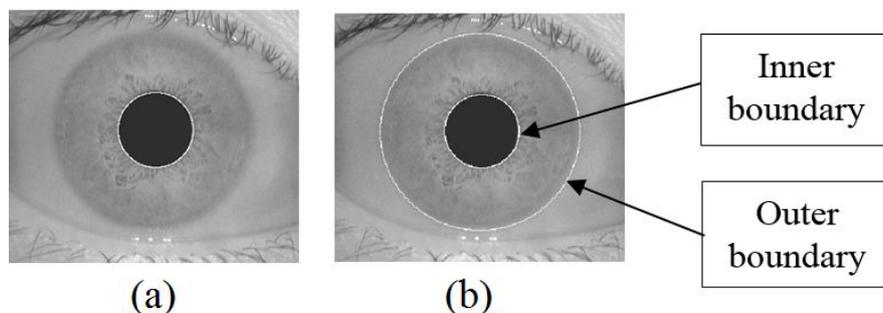


**Figure 2.** The framework for the proposed system for CA classification

Several data-set have been used in this experiment. These data-sets are divided into two groups, which are the normal and abnormal image contain CA presence. For normal data-set the sample acquired from the free database sources such as from Institute of Automation, Chinese Academy of Sciences [19],[20] and IITD databases [21]. While, for second group (i.e. the abnormal eye) obtained from medical websites, and journals related to the CA study[1], [3], [12], [15], [22]–[24].

### *Pre-processing*

The first stage of this work is the pre-processing, for this the images are processed for normalizations from the data-set. The process begins with localization and detection two approximate circles. These circle namely as the inner and the outer circle. The inner circle is the region of pupil boundary, while the outer-circle represents to the sclera-iris boundary. The iris localization process, is to detect the center-point of the iris and pupil. From these center-point of pupil and iris, the inner and outer ring can be used for the next process, which is the segmentation (Figure 3).The segmentation process begins by detect the circle of both circle of pupil and iris. The segmentation is done based on circle radius and center of both boundaries. After successful detection of the circle around the iris and the pupil, the circles are drawn on the both boundaries to mark the pupil and the iris boundary. The process will involve grey scale conversion, enhancement, morphological, and detecting circle using the Daugman Integro-differential operator (DIDO) [25].The process, including detection of the circle or boundary of the iris and the pupil. After detecting of the iris and the pupil region, the unwanted area outside the outer region of the iris is eliminated. This refers to the process called cropped (remove the unwanted area). The segmentation iris will be transformed to a rectangular shape using Daugman rubber sheet technique [25] become the normalization image.



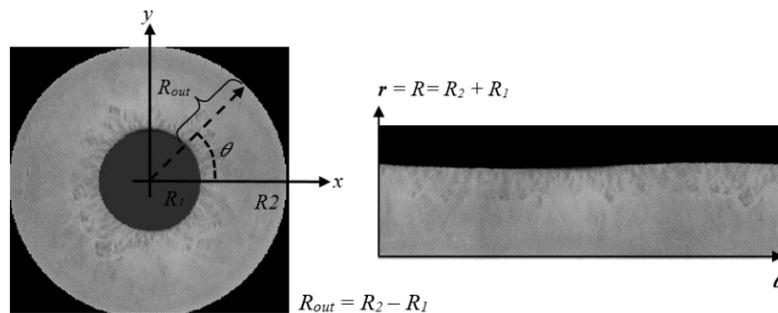
**Figure 3.** The localization and segmentation result of pupil and iris, (a) pupil segmentation, (b) segmentation of inner and outer boundary (both pupil and iris).

In (1), present the DIDO equation where,  $I(x, y)$  is the eye domain for the eye image data-set. The function will find over the image domain of  $x$  and  $y$ , for the maximum in the blurred partial-derivatives referring to the radius  $r$ , of the normalized contour integral,  $I(x, y)$ . This searching involves along a circular arc  $ds$  of radius  $r$  and the center coordinates  $(x_0, y_0)$ . The convolution  $*$  symbol, in (1)

serves as a smoothing function such as a Gaussian with the scale of sigma ( $\sigma$ ). Operator DIDO act as the circular edge detector blur according to scale of sigma  $\sigma$ , which examines repeatedly all over the image space over the parameter set  $(x_0, y_0, r)$  [25], [26].

$$\max_{(r, x_0, y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right| \quad (1)$$

To describe this condition, Figure 4 illustrates how this normalization is performed using Daugman's rubber sheet. All points inside the iris area of this model are remapped to polar coordinates  $(r, \theta)$ . For this  $\theta$  represents as the angle  $[0, 2\pi]$  while  $r$  is the interval  $[0, 1]$ . This remapping process Cartesian coordinates  $I(x, y)$  to polar coordinates  $(r, \theta)$ .



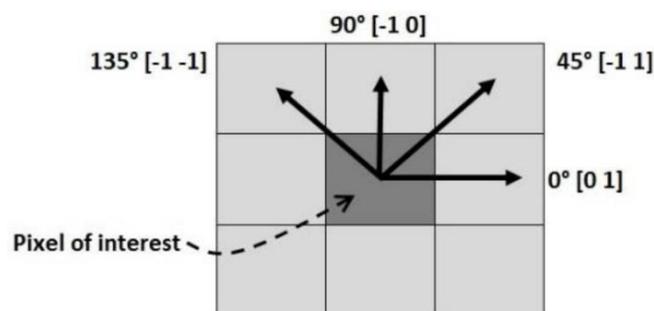
**Figure 4.** The normalization of the segmented image transform from Cartesian to polar coordinates.

#### Features Extraction

The second stage of the experiment, is the features extraction process. Feature' sex traction is related to dimension reduction and dimension selection of images. When the input data (image) processed to the algorithm is too big then it needs to be changed into a set of the texture features that it can also be characterized as a feature-vector. In this experiment the gray level co-occurrence matrix (GLCM) is implemented for feature extraction.

#### Gray Level Co-occurrence Matrix (GLCM)

By using GLCM, the matrix image formed as a result of the distribution of co-occurring values according to the setting of the offset can be obtained. This offset can be set in the following ranges of the orientation constant ( $0^\circ, 45^\circ, 90^\circ, \text{ and } 135^\circ$ ) [27].



**Figure 5.** The Pixels movement by the angle of the array offset in GLCM.

Given,  $I$  as represents the image, of the  $N \times N$ , thus the co-occurrence matrix of  $P$  can be defined as:  $P(i, j | \Delta x = 1, \Delta y = 0)$ . Where the  $(\Delta_x, \Delta_y)$ , is the offset specifying the distance between the pixel-of-interest and its neighbour. The orientation of the angle offset can be represent by the distance (i.e.  $[0 \Delta]$  for  $0^\circ$ ,  $[-\Delta, \Delta]$  for  $45^\circ$ ,  $[-\Delta 0]$  for  $90^\circ$ , and  $[-\Delta -\Delta]$  for  $135^\circ$ ). The GLCM can be divided into three main direction, namely as horizontal ( $0^\circ$ ), vertical ( $90^\circ$ ), and diagonal. And for diagonal it has two directional of analysis, that is bottom left to top right ( $-45^\circ$ ) and bottom right to top left ( $-135^\circ$ ). Each of these directional analysis offsets are denoted as;  $P0^\circ, P45^\circ, P90^\circ, \text{ \& } P135^\circ$  respectively as illustrate in Figure 5.

### Texture Features

The following definition and their equation are referring to the image features. As shown in Figure 2 the notation of  $f_1, f_2, f_3, f_4,$  and  $f_5$  refer to contrast, correlation, energy, homogeneity and entropy.

In contrast ( $f_1$ ), the local variation will be calculated from the GLCM matrix and obtained from calculations in (2).

$$f_1 = \sum_{i,j} |i - j|^2 p(i, j) \quad (2)$$

The correlation ( $f_2$ ), measures the amount of joint probability occurrence of the specified pixel pairs in the images as (3). Where  $\mu$  is the GLCM mean, given by,  $(-\sum_{i,j=0}^{M-1} iP_{ij})$ .

$$f_2 = \sum_{i,j} \frac{(i - \mu i)(j - \mu j)p(i, j)}{\sigma_i \sigma_j} \quad (3)$$

The energy ( $f_3$ ), provides the amount of squared elements contain in the GLCM matrix, also known as the angular second moment or uniformity, as described by (4).

$$f_3 = \sum_{i,j} p(i, j)^2 \quad (4)$$

The homogeneity ( $f_4$ ), measures the nearness of the distribution of components in the GLCM to the diagonal of it.

$$f_4 = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (5)$$

In the image processing, the entropy ( $f_5$ ) is used to calculate the statistical attributes for determined the image texture. It's can be described as:

$$f_5 = - \sum_{k=0}^{M-1} p_k \log_2(p_k) \quad (6)$$

According to [28], feature selection is an important step in data processing where it can reduce irrelevant features [29], and improve the performance of classification. In [30], they used the Kruskal-Wallis, test to the suitable features from the electroencephalogram (EEG) for identify the hand movement.

We used the non-parametric Kruskal-Wallis (KW) test, to ensure that the characteristics of the two groups are of the same distribution or otherwise. This test is a kind of non-parametric statistical measure extension from Wilcoxon rank test. Based on the KW test, ANOVA parameters, each of the individual characteristics of the features can be analysed. The Kruskal-Wallis test assumes that all samples come from populations having the same continuous distribution, apart from possibly different locations due to group effects, and that all observations are mutually independent. By contrast, classical one-way ANOVA replaces the first assumption with the stronger assumption that the populations have normal distributions. The probability ( $P$ ) value obtained by the KW test shown in Table 1 demonstrated that the individual features for the normal and abnormal data are not the same distribution, at a significance level of 5%.

Table 1 shows the statistical analysis using KW for all features proposed in this experiment, namely as the contrast ( $f_1$ ), correlation ( $f_2$ ), energy ( $f_3$ ), homogeneity ( $f_4$ ), and entropy ( $f_5$ ). The p-values in the last column of Table 4 are compared to the Chi-square values for showing the features are significant or not. From this table all feature except feature 3, indicate rejects the null hypothesis that the features

come from the same distribution at a 5% significance level.

Table 1. Kruskal-Wallis ANOVA Table

Features	$\chi^2$	$P > \chi^2$
Contrast ( $f_1$ )	8.9451	0.0027822
Correlation ( $f_2$ )	3.8418	0.04999
Energy ( $f_3$ )	0.0705	0.79062
Homogeneity ( $f_4$ )	16.872	4.00E-05
Entropy ( $f_5$ )	5.1589	0.023128

$\chi^2$ =chi-square,  $P$ =probability

### Classification

The final stage in this proposed system is classification. This classification is the binary-classification of normal and abnormal classification. In this experiment, three classifiers are tested for determining the best classifier among these. The classifiers used are Bayesian regularization (BR), scaled conjugate gradient (SCG), and Lavenberg Marquardt (LM).

The input of the statistical features are fed into neural network for training, testing and validation the data-set. For that the data-set are split in to 70% for training, 15% for testing and remaining 15% for validation. This ratio are implemented in the process of classification.

### Bayesian Regulation Back-propagation Algorithm

The over-fitting problems often interfere in the training process of the neural network. For this reason, Foresee [31] used, the BR implemented within the framework of the Levenberg-Marquardt algorithm to solve this problem.

In Equation (7), a data-set of  $D = \{t, (p_i)_{i=1, \dots, n}\}$ ,  $p_i$  is an input vector of the individual and vector of target of  $i$  and  $t$  respectively.  $W$ , is weight give to the network define a mapping of the input vector  $p_i$  to the output  $\hat{t}_i$ .  $M$ , is a network model in term of the architecture of the network.  $E_D$ , is the sum of squared estimation errors, (used as the typical objective function for network for early stopping) and  $n$ , is input-target pair defining in data-set.

$$E_D(D|w, M) = \sum_{i=1}^n (t_i - \hat{t}_i)^2 \quad (7)$$

In Equation(8),  $E_w$  is the sum squared weights which is added to the objective function of early stopping. It to smooth the interpolation through the training data of the network by penalize large weight.

$$F = \beta E_D(D|w, M) + \alpha E_w(w|M) \quad (8)$$

The  $\alpha E_w$  in (8) is the weight decay and  $\alpha$ , it favors small value of  $w$  and decreases the tendency of a model from over-fit [32], [33]. The posterior distribution of  $w$  given  $\alpha$ ,  $\beta$ ,  $D$  and  $M$  state below as (9)

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)} \quad (9)$$

In (9),  $D$  is the data-set for training,  $P(w|D, \alpha, \beta, M)$  is the posterior probability of  $w$  and  $P(D|w, \beta, M)$  is the likelihood function of  $w$ .  $P(w|\alpha, M)$  is the prior distribution of weight under  $M$  which is the probability of observing the data given  $w$ .  $P(D|\alpha, \beta, M)$  is the factor of normalization for hyper-parameters  $\alpha, \beta$  as stated by [34]. They stated that the  $P(D|\alpha, \beta, M)$  does not depend on

weight  $w$ , the factor of normalization is  $[P(D|\alpha, \beta, M) = \int P(D|w, \beta, M)P(w|\alpha, M)dw]$ . The weights  $w$  are assumed to be identically distributed, each following the Gaussian distribution  $(w|\alpha, M) \sim (0, \alpha^{-1})$ . Given this, the expression of joint prior density of  $w$  as in Equation (9). Assuming noise of training data and prior distribution of weights  $P(w|\alpha, M)$  are Gaussian, (10) and (11) can be established.

$$P(D|w, \beta, M) = \frac{1}{Z_D(\beta)} \exp\left[-\beta E_D\right] \quad (10)$$

$$P(w|\alpha, M) = \frac{1}{Z_w(\alpha)} \exp[-\alpha E_w] \quad (11)$$

Where  $Z_D(\beta) = \left(\frac{\pi}{\beta}\right)^{\frac{n}{2}}$  and  $Z_w(\alpha) = \left(\frac{\pi}{\alpha}\right)^{\frac{n}{2}}$ . Substituting equation (10) and (11) in (9) will obtained  $P(w|D, \alpha, \beta, M) = \frac{1}{Z_F(\alpha, \beta)} \exp[-F(w)]$ . From this  $Z_F$  can be derived, as in (12):

$$Z_F(\alpha, \beta) = Z_D(\beta)Z_w(\alpha).P(D|\alpha, \beta, M) \quad (12)$$

The objective is to determine the weights for minimize  $F(w)$ : this will maximizing  $P(w|D, \alpha, \beta, M)$ . Once again using Bayes' rule, applied  $P(\alpha, \beta|D, M) = \frac{P(D|\alpha, \beta, M)P(\alpha, \beta|M)}{P(D|M)}$ . The Posterior density,  $P(\alpha, \beta|D, M)$  has the same form with Equation (9). If we assume,  $P(\alpha, \beta|M)$  which is the prior density as a constant (uniform), for the regularization parameter of  $\alpha$  and  $\beta$ . Maximizing the posterior is obtained by maximizing the likelihood function, which is  $P(\alpha, \beta|D, M)$ , will give:

$$P(\alpha, \beta|D, M) = \frac{\left(\frac{1}{Z_D(\beta)}\right) \exp\left[-\beta E_D\right] \left(\frac{1}{Z_w(\alpha)}\right) \exp\left[-\alpha E_w\right]}{\left(\frac{1}{Z_F(\alpha, \beta)}\right) \exp\left[-F(w)\right]} \quad (13)$$

$$P(\alpha, \beta|D, M) = \left(\frac{Z_F(\alpha, \beta)}{Z_D(\beta)Z_w(\alpha)}\right) \quad (14)$$

By deriving Equation (14), the  $Z_D(\beta)$ , and  $Z_w(\alpha)$ , are acquired.  $w^{MP}$  is the minimum point of the posterior density, where the gradient is zero. The estimation of  $Z_F(\alpha, \beta)$  can be obtained using Taylor's series expansion. By solving the normalization constant, the Equation (15) can be written, as;

$$Z_F = (2\pi)^{\frac{N}{2}} (\det(H^{MAP})^{-1})^{\frac{1}{2}} \exp(-F(w^{MP})) \quad (15)$$

$H$ , is the Hessian matrix defines by  $H = \beta \nabla^2 E_D + \alpha \nabla^2 E_w$ . Substitute the  $Z_F$  as in Equation (15) in (14) and perform derivative of the log both sides of the equation and comparing them to zero, thus the optimal values of  $\alpha$  and  $\beta$  at  $w^{MP}$  can be obtained as;

$$\alpha^{MP} = \frac{\gamma}{2E_w(w^{MAP})} \text{ and } \beta^{MP} = \frac{n - \gamma}{2E_D(w^{MAP})} \quad (16)$$

$\gamma = N - 2\alpha^{MP} \text{tr}(H^{MP})^{-1}$ , defines as the effective number of parameters. It is the degrees of freedom in the neural network for giving the weights and biases effectively used by the network. The function  $F$ , Levenberg–Marquardt algorithm is minimizing using Levenberg–Marquardt algorithm by [35].

*Scaled conjugate gradient back-propagation algorithm*

According to [36], the scaled conjugate gradient technique is very suitable for the large networked operation because the matrix inverse calculation scales geometrically with the number of weights. Another application of SCG is in solving the issues related to pattern recognition problems.

In [36] Hagan stated, a set of vectors  $\{P_k\}$  is mutually conjugate with respect to a positive definite Hessian matrix. In (17), shows the initial searching in the steepest descent direction with  $P_o$  to be the negative of the gradient ( $-g_o$ ).

$$p_o = -g_o \quad (17)$$

$$x_{k+1} = x_k + \alpha_k g_k \quad (18)$$

$$p_k = -g_k + \beta_k p_{k-1} \quad (19)$$

Møller [38], introduced the optimization approach on SCG, where according to them this technique is more effective than back-propagation, conjugate gradient algorithm with line search (CGL), and Broyden-Fletcher-Goldfarb-Shanno memory-less quasi-Newton algorithm (BFGS). Their idea is to combine the model of Levenberg-Marquardt algorithm with the conjugate gradient to establish the optimization approach of SCG. For this

In (20) where  $s$  is the approximation of Hessian matrix, where  $E'(w_k)$  is the Hessian scalar which is the gradient of  $E$ , while the scaling factor  $\lambda_k$  and  $\sigma_k$  in (20) is used to approximate the Hessian matrix so that their range as follows;  $0 < \lambda_k < 10^{-6}$  and  $0 < \sigma_k < 10^{-4}$ .

The factor of  $\beta_k$  calculation and direction of the searching step can be shown as in (21) and (22) [38].

$$s_k = \frac{E'(w_k + \sigma_k p_k) - E'(w_k)}{\sigma_k} \quad (20)$$

$$\beta_k = \frac{(|g_{k+1}|^2 - g_{k+1}^T g_k)}{g_k^T g_k} \quad (21)$$

$$p_{k+1} = -g_{k+1} + \beta_k p_k \quad (22)$$

*Levenberg-Marquardt back-propagation algorithm*

LM algorithm was designed to approach second-order training speed without having to compute the Hessian matrix [39], [40]. The relationship between Hessian matrix is given by  $H$  and the Jacobian matrix denote as  $J$  can be rewritten as (23), while in (24) shows the gradient vector denote as  $g$  and  $e$  is a vector of network error.

$$H = J^T J \quad (23)$$

$$g = J^T e \quad (24)$$

The LM algorithm uses the derivation of the Gauss-Newton algorithm, as presented in (25), where the scalar  $\mu$  will make this equation merely as a Newton's method if its value is zero. Meanwhile if the scalar  $\mu$  value is greater, (25) becomes a gradient descent.

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (25)$$

The idea of using Newton's method in LM are, to have fast and accurate (minimum error). Thus, the

scalar  $\mu$  will change increase and decrease accordingly at each iteration. This will make the LM technique is better than the conventional gradient descent techniques[36].

### *Evaluation measurements*

A binary classification is implemented in this experiment. Three classifiers are used to classify labels instances as either positive or negative. Subsequently, results are represented in a format known as a confusion matrix.

### *Confusion Matrix*

To evaluate the output performance of this model, we use the confusion matrix (CM) as shown in Table 2. The statistical values are calculated from this table for evaluation the algorithm performance. From these attributes some statistical parameters for showing the algorithm performance can be calculated such as the specificity (Sp), sensitivity (Se), and accuracy (Acc).

Russell [41], called the confusion matrix as error matrix or contingency table. He demonstrated the usage of this error matrix for the classification system, the sampling scheme, the sample size, spatial autocorrelation, and the assessment techniques.

In the confusion matrix, there are four different situations to represent the data. The elements of true positive (TP) in Table 2, represents the number of samples in the classified correctly by the classifier as a positive (abnormal eye). Meanwhile, the elements of true negative (TN), represents the number of samples in the classified correctly by the classifier as a negative (normal eye). The confusion matrix will gain 100% accuracy if both TP and TN detected correctly for each class.

**Table 2.** The confusion matrix with statistical calculation

		Actual	
		Positive	Negative
Predict	Positive	<b>TP</b>	<b>FP</b>
	Negative	<b>FN</b>	<b>TN</b>

The sensitivity (Se) is the measure of the positive elements, which are correctly identified as true by the algorithm, (e.g. the percentage of the eye samples which are correctly identified as having the condition of the disease). The sensitivity calculation is given by (26).

$$Se = TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (26)$$

The specificity (Sp) as shown in (27), is a measure of the percentage the negative rate, which is opposite to sensitivity (e.g., the percentage of the eye samples which are correctly identified as not having the disease).

$$Sp = \frac{TN}{N} = \frac{TN}{TN + FP} \quad (27)$$

The accuracy (Acc) can be calculated as (28), is a measure of the positive and negative classes that are correctly identified by the algorithm.

$$ACC = \frac{TP + TN}{P + N} \quad (28)$$

The positive and negative predictive values (PPV and NPV respectively) are the proportions of

positive and negative results in statistics and diagnostic tests that are true positive and true negative results, respectively. Both are shown in (29) and (30).

$$PPV = \frac{TP}{TP + FP} \quad (29)$$

$$NPV = \frac{TN}{TN + FN} \quad (30)$$

To demonstrate the statistic's values of the algorithm's assessment, as tabulated in Table 3, the equations (26) to (30) are used. From the BR's CM the values of each instance are; TP=49, FP=0, TN=74, and FN=2.

$$Se = \frac{49}{49 + 2} = 96.08\%$$

$$Sp = \frac{74}{74 + 0} = 100\%$$

$$ACC = \frac{49+74}{51+74} = 98.4\%$$

$$PPV = \frac{49}{49+0} = 100\%$$

$$NPV = \frac{74}{74+2} = 97.37\%$$

The full statistical values as tabulated in Table 3 below. In this table, also gives the numbers of CM's attributes of classifiers output.

**Table 3.** The performance evaluation comparison for CA classification of BR.

A	FP	TN	FN	Se (%)	Sp (%)	PPV (%)	NPV (%)	ACC (%)
49	0	74	2	96.1	100	100	97.3	<b>98.4</b>

Note: TP = True Positive, FP = False positive, TN = True negative, FN = False negative, Se = sensitivity, Sp = specificity, PPV = Positive predictive value, NPV= Negative predictive value, ACC = Accuracy

### 3. Experimental and Result Analysis

In this section, the quantitative assessment is presented to classify CA by using three types of classifiers, namely BR, SCG, and LM. In order to evaluate the classifiers performance, the confusion matrix (as explained in sub-section 2.4.1) is used to calculate the statistical values.

Table 4 shows the values of all the performance evaluation by previous researchers, and also three types of the classifiers tested in this experiment. It shows the comparison of classifiers extracted from the confusion matrix presented, such as the *Se*, *Sp*, *Acc*, *Ppv*, and *Npv*.

**Table 4.** The comparison of the proposed methods against the previous techniques for the CA classification

Method	Authors	Se (%)	Sp (%)	Ppv (%)	Npv (%)	Acc (%)
RBF	Acharya U et al. 2007[12]	90	100	100	91	95
ANN	Acharya U et al. 2006[13]	85	100	100	80	90.74
FUZZY		89	100	100	83.33	92.73
ANFIS		89	100	100	83.33	92.73

Method	Authors	Se (%)	Sp (%)	Ppv (%)	Npv (%)	Acc (%)
SVM	Mahesh et al. 2016[14]	94	98	97.91	94.23	96
HNFNN	Nasution et al. 2009 [17]	NA	NA	NA	NA	93.3
HNFNN	Nasution et al. 2015 [15]	NA	NA	NA	NA	93.6
LM	Proposed algorithm	84	95.95	93.48	89.87	92
SCG		75.51	93.24	88.37	84.15	85.6
<b>BR</b>		<b>96</b>	<b>100</b>	<b>100</b>	<b>97.37</b>	<b>98.4</b>

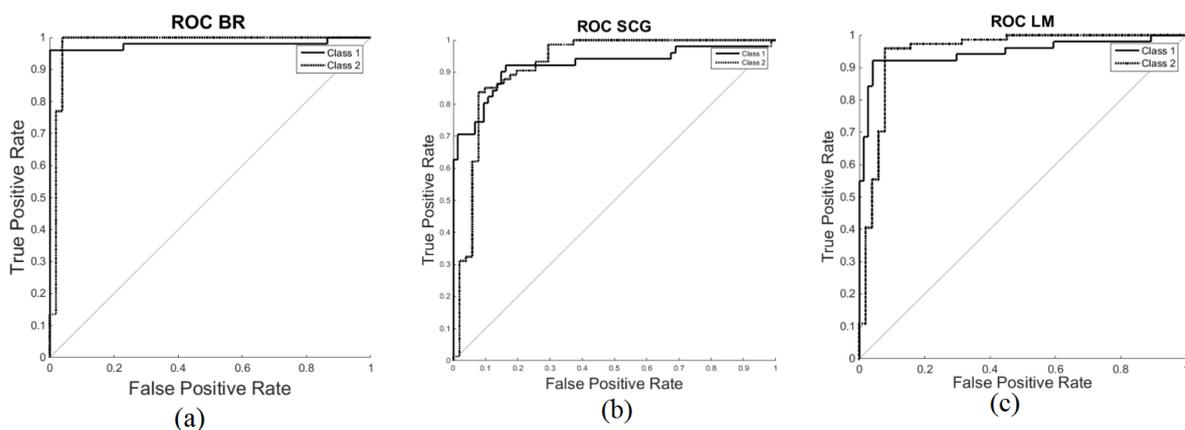
NA = not available.

It highlights, the accuracy results of the classifiers, for comparison. It has shown that the highest percentage of accuracy is by using BR classifier, i.e. by 98.4%. Meanwhile, the lowest percentage of accuracy is to use a classifier SCG, i.e. by 85.6%, where the value still acceptable as a good to the classifier for classifying the CA. From this chart, we can conclude that the classification accuracy using BR classifier (proposed classifier) is the best compared with other types of the classifier.

It's demonstrated the BR classifiers showing superiority as compared to LM, SCG, RBF [12], ANFIS[13], and SVM [14]. We have made a comparative analysis of the proposed system with the methods employed by previous researchers. However, some issues arise with respect to database benchmarking, because between our database and the researchers cited, do not use the same database. In order to get the samples for normal eye, we have to use a database that is publicly available, such as UBARIS[42], CASIA [19], and IITD [21]. Meanwhile, for CA eye we have collected from the images publicly available such [43]–[49] and other medical and iridology websites.

The performance valuation of the classifier using the receiver operating characteristic (ROC) also presented as in Figure 6. For classification purposes, class 1 and class 2 is used to represent the abnormal and normal respectively. The data-sets are trained, tested, and validated by the classifier's algorithm for classification according to each class. Based on the results of this classification, various threshold values of the true positive rate (TPR) and negative positive rate (FPR) will be plotted using ROC curve.

The ROC curve for BR as shown in Figure6 (a), demonstrated the percentage of positive predictive value (PPV) is 98% while for negative predictive value (NPV) is 97.3%. The second ROC graph from SCG performance (Figure6 (b)), the PPV and NPV are 88.4% and 84.1% respectively. Meanwhile, the third ROC graph for LM (Figure6 (c)), the PPV is 93.6% and NPV is 91%.



**Figure 6.** The ROC output of the classifier, (a) BR, (b) SCG, and (c) LM.

#### 4. Conclusion

In this paper, the classifications of the diseased eye using neural network classifiers are implemented. Three types of classifiers namely BR, LM and SCG are used in this experiment to find the best classification the CA. For the inputs of the classifier, we have used the features of GLCM matrix. The Kruskal-Wallis is used to test the inputs data (image features) of the classifier are from the same distribution or not. These image features were divided by the ratio of 70% for training, 15% for testing, and 15% for validation for the process of classification. In this experiment, we found that BR classifier is the best for classification with 98.4% of accuracy.

In the future, this study could be used to classify other types of eye diseases. Research will be done to find the best way to get the image features that will be used in machine learning and neural networks. In addition, the use of larger databases can be taken into account as it can see how relevant the classification process is.

#### Acknowledgements

The authors would like to thank the Universiti Teknikal Malaysia Melaka (UTeM) and Universiti Malaysia Perlis (UniMAP), and all the staffs, who contribute ideas and the materials for this study and submission of the research paper work can be made. The authors would also like to thank the Technical University of Malaysia and the Ministry of Higher Education, which provides guidance, support and sponsorship for this study can be done.

#### References

- [1] F. L. Urbano, "Ocular Signs of Hyperlipidemia," *Hosp. Physician*, pp. 51–54, 2001.
- [2] R. E. C. B. M. P. Wang JJ, "Corneal Arcus and Hyperlipidemia: Findings From an Older Population," *Am J Ophthalmol.*, pp. 363–365, 2002.
- [3] A. Fernández, A. Sorokin, and P. D. Thompson, "Corneal arcus as coronary artery disease risk factor.," *Atherosclerosis*, vol. 193, no. 2, pp. 235–40, Aug. 2007.
- [4] N. Hickey, B. Maurer, and R. Mulcahy, "Arcus senilis: its relation to certain attributes and risk factors in patients with coronary heart disease.," *Br. Heart J.*, vol. 32, no. 4, pp. 449–52, Jul. 1970.
- [5] M. Ang, W. Wong, J. Park, R. Wu, R. Lavanya, Y. Zheng, H. Cajucom-Uy, E. S. Tai, and T. Y. Wong, "Corneal arcus is a sign of cardiovascular disease, even in low-risk persons.," *Am. J. Ophthalmol.*, vol. 152, no. 5, pp. 864–871.e1, Nov. 2011.
- [6] N. T. Cooke, "Significance of arcus senilis in Caucasians.," *J. R. Soc. Med.*, vol. 74, no. 3, pp. 201–4, Mar. 1981.
- [7] I. Bersohn, W. M. Politzer, and D. Blumsohn, "Arcus senilis corneae, its relationship to serum lipids in the South African Bantu.," *South African Med. J.*, vol. 43, no. 33, pp. 1025–1027, Aug. 1969.
- [8] H.-T. Chen, H.-C. J. Chen, C.-H. Hsiao, D. H.-K. Ma, Y.-T. Chen, and K.-K. Lin, "Corneal arcus and cardiovascular risk factors in middle-aged subjects in Taiwan.," *Am. J. Med. Sci.*, vol. 338, no. 5, pp. 268–272, 2009.

- [9] H. S. H. S. Hames CG, "Corneal arcus and coronary heart disease mortality," *Br J Ophthalmol*, vol. 1, no. X, pp. 603–604, 1984.
- [10] H. Z. Pomerantz, "The relationship between coronary heart disease and the presence of certain physical characteristics.," *Can. Med. Assoc. J.*, vol. 86, pp. 57–60, Jan. 1962.
- [11] G. E. Navoyan, "A Case-Control Study of Corneal Arcus and Coronary Heart Disease in Yerevan," University of Armenia, 2003.
- [12] U Rajendra Acharya, L. Y. Wong, E. Y. K. Ng, and J. S. Suri, "Automatic identification of anterior segment eye abnormality," *Itbm-Rbm*, vol. 28, pp. 35–41, 2007.
- [13] U. R. Acharya, N. Kannathal, E. Y. K. Ng, L. C. Min, and J. S. Suri, "Computer-based classification of eye diseases.," *Eng. Med. Biol. Soc. IEEE*, vol. 1, pp. 6121–4, Jan. 2006.
- [14] K. Mahesh and Gunasundari, "Diagnosis of Corneal Arcus Using Statistical Feature Extraction and Support Vector Machine," *Adv. Intell. Syst. Comput.*, vol. 394, 2016.
- [15] A. Nasution, S. Fahdarina, and D. I. Cahya, "System for quantifying the formation stages of corneal arcus," *Proc. SPIE*, vol. 9659, p. 96590M, 2015.
- [16] A. Nasution and D. I. Cahya, "Development of Simple Digital Image Acquisition System for an Accurate Quantification of Corneal Arcus Formation," *Appl. Mech. Mater.*, vol. 771, pp. 112–115, 2015.
- [17] M. T. Nasution and W. W. Kusuma, "Recognition of incremental changes in corneal ring ( arcus-senilis ) using the hybrid N-feature neural network ( HNFNN )," *Proc. 4th Asian Pacific-Rim Symp. Biophotonic (APBP), Jeju-South Korea*, 2009.
- [18] R. A. Ramlee and S. Ranjit, "Using Iris Recognition Algorithm, Detecting Cholesterol Presence," 2009 *Int. Conf. Inf. Manag. Eng.*, pp. 714–717, 2009.
- [19] CASIA, "CASIA iris image database collected by Institute of Automation, Chinese Academy of Sciences," 2003. [Online]. Available: <http://www.sinobiometrics.com>.
- [20] H. Proenc and A. Alexandre, "Iris Recognition : An Analysis of the Aliasing Problem in the Iris Normalization Stage," vol. 2, 2006.
- [21] A. Kumar, "IIT Delhi Iris Database (version 1.0)," 2008. [Online]. Available: [http://web.iitd.ac.in/~biometrics/Database\\_Iris.htm](http://web.iitd.ac.in/~biometrics/Database_Iris.htm).
- [22] L. A. Zech and J. M. Hoeg, "Correlating corneal arcus with atherosclerosis in familial hypercholesterolemia.," *Lipids Health Dis.*, vol. 7, p. 7, 2008.
- [23] P.-B. M., S. F., and A. M., "Corneal arcus as the first presentation of hypercholesterolemia in children," *Pediatr. Pol.*, vol. 89, no. 2, pp. 135–137, 2014.

- [24] O. May, "Detection of High Level Cholesterol in Blood with Iris Analysis," 2016.
- [25] J. Daugman, "How iris recognition works," *Proceedings. Int. Conf. Image Process.*, pp. I-33-I-36, 2004.
- [26] J. G. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence," vol. 15, no. 11, 1993.
- [27] M. M. Mokji and S. a. R. A. Bakar, "Gray Level Co-Occurrence Matrix Computation Based On Haar Wavelet," *Comput. Graph. Imaging Vis. (CGIV 2007)*, no. Cgiv, 2007.
- [28] T. Tanantong, E. Nantajeewarawat, and S. Thiemjarus, "Automatic electrocardiogram signal quality assessment in continuous wireless monitoring," *Maejo Int. J. Sci. Technol.*, vol. 10, no. 02, pp. 127-141, 2016.
- [29] A. Sopharak, B. Uyyanonvara, and S. Barman, "Automated Microaneurysm Detection Algorithms Applied to Diabetic Retinopathy Retinal Images," *Maejo Int. J. Sci. Technol.*, vol. 7, no. 02, pp. 294-314, 2013.
- [30] S. K. Bashar and M. I. H. Bhuiyan, "Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method," *Eng. Sci. Technol. an Int. J.*, 2016.
- [31] F. D. Foresee and M. T. Hagan, "Gauss-Newton Approximation to Bayesian Learning," *Network*, no. 7803, pp. 4122-4128, 1997.
- [32] D. J. C. Mackay and D. J. C. Mackay, *Information Theory , Inference , and Learning Algorithms*. 2003.
- [33] M. Genes, N. Nephropathy, A. A. View, and B. M. Index, "Bayesian Regularized Neural Networks for Small n Big p Data Technology & Medicine," no. October, 2016.
- [34] P. Kumar, S. N. Merchant, and U. B. Desai, "Improving performance in pulse radar detection using Bayesian regularization for neural network training," vol. 14, pp. 438-448, 2004.
- [35] M. T. Hagan and M. B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm," *IEEE Trans. NEURAL NETWORKS*, vol. 5, no. 6, pp. 2-6, 1994.
- [36] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*, vol. 2. 2002.
- [37] P. G. Casazza and G. Kutyniok, "A generalization of gram-schmidt orthogonalization generating all parseval frames," 2005.
- [38] M. F. Møller, "A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning Supervised Learning," *Neural Networks*, vol. 6, pp. 525-533, 1993.
- [39] Ö. Kisi and E. Uncuoglu, "Comparison of three back-propagation training algorithm for two case

- study,” *Indian J. Eng. Mater. Sci.*, vol. 12, no. October, pp. 434–442, 2005.
- [40] O. Baghirli, “Comparison of Lavenberg-Marquardt, Scaled Conjugate Gradient And Bayesian Regularization Backpropagation Algorithms For Multistep Ahead Wind Speed Forecasting Using Multilayer Perceptron Feedforward Neural Network,” no. June, 2015.
- [41] R. G. Congalton, “A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data,” *Remote sens. Env.*, vol. 46, pp. 35–46, 1991.
- [42] H. Proenc and A. Alexandre, “UBIRIS : A noisy iris image database,” *Symp. A Q. J. Mod. Foreign Lit.*, no. 3617, pp. 970–977, 2006.
- [43] T. Deserno, “Medical Image Processing,” in *Optipedia*, SPIE Press, Bellingham, WA, 2009. [Online]. Available: <http://spie.org/publications/optipedia-pages/submitted-content/deserno-medical-image-processing>.
- [44] M. Andrew Lee, “MedicalEye,” Department of Ophthalmology at the University of Iowa, 2005. [Online]. Available: <http://webeye.ophth.uiowa.edu/eyeforum/atlas/pages/Arcus/index.htm>.
- [45] T. Harleen, S. Harbir, and B. Ruby, “Eyecare Center,” 2014. [Online]. Available: <http://www.highstreeteyecare.ca/blog/2014/2/26/the-eye-heart-connection>. [Accessed: 20-Jun-2015].
- [46] Keepyourhearthealthy, “Keep your heart healthy,” 2010. [Online]. Available: <https://keepyourhearthealthy.wordpress.com/2010/05/11/arcus-senilis-its-probably-not-what-you-think/>.
- [47] “MansMed Media Center,” 2002. [Online]. Available: [http://media.mansmed.com/details.php?image\\_id=74](http://media.mansmed.com/details.php?image_id=74).
- [48] “Multidimensional Iridology & Flower Essences,” 2008. [Online]. Available: <http://sivasanta.blogspot.my/2008/08/cholesterol-ring.html>.
- [49] I. H. James, “Herbal Clinic,” 2002. [Online]. Available: <http://www.iridology-swanssea.co.uk/patient-cases/patient-4/>.