

Received June 26, 2020, accepted July 23, 2020, date of publication July 28, 2020, date of current version August 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3012583

Keratoconus Detection Using the Fusion Features of Anterior and Lateral Segment Photographed Images

MARIZUANA M. DAUD¹, W. MIMI DIYANA W. ZAKI², (Member, IEEE),

AINI HUSSAIN², (Member, IEEE), AND HALIZA ABDUL MUTALIB³

¹MIMOS Berhad, Technology Park Malaysia, Kuala Lumpur 57000, Malaysia

²Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia

³Optometry and Vision Sciences Program, School of Healthcare Sciences, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia

Corresponding author: W. Mimi Diyana W. Zaki (wmdiyana@ukm.edu.my)

This work was supported in part by the Research University Grant from Universiti Kebangsaan Malaysia under Grant DIP-2018-020, and in part by the Ministry of Higher Education, Malaysia, under Grant FRGS/1/2016/ICT01/UKM/02/4.

ABSTRACT Keratoconus (KC) is a type of eye disease that involves the thinning of the corneal layer and a change in the semispherical shape of the cornea to a bulging cone shape when viewed laterally. KC is difficult to detect in the early stages of the disease, as the patient does not feel any pain. Hence, the development of a KC detection (KD) method using a digital image processing approach is needed for the early detection of KC so that physicians can provide patients with the subsequent treatment sooner. The objective of this study was to develop a method of KD using a camera from a smart device to capture anterior and lateral segment photographed eye images (A&LSPIs). A total of 280 images comprising 140 KC and 140 normal A&LSPIs were used in this study, and all images were validated by a qualified expert. First, the corneal area of both image views was segmented so the geometric features could be extracted using the automated modified active contour model and the semiautomated spline function for the anterior and lateral images, respectively. Then, the features were selected using infinite latent feature selection (ILFS) by identifying the feature rankings based on the graph weighting that was automatically learned by the probabilistic latent semantic analysis (PLSA). The results showed that the all-combined features, where the proposed and improved features were successfully top ranked, had 96.05% accuracy, 98.41% sensitivity and 93.65% specificity with the $RF_{n=50}$ classifier, outperforming the $7\text{-NN}_{\text{Maha}}$ and SVM_{RBF} classifiers. In conclusion, this study successfully developed a keratoconus detection system based on fusion features using a digital image processing approach for A&LSPIs captured with a smartphone camera.

INDEX TERMS Keratoconus, lateral segment photographed images, anterior segment photographed images, fusion features, digital image processing.

I. INTRODUCTION

Keratoconus (KC) is a non-inflammatory corneal disorder that is associated with corneal steepening and thinning and results in a corneal bulge [1], [2], as shown in Fig. 1. Corneal thinning induces irregular astigmatism with visual distortion and myopia, leading to mild to moderate visual impairment [3], but KC does not lead to blindness.

The aetiology of KC is heterogeneous and multifactorial and is caused by a combination of genetic and environmental

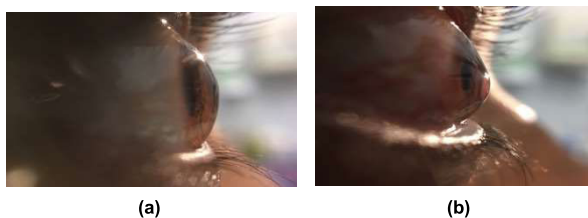
factors. The exact contributions to the aetiology remain unresolved. Factors known to have a positive association with KC include excessive eye rubbing, atopy, eczema, Down syndrome and parental consanguinity [4]. Genetic factors may also be linked to ethnicity; for example, India, Pakistan and Saudi Arabia has been identified as having high prevalence of KC.

This is influenced by the practice of marriage with blood relatives in India and Pakistan [20], [21]. In addition, geographical factors, such as the hot climate and exposure to long-lasting ultraviolet light in Asian countries, contribute to higher prevalence of KC than in colder climates [22], [23].

The associate editor coordinating the review of this manuscript and approving it for publication was Jenny Mahoney.

TABLE 1. Prevalence rate of KC.

Authors (years)	Country	Prevalence rate (%)	Sample size (case)	Men	Women	Prevalence per 100,000
Assiri et al. (2005) [5]	Asir, Saudi Arabia	0.02	654,163 (125)	-	-	20
Nielsen et al. (2007) [6]	Denmark	0.09	5,326,404 (4,600)	-	-	86
Jonas et al. (2009) [7]	Maharashtra, India	2.74	4,667 (128)	-	-	2,743
Agrawal & Vision (2011) [8]	India	-	274	189	85	-
Mohd-Ali et al. (2012) [9]	Malaysia	1.20	13,000 (159)	113	46	1,223
Xu et al. (2012) [10]	Beijing	0.90	3,166 (27)	-	-	853
Mohd-ali et al. (2015) [11]	Malaysia	30.20	629 (190)	104	86	30,207
Shehadeh et al. (2015) [12]	Palestine	1.50	620 (9)	-	-	1,452
Kosker et al. (2016) [13]	Turkey	4	100 (4)	-	4	4,000
Abdu et al. (2016) [14]	Sudan, Saudi arabia	-	208	116	92	-
Althomali et al. (2017) [15]	Taif, Saudi Arabia	8.59	687 (59)	31	28	8,588
Almeida et al. (2018) [16]	Riyadh, Saudi Arabia	4.79	522 (25)	-	-	4,790
Al-amri (2018) [17]	Abha, Saudi Arabia	24	2280 (548)	-	-	24,035
Hashemi et al. (2018) [18]	Iran	4	2,268 (100)	-	-	4,409
Hwang et al. (2018) [19]	South korea	0.04	47,990,761 (17,931)	9,155	8,776	37.4
Sashia et al. (2019) [53]	Denmark	0.003	5,657,361 (183)	-	-	3.23

**FIGURE 1. Pathological changes of (a) normal and (b) KC.**

This is also evidenced by the reduction of aldehyde dehydrogenase (ALDH) and the increased aldehydes that occur in the cornea, which will cause severe oxidative stress and, subsequently, irreversible damage, also known as apoptosis [24].

There have been very few studies on the epidemiology of KC in Asia, especially in Malaysia. This is due to the previous focus of the National Eye Study of 1996 [25] on more prevalent diseases, such as diabetic retinopathy and cataracts, which are the main causes of blindness in Malaysia. The first prevalent study of KC in Malaysia was reported by Mohd-Ali et al. [9] in 2012; the study included 13,000 patients, 159 of which were diagnosed with KC, indicating a prevalence rate of 1.2% (1,223 per 100,000 individuals). The majority of patients were Indian (39%), followed by Malay (37%) and Chinese (19%). In 2015, Mohd-Ali et al. [11] once again reported on KC, now noting an increase in the number of patients, with an estimated prevalence rate of 30.2%, or 30,207 per 100,000 individuals.

KC has been reported in several countries with varying prevalence rates, as shown in Table 1. Most of the studies reported were from Middle Eastern countries such as Saudi

Arabia, Turkey and Palestine, with prevalence rates of 37.4%, 4% and 1.5%, respectively. Starting in 2016, KC prevalence studies have been conducted in many Saudi Arabian regions, such as Riyadh, Abha and Taif, with prevalence rates of 4,790, 24,035 and 8,588 per 100,000 individuals, respectively.

KC is diagnosed by clinical signs such as Fleischer rings, Rizzuti signs, Vogt striae, Munson signs and the retinoscopy scissoring reflex, which are examined using slit lamp examination [2]. However, the most sensitive methods for the early detection of KC are corneal topography and the Scheimpflug principle. Although these imaging tools can be used to detect KC, they have high purchase and maintenance costs. Additionally, these tools are quite large, immobile and heavy, and they need to be controlled by well-trained and experienced optometrists. At the same time, optometrists should also refer to other clinical signs [20] as a guide in diagnosing KC. Hence, with the aid of smart technology, digital image processing and machine learning can fill the role of easy-to-use imaging, especially in rural areas. Furthermore, mobile applications are evidently the current niche in healthcare, even for the screening and analysis of ocular diseases. On this basis, this paper proposed a method of KD by analysing the visual characteristics of eyes with KC; these characteristics were extracted from segmentation of anterior and lateral segment photographed eye images (A&LSPI). According to our collaborated experts, the geometrical features of both the anterior and lateral views may provide significant information about keratoconic eyes. These features were selected and ranked according to their relevance to KC and normal eyes.



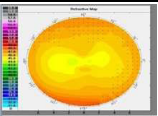


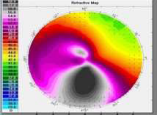


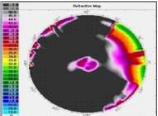
The rest of the paper is outlined as follows. Section 2 reviews the current methods for ocular disease detection using image processing integrated with machine learning. Section 3 provides a detailed explanation of the KD algorithm, which is based on two phases: ASPI and LSPI. Then, the ASPI, LSPI and A&LSPI features are tested to classify the images into KC or normal groups in Section 4. In the last section, the conclusions are presented.

II. KERATOCONUS DETECTION USING A DIGITAL IMAGE PROCESSING APPROACH

To the best of our knowledge, there are only a few related studies on detecting or diagnosing KC using smartphones integrated with machine learning and digital image processing approaches. Most studies conducted were based on images from diagnostic tools such as corneal topography using the topographic map information, where the map images were processed using image processing or machine learning methods. Ali *et al.* [26] classified topographical map images from Pentacam using ANNs, SVMs and decision trees to determine whether an eye was normal or exhibited KC. Arbelaez *et al.* [27] implemented an SVM classifier for identifying KC in eyes using a combination of topographical map images from corneal topography and data from a Scheimpflug camera. Dhaini *et al.* [28] proposed a new automated KD that employs image processing and machine learning methods to detect and measure corneal haze and demarcation lines in OCT images. Most recently, a study was conducted to mimic the concept of Placido's disk. Nabhan [29] patented a system and method for optical imaging using mobile processing devices. The author developed a device that mimics a topographical concept by using cone frustum segments, optimization lenses and vertex-angulation positioning lights. The system also uses image processing and computer vision techniques to extract the corneal feature information from topographical maps that have been captured using smartphones.

Recently, a study was published that used smartphones to capture an eye image and applied digital image processing methods as a tool to help in KD. Askarian *et al.* [30] used smartphones to capture the lateral eye images of 10 participants; 20 eye images were captured, of which 6 were diagnosed as KC and 14 were diagnosed as normal. The study suggested two main steps: preprocessing and classifying images as normal or KC. First, the image is cropped to obtain only the cornea area, and it is rotated so that the corneal curve has a 90-degree horizontal angle. Then, Canny's edge detection technique is applied to detect the corneal curve, and a morphological dilation technique is performed to obtain smooth corneal boundaries. The classification of the cornea is based on the slope of the corneal curvature by dividing the corneal curve into N different points, with each point sampled every 10 degrees. An irregular slope angle is classified as KC, whereas the normal cornea has a round curvature and regular slope variation [30]. The results show that the proposed method is able to detect severe, advanced and moderate

TABLE 2. Example of images for subjects (Sub) normal and KC patients.

Sub	Front eye images	Lateral eye images	Topography map
A (Normal)			
B (KC)			
C (KC)			

KC with accuracies of 93%, 86% and 67%, respectively. To the best of our knowledge, this is the first study to detect KC using images from a smartphone; this study is the most similar to the study proposed in this paper. However, our study was conducted using the combination of 140 ASPIs and 140 LSPIs of KC and normal images that were validated by an experienced optometrist using corneal topography.

III. PROPOSED KERATOCONUS DETECTION USING A&LSPI

Briefly, the proposed KD study consists of three main phases: collection of the A&LSPI data, processing of the ASPIs and LSPIs, and feature selection with classification evaluation, as shown in Fig. 2.

A. A&LSPI DATABASE DEVELOPMENT

The development of the A&LSPI database was conducted at the Ophthalmology Department in Hospital Kuala Lumpur (HKL), Malaysia. The collection of A&LSPI was conducted over a 6-month period from December 2017 to June 2018. All KC cases that were referred to the Ophthalmology Department were first examined by ophthalmologists. The confirmed cases that only presented with KC were randomly filtered and passed to us. For normal cases, healthy subjects were invited to the HKL for screening and validated as normal subjects. The eye images were captured from the anterior and lateral eye views using smartphones.

This data collection process was set up at the Ophthalmology Department in Hospital Kuala Lumpur, Malaysia, with an appointed optometrist to validate the eye images. The topographic map of each subject's eye image was obtained using an OCULUS Easygraph Topographer machine, as shown in Fig. 3 (b). Topographic maps were used by the expert to validate both the keratoconus and normal cases. Table 2 shows examples of anterior and lateral images captured using Huawei P9 smartphones with a 20-megapixel camera and the topography maps of the eyes of KC patients and normal subjects. Subject A represents a normal case of A&LSPIs, whereas subjects B and C are examples of KC patients with different degrees of KC severity. As shown in Table 2, the

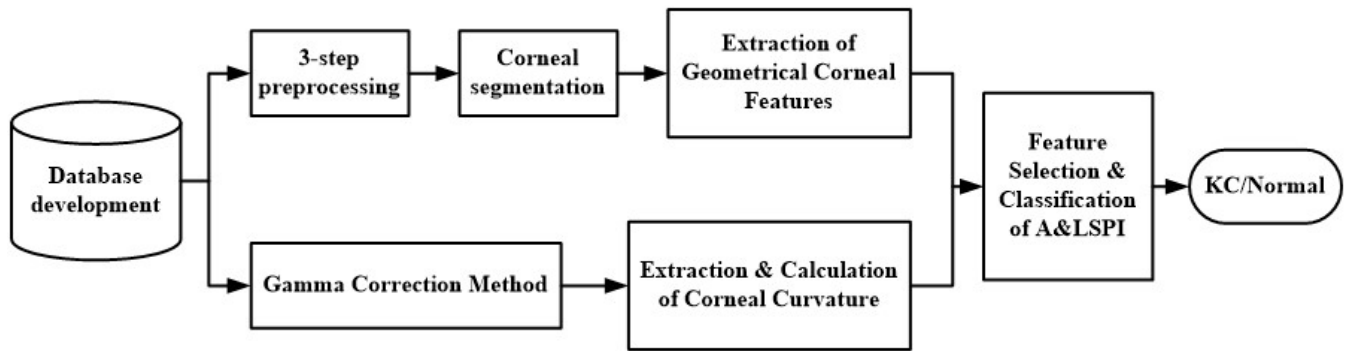


FIGURE 2. Block diagram of proposed the keratoconus disease detection system.

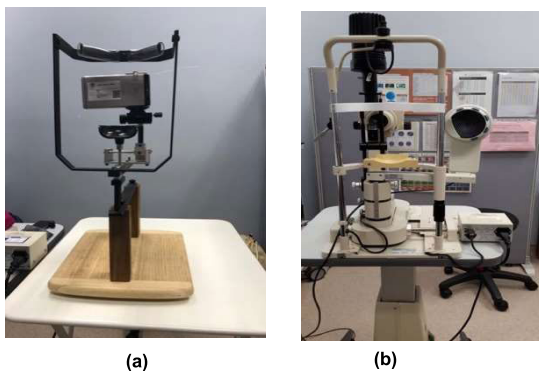


FIGURE 3. Data collection was performed using (a) a chin rest and (b) an Easygraph corneal topographer.

topographic map of each subject differs according to its KC status. In addition, the topography information comprises a colour map that represents the steepness of the cornea curvature. The yellow and green denote a ‘normal’ condition, red and white denote a ‘steeper than normal’ condition, and blue and violet denote a ‘flatter than normal’ condition.

Handling the corneal topographer is quite difficult when obtaining images from a subject with severe KC. This is because it is difficult for patients with KC to keep their eyes open throughout the test duration. Patients with KC tend to blink often, and their eyes easily become watery and red. In the case of subjects with KC, the cornea transparency tends to decrease due to deformation of the cornea shape, which causes light to be scattered. Thus, it is difficult to map the corneal topology. This is evident, as shown in the topographic maps of subjects B and C in Table 2. After the data cleaning process, only 140 KC images and 140 normal A&LSPI images were used in this study. This is because some of the captured images were blurry and unclear, however, images with noises such as eyelashes, eyelid, skin color, undefined corneal edge and reflection is included in this work, as shown in Fig. 4. Some of the rejected images were from patients who had undergone cross-linking procedures or from patients with injured corneas. All images were processed using an Intel (R) Core i7-6700 processor with MATLAB®(licence number: 40699855).



FIGURE 4. Sample images containing noise such as refractive, reflection, eyelid, and eyelashes.

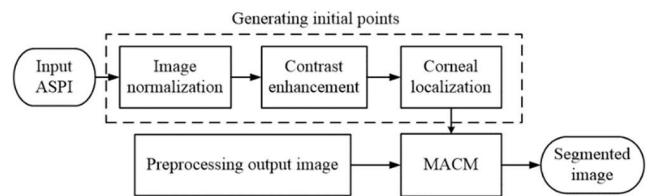


FIGURE 5. The proposed 3-step preprocessing approach for ASPI.

B. PROCESSING OF ASPI

The processing of the ASPIs consists of three modules: (1) preprocessing, (2) segmentation and (3) feature extraction.

1) 3-STEP PREPROCESSING APPROACH

Preprocessing is one of the most important steps for improving the quality of images. Fig. 5 shows the pipeline of the 3-step preprocessing approach for ASPIs involving a) RGB to HSV conversion, b) edge enhancement and c) image smoothing. This 3-step approach is crucial in preparing the ASPIs for the segmentation process, which requires the noise artefacts to be suppressed and the non-region of interest areas, such as the eye lashes and skin colour, to be blurred beforehand while highlighting the edge of the cornea.

First, the RGB input images are converted into HSV colour channels. The HSV colour space has been chosen because HSV is able to discriminate the colour and intensity information [31] and is easier to convert compared to the RGB colour space. After conducting several tests, only the value (V) component is included in the process of increasing the corneal edge because it produces an output that has the best

contrast between the cornea and the sclera. However, the hue (H) and saturation (S) channel images are unusable for maintaining the original corneal structure.

A Gaussian filter with $\sigma = 0.01$ is applied on the value component, $V(x, y)$, to produce an image $I(x, y)$, followed by the image smoothing process using the anisotropic diffusion (AD) of Perona and Malik [32]. This involves the convolution operations of the AD technique to highlight the cornea edges. This convolutional technique was adapted from the study conducted by Kaur and Kaur [33]. Fig. 6 shows the visual results of the proposed 3-step preprocessing approach.

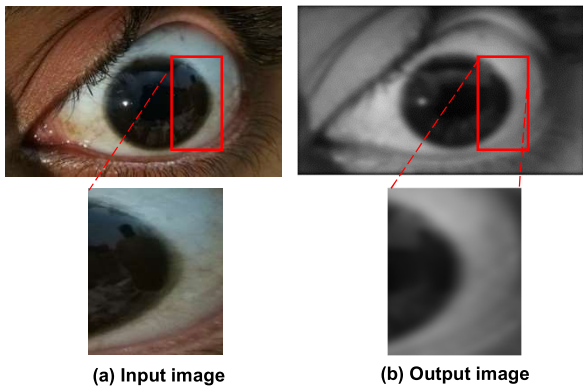


FIGURE 6. Example of the 3-step preprocessing of an output image.

2) CORNEAL SEGMENTATION USING A MODIFIED ACTIVE CONTOUR MODEL

The cornea is segmented so that its features can be extracted. Therefore, this step is important to produce quantitative calculations that will affect the entire developed detection method. The proposed method for segmenting the cornea is the modified active contour model (MACM), as shown in Fig. 7, which has been improved from the active contour model [33], which generates the initial curvature point manually. In this study, the initial curvature point is automatically generated using the estimation of the location of the corneal texture.



FIGURE 7. A block diagram of the proposed MACM.

Image normalization is performed using a single-scale retinex technique [34] for normalizing the eye image illumination; the technique is able to improve the colour consistency under severe illumination variance. Then, relative total variation [35] is adopted to measure and regularize the local pixel variation; this approach offers a resilient ability to discriminate the key image features from the background images. This method enhances the image contrast structure component and minimizes interference from noise. Finally, the corneal region can be localized, and the initial points

can be generated around the corneal region. This improvised method was proposed by Zhao and Kumar [36] to extract the iris structure and has been applied in this work. This method is important, as this step is crucial to localize the corneal region to automate the proposed segmentation method, as detailed in [37].

TABLE 3. Dice similarity index for the segmentation results.

Segmentation methods	P ₁	P ₂	Average Dice index
MACM	0.8607	0.8643	0.8625
CV	0.6319	0.4947	0.5633
CHT	0.5358	0.5394	0.5376
DG	0.7615	0.7617	0.7616

Table 3 tabulates the segmentation results of the four methods, namely, the proposed MACM, Chanvese (CV), circular Hough transform (CHT) and Daugman (DG), using the Dice similarity index measures. Two experts, P₁ and P₂, were assigned to perform the manual segmentation to establish the ground truth data so that a quantitative comparison can be made. From the table, the proposed MACM method obtained the highest average Dice similarity index of 0.8625 compared to the other three methods, which had indexes of 0.5633 for CV, 0.5376 for CHT and 0.7616 for DG. In addition, the segmentation results of both P₁ and P₂ are generally in agreement.

3) GEOMETRICAL CORNEAL FEATURE EXTRACTION

In this study, geometric shapes are the significant features that are extracted from the segmented corneal regions. Sometimes, imaging tools do not provide sufficient assistance during an eye examination; therefore, optometrists and ophthalmologists perform manual examinations based on clinical signs through the observation of the cone shape of the KC cornea. This practice shows that the shape features are important and useful in identifying any changes of the corneal shape in ASPIs. Additionally, features such as eccentricity and solidity can be measured based on a study conducted on the ASPIs of pterygium as reported by Wan Zaki et al. [38]. These features can be used to assess the changes in the corneal shape of the human eye by means of ASPIs.

Seven features representing the shape of the cornea, including the horizontal visible iris diameter (HVID), vertical visible iris diameter (VVID), eccentricity, asphericity, major line, minor lines, and orientation, have previously been extracted in other studies. HVID is the distance between the nasal and temporal imaginary limbal tangent to the corneal circumference; the VVID is calculated similarly but in the vertical orientation [39]. Eccentricity is a parameter that characterizes the course and shape of the cornea [40]. It is measured in terms of the shape factor, sf , which is mathematically expressed in (1) and (2).

$$ecc = \sqrt{1 - sf} \tag{1}$$

$$sf = b^2 / a^2 \tag{2}$$

where a represents the HVID and b represents the VVID. The sf varies between 0 and 1, where 1 is a perfect sphere.

The asphericity of the anterior corneal surface shows how the shape of the cornea differs from a sphere when measured from the centre of the cornea to the periphery, which is mathematically shown in (3). The average asphericity, Q , of the human cornea is between -0.26 and -0.42 , while the average ecc value is between 0.4 and 0.6 [40]. These values can be obtained using corneal topography machines that are commonly used during eye examinations.

$$Q = sf - 1 \tag{3}$$

For an aspherical surface, $Q = 0$ if the surface is a spheroid, $-1 < Q < 0$ if the surface is prolate, and $Q > 0$ if it is oblate. The major line is the line along the longest diameter, and the minor line is the line along the shortest diameter passing through a midpoint. The orientation, major line and minor line are calculated using central moments that provide useful descriptors of shape.

C. PROCESSING OF LSPI

An LSPI is a photographed image that is captured using a smartphone when the eye is viewed laterally. The processing of LSPIs consists of two modules: preprocessing using gamma correction and feature extraction. Clinically, the KC cornea bulges outward in either the straight or downwards direction. Therefore, the corneal protrusion can be identified visually, and the corneal curvature can be measured quantitatively.

1) GAMMA CORRECTION METHOD

The purpose of the preprocessing step of the LSPIs is to produce clearer and sharper pixels of the corneal curvature. The reflection from a 45-degree pen torch light could enhance the brightness of the corneal curvature. In most keratoconic or unhealthy eyes, there is weak contrast between the cornea and sclera because the sclera is yellowish or reddish. Therefore, gamma correction is applied to the RGB image to correct the input image brightness. This step is important to extract the accurate corneal curvature. The gamma correction equation is given by (4).

$$I' = 255 \times \left(\frac{I}{255}\right)^{1/\gamma} \tag{4}$$

where γ is the gamma value and I' is the gamma-corrected output image. The brightness of the image intensity increases when $\gamma \geq 1$ and decreases when $\gamma < 1$. In this study, the gamma value is set as $\gamma = 1.5$. The determination of this value is based on the comparison of image findings through trial-and-error.

2) EXTRACTION AND CALCULATION OF CORNEAL CURVATURE

The feature extraction of the lateral eye view is performed by extracting and then calculating the corneal curvature.

The corneal curvature is extracted using a semi-automated approach, the spline linear interpolation technique, which interpolates the data between points using linear functions to estimate polynomials [41]. The technique is based on the fact that accuracy and sensitivity of the corneal curvature are critical since corneal curves carry important information that can distinguish between keratoconic eyes and normal eyes.

Because the chin rest used during the data collection process is not adjustable, the position of the patient's chin is fixed, as shown in Fig. 3 (a). This results in an inconsistency of corneal curvature due to variations in the sizes of the subjects' faces. Therefore, correction of the corneal curvature should be performed by detecting the first point, t_1 , and the last point, t_2 , at the curvature of the cornea, as indicated in Fig. 8.

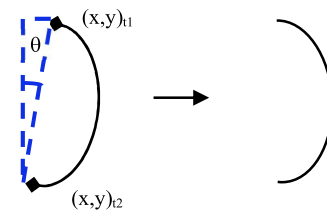


FIGURE 8. Correction of corneal curvature.

The difference in the corneal position is calculated by finding the trigonometric angle, θ , at points t_1 and t_2 as in (5). After determining θ , the new coordinates for t_1 and t_2 can be determined as x' and y' from (6).

$$\theta = \tan^{-1} \left(\frac{x_{t1} - x_{t2}}{y_{t1} - y_{t2}} \right) \tag{5}$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{6}$$

Next, this work proposes three distinct methods to measure the curvature of the cornea to represent the six features of the LSPIs based on a) a ratio calculation (f_{11}), b) a modified template disk (f_{15}, f_{16}, f_{17}) and c) the trapezoidal rule (f_{18}, f_{19}). For comparison purposes, the features extracted from the original template disk method introduced by Aghamohamadian-Sharbat *et al.* [42] will also be discussed.

a: CALCULATION OF RATIO

The normalization process is initiated by finding the extreme point of the corneal curve before the reference ratio, R_R , is calculated. R_R is the ratio between the reference width, R_w , and the image width, I_w , as in (7). Here, R_w was empirically set to 500. I_n is the new value of the curve based on R_w , as shown in (8).

$$R_R = R_w / I_w \tag{7}$$

$$I_n = R_R \times I_H \tag{8}$$

n is the ratio between the width, l , and the length, p , of the corneal curve, as shown in Fig. 9, with l and p equal to R_w and I_n , respectively, in (9).

$$f_{11} = I_n / R_w \tag{9}$$

The ratio method, f_{i1} , is proposed based on the notion that the greater the ratio is, the more convex the eye cornea. This is in accordance with the clinical fact that the convexity of a KC cornea is directly proportional to the severity of the KC [43].



FIGURE 9. The ratio method was measured using the length and width of the curve.

b: MODIFIED TEMPLATE DISC

This method introduces a nonlinear technique, c_{nl} , developed by Inge et al. [44]. However, Aghamohamadian-Sharbat et al. [42] proposed a crossover point technique, c_{cp} , and a trigonometric identification technique, c_{tri} , to improve c_{nl} for the measurement of the retinal blood vessel tortuosity. c_{nl} is introduced to reduce the complexity of the curvature calculation; however, c_{nl} is inaccurate when estimating high curvature. The c_{nl} and c_{cp} techniques are based on the overlapping area between the curve function and a template disk; c_{tri} is based on the angle between the corneal curve and the template disk. The purpose of this method is to relate the area between the curve and the template disk of an appropriate radius and curvature. Therefore, the features that are extracted are the nonlinear feature, $f_{i2} = c_{nl}$; the crossover point feature, $f_{i3} = c_{cp}$; and the trigonometric feature, $f_{i4} = c_{tri}$.

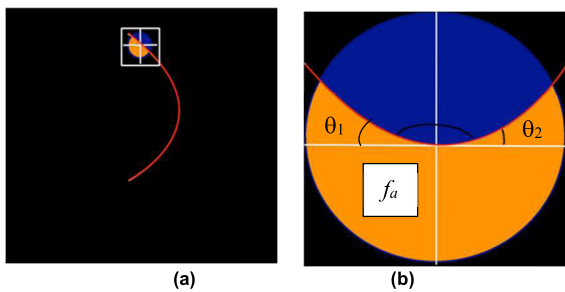


FIGURE 10. (a) Template disk method presented by Aghamohamadian-Sharbat et al. [35] and (b) the proposed modified template disk.

This original method uses a kernel window to calculate the extracted local corneal curvature. The kernel window has been improved in this work by proposing global curvature calculations, that is, using a large radius to cover the whole corneal curvature. This is because small radii and area calculations may contain significant errors [38]. The parameters of the normalized image, $I_N(R_w, N_H)$, calculated from (7) and (8), are used for the global calculation technique. Feature $f_{i5} = f_a$ is the area coloured orange in Fig. 10, while f_{i6} is the inverse square of the area, which is coloured blue. However,

feature f_{i6} is not used in this paper because f_{i6} is just the inverse of f_a .

The second feature, f_{i6} , is the intersection angle between the curve function and the template disk on one side of the curve, and the basis of f_{i7} is the inner angle of the curve, θ_M . This means that f_{i6} considers the corneal curve to be symmetric, whereas f_{i7} considers the KC corneal curve to be asymmetric. Features f_{i6} and f_{i7} are described in (10) to (14).

$$\theta_d = |\theta_1 - \theta_2| \tag{10}$$

$$\theta_M = \pi - \theta_d \tag{11}$$

$$\theta_t = \frac{\theta_M}{2} \tag{12}$$

$$f_{i6} = \frac{2\sin\theta_t}{\cos^2\theta_t} \tag{13}$$

$$f_{i7} = \frac{2\sin\theta_M}{\cos^2\theta_M} \tag{14}$$

c: TRAPEZOIDAL RULE

The trapezoidal rule, f_{trap} , is a common method used to estimate the area under the curve (AUC). The trapezoidal rule and the Simpson rule have been used to calculate the volume of structures [45]. Slavinic & Cvetkovic [45] state that although there are relatively small numbers of ordinates or isopach contours, both numerical integration methods present better results than the detailed cell-based model. In addition, there are few studies that measure the area under the curve using the trapezoidal technique [46], [47]; thus, it is expected to yield beneficial results for corneal curvature measurements.

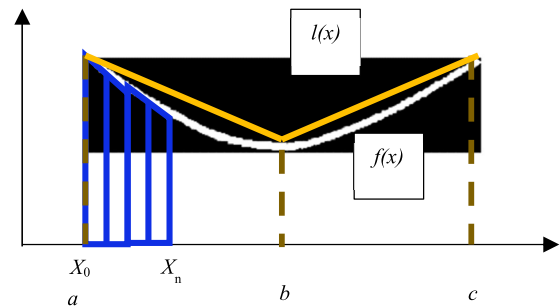


FIGURE 11. Illustration of the trapezoidal rule.

The trapezoidal rule is applied to calculate the corneal curve convexity by identifying the centre point of the curve, $f(x)$, at b , as shown in Fig. 11. Then, the total area under the left curve (as in (15)) and the right curve (as in (16)) are calculated. The difference of the left and right AUC is then measured in (17).

$$\int_c^b f_{irR}(x) = \frac{(b-c)}{2n} [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + 2f(x_n)] \tag{15}$$

$$\int_a^c f_{irL}(x) = \frac{(c-a)}{2n} [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + 2f(x_n)] \tag{16}$$

$$f_{i8} = |f_{irL} - f_{irR}| \tag{17}$$

This approach is in line with the clinical fact that the KC corneal curvature is asymmetrical and droops to the inferior part of the eye, according to the readings from topography maps. Additional feature parameters that can be derived from the trapezoidal rule on the curvature measurement are f_{irR} and f_{irL} , the area under the right and left lines of $l(x)$, respectively.

Equations (18) and (19) represent the results of the trapezoidal rule on $l(x)$, and the absolute difference between the two is shown in (20).

$$f_{irapR} = f_{irR} - f_{irL} \tag{18}$$

$$f_{irapL} = f_{irL} - f_{irR} \tag{19}$$

$$f_{i9} = |f_{irapR} - f_{irapL}| \tag{20}$$

D. FEATURE SELECTION AND CLASSIFICATION OF A&LSPI

ASPI corneal segmentation and LSPI corneal curvature extraction were carried out in the previous modules. The output features were extracted and fused for the purpose of selecting the best features. The feature selection module in the developed KD system helps to evaluate the significance of the proposed corneal and adopted features based on feature ranking. The ASPI corneal features are eccentricity (f_{a1}), asphericity (f_{a2}), major line (f_{a3}), minor line (f_{a4}), orientation (f_{a5}), HVID (f_{a6}), and VVID (f_{a7}), and the LSPI features are the corneal features measured using the ratio calculation (f_{l1}), the original template disk method (f_{l2}, f_{l3}, f_{l4}), the modified template disk method (f_{l5}, f_{l6}, f_{l7}), and the trapezoidal rule technique (f_{l8}, f_{l9}).

The feature selection (FS) techniques are chosen based on several factors, such as stability, scalability and model features. An important aspect in evaluating the stability of FS techniques is feature variability. FS techniques that produce high variability are not suitable for classification purposes. The data size is also contributes to the stability of FS, for example, a small data size with a high-dimensional distribution can cause FA instability. Each feature is weighted and ranked accordingly. However, selecting the optimal number of features to build a feature model is challenging in the FS method. In this work, ReliefF, mutual information feature selection (MutFS) and infinite latent feature selection (ILFS) methods will be used and tested to determine the effectiveness and relevance of the ASPI and LSPI features.

ReliefF is an extended method of the Relief method, which is limited to binary classification and overlooks lost data [48], [49], whereas ReliefF is more robust and can deal with incomplete and noisy data [50]. ReliefF is one of the filter methods used to analyse intrinsic data properties regardless of the type of classifier. ReliefF is similarity-based and is able to identify interactions between features; nevertheless, most ReliefF methods cannot handle feature redundancy and can repeatedly find highly correlated features. However, theory-based filter methods, such as MutFS and ILFS, can be integrated into a probabilistic framework that can consider feature relevancy and feature redundancy. The FS method is also selected based on the suitability of the data size. Wu et al. [51] reported that the embedded random forest (RF)

method appeared to be more suitable with a large number of samples, whereas ReliefF was better for medium-sized data, and the chi-square method outperformed the other methods when the number of samples was small. The ILFS method returns the best results in terms of AUC. However, ILFS is reported to be less stable compared to other techniques as the top- k features increase.

TABLE 4. Comparison of SVM kernel types.

Top- k L/ASPI	Kernel	Acc (%)	Sen (%)	Spe (%)	Err
T _{al5}	Linear	96.43	92.86	100	0.0357
	Quad	89.28	85.71	92.86	0.1072
	Cubic	85.71	78.57	92.86	0.1429
	RBF	100	100	100	0
T _{al9}	Linear	78.57	78.57	85.71	0.2143
	Quad	82.14	78.57	85.71	0.1786
	Cubic	82.14	78.57	85.71	0.1786
	RBF	85.71	78.57	92.86	0.1429
T _{al15}	Linear	96.43	92.86	100	0.0357
	Quad	85.71	78.57	92.86	0.1429
	Cubic	82.14	78.57	85.71	0.1786
	RBF	100	100	100	0
Average					
	Linear	90.48	88.10	95.24	0.0952
	Quad	85.71	80.95	90.48	0.1429
	Cubic	83.33	78.57	88.09	0.1667
	RBF	95.24	92.86	97.62	0.0476

Classification refers to the class-predicting process for the given data. In this study, the classifier will classify keratoconic eyes or normal eyes based on features that have been extracted and ranked. In addition, the classification methods are closely related to the FS methods. Classification performance not only represents the overall performance of the system but also represents the suitability of the data type and data size with each FS method. In this work, the classification analysis includes SVM classifier methods with linear, quadratic and RBF kernels (Table 4), whereas RF considers $n = 10, 50, 100$ and 500 , which are chosen empirically to identify the optimum number of trees (Table 5). k -NN models with suitable numbers for k and Mahalanobis, Chebychev and Euclidean kernels (Table 6 and Table 7) were also tested.

These preliminary experiments are intended to determine the optimum number of features that are appropriate for the optimized classifier because the specific features or properties of the classifier need to be suitable for the size of the dataset, as described by Foster et al. [52]. Accordingly, the classification process is run based on the holdout cross-validation test by holding 70% of the total data as a training data set and using 30% as a random test data set. Then, k -fold cross validation is used to train the training data. For the preliminary experiment, the sets of A&LSPI features used are the top 5, T_{al5} ; top 9, T_{al9} ; and top 15, T_{al15} , of the features ranked $< f_{l5}, f_{l7}, f_{l6}, f_{l1}, f_{l4}, f_{l9}, f_{l8}, f_{a6}, f_{a7}, f_{a4}$,

TABLE 5. Comparison of the number of trees for the RF classifier.

Top- <i>k</i> L/ASPI	Number of tress, <i>n</i>	Acc (%)	Sen (%)	Spe (%)	Err
T _{al5}	10	100.0	100.0	100.0	0
	50	100.0	100.0	100.0	0
	100	100.0	100.0	100.0	0
	500	100.0	100.0	100.0	0
T _{al9}	10	96.43	92.86	100.0	0.0357
	50	92.86	92.86	92.86	0.0714
	100	92.86	92.86	92.86	0.0714
	500	92.86	92.86	92.86	0.0714
T _{al15}	10	85.71	71.43	100.0	0.1429
	50	85.71	78.57	92.86	0.1429
	100	82.14	71.43	92.86	0.1786
	500	85.71	78.57	92.86	0.1429
Average					
	10	94.05	88.10	100.0	0.0595
	50	92.86	90.48	95.24	0.0714
	100	91.67	88.10	95.24	0.0833
	500	92.86	90.48	95.24	0.0714

TABLE 6. Comparison of distance function kernels and number for the K-NN models.

Distance function	Index Metric	<i>k</i> =3 (%)	<i>k</i> =5 (%)	<i>k</i> =7 (%)	<i>k</i> =9 (%)
Euclidean	Acc	85.71	89.29	89.29	96.43
	Sen	71.43	78.57	78.57	92.86
	Spe	100.0	100.0	100.0	100.0
	Err	14.29	10.71	10.71	3.570
Mahalanobis (Maha)	Acc	89.29	96.43	96.43	92.86
	Sen	85.71	92.86	92.86	85.71
	Spe	92.86	100.0	100.0	100.0
Chebychev (Cheby)	Err	10.71	3.570	3.570	7.140
	Acc	92.86	89.29	85.71	92.86
	Sen	85.71	78.57	71.43	85.71
	Spe	100.0	100.0	100.0	100.0
	Err	7.140	10.71	14.29	7.140

$f_{a1}, f_{a2}, f_{i3}, f_{a5}, f_{i2}, f_{a3} >$ to find the optimized parameters for each classifier. The performance of the algorithm is compared using the index metrics of accuracy (Acc), sensitivity (Sen), specificity (Spe) and error rate (Err).

Based on Table 4, the SVM with an RBF kernel shows a 95.24% accuracy. This kernel has consistently outperformed the others in nonlinear data distribution. The testing results of the RF methods, shown in Table 5, indicate that the method that uses $n = 10$ trees yields the highest Acc of 94.05%. However, the Sen is higher when using $n = 50$ trees. Thus, $RF_{n=50}$ will be used in the next experiment; in this study, the Sen is more important than the Acc and Spe because KC patients cannot be misclassified. Table 5 and Table 7 show the best number of nearest neighbours with the best kernel for the k -NN model. The kernel $7\text{-NN}_{\text{Maha}}$ yields the highest values: Acc of 94.04%, Sen of 97.62% and Spe of 90.47%.

TABLE 7. Comparison of the k-optimized distance function kernels for the K-NN models.

Top- <i>k</i> L/ASPI	Kernel with the best <i>k</i>	Acc (%)	Sen (%)	Spe (%)	Err (%)
T _{al5}	Maha _(<i>k</i>=7)	92.86	100.0	85.71	0.0714
	Cheby _(<i>k</i>=9)	78.57	85.71	71.43	0.2143
	Euclidean _(<i>k</i>=9)	89.29	85.71	92.86	0.1071
T _{al9}	Maha _(<i>k</i>=7)	92.86	100.0	85.71	0.0714
	Cheby _(<i>k</i>=9)	71.42	64.29	78.57	0.2858
	Euclidean _(<i>k</i>=9)	82.14	85.71	78.57	0.1786
T _{al15}	Maha _(<i>k</i>=7)	96.40	92.86	100.0	0.0360
	Cheby _(<i>k</i>=9)	64.29	57.14	71.43	0.3571
	Euclidean _(<i>k</i>=9)	75.00	64.29	85.71	0.2500
Average					
	Maha_(<i>k</i>=7)	94.04	97.62	90.47	0.0600
	Cheby _(<i>k</i>=9)	71.43	69.05	73.81	0.2900
	Euclidean _(<i>k</i>=9)	82.14	78.57	85.71	0.1800

IV. PERFORMANCE OF KERATOCONUS DETECTION APPROACH USING FUSION FEATURES

The performance of the KD approach was evaluated based on the preliminary results using the $RF_{n=50}$, SVM_{RBF} , and $7\text{-NN}_{\text{Maha}}$ classifiers. For the preliminary experiment, the tests were performed three times with the same procedures. Table 8 shows the anterior, lateral and anterior-lateral tests indicated by U_{xy} , where $x = \{a, l, al\}$ and $1 \leq y \leq 3$. For the A&LSPI fusion features, ILFS produced five features, $\langle f_{i1}, f_{i5}, f_{i6}, f_{i7}, f_{i9} \rangle$, with repeated ranked positions, whereas ReliefF and MutFS produced just one recurring feature, f_{i5} , as indicated by the red boxes in Table 8. The repetition of features in the same position shows that ILFS is more stable than the other methods.

In addition, the ILFS technique also repeats the same 4 features in the first four positions for all three tests. These recurring positions indicate to domain experts that these algorithms are trustworthy and can be used in other works containing new data samples. The features of the original template disk method did not obtain many votes in the highest ranked positions, with very few votes in the top 3 positions. In addition, the majority of the features from the front view, which were extracted from the ASPIs, are in the bottom positions, especially the features $\langle f_{a1}, f_{a2}, f_{a3}, f_{a4}, f_{a5} \rangle$, which are not among the top k features, where k is the set of the features that have been tested. The comparison between these three techniques clearly shows that ILFS is more stable than the ReliefF and MutFS techniques. However, stability is not the only benchmark or standard that determines the best output feature set. Performance is usually evaluated by the accuracy of classification.

The ILFS technique presents the best classification performance for all three sets of features, ASPI, LSPI and A&LSPI, as shown in Table 9. This could be because ReliefF is more stable with more sample counts (> 200), which would affect the performance of the ReliefF technique in our experiments; this finding is similar to that from Wu *et al.* [51]. Additionally, ILFS is found to have the best and most stable performance suitable for the size of the data and the numbers of features.

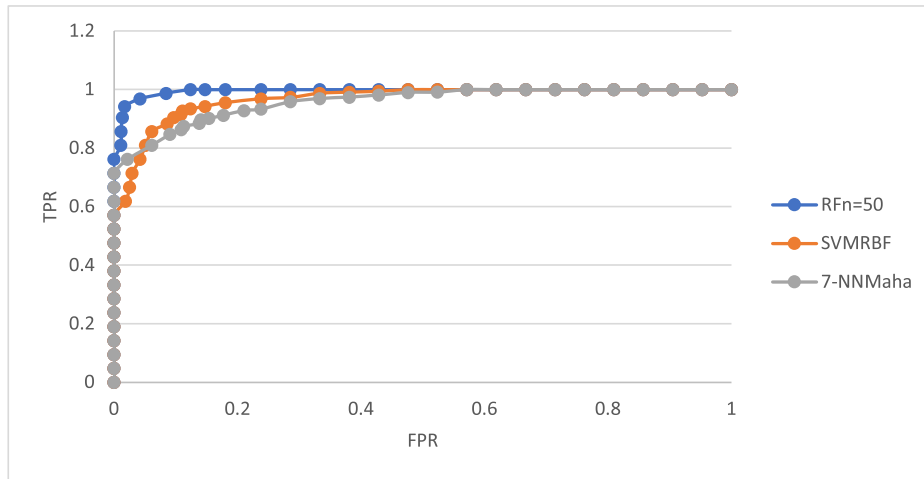


FIGURE 12. ROC curve of the T_{hs16} features set.

TABLE 8. Feature selection performance based on feature ranking.

Tests/ Features	ILFS			ReliefF			MutFS			
	U_{a1}	U_{a2}	U_{a3}	U_{a1}	U_{a2}	U_{h3}	U_{a1}	U_{a2}	U_{a3}	
ASPI	f_{a1}	7	7	7	7	7	7	1	4	4
	f_{a2}	2	2	6	6	6	6	3	3	2
	f_{a3}	1	1	1	3	5	2	4	1	1
	f_{a4}	4	6	2	2	3	3	2	2	3
	f_{a5}	3	3	5	4	2	4	7	7	7
	f_{a6}	6	5	3	1	4	5	6	6	6
	f_{a7}	5	4	4	5	1	1	5	5	5
Test	U_{i1}	U_{i2}	U_{i3}	U_{i1}	U_{i2}	U_{i3}	U_{i1}	U_{i2}	U_{i3}	
LSPI	f_{i1}	5	5	5	3	3	5	6	6	7
	f_{i2}	7	7	7	4	7	3	7	7	8
	f_{i3}	6	6	6	8	4	7	4	8	6
	f_{i4}	1	1	1	7	2	4	8	3	4
	f_{i5}	9	9	9	5	1	8	1	1	3
	f_{i6}	8	8	4	6	6	1	5	4	1
	f_{i7}	2	4	8	1	5	6	3	5	5
	f_{i8}	4	3	3	9	8	2	2	2	2
	f_{i9}	3	2	2	2	9	9	9	9	9
Test	U_{al1}	U_{al2}	U_{al3}	U_{al1}	U_{al2}	U_{al3}	U_{al1}	U_{al2}	U_{al3}	
A&LSPI	f_{a1}	11	10	12	14	13	12	13	11	12
	f_{a2}	12	14	10	11	12	16	16	15	13
	f_{a3}	16	12	9	13	15	13	12	13	15
	f_{a4}	10	7	11	12	14	14	15	16	16
	f_{a5}	14	13	14	15	16	15	9	14	9
	f_{a6}	8	8	5	10	9	9	11	9	10
	f_{a7}	9	11	13	9	11	11	8	8	8
	f_{i1}	4	4	4	3	4	3	6	6	5
	f_{i2}	15	16	15	7	10	8	7	5	6
	f_{i3}	13	15	8	4	2	2	4	2	4
	f_{i4}	5	9	16	5	1	5	5	7	7
	f_{i5}	1	1	1	6	6	6	1	1	1
	f_{i6}	3	3	3	2	5	4	2	3	2
	f_{i7}	2	2	2	1	3	1	3	4	3
	f_{i8}	7	5	7	8	7	7	14	10	11
	f_{i9}	6	6	6	16	8	10	10	12	14

The interesting findings that have been analysed are the number of samples, number of features and signal feature strengths, which are among the elements that affect stability

and FS performance. The MutFS naturally selects features with a univariate approach; thus, the signal extracted is not strong enough. In contrast, ILFS looks for relevancy with the PLSA method utilizing a multivariate approach, causing the extracted signal to be strong. Therefore, the ILFS technique succeeded in selecting the features ranked as follows: $\langle f_{i5}, f_{i7}, f_{i6}, f_{i1}, f_{i4}, f_{i9}, f_{i8}, f_{a6}, f_{a7}, f_{a4}, f_{a1}, f_{a2}, f_{i3}, f_{a5}, f_{i2}, f_{a3} \rangle$; this ranking yielded the highest average accuracies of 95.45%, 88.89% and 88.89% using $RF_{n=50}$, SVM_{RBF} and $7-NN_{Maha}$, respectively, outperforming the ReliefF and MutFS techniques.

However, the optimum feature set must still be tested against combinations of either the ASPI, LSPI or A&LSPI sets and one of the proposed classifiers; the analyses of these combinations are summarized in Table 10. Based on these results, it can be concluded that each classifier has a different suitability with different sets of features. However, all of the classifiers achieved the highest accuracy using the T_{al16} feature set. This is because the fusion features of the top 16 sets are more complex resulting from the combinations of the distinctive types of methods and features.

Using T_{al7} and T_{al15} , the $7-NN_{Maha}$ classifier outperforms $RF_{n=50}$ and SVM_{RBF} , but the accuracy decreases with T_{al9} and T_{al12} . This is due to the nature of the $7-NN_{Maha}$ classifier to search for the feature distance within a neighbourhood, which is measured using a specific distance function. Therefore, attempting to classify KC and normal data with the small distribution ranges of the T_{al9} and T_{al12} feature sets, which contain more ASPI features, could lower the k -NN performance and vice versa. However, the addition of features $\langle f_{i3}, f_{i2}, f_{i3}, f_{a5} \rangle$ can aid the $7-NN_{Maha}$ classifier and improve the classification rate. This condition is proven by the insensitivity of the $7-NN_{Maha}$ classifier; however, the average accuracy is still improved when using the T_{al16} feature set. On the other hand, the $RF_{n=50}$ classifier is built based on the splitting process and would therefore not be well split if it contained overlapping features, which would cause difficulty in determining the threshold value.

TABLE 9. Feature selection performance based on classification accuracy.

Tests	ILFS			ReliefF			MutFS		
	RF _{n=50}	SVM _{RBF}	7-NN _{Maha}	RF _{n=50}	SVM _{RBF}	7-NN _{Maha}	RF _{n=50}	SVM _{RBF}	7-NN _{Maha}
U _{a1}	92.86	83.33	90.48	92.86	78.57	85.71	92.86	88.10	85.71
U _{a2}	95.24	88.10	92.86	85.71	97.62	95.24	85.71	90.48	85.71
U _{a3}	88.10	95.24	88.10	83.33	83.33	76.19	92.86	88.10	85.71
Ave	92.06	88.89	90.48	87.30	86.51	85.71	90.47	88.89	85.71
U _{l1}	92.86	95.24	80.95	83.33	85.71	73.81	85.71	80.95	78.57
U _{l2}	95.24	88.10	88.10	92.86	88.10	83.33	88.10	73.81	88.10
U _{l3}	92.86	80.95	83.33	80.95	83.33	90.48	95.24	85.71	88.10
Ave	93.65	88.10	84.13	85.71	85.71	82.54	89.69	80.16	84.92
U _{al1}	94.24	88.10	90.48	85.71	85.71	78.57	83.33	78.57	76.19
U _{al2}	95.24	90.48	88.10	83.33	85.71	83.33	88.10	78.57	80.95
U _{al3}	92.86	88.10	88.10	95.24	88.10	73.81	90.48	88.10	88.10
Ave	95.45	88.89	88.89	88.09	86.51	78.57	87.30	81.75	81.75

TABLE 10. Feature selection performance.

Features	Top rank	Classifier	Acc (%)	Sen (%)	Spe (%)	Err
ASPI	T _{a3}	RF _{n=50}	61.90	60.32	63.49	0.3810
		SVM _{RBF}	69.05	74.60	63.49	0.3095
		7-NN _{Maha}	70.64	50.79	90.48	0.2936
	T _{a5}	RF _{n=50}	68.25	60.31	76.19	0.3175
		SVM _{RBF}	73.17	74.60	71.43	0.2683
		7-NN _{Maha}	66.02	49.21	82.54	0.3398
	T _{a7}	RF _{n=50}	62.70	66.67	58.73	0.3730
		SVM _{RBF}	70.64	80.95	60.32	0.2936
		7-NN _{Maha}	69.84	49.21	90.48	0.3016
T _{l3}	RF _{n=50}	79.55	80.95	77.78	0.2045	
	SVM _{RBF}	83.33	87.30	79.36	0.1667	
	7-NN _{Maha}	81.74	76.19	87.30	0.1826	
T _{l5}	RF _{n=50}	80.95	79.36	82.54	0.1905	
	SVM _{RBF}	85.72	82.54	88.89	0.1428	
	7-NN _{Maha}	86.51	85.71	87.30	0.1349	
T _{l7}	RF _{n=50}	87.30	77.78	96.83	0.1270	
	SVM _{RBF}	85.71	79.37	92.07	0.1429	
	7-NN _{Maha}	87.30	80.62	93.65	0.1270	
T _{l9}	RF _{n=50}	88.88	84.13	93.66	0.1112	
	SVM _{RBF}	85.70	87.29	84.12	0.1430	
	7-NN _{Maha}	84.93	88.90	80.95	0.1507	
T _{al3}	RF _{n=50}	85.71	95.24	76.19	0.1429	
	SVM _{RBF}	85.71	92.06	79.36	0.1429	
	7-NN _{Maha}	85.71	92.06	79.36	0.1429	
T _{al7}	RF _{n=50}	84.13	82.54	85.71	0.1587	
	SVM _{RBF}	86.51	85.72	87.30	0.1349	
	7-NN _{Maha}	87.30	84.13	90.48	0.1270	
T _{al9}	RF _{n=50}	90.48	87.30	93.65	0.0952	
	SVM _{RBF}	82.54	80.95	84.13	0.1746	
	7-NN _{Maha}	84.92	84.13	85.71	0.1508	
T _{al12}	RF _{n=50}	85.71	74.60	96.83	0.1429	
	SVM _{RBF}	84.13	82.54	85.71	0.1587	
	7-NN _{Maha}	82.54	76.19	88.89	0.1746	
T _{al15}	RF _{n=50}	84.92	80.95	88.89	0.1508	
	SVM _{RBF}	85.71	92.07	79.36	0.1429	
	7-NN _{Maha}	87.30	77.78	96.83	0.1270	
T _{al16}	RF _{n=50}	96.03	98.41	93.65	0.0397	
	SVM _{RBF}	88.09	87.30	88.89	0.1191	
	7-NN _{Maha}	89.69	84.13	95.24	0.1031	

As previously emphasized, accuracy and sensitivity are the most important aspects in determining the performance of the developed KD system. The ROC curve is also used to assess

the performance of each classifier with the TPR and FPR performance. According to Fig. 12, the RF_{n=50} classifier with the top 16 features is able to detect KC with the highest AUC (0.996), whereas SVM_{RBF} yields an AUC of 0.9702, and 7-NN_{Maha} only achieves an AUC of 0.9562. This suggests that the SVM_{RBF} classifier is more sensitive to KD than the 7-NN_{Maha}, but the 7-NN_{Maha} is more sensitive to detecting normal eyes than keratoconic eyes.

Overall, the RF_{n=50} classifier with the top 16 features achieves the best performance when compared to the SVM_{RBF} and 7-NN_{Maha} classifiers. The ROC curve schematic shown in Fig. 12 shows the relationship between TPR and FPR performances, where the ROC curve demonstrates the capabilities of the KD approach in detecting KC, thus determining the effectiveness of the proposed features for normal eyes and keratoconic eyes in clinical situations.

V. CONCLUSION AND FUTURE WORKS

This study has proposed the development of a KD approach using A&LSPIs captured by smartphones. Following a literature review, this study was found to be among the earliest KD methods based on a digital image processing approach. A combination of 3-step preprocessing and the MACM segmentation approach resulted in the successful extraction of features from 140 ASPIs. Additionally, corneal edges were extracted using the spline function approach to measure the features from 140 LSPIs. ILFS techniques were selected based on their stability performance, which was better than of the other FS methods for feature ranking. This technique has succeeded in producing a ranked feature set, $\langle f_{15}, f_{17}, f_{16}, f_{11}, f_{14}, f_{19}, f_{18}, f_{a6}, f_{a7}, f_{a4}, f_{a1}, f_{a2}, f_{a3}, f_{a5}, f_{12}, f_{a3} \rangle$.

The results have shown that the corneal curvature is more distinctly measured using the modified template disk method compared to the original template disk method. The FS performance was measured based on the performance of the classification method. The performances of the RF_{n=50}, SVM_{RBF} and 7-NN_{Maha} classifiers showed a positive correlation between KC disease and the LSPI features. However, the ASPI features also aided in KD classification by yielding fusion feature sets for A&LSPI, which yielded the highest average Acc (96.03%) with the RF_{n=50} classifier, as well

as Sen and Spe values of 98.41% and 93.65%, respectively. In the future, this system can be extended and applied in cloud computing with the implementation of a deep learning approach since the system is based upon a powerful machine learning method. This KD system has good potential to be integrated into computer-based diagnosis systems for KC screening, which would benefit rural areas the most. In addition, the developed KD is easy to use, so clinicians can employ it without intense training.

ACKNOWLEDGMENT

The authors wish to thank our collaborator at the Ophthalmology Department in Hospital Kuala Lumpur, Malaysia, for their kind cooperation and support.

REFERENCES

- Z. Gatziofufas, G. D. Panos, and S. Hamada, "Keratoconus: Is it a non-inflammatory disease?" *Med. Hypothesis, Discov. Innov. Ophthalmol. J.*, vol. 6, no. 1, pp. 6–7, 2017.
- Y. S. Rabinowitz, "Keratoconus major review," *Surv. Ophthalmol.*, vol. 42, no. 4, pp. 297–319, 1998.
- A. Gordon-Shaag, M. Millodot, and E. Shneor, "The epidemiology and etiology of keratoconus," *Int. J. Keratoconus Ectatic Corneal Diseases*, vol. 1, no. 1, pp. 7–15, Apr. 2012.
- A. Gordon-Shaag, M. Millodot, E. Shneor, and Y. Liu, "The genetic and environmental factors for keratoconus," *BioMed Res. Int.*, vol. 2015, pp. 1–19, May 2015.
- A. A. Assiri, B. I. Yousuf, A. J. Quantock, and P. J. Murphy, "Incidence and severity of keratoconus in Asir province, Saudi Arabia," *Br. J. Ophthalmol.*, vol. 89, no. 11, pp. 1403–1406, 2005.
- K. Nielsen, J. Hjortdal, E. Aagaard Nohr, and N. Ehlers, "Incidence and prevalence of keratoconus in Denmark," *Acta Ophthalmologica Scandinavica*, vol. 85, no. 8, pp. 890–892, Jul. 2007.
- J. B. Jonas, V. Nangia, A. Matin, M. Kulkarni, and K. Bhojwani, "Prevalence and associations of Keratoconus in rural Maharashtra in central India: The central India eye and medical study," *Amer. J. Ophthalmology*, vol. 148, no. 5, pp. 760–765, Nov. 2009.
- V. B. Agrawal, "Characteristics of keratoconus patients at a tertiary eye center in India," *J. Ophthalmic Vis. Res.*, vol. 6, no. 2, pp. 87–91, 2011.
- B. Mohd-Ali, M. Abdu, C. Y. Yaw, and N. Mohidin, "Clinical characteristics of keratoconus patients in Malaysia: A review from a cornea specialist centre," *J. Optometry*, vol. 5, no. 1, pp. 38–42, Jan. 2012.
- L. Xu, Y. X. Wang, Y. Guo, Q. S. You, and J. B. Jonas, "Prevalence and associations of steep cornea/keratoconus in greater Beijing. The Beijing eye study," *PLoS ONE*, vol. 7, no. 7, Jul. 2012, Art. no. e39313.
- B. Mohd-Ali, P. Muthiah, S. Retnasabapathy, and N. Mohidin, "Clinical characteristics of keratoconus patients in Malaysia: A 5-year retrospective study from a cornea specialist centre in Selangor (2008–2013)," *Contact Lens Anterior Eye*, vol. 38, p. e25, Feb. 2015.
- M. M. Shehadeh, V. F. Diakonis, S. A. Jalil, R. Younis, J. Qadumi, and L. Al-Labadi, "Prevalence of keratoconus among a palestinian tertiary student population," *Open Ophthalmol. J.*, vol. 9, no. 1, pp. 172–176, Dec. 2015.
- M. Kosker, N. Arslan, M. Y. Alp, C. Ozisler, M. Acar, A. S. Dogan, A. Yesilyurt, and C. Gurdal, "Association between keratoconus and familial mediterranean fever in turkey," *Cornea*, vol. 35, no. 1, pp. 77–80, Jan. 2016.
- M. Abdu, K. H. Binnawi, A. Elaziz, M. Elmadina, and R. Hassan, "Clinical profile of keratoconus patients in Sudan," *Sudan. J. Ophthalmol.*, vol. 8, no. 1, pp. 20–25, 2016.
- T. A. Althomali, I. M. Al-Qurashi, S. M. Al-Thagafi, A. Mohammed, and M. Almalki, "Prevalence of keratoconus among patients seeking laser vision correction in Taif area of Saudi Arabia," *Saudi J. Ophthalmol.*, vol. 32, no. 2, pp. 114–118, Apr. 2018.
- E. A. T. Netto, W. M. Al-Otaibi, N. L. Hafezi, S. Kling, H. M. Al-Farhan, J. B. Randleman, and F. Hafezi, "Prevalence of keratoconus in paediatric patients in Riyadh, Saudi Arabia," *Brit. J. Ophthalmol.*, vol. 102, no. 10, pp. 1–6, 2018.
- A. M. Al-Amri, "Prevalence of keratoconus in a refractive surgery population," *J. Ophthalmol.*, vol. 2018, pp. 1–5, Sep. 2018.
- H. Hashemi, S. Heydarian, A. Yekta, H. Ostadimoghaddam, M. Aghamirsalim, A. Derakhshan, and M. Khabazkhoob, "High prevalence and familial aggregation of keratoconus in an Iranian rural population: A population-based study," *Ophthalmic Physiol. Opt.*, vol. 38, no. 4, pp. 447–455, 2018.
- S. Hwang, D. H. Lim, and T.-Y. Chung, "Prevalence and incidence of keratoconus in South Korea: A nationwide population-based study," *Amer. J. Ophthalmol.*, vol. 192, pp. 56–64, Aug. 2018.
- Y. O. Kok, G. F. L. Tan, and S. C. Loon, "Review: Keratoconus in Asia," *Cornea*, vol. 31, no. 5, pp. 581–593, 2012.
- T. Georgiou, C. L. Funnell, A. Cassels-Brown, and R. O'Conor, "Influence of ethnic origin on the incidence of keratoconus and associated atopic disease in Asians and white patients," *Eye*, vol. 18, no. 4, pp. 379–383, Apr. 2004.
- V. Galvis, T. Sherwin, A. Tello, J. Merayo, R. Barrera, and A. Acera, "Keratoconus: An inflammatory disorder?" *Eye*, vol. 29, no. 7, pp. 843–859, Jul. 2015.
- R. Sharif, S. Bak-Nielsen, J. Hjortdal, and D. Karamichos, "Pathogenesis of keratoconus: The intriguing therapeutic potential of prolactin-inducible protein," *Prog. Retinal Eye Res.*, vol. 67, pp. 150–167, Nov. 2018.
- M. L. Khaled, I. Helwa, M. Drewry, M. Seremwe, A. Estes, and Y. Liu, "Molecular and histopathological changes associated with keratoconus," *BioMed Res. Int.*, vol. 2017, pp. 1–16, 2017.
- M. Zainal, "Prevalence of blindness and low vision in Malaysian population: Results from the national eye survey 1996," *Brit. J. Ophthalmol.*, vol. 86, no. 9, pp. 951–956, Sep. 2002.
- A. H. Ali, N. H. Ghaeb, and Z. M. Musa, "Support vector machine for keratoconus detection by using topographic maps with the help of image processing techniques," *IOSR J. Pharmacy Biol. Sci.*, vol. 12, no. 6, pp. 50–58, 2017.
- M. C. Arbelaz, F. Versaci, G. Vestri, P. Barboni, and G. Savini, "Use of a support vector machine for keratoconus and subclinical keratoconus detection by topographic and tomographic data," *Ophthalmology*, vol. 119, no. 11, pp. 2231–2238, Nov. 2012.
- A. R. Dhaini, M. Chokr, S. M. El-Oud, M. A. Fattah, and S. Awwad, "Automated detection and measurement of corneal haze and demarcation line in spectral-domain optical coherence tomography images," *IEEE Access*, vol. 6, pp. 3977–3991, 2018.
- T. I. Nabhan, "System and method for ophthalmological imaging adapted to a mobile processing device," U.S. Patent 10 129 450, Nov. 13, 2018.
- B. Askarian, T. Fatemehsadat, A. Amin, and J. W. Chong, "An affordable and easy-to-use diagnostic method for keratoconus detection using a smartphone," *Proc. SPIE*, vol. 10575, Feb. 2018, Art. no. 1057512.
- F. H. Seitner and B. C. Lovell, "Pedestrian tracking based on colour and spatial information," in *Proc. Digit. Image Comput., Techn. Appl. (DICTA)*, Dec. 2005, p. 7.
- P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
- I. Kaur and A. Kaur, "Modified active contour snake model for image segmentation using anisotropic filtering," *Int. Res. J. Eng. Technol.*, vol. 3, no. 5, pp. 2151–2157, 2016.
- D. H. Brainard and B. A. Wandell, "Analysis of the retinex theory of color vision," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 3, no. 10, p. 1651, 1986.
- L. Xu, Q. Yan, Y. Xia, and J. Jia, "Structure extraction from texture via relative total variation," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 1–10, Nov. 2012.
- Z. Zhao and A. Kumar, "An accurate iris segmentation framework under relaxed imaging constraints using total variation model," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3828–3836.
- M. M. Daud, W. M. Diyana, W. Zaki, A. Hussain, H. Mutalib, M. H. M. Saad, and A. Razak, "Automated corneal segmentation of anterior segment photographed images using centroid-based active contour model," *Procedia Comput. Sci.*, vol. 163, pp. 330–337, Jan. 2019.
- W. M. D. Wan Zaki, M. Mat Daud, S. R. Abdani, A. Hussain, and H. A. Mutalib, "Automated pterygium detection method of anterior segment photographed images," *Comput. Methods Programs Biomed.*, vol. 154, pp. 71–78, Feb. 2018.
- K. P. Mashige, "A review of corneal diameter, curvature and thickness values and influencing factors," *South African Optom.*, vol. 72, no. 4, pp. 185–194, Dec. 2013.

- [40] P. Benes, S. Synek, and S. Petrová, "Corneal shape and eccentricity in population," *Coll. Antropol.*, vol. 37, no. 1, pp. 117–120, 2013.
- [41] M. M. Daud, W. M. D. W. Zaki, A. Hussain, and H. A. Mutalib, "Detection of keratoconus in anterior segment photographed images using corneal curvature features," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 3, pp. 1191–1198, 2019.
- [42] M. Aghamohamadian-Sharbat, H. R. Pourreza, and T. Banaee, "A novel tortuosity-based algorithm for automatic grading of retinal blood vessel tortuosity," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 2, pp. 586–595, Mar. 2016.
- [43] E. D. C. Sánchez, J. N. Martínez, and F. J. F. Cañavate, "Corneal topography in keratoconus: State of the art," *Eye Vis.*, vol. 3, no. 1, p. 5, 2016.
- [44] O. I. Frette, G. Virnovsky, and D. Silin, "Estimation of the curvature of an interface from a digital 2D image," *Comput. Mater. Sci.*, vol. 44, no. 3, pp. 867–875, Jan. 2009.
- [45] P. Slavinić and M. Cvetković, "Volume calculation of subsurface structures and traps in hydrocarbon exploration—A comparison between numerical integration and cell based models," *Open Geosci.*, vol. 8, no. 1, pp. 14–21, Jan. 2016.
- [46] E. C. Ehman et al., "Immunotherapy of non-Hodgkin's lymphoma with a defined ratio of CD8⁺ and CD4⁺ CD19-specific chimeric antigen receptor-modified T cells," *Sci. Transl. Med.*, vol. 46, no. 5, pp. 1247–1262, 2017.
- [47] S.-T. Yeh, "Using trapezoidal rule for the area under a curve calculation," in *Proc. 27th Annu. SAS User Gr. Int.*, 2002, pp. 1–5.
- [48] H. Liu and H. Motoda, *Computational Methods of Feature Selection*. Boca Raton, FL, USA: CRC Press, 2007.
- [49] R. C. De Amorim, "Computational methods of feature selection," Huan Liu, Hiroshi Motoda, CRC Press, Boca Raton, FL (2007), 440 pp., ISBN: 978-1-58488-878-9," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 490–493, 2009.
- [50] R. P. L. Durgabai and R. B. Y., "Feature selection using ReliefF algorithm," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 3, pp. 8215–8218, Oct. 2014.
- [51] L. Wu, Y. Hu, X. Liu, X. Zhang, W. Chen, A. S. L. Yu, J. A. Kellum, L. R. Waitman, and M. Liu, "Feature ranking in predictive models for hospital-acquired acute kidney injury," *Sci. Rep.*, vol. 8, no. 1, p. 17298, Dec. 2018.
- [52] K. R. Foster, R. Koprowski, and J. D. Skufca, "Machine learning, medical diagnosis, and biomedical engineering research—Commentary," *Biomed. Eng. Online*, vol. 13, no. 1, pp. 1–9, 2014.
- [53] S. Bak-Nielsen, C. H. Ramlau-Hansen, A. Ivarsen, O. Plana-Ripoll, and J. Hjortdal, "Incidence and prevalence of keratoconus in Denmark—an update," *Acta Ophthalmologica*, vol. 97, no. 8, pp. 752–755, Dec. 2019.



MARIZUANA M. DAUD was born in Klang, Selangor, Malaysia, in 1988. She received the B.Eng. degree in electrical and electronics engineering from the University of Malaysia Pahang, Pahang, Malaysia, in 2011, the M.Sc. degree from University Technology PETRONAS, Perak, Malaysia, in 2014, and the Ph.D. degree in electrical and electronics engineering from University Kebangsaan Malaysia, Selangor.

She is currently a Senior Researcher at MIMOS Berhad, Technology Park Malaysia, Malaysia. Her research interests include the development of hierarchy system for compound detection, determine the associated peak in mass spectrum, biomedical imaging, surveillance imaging, feature matching, and feature calculation.



W. MIMI DIYANA W. ZAKI (Member, IEEE) received the bachelor's degree in electronics engineering, the master's degree (Engr. Sc.), and the Ph.D. degree from Multimedia University (MMU), Cyberjaya, Malaysia, in 2000, 2005, and 2013, respectively.

In 2008, she joined Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia, where she is currently a Researcher and an Associate Professor at the Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment. Her research interests include biomedical engineering, bioinformatics, intelligent systems, image processing, and the IoT-related healthcare technology.



AINI HUSSAIN (Member, IEEE) received the B.Sc. degree in electrical engineering from Louisiana State University, USA, the M.Sc. degree in systems and control from The University of Manchester Institute of Science and Technology (UMIST), U.K., and the Ph.D. degree in electrical and electronics engineering from Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia.

She is currently a Professor at the Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, UKM. Her research interests include AI, intelligent systems, machine learning, and signal and image processing. In the big data era, she is also looking into research involving predictive analytics, cloud computing, and the IoT applications.



HALIZA ABDUL MUTALIB received the bachelor's degree (Hons.) in optometry from Universiti Kebangsaan Malaysia (UKM) in 1993, the M.Sc. degree in 2000, and the Ph.D. degree from UMIST, Manchester, U.K., in 2000.

She is currently an Associate Professor in optometry with Universiti Kebangsaan Malaysia. She specializes in contact lenses and corneal morphology with extensive research. She has authored a few books and also a few chapters in books. She is also actively involved in community work and is the Coordinator of the UKM Mobile Optometry Clinic.

...