Received: February 11, 2024. Revised: March 28, 2024.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Automated Cataract Detection and Classification Using Random Forest Classifier in Fundus Images

Esra'a Mahmoud Jamil Al Sariera[1] M. C. Padma[2] Thamer Mitib Al Sariera[3]*

[1]*Department of Computer Science and Engineering, PES College of Engineering Mandya, University of Mysore, Mysore, India*
[2]*Department of Computer Science and Engineering, PES College of Engineering Mandya, Mandya, India*
[3]*Department of Competer Science and Information Systems, Amman Arab University, Amman, Jordan*
* Corresponding author's Email: t.alsariera@aau.edu.jo

Abstract: One of the most common causes of blindness, particularly in the elderly, is cataracts. Nearly half of India's elderly population has cataracts by the age of 80 or has had surgery to treat them. According to surveys conducted by the WHO and NPCB, there are over 12 million blind persons in India, and 80.1% of them are blind due to cataracts. Early detection of cataract cases is necessary to prevent total blindness. The World Health Organization (WHO) states that cataracts are the most frequent cause of blindness and visual loss. The risk of blindness among cataract patients can be decreased with prompt detection and treatment. On the other hand, clinical cataract identification requires the expertise of ophthalmologists. Therefore, the broad adoption of cataract detection to avert blindness may be hampered by the prospective expenses. Researchers are becoming more and more interested in artificial intelligence assisted diagnosis based on medical imagery. This study suggests an automated cataract detection procedure built on image processing and machine learning methods. A set of fundus retinal images serves as the input for the suggested model. The image dataset includes two kinds of images: healthy and images with cataracts to train the algorithm. This research consists of three primary phases: pre-processing, feature extraction and classification. The initial step of the method involves pre-processing the images to make them easier to process by extract the gray scale from the input image, then contrast-limited adaptive histogram equalization (CLAHE) is applied to improve the image and minimize noise, Finally ROI has been extracted. The second phase is feature extraction; by extract four kinds of texture features: (I) grey-level co-occurrence matrix (GLCM) to extract 11 features: 1) difference variance, 2) inverse difference moment, 3,4) information1\2, 5) entropy, 6) difference entropy, 7) correlation, 8) sum entropy, 9) maximal correlation coefficient, 10) contrast, and 11) angular second moment. (II) First Order Statistics (FOS) to extract 5 features: 1) entropy, 2) maximal gray level, 3) kurtosis, 4) skewness, and 5) energy. (III) Statistical Feature Matrix (SFM) to extract 4 features: 1) coarseness, 2) periodicity, 3) contrast, and 4) roughness. And finally (IV) Neighborhood Gray Tone Difference Matrix (NGTDM) to extract 4 features: 1) complexity, 2) coarseness, 3) strength, and 4) contrast. In the last phases, the extracted 24 features are put as input to the classifier, the classification was done by using Random Forest, support vector machines (SVM), Logistic Regression, K Neighbors (KNN), Decision Tree, and Naive Base Classifier. Classifies the retinal fundus images into two classes Cataracts or normal image. When compared to other current approaches, the experimental results of the suggested method show that its accuracy is 95%.

Keywords: Cataract, Computer-aided diagnosis, Retinal image pre-processing, Feature extraction, GLCM features, FOS features, SFM features, NGTDM features, Classification, SVM classifier, KNN classifier, Random forest classifier, Logistic regression classifier, Decision tree classifier, Naive base classifier.

1. Introduction

The eye is a vital component of the human body that consists of around 40 interrelated subsystems, including the optic nerve, retina, iris, pupil, and lens. The optic disc, macular and blood vessels are the three primary components of the retina. The normal objects of retina are shown in Fig. 1. The optic disc, which links to brain nerves and is the highest circular

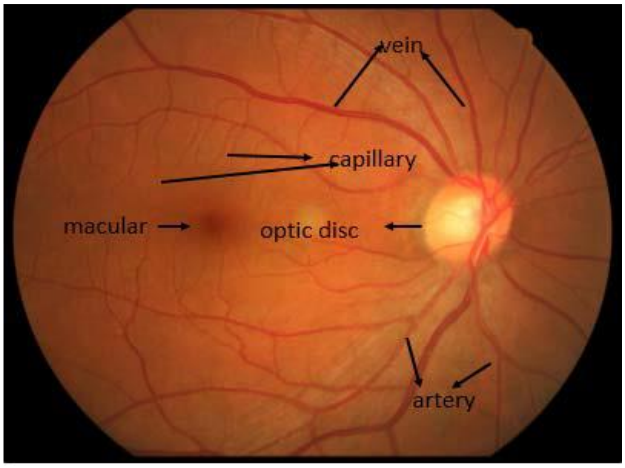


Figure. 1 The objects of retina

disc in the middle of the retina, reflects the health of the entire body, including the brain.

The macular, or optical center of the eye, has a high lutein content, which gives it a yellow appearance. There are two types of blood vessels: arteries and veins. At the optic disc, all blood arteries converge. The lens of the eye functions similarly to a lens in a camera, directing light onto the retina at the rear of the eye to capture an image. To view objects, regardless of their distance, it also regulates the eyes focus. Lenses are clear because they are composed of water and protein; but, occasionally, protein clumps form behind the lens, causing a tiny region of the lens to become clouded with white and impairing vision. As people age, their proteins begin to clump together, clouding the lens of the eye and preventing light from entering. This condition is known as a cataract, according to the World Health Organization (WHO). These proteins only become larger when the older cells compress into the center of the lens, obscuring the retina image [1]. In addition, one of the main causes of the difficulty in removing a cataract at an advanced stage is that it may develop into a hyper maturity cataract, in which the lens solidifies, contracts, wrinkles, or becomes soft and liquid [2]. This can make surgery difficult and risky for the surgeon, as well as result in unsatisfactory outcomes for the patient. The hazards of surgery must thus be weighed against the patient's ability to restore vision by the ophthalmologists. In this instance, it will be advised not to treat the cataract. Early-stage cataract surgery is advised by ophthalmologists. Most people with cataracts are over 50. They can also develop as a result of diabetes mellitus, high blood pressure, smoking, prolonged sun exposure, eye injuries, alcohol use, systemic corticosteroids, excessive oxidant production, prolonged use of steroids and medications, UV radiation, genetics, and disorders related to nutrition or metabolism [3-5]. It also has correlations with several other disorders [2, 4]. Approximately 285 million individuals worldwide suffer from visual impairment, according to the World Health Organization (WHO) [6]. 39 million of them have limited vision, while the remainder of individuals have vision impairments. 33% of visual impairment and 51% of blindness were attributed to cataracts [7]. According to estimations made by Flaxman et al. [8], there were 38.5 million blind persons and 237.1 million people with moderate to severe vision impairment (MSVI) in 2020. Out of them, 13.4 million (35%) and 57.1 million (24%) would have cataracts. By 2025, there will be more than 40 million blind people globally [9]. The most common cause of permanent vision loss is cataracts, an eye condition characterized by lenticular opacity. Above all, patients may find it difficult to detect cataracts at an early stage because they don't exhibit any obvious symptoms. However, if screening and treatment are put off, the severity of the visual impairment will worsen. As a result, early screening is crucial to enhancing the standard of care for cataract patients [10, 11]. There are more than 12 million blind persons in India. Ninety percent of those who live there are from rural regions. In India, the frequency of blindness is 1.1%. To the best of our knowledge, India was the first nation to implement blindness prevention measures. However, they are typically overloaded with routine eye exams, most ophthalmologists don't have much time to perform procedures that prevent blindness. If skilled ophthalmologists conduct a sufficient number of procedures, cataract related blindness can be eliminated. Nonetheless, it remains difficult to perform the necessary number of cataract procedures with the ophthalmologists on hand. The immediate consequences of a cataract, include myopia, and the potentially fatal long-term effects, include blindness. The main obstacles to the prevention of blindness are low public awareness, restricted accessibility, high treatment costs, and subpar surgical results [12, 13]. According to Fig. 2, cataracts account for 51% of all cases of blindness. In 20 of the 21 locations in the world where cataracts occur, uncorrected refractive error (URE) (3%), glaucoma (8%), age-related macular degeneration (AMD) (5%), corneal opacity (4%), trachoma (3%), diabetic retinopathy (DR) (1%), and other causes (25%) are unquestionably the primary causes of blindness [14]. Without a doubt, one of the main causes of blindness is cataracts. Cataracts can be prevented and significant advantages can be obtained quickly if they are properly identified and diagnosed in their early stages. This raises the need for more study in the non-invasive computer-aided cataract diagnosis.

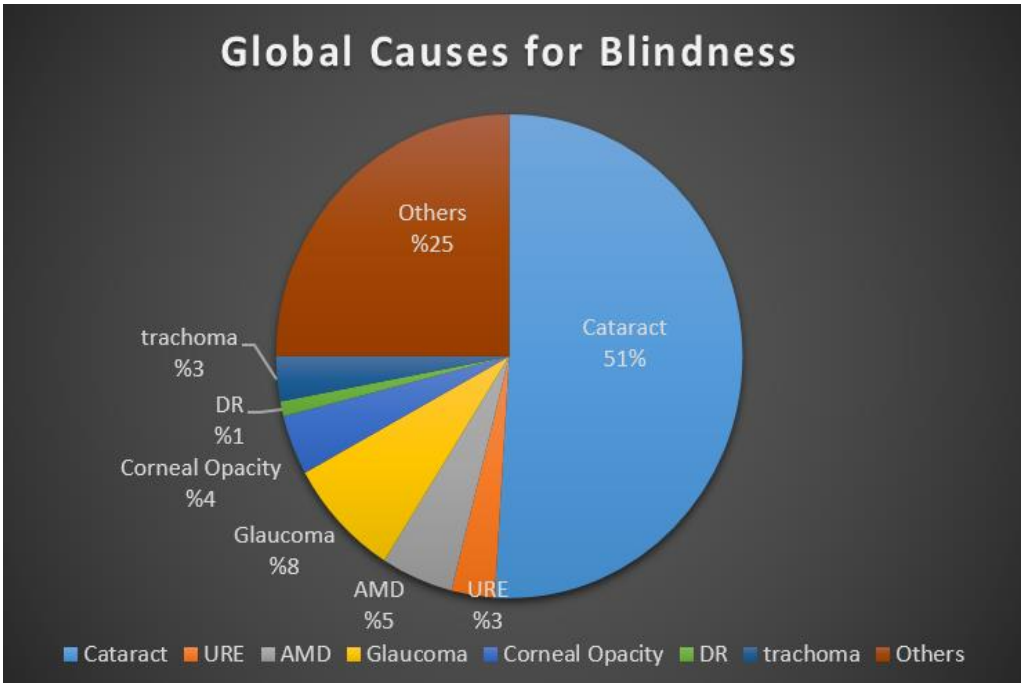
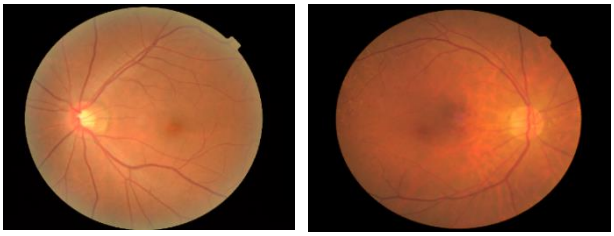


Figure. 2 Global causes of vision impairments



(a) (b)

Figure. 3. Retinal Fundus Image from cataract dataset: (a) Normal Retina Image and (b) Cataract affected Retina Image

Traditional diagnostic techniques rely on the slit lamp, which is an intricate and costly tool. Fundus images are a simpler way to diagnose cataracts than slit-lamp images. Ophthalmologists can identify cataracts by evaluating fundus images for clarity. The retinal images are used by the ophthalmologist to screen for eye problems. Drops are used to dilate the pupil of the eye, improving the vision of the retina [15]. These techniques need expensive equipment and are limited to usage by skilled ophthalmologists for eye screening. Early cataract identification can lower the significant risks of vision impairment developing into blindness. The ophthalmologist uses the retinal images to determine whether a cataract is present. The structural changes occurring in the retina are used to make this diagnosis.

In this study, machine learning and image processing techniques are used to construct an automatic model for cataract identification. As seen in Fig. 3, a set of retinal images obtained using a fundus camera are used as inputs to construct the model. The images are divided into two categories: healthy and cataract-affected retinas. Retinal images are enhanced by pre-processing techniques including grayscale image extraction and histogram equalization. Then the texture features are extracted. Finally, the classification is performing by machine learning classifiers.

The research goals are outlined below:

1. To examine the input fundus image and apply pre-processing methods for smoothing.
2. To take the grayscale images components features and extract them.
3. To use the appropriate classifiers to identify retinal images as normal or cataracts.
4. To examine the accuracy, sensitivity, specificity, F1-score, precision, and Correlation Coefficient of the classification performance metrics.

This paper strategy has the potential to greatly increase ophthalmologists diagnostic efficiency while lessening the financial and physical load on patients and society.

The organization of the remaining paper is as follows: Section 2 describes related works in cataract detection. Pre-processing of retinal images, automatic feature extraction using texture feature model and classification of features using support vector machines (SVM), Random Forest, Logistic Regression, K Neighbors (KNN), Decision Tree, and Naive Base Classifier are explained in Section 3. The experimental results are discussed in Section 4. Conclusion and Future work are presented in Sections 5.

2. Related work

In recent years, an increasing percentage of researchers have focused on the analysis of fundus images, and it is increasingly being utilized in clinical settings for automated cataract grading. The four main components of the cataract classification approach are pre-processing, feature extraction, feature selection, and classifier. According to a study, fundus images may be analyzed using image processing techniques to detect cataracts in the eyes. This section provides a quick overview of the several methods used to identify cataract disease in patients utilizing a cataract image. Then, a variety of methods used to extract features from cataract fundus images are also given. To determine the weight and bias value as processing to identify the optimal model. In [16] employed CNN to classify image the cataract. It has achieved 77.5% accuracy using the cataract data set from Kaggle. The disadvantage of this study is that it employs multiple levels of processing to determine the optimal weights and bias values for the model. The authors in [17] presented a deep learning method for concurrently learning multilayer feature representations of the fundus image. To tackle the cataract classification task, a global-local attention network (GLA-Net) is specifically suggested. This network is composed of two levels of subnets: the global-level attention subnet attends to the global structure information of the fundus image, while the local-level attention subnet concentrates on the local discriminative features of the designated regions. The final cataract categorization is achieved by combining the retinal information extracted by these two types of subnets at varying attention levels. The model used in this work is trained primarily on a large amount of labeled data, which poses a significant challenge for clinical applications. In [18] proposed a unique approach to cataract grading that uses the RF classifier for grading and the MobileNet-V2 architecture for feature extraction. This work aims to present a minimal computational processing, high-performance cataract grading approach that is suited for mobile devices. A dataset of 590 fundus images is used for the assessment, and an accuracy of 90.68% is attained. The work in [19] suggested an ensemble learning method for cataract grading that stacks three convolutional deep neural networks to maximize grading performance. The main contributions are as follows: (1) fundus images are pre-processed and data augmented to ensure the robustness of the cataract grading; (2) established deep learning architectures (InceptionV3, MobileNet-V2, and NasNet-Mobile) are optimized and trained as base classifiers; and (3) a stacking technique is suggested to combine the features of base classifiers. The study makes use of a dataset of 590 fundus images selected from Cataract dataset and ODIR dataset. It can be necessary to increase or modify pre-processing in order to properly handle the amount and complexity of the data. The authors in [20] used a dataset of images of both normal and cataract-affected eyes, Convolution Neural Networks (CNN) and Support Vector Machines (SVM) were employed in this study to detect cataracts. The research has used a variety of strategies as part of the model construction process, including feature extraction, label encoding, and data augmentation as a pre-processing step. In this study [21] presented an automated cataract diagnostic and grading method based on Convolutional Neural Networks (CNNs) that classifies fundus images into four categories: Normal, Mild, Moderate, and Severe. With 92.7% accuracy across 4 classes. In this study, all randomly chosen retinal images from fundus image databases that are available to the public (HRF, STARE, MESSIDOR, DRIVE, cataract, and DRIONS_DB) were used. This method required to reduce training and testing time while increasing classification efficiency. In [22] presented a vessel attention-guided multi-granularity network (VAM-Net) with three main components. First and foremost, the vessel-level segmentation subnet is created using the U-shape segmentation network. These stages improved vessel information may be regarded as visual attention, allowing deep network extraction of discriminative features. It is also recommended that the global-local classification subnet be used to enhance multi-granularity feature learning (which includes both local subtle features and global structural features) and eliminate extraneous information interference while keeping vessel attention under check. Finally, an ensemble network with many granularities is built, and the final cataract diagnosis is based on the predictions from various levels. The real-world cataract dataset was used to evaluate the efficacy of the suggested VAM-Net (92.78% detection accuracy and 87.68% grading accuracy). The study in [23] aimed to offer strategies for applying deep learning models to detect important eye problems. The data set utilized for identifying diseases consists of 2748 retinal fundus images, collected from 1374 normal individuals and 1374 individuals who belong to different disease categories. To compare the classification performances and achieve better performance, a total of 5 different CNN designs were employed in the quest for a solution to the disease identification challenge. This paper focused to classify three types of diseases and the result accuracy achieved to classify the cataract were less than 86%.

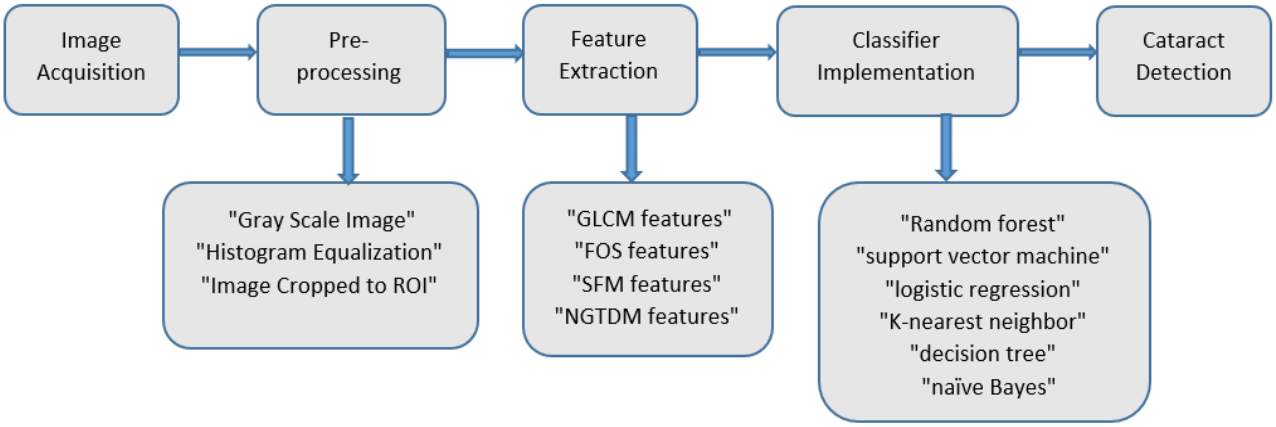


Figure. 4 The block diagram of cataract detection

Table 1. Fundus image material

Class	Label	Image Train	Image Validation
Normal	0	240	60
Cataract	1	80	20
Total Image		320	80

According to the review of the related work, we noted the following points. In such a classification problem, separating the close two-class retinal images is the most challenging decision-making technique because of their similar features and the difficulties of successfully recording their distinctions. Also, feature extraction techniques can also perform satisfactorily in comparison to deep techniques, especially when working with insufficient training data. Finally, the majority of the retinal structures are preserved in the grayscale of the image, and all of the actions in our method are carried out on the grayscale.

3. Proposed method

This study is comprised of four major phases: image acquisition, pre-processing, feature extraction, and classification. The suggested architecture design is shown in Fig. 4. The dataset was taken from the cataract dataset. The input RGB image is first converted to grayscale image. Next, the histogram equalization approach is applied to improve the image, and finally, the image is cropped to a region of interest (ROI). Four texture feature models are used to extract the features from grayscale images. In the end, each classifier was trained using a mixture of the four features. It was determined that integrating features rather than using each kind of feature separately will result in greater accuracy.

3.1 Data collection

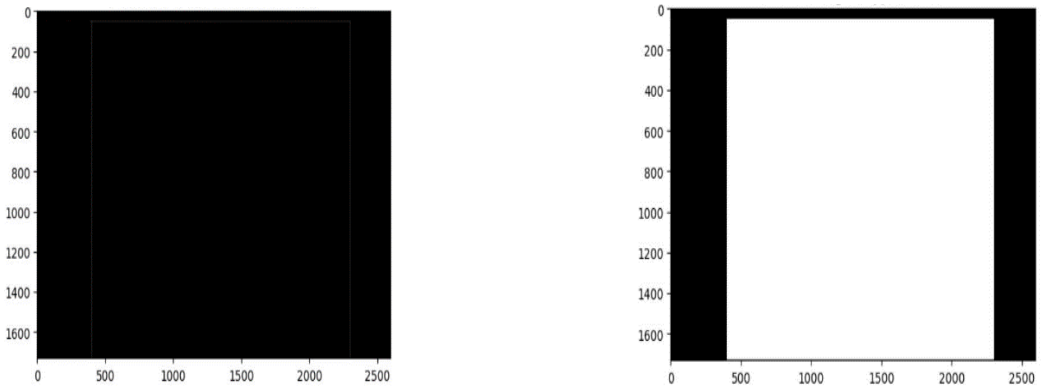
We construct a dataset of retinal images selected from public databases recently published in the Kaggle platform retrieved on January 2020. The Cataract dataset [24] is composed of 600 fundus images having a size of 2592*1728, where 100 are affected by the cataract disease. The retina dataset containing four categories: normal, cataract, glaucoma, and retina disease. The data would be split into training data and validation data. The content of the data split can be seen in Table 1.

3.2 Pre-processing

The proposed system dataset combines images of normal, glaucoma, cataract, and retina disease. In the first phase, we have divided all fundus images aside from cataract and regular fundus images. Given that we are working with Fundus retinal images as input, the models accuracy depends only on the way we pre-process the data. To construct the model, the RGB image dataset needs to go through some basic image processing operations, such as converting the image to grayscale, which can improve the contrast between the background and the subject while preserving the majority of the original images details, supervised machine learning is used. Then, to increase the divergence, Contrast-Limited Adaptive Histogram Equalization (CLAHE) is used. The intensities are frequently more widely scattered on the histogram as a result of this modification. By using this technique, a region with lesser contrast can attain better contrast, and the image is finally cropped to ROI.

3.2.1 Gray scale image

The first stage of image pre-processing involves converting RGB images to grayscale forms. For



(a) (b)

Figure. 5 the process of ROI extraction: (a) Perimeter of image cropped to ROI and (b) Mask of image cropped to ROI

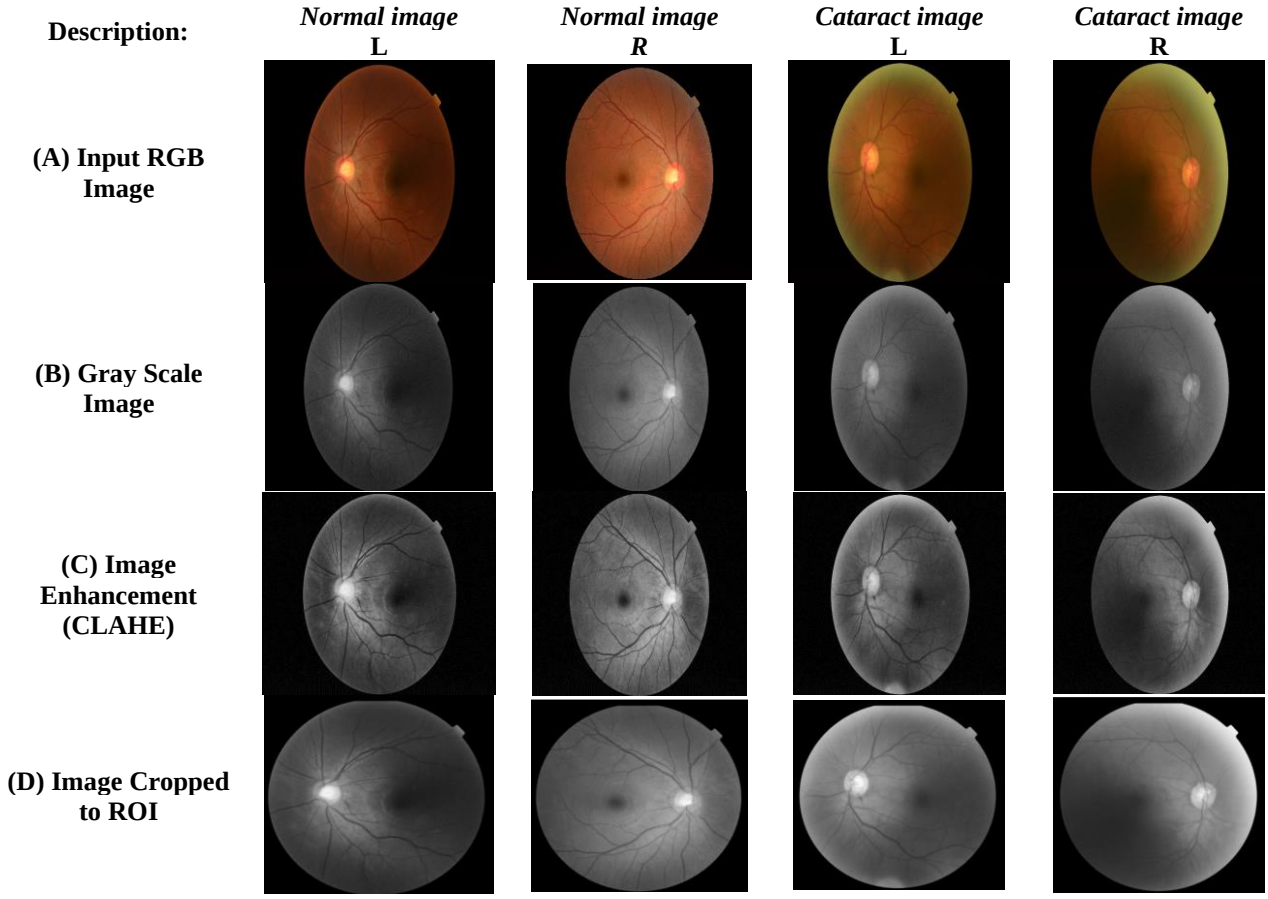


Figure. 6 Step wise pre-processing output results for the input fundus images

various items, grayscale may be adequate for differentiation. Color images can add more complexity and need more memory space since they include more information than black and white ones. The number of pixels that need to be processed is reduced when a color retinal image is converted to grayscale. A grayscale image is one that only contains information on intensity and has one sample pixel value. The method of turning multiple color channels (R, G, and B) into a single gray scale value via a weighted summation is known as gray scale conversion. Below is the formula for converting grayscale to brightness based on characteristics.

$$G = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \quad (1)$$

The result of converting the color image into gray scale seen in Fig. 6 (B).

Table 2. Features extracted using Grey Level Co-occurrence Matrix (GLCM), First Order Statistics (FOS), Statistical Feature Matrix (SFM) and Neighborhood Gray Tone Difference Matrix (NGTDM)

GLCM Feature	FOS Feature	SFM Feature	NGTDM Feature
Difference variance	Entropy	Coarseness	Complexity
Inverse Difference Moment	Maximal Gray Level	Periodicity	Coarseness
Information	kurtosis	Contrast	Strength
Entropy	Skewnewss	Roughness	Contrast
Difference Entropy	Energy		
Correlation			
Sum Entropy			
Maximal Correlation Coefficient			
Contrast			
Angular Second Moment			

3.2.2 Image enhancement (CLAHE)

Contrast Limited Adaptive Histogram Equalization: this could reduce the impact of varying image brightness on the results of classification and make the improved image insensitive to variations in local brightness. The improved image will have more contrast and a wider dynamic range in its gray level of pixels once equalization has occurred. The pixels in the image will occupy as many gray levels as feasible and be distributed equally. The histogram equalization procedure is implemented in Eq. (2). (In the equation, G_f is the gray value after conversion, G_{max} is the maximum gray value in the image, N_0 is the total number of pixels, G_A is the gray value prior to conversion, P_i is the number of pixels in level i gray scale, L is the largest pixel level in the image's pixels.) Because it is simple to apply, fast, and efficient, histogram equalization is frequently employed in image improvement. The result of the Histogram Equalization shown in Fig. 6 (C).

$$G_f = \frac{G_{max}}{N_0} \sum_{i=0}^{G_A}(P_i) \quad (G_A = 0,1,2,\dots,L-1) \quad (2)$$

3.2.3 Image cropped to ROI

To manage the various resolutions of the images, each one will be automatically resized to the same size. It is important to understand that the ROI masks perimeter is normalized instead of the input images size. An ROI mask is obtained in order to acquire this perimeter. This technique is basically used to distinguish between the retina and the dark background. To avoid misclassifying areas with low illumination in the retina as background, Fig. 5 shows the process of ROI extraction. The Gray Scale Image is taken from the RGB color space and binarized because it shows a stronger contrast between the ROI contour and the background. In our method, the mask is an image with two values, 0 (zero) and 1 (one), where 1 denotes the region-of-interest (ROI), around which the features are to be computed (values outside the ROI are ignored). The perimeter, on the other hand, is similar to the mask but shows the ROI perimeter. Fig. 6 (D) shows the results of this process.

3.3 Feature extraction

In computer vision problems, the initial step for all machine learning algorithms is to extract features from the provided images. An essential first step in any system for classifying patterns is feature extraction. Patterns must be transformed into features, which are streamlined representations of the patterns that only include important information, in order for the pattern recognition process to be tractable. Features enable pattern classification by encapsulating a pattern's attributes in a similar format. In fundus imaging features related to the cataract are identified by texture feature extraction. This procedure is crucial to our cataract detection methodology. The texture-based extract features from fundus images that are appropriate for cataract diagnosis are used in this work.

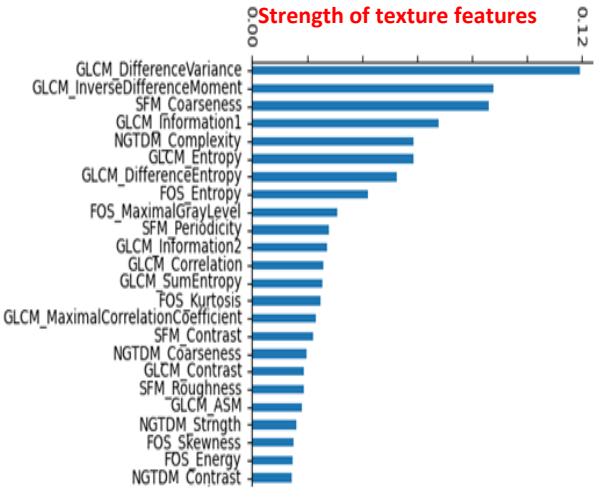


Figure. 7 The most robust texture feature that use to detect cataract diseases

In image processing, texture refers to a collection of metrics that are computed to measure the perceived texture of an image. Texture provides information about the spatial distribution of color and intensity in an image or a specific area within an image [25]. In this paper, we select the Gray-Level Co-occurrence Matrix (GLCM), First Order Statistics (FOS), Statistical Feature Matrix (SFM), and Neighborhood Gray Tone Difference Matrix (NGTDM) as the texture feature matrix. Before calculating the eigenvalues of the four texture matrixes, we have to obtain the four matrixes from original fundus images. And according to our methodology the most robust texture features in detect cataracts disease shown in Fig. 7 and Table 2.

3.3.1 Gray level Co-occurrence matrix (GLCM)

According to [26], the Gray Level Co-occurrence Matrix (GLCM) is based on the estimate of the joint conditional probability density functions of the second order. This study presents the extraction of a statistical texture feature based on the GLCM. The statistical characteristics of the texture region will be computed, and the feature is based on the gray qualities of the pixels and the areas around them. A pixel's joint probability density pertaining to two places defines the GLCM. It has the ability to display integrated data that includes the range, direction, and closeness of changes in gray value. In this paper, 11-types of feature values are employed to extract the texture features of the image (Table 4). 1) difference variance, 2) inverse difference moment, 3) information1, 4) entropy, 5) difference entropy, 6) information2, 7) correlation, 8) sum entropy, 9) maximal correlation coefficient, 10) contrast, and 11) angular second moment. Are calculated, these typical parameters can directly reflect the texture situation of co-occurrence matrix [27].

The second-order combined conditional probability density can be calculated using the following equation:

$$P(i, j) = g\{(x_1, y_1), (x_2, y_2) \epsilon\ m \times n \mid f(x_1, y_1) = i, f(x_2, y_2) = j\} \quad (3)$$

Where P be the probability distribution for gray-level co-occurrences, i and j are the gray values of the gray image f(x, y) at the (x_1, y_1) and (x_2, y_2) coordinate positions in Eq. (3).

Notes: according to Haralick Ng is the number of distinct grey levels in the quantized image, Nevertheless, the calculations showed that Ng is really the highest possible quantized value (referred to as Gmax in the following formulas). where p(i, j) is the normalized GLCM value at the (i, j) coordinate point; μ_x and μ_y are the corresponding means of $p_x(i)$ and $p_y(j)$; and σ_x and σ_y are the corresponding variances of $p_x(i)$ and $p_y(j)$. Table 3. Include the Eqs (4)-(15) that can be used to compute the necessary intermediate parameters in the GLCM texture feature formulas.

3.3.2 First order statistics (FOS)

First-order statistics (FOS) are computed using the probability of observing a specific pixel value at an artificially chosen area within the image. They are independent of the interactions between neighboring pixel values and solely rely on the values of individual pixels. Are computed using the images histogram, which represents the empirical probability density function for individual pixels [43]. The histogram shows the number of pixels with gray level intensity i for each intensity level in an image I with N unique gray levels and a total of M pixels. The probability density of occurrence of that intensity level computed as follows:

$$P(i) = \frac{N(i)}{M} \quad (27)$$

Let: X be a set of N_p voxels included in the ROI, Xi be i_{th} element of X, N_g be the number of discretized intensity values in the image, P_j be the occurrence probability of discretized gray level j. which is equal to (the number of voxels with discretized gray level j) / (total number of voxels N_p).

The FOS features that used in our method are the following: 1) entropy, 2) maximal gray level, 3) kurtosis, 4) skewnewss, and 5) energy shown in Table 5.

3.3.3 Statistical feature matrix (SFM)

The authors in [28] proposed the foundation of the statistical feature matrix. It is employed in texture feature extraction based on visual perception. Measuring the distance between two statistical feature matrices defines the correlation between them. The SFM approach allows for easy matrix expansion, and the size of the matrix varies based on the distance between pixels rather than the number of gray levels. Based on visual perception, four characteristics have been retrieved in this instance. For each ROI, the four statistical matrix properties of coarseness, periodicity, contrast, and roughness were determined, as shown in Table 6.

Let S = {x, y} be the spatial coordinates of a Ly×Lx array of points and Z(x, y) be the intensity at a point in S. Let δ be the inter sample spacing distance vector,

Table 3. Description the GLCM parameters

$P_x(i) = \sum_{j=1}^{G_{max}} \mathrm{P(i,j)}$ (4)	$\mu_y = \sum_{j=1}^{G_{max}} \{ j \bullet \sum_{i=1}^{G_{max}} \mathrm{P(i,\ j)}\}$ (10)
$P_y(j) = \sum_{j=1}^{G_{max}} \mathrm{P(i,j)}$ (5)	$\sigma_x = \left(\sum_{i=1}^{G_{max}} \left\{ (i - \mu_x)^2 \bullet \sum_{j=1}^{G_{max}} \mathrm{P(i,j)}\right\}\right)^{1\backslash 2}$ (11)
$P_{x+y}(n) = \sum_{i+j=n}^{G_{max}-1} \mathrm{P(i,j)}; n \epsilon \{ 0,1,2, \dots, G_{max} - 1\}$ (6)	$\sigma_y = \left(\sum_{i=1}^{G_{max}} \left\{ \left(j - \mu_y \right)^2 \bullet \sum_{i=1}^{G_{max}} \mathrm{P(i,j)}\right\}\right)^{1\backslash 2}$ (12)
$P_{x-y}(n) = \sum_{\lvert i-j\rvert=n}^{G_{max-1}} \mathrm{P(i,j)}; n \epsilon \{ 0,1,2, \dots, G_{max} - 1\}$ (7)	$H_{XY} = -\sum_{i=1}^{G_{max}} \sum_{j=1}^{G_{max}} \mathrm{P(i,j)}\ log\ \{\mathrm{P(i,j)}\}$ (13)
$\mu_{x-y} = \sum_{n=0}^{G_{max}-1} \{ n \bullet P_{x-y}(n)\}$ (8)	$H_{XY1} = -\sum_{i=1}^{G_{max}} \sum_{j=1}^{G_{max}} \mathrm{P(i,j)}\ log\ \{P_x(i)\ P_y(j)\}$ (14)
$\mu_x = \sum_{i=1}^{G_{max}} \{ i \bullet \sum_{j=1}^{G_{max}} \mathrm{P(i,\ j)}\}$ (9)	$H_{XY2} = -\sum_{i=1}^{G_{max}} \sum_{j=1}^{G_{max}} P_x(i)\ P_y(j)\ log\ \{P_x(i)\ P_y(j)\}$ (15)

and η be the average gray-level of an image I, the contrast, covariance, and dissimilarity are defined as:

$$CON = E\{[I(x,y)] - I(x+\Delta x, y+\Delta y)]^2\} \quad (33)$$

$$COv = E\{[I(x,y) - \eta][I(x+\Delta x, y+\Delta y) - \eta]\} \quad (34)$$

$$DSS = E\{|I(x,y) - I(x+\Delta x,\ y+\Delta y)|\} \quad (35)$$

$$N_r = \{(i,j):\ \ |i|,|j| \le r\} \quad (36)$$

Where: **E {.}** denotes the expectation operation, **C** is normalizing factor, **n** is the number of elements in the set, $\overline{\boldsymbol{M_{DSS}}}$ is the mean of all elements of $\boldsymbol{M_{DSS}}$, $\boldsymbol{M_{DSS}(valley)}$ is the deepest valley in the matrix, $\boldsymbol{D_f^h}$ and $\boldsymbol{D_f^v}$ are the estimated fractal dimensions in horizontal and vertical directions.

3.3.4 Neighborhood gray tone difference matrix (NGTDM)

The authors in this work [29] presented this feature extraction technique as a substitute for the GLCM. This texture analysis technique is applied in various domains, such as medical imaging. The primary distinction is that the NGTDM looks at the nearest neighboring pixels/voxels up to a given Chebyshev distance rather than the pixel pairs in particular directions. Our study indicates that the characteristics that may be estimated from the resultant NGTDM matrix are complexity, coarseness, strength, and contrast as shown in Table 7.

Let **s** be the NGTDM vector, indexed **si**, and **pi** be the probability of a voxel value for voxels that are used in the computation of the NGTDM. **Ng** is the number of unique grey levels present in the image (not necessarily equal to the highest grey level value $\mathbf{G_{max}}$, since some values may not be present in the image). When a grey level is not present, the corresponding **si** is zero.

3.4 Classification

The classifiers used in the comparison of results are Random Forest, SVM, Logistic Regression, K-nearest neighbor (KNN), decision tree, and Naive Bayes. Among these classifiers, the classifier is selected depending on the performance in terms of accuracy. The Random Forest classifier performed well in classifying of cataract. These accuracies are plotted in Fig. 8.

3.4.1 Random forest classifier

As the algorithms work together, many decision

Table 4. Description and formulas to calculate texture features from GLCM

GLCM Features	Qualitative Description	Equation
Difference variance (DV)	This statistic, which quantifies heterogeneity, has a significant correlation with first-order statistical variables like standard deviation. As a gray level value deviates from its mean, variance rises.	$f1 = -\sum_{n=0}^{G_{max}-1}\{(n-\mu_{x-y})^2 \bullet P_{x-y}(n)\}$ (16)
Inverse Difference Moment (IDM)	The maximum value is reached when every element in the image is the same. In a sample of pixel pairs, there is a substantial but inverse correlation between GLCM contrast and homogeneity in terms of equivalent distribution.	$f2 = \sum_{i=1}^{G_{max}}\sum_{j=1}^{G_{max}}\{\frac{1}{1+(i-j)^2} \bullet P_{i,j}\}$ (17)
Information of correlation1\ 2 (I)	Enables the characterization of texture patterns in an image by providing details about the linear connection between pixel values in one direction and distance.	$f3 = \frac{H_{XY}-H_{XY1}}{max(H_{XY})}$ (18) $f4 = 1 - e^{(-2(H_{XY2}-H_{XY}))^{\frac{1}{2}}}$ (19)
Entropy (E)	The most entropy occurs when all the values in the co-occurrence matrix are equal or when the pixel values show the highest degree of randomness. The complexity of the images gray distribution is indicated by the entropy value.	$f5 = -\sum_{i=1}^{G_{max}} \sum_{j=1}^{G_{max}} \mathrm{P(i,j)} Ln\ \{\mathrm{P(i,j)}\}$ (20)
Difference Entropy (DE)	Measurement of the unpredictability or randomness of the variations in pixel values. It offers details on the differences in intensity between neighboring pixels in various orientations.	$f6 = -\sum_{n=0}^{G_{max}-1} P_{x-y}(n)\ Ln\ \{P_{x-y}(n)\}$ (21)
Correlation Coefficient (CC)	The correlation value of the GLCM is high if there is a texture distribution along a particular direction.	$f7 = \sum_{i=1}^{G_{max}}\sum_{j=1}^{G_{max}} \frac{(i-\mu_x)(j-\mu_y)\mathrm{P(i,j)}}{\sigma_x\sigma_y}$ (22)
Sum Entropy (SE)	The unpredictability or disorder in the distribution of the summed pixel values is quantified by sum entropy. Greater disorder or variety in the sums of pixel values is indicated by a higher sum entropy value.	$f8 = -\sum_{n=2}^{2\bullet G_{max}} P_{x+y}(n)\ Ln\ \{P_{x+y}(n)\}$ (23)
Maximal Correlation Coefficient (MCC)	The textures complexity is measured by the maximum correlation coefficient.	$f9 = \sum_{n=0}^{G_{max}} \frac{\mathrm{P(i,n)P(j,n)}}{P_x(i)P_y(n)}$ (24)
Contrast (C)	The distribution of the matrix and the extent of the local change are reflected in contrast. It displays the images sharpness and the textures depth.	$f10 = -\sum_{n=0}^{G_{max}-1} n^2 \bullet P_{x-y}(n)$ (25)
Angular Second Moment (ASM)	ASM is called Energy or Uniformity. It displays the texture thickness and the regularity of the images gray distribution.	$f11 = \sum_{i=1}^{G_{max}}\sum_{j=1}^{G_{max}}\left(\mathrm{P(i,j)}\right)^2$ (26)

Table 5. Description and formulas to calculate features from FOS

FOS Features	Qualitative Description	Equation
Entropy (ET)	The uncertainty or unpredictability in the images values is specified by entropy. It calculates the average amount of data needed to encode the images data.	$f12 = -\sum_{j=1}^{N_g} P_i \, log_2 \, P_i$ (28)
Maximal gray level (MGL)	The maximum gray level intensity within the ROI.	$f13 = \max(x)$ (29)
Kurtosis (K)	Kurtosis is a metric used to quantify how "peaked" the value distribution in the image ROI. Higher kurtosis values indicate that the distributions mass is more concentrated in the tails than in the mean. The opposite is implied by a smaller kurtosis: that the distributions majority is focused towards a peak close to the mean value.	$f14 = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(x_i - \bar{x})^4}{\left(\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(x_i - \bar{x})^2}\right)^4}$ (32)
Skewnewss (S)	The asymmetry of the value distribution around the mean is measured by skewness. This number might be positive or negative, depending on where the tail extends and the distributions mass is concentrated.	$f15 = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(x_i - \bar{x})^2}{\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(x_i - \bar{x})^2}}$ (33)
Energy (EN)	The energy features value is expressed in cubic units of voxel volume.	$f16 = V_{voxel} \sum_{j=1}^{N_p}(\mathrm{x(i)} + c)^2$ (34)

Table 6. Description and formulas to calculate features from SFM

SFM Features	Qualitative Description	Equation
Coarseness (CO)	The texture is considered coarser when most of the points in an image have somewhat differing gray levels from their neighbors.	$f17 = \frac{c}{\sum_{(i,j)\in N_r} \frac{DSS(i,j)}{n}}$ (37)
Periodicity (P)	The periodicity in the SFM means that the data contains cycles or repeated patterns. It records data on a signals periodicity or regularity.	$f18 = \frac{\overline{M_{DSS}} - M_{DSS}(valley)}{\overline{M_{DSS}}}$ (38)
Contrast (C)	The sharpness of the edges is one of the most crucial components of contrast.	$f19 = \sqrt{\sum_{(i,j)\in N_r} \frac{CON(i,j)}{4}}$ (39)
Roughness (R)	The roughness of natural surfaces is well-represented by the fractal concepts.	$f20 = \frac{(D_f^h + D_f^v)}{2}$ (40)

Table 7. Description and formulas to calculate features from NGTDM

NGTDM Features	Qualitative Description	Equation
Complexity (COP)	The frequency of the non-uniform and quick shifts in the gray levels is described by the NGTDM Complexity feature.	$f21 = \sum_{i=1}^{G_{max}} \left\{ \sum_{j=1}^{G_{max}} \{ \frac{\lvert i-j \rvert \cdot (P_i \cdot s_i + P_j \cdot s_j)}{n \cdot (P_i + P_j)} \} \right\}$ (41)
Coarseness (CO)	The degree of the spatial rate of intensity change is indicated by the NGTDM Coarseness feature.	$f22 = \left[\sum_{i=1}^{G_{max}} \{ P_i \cdot s_i \} \right]^{-1}$ (42)
Strength (ST)	The total of the absolute variations between a voxels gray level and its neighbors average gray level. It offers details on the general intensity variation in a voxels immediate neighborhood.	$f23 = \frac{\sum_{i=1}^{G_{max}} \{ \sum_{j=1}^{G_{max}} \{ (P_i + P_j) \cdot (i-j)^2 \} \}}{\sum_{i=1}^{G_{max}} \{ s_i \}}$ (43)
Contrast (C)	The spatial frequency and the dynamic range of intensity variations affect the NGTDM Contrast feature.	$f24 = [\frac{1}{N_g \cdot (N_g - 1)} \cdot \sum_{i=1}^{G_{max}} \{ \sum_{j=1}^{G_{max}} \{ P_i \cdot P_j \cdot (i-j)^2 \} \}] \cdot [\frac{1}{n} \cdot \sum_{i=1}^{G_{max}} \{ s_i \}]$ (44)

trees are constructed. The cornerstones of this method are decision trees. A random forest is a set of decision trees whose nodes are defined at the pre-processing stage [30]. The best feature is chosen from the random subset of features following the construction of many trees. An additional idea that is developed via the use of the decision tree method is to construct a decision tree. These trees comprise the random forest, which is utilized to classify new objects based on input vectors. Every constructed decision tree is put to use in classification.

3.4.2 SVM classifier

SVM is a supervised machine learning algorithm. Its learning output is robust, it can provide good results with little training sets, and its prediction accuracy is excellent. The purpose of SVM is to determine an ideal border that separates the classes based on data in the training set [31].

3.4.3 Logistic regression classifier

Logistic regression comes under the supervised classification algorithm. The significance of this method has grown recently, and its use has grown significantly. Individuals are classified using this approach according to their logistic function [32].

3.4.4 KNN classifier

Due to its simplicity and cheap computing requirements, K-Nearest Neighbor (K-NN), a classification approach for learning a collection of data, has gained a lot of popularity since its introduction by Fix and Hodges. Finding the parameters for the number of k nearest neighbors was the first step in the classification procedure. The shortest distance between the training and testing sets of data is how K-NN operates. During the training phase, the training data is projected onto a multidimensional space, with each dimension denoting the outcome of the training data feature extraction. Based on the classification of the training data, the space is split into classes; in this instance, there are two classes: normal and cataract. The feature extraction result of the testing data was also projected into some vectors in multidimensional space during the classification process, and the similarity between the test and training data was determined by utilizing the euclidean distance to obtain the closest distance.

3.4.5 Decision tree classifier

One of the most popular categorization methods in real-world machine learning applications is decision trees [33]. Decision tree approaches,

sometimes termed recursive partitioning. To create a decision tree, the training data is divided into subsets that are as pure as feasible. Its central concept is a tree composed of instance attributes. An instance can be classified based on the values of the attributes from the root node to the leaf node. It resembles a tree. Instead, non-leaf nodes are attributes of an instance, while leaf nodes are class sets. The most crucial problem in building the decision tree is figuring out how to choose the node with the greatest attributes each time. Starting with the root node, which represents the top of the dataset, the algorithm searches for the most significant attribute that separates the data into the most different groups. It then breaks the data down into smaller subsets, forming new branches. Every branch in the tree symbolizes a potential path. At every branch, the algorithm keeps dividing the input until it reaches the final leaf nodes, which stand for the predicted results or classifications.

3.4.6 Naive base classifier

The Naïve Bayes classifier is a basic probabilistic algorithm that gains knowledge from training data and use the maximum posterior probability to predict the test instance class. It's a simple yet effective pattern identification technique that assumes each attribute is independent. In Naive-Bayes, the inducer determines which class has a greater posterior probability by computing the conditional probabilities of several classes given the instance [34]. It permits any attribute to influence the other independent attributes final decision. It is also extensively utilized in machine learning for classification.

4. Experimental results and discussion

4.1 Experiment setup

All of our tests have been conducted using an Ubuntu 16.04 operating system, a Quadro M2000 GPU, an E5-1620 CPU, 16GB of RAM, and the Keras Python deep learning package. The major evaluation measure for test data, which consists of Accuracy, Specificity, Sensitivity, Precision, F1 Score, and Correlation Coefficient was used to validate the final results.

4.2 Results and discussion

After running the code successfully, the desired output was produced.

Table 8. Some sample of texture feature extracted using GLCM model

Sample	DV	IDM	I	E	DE	SE	C	ASM
1	0.021243	0.054469	0.86181	0.970049	0.929918	0.88522	0.72936	0.05384
50	0.096991	0.24729	0.78103	0.685823	0.732828	0.56770	0.24693	0.90503
120	0.121808	0.291426	0.75177	0.661173	0.697105	0.56035	0.23322	0.84671
150	0.07006	0.212251	0.82304	0.721323	0.775871	0.58984	0.33888	0.81506
190	0.161529	0.394744	0.66277	0.622502	0.602755	0.58062	0.33606	0.68996
230	0.129359	0.303854	0.74842	0.656989	0.690376	0.56013	0.25269	0.80802
300	0.23058	0.503688	0.57328	0.505765	0.496859	0.47988	0.18735	0.77936

Table 9. Some sample of texture feature extracted using FOS model

Sample	Co	P	C	R
1	0.05345	0.30586	0.79948	0.68630
50	0.14920	0.53951	0.3504	0.38521
120	0.16543	0.63395	0.33447	0.33779
150	0.11256	0.60948	0.45072	0.36618
190	0.21350	0.71995	0.44785	0.28079
230	0.16335	0.66004	0.35698	0.31067
300	0.32870	0.64231	0.27910	0.30866
350	0.45927	0.65292	0.24613	0.33533
399	0.720078	0.89228	0.13897	0.19287

Table 10. Some sample of texture feature extracted using SFM model

Sample	ET	MGL	K	S
1	0.87050	0.91891	0.11422	0.34116
50	0.55157	0.92792	0.11002	0.27077
120	0.53994	0.90090	0.06158	0.16718
150	0.56830	1	0.07359	0.18746
190	0.55348	0.96396	0.05152	0.10744
230	0.53673	0.89189	0.05833	0.15473
300	0.45836	0.91891	0.08552	0.15826
350	0.66338	0.77477	0.09449	0.29133
399	0.243574	0.648649	0.043456	0.05220

Table 11. Some sample of texture feature extracted using NGTDM model

Sample	COP	CO	ST	C
1	0.89699	0.11022	0.22510	0.64908
50	0.50523	0.44276	0.42235	0.42947
120	0.46096	0.46907	0.42423	0.48018
150	0.60102	0.37600	0.43342	0.56546
190	0.45976	0.38017	0.52383	0.54380
230	0.45468	0.42158	0.41948	0.48363
300	0.36292	0.41554	0.56334	0.33264
350	0.231004	0.708095	0.519718	0.23477
399	0.128005	0.494431	0.580859	0.17835

A sample of the retrieved texture feature results are displaypre-processed, Tables 8-11. The Kaggle dataset provided the data utilized in this research. There were 400 images in all, with 300 of those being normal images and up to 100 being images with cataracts. 80 test images are used to evaluate the model. 60 images are normal, while 20 are impacted by cataracts. Accuracy, sensitivity, specificity, F1-score, precision, and correlation coefficient are the six commonly used criteria that are utilized in the experiment to assess the quality of image classification. In the medical world, sensitivity is defined as the percentage of cataract patients that have a positive test result. A percentage of patients who have a negative result but do not have a cataract is known as specificity. The values of false positive (FP) normal wrongly recognized as cataract, false negative (FN) cataract incorrectly identified as normal, true positive (TP) cataract properly identified as cataract, and true negative (TN) normal correctly rejected as normal are used to generate these criteria. Defined by the following formula:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (45)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (46)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (47)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (48)$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (49)$$

$$\text{CC} = \frac{(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})}{\left((\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})\right)^{\frac{1}{2}}} \quad (50)$$

Table 12 displays the classifiers performance in our suggested model. As seen in Fig. 10, the bar plot is used to display and compare the classifiers performances. As shown in Table 12, the result for Random Forest classifier shows the highest accuracy rate at 95%. The Accuracy rate of SVM classifier was 93.57%. The accuracy rate of Logistic Regression classifier was 90%. The algorithm yielded good results on the base learning models.

A receiver operating characteristic curve (ROC) visually depicts the difference between a tests sensitivity and specificity. Samples that are assessed as true or false are calculated according to the test in order to produce a ROC curve. For certain cut-off values, the TPR (sensitivity) is mapped against the FPR (1 - specificity). To reduce false positives and increase real positives, the perfect point is one that lies over the curves shoulder. The efficiency of the suggested framework is evaluated and compared with the results that each classifier separately produced. To do this, the features that were extracted to identify images are supplied to each classifier in the algorithm. The details on the cataract detection performance of each classifier are provided by the confusion matrices and ROC in Fig. 9. In Table 13 provides an overview of the published research on cataract diagnosis using

Figure. 8 Comparison of accuracies for different classifiers evaluated on fundus images

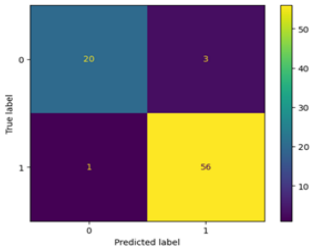
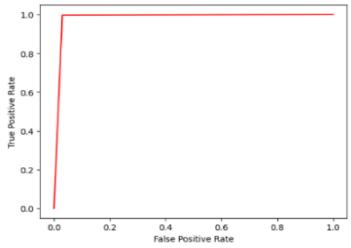
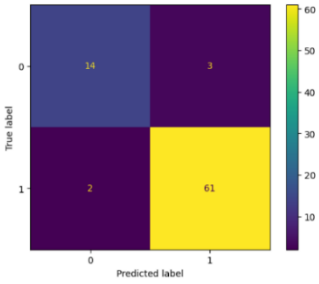
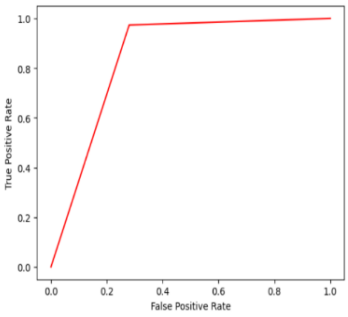
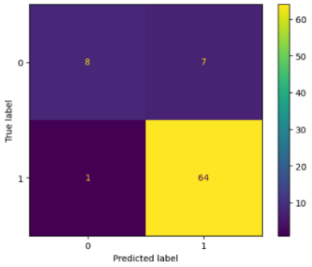
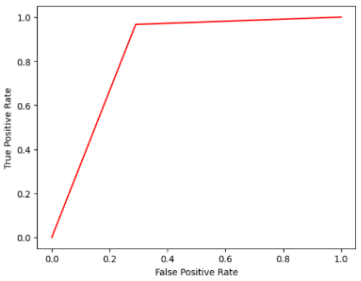
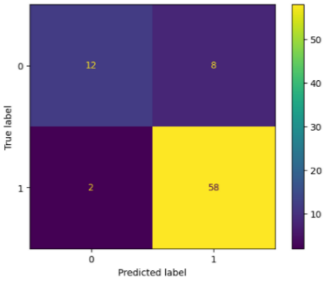
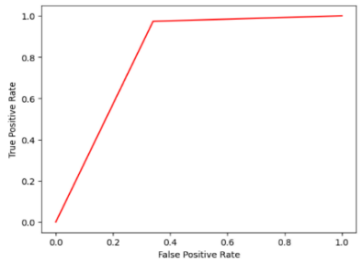
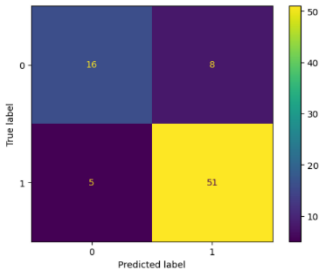
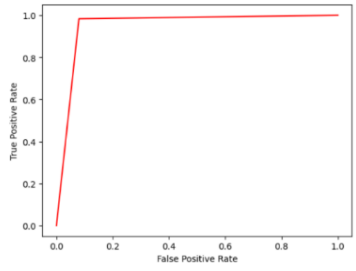
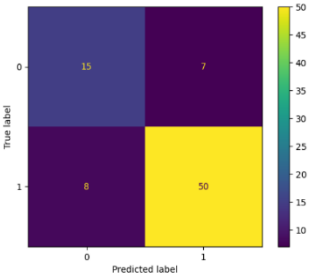
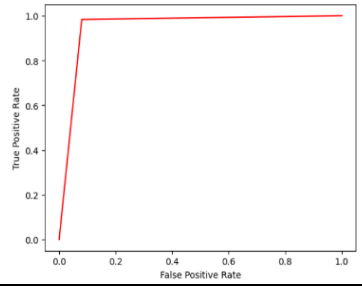
Classifier:	Confusion Matrix:	ROC Curve:
Random Forest:		
SVM:		
Logistic Regression:		
KNN:		
Decision Tree:		
Naive base:		

Figure. 9 Confusion Matrix and Receiver operating characteristic curve **(**ROC) obtained by proposed classifier

Table 12. Average performance of the proposed model for machine learning classifiers

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	Precision (%)	Correlation Coefficient (%)
Random Forest	95	86.96	98.25	91	95.24	87.64
SVM	93.57	82.35	96.83	84.85	87.5	81
Logistic Regression	90	53.33	98.46	66.66	88.88	63.98
KNN	87.5	60.0	96.67	70.59	85.71	64.58
Decision Tree	83.75	66.66	91.07	71.11	76.19	60.13
Naive base	82.5	68.18	86.21	66.66	65.22	53.66

Table 13. Performance comparison between existing methods of cataract grading

S. N o	**Approaches**	**Year**	**Method**	**Dataset**	**Accuracy %**
1	Sariah et al. [16]	2020	CNN AlexNet	Cataract Dataset	77.5%
2	Xu X et al. [17]	2021	Global Local Attention Network	Cataract Dataset	90.65%
3	Elloumi et al. [18]	2021	MobileNet-V2, random forest classifier	Cataract Dataset ODIR Dataset	90.68%
4	Elloumi et al. [19]	2022	DCNN	Cataract Dataset ODIR Dataset	93.97%
5	Pahuja et al. [20]	2022	SVM CNN	Cataract Dataset	87.5% 85.5%
6	Varma et al. [21]	2023	CNN	Cataract Dataset, HRF, STARE, MESSIDOR DRIVE,	92.7%
7	Xu et al. [22]	2023	VAM-Net	Cataract Dataset	92.78%
8	Arslan et al. [23]	2023	Dense Net	Cataract Dataset, IDRID, Oculus, Mendeley dataset,	85.35%
9	**Proposed Method**	**2024**	**Texture features, Random Forest Classifier**	**Cataract Dataset**	**95%**

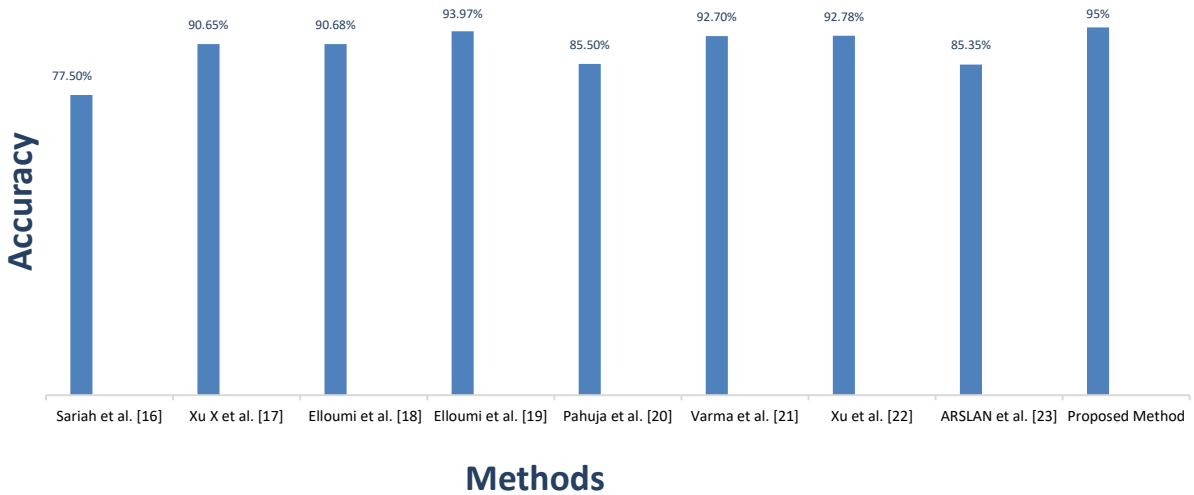


Figure. 10 Comparative analysis with different method

image processing, machine learning, and deep learning, together with information on their methodologies, datasets, and accuracy. From Table 13, Sariah et al. [16], Xu et al. [17], Elloumi et al. [19], proposed deep learning-based approaches with accuracies of 77.50.3%, 90.65%, and 93.97% respectively. Elloumi et al. [19], and Pahuja et al. [20] proposed machine learning techniques with accuracies 90.68% and 85.50% respectively.

5. Conclusion and future work

In this study we have developed a new model to diagnosis cataract diseases automatically. We obtained the combination of texture features analysis: Gray Level Co-occurrence Matrix (GLCM), First Order Statistics/Statistical Features (FOS/SF), Statistical Feature Matrix (SFM), and Neighborhood Gray Tone Difference Matrix (NGTDM). The experimental results on a dataset of 400 fundus images show the performance of the texture-based method. Comparative analysis is performed for the proposed and existing research to identify effective results. Our model outperformed the state-of-the-art cataract detection approaches in terms of accuracy (95%), precision (95.24%), recall (89%), specificity (98.25%), f1-score (91%), and Correlation Coefficient (87.64%). Our results demonstrate that the suggested techniques are effective for dividing the two types of eye conditions: normal and cataract. We plan to apply our approach to a different larger retinal dataset in the future. We intend to carry out more studies in the field of deep learning. To optimize our study, we will also take inspiration from Googles artificial intelligence framework. Furthermore, we want to expand our framework to take into account additional ocular diseases such as aged macular degeneration, glaucoma, and diabetic retinopathy. Additionally, the framework may be included in smartphone apps to function as a mobile-aided cataract grading screening method, greatly expanding the accessibility of eye care. Android-based smartphone the method will be able to identify and categorize three types of cataracts for future study: immature, mature, and hyper maturity.

Conflicts of Interest

Authors declare no conflict of interest.

Author Contributions

The main author of the paper, Esra'a Al sariera, was responsible for the paper planning, data collection, execution. M.C. Padma and Thamer Al Saraiera examined the work, made revision suggestions, and confirmed the results.

References

[1] J. Seddon, D. Fong, S. West, S. K, and C.T. Valmadrid, “Epidemiology of risk factors for age- related cataract”, *Survey of Ophthalmology*, Vol. 39, No. 4, pp.323-334, 1995.

[2] S. Hu, X. Wang, H. Wu, X. Luan, P. Qi, Y. Lin, H. Xiangdong, and H. Wei, “Unified Diagnosis Framework for Automated Nuclear Cataract Grading Based on Smartphone Slit Lamp Images”, *IEEE Access*, Vol. 8, pp. 174169-174178, 2020.

[3] W. Song, Y. Cao, Z. Qiao, Q. Wang, and J. Yang, “An Improved Semi-Supervised Learning Method on Cataract Fundus Image Classification”, In: *Proc. of 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, Vol. 2, pp. 362-367, 2019.

[4] V. Agarwal, V. Gupta, V. M. Vashisht, K. Sharma, and N. Sharma, “Mobile Application Based Cataract Detection System”, In: *Proc. of 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 780-787, 2019.

[5] R. Sigit, M. Kom, M. B. Satmoko, D. K. Basuki, S. Si, and M. Kom, “Classification of Cataract Slit-Lamp Image Based on Machine Learning”, In: *Proc. of 2018 International Seminar on Application for Technology of Information and Communication*, pp. 597-602, 2018.

[6] D. Allen and A. Vasavada, “Cataract and surgery for cataract”, *BMJ*, Vol. 333, No. 7559, pp. 128_132, 2006.

[7] J.-J. Yang, J. Li, R. Shen, Y. Zeng, J. He, J. Bi, Y. Li, Q. Zhang, L. Peng, and Q. Wang, “Exploiting ensemble learning for automatic cataract detection and grading”, *Comput. Methods Programs Biomed*, Vol. 124, pp. 45-57, 2016.

[8] S. Flaxman, R. R. A. Bourne, S. Resnikoff, P. Ackland, T. Braithwaite, M. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg, K. Naidoo, K. Pesudovs, A. Silvester, G. A. Stevens, N. Tahhan, T. Y. Wong, and H. R. Taylor, “Global causes of blindness and distance vision impairment 1990_2020: Asystematic reviewand metaanalysis”, *Lancet Global Health*, Vol. 5, No. 12, pp. 1221-1234, 2017.

[9] L. Cao, H. Li, Y. Zhang, L. Zhang, and L. Xu, “Hierarchical method for cataract grading based

on retinal images using improved Haar wavelet", *Inf. Fusion*, Vol. 53, pp. 196-208, 2020.

[10] X. Han, J. Zhang, Z. Liu, X. Tan, G. Jin, M. H, L. Luo, and Y. Liu, "Real-world visual outcomes of cataract surgery based on population-based studies: a systematic review", *British Journal of Ophthalmology*, Vol. 107, No. 8, pp. 1056- 1065, 2023.

[11] X. Zhang, Z. Xiao, L. Hu, G. Xu, R. Higashita, W. Chen, J. Yin and J. Liu, "CCA-Net: Clinical-awareness attention network for nuclear cataract classification in AS-OCT", *Knowledge-Based Systems*, 250, pp. 109109, 2022.

[12] R. Bourne, S. Flaxman, T. Braithwaite, M. Cicinelli, A. Das, J. Jonas, J.Keeffe, J. Kempen, J. Leasher, H. Limburg, and K. Naidoo, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and nearvision impairment: a systematic review and meta-analysis", *Lancet GlobalHealth*, Vol. 5, No. 9.pp. 888 -897, 2017.

[13] M. Deshpande, "Vision 2020: right to sight-india, Med", *J. Armed Forces India*, Vol. 64, No. 4, pp. 302-303, 2008.

[14] D. Pascolini, S. Mariotti, "Global estimates of visual impairment: 2010, Br", *J. Ophthalmol.* Vol. 96, No. 5, pp. 614-618, 2012.

[15] P. Louis, A. Adenike, F. Timothy, D. Rainaldo, R. Thulasiraj, T. Hugh, F. Hannah, R. Gullapali, K. Ivo, and R. Serge, "Vision 2020: The right to sight: a global initiative to eliminate avoidable blindness", *Arch Ophthalmol*, Vol. 122, No. 4, pp. 615-620, 2004.

[16] M. Sariah, A. Bustamam, and P. Tampubolon, "Cataract classification based on fundus image using an optimized convolution neural network with lookahead optimizer", In: *AIP Conference Proceedings*, Vol. 2296, No. 1, 2020.

[17] X. Xu, J. Li, Y. Guan, L. Zhao, Q. Zhao, L. Zhang, and L. Li, "GLA-net: a global-local attention network for automatic cataract classification", *J Biomed Inform*, 124, pp. 103939, 2021.

[18] Y. Elloumi, "Mobile aided system of deep-learning based cataract grading from fundus images", In: *Proc. of Artificial Intelligence in Medicine: 19th International Conference on Artificial Intelligence in Medicine, AIME 2021, Virtual Event*, pp. 355-360, 2021.

[19] Y. Elloumi, "Cataract grading method based on deep convolutional neural networks and stacking ensemble learning", *International Journal of Imaging Systems and Technology*, Vol. 32, No. 3, pp. 798-814, 2022.

[20] R. Pahuja, U. Sisodia, A. Tiwari, S. Sharma, and P. Nagrath, "A Dynamic Approach of Eye Disease Classification Using Deep Learning and Machine Learning Model", In: *Proc. of Proceed Data Anal Manage, ed: Springer*, pp. 719-736, 2022.

[21] N. Varma, S. Yadav, and J. Yadav, "Fundus image-based automatic cataract detection and grading system", In: *Proc. of AIP Conference Proceedings*, Vol. 2724, No. 1, 2023.

[22] X. Xu, L. Zhao, J. Li, and L. Li, L, "Incorporating medical domain knowledge into data-driven method: A vessel attention guided multi-granularity network for automatic cataract classification", *Expert Systems with Applications*, Vol. 241, pp.122671, 2023.

[23] G. ARSLAN, and C. Erdaş, "Detection of Cataract, Diabetic Retinopathy and Glaucoma Eye Diseases with Deep Learning Approach", *Intelligent Methods in Engineering Sciences*, Vol. 2, No. 2, pp. 42-47, 2023.

[24] Cataract Dataset, 2020. https://www.kaggle.com/andrewmvd/ocular-disease-recognitionodir5k, (Accessed on 22/12/2023).

[25] B. Li, and Q.H. Max, "Texture analysis for ulcer detection in capsule endoscopy images", *Image Vis. Comput*, Vol. 27, No. 9, 2009.

[26] RM. Haralick, K. Shanmugam, I. Dinstein, "Textural features for image classification", *IEEE Trans Syst, Man, Cybern*, Vol. 6, pp. 610-621, 1973.

[27] A. Materka, and M. Strzelecki, "Texture analysis methods-a review", *Brussels: Technical University of Lodz, Institute of Electronics*, Vol. 10, pp. 4968, 1998.

[28] W. Chung-Ming, and Y. Chen Y, "Statistical feature matrix for texture analysis" *Comput Vis Graph Image Process*, Vol. 54, No. 5, pp. 407-419, 1992.

[29] M. Amadasun, and R. King, "Textural features corresponding to textural properties", *IEEE Trans Syst Man Cybern*, Vol. 19, No. 5, pp.1264-74, 1989.

[30] M. Kabir, M. Jahangir, S. Xu, and B. Badhon, "An empirical research on sentiment analysis using machine learning approaches", *Int J Comput Appl*, Vol. 43, No. 10, pp.1011-1019, 2021.

[31] S. Ding, B. Qi, and H. Tan, "An overview on theory and algorithm of support vector machines", *Journal of University of Electronic Science and Technology of China*, Vol. 40, No. 1, pp.2-10, 2011.

[32] Y. Liu, M. Yang, M. Ramsay, L. Coid, and W. Jeremy, "A comparison of logistic regression, classification and regression tree, and neural networks models in predicting violent re-offending", *J Quant Criminol*, Vol. 27, No. 4, pp. 547-553, 2011.
[33] E. Paul and Utgoff, "Incremental induction of decision trees", *Machine learning*, Vol. 4, No. 2, pp. 161-186, 1989.
[34] H. Witten, and E. Frank, "Data mining: practical machine learning tools and techniques with java implementations", *ACM Sigmod Record*, Vol. 31, No. 1, pp.76-77, 2002.