Decision Support System for Age-related Macular Degeneration using Discrete Wavelet Transform

Muthu Rama Krishnan Mookiah* ·
U. Rajendra Acharya · Joel E.W. Koh ·
Chua Kuang Chua · Jen Hong Tan ·
Vinod Chandran · Choo Min Lim · Kevin
Noronha · Augustinus Laude · Louis Tong

Abstract Age-related Macular Degeneration (AMD) affects the central vision and subsequently may lead to visual loss in people over 60 years of age. There is no permanent cure for AMD, but early detection and successive treatment may improve the visual acuity. AMD is mainly classified into dry and wet type, however dry AMD is more common in ageing population. AMD is characterized by drusen, yellow pigmentation and neovascularisation. These lesions are examined through visual inspection of retinal fundus images by ophthalmologists. It is laborious, time consuming and

*Muthu Rama Krishnan Mookiah (Corresponding author) Joel E.W. Koh · Chua Kuang Chua ·
Jen Hong Tan · Choo Min Lim
Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
599489,
E-mail: mkm2@np.edu.sg, kje2@np.edu.sg, ckc@np.edu.sg, tjh6@np.edu.sg, lcm@np.edu.sg,

U. Rajendra Acharya
Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
599489
Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Malaysia
50603,
E-mail: aru@np.edu.sg

Vinod Chandran
School of Electrical Engineering and Computer Science, Queensland University of Technology,
Brisbane QLD 4000, Australia,
E-mail: v.chandran@qut.edu.au

Kevin Noronha
Department of Electronics & TeleCommunication, St Francis Institute of Technology, Mumbai,
India 400103,
E-mail: kevinkurkal@yahoo.co.in

Augustinus Laude
National Healthcare Group Eye Institute, Tan Tock Seng Hospital, Singapore 308433,
E-mail: draugustinuslaude@yahoo.com.sg

Louis Tong
Singapore National Eye Center, Singapore 168751
Ocular Surface Research Group, Singapore Eye Research Institute, Singapore 168751
Duke-NUS Graduate Medical School, Singapore 169857
Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597
E-mail: louis.tong.h.t@snec.com.sg

resource intensive. Hence, in this study we have proposed an automated AMD detection system using Discrete Wavelet Transform (DWT) and feature ranking strategies. The first four order statistical moments (mean, variance, skewness and kurtosis), energy, entropy and Gini index-based features are extracted from DWT coefficients. We have used five (t-test, Kullback-Lieber Divergence (KLD), Chernoff Bound and Bhattacharyya Distance (CBBD), Receiver Operating Characteristics Curve (ROC)-based and Wilcoxon) feature ranking strategies to identify optimal feature set. A set of supervised classifiers namely Support Vector Machine (SVM), Decision Tree (DT), k-Nearest Neighbour (k-NN), Naive Bayes (NB) and Probabilistic Neural Network (PNN) were used to evaluate the highest performance measure using minimum number of features in classifying normal and dry AMD classes. The proposed framework obtained an average accuracy of 93.70%, sensitivity of 91.11% and specificity of 96.30% using KLD ranking and SVM classifier. We have also formulated an AMD Risk Index (AMDRI) using selected features to classify the normal and dry AMD class using one number. The proposed system can be used to assist the clinicians and also for mass AMD screening programs.

Keywords Age-related macular degeneration; Energy; Entropy; Gini index; Feature ranking; Classifier tuning; Computer aided diagnosis

1 Introduction

AMD is one of the most common retinal disease, caused by deterioration of cells in the macula[1]. Macula is the small part of the retina responsible for central vision and color[2]. AMD has several risk factors such as age, hypertension, smoking and family history[3]. It is one of the leading cause of vision loss for people aged 60 years and older[1]. The World Health Organization (WHO) report revealed that 8 million people have affected with severe vision loss due to AMD [1]. Moreover, globally 20-25 million people are affected by AMD[4]. Depending on the presence of clinical signs such as drusen or hyper-pigmentations or small hypo-pigmentations, AMD is mainly categorized into early, intermediate and late stage[1]. The clinical signs, symptoms, diagnostic technique and treatment of AMD are briefly described in Table 1.

Table 1: Stages of AMD[1,3,4].

Clinical signs	Early AMD	Intermediate AMD	Late AMD	
			Dry	Wet
Clinical signs	-Presence of drusen ($63\mu m$ to $124\mu m$ in size) or hyper-pigmentations or small hypo-pigmentations. -Without visible choroidal vessels.	-Presence of minimum one drusen ($124\mu m$ in size) or geographic atrophy not in centre of the macula (Figure 1b).	-Presence of fatty deposits or drusen in centre of the macula.	-Filling of fluid under the retina. -Choroidal neovascularisation (Figure 1c).
Symptoms	-Blurring of central vision, Metamorphopsia and Reduced vision.			
Diagnostic technique	-Opththalmological examination.	-Retinal and choroidal angiography.		
Treatment	-Thermal laser or verteporfin photo-dynamic therapy.			

Early detection of AMD can be achieved by identifying drusen and fatty deposits or exudates using retinal fundus images of the affected patients. Drusen can be mainly classified as hard or soft drusen according to the visibility of the border. Hard drusen have well defined boundaries compared to soft drusen[2]. Most of the previous works have reported that automated drusen segmentation is necessary for AMD classification. Ben Sbeh et al.[5] have proposed maxima and minima-based mathematical morphology for automated drusen segmentation and classified various classes of drusen. Pixel and region-wise classification methods were introduced in[6] to identify drusen. Brandon et al.[7] have proposed wavelet-based drusen segmentation and multi-level pixel classification to discriminated drusen. Rapantzikos et al.[8] employed multilevel histogram equalization to enhance the contrast of the local retinal image structures and Histogram-based Adaptive Local Thresholding (HALT) applied to detect the drusen. Köse et al.[9] proposed an inverse segmentation method to identify the unhealthy regions in the retinal fundus images. Initially, optic disk is detected to determine the macula. Further, the healthy area is chosen around the macula to compute the statistical properties and region growing algorithm applied to segment the healthy area. Finally, inverse segmentation can be performed to identify drusen. Inverse segmentation and statistical properties were used to segment out drusen in[10]. The statistical information of the reference Characteristic Images (CIs) are compared with Sample Images (SIs) to identify the drusen[10]. Multi-scale analysis using Mexican hat wavelet and Support Vector Data Description (SVDD) are described for automated detection of AMD anomaly in[11]. Santos-Villalobos et al.[12] proposed a statistical method to segment drusen. Their method assumes the class conditional density from color features of the drusen lesions. Liang et al.[13] proposed a maximal region-based pixel intensity method to detect drusen. Their method identified local regions with high intensity pixels and fixed threshold value to segment the drusen.
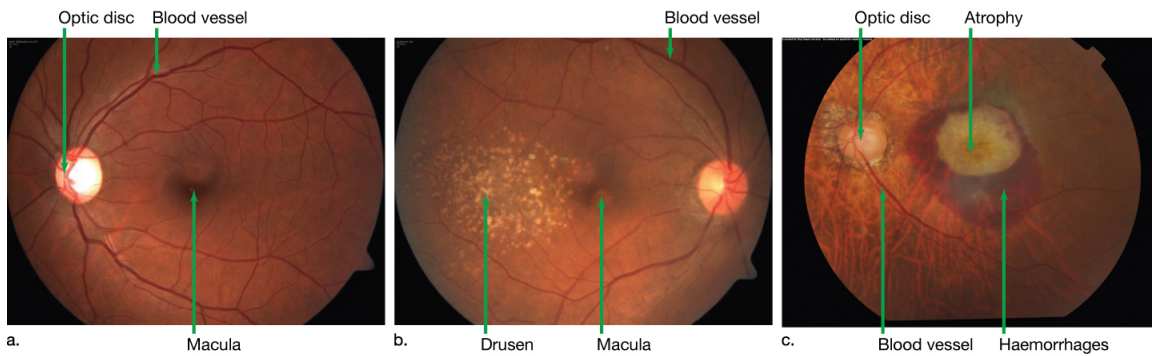
Fig. 1: Typtical fundus images: (a) Normal; (b) Dry AMD; (c) Wet AMD.

All above mentioned work mainly focuses on the segmentation or identification of drusen, rather than AMD screening. However, very few work [2,14–18] reported the application of image analysis and machine learning methods for automated AMD diagnosis. Cheng et al. [14] proposed automated AMD detection using Biologically Inspired Features (BIF) based on Gabor filters. Independent Component Analysis (ICA) is used to extract different drusen phenotypes to perform automated AMD screening in [15]. Amplitude Modulation (AM)-Frequency Modulation (FM)-based multi-scale features are computed in [16,17,19] to classify pathological structures (drusen) for automated AMD screening. Hijazi et al. [18] proposed Case Based Reasoning (CBR) and Dynamic Time Warping (DTW) approaches to quantify the similarity between normal and AMD image histogram to identify AMD. Also Hijazi et al. [2] have compared the performance of spatial histograms and hierarchical decomposition methods for automated screening of AMD. Zheng et al. [20] have used quadtree decomposition and Weighted Frequent SubGraph (WFSG) mining for automated detection of AMD.

The aforementioned approaches, except [2,18] require drusen segmentation, but it is a challenging task due to poor visual appearance of drusen and associated lesions [2]. Hence, we propose a method to classify normal and dry AMD retinal fundus images using energy and entropy features of DWT coefficients without using drusen segmentation. The overall block diagram of the proposed approach is shown in Figure 2. Initially, the fundus images are subjected to preprocessing to enhance the contrast using adaptive histogram equalization. Further, we have performed three level wavelet decomposition to obtain fine and coarse changes in the form of coefficient matrices. We have computed 184 features such as energy, entropy, relative wavelet energy & entropy, probability of energy, entropy, statistical moments and Gini index from the approximate, horizontal, vertical and diagonal coefficients. All these features were fed to ttest, KLD, CBBD, ROC-based and Wilcoxon ranking methods to rank the features. Subsequently, these ranked features are fed to the supervised classifiers (SVM, DT, k-NN, NB and PNN) to discriminate normal and AMD classes using minimum number of features. We have employed 10-fold cross validation for data resampling. During training phase the proposed system uses significant features which are extracted from preprocessed normal and dry AMD fundus images. Once the system is trained, the trained model classifies the test set or unknown image into either normal or dry AMD class automatically using

the significant features are extracted from the test set or unknown image. Hence, the system does not require any user intervention.

Fig. 2: The proposed automated AMD detection system.

The paper is organized as follows: Materials and methods such as study subject selection and fundus image acquisition, preprocessing, DWT method, feature extraction, feature ranking, selection, AMD Risk Index and AMD classification are discussed in Section 2. The results of the proposed system is presented in Section 3. The obtained results are described in Section 4 and the paper is concluded in Section 5.

2 Materials and methods

This section describes with the details of fundus image acquisition, contrast enhancement of the acquired fundus images using Contrast Limited Adaptive Histogram Equalisation (CLAHE), feature extraction using DWT, choosing the best feature ranking method, AMD risk index computation and comparison of various supervised classifiers to discriminate normal and AMD classes.

2.1 Study subject selection and fundus image acquisition

All the participants in this study were screened and clinically confirmed for AMD at the Department of Ophthalmology, Kasturba Medical College, Manipal, India. After signing an informed consent, a total of 135 patients with AMD and 135 normal subjects were participated in this study. The age group of subjects ranged from 50 to 80 years. The images were acquired from both the eyes using TOPCON non-mydriatic retinal camera (TRC-NW200) from Department of Ophthalmology, Kasturba Medical College, Manipal, India. The images were acquired and labelled by the clinicians. The institute ethics committee has approved the images for research purpose. Several images are taken from each study subjects, however in this work five hundred and forty images (Normal-270 and AMD-270) are used and all these images are stored in a 24-bit uncompressed JPEG format. The resolution of the image is 480×364 pixels.

2.2 Preprocessing

The contrast enhancement of the retinal fundus images are performed using CLAHE[21]. Initially, the green band was separated from the Red Green Blue (RGB) color fundus image. Further, CLAHE is applied on the green band image. This method separates

the images into contextual regions and applies adaptive histogram equalization on the separated regions. Hence, each pixel intensity in the image is mapped to a new intensity value within the display range which is proportional to the pixel intensity rank in the local intensity histogram[21]. This method stretches the grayscale intensity distribution to the whole image which makes the image features such as blood vessels, Optic Disk (OD), macula, and drusen are more visible[21].

2.3 Discrete Wavelet Transform method

DWT analyses a temporal signal in the scale (or frequency) and time dimensions[22]. It is used in several automated disease detection systems such as cardiac arrhythmias[23], epilepsy[24] and diabetic neuropathy[25]. It decomposes an image in scale and space dimensions (not time). It is capable of capturing both high frequency (detail) and low frequency (approximate) information and is better at capturing transient or localized features than the Fast Fourier Transform (FFT). The transfer functions H_L and H_H are used to define the low pass and high pass filters respectively. The high pass filter output contains detail coefficients and low pass filter output contains approximate coefficients. The detail coefficients ($D[n]$) are obtained using the following equation

$$D[n] = \sum_{k=-\infty}^{\infty} x[k]H_H[2n-k]$$

(1)

The approximate coefficients ($A[n]$) are obtained using the following equation

$$A[n] = \sum_{k=-\infty}^{\infty} x[k]H_L[2n-k]$$

(2)

Where $x[k]$ is original input signal, k is an integer and n is denoted as sampling factor.

The frequency resolution is further increased by cascading the low and high pass filter operations where the first level low pass filter output is fed into the same low and high pass filter combination[22]. The above decomposition is repeated recursively and it is called as Mallat-tree decomposition[26] (Figure 3) shows the two dimensional DWT structure[22].
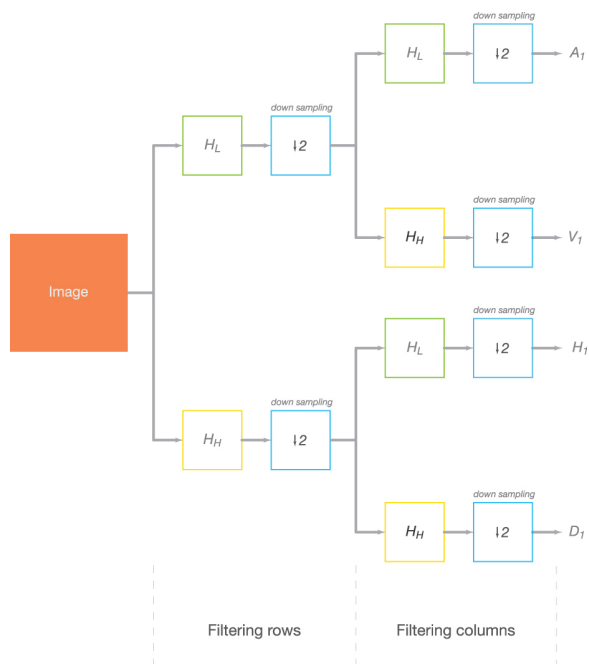
Fig. 3: Two-dimensional DWT decomposition structure (one-level).

In this work, the normal and dry AMD fundus images are subjected to three-level DWT decomposition using bi-orthogonal 3.7 mother wavelet[27]. Initially, rows are fed to the low pass and high pass filters to obtain H_L and H_H and these down sampled images along the columns were fed to both low pass and high pass filters to obtain approximation (A_1), horizontal (H_1), vertical (V_1) and diagonal (D_1) coefficients respectively. Figure 4 shows the result of 2D three level DWT decomposition on normal and dry AMD fundus images. It can be seen from the figure that, the subtle changes are visible in the high frequency sub-bands of dry AMD images.

(a)

(b)
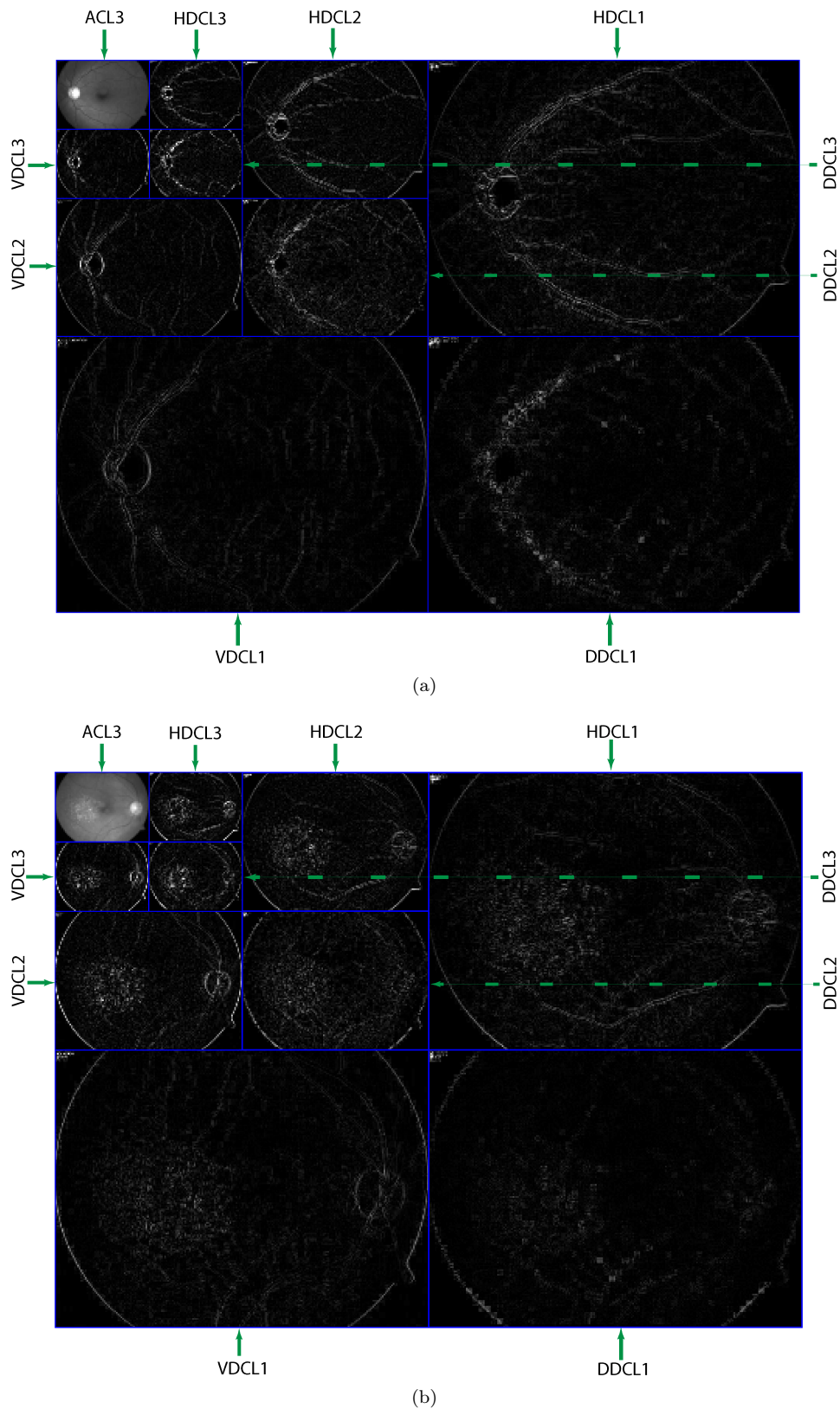
Fig. 4: Discrete Wavelet Transform decomposition of fundus image (a) Normal; (b) Dry AMD.
(ACL3: Approximation coefficient at level 3; HDCL1: Horizontal detail coefficient at level 1; HDCL2: Horizontal detail coefficient at level 2; HDCL3: Horizontal detail coefficient at level 3; VDCL1: Vertical detail coefficient at level 1; VDCL2: Vertical detail coefficient at level 2; VDCL3: Vertical detail coefficient at level 3; DDCL1: Diagonal detail coefficient at level 1; DDCL2: Diagonal detail coefficient at level 2; DDCL3: Diagonal detail coefficient at level 3)

2.4 Feature extraction

Features such as mean, variance, skewness. kurtosis, Shannon entropy, Renyi entropy, Kapur entropy, relative energy, relative entropy, probability of energy, entropy and Gini index are computed from the wavelet coefficients. These features are briefly described below. The list of symbols and mathematical notations used in the following equations are explained in the Table 2.

Table 2: List of symbols and mathematical notations.

Symbol or mathematical notation	Description
$f(x,y)$	Probability density function
M and N	Rows and columns of wavelet coefficients
m	*mean* value of wavelet coefficients
σ	Standard deviation
C_N^k	C is either horizontal (H), vertical (V), diagonal (D) or approximate (A) coefficients; N is the number of levels of decomposition (in this work 3); k is the number of coefficients at each level of decomposition
$Eng_N^{C^k}$	Energy of either approximate, horizontal, vertical or diagonal coefficients
$Eng_N^{A^k}$	Energy of the approximate coefficients
$Eng_N^{H^k}$	Energy of the horizontal coefficients
$Eng_N^{V^k}$	Energy of the vertical coefficients
$Eng_N^{D^k}$	Energy of the diagonal coefficients
$Ent_N^{C^k}$	Entropy of either approximate, horizontal, vertical or diagonal coefficients
$Ent_N^{A^k}$	Entropy of the approximate coefficients
$Ent_N^{H^k}$	Entropy of the horizontal coefficients
$Ent_N^{V^k}$	Entropy of the vertical coefficients
$Ent_N^{D^k}$	Entropy of the diagonal coefficients
α and β	Diversity indices
c	Ordered sparsity vector
$\|c\|_1$	Manhattan distance of vector c
i	Integer
V_{Energy}	Energy of vertical coefficients
PRW_{Energy}	Probability of energy and relative wavelet energy
H_{Energy}	Energy of horizontal coefficients
$\delta_{Entropy}$	Probability of entropy and relative wavelet entropy
λ	Bias
$CV2Eng$	Energy of vertical coefficient at level 2
$CV2AvgEng$	Average energy of vertical coefficient at level 2

Continued on next page

Table 2 – *Continued from previous page*

Symbol or mathematical notation	Description
$CV1Eng$	Energy of vertical coefficient at level 1
$CH1Eng$	Energy of horizontal coefficient at level 1
$CH3Eng$	Energy of horizontal coefficient at level 3
$CH3AvgEng$	Average energy of horizontal coefficient at level 3
$PH1Eng$	Probability of horizontal coefficient energy at level 1
$CH1RWEng$	Relative wavelet energy of horizontal coefficient at level 1
$PH3Eng$	Probability of horizontal coefficient energy at level 3
$CH3RWEng$	Relative wavelet energy of horizontal coefficient at level 3
$ARWEng$	Average relative wavelet energy
$CH2RWEnt$	Relative wavelet entropy of horizontal coefficient at level 2
$PH2Ent$	Probability of horizontal coefficient entropy at level 2

2.4.1 Statistical moments

Statistical moments are used to define the probability density function of the wavelet coefficients. The first four order statistical moments [26] are computed using below equations:

$$\text{Mean} = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \{f(x,y)\}}{M \times N} \tag{3}$$

$$\text{Variance} = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \{f(x,y) - m\}^2}{M \times N} \tag{4}$$

Mean and variance provides the information on the spread or scale of the data distribution [28,29].

$$\text{Skewness} = \frac{1}{M \times N} \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \{f(x,y) - m\}^3}{\sigma^3} \tag{5}$$

Skewness is mainly used to measure the shape of a distribution, whether it follows symmetry or not. If the value is zero the distribution is symmetry[28,29].

$$\text{Kurtosis} = \frac{1}{M \times N} \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \{f(x,y) - m\}^4}{\sigma^4}$$

(6)

Kurtosis is mainly used to measure the flatness of the distribution[28,29].

2.4.2 Energy and entropy features

Energy is used to define the amount of information present in the image and entropy is used to measure the uncertainty associated with randomness[30]. In this work, relative wavelet energy (RW_{Energy}), relative wavelet entropy ($RW_{Entropy}$), probability of energy (P_{Energy}) and entropy ($P_{Entropy}$) features were computed on the wavelet coefficients[31]. It can be defined as follows

$$RW_{Energy}(C_N^k) = \frac{\sum_N Eng_N^{C^k}}{\sum_N Eng_N^{A^k} + \sum_N Eng_N^{H^k} + \sum_N Eng_N^{V^k} + \sum_N Eng_N^{D^k}}$$

(7)

$$RW_{Entropy}(C_N^k) = \frac{\sum_N Ent_N^{C^k}}{\sum_N Ent_N^{A^k} + \sum_N Ent_N^{H^k} + \sum_N Ent_N^{V^k} + \sum_N Ent_N^{D^k}}$$

(8)

$$P_{Energy}(C_N^k) = \frac{Eng_N^{C^k}}{[\sum_N Eng_N^{A^k}] + [\sum (Eng_N^{H^k} + Eng_N^{V^k} + Eng_N^{D^k})]}$$

(9)

$$P_{Entropy}(C_N^k) = \frac{Ent_N^{C^k}}{[\sum_N Ent_N^{A^k}] + [\sum (Ent_N^{H^k} + Ent_N^{V^k} + Ent_N^{D^k})]}$$

(10)

where $Eng_N^{C^k} = \sum_k |C_N^k|^2$; $Ent_N^{C^k} = -\sum_k (C_N^k) log_2(C_N^k)$

In addition to above mentioned features, Shannon, Renyi and Kapur entropies were also computed to measure the uncertainty in the wavelet coefficients. The basic idea of

the Shannon entropy is the information gain of an event and is inversely related to its probability of occurrence[30,32]. It can be computed using the following formula[30,32]

$$\text{Shannon entropy} = -\sum_{i=1}^{k}(C_N)log_2(C_N)$$

(11)

Renyi and Kapur entropy have high dynamic range in scattering conditions compared to Shannon entropy. It can be computed[30,32] as follows:

$$\text{Renyi entropy} = \frac{1}{1-\alpha}log\sum_{i=0}^{k}C_N^{\alpha}, \text{ For } \alpha \neq 1, \alpha > 0$$

(12)

Here, we have chosen $\alpha = 3$.

The Kapur entropy is further generalized version of Renyi entropy[30,32] and is given by:

$$\text{Kapur entropy} = \frac{1}{1-\alpha}ln\frac{\sum_{i=0}^{k}C_N^{\alpha}}{\sum_{i=0}^{k}C_N^{\beta}}, \text{ For } \alpha \neq \beta, \alpha + \beta - 1 > 0, \beta > 0$$

(13)

Here, we have chosen $\alpha = 0.5$ and $\beta = 0.7$.

2.4.3 Gini index

Gini index is used to measure the dispersion of wavelet coefficients[33]. It is computed using the formula:

$$G = 1 - 2\sum_{i=1}^{k}\frac{c_i}{\|c\|_1}(\frac{k-i+0.5}{k})$$

(14)

Using above mentioned feature extraction methods, total 184 features are extracted.

2.5 Feature ranking and selection

Feature ranking step is used to select the minimum number of best possible features which can discriminates the two classes[34]. In this work we have compared five feature ranking methods namely t-test, KLD, CBBD, ROC-based and Wilcoxon ranking methods. t-test compares the population *means* of the two groups to identify the correlation among the features[35]. KLD rank the features are based on summation of the divergence measure for each features between normal and dry AMD classes. It has a

direct relation with the Bayes error[36]. Chernoff bound provides exponential decay on tail distributions using independent variables. Hence it is used with Bhattacharyya distance for feature ranking[37]. ROC-based feature ranking selects the features using area between ROC and the random classifier slope[37]. Wilcoxon method is a non-parametric test, which does not assume normal distribution. The features are ranked based on U statistic[35].

2.6 AMD Risk Index

In addition to automated classification we have formed single-valued index using a combination of ranked features (See Figure 5) extracted from 270 normal and 270 dry AMD images. This index is computed using Eq. (15) such that the combination of the ranked features resulted in unique range for normal and dry AMD class with maximum separation. Since we noticed a uniform pattern in the selected features (See Table 3) between normal and dry AMD, we have used selected these features to formulate AMDRI. In the clinical context, a threshold can be set for these index to enable objective discrimination by the doctors as to whether the input retinal fundus image is belongs to normal or dry AMD class. The range of index may provide the information about the classes. AMDRI is computed using the following formula.

Table 3: Summary statistics of features used for AMDRI formulation.

Features	Normal (Mean± Standard Deviation (SD))	Dry AMD (Mean±SD)	p-value
$CV2Eng$	$(1.59 \pm 1.19)10^7$	$(8.83 \pm 5.24)10^6$	<0.0001
$CV2AvgEng$	$(7.97 \pm 5.96)10^6$	$(4.41 \pm 2.62)10^6$	<0.0001
$CV1Eng$	$(1.36 \pm 0.62)10^6$	$(9.03 \pm 3.99)10^5$	<0.0001
$CH1Eng$	$(1.67 \pm 1.04)10^6$	$(9.00 \pm 3.78)10^5$	<0.0001
$CH3Eng$	$(5.31 \pm 2.83)10^7$	$(3.16 \pm 1.60)10^7$	<0.0001
$CH3AvgEng$	$(1.77 \pm 0.94)10^7$	$(1.05 \pm 0.53)10^7$	<0.0001
$PH1Eng$	0.001 ± 0.0005	0.0006 ± 0.0002	<0.0001
$CH1RWEng$	0.0003 ± 0.0002	0.0002 ± 0.0001	<0.0001
$PH3Eng$	0.0254 ± 0.0096	0.0164 ± 0.0063	<0.0001
$CH3RWEng$	0.0095 ± 0.0038	0.0061 ± 0.0024	<0.0001
$ARWEng$	0.9703 ± 0.0104	0.9790 ± 0.0082	<0.0001
$CH2RWEnt$	0.0610 ± 0.0010	0.0591 ± 0.0022	<0.0001
$PH2Ent$	0.2103 ± 0.0029	0.2039 ± 0.0069	<0.0001

$$\text{AMDRI} = \frac{(V_{Energy} \times (PRW_{Energy} \times H_{Energy}))}{\delta_{Entropy} \times \lambda}$$

(15)

where

$$V_{Energy} = CV2Eng + CV2AvgEng + CV1Eng$$

$$H_{Energy} = CH1Eng + CH3Eng + CH3AvgEng$$
$$PRW_{Energy} = (PH1Eng \times CH1RWEng \times PH3Eng \times CH3RWEng) + (ARWEng)$$
$$\delta_{Entropy} = CH2RWEnt + PH2Ent$$
$$\lambda = 10^{15}$$

2.7 Classification

In this work, we have used five supervised classifiers (NB, DT, k-NN, PNN and SVM) to compare the performance of various ranking schemes. The NB is a simple classifier based on Bayes theorem with independent assumptions between predictors. The posterior probability is calculated by building frequency table for each feature against the class. Then, the frequency table is transformed into likelihood table to calculate posterior probability for each class. The class with highest posterior probability is the desired output[37]. In the DT, complex decision is divided into a union of several simpler decisions, that are derived from the training data[38]. k-NN is a simple, non-parametric lazy learning algorithm. It performs classification by computing distance between k-nearest neighbours, where majority vote of its neighbours decides the class. In this work, we have chosen $k = 1$[37]. PNN is a feed forward neural network developed using probability density function based on Parzens' result. It is a three layer (input, pattern and summation) network, where summation layer *compete* transfer function decides the output using distance vector probabilities[39]. The spread (σ) value of the summation layer is obtained using bootstrap[40] approach (varying σ value from 0.01 to 1 with the increment of 0.01). SVM is developed based on statistical learning and uses the essence of structural risk minimization principle. The basic idea of SVM is to find an optimal hyperplane in the high dimensional feature space that can separate the data in the best possible way[37,41]. In this work, we have evaluated various kernels of SVM such as linear, quadratic, Radial Basis Function (RBF) and polynomial. The RBF kernel width (σ) is varied between 0.1 to 5 with the increment of 0.1 using bootstrap[40] method. The optimal spread (σ) value and RBF kernel width (σ) for PNN and SVM classifiers are 0.23 and 4.2 respectively.

3 Results

In this work, 184 features are extracted using statistical moments, wavelet energy, entropy and gini index from the wavelet coefficients. Statistical moments captures the shape variations in the images[26]. Energy and entropy measures quantify the uniformity and randomness in the images[31]. Gini index measures the dispersion of the wavelet coefficients[33], which reflects the pixel pattern variation in the normal and dry AMD. The statistical summary ($Mean \pm$ SD) of features is shown in Figure 5(a)–(d).
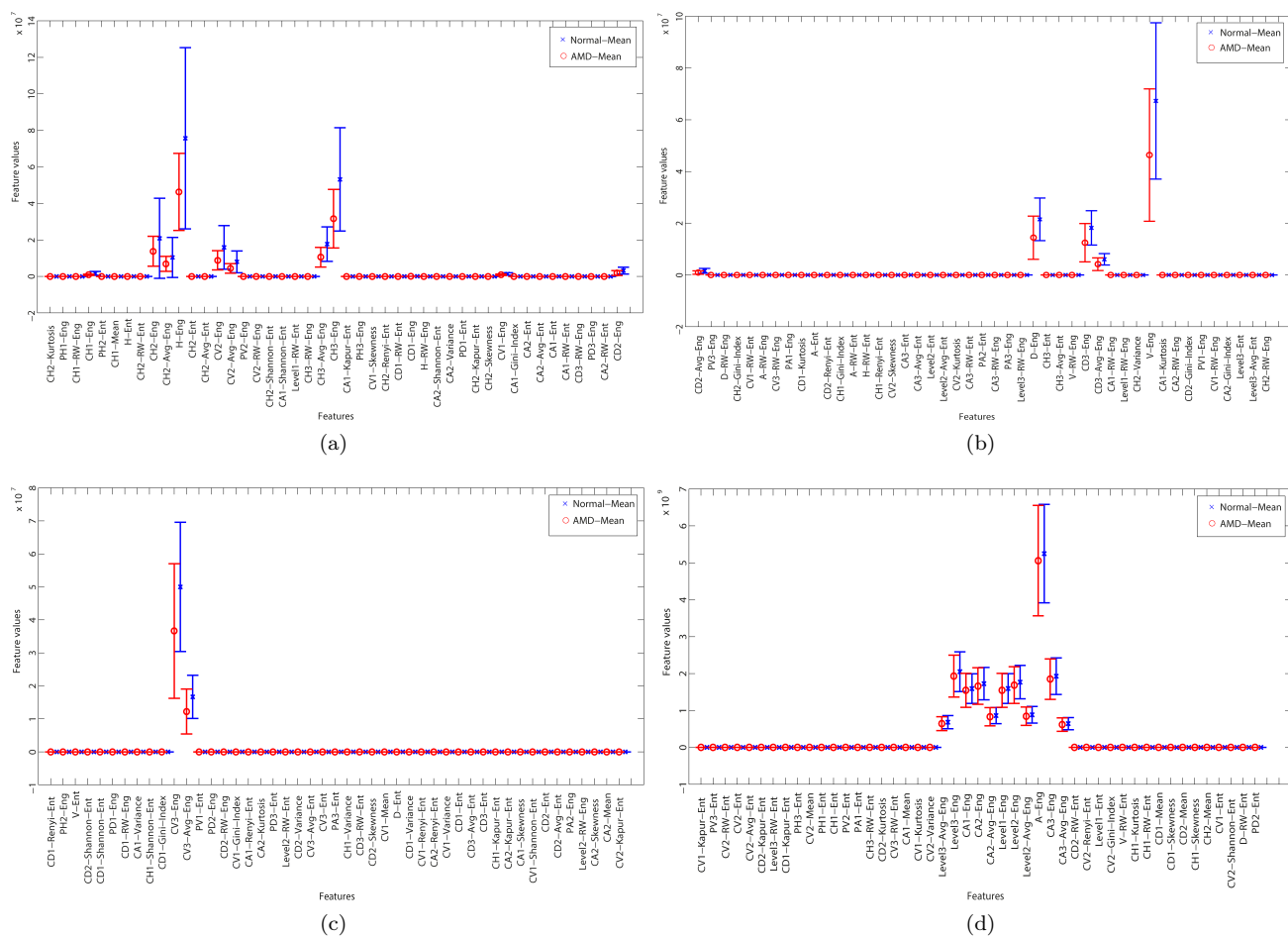
Fig. 5: (a)–(d) Statistical variation (mean ± standard deviation) of features for normal and dry AMD classes.

The results show that the *mean* values of all features are low for dry AMD and high for normal. This is due to the presence of drusen and yellow pigmentation in the dry AMD images that changes the pixel patterns. These subtle changes are clearly captured by the extracted features (Figure 6 and 7).
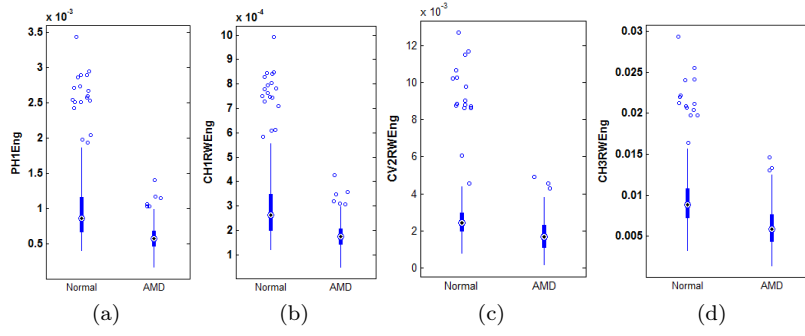
Fig. 6: Box plots: (a) $PH1_{Energy}$; (b) $CH1RW_{Energy}$; (c) $CV2RW_{Energy}$; (d) $CH3RW_{Energy}$

The box plots (Figure 6) also show that the energy values are low for dry AMD and high for normal. The *medians* of the box plots are significantly different for normal and dry AMD classes.
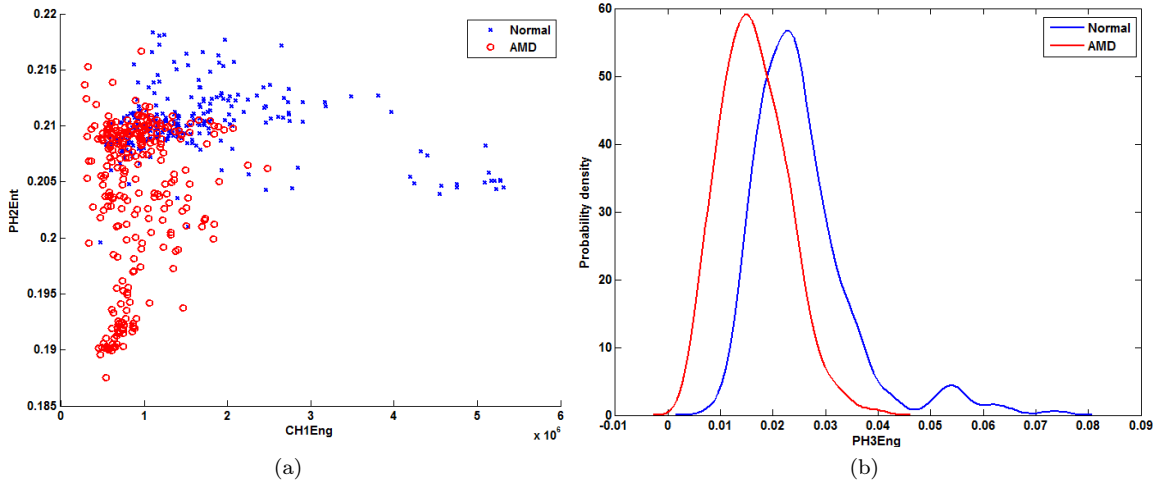

Fig. 7: (a) Scatter plot of features $CH1_{Energy}$ versus $PH2_{Entropy}$; (b) Density estimation plot of $PH3_{Energy}$ for normal and dry AMD classes.

The group scatter and density estimation plots (Figure 7) show that the features are distinct for normal and dry AMD classes. All these plots reveal that the extracted features can discriminate the normal and dry AMD classes with higher accuracy.

The extracted features are ranked using five ranking methods (ttest, KLD, CBBD, ROC-based and Wilcoxon) to identify the optimal feature set for the classification

task. The ranked features are fed to NB, DT, k-NN, PNN and SVM classifiers using bootstrapping[40] and each classifier was evaluated using 10-fold cross validation to measure the accuracy. Finally, the optimal feature set and ranking method are identified using accuracy measure where the ranking method should maximize the accuracy using minimum number of features (See Figure 8). Among these five ranking methods KLD method used 121 optimum features to achieve highest average accuracy of 93.70%. The performance of the different ranking methods for different combination of features is shown in Figure 8. The performance measures like average accuracy, sensitivity, specificity and Positive Predictive Value (PPV) are tabulated in Table 4.

Table 4: Performance measures of various classifiers (features are ranked using KLD ranking method).

Classifier	Number of features used	Accuracy (%)	PPV (%)	Sensitivity (%)	Specificity (%)
NB	63	83.70	82.88	85.19	82.22
DT	55	84.44	84.70	84.81	84.07
k-NN	117	87.78	89.28	85.93	89.63
PNN	123	89.07	96.41	81.11	97.04
SVM-Linear	**121**	**93.70**	**96.09**	**91.11**	**96.30**
SVM-Quadratic	30	88.52	89.32	87.78	89.26
SVM-Polynomial	11	86.11	87.92	84.07	88.15
SVM-RBF	121	91.85	91.45	92.22	91.48

The classification results (Table 4 and Figure 8) show that SVM linear kernel obtained highest accuracy (93.70%) using 121 features ranked by KLD ranking method.
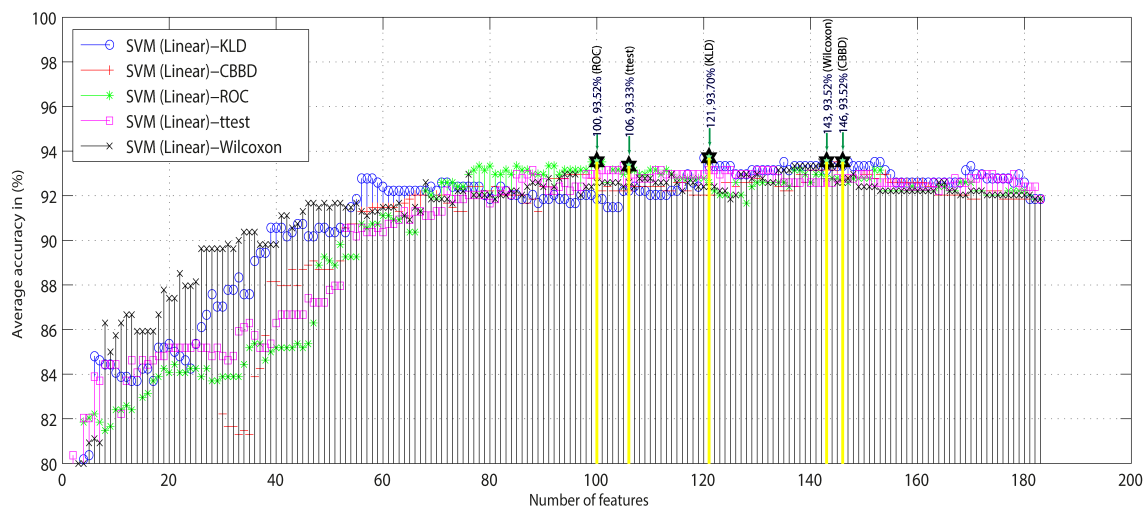

Fig. 8: Plot of average accuracy versus number of features for various ranking methods using SVM classifier.

In this work, the SVM-RBF kernel and PNN classifier parameters are tuned using bootstrap[40] method to obtain better performance. SVM-RBF kernel yielded the highest accuracy (91.85%) with $\sigma = 4.9$ and PNN obtained highest accuracy (89.07%) with $\sigma = 0.32$ (Figure 9) using 121 and 123 ranked features respectively(Table 4).

(a)   (b)
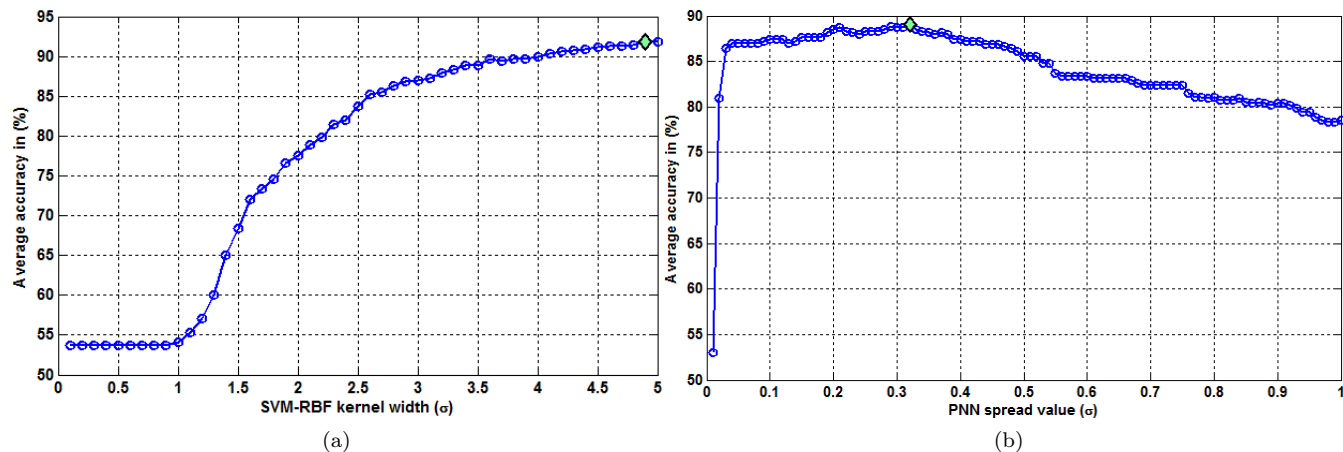
Fig. 9: Classifier parameter (σ) tuning : (a) SVM-RBF kernel; (b) PNN classifier

In addition to classification, we have formulated AMDRI to discriminate normal and dry AMD classes. The range ($mean \pm$SD) of AMDRI for the normal and dry AMD classes are 6.539 ± 1.911 and 2.264 ± 0.162 respectively. The values are significantly different for both the classes (Figure 10), it is low for dry AMD and high for normal.
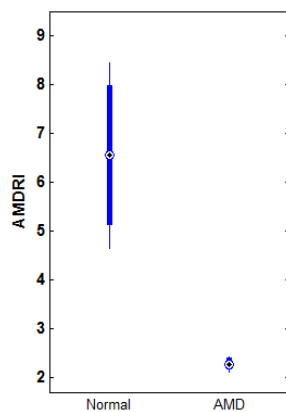

Fig. 10: Box plot of AMDRI.

4 Discussion

Automated retinal fundus image processing and machine learning methods help in early detection of Diabetic Retinopathy (DR), glaucoma and AMD [42,43]. In this work, features such as statistical moments, energy, entropy and Gini index are extracted from wavelet coefficients. DWT decomposes the images in both rows and columns using low and high pass filter. Hence, it can capture local and global features in the normal and dry AMD fundus images. This DWT property [44] captures the morphological structures of fundus images and also captures the drusen, exudates and neovascularization [45]. Hence, the extracted features are robust (See Figure 6 and 7) and resulted in highest classification accuracy (See Table 4).

The results (See Table 4) show that SVM linear kernel yielded highest accuracy (93.70%) with 121 features compared to other kernels and classifiers. However, other kernel functions such as quadratic, polynomial and RBF provided average accuracies of 88.52%, 86.11% and 91.85% using 30, 11 and 121 features respectively. The difference in accuracies is approximately 4%-7%-2% compared with linear kernel using minimum number of features. It shows that the complex kernels convert the non-linearly separable features into linearly separable one [46]. Other classifiers such as NB, DT, k-NN and PNN obtained accuracies of 83.70%, 84.44%, 87.78% and 89.07% respectively. Hence, we can infer that the linear kernel achieved highest accuracy in classifying normal and dry AMD classes. We have compared our results with existing automated AMD detection literatures in Table 5 and are briefly summarized as follows.

Murray et al. [16] have developed AM-FM-based methods to characterize AMD lesions in retinal fundus images. They used four-scale filterbank and AM-FM-based estimates to characterize the retinal images having soft and hard drusen. Their method are tested using both synthetic and AMD retinal images. However, the accuracy of their method is not reported. Their method can be used to encode the AMD lesions with certain frequency component.

Soliz et al. [15] have presented a novel approach for extracting image-based features to discriminate AMD in retinal fundus images. Initially, the images are categorized into 12 drusen phenotypes. The features are extracted from these phenotypes using ICA. Further, the phenotypes are classified by partial least square regression using ICA derived features. Their approach achieved 100% accuracy using 12 images. To train their algorithm, independent components of drusen needs to be selected using Region Of Interest (ROI) at different scales.

Barriga et al. [17] have proposed AM-FM-based multi-scale features to characterize intensity variations in the normal and AMD images. The dimension of the AM-FM features are reduced using Principle Component Analysis (PCA) and obtained the classification accuracy of more than 90% using Mahalanobis distance. Their method encodes the AMD lesions with certain frequency component and train their algorithm hard and soft drusen, vessels and background needs to be selected by the trained grader.

Hijazi et al. [18] have proposed two stage CBR technique which is applied to histogram of the retinal image to detect AMD. Their approach involves two stages: (i) case base generation and (ii) classification of unknown image. Their algorithm achieved a mean sensitivity of 82%, specificity of 65% and accuracy of 75% compared with graders observation. The anatomical structures such as OD and retinal blood vessels needs to be removed before applying their method.

The Instantaneous Frequency (IF) and Instantaneous Amplitude (IA) of AM-FM is used in [19] for automated DR and AMD detection. Initially, the green channel is

separated from RGB color image and 140×140 pixels ROI is chosen for AM-FM decomposition. The features such as statistical moments and histogram percentiles are extracted from the decomposed images. The extracted features are fed to partial least square method the discriminate DR and AMD. Their algorithm is tested on two datasets namely Retina Institute of South Texas (RIST) and University of Texas Health Science Centre in San Antonio (UTHSCSA) and obtained Area Under receiver operator characteristics Curve (AUC) of 0.84 and 0.77 respectively. Their method encodes the AMD lesions such as drusen, abnormal pigmentation and Geographic Atrophy (GA) with certain frequency component.

Cheng et al.[14] have proposed an automated AMD detection system using BIF. Initially, centre of the macula is located using particle tracing algorithm. Further, BIF features are computed around the centre using 5×5 grid. These features are fed to the SVM classifier to diagnose AMD. Their method obtained a sensitivity of 86.3% and specificity of 91.9%. Before applying their method OD, blood vessels and fovea needs to be segmented.

Hijazi et al.[2] have presented spatial histograms and hierarchical image decomposition methods to differentiate normal and AMD images. Spatial histogram method generates signature histogram using shape and texture information of the normal and AMD images. These histogram features are fed to CBR and DTW approaches to classify normal and AMD classes. Hierarchical decomposition provides collection of trees and frequently occurring sub-trees are identified using weighted frequent sub-graph mining algorithm. Further, these trees were fed to SVM classifier to discriminate normal and AMD classes. Their method yielded an accuracy of 74% and 100% using spatial histograms and hierarchical image decomposition methods respectively. Authors have improved the classification accuracy by removing the retinal anatomy such as OD and blood vessels. Also, their method do not require to locate and segment drusen.

Table 5: Summary of literatures on automated AMD detection.

Authors	Method	Classifier (Image data size)	Detected feature	Advantages/ Disadvantages	Performance measure
Murray et al. (2008)[16]	Four scale filter bank AM-FM estimates	Not Applicable (Not Mentioned)	Instantaneous amplitude and frequency	–Encodes AMD lesions with certain frequency component	Not Mentioned
Soliz et al. (2008)[15]	ICA	Linear regression (12)	Independent Components	–Independent components of drusen needs to be selected using ROI for training	Accuracy-100%
Barriga et al. (2009)[17]	AM-FM	Linear regression (5)	Instantaneous amplitude and frequency	–Encodes AMD lesions with certain frequency component –Manual grading of lesions and anatomical structures needed for training	Accuracy-90%
Hijazi et al. (2010)[18]	DTW	CBR (144)	Histogram features	–Isolation and segmentation of drusen is not required –OD and retinal blood vessels needs to be removed before applying DTW	Sensitivity-82% Specificity-65% Accuracy-75%

Table 5 – *Continued from previous page*

Authors	Method	Classifier (Image data size)	Detected feature	Advantages/ Disadvantages	Performance measure
Agurto et al. (2011)[19]	AM-FM	Partial least square (507)	Statistical moments and histogram percentiles	–Encodes lesions such as drusen, abnormal pigmentation and GA with certain frequency component	AUC-0.84 (RIST) AUC-0.77 (UTHSCSA)
Cheng et al. (2012)[14]	Dyadic Gaussian pyramids convolution	SVM (350)	Biologically inspired features	–Segmentation of OD, blood vessels and fovea needed	Sensitivity-86.3% Specificity-91.9%
Hijazi et al. (2012)[2]	Spatial histogram and hierarchical image decomposition	CBR and DTW (161)	Image feature	–Isolation and segmentation of drusen is not required –Removal of OD and blood vessels are needed to improve the accuracy	Spatial histogram accuracy-74%; Hierarchical decomposition accuracy-100%
Zheng et al. (2012)[20]	Quadtrees and WFSG mining	SVM and NB (258)	Image feature	–Removal of blood vessels is required to improve the accuracy	Accuracy-99.6% (SVM) Accuracy-78.7% (NB)
The current study	DWT	SVM (540)	Statistical moments, energy, entropy and Gini index on wavelet coefficients	–Segmentation and removal of retinal anatomy such as OD, blood vessels and fovea are not needed –Isolation and segmentation of drusen is not required –The method was evaluated using 10-fold cross validation –Extracted features were evaluated using five ranking methods –Five supervised classifiers were used to compare the discrimination ability of the extracted features	Accuracy-93.70% Sensitivity-91.11% Specificity-96.30% PPV-96.09%

The WFSG mining algorithm is used in[20] to develop automated AMD grading system. During preprocessing the gray level images are subjected to histogram specification and equalization to normalize the color and contrast. Further, quadtree-based image decomposition is used to extract features. These features are fed to SVM and NB classifiers to discriminate normal and AMD classes. Their method yielded an accuracy of 99.6% and 78.7% using SVM and NB classifiers respectively. The classification accuracy of their method can be improved by removing the blood vessel.

The salient features of the proposed framework are summarized below :

1. The proposed system does not require any segmentation such as retinal anatomy and abnormal lesions.

2. Compared to available literatures reported in Table 5 our AMD detection system shows the highest average accuracy of 93.70%, sensitivity of 91.11% and specificity of 96.30% using 540 images.
3. The proposed system is evaluated over 10-fold cross validation.
4. The significance of the features are evaluated using five ranking methods and the discrimination ability is tested using five supervised classifiers, hence the obtained results are robust and reproducible.
5. In this work, we used multi-resolution analysis for feature extraction and hence captures subtle variations in the image pixels and provided the highest accuracy (See Table 4 and 5).
6. The devised integrated index (AMDRI) is distinct for normal and dry AMD classes (See Figure 10), hence it can help the clinicians in faster and accurate screening of disease.
7. The proposed framework can be extended to diagnose DR, maculopathy and glaucoma diseases also.
8. The execution time of the optimal system for feature extraction (image data size=540), training (10-Fold) and testing (10-Fold) are 87.5131 sec, 0.323 sec and 0.2692 sec respectively using Intel i7-4770 3.47GHz processor, 16GB 1600 MHz CL9 DDR3-RAM and MATLAB 2012b computational environment. Hence our proposed system is fast and less complex.

5 Conclusion

In this work, an automated AMD detection system is proposed using DWT and feature ranking frame work. The statistical moments, energy, entropy and Gini index features are extracted on DWT coefficients to capture the minute changes in the normal and dry AMD images. The t-test, KLD, CBBD, ROC-based and Wilcoxon ranking methods are evaluated to rank the features and subsequently features were nested for automated classification. A set of supervised classifiers are used on the nested features to obtain the maximum classifier performance with minimum number of features. The SVM classifier with linear kernel yielded an average accuracy of 93.70%, sensitivity of 91.11% and specificity of 96.30% using 121 optimal features over ten-fold cross validation. Hence, this system can be used as an adjunct tool for AMD screening. Moreover, the developed integrated index AMDRI discriminates normal and AMD fundus images accurately using a single number. Hence, it may help the clinicians to make faster decisions during mass screening of the disease. The proposed system was evaluated using 540 images and can be further tested using more diverse images to evaluate the robustness of the system.

Conflicts of interest

The authors do not have any related conflict of interest.

Acknowledgements

A Bibliography

References

1. Paulus T.V.M. de Jong. Age-related macular degeneration. *New England Journal of Medicine*, 355(14):1474–1485, 2006.
2. Mohd Hanafi Ahmad Hijazi, Frans Coenen, and Yalin Zheng. Data mining techniques for the screening of age-related macular degeneration. *Knowledge-Based Systems*, 29:83–92, 2012.
3. Jennifer R Evans. Risk factors for age-related macular degeneration. *Progress in retinal and eye research*, 20(2):227–253, 2001.
4. Amresh Chopdar, Usha Chakravarthy, and Dinesh Verma. Age related macular degeneration. *BMJ: British Medical Journal*, 326(7387):485, 2003.
5. Zakaria Ben Sbeh, Laurent D Cohen, Gérard Mimoun, and Gabriel Coscas. A new approach of geodesic reconstruction for drusen segmentation in eye fundus images. *Medical Imaging, IEEE Transactions on*, 20(12):1321–1333, 2001.
6. Yuanjie Zheng, Brian Vanderbeek, Ebenezer Daniel, Dwight Stambolian, Maureen Maguire, David Brainard, and James Gee. An automated drusen detecion system for classifying age-related macular degeneration with color fundus photographs. In *IEEE 10th International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1440–1443, San Francisco, CA, USA, April 2013. IEEE.
7. Lee Brandon. *Automated Drusen Detection in a Retinal Image using Multi-Level Analysis*. PhD thesis, Clemson University, May 2003.
8. Konstantinos Rapantzikos, M Zervakis, and K Balas. Detection and segmentation of drusen deposits on human retina: Potential in the diagnosis of age-related macular degeneration. *Medical image analysis*, 7(1):95–108, 2003.
9. Cemal Köse, Uğur Şevik, and Okyay Gençalioğlu. Automatic segmentation of age-related macular degeneration in retinal fundus images. *Computers in Biology and Medicine*, 38 (5):611–619, 2008.
10. Cemal Köse, Uğur Şevik, Okyay Gençalioğlu, Cevat İkibaş, and Temel Kayıkıçıoğlu. A statistical segmentation method for measuring age-related macular degeneration in retinal fundus images. *Journal of medical systems*, 34(1):1–13, 2010.
11. David E Freund, Neil Bressler, and Philippe Burlina. Automated detection of drusen in the macula. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 61–64. IEEE, 2009.
12. Hector Santos-Villalobos, Thomas Paul Karnowski, Deniz Aykac, Luca Giancardo, Yaqin Li, T Nichols, KW Tobin, and Edward Chaum. Statistical characterization and segmentation of drusen in fundus images. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 6236–6241. IEEE, 2011.
13. Ziyang Liang, Damon WK Wong, Jiang Liu, Kap Luk Chan, and Tien Yin Wong. Towards automatic detection of age-related macular degeneration in retinal fundus images. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4100–4103. IEEE, 2010.
14. Jun Cheng, Damon Wing Kee Wong, Xiangang Cheng, Jiang Liu, Ngan Meng Tan, Mayuri Bhargava, Chui Ming Gemmy Cheung, and Tien Yin Wong. Early age-related macular degeneration detection by focal biologically inspired feature. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 2805–2808. IEEE, 2012.
15. P Soliz, SR Russell, MD Abramoff, S Murillo, M Pattichis, and H Davis. Independent component analysis for vision-inspired classification of retinal images with age-related macular degeneration. In *Image Analysis and Interpretation, 2008. SSIAI 2008. IEEE Southwest Symposium on*, pages 65–68. IEEE, 2008.

16. Victor Murray, Marios Pattichis, and Peter Soliz. New am-fm analysis methods for retinal image characterization. In *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, pages 664–668. IEEE, 2008.
17. Eduardo S Barriga, Víctor Murray, Carla Agurto, Marios S Pattichis, S Russell, MD Abramoff, Herbert Davis, and Peter Soliz. Multi-scale am-fm for lesion phenotyping on age-related macular degeneration. In *Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on*, pages 1–5. IEEE, 2009.
18. Mohd Hanafi Ahmad Hijazi, Frans Coenen, and Yalin Zheng. Retinal image classification using a histogram based approach. In *IEEE International Joint Conference on Neural Networks*, pages 3501–3507, 2010.
19. Carla Agurto, E Simon Barriga, Victor Murray, Sheila Nemeth, Robert Crammer, Wendall Bauman, Gilberto Zamora, Marios S Pattichis, and Peter Soliz. Automatic detection of diabetic retinopathy and age-related macular degeneration in digital fundus images. *Investigative ophthalmology & visual science*, 52(8):5862–5871, 2011.
20. Yalin Zheng, Mohd Hanafi Ahmad Hijazi, and Frans Coenen. Automated disease/no disease grading of age-related macular degeneration by an image mining approach. *Investigative ophthalmology & visual science*, 53(13):8310–8318, 2012.
21. Etta D Pisano, Shuquan Zong, Bradley M Hemminger, Marla DeLuca, R Eugene Johnston, Keith Muller, M Patricia Braeuning, Stephen M Pizer, et al. Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *Journal of Digital Imaging*, 11(4):193–200, 1998.
22. J Bossu, Ch Gée, G Jones, and F Truchetet. Wavelet transform to discriminate between crop and weed in perspective agronomic images. *computers and electronics in agriculture*, 65(1):133–143, 2009.
23. L. Khadra, A.S. Al-Fahoum, and H. Al-Nashash. Detection of life-threatening cardiac arrhythmias using the wavelet transformation. *Medical and Biological Engineering and Computing*, 35(6):626–632, 1997. ISSN 0140-0118.
24. O.A. Rosso, A. Figliola, J. Creso, and E. Serrano. Analysis of wavelet-filtered tonic-clonic electroencephalogram recordings. *Medical and Biological Engineering and Computing*, 42 (4):516–523, 2004. ISSN 0140-0118.
25. M. Puri, K.M. Patil, V. Balasubramanian, and V.B. Narayanamurthy. Texture analysis of foot sole soft tissue images in diabetic neuropathy using wavelet transform. *Medical and Biological Engineering and Computing*, 43(6):756–763, 2005. ISSN 0140-0118.
26. Stéphane Mallat. *A wavelet tour of signal processing*. Access Online via Elsevier, 1999.
27. Muthu Rama Krishnan Mookiah, U Rajendra Acharya, Choo Min Lim, Andrea Petznick, and Jasjit S Suri. Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features. *Knowledge-Based Systems*, 33:73–82, 2012.
28. R.M. Haralick, K. Shanmugam, and Its'Hak Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3(6):610–621, 1973.
29. Fritz Albregtsen et al. Statistical texture measures computed from gray level coocurrence matrices, 1995.
30. Amar Partap Singh Pharwaha and Baljit Singh. Shannon and non-shannon measures of entropy for statistical texture feature extraction in digitized mammograms. In *proceedings of the World Congress on Engineering and Computer Science*, volume 2, pages 20–22, 2009.
31. G.A. Giannakakis, N.N. Tsiaparas, M.-F.S. Xenikou, C. Papageorgiou, and K.S. Nikita. Wavelet entropy differentiations of event related potentials in dyslexia. In *BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on*, pages 1 –6, oct. 2008.
32. M. R. K. Mookiah, U. Rajendra Acharya, Roshan Joy Martis, Chua Kuang Chua, C. M. Lim, E. Y. K. Ng, and Augustinus Laude. Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach. *Knowledge-Based Systems*, 39(0):9 – 22, 2013.
33. Niall Hurley and Scott Rickard. Comparing measures of sparsity. *Information Theory, IEEE Transactions on*, 55(10):4723–4741, 2009.
34. Ron Kohavi and George H John. Wrappers for feature subset selection. *Artificial intelligence*, 97(1):273–324, 1997.
35. Atindra Mohan Goon, Milan Kumar Gupta, and Bhagabat Dasgupta. *An outline of statistical theory*. World Press Private, 1973.

36. Marco Bressan and Jordi Vitria. On the selection and classification of independent features. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(10):1312–1317, 2003.
37. Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.
38. S Rasoul Safavian and David Landgrebe. A survey of decision tree classifier methodology. *Systems, Man and Cybernetics, IEEE Transactions on*, 21(3):660–674, 1991.
39. Donald F Specht. Probabilistic neural networks. *Neural networks*, 3(1):109–118, 1990.
40. Kesar Singh and Minge Xie. Bootstrap: A statistical method, 2008.
41. Vladimir N Vapnik. Statistical learning theory. 1998.
42. T. Teng, M. Lefley, and D. Claremont. Progress towards automated diabetic ocular screening: A review of image analysis and intelligent systems for diabetic retinopathy. *Medical and Biological Engineering and Computing*, 40(1):2–13, 2002. ISSN 0140-0118.
43. Muthu Rama Krishnan Mookiah, U. Rajendra Acharya, Chua Kuang Chua, Choo Min Lim, E.Y.K. Ng, and Augustinus Laude. Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in Biology and Medicine*, 43(12):2136 – 2155, 2013.
44. Wei-Ying Ma and Bangalore S Manjunath. Texture features and learning similarity. In *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on*, pages 425–430. IEEE, 1996.
45. Sumeet Dua, U Rajendra Acharya, Pradeep Chowriappa, and Subbhuraam Vinitha Sree. Wavelet-based energy features for glaucomatous image classification. *Information Technology in Biomedicine, IEEE Transactions on*, 16(1):80–87, 2012.
46. Nello Cristianini and John Shawe-Taylor. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.