Glaucoma is a diffuse disease: it is the first cause of unemendable visual disability and blindness worldwide and a recent epidemiological review concludes that 1 in 40 adults over 40 years of age suffers from glaucoma with visual loss. Estimating a prevalence of 2.65% in the population over 40, the overall number of glaucomatous subjects is expected to increase in the course of the present decade, owing both to demographic expansion and population ageing, from about 60 million in 2010 to nearly 80 million in 2020. GLAUCOMA is caused by increased intraocular pressure (IOP) due to the malfunction of the drainage structure of the eyes. Approximately 120,000 are blind from glaucoma, thus accounting for 9%–12% of all cases of blindness in the U.S. About 2% of the population between 40–50 years old and 8% over 70 years old have elevated IOP, which increases their risk of significant vision loss and even blindness. The scheme proposed in this study is shown in Fig. 1.

Higher order spectra (HOS) based and texture-based features are commonly used in many medical image-processing areas. However, such studies have not yet been done on glaucoma images. Therefore, these features were extracted in our study. After preprocessing the acquired fundus images, HOS-based and texture-based features are extracted from the preprocessed images. Subsequently, these features are fed to Minimum distance, random forest, and naïve Bayesian (NB) classifiers for classification. Feature ranking is also performed to highlight and employ the discriminatory ability of the features in the classification process.
IV. FEATURE EXTRACTION

In this study, we have extracted two types of features: 1) HOS parameters 2) texture descriptors. Brief explanations of these features are given in the following.

A. Higher Order Spectra

HOS elicits both amplitude and phase information of a given signal. It offers good noise immunity and yields good results, even for weak and noisy signals. HOS consist of moment and cumulant spectra and can be used for both deterministic signals and random processes. We derived the features from the third-order statistics of the signal, namely, the bispectrum. The bispectrum is given by 
\[ B(f_1, f_2) = E[X(f_1)X(f_2)X^*(f_1 + f_2)] \]
where \( X(f) \) is the Fourier transform of the signal \( x(nT) \), and \( E[.] \) stands for the expectation operation. Features are calculated by integrating the bispectrum along the dashed line with slope \( a \). Frequencies are normalized by the Nyquist frequency (see Fig. 3). These bispectral invariants contain information about the shape of the waveform within the window and are invariant to shift and amplification and robust to time-scale changes. In this study, we used these bispectral invariant features for every 20\(^{th} \) Bispectral entropies have been derived from bispectrum plots to find the rhythmic nature of the heart rate variability and electroencephalogram signals. The equations used to determine the various HOS features. The normalization in the equations ensures that entropy is calculated for a parameter that lies between 0 and 1 (as required of a probability) and, hence, the entropies (Ent1, Ent2, and Ent3) computed are also between 0 and 1.

B. Texture Features

Texture descriptors provide measures of properties, such as smoothness, coarseness, and regularity, which indicate a mutual relationship among intensity values of neighboring pixels repeated over an area larger than the size of the relationship. Such properties can be used as features for pattern recognition.

Co-Occurrence Matrix: A gray-level co-occurrence matrix (GLCM) depicts how often different combinations of pixel brightness values (gray levels) occur in an image. It is a second order measure because it measures the relationship between neighborhood pixels. For an image of size \( m \times n \), we performed a second-order statistical textural analysis by constructing the GLCM as
\[ \phi(i,j) = \frac{\sum_{x,y} f(x,y)}{\sum_{x,y} f(x,y)} \]
where \( f(x,y) = 1 \) if \( (p+\Delta x, q+\Delta y) \in M \times N \) and \(|.| \) denotes the cardinality of a set. For a pixel in an image having a gray level \( i \), the probability that the pixel at a distance \( (\Delta x, \Delta y) \) away is \( j \), which is defined as
\[ P_{ij} = \sum_{k \in U} \mu(k) \]

Using (2) and (3), the following features were calculated: energy, contrast, homogeneity, entropy, and moments. The difference vector, which represents the gray-level difference statistics that can be obtained from the co-occurrence matrix can be derived using the following equation:
\[ P_{ij} = \sum_{k \in U} \mu(k) \]

Run Length Matrix: In the run-length matrix \( P \) (i, j), each cell in the matrix consists of the number of elements in which gray level \( i \) successively appears \( j \) times in the direction \( \theta \), and the variable \( j \) is termed as run length. The resultant matrix characterizes the gray-level runs by the gray tone, length, and direction of the run. As a common practice, run-length matrices of \( 0 \) equal to \( 0^\circ, 45^\circ, 90^\circ, \) and \( 135^\circ \) were calculated to determine the following features [20]: short-run emphasis, long-run emphasis, gray-level non uniformity, run-length non uniformity, and run percentage.
V. CLASSIFIERS USED

Our feature vector comprises heterogeneous features with densely distributed values. Figs. 4 and 5 demonstrate the distribution of feature values for the two classes of features (HOS and texture). HOS features are more widely distributed and have very limited correlation among them. The texture features are relatively correlated, but do exhibit discriminatory character for each of the images. The classifiers were chosen based upon their effectiveness in capturing the discriminative properties of these features, the impact of the ranking of features and efficiency, and the efficacy of the classification results. Four classifiers were employed for supervised learning and testing: random forest, and NB. Hardware consistency was maintained during the evaluation of these classifiers. The algorithmic control parameters for different classifiers are provided in Table I, and a succinct discussion on these classifiers is presented in the following.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Parameters Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Vector</td>
<td>SVM type: C-SVC, Kernel type: Radial basis function:</td>
</tr>
<tr>
<td>Machines</td>
<td>exp(-gamma(u-v)^2), Cost parameter 1, Kernel</td>
</tr>
<tr>
<td>(LibSVM)</td>
<td>degree: 3, and weights of classes: neutral (1).</td>
</tr>
<tr>
<td>Sequential Minimal Optimization</td>
<td>Kernel: Polynomial, Complexity parameter C:1.0, and Optimizations: Epsilon for round-off error: 1.0e-12.</td>
</tr>
<tr>
<td>Random Forest</td>
<td>Maximum tree depth: unlimited, Number of trees: 10.</td>
</tr>
</tbody>
</table>

Naive Bayesian (NB) is a statistical classifier based on Bayes rules. The naive Bayesian (NB) classifier can predict class membership probabilities, such as the probability that a given tuple belongs to a particular class, which is then employed for classification purposes. Since each of the features in the feature vector does not contribute equally to the classification process, we employ feature ranking using the methods of chi-square, gain ratio, and information gain (IG). Consequently, each of the features is weighted by their corresponding rank (normalized between 0–1). These weighted features are then used for training and testing of instances. In an additional study, we investigated the impact of normalization on the nonranked features. Lack of adequate normalization can skew the classification results causing significant false alarms and dismissals. z-Score normalization (standard norm) and min–max normalization were used to normalize the features. In z-score normalization, the values of feature A are normalized based on the mean and standard deviation of A. The z-score normalization maps a value $v$ of A to $v'$ using the formula

$$v' = \frac{v - \mu_A}{\sigma_A},$$

where $\mu_A$ is the mean and $\sigma_A$ is the standard deviation of the attribute. We also performed minmax normalization to represent the data in a new range using the formula

$$v = \frac{v - \min_A}{\max_A - \min_A},$$

where new min was chosen to be 0.0 and new max was chosen to be 1.0. Normalization was first employed for each feature independently and then on the aggregate of features, and classification accuracy was then investigated on the derived values of features.

VI. RESULTS

Table II shows the HOS features and the p-values. In this study, we extracted bispectrum invariants for each radon transformed eye image. Among them, we have used $P (6/20)$ feature as it was clinically significant ($p < 0.005$) in this study. Other entropy and bispectrum magnitudes at different angles were also chosen as the input vector for the classifier ($p < 0.005$). A two-sample t-test was used to estimate whether the mean value of each HOS feature was significantly different between the two classes. We assumed a null hypothesis that there is no difference between the two means. A $p$-value is a measure of probability that a difference between the two means happened by chance. In general, the null hypothesis is rejected, if the $p$-value is less than 0.05 or 0.01, corresponding to a 5% or 1% chance, respectively, of the null hypothesis being true. It can be seen from the table that all the features show low $p$-values, which indicate that there is a clinically significant difference between the means of the two classes. In addition, Table III shows the features extracted from the texture of the fundus image. These features also show significantly low $p$-values. All the texture features show higher variation for glaucoma images as compared to the normal fundus image. However, most of the HOS parameters show lower values for the glaucoma than the normal images. The ratio of the diameter of the optic cup to that of the optic disc in a healthy eye is generally less than 0.5. When the optic nerve is damaged by glaucoma, many of the individual fibers that make up the nerve are lost and the optic nerve becomes excavated. As glaucoma progresses, more optic nerve tissue is lost and the optic cup grows larger. Thus, the cup-to-disc ratio is higher for glaucoma subjects than for normal subjects, thus leading to differences in the respective fundus images. Moreover, in the case of optic nerve hemorrhage, another sign of glaucoma-related damage, the blood typically collects along the individual nerve fibers that radiate outward from the nerve. Such physiological changes are manifested in the fundus images, and our experiments show that the HOS and texture features are able to quantify such differences in eye physiology.
Table IV summarizes our classification results with and without feature ranking, and with and without feature selection. We applied fivefold cross validation for training and testing purposes. LibSVM outperformed the other classifiers without the ranking of features, while random forest performed better than other methods with IG feature ranking. However, feature ranking methods do not improve the performance of other classifiers. Significantly, feature selection based on any of the three feature-ranking methods outperformed feature-ranking methods. This result demonstrated that some of the features do not contribute to the classification, or even hamper the performance of classifiers on the diagnosis of glaucoma. The feature-selection methods were subsequently adopted for normalization-based studies. Table V summarizes the classification results obtained as a result of normalization. The classifiers generally maintained their performance or presented improved accuracies when the normalization was performed on features as an aggregate. The z-score normalization based method outperformed the min–max normalization-based method without feature ranking and feature selection. We attribute this performance to the fact that z-score normalization reassigns values based on the variance around the mean (resulting in a mean of zero for the normalized values), while the min–max normalization exploits the boundary values. Consequently, min–max normalization is sensitive to the outliers and leads to increased skewness. Thus, we applied feature-selection methods on the z-score normalization result. All of the four classifiers performed the same using different feature-selection methods. Random forest obtained the highest classification accuracy at 91.7% for z-score normalized (all features) data. We also demonstrated that the combination of min–max normalization and chi-squared feature selection did not improve the classification accuracy using chi-squared feature selection on the original features.

**CONCLUSIONS**

In this study, we developed an automatic glaucoma diagnosis system that combines texture and HOS features extracted from fundus images for diagnosis. We found that the texture and HOS-based features were clinically significant, i.e., these features had a low p-value, which means that there is more chance that these features have very different values for the normal and abnormal classes, and, hence, better discriminate the two classes. We, therefore, used these features for classification. The performances of four supervised classifiers were evaluated. We found that the random-forest classifier, combined with z-score normalization and feature-selection methods, performed the best among the four classifiers with a classification accuracy of more than 91%. Our technique is of clinical significance, as the accuracy obtained is comparable to the accuracies achieved so far by other studies and the equipment used is the most commonly used fundus-imaging equipment. Therefore, our proposed technique can be easily incorporated into existing medical infrastructures, thus making it clinically a viable option. The classification accuracy can be further improved by increasing the number of diverse training images, choosing better features and better classifiers and using controlled environmental lighting conditions during image acquisition. Using more diverse digital fundus images from normal and glaucoma subjects can further enhance the percentage of correct diagnosis. Physicians and other clinical practitioners can employ our proposed approach within a decision support system that offers secondary opinion in the clinical diagnosis of glaucoma.

**APPENDIX**

HOS features:

Mean of magnitude: \( M_{\text{ave}} = \frac{1}{L} \sum_{l=1}^{L} |B(f_1, f_2)| \)

Phase entropy: \( P_e = \sum_{n} p(y_n) \log p(y_n) \).

Normalized bispectral entropy (BE 1): \( \text{Ent1} = -\sum_{n} p_n \log p_n \)

where \( p_n = (|B(f_1, f_2)|)/(\sum_{l} |B(f_1, f_2)|) \), and \( W \) is the region as shown in Fig. 3.

Normalized bispectral squared entropy (BE 2):

\( \text{Ent2} = -\sum_{n} q_n \log q_n \)

where \( q_n = (|B(f_1, f_2)|^2)/(\sum_{l} |B(f_1, f_2)|^2) \).

Normalized bispectral cubic entropy (BE 3):

\( \text{Ent3} = -\sum_{n} r_n \log r_n \)

where \( r_n = (|B(f_1, f_2)|^3)/(\sum_{l} |B(f_1, f_2)|^3) \).
Co-occurrence matrix based features:

\[
\text{Energy} = \sum_i \sum_j N^2_d(i,j)
\]

\[
\text{Entropy} = -\sum_i \sum_j N_d(i,j) \log_2 N_d(i,j)
\]

\[
\text{Contrast} = \sum_i \sum_j (i-j)^2 N_d(i,j)
\]

\[
\text{Homogeneity} = \sum_i \sum_j \frac{N_d(i,j)}{1+|i-j|}
\]

\[
\text{Correlation} = \frac{\sum_i \sum_j (i-\mu_i)(j-\mu_j) N_d(i,j)}{\sigma_i \sigma_j}
\]

Run-length-matrix-based features:

**Short-run emphasis:**

\[
\frac{\sum_i (P_{E}(i,j)/R)^2}{\sum_i P_{E}(i,j)}
\]

**Long-run emphasis:**

\[
\frac{\sum_i \sum_j (P_{E}(i,j))^2}{\sum_i \sum_j P_{E}(i,j)}
\]

**Gray-level non-uniformity:**

\[
\frac{\sum_i \sum_j P_{E}(i,j)^2}{\sum_i \sum_j P_{E}(i,j)}
\]

**Run-length non-uniformity:**

\[
\frac{\sum_i \sum_j P_{E}(i,j)}{\sum_i P_{E}(i,j)}
\]

**Run percentage:**

\[
\frac{\sum_i \sum_j P_{E}(i,j)}{A}
\]

Where A is the area of the image of interest.

REFERENCES


