Decision Support for Automated Screening of Diabetic Retinopathy

Pallavi Kahai  Kamesh Rao Namuduri  ECE, Wichita State University
Wichita, KS 67208

Hilary Thompson  LSU Eye Center, Louisiana State University
New Orleans, LA 2234

Abstract—Diabetic Retinopathy (DR) is the leading cause of blindness. DR results in retinal disorders that include: microaneurysms, drusens, hard exudates and intra-retinal microvascular abnormalities (IRMA). The early signs of DR are depicted by microaneurysms among other signs. A prompt diagnosis when the disease is at the early stage can help prevent irreversible damages to the diabetic eye. This paper presents a decision support framework for automated screening of early signs of DR and classification schemes for deducing the presence or absence of microaneurysms are developed and tested under a univariate environment. The detection rule is based on binary hypothesis testing problem which simplifies the problem to yes/no decisions. An analysis of the performance of the Bayes optimality criteria is also presented in the paper.

I. INTRODUCTION

According to the American Diabetes Association, 18.2 million of the American population which constitutes 6.3% of the total population, has diabetes. In the United States alone, diabetes is responsible for 8% of legal blindness, making it the leading cause of blindness in people 20 to 74 years of age [1], [2], [3]. Because of the asymptotic nature of the disease and severe aftermath, it is essential that identification and treatment starts as early as possible. If the disease remains undetected for long it leads to permanent complications in the retina. Researchers including the authors of this paper [4] have therefore suggested an automated screening system for diabetic retinopathy for prompt diagnosis. Since the disorders exhibited in the early stage do not affect the vision, detection of the disease right at it’s onset can be done only if regular eye examination of the diabetic patients are performed.

This paper suggests an automated screening system that would detect early signs of non-proliferative diabetic retinopathy (NPDR). The focus of the paper is on microaneurysms as these are the early signs of DR and are present at all the stages as the disease progresses from mild to severe NPDR. Univariate approach has been devised to test the suitability of the classification mechanism with respect to the detection of retinopathy. The entire process is composed of recognition, training and classification mechanisms. Recognition involves developing algorithms for the detection of disorders. Training uses supervised learning and training data is obtained by applying the detection algorithms to retinal images. Classification is performed by test data subjected to unsupervised learning. In order to test the performance of the machine the results obtained are compared with the physician’s diagnosis. This approach has been developed for one particular feature but the feature space can be extended depending on the number of disorders needed to be detected.

The organization of the paper is as follows: Section II discusses the methodologies employed for detection of disorders related to diabetic retinopathy. Section III presents the decision framework adopted for the detection of microaneurysms followed by the experimental results and conclusions in section IV and V respectively.

II. DETECTION OF FEATURES USING IMAGE-PROCESSING TECHNIQUES

Researchers [5], [6], [7], [8] have approached the problem of feature detection in varied ways. A modular system developed by [9] makes use of a large database of images where features have been identified by the physicians and which is later used to detect similar features in new images. The recognition process employs unsupervised learning mechanism and the classification phase uses supervised learning.

Another technique developed by [10] automatically detects and distinguishes between different lesions (hard exudates, cotton wool spots and haemorrhages) after image enhancement. The image is enhanced by taking the difference between the background illumination and an edge detection operator.

The detection technique used to identify a pattern depends on the feature exhibited by it. Thus, different methodologies are employed for detecting different features. We are working on algorithms for detection of hard exudates, microaneurysms, drusens and venous beading. Wavelet transformation which helps in viewing the image at multiple resolutions is being used for detecting the hard exudates. Matched filters are used to detect microaneurysms and drusens. This paper lays emphasis on the detection of the microaneurysms and develops decision support framework and tests it’s validity using several retinal images. This framework can be easily extended to include multiple features.

1) Filters used for detecting microaneurysms: Microaneurysms are red spots of very small diameter, therefore, a center surround filter with a dark pattern inside and a brighter one surrounding it is used as the matched filter. The raw image is convolved with the filter and a threshold is set in order to determine the location of the microaneurysms. The size of the filter provides an estimate for the size of the microaneurysms. Fig. 1 shows the image of a retina infected with microaneurysms and the response of the matched filter tuned to detect microaneurysms of size 2X2. Several microaneurysms
have been identified on the blood vessels as well, these can be excluded by applying appropriate constraints. For instance, detection of blood vessel can be performed first and a rule that microaneurysms may not lie on the blood vessel can be imposed.

![Retinal image](image)

Fig. 1. Retinal image (left) infected with microaneurysms and the matched filter response (right) that detects the microaneurysms of size 2X2

III. DECISION SUPPORT FRAMEWORK FOR DIABETIC RETINOPATHY

Pattern classification can be performed using decision rules. This paper analyzes the performance of Bayes optimality criteria.

Diabetic Retinopathy is characterized by the presence of one or all of the features discussed previously. Thus, a person suffering from the disease might manifest one of the several patterns. Also, as the patterns are independent of each other, each of the features can be studied individually. We would investigate the presence or absence of one particular feature, that is, microaneurysms. Thus, the problem reduces to yes/no decision making. Training data is obtained from the images of the retina provided by the physician. Each training instance is composed of all the attributes and belongs to exactly one class among all the possible classes. For medical diagnosis, the training instances are the descriptions of patients with known final diagnosis, attributes are the symptoms and classes are the possible diagnosis. For the problem under consideration the possible classes are affected retinal images (abnormal case) and unaffected retinal images (normal case). Probability density functions are obtained for each of the states of nature, that is, for the case where a symptom is absent and for the case where it is present.

There is a cost associated with each decision that depends on whether the decision coincides with the true diagnosis or not. The probable classifications are thus given as:

- Correct accept/hit: A person affected with diabetic retinopathy has a true diagnosis.
- False accept/false alarm: A normal person diagnosed with diabetic retinopathy.
- False reject/miss: An affected person diagnosed as normal.
- Correct reject: A normal person classified as unaffected.

The erroneous classifications are given by false accept (FA) and false reject (FR) which are referred to as type I and type II errors [11]. In relation to diabetic retinopathy application, the cost associated with each of these would be governed by the amount of harm caused by a misdiagnosis. The repercussions incurred in categorizing a person affected with diabetic retinopathy as normal is obviously more than the converse. Some statisticians refer to the cost function as the loss function. The cost or the loss associated with any misclassification is directly proportional to the severity induced by the error.

Considering a univariate case and analyzing each feature independently the probability density functions obtained for the case wherein microaneurysms are present and the case where they are absent is shown in Fig. 2. The data is composed of 25 normal images and 23 affected images of the retina selected from a large pool of images.

![Class conditional probability density function](image)

Fig. 2. Class conditional probability density function for normal category (unaffected retina) and abnormal category (affected retina) with discriminability=2.7528, False alarm probability=9.8374e-007 and Detection probability=0.7071

The boundary value occurs at feature value of 0.3. At the boundary value decision can be taken for either case. But the loss function can be such that the loss associated with normal classification is more than the abnormal classification. Thus, in order to minimize the risk involved in misclassification, all the cases at the boundary value can be classified as abnormal. The discriminability is directly proportional to the difference in the means and varies inversely with the variance. Therefore, the farther apart are the two density function and the lesser the variance they exhibit, the greater is the discriminability and lesser the overlap. If the discriminability is infinity then the two density functions do not exhibit any overlap and the type I and type II errors are zero and the detection is perfect. The discriminability is a quantitative measure of decidability and is independent of the chosen decision criteria. This factor is given as:

$$d = \frac{|\mu_1 - \mu_0|}{\sqrt{(\sigma_1^2 + \sigma_0^2)}}$$

The discriminability for the case shown in Fig. 2 was found to be $d = 2.7528$. When a feature value is presented to the
system then it can be either classified as normal or abnormal. This detection process can be treated as a binary hypotheses testing problem. The hypotheses supported are the null hypothesis, \( H_0 \) (specifies the absence of the microaneurysms) and the alternative hypothesis, \( H_1 \) (specifies the presence of the microaneurysms). Each of the hypothesis has a probability density function associated with it. In the case of diabetic retinopathy, each of the two categories, that is, the affected and the unaffected, represent a Gaussian distribution where the feature values tend to a particular value (given by the mean) but with a little variation (given by the variance). For the aforementioned data maximum likelihood estimation was carried out in order to determine the conditional probability density functions and the problem was reduced to the detection of following hypotheses:

\[
H_0 : p_0 \sim N(\mu_0, \sigma)
\]

versus

\[
H_1 : p_1 \sim N(\mu_1, \sigma)
\]

where, \( p_0 \) and \( p_1 \) are the density functions for the normal and the affected case respectively. The variance obtained for the normal and the abnormal case for the data under consideration was found to be the same. The respective values obtained for these two sample datasets are: \( \mu_0 = 0.0475 \) and \( \mu_1 = 0.5842 \) for a variance of \( \sigma = 0.0832 \). Since the variance for the two categories is the same and the difference between the means is small, the discriminability is small. Decision rule, \( \delta \) for \( H_0 \) versus \( H_1 \) is a function of the feature and can take either a value of 1 or 0 depending on whether the feature belongs to \( H_1 \) or \( H_0 \). Also, if the feature does not belong to \( H_1 \) then it belongs to \( H_0 \), i.e., \( H_0 = H_1^c \). Therefore, the decision rule can be mathematically represented as:

\[
\delta(x) = \begin{cases} 
1 & \text{if } x \in H_1 \\
0 & \text{if } x \in H_0 = H_1^c 
\end{cases}
\]

where \( x \) is the feature. Detection can be performed by employing any optimal decision criteria. This paper considers the following detection methods.

A. Likelihood Ratio Test

The conditional p.d.f, \( p_i(x) \) gives the likelihood that a feature value, \( x \) belongs to a particular state of nature \( H_i \). The likelihood ratio is given by:

\[
l(x) = \frac{p_1(x)}{p_0(x)}
\]

Detection method that compares the likelihood ratio with a certain threshold value is called the likelihood ratio test. For the Gaussian distributions considered for the two possible classes, the likelihood function is given as:

\[
l(x) = e^{\frac{\mu_1 - \mu_0}{\sigma^2}(x - \frac{\mu_0 + \mu_1}{2})}
\]

Fig. 5 shows the likelihood ratio. Since the two distributions are independent of each other the numerator and the denominator can be interchanged.

A threshold, \( \tau \) is fixed such that decision can be made on \( \tau \). Thus, the decision rule becomes:

\[
\delta(x) = \begin{cases} 
1 & l(x) \geq \tau \\
0 & l(x) < \tau 
\end{cases}
\]

The threshold depends on the prior knowledge about each hypothesis and cost or loss associated with each classification. The prior probability does not depend on the feature value. The prior probabilities for diabetic retinopathy can be demographically based on the percentage of population affected by the disease. We have assumed that 60% of the diabetic population is prone to diabetic retinopathy. Thus, the prior probability for the abnormal case, \( P(abnormal) = 0.6 \) and for the normal case would be, \( P(normal) = 1 - P(abnormal) = 0.4 \). Loss associated with each classification is determined by a loss matrix, \( L \) where an element \( L_{ij} \) represents the loss associated in choosing a hypothesis \( H_i \) when \( H_j \) is true. For a binary hypothesis testing problem \( i \) and \( j \) can only take values of 0 or 1. Also, the range of loss would be from 0 to 1. Zero corresponds to no loss and 1 corresponds to maximum loss. Zero loss would be attributed to the case where detection occurs, that is, type I (FA) and type II (FR) errors are absent. As mentioned earlier the loss associated with type I error should be more than type II error. In other words loss incurred in classifying an abnormal person as normal given by \( L_{01} \) should be greater than the converse, i.e., \( L_{10} \). In our calculations we have made use of the following loss matrix.

\[
Loss \ matrix \ L = \begin{pmatrix} 
0 & 0.8 \\
0.3 & 0 
\end{pmatrix}
\]

The threshold \( \tau \) is given as:

\[
\tau = \frac{P(normal)(L_{10} - L_{00})}{P(abnormal)(L_{01} - L_{11})}
\]

By substituting, \( P(normal) = 0.4, P(abnormal) = 0.6, L_{11} = L_{00} = 0, L_{01} = 0.8 \) and \( L_{10} = 0.3 \) we get a threshold value of 1.7778. Thus, if the likelihood ratio is more than 1.7778 then the feature should be classified as abnormal and normal otherwise. Instead of comparing the likelihood ratio with the threshold, \( \tau \) the feature value can be compared...
with a threshold, \( \tau' \) which can be derived from (5). Threshold \( \tau \) was found to be

\[
\tau' = \frac{\mu_0 + \mu_1}{2} + \frac{\sigma_0^2 \log \tau}{\mu_1 - \mu_0} \quad (9)
\]

Thus, the detector becomes:

\[
\delta(x) = \begin{cases} 
1 & x \geq \frac{\mu_0 + \mu_1}{2} + \frac{\sigma_0^2 \log \tau}{\mu_1 - \mu_0} \\
0 & x < \frac{\mu_0 + \mu_1}{2} + \frac{\sigma_0^2 \log \tau}{\mu_1 - \mu_0}
\end{cases} \quad (10)
\]

By substituting the appropriate value in the above equation we get \( \tau' = 0.3233 \). Thus, if a feature value is above 0.3233 then it can be classified as abnormal and normal otherwise.

**B. Maximum a posteriori (MAP detector)**

The posterior probability, unlike the prior probability, is the conditional probability that a particular state of nature exists for a given feature value and is represented as \( P(H_k|x) \) where \( k \) represents the class index. For better comprehension, we would represent the posterior probability for the normal category as, \( P(\text{normal}|x) \) and for the abnormal case as \( P(\text{abnormal}|x) \). In general, the posterior probability is obtained by the Bayes formula as follows:

\[
P(H_k|x) = \frac{p(x|H_k)P(H_k)}{\sum_{l=1}^{c} p(x|H_l)P(H_l)} \quad (11)
\]

where, \( p(x|H_k) \) is the conditional probability or the likelihood, \( P(H_k) \) is the prior probability and \( c \) is the total number of possible hypotheses. The posterior probability for the normal case is thus given as follows:

\[
P(\text{normal}|x) = \frac{p_0(x) * P(\text{normal})}{p(x)} \quad (12)
\]

where, \( p_0(x) = p(x|H_0) \).

Similarly, the posterior probability for the abnormal case is:

\[
P(\text{abnormal}|x) = \frac{p_1(x) * P(\text{abnormal})}{p(x)} \quad (13)
\]

where, \( p_1(x) = p(x|H_1) \).

Here, \( p(x) \) is a scaling factor and is given as:

\[
p(x) = p_0(x) * P(\text{normal}) + p_1(x) * P(\text{abnormal})
\]

Now, the decision rule is based on choosing that hypothesis whose posterior probability is maximum and hence, it is called maximum a posterior estimator. So, the decision rule is represented as:

Decide \( H_1 \) if \( P(\text{normal}|x) > P(\text{abnormal}|x) \)

Decide \( H_0 \) otherwise

Classification using MAP detector is depicted in Table I. It can be deduced from the table that a retinal image that presents 4 or more microaneurysms is treated as abnormal by the MAP detector. The final diagnosis rests with the physician and is shown in section IV.

**TABLE I**

**Table showing the posterior probabilities for different counts of microaneurysms and the corresponding categorization obtained from the MAP detector**

| Feature Value | \( P(\text{normal}|x) \) | \( P(\text{abnormal}|x) \) | Classification |
|---------------|-------------------------|-------------------------|---------------|
| 0             | 8.4717683e-001          | 1.5823744e-001          | Normal        |
| 1             | 7.6051899e-001          | 2.3948108e-001          | Normal        |
| 2             | 6.5458190e-001          | 3.4541814e-001          | Normal        |
| 3             | 5.3060895e-001          | 4.6939116e-001          | Normal        |
| 4             | 4.0265114e-001          | 5.9734986e-001          | Abnormal      |
| 5             | 2.866279e-001           | 7.1373219e-001          | Abnormal      |
| 6             | 1.9316530e-001          | 8.0683472e-001          | Abnormal      |
| 7             | 1.2480500e-001          | 8.7519500e-001          | Abnormal      |
| 8             | 7.826310e-002           | 9.2173690e-001          | Abnormal      |
| 9             | 4.810631e-002           | 9.849791e-001           | Abnormal      |
| 10            | 2.919115e-002           | 9.708086e-001           | Abnormal      |
| 11            | 1.756968e-002           | 9.824303e-001           | Abnormal      |
| 12            | 1.052091e-002           | 9.894791e-001           | Abnormal      |
| 13            | 6.279664e-003           | 9.937203e-001           | Abnormal      |
| 14            | 3.740336e-003           | 9.962597e-001           | Abnormal      |
| 15            | 2.224729e-003           | 9.977753e-001           | Abnormal      |
| 16            | 1.321955e-003           | 9.986780e-001           | Abnormal      |

**C. Bayes detector**

The Bayes detector is based on the optimality criteria that minimizes the Bayes risk or average risk, \( r(\delta) \). The Bayes risk is the overall cost incurred by a decision rule, \( \delta \) and it depends on the conditional risk or expected loss. The conditional risk, \( R_k(\delta) \) is the average cost incurred by a decision rule when hypothesis \( H_k \) is true. Suppose that a particular observed feature is classified as \( H_i \) whereas, the true hypothesis happens to be \( H_j \), then the loss incurred, as shown in the previous section, would be \( L_{ij} \). Because \( P(H_j|x) \) is the probability that the true state of nature is \( H_j \) then the expected loss in classifying the feature as \( H_i \) would be [12]:

\[
R_i(\delta) = \sum_{j=1}^{c} L_{ij}P(H_j|x) \quad (14)
\]

where \( c \) is the total number of possible hypothesis. To minimize the overall risk, compute the conditional risk given as above for each possible hypothesis and select that hypothesis for which the risk is minimum. This minimum conditional risk is called the Bayes risk. The average cost incurred by the decision rule when hypothesis \( H_0 \) is true, that is, the person under consideration is normal given by, \( R_0(\delta) \), is:

\[
R_0(\delta) = L_{00}P(\text{normal}|x) + L_{10}P(\text{abnormal}|x) \quad (15)
\]

Similarly, average cost incurred in classifying a person as abnormal is,

\[
R_1(\delta) = L_{01}P(\text{normal}|x) + L_{11}P(\text{abnormal}|x) \quad (16)
\]

The Bayes risk is given as:

\[
r(\delta) = \min[R_0(\delta), R_1(\delta)] \quad (17)
\]

The bayes decision rule is:

Decide \( H_1 \) if \( R_1(\delta) < R_0(\delta) \)

Decide \( H_0 \) otherwise
The Bayes detector does not provide any condition at the boundary. Therefore, the boundary feature values at which the probability of classifying in to each of the categories is the same, the feature can be classified into any category. Specifically, at the boundary we should be able to classify each feature as abnormal so that a person for whom the true hypothesis is abnormal is not misdiagnosed.

IV. EXPERIMENTAL RESULTS

We have made use of 143 retinal images provided by the Louisiana State University Eye Center. Supervised learning was performed for training whereas unsupervised learning was used to test the system. The system is trained for NPDR. In our experiments we have compared the retinal images of the diabetic patients which do not manifest microaneurysms with those which do. Moderate to severe cases were considered for the case wherein microaneurysms are present. A YES decision (abnormal) corresponds to the presence of microaneurysms for the moderate and severe cases of NPDR and NO decision (normal) relates to the absence of microaneurysms. Each decision has an associated cost that is represented by the Bayes risk. The results obtained are given in Table II.

<table>
<thead>
<tr>
<th>Feature Value</th>
<th>Classification</th>
<th>Bayes Risk</th>
<th>Physician’s Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal</td>
<td>1.265854e-001</td>
<td>Normal</td>
</tr>
<tr>
<td>1</td>
<td>Normal</td>
<td>1.915849e-001</td>
<td>Normal</td>
</tr>
<tr>
<td>2</td>
<td>Normal</td>
<td>2.763344e-001</td>
<td>Normal</td>
</tr>
<tr>
<td>3</td>
<td>Normal</td>
<td>3.755129e-001</td>
<td>Normal</td>
</tr>
<tr>
<td>4</td>
<td>Abnormal</td>
<td>1.207953e-001</td>
<td>Abnormal</td>
</tr>
<tr>
<td>5</td>
<td>Abnormal</td>
<td>8.598838e-002</td>
<td>Normal</td>
</tr>
<tr>
<td>6</td>
<td>Abnormal</td>
<td>5.794959e-002</td>
<td>Abnormal</td>
</tr>
<tr>
<td>7</td>
<td>Abnormal</td>
<td>3.744149e-002</td>
<td>Abnormal</td>
</tr>
<tr>
<td>8</td>
<td>Abnormal</td>
<td>2.347893e-002</td>
<td>Abnormal</td>
</tr>
<tr>
<td>9</td>
<td>Abnormal</td>
<td>1.443189e-002</td>
<td>Abnormal</td>
</tr>
<tr>
<td>10</td>
<td>Abnormal</td>
<td>8.757346e-003</td>
<td>Abnormal</td>
</tr>
<tr>
<td>11</td>
<td>Abnormal</td>
<td>5.207903e-003</td>
<td>Abnormal</td>
</tr>
<tr>
<td>12</td>
<td>Abnormal</td>
<td>3.156274e-003</td>
<td>Abnormal</td>
</tr>
<tr>
<td>13</td>
<td>Abnormal</td>
<td>1.883899e-003</td>
<td>Abnormal</td>
</tr>
<tr>
<td>14</td>
<td>Abnormal</td>
<td>1.122101e-003</td>
<td>Abnormal</td>
</tr>
<tr>
<td>15</td>
<td>Abnormal</td>
<td>6.674187e-004</td>
<td>Abnormal</td>
</tr>
<tr>
<td>16</td>
<td>Abnormal</td>
<td>3.965865e-004</td>
<td>Abnormal</td>
</tr>
</tbody>
</table>

Feature value corresponds to the number of microaneurysms exhibited by the test image. The system classifies affected retinal image as abnormal whereas unaffected retinal image is classified as normal. The Bayes risk as shown in (17) is given by the minimum of the two expected losses. From Table I, it can be deduced that for an image that exhibits less than 3 microaneurysms, the expected loss associated with normal classification is less than that of abnormal classification. Hence a feature value of less than 3 has been categorized as normal which is rational. The threshold provided by the likelihood ratio test, that is, $\tau = 0.3233$ corresponds to a feature value of 4. If the posterior probability of classifying a feature as abnormal is same as that of normal then the classification can be made either way. But classifying it as abnormal would reduce the risk involved in misclassification. Thus, if the retinal image presented exhibits 4 or more microaneurysms, it has been classified as abnormal. The system classifies a feature value of 4 and 5 as normal whereas the physician’s diagnosis is otherwise. This is because of the asymmetric costs associated with each classification. The final decision rests with the physician, therefore, if the system treats a normal image as abnormal the cost incurred is less than that of the converse. The detection probability for Fig. 2 was found to be 0.7071 and false alarm probability was 9.8374e-007.

V. CONCLUSIONS AND FUTURE WORK

This paper proposed a decision support framework for automated screening of DR for the univariate case. This model can be extended to multiple disorders that would include the covariance associated with all the signs of DR. The experiments support the feasibility of a complete automated screening mechanism that includes all the disorders related to DR. The machine can be made adaptable by including Bayesian Learning Mechanism that would improve the accuracy of the classifier as a new feature value is presented to it by modifying the priors.

REFERENCES