MEDICAL DIAGNOSIS SYSTEM FOR THE DETECTION OF DIABETES RETINOPATHY BASED ON RETINAL IMAGES USING DATA MINING TECHNIQUES

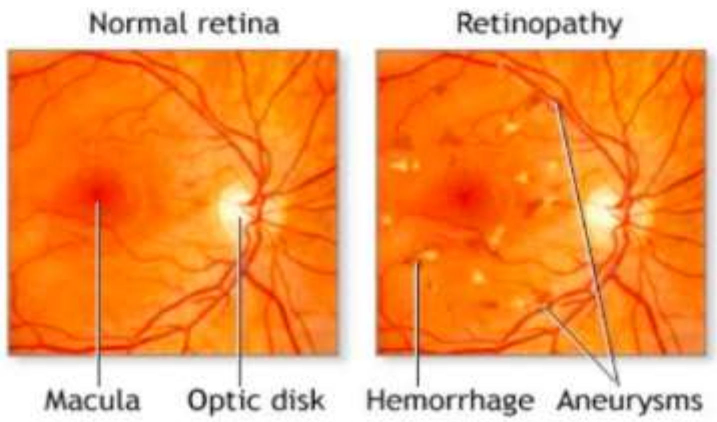
G.PARY[1] , Dr.M.ARAMUDHAN[2] and Dr.N.THIRUMOORTHY[3]
[1]Reseach scholar, Manonmaniam Sundaranar University, Tirunelveli.
[2]Associate professor, Department of Information Technology, Perunthanlaivar Kamarajar Institute of Engineering and Technology, Nedungadu.
[3]Associate professor, Mahindra college of Engineering, Salem

ABSTRACT:
This paper considered the retinal images and to detect the Diabetic retinopathy disease is either affected the eye or not using data mining technique. Th early detection of Diabetic retinopathy disease from retinal images are to help the ophthalmologist to be diagnosed to detect diabetic retinopathy by using data mining techniques such as Support Vector Machine, page rank, Random forest Algorithm, means and Apriori algorithm and also It supports for future implementation to know the people of their Diabetic retinopathy disease condition from any ATM machine.

Keywords: Diabetic retinopathy, mean, standard deviation and entropy, Gray level co-occurrence. Support Vector Machine, page ranking, random forest Algorithm, K means and Apriori algorithm

Introduction:
The retina is the one of the nerve layer that is in the back of the eye that senses light, and creates impulses to travel through the optic nerve to the brain. The retinal images are taking important role for proper diagnosis of several retinal diseases affect the eye like diabetic retinopathy, hypertensive retinopathy, Glaucoma, central serous retinopathy and cardiovascular diseases etc. The glucose is an important source of energy for the body's cells .The too much of glucose in the blood for a long time can cause damage in blood vessels and the small blood vessels in the eyes. When high blood sugar levels cause damage to blood vessels in the retina and swell, leak or close off completely or new blood vessels grow on the surface of the retina is called diabetic retinopathy. That can cause blood vessels in the retina to leak fluid or hemorrhage (bleed), distorting vision, leak fluid or hemorrhage (bleed), distorting vision, reduced vision or impaired vision - sometimes sudden or often a gradual degradation dark spots in front of the eyes, double visions. According to the Centers for Disease Control and Prevention (CDC), about 90 percent of diabetes-related vision loss can be prevented, for that the early detection is important. There are four types of diabetic retinopathy diseases progress. The mild non proliferative retinopathy in that the Small areas of balloon-like swelling in the retina's tiny blood vessels, called micro aneurysms that occur at this earliest stage of the disease. The micro aneurysms may leak fluid into the retina. Moderate no proliferative retinopathy as the disease progresses, blood vessels that nourish the retina may swell and distort. And also lose their ability to transport blood. The appearance of the retina that can contribute to DME. Sevier non proliferative retinopathies in that many more blood vessels are blocked, depriving blood supply to areas of the retina. This area's growth factors that signal the retina to grow new blood vessels. Proliferative diabetic retinopathy (PDR) is a advanced stage, growth factors secreted by the retina trigger the proliferation of new blood vessels, this grow in the surface of the retina and the vitreous gel, that fluid to fills the eye. The new blood vessels are fragile; it makes them to leak and blood. The scar tissue can contract and cause retinal detachment—the pulling away of the retina from underlying tissue, like wallpaper peeling away from a wall and the Retinal detachment can lead to permanent vision loss. Many people are daily affected by those diseases in the world. Manual analysis of retinal images is quietly time-consuming process and accuracy of the diseases depends on the ophthalmologist. There is limited number of medical experts for mass screening of patients in ophthalmology field in various countries. so the development of " the medical diagnosis system for the detection of diabetes Retinopathy from retinal images using data mining techniques" are used for mass screening of patients to help the medical expert for the early detection of diabetic retinopathy and also used in ATM to know the people of their diseases to prevent from their risk condition.

2. **System implementation:**
This system shows that the Normal retinal and Diabetic Retinopathy (DR) retinal images are extracted by fundus camera from data set. the pre-processing ,conversion of RGB to HIS of retinal image and filters are used to enhance the quality of image by reducing noisy and poorly contrast of retinal images for sugar level prediction like normal ,mild, moderate, and Sevier etc.

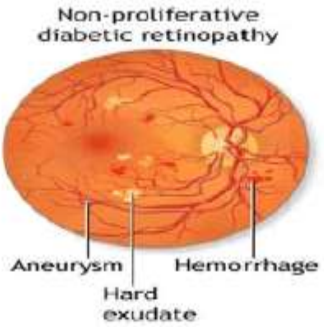


3. Major Types of Diabetic Retinopathy (DR)

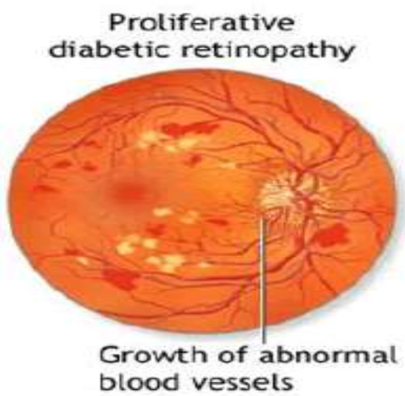
With respect to the levels of retinal image damage, the diabetic retinopathy diseases have been classified. That is called some type of "diabetic retinopathy".

3.1."Non-proliferative diabetic retinopathy" (NPDR)

The early stage of diabetic eye disease is called “Non-proliferative diabetic retinopathy" (NPDR). In this tiny blood vessels leak, making the retina swell. the, macula swells, this is called macular edema. The Macular edema is the leading cause of visual impairment of patients with diabetes. It result may lead from functional damage and necrosis of retinal capillaries. And also blood vessels in the retina can close off is called macular ischemia. By this blood cannot reach the macula. In particular times the tiny particles called exudates can form in the retina that can affect the vision.



3.2. Proliferative Diabetic Retinopathy (PDR)
Advanced diabetic retinopathy is called “proliferative Diabetic retinopathy" and is the most serious type that can seriously impair vision. It is classified as early or high risk. In early PDR, new vessels are appearing, but they do not meet the criteria for high-risk PDR. In high-risk PDR, revascularization of the disc [NVD]) is one-third to one-half, or greater, of disc area



4. Symptoms of Diabetic Retinopathy (DR)

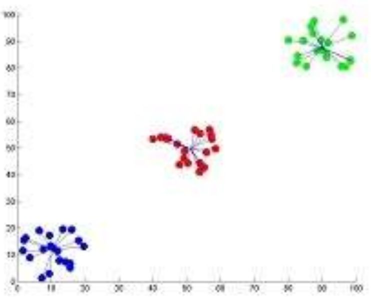
The diabetic retinopathy is a diabetes complication that affects eyes And also damages the blood vessels of the light-sensitive tissue at the back of the eye or retina. In the initial stage of the diabetic retinopathy may cause no symptoms or only mild vision problems. Eventually, it causes to develop any eye complication and blindness. As the diabetic retinopathy condition progresses, this symptom may include Spots or dark strings floating in the vision (floaters), Blurred vision, Fluctuating vision, Impaired color vision, Dark or empty areas in your vision, Vision loss usually affect the both eyes

4.1. Diabetic Retinopathy detection

Diagnosis of Diabetic Retinopathy can be done manually based upon a variety of clinical findings by different eye care experts. It is not a single disease, but rather a group of diseases with some common characteristics. The manual detection and prediction of Diabetic Retinopathy are very difficult. It is time-consuming process and accuracy of the diseases depends on the ophthalmologist. So the medical diagnosis system is developed by using data mining technique for the detection and prediction of diabetic retinopathy

4.1.1. K-MEANS

K-means is one of the data mining algorithm .It is an unsupervised learning algorithm. The goal of k-means is to group data points together based on similar groups are called "Clusters". The 'K' simply tells that can be any number of groups. The below figure shows a data set that has been grouped using this technique.



K-means works by finding points called "centroids", then assigns each point in the data to a cluster based on its closest centroid and to find the optimum number and position of centroids to group the data properly. This is done in a very best way in this algorithm Choose the number K of centroids Randomly select positions for the centroids in multidimensional space Assign each data point to its closest centroid, creating K clusters Readjust the positions of the centroids to be the average (or mean) position of these new clusters this is called K-"means". Repeat steps 3 and 4 until the centroids no longer move! Once the centroids stop moving, and the data has been clustered. For best results, k-means is usually run multiple times with different random starting points with the tightest clusters.

4.1.2. SUPPORT VECTOR MACHINES (SVM)

Support vector machines (SVM) are consider one of the most robust and accurate methods among all well-known algorithms. And efficient methods for training SVM are also being developed at a fast pace.In a class learning task, the main intention of SVM is to find the best classification function to distinguish between members of the two classes in the training data. The metric of the concept for the "best" classification function can be realized geometrically for the linearly separable dataset, a linear classification function corresponds to a separating hyperplane f (x) this passes through the middle of the two classes can separating as a two. If the function determined once the new data instance x_n can be classified by simply testing the sign of the function f (x_n); x_n belongs to the positive class if f (x_n) >0.Because there are many such linear hyperplane, what SVM additionally guarantee is that The best function is found by maximizing the margin between the two classes. Firstly the margin is defined as the amount of space, or separation between the two classes as defined by the hyperplane. The margin is corresponds to the shortest distance between the closest data points to a point on the hyperplane. Having this geometric definition allows us to explore to maximize the margin, and are an infinite number of hyper planes, only a few qualify as the solution to SVM.

The SVM insists because of finding the maximum margin hyper planes this can offer the best generalization ability. And also allows not only the best classification performance (e.g., accuracy) on the training data, but also leaves much room for the correct classification of the future data. To ensure for the maximum margin hyperplane are actually found, an SVM classifier attempts to maximize the following function with respect to w and b:

LP=1/2w- i=0^{t} ai.yi(w.xi+b)+ t i=0 ai1

Where t is the number of training examples, and α_i, $i = 1, \ldots, t$, are non-negative numbers such that the derivatives of L_P with respect to α_i are zero. A_i *are* the Lagrange multipliers and L_P is called the Lagrangian. The equation shows, the vectors **w** a n d constant *b to* define the hyperplane.

There are several important questions and related extensions on the e basic formulation of support vector machines. The list of some questions and extensions are given below

1. Can understand the meaning of the SVM through a solid theoretical foundation?
2. Can extend the SVM formulation to handle cases where allow errors to exist

Even though the best hyperplane must admit some errors on the training data?

3. Can extend the SVM formulation so that it works in situations where the training data are not linearly separable

. Can extend the SVM formulation so that the task is to predict numerical values or to rank the instances in the likelihood of being a positive class member, rather than classification?

Can scale up the algorithm for finding the maximum margin hyperplane to thousands and millions of instances?

Question 1 can understand the meaning of the SVM through a solid theoretical foundation?

There are Several important theoretical results exist to answer this question for learning machine, like a SVM, can be modeled as a function class based on some parameters α. The Different function classes can have different capacity in learning, this is represented by a parameter h known as the VC dimension The VC dimension measures the maximum number of training sample for the function class can be used to learn perfectly, by obtaining zero error rates on the training data, for any assignment of class labels on these points. This can be proven that the actual error on the future data is bounded by a sum of two terms. The first term is the training error, and the second term if proportional to the square root of the VC dimension h. Thus, if it can minimize h that can minimize the future error, as long as it also minimize the training error. In fact, the above maximum margin function learned by SVM learning algorithms is one such function. Thus, theoretically, the SVM algorithm is well founded

2 Can we extend the SVM formulation to handle cases where we allow errors to exist, when even the best hyperplane must admit some errors on the training data?

To answer this question, imagine that there are a few points of the opposite classes that cross the middle. These points represent the training error that existing even for the maximum margin hyperplane. The "soft margin" idea is aimed at extending the SVM algorithm [83] so that the hyperplane allows a few of such noisy data to exist. In particular, introduce a slack variable ξ_i *to* account for the amount of a violation of classification by the function $f(x_i)$; ξ_i has a direct geometric explanation through the distance from a mistakenly classified data instance to the hyperplane $f(x)$. Then, the total cost introduced by the slack variables can be used to revise the original objective minimization function

Question 3 Can extend the SVM formulation so that it works in situations where the training data are not linearly separable?

The answer to this question depends on an observation on the objective function where the only appearances of $\mathbf{x_i}$ is in the form of a dot product. Thus, if we extend the dot product $\mathbf{x_i} \cdot \mathbf{x_j}$ through a functional mapping $(\mathbf{x_i})$ of each $\mathbf{x_i}$ to a different space H of larger and even possibly infinite dimensions, then the equations still hold. In each equation, where had the dot product $\mathbf{x_i} \cdot \mathbf{x_j}$, we now have the dot product of the transformed vectors $(\mathbf{x_i}) \cdot (\mathbf{x_j})$,which is called a kernel function.

The kernel function this can be used to define a many type of nonlinear relationship between those inputs. For example, besides linear kernel functions, that can define quadratic or exponential kernel functions. Much study in recent years has gone into the study of various kernels for SVM classification and for many other statistical tests. This can also extend the above descriptions of the SVM classifiers from binary classifiers to problems that involve more than two classes. This can be done by repeatedly using one of the classes as a positive class, and the rest as the negative classes (thus, this method is known as the one-against-all method).

The SVM can be easily extended to perform the numerical calculations. Here this discuss two such extensions. The first step is to extend SVM to perform regression analysis, and the goal is to produce a linear function that can approximate to target function. The Careful consideration goes into the choice of the error models; in support vector regression, or SVR, the error is defined to be zero when the difference between actual and predicted values is within a epsilon amount. the epsilon insensitive error will grow linearly. The support vectors can be learned through the minimization of the Lagrangian. An advantage of support vector regression is given to be its insensitivity to outliers.

The next extension is to learn and to rank elements rather than producing a classification for individual elements. Ranking can be reduced to comparing pairs of instances and producing a +1 estimate if the pair is in the correct ranking order and −1 otherwise. Thus, a way to reduce this task to SVM learning is to construct new instances for each pair of ranked distance in the training data, and to learn a hyperplane on this new training data. this pattern can be applied to many areas where ranking is important, such as in document ranking in information retrieval areas.

Question 5 Can scale up the algorithm for finding the maximum margin hyperplane to thousands and millions of instances?
The computational inefficiency is one of the primary drawbacks of SVM this problem is being solved with great success. One approach is to break a large optimization problem into a series of smaller problems; here each problem only involves a couple of carefully chosen variables. The optimization can be done efficiently. This process iterates until all the decomposed optimization problems are solved successfully. the recent approach is to consider the problem of learning an SVM as that of finding an approximate minimum enclosing ball of a set of instances.
This instance when mapped to an N-dimensional space to represent a core set .this can be used to construct an approximation to the minimum enclosing ball. The Solving of SVM learning problem on these core sets can produce a good approximation solution in very speed manor. For example, the core-vector this produced can learn an SVM for millions of data in seconds.

4.1.3 The Apriori algorithm

The data mining approaches is the most popular one to find frequent item sets from a transaction dataset and derive association rules. The finding frequent item set with frequency are larger than or equal to a user specified minimum support is not trivial because of its combinatorial explosion. After the frequent item sets are obtained once then it is straightforward to generate association rules with confidence larger or equal to a user specified minimum confidence.
The Apriori algorithm is a seminal one for finding frequent item sets using candidate generation. This is characterized as a level-wise complete search algorithm using anti-monotonicity of item sets, "if an item set is not frequent, then any of its superset is never frequent". By convention, Apriori algorithm assumes that items within a transaction or item set are sorted in lexicographic order. If the set of frequent item sets of size k be F_k and their candidates will be C_k. The Apriori first scans and Generate C_{k+1}, candidates of frequent item sets of size $k+1$, from the frequent item sets of size k. The Scanned database calculates the support of each candidate of frequent item sets. And ad d those item sets to satisfy the minimum support requirement for F_{k+1}.

The Apriori algorithm is shown below and Function of Apriori in line 3 generates C_{k+1} From F_k in the following two step process:1.Joining step: Generate R_{K+1}, the initial candidates of frequent item sets of size $k+1$ by taking the union of the two frequent item sets of size k, P_k and Q_k that have the first $k-1$elements in common.

$$R_{K+1} = P_k \cup Q_k = \{item1, \ldots, item_{k-1}, item_k, item_{k'}\}$$

$$P_k = \{item1, item2, \ldots, item_{k-1}, item_k\}$$

$$Q_k = \{item1, item2, \ldots, item_{k-1}, item_{k'}\}$$

Where, $item1 < item2 < \cdots < item_k < item_{k'}$.

Second one is the Prune step: this Check all the item sets of size k in R_{k+1} are frequent and generate C_{k+1} by removing those that do not pass this requirement from R_{k+1}.This is because any subset to size k of C_{k+1} that is not frequent cannot be a subset of a frequent item set of size $k+1$.

Function subset in line 5 finds all the candidates of the frequent item sets included in trans- actions t then the Apriori calculates frequency only for those candidates generated this way by scanning the database.
This is a evident for the Apriori to scans the database at most k_{max+1} times when the maximum size of frequent item sets is set at k_{max}.
The Apriori achieves good performance by reducing the size of candidate sets at in any situations with very many frequent item sets, large item sets, or very low minimum support; this suffers from the cost of generating with a large number of candidate sets

4.1.4 Page rank

Page Rank was presented and published by Sergey Brin and Larry Page at the Seventh International World Wide Web Conference (WWW7) in April 1998. this is a search ranking algorithm by using hyperlinks on the Web. Based on this algorithm, the search engine Google has been built, which has been a huge success. Now a days every search engine has its own hyperlink based ranking method.
The Page Rank produces a static ranking of Web pages in the sense that a Page Rank value is computed for each page off-line and it does not depend on search queries. This algorithm relies on the democratic nature of the Web by using its vast link structure as an indicator of an individual page's quality in web for essence and Page Rank interprets a hyperlink from page x to page y as a vote, by page x, for page y. However, Page Rank looks at more than just the sheerer of votes, or links that a page receives. This analyzes the page and casts the vote. The Votes are casted by pages that are themselves "important" weigh more heavily and help to make other pages more "important". This is exactly the idea of rank prestige in social networks .This algorithm introduce the Page Rank formula. First it states some main concepts in the Web context In-links of page. The hyperlinks that can point to page from i to other pages. Normally the hyperlinks from the same site are not considered.

Out-links of page. the hyperlinks that can point out to other pages from page i. Usually y, links to pages of the same site are not considered.

The following ideas based on rank prestige are used to derive the Page Rank algorithm:

1. The hyperlink pointing from a page to another page is an implicit conveyance of authority to the target page. This is more in-links that page i receives, and more prestige the page i has. The Pages that point to page i also have their own prestige scores. the page with a higher prestige score pointing is to i is more important than a page with a lower prestige score pointing to i.Otherwise, a page is important if it is pointed to by other important pages.

According to rank prestige in social networks, the importance of page i (i's Page Rank score) is determined by summing up the Page Rank scores of all pages that point to i this a page can point to many other pages, it's prestige score should be shared among all the pages that it points to.

To formulate the above ideas, this treat the Web as a directed graph $G = (V, E)$, here V is the set of vertices or nodes, this means it is set of all pages, and E is the set of directed edges in the graph, that is, hyperlinks. Let the total number of pages on the Web be n (i.e., $n = |V|$).The Page Rank score of the page i (denoted by $P(i)$) is defined by

Where O_j is the number of out-links of page j.Mathematically, this system has of n linear equations (12) with n unknowns. This can use a matrix to represent all the equations. Let P be a n-dimensional column vector of Page Rank values, i.e.,

the equation $p = (p(1), p(2), p(n))$ T.

Let A be the adjacency matrix of our graph with this can write the system of n equations with $P = A^T P$ is the characteristic equation of the Eigen system, where the solution to P is aneigenvector with the corresponding Eigen value of 1. Since this is a circular definition, aniterative algorithm is used to solve it. It turns out that if some conditions are satisfied, 1 is the largest Eigen value and the Page Rank vector P is the principal eigenvector. One of the well known mathematical techniques called power iteration this can be used to find P. However; the problem is that Eq. this is not quite sufficed because the Web graph does not meet the conditions. The equation can also be derived based on the Markov chain. There are some theoretical results from Markov chains can be applied. After augmenting the Web graph to satisfy the conditions, the following Page Rank equation is produced:

$$P = (1 - d)e + dA^T P,$$

Here e is a column vector of all 1's. This gives us the Page Rank formula for each page i:

$$P(i) = (1 - d) + d \sum_{j=1}^{n} A_{ji} P(j),$$

Which s equivalent to the formula given

$$P(i) = (1 - d) + d \sum_{(j,i)\in E} \frac{P(j)}{O_j}.$$

The parameter d is called the damping factor which can be set to a value between 0 and1. d = 0.85 is used in [.The computation of PageRank values of the Web pages can be done using the power iteration method which produces the principal eigenvector with the eigen value of 1. The algorithm is simple, and is given in. One can start with any initial assignments of Page Rank values. The iteration ends when the Page Rank values do not change much or converge. The iteration ends after the 1-norm of the residual vector is less than a pre-specified threshold. Since in Web search, this are only interested in the ranking of the pages, the actual convergence may not be necessary. Thus, less iteration is needed. This is reported on a database of above 322 million links the algorithm converges to an acceptable tolerance in roughly 52 iterations.

4.2. Automated Diabetic Retinopathy Detection Process

The automatic medical diagnosis system shown below in the figure. The retinal image of the patients is taken by using fundus camera. The images are in irregularities shape. So the preprocessing is performed in retinal image to enhance the quality of the image by equalizing the irregularities of the image.

Feature extraction is a dimensionality reduction process to extract the various retinal images for identifying of interpreting. The some of the features are extracted from retinal images are vascular network, exudates and optic disc. Using classifier PCA and Bays Classier are used to detect the condition of the diabetic retinopathy and the sugar level.

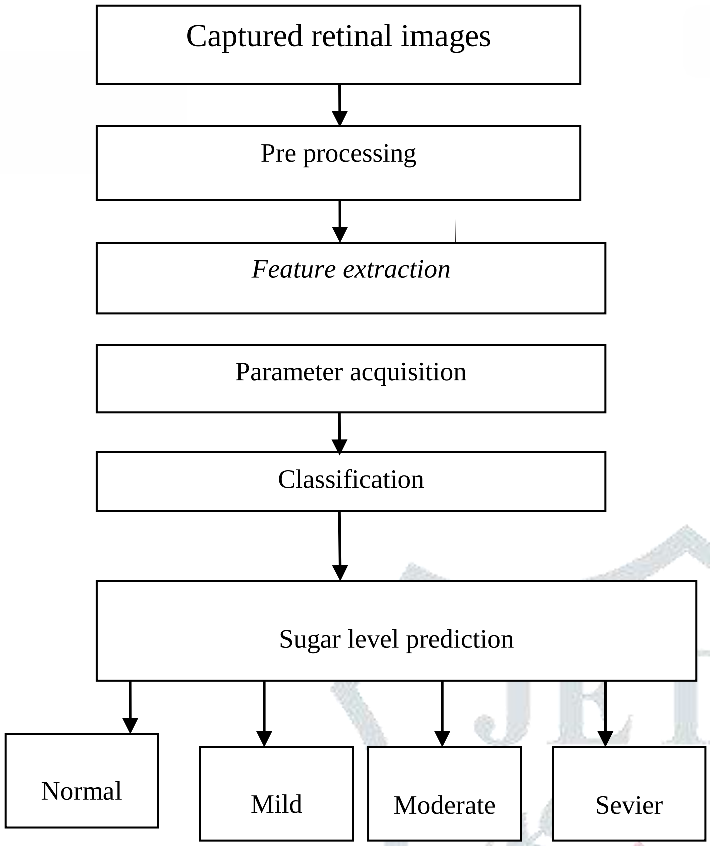


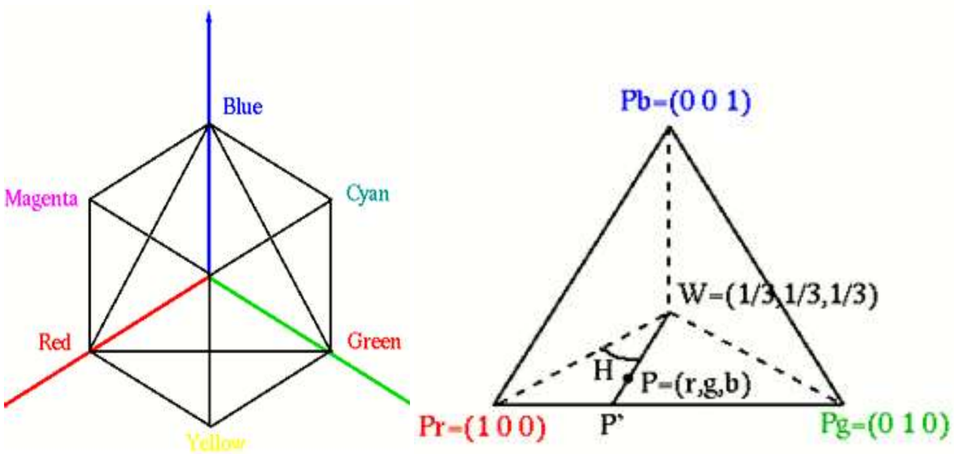
Figure 1.the automatic process for diabetic retinopathy detection

5. Methodologies used to detect diabetic retinopathy

5.1. Preprocessing: these steps transformed each original RGB image into the HSI (hue, saturation, and intensity) color space. To reduce noise and improve image contrast, and applied median filtering for Contrast-Limited Adaptive Histogram Equalization

5.2 Conversion of RGB image to HIS

The RGB Cubes defines RGB model, the HIS is represented by triangular color .The intensities of the three primaries R, G, and B of a color, its HSV representation using different models. Here the RGB plane of the cube to find the corresponding HSV. The three vertices are represented pr= (1, 0, 0), Pg= (0, 1, 0) and Pb= (0, 0, 1) and the three components of the given color is represented by a point P=(R, G, B) in the RGB 3-D space. Here the intensities are normalized so that the R, G and B values are between 0 and 1, and point P is inside or on the surface of the color cube.



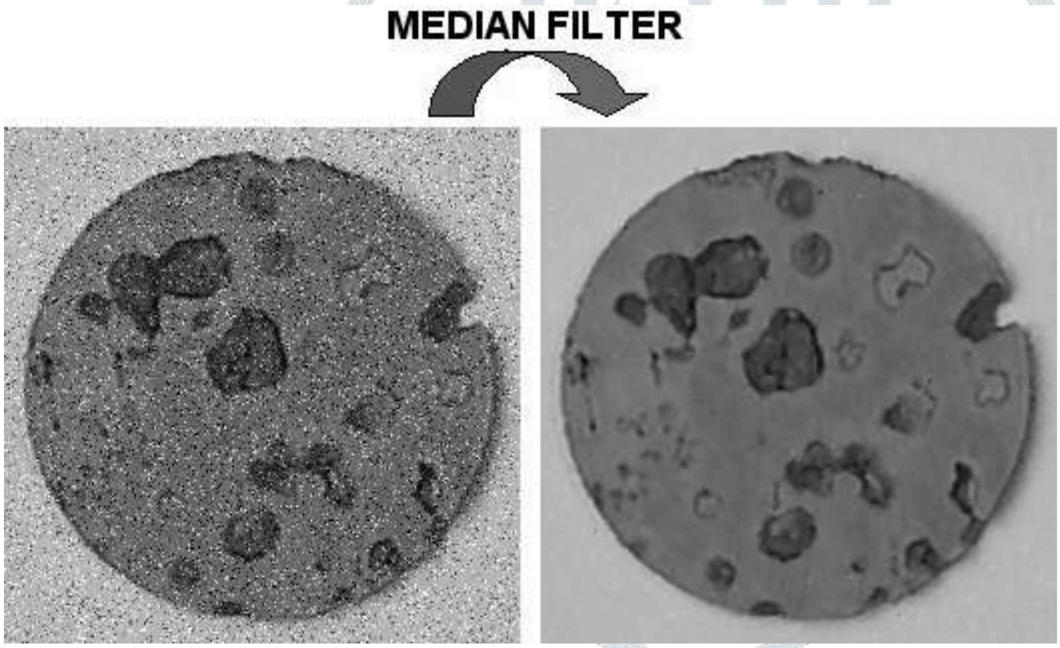
Determine the intensity I: The intensity I can be defined as: I=1/3(R+G+B), i.e. R+G+B=3I. Determine the hue H : first find the intersection of the color vector P=(R,G,B) with the RGB plane triangle determined by equation: R+G+B=1 .Determine S: Determine S: The saturation of the colors on any of the three edges of the RGB triangle is defined as 1 (100% saturated), and the saturation of w=(1/3,1/3,1/3) is zero. That is denoted as p1 the intersection of the extension of line WP With edge. The normalized color is p=w,s=0,and p=p1,s=1. The saturation of any color point p between w. the given the R, G, and B components of a color in the cubic color model, that can find the H, S, and I values by: I=(R+G+B)/3. The RGB to HSI conversion, need to normalize the R, G, and B components of a pixel in a color image from the range of (0,255) to(0,1).

5.3 Histogram Equalization

Digital images are composed of discrete values of intensity they are between 0 and 255. Histogram of images shows frequency of pixel's intensity values of x axis and y axis are represented by the gray level intensities. The x axis of the histogram pixel values range is an 8bp image that means it has 256 levels of gray level. The range of x axis value starts from 0 to 255 with a gap of 50. And also the y axis is the count of these intensities.

5.4. Median Filtering

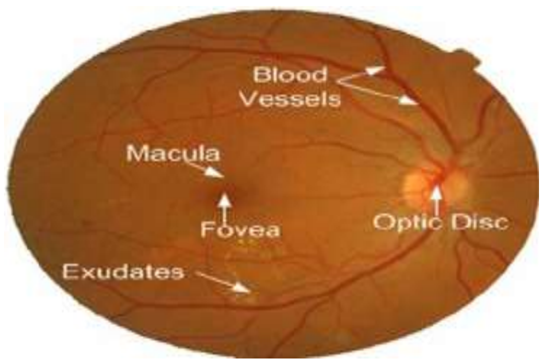
Filtering is a technique for modifying or enhancing the quality of image and also making them for some other function. Image enhancement technique has two domains: 1. Spatial Domain 2.Frequency Domain Filtering. Spatial filtering operations are performed on the pixels inside the neighborhood output image. A filtered image moves to every pixel in the input image. output image Frequency filters process an image with respect to the operator are taken an image and a filter function is in a pixel-by-pixel fashion and smoothing images by reducing the intensity variation between one pixel to the next and to reduce noise (removing 'salt and pepper' type noise)in images.



6. Detection of Main component in retinal images

6.1. Introduction:

The features of diabetic retinopathy from the retina are optic disc, blood vessels, macula, fovea, and exudates



6.2. Detection of Optic Disc

Optic disc is a bright area on the retina where blood vessels converge. The optic disc represents the beginning of the optic nerve and is the point where the axons of retinal ganglion cells come to the optic disc is also the entry point for the major blood vessels that supply in the retina. The optic disc in a normal human eye carries 1–1.2 millionth optic disc is placed 3 to 4 mm to the nasal side of the fovea It is a vertical oval shape, with dimensions of 1.76mm horizontally by 1.92mm vertically and There is a central depression, of variable size, called the optic cup. This depression can be a variety of shapes from a shallow indentation to a bean pot. This shape will be significant for diagnosis of some retinal image disease. The optic disc or optic nerve head is the exit point of ganglion cell axons leaving the eye. Because of there is no rods or cones overlying the optic disc, it corresponds to a small blind spot in each eye. A normal optic disc is orange to pink in color. A pale disc is an optic disc this is varying in color from a pale pink or orange color to white. A pale condition of disc is an indication of a disease

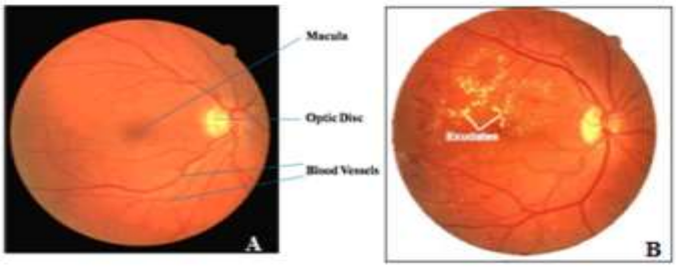
6.3. Detection of blood vessels

The retinal blood vessels are the major vascular network of a retina. The retinal blood vessels originate from the center of OD and spreads over the region of the retina. The blood vessels are taking response for supplying the blood throughout the entire region of the retina. Micro aneurysms, hemorrhages, and exudates lesions are formed in retinal image due to the damage in retinal blood vessels. The retinal blood vessel detection and segmentation are used to detect and diagnosis of the retinal images for the ranking of the diseases severity.

6.4. Detection of Exudates

The Exudates are bright yellow spots on the surface of the retina and are the primary sign of DR. the detection of exudates is very important in diagnosis of diabetic retinopathy. Exudates are randomly spread over the retina and appear as yellow-white patches of in different sizes and shapes that are the primary sign of diabetic retinopathy. The exudates are classified as two types. They are soft exudates and hard exudates.

The cotton wool spot is called soft exudates. They are flat-reddish white in color and obscure underlying the retinal blood vessels.. The hard exudates are intra-retinal fatty which are the important sign of the DR and macula edema. the detection of exudates are main aim which correlates the DR and macula edema to prevent the earlier vision loss in diabetic patients. By analyzing the exudates and fovea region, the severity of DR can be easily identified to prevent vision loss in the diabetic patients.



A. Original retinal image B. Exudates retinal image

7. Conclusion:

Early detection of diabetic retinopathy is identified based on retinal images by using data Mining technique to help ophthalmology and also it supports for future implementation from ATM Machine the people may know their sugar level.

References

1. He Huang He Ma, Han JW van Triest ,Yinghua Wei, Wei Qian ,"Automatic detection of neovascularization in retinal mages using extreme learning machine" : Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang 110004, PR China,2017.

2. .Bo Wu, Weifang Zhu, Fei Shi, Shuxia Zhu, Xinjian Chen, "Automatic detection of microaneurysms in retinal fundus images": Computerized Medical Imaging and Graphics, ,2017.

3. Evangelia Kotsiliti, Bashir Al-Diri, Xujiong Ye, Maged Habib, Andrew Hunter,"A classification model for predicting diabetic retinopathy based onpatient characteristics and biochemical measures":Journal for Modelling in Ophthalmology, Vol 1,2017.

4. Gargeya.R, Leng .T,"Automated identification of diabetic retinopathy using deep learning": Opthalmology 2017 Jul;124(7):962-969. doi: 10.1016/j.ophtha.2017.02.008. Epub 2017 Mar 27.

5. Rezty Amalia Aras*, Tri Lestari, Hanung Adi Nugroho, Igi Ardiyanto," Segmentation of retinal blood vessels for detection of diabetic retinopathy": Communications in Science and Technology 1 (2016).

6. Karan Bhatia ,Shikhar Arora ,Ravi Tomar," Diagnosis of diabetic retinopathy using machine learning classification algorithm": IEEE Conference on Next Generation Computing Technologies,2016.

7. [8]Md.Jahiruzzaman, A.B.M Aowlad Hossain, "Detection and classification of diabetic retinopathy using K-means clustering and fuzzy logic":Computer and Information Technology,2015.

8. Gowsalya P, Member IEEE, Mrs.S.Vasanthi, " Segmentation and Classification of Features in Retinal Images", International Conference on Communication and Signal Processing, April 3-5, 2014.

9. Sumathy.B1 , Dr.Poornachandra.S2,"Feature Extraction in Retinal Fundus Images":IEEE,2013.

10. Ahmad Zikri Rozlan, Hadzli Hashim, Syed Farid Syed Adnan, Chen Ai Hong Miswanudin Mahyudin," A Proposed Diabetic Retinopathy Classification Algorithm with Statistical Inference of Exudates Detection":Electrical , Electronics and System Engineering,2013.

11. Pilar Pérez Conde, Jorge de la Calleja, Antonio Benitez,"Image-based classification of diabetic retinopathy using machine learning",Intelligent Systems Design and Applications(ISDA),2013.

12. .Photocoagulation for diabetic macular edema. Early Treatment Diabetic Retinopathy Study report number 1. Early Treatment Diabetic Retinopathy Study research group **Arch Ophthalmol, 103 (12) (1985),**

13 Photocoagulation for diabetic macular edema. Early Treatment Diabetic Retinopathy Study report number 1. Early Treatment Diabetic Retinopathy Study research group Arch phthalmol, 103 (12) (1985),

14. Cheung, I.Y. Wong, T.Y. Wong**Ocular anti-VEGF therapy for diabetic retinopathy: overview of clinical efficacy and evolving applications**Diabetes Care, 37 (4) (2014),

15 .L.P. Aiello, T.W. Gardner, G.L. King, *et al.***Diabetic retinopathy** Diabetes Care, 21 (1) (1998),

15. .Q. Mohamed, M.C. Gillies, T.Y. Wong**Management of diabetic retinopathy: a systematic review**JAMA, 298 (8) (2007),

17. World Health Organization**Global status report on noncommunicable diseases 2014** Accessed June 2016

18..N. Congdon, B. O'Colmain, C.C. Klaver, *et al.***Causes and prevalence of visual impairment among adults in the United States**Ophthalmol, 122 (4) (2004),

19. International Diabetes Federation. IDF Diabetes Atlas. 6 th ed. Brussels, Belgium: International DiabetesFederation; 2013.

20. Haritoglou C, Kernt M, Neubauer A, Gerss J, Oliveira CM, Kampik A, *et al.* Microaneurysm formation rate as a

21.predictive marker for pogression to Clinically significant macular edema in nonproliferative diabetic retinopathy. Retina 2014

22. Ribeiro ML, Nunes SG, Cunha-Vaz JG. Microaneurysm turnover at the macula predicts risk of development of clinically significant macular edema in Persons with mild nonproliferative diabetic retinopathy. Diabetes Care 2013

23. Thomas RL, Dunstan F, Luzio SD, Roy Chowdury S, Hale SL, North RV, *et al.* Incidence of diabetic retinopathy in people with type 2 diabetes mellitus attending the Diabetic Retinopathy Screening Service for Wales: etrospective analysis. BMJ 2012