A Survey on Machine Learning Based Techniques For Detection of Glaucoma

Nidhi Pareek[1], Prof. Virendra Verma[2]

Abstract- *Glaucoma is often termed as the silent snatcher of eyesight. It is one of the leading causes for blindness which is caused due to abnormally high pressure created in the eye, causing damage to the optic nerves. With the advancements in image processing and machine learning techniques, automated detection of glaucoma has been an active area of research. The retinal fundus images is analysed for detection of glaucoma which is critical to prevent loss of vision. An early detection of glaucoma can help to arrest the condition and avoid further degradation to the optic nerves and vision quality. This paper analyses the various techniques pertaining to retinal image processing, noise removal, segmentation, feature computation and classification. Different machine learning based classifiers are evaluated in terms of the performance metrics such as accuracy, sensitivity, precision and recall. A comparative analysis of the noteworthy contribution in the field is also presented which can lay the foundation for the development of novel algorithms to address the challenge of early and accurate detection of glaucoma.*

Keywords- Fundus Image, Optic Disc, Image Processing, Segmentation, Feature Extraction, Cup to Disc Ratio (CDR), Machine Learning, Accuracy.

I. INTRODUCTION

The retina is a layered tissue which is responsible for conversion of light signal into electrical signals for the brain to garner a sense of sight. Thus the retina is responsible for the creation of the optic images in the eye which are perceived by the brain as equivalent electrical signals giving a view of the outside world. In general, several physical ailments can result in the damage to the optic nerve and retinal function, among which hypertension, diabetes and other cardiovascular diseases are prominent [1]. A significant part of the above causes is diabetic retinopathy which is caused by excess blood sugar levels in the blood streams causing irreversible damage to the heart, lungs, kidneys and eyes. Typically, patients suffering from type-2 diabetes often suffer from diabetic retinopathy and this causes serious damage to the peripheral vision in humans. In developed countries, glaucoma is also termed as the silent snatcher of eyesight, which can only be arrested but not reversed making its early and accurate detection extremely crucial. With sedentary lifestyles, poor eating habits, depleting nutritional values in food intake, these diseases are often interconnected and are also progressing fast worldwide. A typical retinal image is depicted in figure 1 which consists of blood vessels, optic disc (OD), Optic Cup (OC) and the macular region. The blood vessels perform the task of supplying blood to the entire retinal area by spreading across the length and breadth of the retinal region. The damage to the optic nerves and vessels are caused by excessive blood pressure to the vessels primarily caused by:

1) Hyper tension
2) High blood glucose levels.

The additional pressure of the blood modifies the elasticity of the blood vessels and also their texture resulting in macular degradation [2]-[3]. Often this can be perceived as flashes and floaters, loss of peripheral vision and erroneous perception of depth.

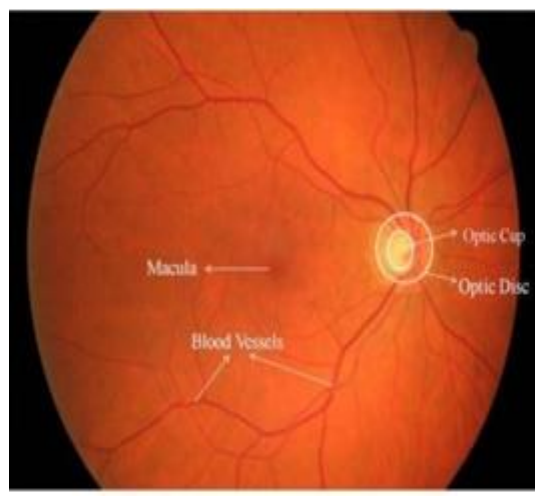

Fig. 1 A typical retinal image
(Source: https://www.aao.org/eye-health)

The pivotal aspect in the detection and treatment of glaucoma is the process of retinal imaging which is 2-Dimensional projection of the 3-Dimensional retinal image on an imaging plane. Such images are also termed as fundus images [4]. The different retinal imaging techniques commonly employed are:

1.1 Fundus Photography:

In the case of fundus photography, a wideband (in visible spectrum) reflected light representation is captured which constitutes the fundus image.

1.2 Colour Fundus Photography:

This is a modified version of the fundus photography wherein a three channel representation of the reflected light is images in terms of the R-G-B values in the waveband.

1.3 Stereo Fundus Photography:

This imaging technique comprises of combining light waves form two or more visual angles to garner information regarding the depth of resolution. Lower resolution results in the patients impairment to distinguish between different objects and considering them as a composite object.

1.4 Hyper Spectral Imaging:

This technique consist of capturing a composite image created from light reflections from specific wavelengths. The intensity response of different wavelengths represents the capturing capability of the retina to respond to different colours.

1.5 Scanning LASER Ophthalmoscope (SLO):

The SLO technique analyses a temporal pattern of the reflection intensities to a fixed LASER. The process is based on the phenomenon of back scattering causing back scattering of light from the LASER source. A modified version of the technique constitutes Laser Scanning Tomography (SLT) which tries to gather multi-wavelength resolution by changing the wavelength of the Laser source.

1.6 Fluorescein angiography (FA):

This procedure consists of injecting a fluorescent dye into the blood which results in clear photography of the blood vessels in the eye. This type of imaging is very useful for diabetic retinopathy and can help to assess the amount of proliferation.

There are other reasons of vision loss among patients which are typically age related macular degradation, diabetic retinopathy, cardio vascular diseases affecting the vision and cataract. While glaucoma is not necessarily an outcome of the aforesaid health problems, yet studies show a dependence and close correlation among these common ailments and the onset of glaucoma among patients.

II. AUTOMATED DETECTION OF GLAUCOMA

The automated detection of glaucoma has gained prominence with advanced image processing techniques coupled with machine learning. The detection of glaucoma is generally based on three major steps which are image processing and data preparation, feature extraction and final classification. Each of these processes are discussed in this section [6]-[7]:

2.1 Image Processing:

Prior to computing important parameters or feature of the fundus image, which lays the foundation for the final classification, it is necessary to process the image for the following reasons:

Illumination Correction: In this part, the inconsistencies in the image illumination are corrected so as to make the image background uniform and homogenous. Illumination inconsistencies occur due to capturing the image from different angles which makes the reflection from the retinal image variable rendering inconsistencies. Inconsistencies in the illumination can be caused due to the position and orientation of the source, the non-homogeneity of wavelengths of the source, the nature of the surface such as smoothness, orientation and material characteristics and finally the characteristics of the sensing device such as resolution, capturing capability and sensitivity. Typically, illumination correction is done based on the computation of the correlation co-efficient given by:

$$Corr2(x,y) = \frac{O(x,y)-D(x,y)}{B(x,y)-D(X,y)}.N \quad (1)$$

Here,
Corr2 represents the 2 dimensional cross correlation,
N is termed as the normalizing factor
O represents the original image
D represents the dark image
B represents the bright image

The normalizing term 'N' is computed as:

$$N = \frac{mean\{O(x,y)\}}{\frac{O(x,y)-D(x,y)}{B(x,y)-D(X,y)}} \quad (2)$$

Here,
Mean represents the average value of the random variables (x,y) which are the pixel values of the images.

Another common approach is the use of low pass filtering which considers the inconsistencies to occur in the low spectral range of the image and are filtered out using a low pass filter (LPF), given by:

$$N_I = O(x,y) - LPF\{O(x,y)\} + mean[LPF\{O(x,y)\}] \quad (3)$$

Here,
N_I represents the normalized image
O represents the original image
LPF stands for the low pass filter

Segmentation, Vessel Removal and Optic Nerve Normalization: This process typically employs a process called image inpainting. In this process, the vessels are removed using threshold based segmentation and then the void regions are inpainted using the image statistics. Further, the optic nerve is normalized to compute the cup to disc ratio (CDR) accurately [8]. The segmentation is typically a threshold based segmentation since the parts to be separated are not generally regular shapes. The segmentation is generally done adopting the sudden change in pixel characteristics given by the gradient:

$$G = max(r, x_0, y_0)\left|K_\sigma(r)\frac{\partial}{\partial r}\oint_{r,x_0,y_0}^{r,x_f,y_f}\frac{I(x,y)}{2\pi r}ds\right| \quad (4)$$

Here,
G is the gradient
(x,y) are the image pixels
r represents the image radius
ds is the surface integral
K is a typically a Gaussian kernel

The gradient based method allows to find the maximum change in the pixel intensities to perform the thresholding so as to separate out the vessels. Further the inpainting can be performed based on the neighbouring pixel information and the stochastic characteristics utilizing the fact that image regions generally comprise of highly redundant values. Considering the region to be inpainted as S and its boundary to be δS, inpainting is performed by computing the stochastic parameters around δS and gradually moving towards S as:

$$\frac{\partial I}{\partial(x,y)} = \nabla(\Delta I).\nabla^{O} I \quad (5)$$

Here,
$\frac{\partial I}{\partial(x,y)}$ represents the partial derivative of the Image w.r.t. x & y
ΔI represents the Laplacian of I
∇^{O} represents the orthogonal gradient.

Finally the optical nerve is normalized to compute the CRD of the image.

2.2 Feature Extraction:

Feature Extraction: The classification of glaucoma is to be done using any automated classifier but machine learning classifiers need to be fed with image features or parameters which can help the classifier to learn the differences between the glaucoma positive and glaucoma negative images [9]. Typically, the features care stochastic features of the image such as the mean, variance, standard deviation, skewness, correlation etc. The feature extraction is critically important since the classifier would decide based on the feature vector whether any new image is glaucoma positive or negative [10]-[11]. The GLCM features thus help to evaluate the co-occurrence of the pixel values in an image and hence can be used to judge the similarity or redundancies in the image pixel regions.

2.3 Classification:

Based on the image processing and feature extraction, the classification is done. Automated classification requires training a classifier with the pre-defined and labelled data set and subsequently classifying the new data samples [12]. Off late machine learning based classifiers are being used for the classification problems [13]. Machine learning based classifiers are typically much more accurate and faster compared to the conventional classifiers. They render more robustness to the system as they are adaptive and can change their characteristics based on the updates in the dataset [14]. The common classifiers which have been used for the classification of glaucoma cases are:

Regression Models: In this approach, the relationship between the independent and dependent variable is found utilizing the values of the independent and dependent variables. The most common type of regression model can be thought of as the linear regression model which is mathematically expressed as [15]:

$$y = \theta_1 + \theta_2 x \quad (6)$$

Here,
x represents the state vector of input variables
y represents the state vector of output variable or variables.
Θ1 and Θ2 are the co-efficients which try to fit the regression learning models output vector to the input vector.

Often when the data vector has large number of features with complex dependencies, linear regression models fail to fit the input and output mapping. In such cases, non-linear regression models, often termed as polynomial regression is used. Mathematically, a non-linear or higher order polynomial regression models is described as:

$$y = \theta_0 + \theta_1 x^3 + \theta_2 x^2 + \theta_3 x \quad (7)$$

Here,
x is the independent variable
y is the dependent variable
$\theta_1, \theta_2 \ldots\ldots \theta_n$ are the co-efficients of the regression model.
Typically, as the number of features keep increasing, higher order regression models tend to fit the inputs and targets better. A typical example is depicted in figure 2

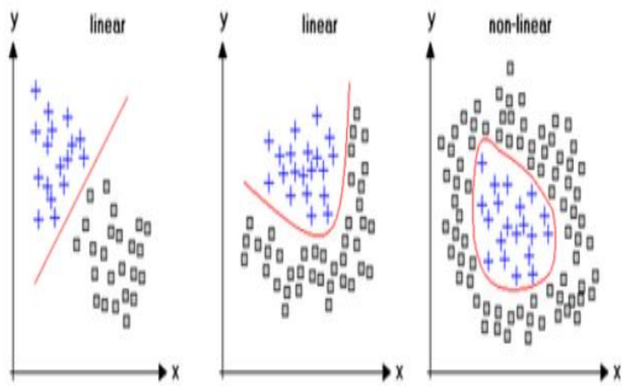

Fig. 2 Linear and Non-Linear Regression fitting [14]

Support Vector Machine (SVM): This technique works on the principle of the hyper-plane which tries to separate the data in terms of 'n' dimensions where the order of the hyperplane is (n-1). Mathematically, if the data points or the data vector **'X'** is m dimensional and there is a possibility to split the data into categories based on **'n'** features, then a hyperplane of the order **'n-1'** is employed as the separating plane. The name plane is a misnomer since planes corresponds to 2 dimensions only but in this case the hyper-plane can be of higher dimensions and is not necessarily a 2-dimensional plane. A typical illustration of the hyperplane used for SVM based classification is depicted in figure 3.

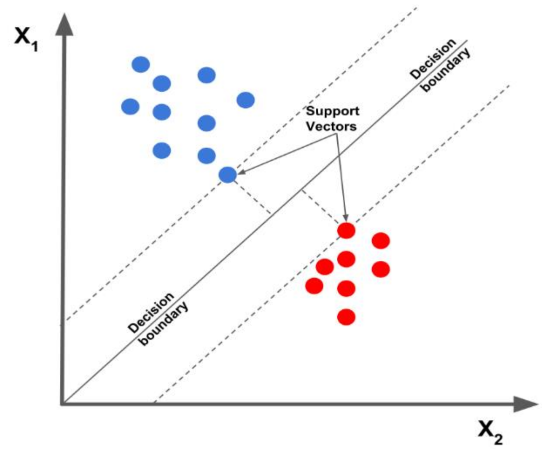

Fig. 3 Separation of data classes using SVM [15]

The selection of the hyperplane H is done on the basis of themaximum value or separation in the Euclidean distance d given by:

$$d = \sqrt{x_1^2 + \cdots \ldots \ldots . x_n^2} \quad (8)$$

Here,
x represents the separation of a sample space variables or features of the data vector,
n is the total number of such variables
d is the Euclidean distance

The (n-1) dimensional hyperplane classifies the data into categories based on the maximum separation. For a classification into one of **'m'** categories, the hyperplane lies at the maximum separation of the data vector **'X'.** The categorization of a new sample **'z'** is done based on the inequality:

$$d_x^z = Min(d_{C1}^z, d_{C2}^z \ldots d_{C2=m}^z) \quad (9)$$

Here,
d_x^z is the minimum separation of a new data sample from **'m'** separate categories
$d_{C1}^z, d_{C2}^z \ldots d_{C2=m}^z$ are the Euclidean distances of the new data sample **'z'** from m separate data categories.

Neural Networks: Owing to the need of non-linearity in the separation of data classes, one of the most powerful classifiers which have become popular is the artificial neural network (ANN). The neural networks are capable to implement non-linear classification along with steep learning rates. The neural network tries to emulate the human brain's functioning based on the fact that it can process parallel data streams and can learn and adapt as the data changes. This is done through the updates in the weights and activation functions. The mathematical model of the neural network is depicted in figure 4.

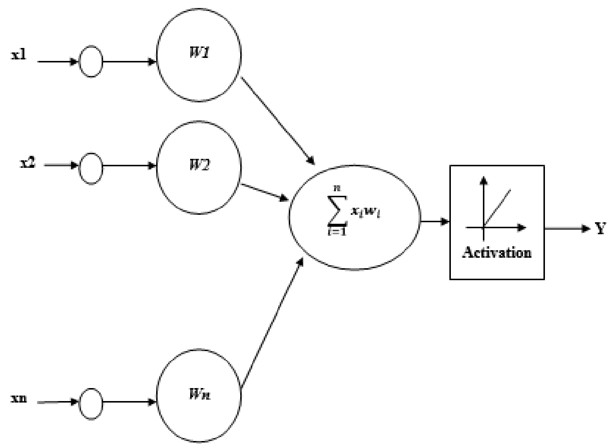

Fig. 4 Mathematical Model of Single Neuron [13]

The mathematical equivalent of an artificial neuron is depicted in figure 4 where the output can be given by:

$$y = f\left(\sum_{i=1}^{n} x_i w_i + b\right) \quad (10)$$

Here,
x denote the parallel inputs
y represents the output
w represents the bias
f represents the activation function

The neural network is a connection of such artificial neurons which are connected or stacked with each other as layers. The neural networks can be used for both regression and classification problems based on the type of data that is fed to them. Typically the neural networks have 3 major conceptual layers which are the input layer, hidden layer and output layer. The parallel inputs are fed to the input layer whose output is fed to the hidden layer. The hidden layer is responsible for analysing the data, and the output of the hidden layer goes to the output layer. The number of hidden layers depends on the nature of the dataset and problem under consideration. If the neural network has multiple hidden layers, then such a neural network is termed as a deep neural network. The training algorithm for such a deep neural network is often termed as deep learning which is a subset of machine learning. Typically, the multiple hidden layers are responsible for computation of different levels of features of the data. Several categories of neural networks such as convolutional neural networks (CNNs), Recurrent Neural Network(RNNs) etc. have been used as effective classifiers.

Previous Work

This section cites the various contemporary approaches developed for the automated classification of glaucoma.

Table I. Previous Work.

Authors	**Technique**	**Advantages**	**Limitations**
Memari et al. [1]	Fuzzy C-means clustering.	Vessel center lines clearly enhanced with labelled Fuzzy training.	Saturation of performance after which adding training data doesn't improve performance.
Mendonca et al. [2]	Multi-scale morphological enhancement technique	Vessel center lines were clearly enhanced	Not suitable for low resolution retinal images
Perez et al. [3]	Morphological Reconstruction.	Low complexity.	Image inpainting not incorporated after segmentation of vessel structures.
Marin et al. [4]	Gray level and moment feature based supervised classifier	Background enhancement, with high contrast difference between vessel and Background.	Relatively low accuracy and sensitivity, due to computation of moment features only.
Fraz et al. [5]	Ensemble classifier	Suitable for both low and high resolution retinal images	Relatively higher computational complexity.
Xiao et al. [6]	Bayesian classifier	Clear segmentation of vessels from the background pixels	Relatively low sensitivity and saturation of performance with adding data to training set.

Bansal et al. [7]	Fuzzy Classifier	Relatively high accuracy.	Relative high computational complexity with increasing features and performance saturation for less number of features.
Budai et al. [8]	Frangi algorithm	Low computational complexity.	Relatively low accuracy owing to lesser number of features extracted.
Wong et al. [9]	Seeded mode tracking approach with feature computation.	Relatively high accuracy.	Not applicable for low resolution images.
Khawaja et al. [10]	Directional Multiscale Line Detectors	Robust multi-variate classifier.	Relatively low accuracy
Islam et al.[11]	Deep-learning-based approach	Relatively high accuracy with low and high level features extracted through hidden layers of Deep Neural Network.	Relatively high computational complexity.
Ceccon et al. [12]	Naïve Bayes Classifier	Probabilistic approach robust for overlapping features.	Background enhancement and noise removal not explored.

The performance metrics of the classifiers are generally computed based on the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) values which are used to compute the accuracy and sensitivity of the classifier, mathematically expressed as:

$$Ac = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

The aim of any designed approach is to attain high values of accuracy of classification along with other associated parameters. The computation complexity of the system often evaluated in terms of the number of training iterations and execution time is also a critically important metric which decides the practical utility of any algorithm on hardware constrained devices.

III. CONCLUSIONS

This paper introduces the background behind the occurrence of glaucoma and its repercussions. The need for image processing and automated classification has also been explained. The working of automated classifiers along with their attributed anddependence on feature extraction has been explained in detail. Different stages of the image processing and segmentation have been enlisted. The significance of different image features and extraction techniques have been clearly mentioned with their utility and physical significance. Various machine learning based classifiers and their pros and cons have been highlighted. The mathematical formulations for the feature extraction and classification gave been furnished. A comparative analysis of the work and results obtained has been cited in this paper. It can be concluded that image enhancement and feature extraction are as important as the effectiveness of the automated classifier, hence appropriate data processing should be applied to attain high accuracy of classification.

REFERENCES

[1] N. Memari, A. Ramli, B Saripan, S Mashohor and M Moghbe, “Retinal Blood Vessel Segmentation by Using Matched Filtering and Fuzzy C-means Clustering with Integrated Level Set Method for Diabetic Retinopathy Assessment”, *Journal of Medical and Biological Engineering*, 2019 vol. 39, pp. 713–731.

[2] A.M. Mendonca and Campilho,A. “Segmentation of Retinal blood vessels by combining the detection of centre lines and morphological reconstruction*”, IEEE Transactions on Medical Imaging*, 2006, vol. 25, no. 9, pp. 1200-1213.

[3] P.Perez, M Perez, ME. Perez and O. Arjona, “Parallel multiscale feature extraction and region growing: Application in retinal blood vessel detection”, *IEEE Transactions on Information Technology in Biomedicine*, 2010, vol. 14, no. 2, pp. 500-506.

[4] D. Marin, A. Aquino, E. Arias, and M. Bravo, “A new supervised method for blood vessel segmentation in retinal images by using gray–level and moment invariants–based features”, *IEEE Transactions on Medical Imaging*, 2011, vol. 30, pp. 146-158.

[5] M. Fraz, P. Remagnino, A.Hoppe, B.Uyyanonvara, A.Rudnicka, C.Owen andS.Barman, “An ensemble classification–based approach applied to retinal blood

vessel segmentation", *IEEE Transactions on Biomedical Engineering*, 2012 vol. 59, pp. 2538-2548.

[6] Z.Xiao, M.Adel and S.Bourennane, "Bayesian method with spatial constraint for retinal vessel segmentation",*Computational and Mathematical Methods in Medicine*, 2013, vol. 13, no. 401413, pp. 1-9.

[7] N.Bansal and M. Dutta, "Retina vessels detection algorithm for biomedical symptoms diagnosis", *International Journal of Computer Applications,*2013, vol. 71, no. 20, pp. 41-46.

[8] A.Budai, R.Bock, A.Maier, J.Hornegger, and G.Michelson, "Robust vessel segmentation in fundus images", *International Journal of Biomedical Imaging*, 2013, vol. 20, pp. 1-11.

[9] D.Wong,J.Liu, N.Tan and F.Yin, "Automatic detection of the macula in retinal fundus images using seeded mode tracking approach", *IEEE Annual International Conference of Engineering in Medicine and Biology Society (EMBC)* 2012, pp. 4950-4953.

[10] A. Khawaja, T. Khan, M. Khan, and S. Nawaz, "A Multi-Scale Directional Line Detector for Retinal Vessel Segmentation", *Sensors (Basel),*2019, vol. 19, issue 22, pp. 4949.

[11] M. Islam, J. Wang, S. Johnson, M. Thurtell, R. Kardon and M. Garvin, "A Deep-Learning Approach for Automated OCT En-Face Retinal Vessel Segmentation in Cases of Optic Disc Swelling Using Multiple En-Face Images as Input", *arvo journal*, 2020, vol. 9, issue 2.

[12] S. Ceccon, F. David, G.Heath, P.David Crabb and A. Tucker, "Exploring Early Glaucoma and the Visual Field Test: Classification and Clustering Using Bayesian Networks*", IEEE Journal of Biomedical and Health Informatics*, 2014, vol. 18, pp. 1008-1014.

[13] S.Haykin, "Neural Networks and Learning Machines", 3rd Edition, *Pearson Publications*.

[14] M.Hagan, "Neural Network Design", 2nd Edition, *Cengage Publication*.

[15] Machine Learning Notes: Stanford University: http://cs229.stanford.edu/materials.html