**IOSR Journal of Engineering (IOSRJEN)** **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 09, Issue 5 (May. 2019), ||S (II) || PP 68-73

# Detection of Glaucoma Using Machine Learning Algorithms

[1]Subhadra Kompella
*[1]Department of Computer Science And Engineering, GITAM Deemed To Be University, Visakhapatnam*
*Corresponding Author: subhadra Kompella*

**Abstract:** This paper aims at detecting the presence of glaucoma based on retinal nerve fibre layer thickness and visual field. To perform this data has been collected from various hospitals having numerous candidate features. These features are reduced by preprocessing and dimensionality reduction in order to remove noise.Glaucoma is characterized by dysfunction and loss of retinal ganglion cells (RGCs), with resulting structural changes to the optic nerve head, retinal nerve fiber layer (RNFL) thickness, and ganglion cell and inner plexiform layers as well as loss of the visual field.The diagnosis of glaucoma in its early stages is challenging.Diagnosis of glaucoma in myopic eyes and patients with brain diseases such as brain tumor is known to be difficult due to those eye's characteristic disc shape and visual field defect. A more effective glaucoma – detection machine learning model would be very helpful to medical practitioners. Several machine learning algorithms such as decision trees, k- nearest neighbourhood, random forest algorithm have been implemented to compare the efficiency. This paper aims at identifying most effective machine learning algorithm in detecting the presence of glaucoma.

**Keywords:** glaucoma, machine learning, dimensionality reduction, random forest, decision trees.

---

Date of Submission: 15-05-2019 Date of acceptance: 31-05-2019

---

## I. INTRODUCTION

Glaucoma is characterized by dysfunction and loss of retinal ganglion cells (RGCs), with resulting structural changes to the optic nerve head, retinal nerve fiber layer (RNFL) thickness, and ganglion cell and inner plexiform layers as well as loss of the visual field.The diagnosis of glaucoma in its early stages is challenging. Misdiagnosis can lead to failure to identify individuals with the condition during its early stages until significant functional loss has occurred. Thus, early detection of glaucoma allows for early treatment to delay vision loss. Diagnosing glaucoma is problematic, especially when it is in the earliest stage of glaucoma. Diagnosis of glaucoma in myopic eyes and patients with brain diseases such as brain tumor is known to be difficult due to those eye's characteristic disc shape and visual field defect. A more effective glaucoma – detection machine learning model would be very helpful to medical practitioners.The classification scheme in machine learning is suitable for diagnosis glaucoma. Various classification algorithms were tested. The classification algorithms used were theK – Nearest Neighbor, Decision Tree, Random Forest Classifier.

Machine learning (ML) is a subset of artificial intelligence in the field of computer science that often uses statistical techniques to give computers the ability to learn with data, without being explicitly programmed.Data is being produced and stored continuously ("big data"), which is used in various fields:science: genomics, astronomy, materials science, particle accelerators, sensornetworks: weather measurements, people: social networks, blogs, mobile phones, purchases, bank transactions.Data is not random, it contains structure that can be used to predict outcomes, or gain knowledge in some way.ML is programming computers using data (past experience) to optimize a performance criterion.ML relies on:Statistics: making inferences from sample data and also Numerical algorithms (linear algebra, optimization): optimize criteria, manipulate models.Machine Learning is an idea to learn from examples and experience, without being explicitly programmed. Instead of writing code, data is fed to the generic algorithm, and it builds logic based on the data given.For example, one kind of algorithm is a classification algorithm. It can classify data into different groups. The classification algorithm used to detect handwritten alphabets could also be used to classify emails into spam and not-spam.

**k-NEAREST NEIGHBOR (k-NN):**The k-Nearest-Neighbors (k-NN) method of classification is one of the simplest methods in machine learning. At its most basic level, it is essentially classification by finding the most similar data points in the training data, and making an educated guess based on theirclassifications. This method has seen wide application in many domains, such as in recommendation systems, semantic searching, and anomaly detection.k-NN based classification is a type of lazy learning as it does not attempt to construct a general internal model, but simply stores instances of the training data. Classification is computed from a simple majority votes of the k-nearest neighbors of each point.

**DECISION TREE:** Decision tree classifier is a simple yet widely used classification technique. The decision tree is a hierarchical structure consisting of nodes and directed edges. The tree has three types of nodes:
- A root node that has no incoming edges and zero or more outgoing edges.
- Internal nodes, each of which has exactly one incoming edge and two or more outgoing edges.
- Leaf or terminal nodes, each of which has exactly one incoming edge and no outgoing edges.

In a decision tree, each leaf node is assigned a class label. The non- terminal nodes, which include the root and other internal nodes, contain attribute test conditions to separate records that have different characteristics.

**RANDOM FOREST CLASSIFIER:** Random Forests are an example of an ensemble method, in which we combine multiple machine learning algorithms to obtain better predictive performance. We'll run multiple models on the data and use the aggregate predictions, which will be better than a single model alone. Random Forest is one of the most common ensemble methods, which consists of a collection of Decision Trees.The idea behind a Random Forest is to repeatedly select data from the data set (with replacement) and build a Decision Tree with each new sample. It is important to note that since we are sampling with replacement, many data points will be repeated and many won't be included as well. This is important to keep in mind when we talk about measuring error of a Random Forest. Another important feature of the Random Forest is that each node of the Decision Tree is limited to only considering splits on random subsets of the features.In the case of classification with Random Forests, we use each tree in our forest to get a prediction, then the label with the most votes becomes the predicted class for that data point. The usual parameters when building a forest (standard defaults used in the SciKit-Learn library) are 10 trees and only considering the square root of 'n' features, where n is the total number of features.
This paper implements the above mentioned algorithms and compares the efficiency of the three to predict and detect the presence of glaucoma.

## II. RELATED WORK

Chan et al [1] tested various classification algorithms based on the examination of visual fields. The machine learning algorithms studied included multilayer perceptron (MLP), support vector machine (SVM), and linear (LDA) and quadratic discriminant analysis (QDA), Parzen window, mixture of Gaussian (MOG), and mixture of generalized Gaussian (MGG). MLP and SVM are classifiers that work directly on the decision boundary and fall under the discriminative paradigm. Generative classifiers, which first model the data probability density and then perform classification via Bayes' rule, usually give deeper insight into the structure of the data space. They have applied MOG, MGG, LDA, QDA, and Parzen window to the classification of glaucoma from SAP. Performance of the various classifiers was compared by the areas under their receiver operating characteristic curves and by sensitivities (true-positive rates) at chosen specificities (true-negative rates). The machine-learning-type classifiers showed improved performance over the best indexes from STATPAC. Forward-selection and backward-elimination methodology further improved the classification rate and also has the potential to reduce testing time by diminishing the number of visual-field location measurements.

Goldbaum et al [2] also compared machine learning classifiers and suggested a mixture of of Gaussian as the best classifier. Multilayer perceptrons (MLP), support vector machines (SVM), mixture of Gaussian (MoG), and mixture of generalized Gaussian (MGG)classifiers were trained and tested by cross validation on the numerical plot of absolute sensitivity plus age of 189 normal eyes and 156 glaucomatous eyes, designated as such by the appearance of the optic nerve. The authors compared performance of these classifiers with the global indices of STATPAC, using the area under the ROC curve. Two human experts were judged against the machine classifiers and the global indices by plotting their sensitivity-specificity pairs.

Bizios et al [3] tested the artificial neural network (ANN) and support vector machine (SVM) based on RNFL thickness parameters. Analyzed Stratus OCT data from 90 healthy persons and 62 glaucoma patients. Performance of MLCs was compared using conventional OCT RNFLT parameters plus novel parameters such as minimum RNFLT values, 10th and 90th percentiles of measured RNFLT, and transformations of A-scan measurements.

Barella et al [4] investigated the diagnostic accuracy of machine learning classifiers (MLCs) and random forest (RF) using RNFL and optic nerve data. They got 0.877 of area under the ROC value using RF. Purpose: To investigate the diagnostic accuracy of machine learning classifiers (MLCs) using retinal nerve fiber layer (RNFL) and optic nerve (ON) parameters obtained with spectral domain optical coherence tomography (SD-OCT). Fifty-seven patients with early to moderate primary open angle glaucoma and 46 healthy patients were recruited. All 103 patients underwent a complete ophthalmological examination, achromatic standard automated perimetry, and imaging with SD-OCT. Receiver operating characteristic (ROC) curves were built for RNFL and ON parameters. Ten MLCs were tested. Areas under ROC curves (a ROCs) obtained for each SD-OCT parameter and MLC were compared. MLCs showed good accuracy but did not improve the sensitivity and specificity of SD-OCT for the diagnosis of glaucoma.

Recently, Silva et al [5] tested almost all of the classifiers using Spectral Domain optical coherence tomography (OCT) and standard automated perimetry. They got 0.946 as the best a ROC value using RF. Previous studies show that SVM and RF have good prediction power. Focal RGC loss was believed to account for clinical patterns of visual loss caused by glaucoma. Sixty two glaucoma patients and 48 healthy individuals were included. All patients underwent a complete ophthalmologic examination, achromatic standard automated perimetry (SAP) and retinal nerve fiber layer (RNFL) imaging with SD-OCT.

[6] investigated Six glaucomatous retinas (67-83 years old) and six age-matched control retinas (61-89 years old) were prepared as whole amounts and stained by 4′,6-diamidino-2-phenylindole (DAPI) solution (3µg/mL in PBS). An area corresponds to central 14 degree of the visual field was imaged. The nearest neighbor distribution was determined for cells in both normal and glaucomatous RGCL.

Ohkubo [7] considered Sixty eyes of 60 subjects with glaucoma were included. Sensitivity of each test point of 10-2 standard automated perimetry was compared with the thickness of the retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), GCL+ inner plexiform layer (IPL), and RNFL+GCL+IPL (GCC), with and without RGC displacement, using Spearman's rank correlation coefficients. Visual sensitivity was evaluated by unlogged 1/Lambert (1/L) values. The GCC is the most useful parameter to evaluate structure and function within the central 10° in glaucoma. Adjusting for RGC displacement is essential to evaluate the relationship between structure of the GCL-related layer and function at the central macula.

## III. METHODOLOGY

Once you know exactly what you want and the equipment's are in hand, it takes you to the first real step of machine learning, is Gathering Data. This step is very crucial as the quality and quantity of data gathered will directly determine how good the predictive model will turn out to be. The data collected is then tabulated and called as Training Data.The raw data of several patients has been collected from the 'Visakha Eye Hospital'. The data collected is for the duration of two months. The raw data consisting of around 1000 records in the form of images of different patients having different eye diseases has been collected. Under the guidance of doctor, from these 1000 records, 212 records of patients anticipated with glaucoma has been separated manually. These 212 records has been used for the analysis.

The attributes represented in the dataset are:
Attribute 1. Age
Attribute 2. Intra ocular pressure of right eye (IOP OD)
Attribute 3. Intra ocular pressure of left eye (IOP OS)
Attribute 4. Cornea thickness of right eye(CT OD)
Attribute 5. Cornea thickness of left eye(CT OS)
Attribute 6. Retinal nerve fiber layer thickness of right eye (R OD)
Attribute 7. Retinal nerve fiber layer thickness of left eye (R OS)
Attribute 8. Difference of retinal nerve fiber layer thickness of two eyes (R OD-OS)
Attribute 9. Cup disc vertical ratio of right eye (C.V.OD)
Attribute 10. Cup disc vertical ratio of left eye (C.V.OS)
Attribute 11.Difference of Cup disc vertical ratio of two eyes (C.V.OD-OS)
Attribute 12. Cup disc horizontal ratio of right eye (C.H.OD)
Attribute 13. Cup disc horizontal ratio of left eye (C.H.OS)
Attribute 14. Difference of Cup disc horizontal ratio of two eyes (C.H.OD-OS)
Attribute 15. Cup disc area ratio of right eye (C.A.OD)
Attribute 16. Cup disc area ratio of left eye (C.A.OS)
Attribute 17. Difference of Cup disc area ratio of two eyes (C.A.OD-OS)
Attribute 18. Ganglion Attribute on cell complex of right eye (G.OD)
Attribute 19. Ganglion cell complex of left eye (G.OS)
Attribute 20. Difference of ganglion cell complex of two eyes (G.OD-OS)
Attribute 21. Focus loss volume of right eye (FLV OD)
Attribute 22. Focus loss volume of left eye (FLV OS)
Attribute 23. Global loss volume of right eye (GLV OD)
Attribute 24. Global loss volume of left eye (GLV OS)
Attribute 25. Rim area of right eye
Attribute 26. Rim area of left eye

Attribute 27. Disc area volume of right eye
Attribute 28. Disc area volume of left eye
Attribute 29. Cup volume of right eye
Attribute 30. Cup volume of left eye

The ‘Label’ Glaucoma is denoted as ‘G’ in the dataset. The patient diagnosed with glaucoma is denoted by ‘1’ and the patient not having glaucoma is denoted by ‘0’. After preprocessing the data, required features are selected from the data set for further processing of detection of glaucoma.

Feature selection is the process of selecting a subset of relevant features (variables, predictors) for use in machine learning model building.A feature selection algorithm can be seen as the combination of a search technique for proposing new feature subsets, along with an evaluation measure which scores the different feature subsets. Different feature subsets render optimal performance for different machine learning algorithms.Finally using feature selection, 30 features are reduced to 10 features. The features reduced are visualized using pair plots and box plots.

Among many classifiers, the machine learning classifiers that are used for the analysis are k- NN, Decision tree and ensemble methods such as Random Forest algorithm. The efficiency is compared using several metrics as discussed below:

The accuracy, sensitivity, specificity have been widely used as criteria for evaluating a diagnosis model. The following terms are fundamental to understanding the utility of them. True Positive(TP): The patient has a disease and prediction is positive.
False Positive(FP): The patient does not have a disease but the prediction is positive.
True Negative(TN): The patient does not have a disease and the prediction is negative.
False Negative(FN): The patient has a disease but the prediction is negative.
**Accuracy** is the ratio of number of correct predictions to the total number of input samples.
**Sensitivity** corresponds to the proportion of positive data points that are correctly considered as positive, with respect to all positive data points.
**Specificity** corresponds to the proportion of negative data points that are mistakenly considered as positive, with respect to all negative data points.
**Precision** is the number of correct positive results divided by the number of positive results predicted by the classifier.
The **recall** of a diagnosis model refers to the ability of the test to correctly identify those patients without the disease.

In addition to the above metrics the results in this paper are recorded using ROC, A receiver operating characteristic curve, is a graphical plot that illustrates the diagnostic ability of a binary classifiers system as its discrimination threshold is varied. The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The TPR defines how many correct positive results occur among all positive samples available during the test. FPR, on the other hand, defines how many incorrect positive results occur among all negative samples available during the test.An ROC space is defined by FPR and TPR as x and y axes, respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs). The best possible prediction method would yield a point in the upper left corner or coordinate (0,1) of the ROC space, representing 100% sensitivity (no false negatives) and 100% specificity (no false positives). The (0,1) point is also called a perfect classification. A random guess would give a point along a diagonal line (the so-called line of no-discrimination) from the left bottom to the top right corners (regardless of the positive and negative base rates). The diagonal divides the ROC space. Points above the diagonal represent good classification results (better than random); points below the line represent bad results (worse than random).

**Confusion matrix:** A confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm. Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class (or vice versa). It is a special kind of contingency table, with two dimensions ("actual" and "predicted"), and identical sets of "classes" in both dimensions (each combination of dimension and class is a variable in the contingency table). It is a table with two rows andtwo columns that reports the number of false positives, false negatives, true positives, and true negatives.

The results obtained are tabulated and compared ass shown below:

| Classifiers | Precision | Recall | F1-Score | Sensitivity | Specificity | Roc |
|---|---|---|---|---|---|---|
| Decision Tree | 0.98 | 0.98 | 0.98 | 1 | 0.97 | 0.99 |
| k- NN | 0.98 | 0.98 | 0.98 | 0.94 | 1 | 0.97 |
| Random Forest | 0.97 | 0.97 | 0.97 | 0.88 | 1 | 0.94 |

**Table1:** Comparison of different Classifiers

| | Predicted:No | Predicted:Yes |
|---|---|---|
| Actual: No | TN=16 | FP=0 |
| Actual: Yes | FN=1 | TP=43 |

**Table2:** k- NN Confusion Matrix

| | Predicted:No | Predicted:Yes |
|---|---|---|
| Actual: No | TN=16 | FP=0 |
| Actual: Yes | FN=1 | TP=43 |

**Table3:**Decision Tree Confusion Matrix

| | Predicted:No | Predicted:Yes |
|---|---|---|
| Actual: No | TN=16 | FP=0 |
| Actual: Yes | FN=2 | TP=42 |

Table 4:Random Forest Confusion Matrix

Table 1. shows various classification evaluation metrics including 'Accuracy', 'Precision', 'Recall', 'F1- Score', 'Sensitivity', 'Specificity' and 'ROC Curve' of different classifiers. Some of them are outperformed in some aspects but all the classifiers disagree on different metrics.Random Forest has also shown equally good performance. Among all the classifiers Decision Tree has shown better performance in most of the metrics.Table 2. shows the confusion matrix of k- NN, it predicted all the non- glaucoma cases correctly. Among the glaucoma cases, one case was incorrect i.e though the patient is suffering from glaucoma it classified it as the patient is not suffering from glaucoma.Table 3.shows the confusion matrix of decision tree, it predicted all the glaucoma cases correctly.Among the non- glaucoma cases, one case was incorrect i.e. the person is not having glaucoma but classified as glaucoma case.Table 4.shows the confusion matrix of Random forest, two glaucoma cases were predicted incorrectly.In medical field, False negative ratio is more important than False positive ratio. Decision tree when compared to other classifiers has performed well as it only misclassified only one non-glaucoma case.Hence Decision tree is used as final Prediction model.

## IV. CONCLUSION & FUTURE SCOPE

Glaucoma is a serious disease that can cause complete, permanent blindness, and its early diagnosis is very difficult. In recent years, computer- aided screening and diagnosis of glaucoma had made progress. The prediction model developed using Decision Tree can be used for the diagnosis of glaucoma. It performed better when compared to other classifiers such as k- NN, Random Forest which is evaluated using various performance metrics. The model misclassified one non-glaucoma case as glaucoma, this is less serious than False Negative. The error would improve if number of cases are increased.The other classifiers also performed well and the performance variation would be seen if more number of cases are considered. The models would perform more better if other features like Glaucoma Hemifield test, Pattern Deviation and Mean Deviation are also considered.k-NN showed less performance when compared to other classifiers in terms of 'Accuracy'. This paper also considers prediction results from the three learning models, take the majority of results as a final decision. The same methods should apply for large glaucoma datasets will give better results.

## REFERENCES


[1]. Kim SJ, Cho KJ, Oh S (2017) Development of machine learning models for diagnosis of glaucoma.PLoS ONE 12(5): e0177726. https://doi.org/10.1371/journal.pone.0177726.

[2]. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. Jama 2014; 311(18):1901±11. https://doi.org/10.1001/jama.2014.3192 PMID: 24825645.

[3]. Tay E, Seah SK, Chan SP, Lim ATH, Chew SJ, Foster PJ. Optic disk ovality as an index of tilt and its relationship to myopia and perimetry. Am J Ophthalmol 2005; 139(2):247±52. https://doi.org/10.1016/j.ajo.2004.08.076 PMID: 15733984,

[4]. Barella KA, Costa VP, GoncËalvesVidotti V, Silva FR, Dias M, Gomi ES. Glaucoma diagnostic accuracy of machine learning classifiers using retinal nerve fiber layer and optic nerve data from SD-OCT. J Ophthalmol 2013.
[5]. Bizios D, Heijl A, Hougaard JL, Bengtsson B. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by Stratus OCT. ActaOphthalmol 2010.
[6]. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Continuing Education in Anaesthesia, Critical Care & Pain 2008.
[7]. Kang SH, Hong SW, Im SK, Lee SH, Ahn MD. Effect of myopia on the thickness of the retinal nerve fiber layer measured by Cirrus HD optical coherence tomography. Invest Ophthalmol Vis Sci 2010.
[8]. Seol BR, Jeoung JW, Park KH. Glaucoma Detection Ability of Macular Ganglion Cell-Inner Plexiform Layer Thickness in Myopic PreperimetricGlaucomaGlaucoma Detection Ability inMyopic Glaucoma. Invest Ophthalmol Vis Sci 2015.

IOSR Journal of Engineering (IOSRJEN) is UGC approved Journal with Sl. No. 3240, Journal no. 48995.

subhadra Kompella. “Detection of Glaucoma Using Machine Learning Algorithms.” IOSR Journal of Engineering (IOSRJEN), vol. 09, no. 05, 2019, pp. 68-73.