Systematic Review of Machine Learning Methods for the Eye Illness Detection

Subhasis Nayak
Department of CSE
Silicon Institute of Technology (of Aff.)
Bhubaneswar, India
cse.20bcsd29@silicon.ac.in

Sagar Verma
Department of CSE
Silicon Institute of Technology (of Aff.)
Bhubaneswar, India
cse.20bcse64@silicon.ac.in

***Abstract*—The essential objective of this paper is to survey various approaches for identifying eye problems, especially glaucoma, by utilizing deep learning procedures to get an early diagnosis based on pictures. The paper talks about the people over the age 40 during the training. Recent articles have employed quantitative analysis to assess particular criteria such as accuracy, affectability, and more specificity at the groud level of patient. Our paper mainly highlights the foremost viable handle and different methodology for the future angle of investigation procedures by utilizing different picture processing tools such like U-net,Kaggle and others, all of which are mainly associated with deep learning through a diffrent classification application and have a organized database. These technique mainly encourage way better learning and prediction of patient' s illness, coming about in broader scope and it reduced determination time, eventually making patients life better.**

***Index Terms*—Deep learning, Glaucoma, image analysis,CNN,KNN.**

I. INTRODUCTION

Retinal imaging is a useful diagnostic tool and, when properly evaluated, can provide reliable diagnosis of many eye diseases such as macular degeneration and herpes zoster,But the challenge is to achieve a good accuracy and reduce the most important that is processing time. This can be handled using methods like segmentation and deep learning, as well as convolutional neural networks.

In addition to segmentation and deep learning diffrent techniques, other methods that can help in image processing methods like feature extraction and classification algorithms can also be applied to enhance the accuracy of pre-diagnoses. texture analysis is one of the methods ,which involves the quantification of the spatial distribution of gray levels in an image to identify subtle patterns that may be indicative of disease. The development of portable and non-invasive devices for retinal imaging and analysis has the potential to improve the accessibility of these diagnostic tools, especially in remote or under-served areas. This could potentially help address the issue of undetected eye diseases and significantly improve overall eye health outcomes. Collaborative efforts among ophthalmologists, engineers, and data scientists can result in the creation of innovative and efficient diagnostic tools for various eye diseases. By working together, these experts can develop cutting-edge solutions that enhance patient outcomes and the overall quality of life. This collaborative approach can also address the issue of undetected eye diseases and improve accessibility to diagnostic tools, particularly in remote or underserved areas. Ultimately, such efforts have the potential to make a significant impact on the field of ophthalmology, leading to improved diagnosis, treatment, and care for individuals with various eye conditions.

Early detection and diagnosis of glaucoma are important in preventing vision loss, as the damage caused by the disease is irreversible. In addition to that, glaucoma often has no symptoms in its early stages, which makes it very difficult for patients to know that they are affected. The primary risk factors for glaucoma include age, genetics that boils down to family history, high intraocular pressure, and thin corneas. It is important to note that even individuals that has normal intraocular pressure can develop glaucoma.

Since these diseases are often asymptomatic in the early stages, they often go unnoticed. Experienced ophthalmologists may be able to identify some symptoms, but about 90% of these diseases remain undetected in adults above the age of 45. Glaucoma, specifically, is the main cause of irreversible loss of vision and is expected to affect approx 111.8 million people by 2040, according to some conducted studies. Therefore, here it is very important to compare all the different methods that are present for early detection of glaucoma so that we can determine the most efficient approach and its feasibility to detect Glaucoma.

II. SYSTEMATIC REVIEW

A systematic review was conducted using the PICOC methodology to select and analyze sources related to the use of deep learning in the detection of glaucoma. The review was restricted to research articles published in Q1 journals between 2019 and 2022 and searched on databases such as Scopus, ScienceDirect, and IEEE. A total of around 400 articles were initially retrieved, and the PICOC methodology was used to filter out the most relevant information.

Deep learning has the ability to perform complex classification procedures which is done by detecting different patterns with or without any help of extensive databases, depending on the selected algorithm. This method can provide more accurate non-invasive diagnoses for early detection of eye diseases, leading to better quality of life and appropriate treatment. Deep learning can also overcome the limitation of many eye diseases which do not present symptoms in the early stages. However, the success of deep learning heavily depends on the quality of the database used. Several databases such as ”U-Net, KAGGLE, DRIVE, and STARE” have been used, but the images in these databases have varying sizes, which may affect the training of deep learning algorithms due to entropy analysis.

Segmentation techniques are used in image processing to identify certain patterns or features and enable deep learning to be trained effectively. Many methods and techniques have been proposed, but the most recent methods include graph-based methods that classify regions and represent them with weighted images. Although this method requires effort and an understanding of the structure, it is still widely used. Learning-based methods involve extracting features for each pixel and classifying them.

To obtain optimal results, this method requires a significant amount of data and can impose a heavy calculation and memory load. The methods are based on deep learning involves utilizing a deep convolutional network structure, such as U-Net or SegNet, which allows for the use of inputs of any size and extracts features through convolutional layers. This approach has been shown to be the most current and efficient method during 2021-2022, according to recent literature [6].

The majority of studies on the detection of eye diseases using deep learning have utilized retinal fundus images or fundus photography, as described in reference [5]. Another potential source is the optical coherence tomography image, which was discussed in reference [6]. To detect eye diseases using these images, segmentation methods can be applied, particularly for the optic disc, as outlined in reference [5]. Reference [7] presented a three-stage process for detecting diabetic retinopathy, which involved image processing to remove noise, segmenting the region using the ”U-Net” and ”Squeeze-and-excitation block” techniques, and classifying the images based on macular degeneration for advanced age.

III. METHODOLOGY

A. *Image Classification Techniques*

There are a lot of methods and techniques for image classification. These techniques are classified into Automatic, Manual and Hybrid techniques. Automatic classification methods are further divided into methods, supervised and unsupervised classification. Variations in the above classification methods are discussed in Section below.

1) Automatic Classification: Algorithms are used systematically in automatic image classification methods to group pixels into meaningful categories across the entire set of images. This category of image classification methods is more common. There are two main categories of automatic image classification methods: supervised and unsupervised.

a. Supervised Classification method

In supervised classification methods, an analyst is required to provide input, which is known as the training set. The accuracy of the classification relies heavily on the quality of the training samples. These samples are classified into two categories, one for classification and the other for evaluating the accuracy of the supervised classification. The training set is provided prior to executing the classification process.

b. Unsupervised Classification method

Unsupervised classification is a technique where the software analyses an image and groups related pixels into classes without the user providing sample classes. The classification outcome is based on the computer’s determination of which pixels are related to one another. Although the technique doesn’t require the user to provide any prior knowledge of the image, it’s essential for the user to have an understanding of the area being classified. This knowledge comes in handy when the groupings of pixels produced by the computer need to be related to actual features on the ground. Therefore, while unsupervised classification can be a powerful tool for image analysis, it still requires human expertise and involvement to accurately interpret and relate the pixel groupings to real-world features.

2) Manual Classification: Although manual image classification methods are robust and effective, they tend to be time-consuming. These methods require the analyst to have a good understanding of the area covered by the image. The classification accuracy and efficiency are closely linked to the analyst’s knowledge and familiarity with the field of study.

3) Hybrid Classification: Hybrid image classification methods leads to combining the advantages of automated and manual methods. This hybrid approach uses automated image classification methods to do initial classification and further usees manual methods to refine classification and make the errors correct.

B. *Convolution Neural Networks*

A Convolutional Neural Network (CNN) is an interconnected system of artificial neurons that contains numerous learnable weights and biases. In comparison to other image classification algorithms, CNNs require relatively minimal pre-processing. During the training process, the connections between neurons are assigned numerical weights that are adjusted to ensure that a properly trained network can accurately identify and classify images or patterns.

A Convolutional Neural Network is composed of interconnected artificial neurons that can adjust their weights and biases during the training process. These neurons communicate with one another through weighted connections to recognize images or patterns. The network contains multiple layers of neurons that detect different features. The first layer detects basic patterns, while the subsequent layers detect more complex combinations of features. The layers are constructed in a hierarchical manner, where each layer detects patterns of patterns from the previous layer. This approach enables the network to recognize increasingly complex visual representations in the input data.

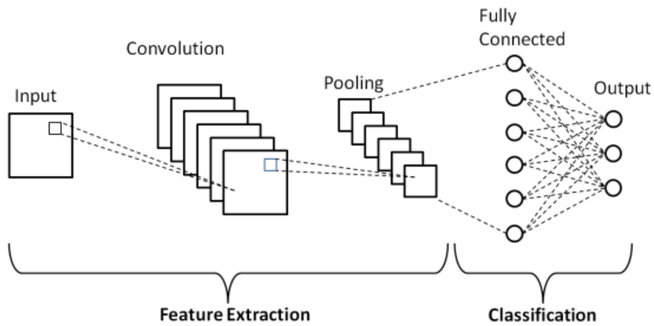


Fig. 1. FIGURE: basic CNN

1) Layers of CNN: For classification problems, a complex architectures are built by stacking multiple and different kind of layers in a CNN. The four diffrent types of layers are convolution layer, pooling/subsampling layers, non-linear (ReLU) layers, and fully connected layers. Figure 3.2 shows all the various layers of CNN. A portion of the input image is that are fed to the convolution layer. The output of this layer is then fed to the next layer that is pooling layer. This is repeated again and followed by a fully connected layer which performs classification.

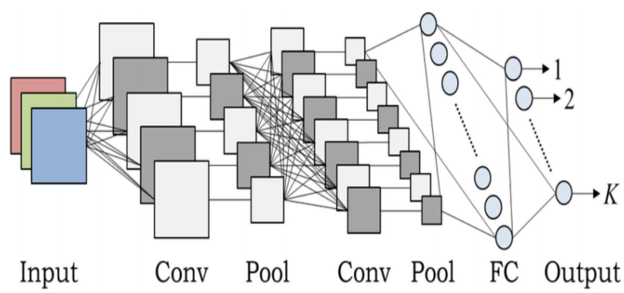


Fig. 2. FIGURE: Architecture of CNN

2) Convolutional Layer: By utilizing the process of convolution, unique characteristics of the input data can be extracted and emphasized. Low-level features such as edges, corners, and lines are identified through the initial convolution layer. As we progress to higher-level layers, more complex features are extracted. A 3D convolution process is used in Convolutional Neural Networks (CNNs), where the input data of size N x N x D is convoluted with H kernels, each of size k x k x D, independently. Each input component is convoluted with one kernel, producing one output feature, and with H kernels independently, it produces H features. The input data is moved from left to right, one element at a time, starting from the top-left corner, until the top-right corner is reached. The process then moves downwards one element at a time, and once again, the data is moved from left to right until the bottom-right corner is reached. This process is repeated until the entire input data has been convoluted.

For instance, let's consider a scenario where N equals 5 and k equals 5. The kernel can be positioned at 5 different

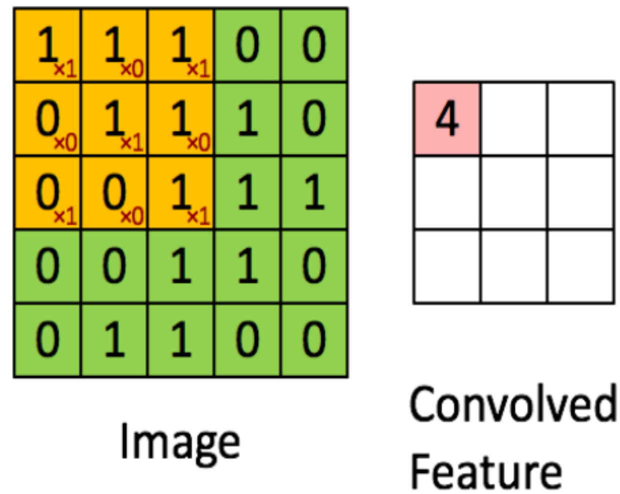


Fig. 3. FIGURE : Representation of Convolutional Process

locations from left to right and 5 different locations from top to bottom. Each output feature will have 28x28 elements, which is equivalent to (N-k+1) x (N-k+1). To generate one element of one output feature, a sliding window process is applied to each position of the kernel. This involves the element-by-element multiplication of k x k x D input and k x k x D kernel, which is then accumulated. As a result, creating one element of one output feature requires k x k x D multiply-accumulate operations.

3) Pooling Layer: To reduce the resolution of the features, CNNs use a pooling (sub-sampling) layer. These features are generally resilient to noise and distortion. Pooling can be done in two ways: max pooling and average pooling. In both cases, the input is divided into non-overlapping sub-regions.

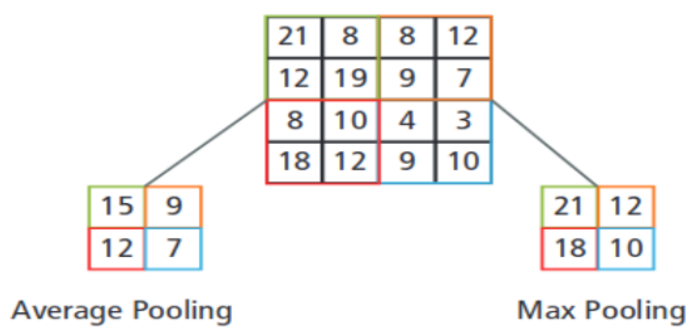


Fig. 4. FIGURE : Representation of Max Pooling and Average Pooling

4) Non Linear Layer: The ReLU layer implements the function: y=max(x,0).Now the input and output sizes of this layer are the nearly same. This operation increases the nonlinear properties of the decision function including the overall network. This don't affect the respective fields of the convolution layer. Compared to the other non-linear functions that exist used in CNNs (e.g., Sigmoid, hyperbolic tangent and another one is absolute of hyperbolic tangent), the most important advantage of a ReLU is that the network trains much faster compare to other ones. The functionality of ReLU is illustrated in Figure 5. The transfer function plotted above the arrow. All the positive values(15, 20, 35, 18, 25, 100, 20, 25, 101, 75, 18, 23) are retained as such and the other negative values (-10, -110, -15, -10) are proselytized to zero

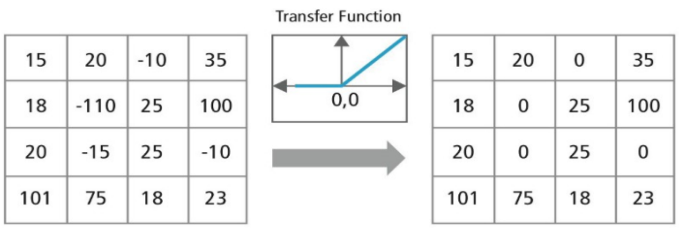


Fig. 5. FIGURE : Representation of ReLU Functionality

5) Fully Connected Layer: The final layers of a CNN are commonly referred to as fully connected layers. These layers sum up a weighted combination of the previous layer's features, resulting in a precise mixture of ingredients to produce a specific output target. In a fully connected layer, all elements of all previous layer's features are utilized in computing each output feature element. Figure 3.6 depicts layer L, which is fully connected. Layer L-1 has two features, each of which comprises four elements. To illustrate, the first feature in the (L-1) layer is obtained by multiplying each feature element by two sets of weights and summing the resulting products. Layer L includes two features, each having a single element. The input images for the CNN model were obtained from the publicly available RIGA dataset and had a size of 256 x 256. Prior to training the CNN model, data augmentation was performed on the images. Ultimately, the trained model was utilized to predict the classes of the test data.

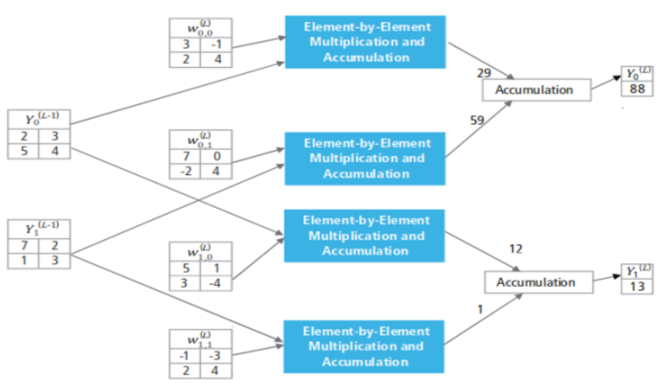


Fig. 6. FIGURE : Processing of Fully Connected Layer

6) L2 Regularization: Overfitting is one of the main common problem of neural network approaches. Over-fitting is basically the classification results that can be very good on the training data set but poor on the test data set. To avoid the problem like overfitting, it is necessary to adopt additional techniques such as regularization. Regularizers allow to apply penalties on these layer parameters or layer activity during optimization. These are penalties that are incorporated in the loss function that the network optimizes.A widely used regularization technique is L2 regularization, which can be applied by incorporating the squared magnitude of all the parameters directly into the objective. Specifically, a penalty term of 1/2 * lambda * w*w is added to the objective for each weight w in the network, where lambda is the regularization strength. The factor of 1/2 is often included to simplify the gradient of the term with respect to w. L2 regularization is known to penalize high-magnitude weights while encouraging more distributed weight vectors.

C. Architecture for Glaucoma Detection

The system architecture shown in Figure 7 gives overall view about all the modules in the proposed system of Glaucoma and the flow of the process right from data collection to detection.

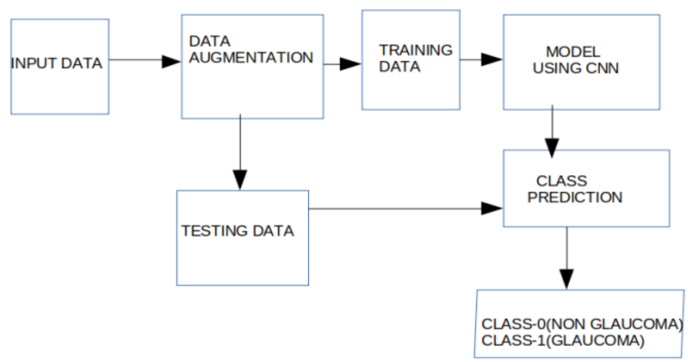


Fig. 7. FIGURE : System Architecture for Glaucoma

IV. IMPLEMENTATION

A. Image Classification Process

Input: The input set consists of all the images, each images are labelled with different classes.This is what called training set.

Training: During the training phase, a classifier or model is learned using the training dataset to identify the important details of each class and its location within the image.

Testing: In the testing phase, the classifier's accuracy is evaluated by predicting labels for a new dataset of images that it has not encountered before. The accuracy of the classifier is determined by the number of correctly predicted labels, indicating a higher level of accuracy when more values match.

V. RESULTS

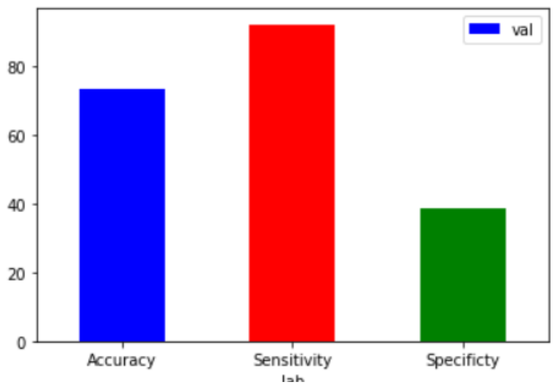
- **Accuracy** Accuracy refers to how closely a prediction or measurement matches the true or expected value. In other words, accuracy is a measure of how well a model or system is able to produce correct results.
  Accuracy = (TP+TN)/(TP+FP+FN+TN)
- **Specificity**
  Specificity or the true negative rate is the measure of the proportion of True Negatives Vs Sum of Predicted False

Positives and Predicted True Negatives.
Specificity = TN/(TN+FP)

- **Sensitivity** Sensitivity in Machine Learning can be described as the metric used for evaluating a model's ability to predict the true positives of each available category. In literature, this term can be also recognized as a true positive rate and it can be calculated with the following equation:
Sensitivity = TP / TP+FN
(True Positive/True Positive + False Negative)

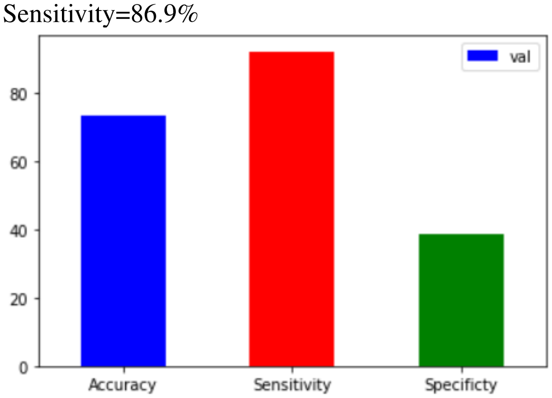
A. KNN

Accuracy=73.1%
Specificity=38.5%
Sensitivity=92%



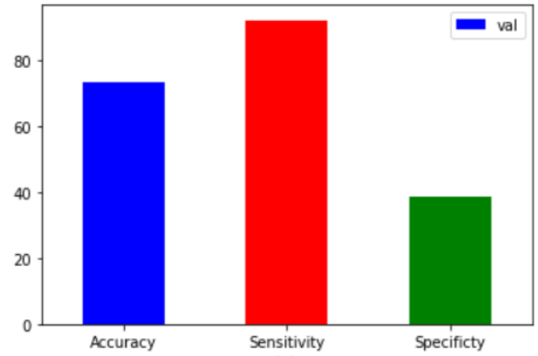
B. SVM

Accuracy==75%
Specificity==53.1%
Sensitivity=86.9%



C. RANDOM FOREST

Accuracy==77.6%
Specificity=32.5%
Sensitivity=92%



D. CNN

VI. GLAUCOMA PROBLEMS AND CHALLENGES

Glaucoma is a eye diseases that cause damage to the optic nerves, which is very much responsible for transmitting visual information from the eye to the main source that is brain. If it is left untreated, Glaucoma can lead to a permanent vision loss can cause blindness. It is estimated that Glaucoma affects nearly over 70 million people worldwide, making it one of the main reason for the causes of blindness.

The biggest challenges in the diagnosis and management of Glaucoma is the variability in the disease's progression and the difficulty in detecting it in its early stages. The damage to the retina caused by Glaucoma makes it a relevant cause of blindness [3]. In order to approach this issue in a methodical manner, it is important to take into account several dependent variables. These include evaluating the fundus, assessing veins and blood vessels, measuring entropy, comparing different methods, considering the size of training images, analyzing image processing time, and assessing algorithm accuracy. By systematically considering these variables, it is possible to develop a more accurate and effective system for classifying disease based on differences in eye shape, irrigation, border, and other factors. However, care must be taken to avoid similarities in phrasing or word choice that might be flagged as plagiarism.

Improving the detection and accuracy of Glaucoma through Deep Learning is the expected outcome of the independent variable.

VII. CONCLUSION

In conclusion, our machine learning project aimed to develop a model that could accurately detect the presence of glaucoma in patients. Through the use of convolutional neural networks and the ResNet50 architecture, we were able to train a model that achieved high accuracy in detecting glaucoma in retinal images.

Our results indicate that machine learning has the potential to be a valuable tool in the early detection of glaucoma, allowing for earlier intervention and improved outcomes for patients. While our model achieved high accuracy, there is still room for improvement and further research in this area.

Overall, we believe that our project demonstrates the power of machine learning in healthcare and highlights the potential for continued innovation and development in this field. With further refinement and optimization, machine learning could become an important tool in the fight against glaucoma and other eye diseases.

REFERENCES

[1] Glaucoma Research Foundation, ”Types of Glaucoma,” Glaucoma Research Foundation, 18 August 2016. [Online]. Available: http://www.glaucoma.org/glaucoma/types-of-glaucoma.php. [Accessed 11 February 2017]

[2] Glaucoma Research Foundation, ”Glaucoma Facts and Stats,” Glaucoma Research Foundation, 18 November 2016. [Online]. Available: http://www.glaucoma.org/glaucoma/glaucoma-facts-and-stats.php. [Accessed 10 February 2017]

[3] https://www.kaggle.com/datasets/keras/resnet50

[4] ‘Glaucoma Dataset RIM-ONE’, https://medimrg.webs.ull.es/research/retinalimaging/rim-one, January 27, 2019.