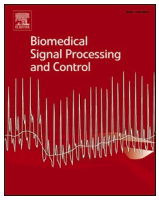
Biomedical Signal Processing and Control 89 (2024) 105770

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



An efficient framework for optic disk segmentation and classification of Glaucoma on fundus images

Jignyasa Sanghavi *, Manish Kurhekar

Visvesvaraya National Institute of Technology, Nagpur, M.S., India

ARTICLE INFO

Keywords:
Glaucoma
CNN
Optic Disk segmentation
Transfer learning
SLIC
Normalized graph cut

ABSTRACT

Accurately segmenting the Optic Disk is a crucial step in classifying Glaucoma using Fundus images. Machine learning and artificial intelligence techniques are widely used in Glaucoma detection, and the main indicators observed in Fundus images are the presence of Papillary Atrophy, Cup to Disc Ratio values, diminishing Neural Retinal Rim (NRR), the Inferior Superior Nasal Temporal (ISNT) rule, and Cup Diameter. In this research, we investigated various segmentation and classification techniques that can be applied to Optic disk segmentation and classification of normal and glaucomatous eyes. The proposed method will be beneficial to clinicians and healthcare workers in facilities with limited resources. In this paper, histogram processing is used to determine the type of image, and based on this information; we decide whether the image requires segmentation. Some datasets in the standard dataset contain complete retinal images while others include segmented optic disks. The segmented images are directly given as input for classification using the proposed Convolutional Neural Network (CNN). For complete retinal images, segmentation is performed using the Simple Linear Iterative clustering (SLIC) and normalized graph cut algorithm. The proposed framework's performance is compared with that of pretrained neural networks, including VGG19, InceptionV3, and ResNet50V2, using major metrics. We trained and tested these architectures with 3115 images from six standard datasets. Our proposed framework outperforms all with an accuracy of 96.33 %.

1. Introduction

Glaucoma is a prominent reason of blindness globally. As per the World Health Organization's estimation; about 4.5 million individuals suffer from blindness due to this ailment worldwide. In India, glaucoma stands as the primary cause of irreversible blindness, affecting 12 million people [1].

Glaucoma arises due to the gradual deterioration of the optic nerve fibers, leading to structural changes in the optic nerve head and a gradual reduction in the neuro-retinal rim. The loss of ganglion cells increases intraocular pressure in the eyes. There are four main types of glaucoma - Open Angle Glaucoma, Angle-closure Glaucoma, Normal Tension Glaucoma, and Congenital Glaucoma. Artificial intelligence-based research is predominantly focused on open-angle glaucoma. In its early stages, Glaucoma shows no symptoms, but as it progresses, it can cause irreversible damage to vision. While vision loss cannot be cured, early detection can help slow down its progression with suitable treatment. A variety of imaging techniques are now available to obtain an accurate diagnosis of Glaucoma, including Fundus imaging, Optical Coherence Tomography (OCT) angiography, Doppler OCT, Laser Speckle Flowgraphy, Polarization-sensitive OCT, Adaptive optics OCT, adaptive optics imaging, ultrasound imaging, spectral and Hyperspectral fundus imaging, and Heidelberg Retinal Tomography [2]. One major drawback of these techniques is their dependence on experts.

Fundus imaging relies on various parameters, such as the presence of Parapapillary Atrophy, Cup to Disc Ratio values, diminishing Neural Retinal Rim (NRR), Inferior Superior Nasal Temporal (ISNT) rule, and Cup Diameter, to diagnose Glaucoma [2]. Although fundus imaging is an affordable solution, image acquisition can be performed using a smartphone attached to a power lens, and it can be particularly useful in developing countries, where healthcare facilities and healthcare practitioners may be limited in remote areas. Artificial intelligence, when integrated with these imaging techniques, can be a promising solution for screening Glaucoma. Initial screening can alert the patient to consult an ophthalmologist at an early stage, potentially leading to better treatment outcomes.

One research has shown automated glaucoma diagnosis using empirical wavelet-based correntropy features from fundus images. By

* Corresponding author.
E-mail address: sjignyasa8588@gmail.com (J. Sanghavi).

https://doi.org/10.1016/j.bspc.2023.105770
Received 11 April 2023; Received in revised form 3 November 2023; Accepted 17 November 2023
Available online 21 November 2023

extracting correntropy features with high t-values for classification, the authors observed that employing RBF and Morlet wavelet kernels provided the highest accuracies. Green channel images proved most effective in comparison to other channels for accuracy [3].

In one research, variational mode decomposition (VMD) is employed for image decomposition, and features such as Kapoor entropy, Renyi entropy, Yager entropy, and fractal dimensions are extracted and fed to the least squares support vector machine (LS-SVM) for classification. The researchers claim high accuracy, achieving 95.19 % and 94.79 % using three-fold and ten-fold cross-validation, respectively [4].

The type of Glaucoma diagnosis is outlined by a method employing an advanced fusion ensemble model incorporating Order-one 2D-FBSE-EWT and ResNet-50. Through this approach, fundus images are divided into sub-images using Order-one 2D-FBSE-EWT, enabling deep feature extraction by ResNet-50. The resulting deep features from these sub-images are then combined, streamlined using principal component analysis, and finally assessed through a Softmax classifier for classification. The investigation assesses various channel grouping strategies, presenting their performance in 5-fold and 10-fold cross-validation tests, with the 3-channel technique demonstrating a 93 % accuracy for a balanced database and 78.3 % for an unbalanced one [5].

2. Related work

The use of a CAD system can aid in the early screening of Glaucoma, potentially preventing blindness in many patients. Funduscopy, via regular checkups for individuals, particularly those over the age of 40, can be employed for early screening. Optic disk segmentation is a crucial step for automatic Glaucoma detection with fundus images. In the past decade, CNN has proven to be a valuable tool in solving many real-world problems. A Convolutional Neural Network (CNN or ConvNet) is a network architecture for deep learning that learns directly from data, eliminating the need for manual feature extraction [6]. Transfer learning, a deep learning approach, uses a model trained for one task as a starting point for a similar task, allowing for the updating and retraining of a network [7]. CNN, a type of Artificial Neural Network, comprises layers for feature extraction and a fully connected network that leads to the output. In Glaucoma detection, CNN is used for feature extraction and classification.

Researchers have successfully used CNNs for Glaucoma detection, achieving high accuracy rates. For instance, an eighteen-layer CNN was developed to extract robust features from digital fundus images and classify them as normal or glaucomatous [8]. The study used retinal fundus images (589 normal + 837 glaucomatous) from Kasturba Medical College, Manipal, India, and achieved an accuracy of 98.13 %. However, this study used only one dataset from a single source with images of the same quality.

Another approach that has shown promising results by combining CNN with machine learning algorithms is the ensemble learning-based CNN architecture [9]. This method uses the circular Hough transform on the green channel for optic disc segmentation and the entropy sampling technique for selecting informative points, thus reducing computational complexity. These selected points are then input to a CNN based on boosting for classification. This approach can be beneficial in situations where large amounts of training data are not available and CNN is trained with extracted features.

Some researchers have utilized CNN for OD-OC segmentation. A semi-auto deep-level set method is proposed for the segmentation of the optic cup (OC) and optic disc (OD) in fundus images [10]. This method uses a multiscale convolution neural network for the prediction network and an evolution model is used to refine the contours. The network is integrated with augmented prior knowledge and supervised by active contour loss. This results in more accurate shape and boundary details.

An Expert System with Deep CNN architecture named Glaucoma Network (G-Net) [11] was designed for automatic Glaucoma detection. This expert system performs OD-OC segmentation using two neural networks in conjunction and manually evaluates the CDR for classification.

An efficient glaucoma master framework consisting of two CNN architectures of 39 layers is used for OD-OC segmentation [12]. The preprocessing includes morphological operations and CLAHE, while Sobel edge detection is used for initial segmentation and Watershed algorithm for localization of optic nerve. The optic nerve image is given as input to two individual CNNs for OD-OC segmentation. For classification, CDR is calculated manually after segmentation.

The OD-OC segmentation is also done using a graph convolutional network (GCN) [13]. A multi-scale CNN is used to create the feature map, which is concatenated with graph nodes and given as input to GCN for predicting OC and OD contours. The REFUGE and Dhrishti-GS1 datasets were used for experiments, and the proposed approach achieved good accuracy.

A novel deep-learning architecture, ResFPN-Net is used for OD-OC segmentation at once [14]. This architecture includes a multi-scale feature extractor for boundary extraction, multi-scale segmentation transition for retaining features of different scales, and attention pyramid architecture for recognizing conjoint connections among OD-OC. The accuracy achieved on Dhrishti-GS and RIM-ONE datasets was above 98 %.

A deep learning architecture was proposed for joint OD-OC segmentation based on the U-Net model [15]. The modified U-Net model is called FAU-Net (feature fusion and attention U-Net), which includes a feature fusion module to minimize information loss during feature extraction. The FAU-Net combines channel and spatial attention mechanisms to emphasize the relevant features and suppress irrelevant ones for the segmentation task. Additionally, a multi-label loss is utilized to produce the final joint segmentation of OD and OC. The experimental results demonstrate that FAU-Net surpasses the state-of-the-art segmentation methods for OD and OC on various datasets, including Drishti-GS1, REFUGE, RIM-ONE, and ODIR.

Glaucoma detection is performed using an automatic CADx framework consisting of the deep-belief network (DBN) and a contextualizing DL structure [16]. The DBN consists of automatic detection, feature extraction, feature optimization, and classification. The proposed model was compared with three models, SVM, RF, and k-NN, and an accuracy of 98 % was claimed for the proposed model.

The innovative feature extraction technique called evolutionary convolutional network (ECNet) was proposed for Glaucoma detection [17]. The technique uses a real-coded genetic algorithm (RCGA) to optimize weights at different layers, generating a feature vector that is then tested with several classifiers, including KNN, backpropagation neural network (BPNN), SVM, extreme learning machine (ELM), and kernel ELM (K-ELM). Results show that ECNet works best with SVM for fundus image classification.

Another Glaucoma classification method is based on a Perceptron-based Convolutional Multi-Layer Neural network [18]. This approach uses entropy-based estimation for OD-OC boundary detection and the Weighted Least square fit method for disc ratio calculation. The method showed accurate prediction in detecting the severity of Glaucoma.

A unified OD detection method that eliminates the need for a localization step was introduced for OD segmentation [19]. The approach combines a fully convolution neural network (FCNN) with a structured matrix decomposition (SMD) model to decompose a feature matrix created using the Simple Linear Iterative Clustering (SLIC) algorithm. The SLIC algorithm extracts color, texture, and boundary features, and a hierarchical segmentation tree is developed based on spatial connectivity and feature similarity. The SMD model and FCNN are then used to decompose the feature matrix into a sparse matrix representing pixels of OD.

A mixed training strategy for OD-OC segmentation was proposed using the GDCSeg-Net architecture. It consists of a multi-scale weight-shared attention (MSA) module and a densely connected depth-wise separable convolution (DSC) module [20]. The architecture also

consists of the encoder-decoder-based general OD-OC segmentation network. The model was evaluated on REFUGE [21], MESSIDOR [22], RIM-ONE-R3 [23], Drishti-GS [24] and IDRiD [25] datasets.

Self-Organized Operational Neural Networks (Self-ONNs) [26] were utilized for Glaucoma detection, based on Taylor series functions, and claimed as fewer complexes than CNNs. They achieved good accuracy on three benchmark datasets: ACRIMA [27], RIM-ONE [23], and ESOGU.

An offline CAD system designed using a combination of image processing, deep learning, and machine learning for Glaucoma detection [28]. Region of interest extraction was performed using the brightest spot algorithm, and classification was performed using machine learning classifiers like SVM, Neural Network, and AdaBoost, as well as deep learning models like Le-Net architecture [29] and U-Net architecture [30] for image validation and segmentation. The author claimed 100 % accuracy achieved by an ensemble of all classifiers on six datasets: RIM-ONE [23], DRISHTI-GS [24], DRIONS-DB [31], JSIEC, and DRIVE [32].

Transfer learning refers to a form of deep learning in which a previously trained model, designed for a specific task, is modified and retrained to perform a similar task. The primary benefit of transfer learning is that it allows models to be reused, eliminating the need to build them from scratch. In the context of detecting glaucoma using Fundus images, many researchers have developed convolutional neural networks inspired by pre-trained models.

A real-time cloud-based framework has been proposed for screening glaucoma using the EfficientNet model [33]. In this framework, the EfficientNet model serves as the encoder for OD-OC segmentation, while the UNet++ model [30] functions as the decoder in the pipeline. The CDR ratio is calculated manually for classification purposes. The researchers evaluated the performance of this model using benchmark datasets such as RIM-ONE and DRISHTI-GS1. The authors claim that this system is scalable, location-independent, and easily accessible.

The multilabel classifier called MCGL-Net was proposed for classifying multiple retinal diseases [34]. The MCGL-Net is a composite model based on a hybrid graph. It utilizes EfficientNet-B4 for feature extraction and LightGBM for classification. According to the author, the proposed model outperformed all other methods on the ODIR dataset, achieving an F1 score of 91.60 % on the test set.

A framework based on CNN is proposed for segmenting the optic cup and disc, combining the EfficientNet and U-Net models [35]. To enhance the local contrast, CLAHE is used for image enhancement. The classification of the image is based on manually calculated CDR.

An asymmetric deep learning network is designed for accurate optic disc segmentation from fundus images. It has a unique sub-network and a decoding convolutional block based on the classical U-Net model [36]. The model uses a feature detection sub-network (FDS) and a cross-connection sub-network (CCS) in combination. The performance of the model is compared with that of U-Net [30], M−Net [37], and the Deeplabv3 model [38] on the MESSIDOR [22], ORIGA [39], and REFUGE [21] datasets, and is found to be the best among them.

For the joint cup and disc segmentation, a modified M−Net with bidirectional convolution long short-term memory (LSTM) based on the UNet model is proposed [40]. Images are pre-processed by converting them to polar coordinates, which are then given as input to the deep learning model. According to the author, the proposed model accurately observes abnormalities of the retinal image on the REFUGE2 dataset.

The visual saliency detection technique is proposed [41] for optic disc detection. The proposed method utilizes a morphological reconstruction-based object detection approach for localizing OD. To segment the image patch, the Simple Linear Iterative Cluster (SLIC) method is employed, which divides the image patch into numerous superpixels. Each superpixel is then assigned an initial saliency value based on background prior information. The deep features of each superpixel are extracted using a pre-trained fully convolutional network (FCN) and are used as input to the single-layer cellular automata (SCA) framework for accurate OD detection.

An innovative approach for optic disc segmentation is proposed using graph-based saliency [42]. For glaucoma classification, an ensemble of three deep learning architectures AlexNet [43], ResNet-50 [44], and ResNet-152 [45] are used. In the final stage, the three CNNs are fused with probabilistic models to improve glaucoma detection accuracy. The researchers tested the model's performance with and without saliency features, and with individual and ensemble CNNs. They found that the ensemble model with saliency features outperformed all other combinations.

A novel ensemble model is proposed for the classification of glaucoma detection [46]. The model comprises three individual architectures, ResNet-50 [44], VGGNet-16 [47], and GoogLeNet [48]. The output of each architecture is used to create a prediction vector, and majority voting is employed for the final prediction. The authors evaluated the performance of individual architectures and the ensemble of these architectures for glaucoma prediction. They claimed that their proposed method outperformed all individual architectures.

A new model called Optic Disk localization and Glaucoma Diagnosis Network (ODGNet) was proposed to segment the optic disk [49]. The model includes a visual saliency map incorporated with a shallow CNN for optic disk localization and transfer learning-based pre-trained models for glaucoma diagnosis. The performance of ODGNet with transfer learning-based models such as AlexNet [43], ResNet, and VGGNet incorporating saliency maps was evaluated on five public retinal datasets (ORIGA [39], HRF [50], DRIONS-DB [31], DR-HAGIS [51], and RIM-ONE [23]). The results indicate that incorporating a saliency map with a shallow CNN improves the performance of ODGNet.

A methodology that advances explainable deep learning and vertical cup-disc ratio (VCDR) was proposed for glaucoma detection [52]. The focus of this research is on the cropping policy of fundus images. To achieve accuracy in cropping, a series of operations were performed on the image, and the inverse of the cropping mask was also performed. The UNet model was used for segmentation and the modified ResNet was used for classification. The fully connected layers were converted to fully convolution layers by a global average polling operation. The authors claimed that this model can achieve significant performance even in the absence of the Optic Nerve Head.

A glaucoma assessment pipeline that runs offline on mobile devices was designed using CNN [53]. The system is based on the MobileNetV2 [54] architecture and the GFI-ASPP-Depth model. This U-shaped CNN model includes an encoding and decoding module for joint segmentation of the optic disc and optic cup. Morphological operations are performed to extract features such as area, vertical length, and rim disc area from the segmented image. A classification network called GFI-C was created using MobileNetV2 to classify images as glaucomatous or normal.

In a recent study, researchers evaluated the performance of five distinct CNN architectures that were previously trained on ImageNet (VGG16, VGG19, InceptionV3, ResNet50, and Xception) for glaucoma classification across five publicly available datasets (ACRIMA, HRF, Dhrishti-GS1, RIM-ONE, and sjchoi86-HRF) [27]. Their findings indicate that the Xception architecture demonstrated the most superior performance among all five models in terms of glaucoma classification accuracy.

3. Methodology

This section outlines the proposed segmentation and classification framework for Glaucoma screening, as illustrated in Fig. 1. Prior to classification, it was observed that some fundus image datasets contain complete retina images while others contain segmented optic disks. To determine whether the input image is segmented or not, the image is first converted to grayscale, and a histogram is created. In complete retinal images, the background is black with an intensity of '0'. If more than 20 % of the pixels in the image are black, the image is segmented to obtain the region of interest, which is then resized and given as input to

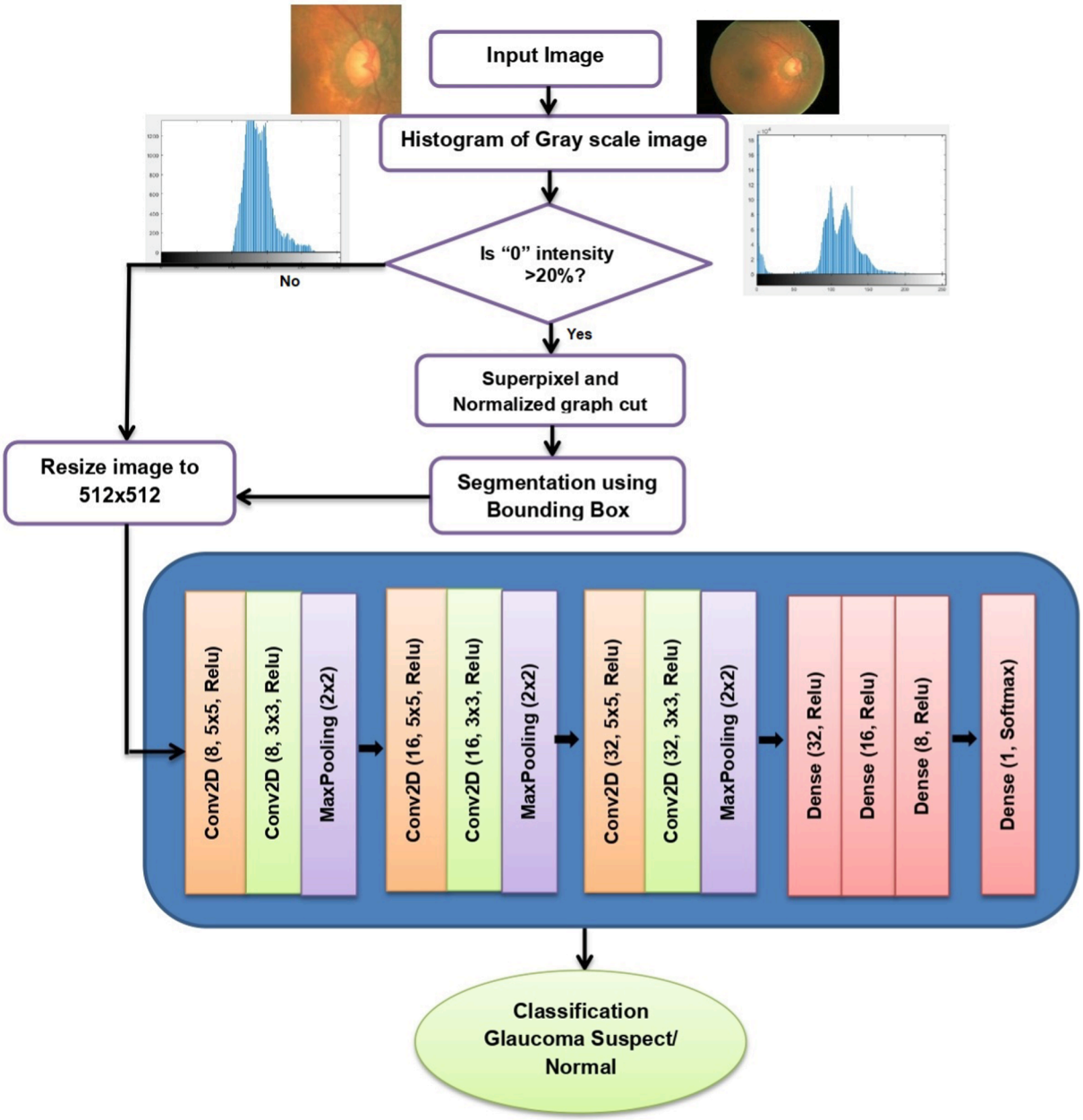


Fig. 1. Overview of the proposed framework for Glaucoma classification.

the proposed convolution neural network for classification. If the image is already segmented, the segmented optic disk is directly given as input to the proposed CNN. Optic disk segmentation is performed using the Simple Linear Iterative Clustering (SLIC) and normalized graph cut algorithm. After segmentation, the optic disk image is resized and given as input to the thirteen-layer convolution neural network for classification.

3.1. Image pre-processing

The preprocessing phase involves converting the input image to greyscale to reduce computational requirements and simplify processing. The complete retinal image in the dataset has a black background, and a histogram is created to determine the type of image (complete or OD segmented). The background of the greyscale image contains black pixels with an intensity of "**0**". If the input image is already segmented, the proportion of "**0**" intensity pixels will not exceed 20%. Thus, the presence of "**0**" intensity pixels is used as a criterion to determine whether segmentation of the image is required. If the image needs to be segmented to the optic disk, the fundus image is converted to the L*ab* color space. The L* component denotes lightness, while the a* and b* components represent the red/green and yellow/blue coordinates, respectively. This conversion is effective since the region of interest has varying color intensities of red, yellow, and orange. If the image is already segmented, it is resized to 512x512 and given as input to the proposed CNN.

3.2. Image segmentation

The imperative significance of accurate optic disk segmentation in the realm of ocular disease analysis has been underscored, given its pivotal role in facilitating early detection and effective treatment. Throughout the study, diverse methodologies encompassing image processing, machine learning, and deep learning have been meticulously examined, each offering unique advantages and challenges. The proposed approach, employing a synergistic fusion of SLIC and graph cut normalization, emerges as a noteworthy contender in the optic disk

segmentation landscape. Notably, the method's performance closely approaches that of the deep learning-based UNet model, signifying its potential as a viable solution.

3.2.1. Simple linear iterative clustering (SLIC)

The Simple Linear Iterative Clustering (SLIC) method is an effective and straightforward approach for generating meaningful and compact clusters in the image structure that represents an object. In this technique, each pixel is represented by a five-dimensional vector $C_i = [l_i\ a_i b_i x_i y_i]^T$, which provides good similarity perception in color images. The SLIC superpixels correspond to the clusters in the image plane space. To initialize the cluster centers $C_k = [l_k\ a_k b_k x_k y_k]^T$, the pixels are sampled at regular grid steps S, where $S = \sqrt{(N/k)}$, N is the number of pixels, and k is the desired number of clusters. The cluster centers are moved to the lowest gradient position in a 3x3 neighborhood to obtain the final position of the pixel [55]. The similarity index for forming the clusters is defined as follows:

$$D_{ij} = m\frac{d_s(i,j)}{s} + d_c(i,j)$$

where m controls compactness and s is used for normalizing the spatial distance $d_s(i, j)$. $d_s(i,j)$ and $d_c(i,j)$ are spatial and color differences respectively and are defined as

$$d_s(i,j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

$$d_c(i,j) = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2}$$

After conducting experiments, a value of ten was chosen for the compactness parameter, and the SLIC technique was used to create eighty clusters based on color variations in the input image. The input image is shown in Fig. 2a, while Fig. 2b displays the same image converted to the L*a*b color space. Finally, the output of the SLIC technique with the created clusters is shown in Fig. 2c.

3.2.2. Normalized graph cut algorithm

A partitioning technique was proposed based on normalized cuts, which is effective in minimizing the cut value and achieving a better partition of an image represented as a graph [56]. To obtain a region of interest, the superpixel clusters obtained in the previous step are input into the algorithm. In the normalized graph cut approach, the image is represented as a weighted undirected graph G = (V, E), where V represents the superpixel clusters as nodes, and E represents the edge between them. The partition is based on the dissimilarity between the clusters, which can be computed as the total weight of the edges that have been removed. This degree of dissimilarity between the two parts is called the cut in graph-theoretic language. It is represented as:

$$cut(A,B) = \sum_{u\in A, v\in B} w(u,v)$$

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)}$$

In this step, the clusters obtained from the SLIC algorithm are given as input to the normalized graph cut algorithm to generate a region of interest. The normalized graph cut algorithm represents the image as an undirected weighted graph G = (V, E), where V represents the superpixel clusters as nodes and E represents the edges connecting them. The cut between two clusters is the sum of the weights of the edges removed. In mathematical notation, assoc (A, V) represents the sum of the weights of all edges from A to all nodes (V) or pixels in graph G, and assoc (B, V) is computed similarly [57]. Fig. 3(a) shows the SLIC image, where the clusters formed are given to the normalized graph cut algorithm to form three clusters, namely the background, retinal image, and optic disk, as shown in Fig. 3(b). The region of interest is segmented using a bounding box, and the optic disk is segmented from the input image, as shown in Fig. 3(c).

3.3. Proposed CNN for classification

CNNs are an effective solution for image classification as they combine feature extraction and classification into one structure. In medical image classification, CNNs have shown promising results using complex networks. The proposed CNN in this study is detailed in the following section.

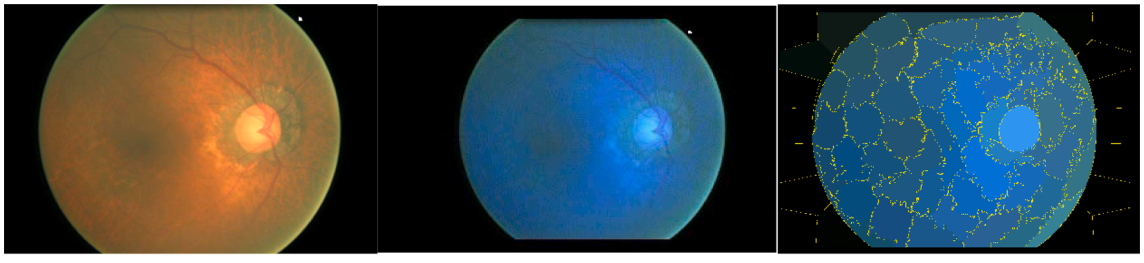
The CNN baseline deep neural network model in this study is tailored specifically to the domain of Fundus images for Glaucoma detection. Crafted from the ground up, it is meticulously designed to enable fine-tuning at each layer. This approach to model design is built upon extensive experimentation. The CNN comprises different layers, including the input layer, convolution layer, pooling layer, and dense or fully connected layer. Each layer is followed by activation functions such as ReLU and Softmax. Fig. 4 illustrates the architecture of the proposed CNN for the classification of fundus images into Glaucoma suspect and normal eye, and Table 1 provides a detailed description of the proposed CNN, including the number of kernels, activation functions, and the size of the output generated in each layer.

Input Layer: The input layer of the CNN receives the segmented RGB image of the Optic disk, which consists of pixel values of a 3-dimensional input image of size 512x512. Thus, the input size of this proposed network is 512x512x3.

Convolution Layer: This layer convolutes learnable kernels on the input image, generating feature maps from the input image's depth. Each kernel is integrated with a corresponding activation function, which is stacked along the depth dimension to form the full output volume from the convolutional layer [58].

Pooling Layer: This layer can perform down-sampling, along the spatial dimensionality of the input. It focuses on selecting features and reducing the number of parameters within that activation [58]. The user selects the value of stride, and the output is generated accordingly.

Dense Layer or Fully Connected Layer: In a fully connected layer, each neuron applies a linear transformation to the input vector through



(a) (b) (c)

Fig. 2. (a) Input image (b) L*a*b image (c) Clustered image after SLIC.

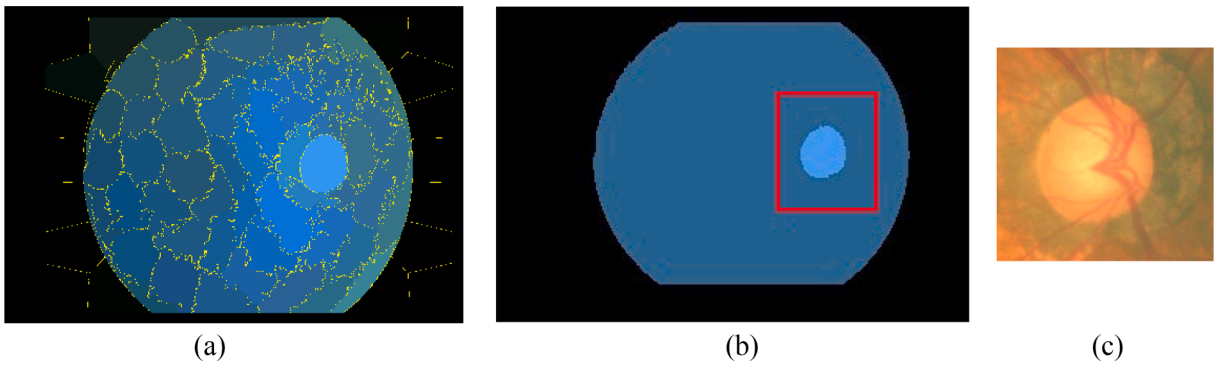


Fig. 3. (a) SLIC image (b) Normalized graph cut image (c) Segmented Optic Disk.

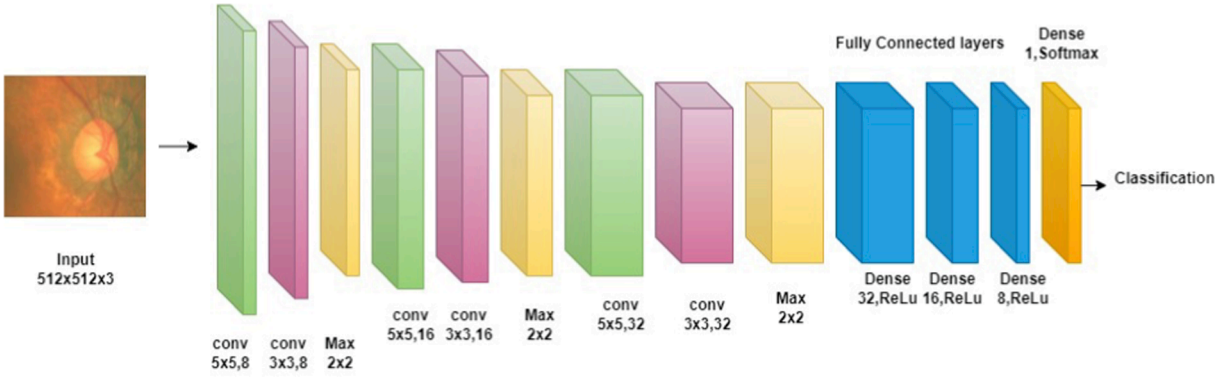


Fig. 4. Thirteen-layer convolution neural network.

Table 1
Description of 13-layer CNN architecture.

Sr. No.	Layer Name	Activation Function	Output shape
1	Input layer	–	512 x 512 x 3
2	Conv2D_0 (8 filters)	ReLu	508x508x8
3	Conv2D_1(8 filters)	ReLu	506x506x8
4	MaxPooling_1(2x2)	–	253x253x8
5	Conv2D_2 (16 filters)	ReLu	249 x 249 x 16
6	Conv2D_3(16 filters)	ReLu	247 x 247 x 16
7	MaxPooling_2(2x2)	–	123 x 123 x16
8	Conv2D_4 (32 filters)	ReLu	119 x 119 x32
9	Conv2D_5(16 filters)	ReLu	117 x 117x 32
10	MaxPooling_3(2x2)	–	58 x 58x 32
11	Dense_0(32 layers)	ReLu	58 x 58 x 32
12	Dense_1(16 layers)	ReLu	58 x 58 x 16
13	Dense_2(8 layers)	ReLu	58 x 58 x 8
14	Dense_3 (1 Layer)	Softmax	1

a weights matrix, resulting in all possible connections layer-to-layer being present, which are densely connected [58].

4. Activation functions

1) The Rectified Linear Unit (ReLU) is a straightforward mathematical function that behaves as the identity function for positive inputs and returns zero for negative inputs. This activation function is commonly utilized in convolutional neural networks. [59].

$$\boldsymbol{ReLU}(\boldsymbol{x}) = \begin{cases} x, & if\ x \geq 0 \\ 0, & otherwise \end{cases}$$

2) Softmax function: The function is a composite of several sigmoid functions and produces outputs between 0 and 1. These values can be interpreted as probabilities representing the likelihood of a given data point belonging to a particular class. [59].

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{k} e^{z_k}} \quad for \quad j = 1, ..., k$$

5. Experiments

The complete framework is developed in Python Jupyter Notebook and executed using a Docker image jupyter/Tensorflow-notebook with system configuration of Intel CPU i3 6600, 3.8 GHz, 24 GB RAM with GPU. This section covers the dataset used, evaluation metrics, and implementation details.

5.1. Dataset

This study utilized six widely recognized datasets of fundus retinal images, as listed in Table 2. The images were in PNG, JPEG, and JPG formats, and had high resolution. The datasets were obtained from various countries, and the table lists the number of images for each class used in the research.

Table 2
List of Dataset used in Research.

Sr. No.	Name of Dataset	Total Images	Normal Images	Glaucoma images
1	Dhrishti GS-1 [24]	101	31	70
2	ORIGA [39]	650	482	168
3	ACRIMA [27]	705	309	396
4	PAPILA [60]	480	333	147
5	G1020 [61]	1020	724	296
6	RIM-ONE r3 [23]	159	85	74
TOTAL		**3115**	**1964**	**1151**

5.2. Evaluation metrics

To evaluate the performance of the models, the research considered the following evaluation metrics. The dataset was divided into 80 % for training the models and 20 % for testing the models. The following performance metrics were used to evaluate the performance of the research:

The confusion matrix depicts True positive (T_{+ve}), true negative (T_{-ve}), false positive (F_{+ve}) and false negative (F_{-ve}) values based on the performance of a classifier [62].

Accuracy defines how well the model works and it is a ratio of correct predictions with total predictions. While calculating accuracy, the ratio of true positive (T_{+ve}) and true negative (T_{-ve}) with total predictions taken [62]. It is formulated as follows:

$$\text{Accuracy} = \frac{T_{+ve} + T_{-ve}}{T_{+ve} + T_{-ve} + F_{+ve} + F_{-ve}}$$

For instance in medical diagnosis, complete treatment is dependent on the predicted diagnosis. So parameters like Precision, Recall, and F-measure also need to be considered for evaluating the models. Precision is about measuring accuracy for positive results. It is the ratio of the total number of positive samples predicted correctly as positive samples to the total samples predicted as positive (positive or negative) by the model [62].

$$\text{Precision} = \frac{T_{+ve}}{T_{+ve} + F_{+ve}}$$

The recall deals with the efficiency in classifying positive samples. It is the ratio of the number of positive samples predicted positive to the total number of positive samples [62].

$$\text{Recall} = \frac{T_{+ve}}{T_{+ve} + F_{-ve}}$$

The F-measure is the harmonic mean metric that gives equal weightage to both precision and recall [62].

$$\text{F} - \text{measure} = \frac{2 * Precision * Recall}{Precision + Recall}$$

5.3. Implementation details

5.3.1. Segmentation performance

The accurate extraction of a region of interest is crucial for image classification. However, basic image processing techniques such as CLAHE, morphological operations, and thresholding are insufficient to produce accurate segmentation results. Studies have shown that these techniques produced results that deviated by almost 20 % from the ground truth in images with low contrast and illumination problems. By converting images to the L*a*b format, it was found that even in low illumination, segmentation results matched the ground truth. To achieve accurate segmentation, the SLIC technique is used to form multiple clusters, which are then processed using a normalized graph cut to obtain the region of interest. The proposed method outperformed basic image processing techniques on RGB images.

5.3.2. Classification performance

In this experiment, we have used a 10-fold cross-validation approach to train and test the proposed model on six standard image datasets. In each dataset, we maintained the same ratio of normal and glaucoma images in the subset of the cross-validation experiment. During the testing phase, we combined the classification results of the networks for each fold dataset, and all the testing results were used to generate confusion matrices for the three pre-trained networks and our proposed framework. The classification performance of the CNN models is shown in the form of a confusion matrix in Fig. 5.

We compared the performance of our proposed model with three standard CNN architectures: VGG19 [63], InceptionV3 [64], and ResNet50V2 [48] selected to distinguish the architecture design. These pretrained networks were originally trained on the ImageNet dataset, which encompasses over a million images for 1,000 distinct object classes. Our initial experimentation shows that these networks on fundus images, without fine-tuning, revealed an accuracy between 60 % and 70 %. We used the same dataset for training and testing these models and reserved 20 % of the images for testing. Table 3 presents the

Table 3
Comparative performance analysis of four CNN.

Method	Accuracy (Average %)	Precision (Average)	Recall (Average)	F-measure (average)
Proposed System	96.33	0.99	0.95	0.97
VGG19	91	0.94	0.91	0.92
InceptionV3	92.5	0.94	0.93	0.93
ResNet50V2	87.5	0.89	0.9	0.89

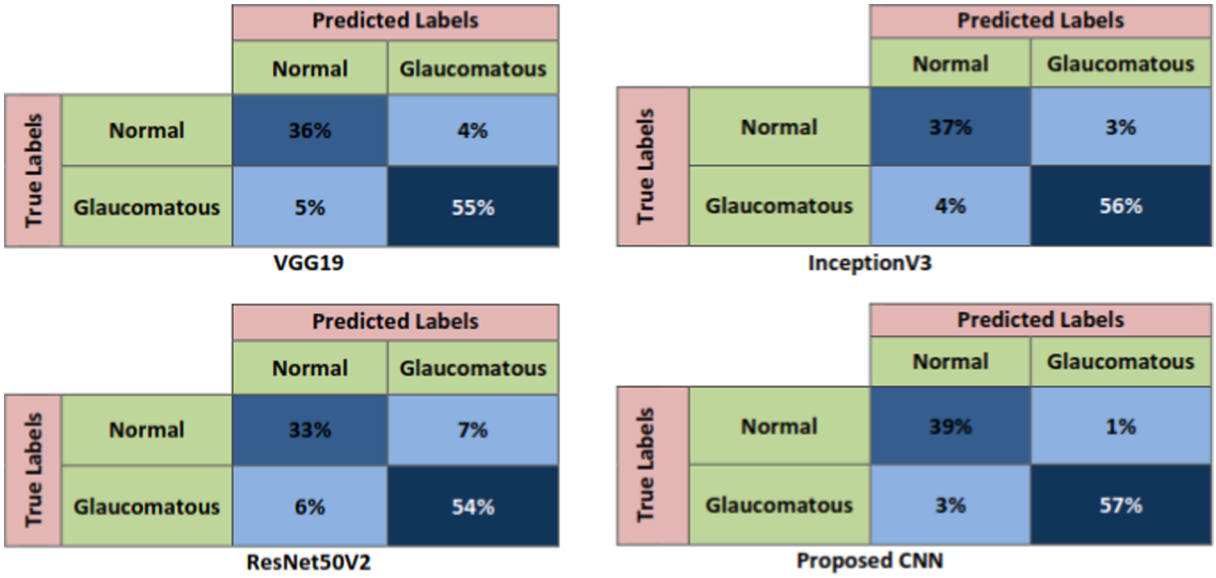


Fig. 5. Confusion matrices for four 2D –CNN in classifying Normal and Glaucomatous suspect.

performance of the proposed CNN and three fine-tuned architectures in terms of glaucoma disease classification, with evaluation metrics such as Accuracy, Precision, Recall, and F-measure. The results show that our proposed model outperforms all the other architectures, achieving an accuracy of 96.33 %.

6. Conclusion

In recent times, there has been a gush in the development of methods for medical image segmentation and classification. Researchers have demonstrated the potential of machine learning algorithms and deep learning algorithms in revolutionizing the medical field and accelerating the diagnostic process. This paper proposes an image sequence-based framework that can perform segmentation when necessary or classify input images into normal or Glaucoma suspects using the proposed CNN. In datasets where the segmented optic disk image is provided, segmentation is unnecessary, and these images are directly given as input to the proposed CNN. To segment, the optic disk from the retinal fundus image, SLIC, and the normalized graph-cut algorithm are employed. The basic image processing method, which consists of CLAHE, morphological operations, and thresholding, was found to be ineffective in low contrast or illumination problem scenarios when compared to the proposed method. Based on experimental results, a thirteen-layer convolutional neural network was designed for the classification of fundus images. To train and test these models, 3115 images from six standard datasets were used. The proposed method can accurately classify fundus images into normal and Glaucoma with an accuracy of 96.33 %. The performance of the proposed CNN was compared with pre-trained neural networks, such as VGG19, InceptionV3, and ResNet50V2, on six standard datasets. The results reveal that the proposed method achieves superior performance when compared to these 2D deep networks. The proposed model is expected to serve as an aid for clinicians or healthcare workers in the early screening of Glaucoma.

CRediT authorship contribution statement

Jignyasa Sanghavi: Conceptualization, Methodology, Software, Validation, Writing – original draft. **Manish Kurhekar:** Validation, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Online available standard dataset are used for research work. All dataset are cited in the article

References

[1] D. A. Rastogi, "World Glaucoma Week 2022," 13 March 2022. [Online]. Available: https://www.nhp.gov.in 〉 world-glaucoma-week-2022_pg. [Accessed 30 June 2022].

[2] L.S. Michaël, J.A. Girard, "Artificial intelligence and deep learning in glaucoma: Current state and future prospects", in Progress in Brain Research Sci. Direct. (2020) 37–64.

[3] S. Maheshwari, R.B. Pachori, U. Rajendra Acharya, "Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images", IEEE J. Biomed. Health Inform. 21 (3) (2017) 803–813.

[4] S. Maheshwari, Ram Bilas Pachori, Vivek Kanhangad, Sulatha V. Bhandary, U. Rajendra Acharya, "Iterative variational mode decomposition based automated detection of glaucoma using fundus images", Comput. Biol. Med. 88 (2017) 142–149.

[5] P.K. Chaudhary, S. Jain, T. Damani, S. Gokharu, R.B. Pachori, "Automatic diagnosis of type of glaucoma using order-One 2D-FBSE-EWT," in *24th International Conference on Digital Signal Processing and its Applications (DSPA)*, Russian Federation, Moscow, 2022.

[6] "Convolution Neural Network," [Online]. Available: https://in.mathworks.com/discovery/convolutional-neural-network-matlab.html. [Accessed 30 06 2022].

[7] "Transfer Learning for Training Deep Learning Models," [Online]. Available: https://in.mathworks.com/discovery/transfer-learning.html. [Accessed 30 06 2022].

[8] U. Raghavendra, Hamido Fujita, Sulatha V. Bhandary, Anjan Gudigar, Jen Hong Tan, U. Rajendra Acharya, "Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images", Inf. Sci. 441 (2018) 41–49.

[9] J. Zilly, J.M. Buhmann, D. Mahapatra, "Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation", Comput. Med. Imaging Graph. 55 (2017) 28–41.

[10] Y. Zheng, X. Zhang, X.u. Xiayu, Z. Tian, D.u. Shaoyi, "Deep level set method for optic disc and cup segmentation on fundus images", Biomed. Opt. Express. 12 (11) (2021) 6969–6983.

[11] M. Juneja, S. Singh, N. Agarwal, S. Bali, S. Gupta, N. Thakur, P. Jindal, "Automated detection of Glaucoma using deep learning convolution network (G-net)", Multimed. Tools Appl. 79 (2020) 15531–15553.

[12] H.N. Veena, A. Muruganandham, T. Senthil Kumaran, "A novel optic disc and optic cup segmentation technique to diagnose glaucoma using deep learning convolutional neural network over retinal fundus images", J. King Saud University – Comp. and Information Sci. (2021).

[13] Z. Tian, Y. Zheng, X. Li, Du. Shaoyi, Xu. Xiayu, "Graph convolutional network based optic disc and cup segmentation on fundus images", Biomed Optics Express. 11 (6) (2020) 3043–3057.

[14] G. Sun, Z. Zhang, J. Zhang, M. Zhu, X.-R. Zhu, J.-K. Yang, Y.u. Li, "Joint optic disc and cup segmentation based on multi-scale feature analysis and attention pyramid architecture for glaucoma screening", Neural Comput. & Applic. 35 (22) (2021).

[15] X. Guo, J. Li, Q. Lin, T.u. Zhenchuan, H.u. Xiaoying, S. Che, Joint optic disc and cup segmentation using feature fusion and attention, Comput. Biol. Med. 150 (2022) 106094.

[16] N. Patil, P.N. Patil, P.V. Rao, "Convolution neural network and deep-belief network (DBN) based automatic detection and diagnosis of Glaucoma", Multimed. Tools Appl. 80 (2021) 29481–29495.

[17] D.R. Nayak, D. Das, B. Majhi, S.V. Bhandary, U. Rajendra Acharya, "ECNet: An evolutionary convolutional network for automated glaucoma detection using fundus images", Biomed. Signal Process. Control. 67 (2021).

[18] F. Mansour Romany, Abdulsamad Al-Marghilnai, "Glaucoma detection using novel perceptron based convolutional multi-layer neural network classification,", Multidim. Syst. Sign. Process. 32 (2021) 1217–1235.

[19] Y. Wang, Y.u. Xiaosheng, W.u. Chengdong, "Optic disc detection based on fully convolutional neural network and structured matrix decomposition", Multimed. Tools Appl. 81 (2022) 10797–10817.

[20] Q. Zhu, X. Chen, Q. Meng, J. Song, G. Luo, M. Wang, F. Shi, Z. Chen, D. Xiang, L. Pan, Z. Li, W. Zhu, "GDCSeg-Net: General optic disc and cup segmentation network for multi-device fundus images", Biomed Optics Express. 12 (10) (2021) 6529–6544.

[21] Huazhu Fu Fei Li José Ignacio Orlando Hrvoje Bogunović Xu Sun Jingan Liao Yanwu Xu Shaochong Zhang Xiulan Zhang "REFUGE: Retinal Fundus Glaucoma Challenge" IEEE Dataport 2019. [Online]. Available: https://ieee-dataport.org/documents/refuge-retinal-fundus-glaucoma-challenge. [Accessed 30 06 2022].

[22] "MESSIDOR," [Online]. Available: https://www.adcis.net/en/third-party/messidor/. [Accessed 30 06 2022].

[23] F. Fumero, S. Alayon, J.L. Sanchez, J. Sigut, M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation", in 24th International Symposium on Computer-Based Medical Systems (CBMS) (2011).

[24] J. Sivaswamy, S.R. Krishnadas, G. Datt Joshi, M. Jain, A.U. Syed Tabish, "Drishti-GS: Retinal image dataset for optic nerve head(ONH) segmentation", in IEEE 11th International Symposium on Biomedical Imaging (ISBI) (2014).

[25] Prasanna Porwal Samiksha Pachade Ravi Kamble Manesh Kokare Girish Deshmukh Vivek Sahasrabuddhe Fabrice Meriaudeau "INDIAN DIABETIC RETINOPATHY IMAGE DATASET (IDRID) " 24-04-2018. [Online]. Available: https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid. [Accessed 30-06-2022].

[26] Ozer Can Devecioglu, Junaid Malik, Turker Ince, Serkan Kiranyaz, Eray Atalay, Moncef Gabbouj, "Real-time glaucoma detection from digital fundus images using self-ONNs", IEEE Access. 9 (2021) 140031–140041.

[27] A. Diaz-Pinto, S. Morales, V. Naranjo, T. Köhler, J.M. Mossi, A. Navea, "CNNs for automatic glaucoma assessment using fundus images: An extensive validation", Dataset. (2019).

[28] R. Shinde, "Glaucoma detection in retinal fundus images using U-Net and supervised machine learning algorithms", Intelligence-Based Medicine. 5 (2021).

[29] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition", Proc. IEEE. 86 (11) (1998) 2278–2324.

[30] N. Siddique, P. Sidike, C. Elkin, V. Devabhaktuni, "U-Net and its variants for medical image segmentation: Theory and applications", IEEE Access. 9 (2021) 82031–82057.

[31] Carmona, Enrique J., Rinc, "DRIONS-DB Retinal Image Database," [Online]. Available: https://www.idiap.ch/software/bob/docs/bob/bob.db.drionsdb/master/index.html. [Accessed 30 06 2022].

[32] "DRIVE: Digital Retinal Images for Vessel Extraction," [Online]. Available: https://drive.grand-challenge.org/. [Accessed 30 06 2022].

[33] H. Garg, N. Gupta, R. Agrawal, S. Shivani, B. Sharma, "A real-time cloud-based framework for glaucoma screening using EfficientNet", Multimed. Tools Appl. 81 (2022) 34737–34758, https://doi.org/10.1007/s11042-021-11559-8.

[34] K. Sun, M. He, X.u. Yao, W.u. Qinying, Z. He, W. Li, H. Liu, X. Pi, Multi-label classification of fundus images with graph convolutional network and LightGBM", Comput. Biol. Med. 149 (2022) 105909.
[35] N. Gupta, H. Garg, R. Agarwal, "A robust framework for glaucoma detection using CLAHE and EfficientNet", Vis. Comput. 38 (2022) 2315–2328.
[36] L. Wang, Gu. Juan, Y. Chen, Y. Liang, W. Zhang, Pu. Jiantao, H. Chen, Automated segmentation of the optic disc from fundus images using an asymmetric deep learning network", Pattern Recogn. 112 (2021).
[37] Raghav Mehta, Jayanthi Sivaswamy, "M-net: A convolutional neural network for deep brain structure segmentation", in IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017) (2017).
[38] Liang-Chieh Chen, George Papandreou, Florian Schroff, Hartwig Adam, "Rethinking atrous convolution for semantic image segmentation", arXiv: 1706.05587 (2017).
[39] Zhuo Zhang, Feng Shou Yin, Jiang Liu, Wing Kee Wong, Ngan Meng Tan, Beng Hai Lee, Jun Cheng, Tien Yin Wong, "ORIGA-light: An online retinal fundus image database for glaucoma analysis and research", in Annual International Conference of the IEEE Eng. in Medicine and Biology (2010).
[40] Maleeha Khalid Khan Syed Muhammad Anwar "M-Net with Bidirectional ConvLSTM for Cup and Disc Segmentation in Fundus Images" *arXiv:2104.03549* 2021.
[41] Y.u. Xiaosheng, Y. Wang, S. Wang, H.u. Nan, "Fully convolutional network and visual saliency-based automatic optic disc detection in retinal fundus images", J. Healthcare Eng. 2021 (2021) 11.
[42] S. Serte, A. Serener, Graph-based saliency and ensembles of convolutional neural networks for glaucoma detection, IET Image Proc. 15 (2021) 797–804.
[43] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "ImageNet classification with deep convolutional neural networks", Adv. Neural Inf. Proces. Syst. (2012) 1097–1105.
[44] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016.
[45] L.D. Nguyen, D. Lin, Z. Lin, J. Cao, "Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, Florence, Italy, 2018.
[46] B. Shubham Joshi, W.A. Partibane, H.T. Hatamleh, C.S. Yadav, D. Krah, "Glaucoma detection using image processing and supervised learning for classification", J. Healthcare Eng. 2022 (2022) 12.
[47] Karen Simonyan Andrew Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition" in *The 3rd International Conference on Learning Representations (ICLR2015)* 2015.
[48] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, "Going Deeper with Convolutions" in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015 pp. 1-9.doi: 10.1109/CVPR.2015.7298594.
[49] J. Latif, T.u. Shanshan, C. Xiao, S.U. Rehman, A. Imran, Y. Latif, "ODGNet: A deep learning model for automated optic disc localization and glaucoma classification using fundus images", SN Applied Sci. 4 (2022) 98.
[50] A. Budai "High-Resolution Fundus (HRF) Image Database," [Online]. Available: https://www5.cs.fau.de/research/data/fundus-images/. [Accessed 30 06 2022].
[51] S. Holm, G. Russell, V. Nourrit, N. McLoughlin, "DR HAGIS-a fundus image database for the automatic extraction of retinal surface vessels from diabetic patients", J. Medical Imaging. 4 (1) (2017) pp.
[52] R. Hemelings, B. Elen, J. Barbosa-Breda, M.B. Blaschko, P. De Boever, I. Stalmans, "Deep learning on fundus images detects glaucoma beyond the optic disc", Sci. Rep. 11 (2021) 20313.
[53] J. Martins, J.S. Cardoso, F. Soares, Offline computer-aided diagnosis for Glaucoma detection using fundus images targeted at mobile devices", Comput. Methods Programs Biomed. 192 (2020) 105341.
[54] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks", in 2018 IEEE/CVF Conference on Comp. Vision and Pattern Recognition (2018).
[55] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods", IEEE Trans. Pattern Anal. Mach. Intell. 34 (11) (2012) 2274–2282.
[56] Malik, Jianbo Shi, Jitendra, "Normalized cuts and image segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence. 22 (8) (2000) 888–905.
[57] R. Szeliski, "Segmentation," in Computer Vision:Algorithms and Applications, Springer. (2010) 267–306.
[58] Keiron O'Shea, Ryan Nash, "An introduction to convolutional neural networks", arXiv preprint. (2015).
[59] Shiv Ram Dubey, Satish Kumar Singh, Bidyut Baran Chaudhuri, "Activation functions in deep learning: A comprehensive survey and benchmark,", Neurocomputing. 503 (2022) 92–108.
[60] O. Kovalyk, J. Morales-Sánchez, R. Verdú-Monedero, et al., PAPILA: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment", Sci. Data 9 (291) (2022).
[61] Muhammad Naseer Bajwa, Gurbinder Singh, Wolfgang Neumeier, Muhammad Imran Malik, Andreas R. Dengel, Sheraz Ahmed, "G1020: A benchmark retinal fundus image dataset for computer-aided glaucoma detection,", in International Joint Conference on Neural Networks (2020).
[62] Mohammad Hossin, M.N. Sulaiman, "A review on evaluation metrics for data classification evaluations,", in International J. Data Mining & Knowledge Management Process. (2015).
[63] M. Mateen, J. Wen, S.S. Nasrullah, Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD", Symmetry. 11 (1) (2019).
[64] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the inception architecture for computer vision", in IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016).