Research on Biomedical Engineering manuscript No.
(will be inserted by the editor)

A Comprehensive Study of Optic Disc Detection in Artefact Retinal Images using a Deep Regression Neural Network for a fused Distance-Intensity Map

Ambika Sharma · Monika Agrawal · Sumantra Dutta Roy · Vivek Gupta

Received: date / Accepted: date

Abstract

Purpose Optic Disc (OD, hereafter) detection is often the first step to detect other retinal landmarks for analysis of conditions such as Glaucoma and Diabetic Retinopathy. It is often not possible to localize the OD based on colour/pixel information alone, especially for poor-contrast and low-resolution images. Community camp-based images under poor lighting conditions and hand-held ophthalmoscopes also induces imaging artefacts.

Methods The paper proposes an automatic OD detection method using a U-Net-based regression, with a Distance-Intensity map. The regression network uses Tukey's biweight loss function to make it robust to outliers, and improve the overall rate of convergence. The method localizes OD coordinates using a Generalized Laplacian-of-Gaussian (gLoG) operator on the predicted Distance-Intensity map. The system shows encouraging experimental results on poor resolution images with non-uniform illumination, noise, motion, blurring, and various imaging artefacts.

Results The experiments with intra- and inter-dataset performance (training and testing on different datasets) are with the following datasets: Messidor, Kaggle, DRIVE, DRIONS, STARE and Drishti-GS. The method shows excellent qualitative and quantitative results on Messidor and the challenging AIIMS Community Camp dataset, as well. The achieved detection accuracy is of 99.67% and 98.83% respectively.

Ambika Sharma
Indian Institute of Technology Delhi
Orchid ID0000-0001-5243-9128
E-mail: ambika.sharma@dbst.iit.ac.in

Monika Agrawal
Indian Institute of Technology Delhi
E-mail: maggarwal@care.iitd.ac.in

Sumantra Dutta Roy
Indian Institute of Technology Delhi
E-mail: sumantra@ee.iitd.ac.in

Vivek Gupta
All India Institute of Medical Sciences, New Delhi
E-mail: vgupta@aiims.ac.in

Conclusion The U-Net-based regressor with the novel loss function is geared towards getting good optic disc detection performance across a large number of datasets. The network shows robust detection performance on challenging images with various retinal artifacts (blurring, poor illumination, and clinical pathologies).

Keywords Retinal images · Optic Disc (OD) · Glaucoma diagnosis · Deep learning CNN models · Image processing

1 Introduction

Glaucoma is an irreversible disease which can lead to blindness. A simple non-invasive test for glaucoma for instance, considers estimating the cup-to-disc ratio from retinal images. A detailed examination of the OD in terms of its shape, size, color, sharpness of disc boundary, and swelling helps to diagnose optic neuropathies such as glaucoma, optic neuritis, drusden, peripapillary atrophy (PPA) and other anomalies (Dada and Coote 2010). An increased cup-to-disc ratio (enlarged optic cup) points to optic nerve fibre damage. Identification of the OD forms the first step in identifying the optic cup (Fig. 1) inside it. OD detection is also the start point for detecting blood vessels: the brightness lends better contrast to the darker vessels. The OD itself is an important ophthalmological landmark: this is the starting point for blood vessels to the retina, and where axons of neural ganglion cells come together (Shukla et al. 2020). A knowledge of the OD center allows the complete retinal vascular structure to be easily tracked.

OD detection forms a major step in identifying important features in the eye (Fig. 1) such as the blood vessels (veins and arteries), the optic cup and the macula. Ophthalmologists search for the macula as a slightly dark region lying within 3-4mm from the OD, as one goes away from the nasal side (Huang et al. 2020). Macula detection is an indicator for conditions such as Age-related Macular Degeneration. Macula-centered images are further used for detection of Diabetic Retinopathy.

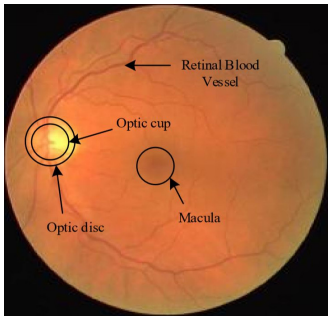


Fig. 1: A representative RGB retinal image, with the marked bright yellow OD and other landmarks of the eye.

OD detection is not an easy task, however. In local health centers and community camps, eye experts use a portable hand-held fundus cameras/low-resolution ophthalmoscopes to capture retinal images. Community camp images typically

suffer from non-uniform lighting, motion blur, low resolution, and other imaging artefacts (Dietter et al. 2019). Even if retinal imaging conditions were better than those in a community camp, there are additional difficulties in OD detection. Images are OD-centered only for Glaucoma analysis (Pathan et al. 2021). Apart from location variation of the OD in a retinal scan, a major challenge is posed by Macula-centered images (which are are used for Diabetic Retinopathy analysis). This leads to blurriness at the periphery of field of view where the OD is typically present. Fig. 2 shows a sample Macula-centered retinal image from the Messidor dataset (Messidor Retinal Database 2016) and an OD-centered retina from the DRIONS dataset (Maninis and Pont-Tuset 2010) respectively.

(a) (b)

Fig. 2: OD detection is difficult even in non-community camp retinal images. (a) Diabetic Retinopathy analysis needs a Macula-centered retinal image (Messidor Retinal Database 2016). Focusing on the Macula typically blurs the OD region. (b) Glaucoma analysis requires an OD-centered retina (Maninis and Pont-Tuset 2010).

This paper presents a robust OD detection mechanism, which is relatively insensitive to the specific shape, size, color and position/rotation of the OD in the entire retinal image region. Both types of retinal acquisitions (OD- and Macula-centered) are taken into account. The paper proposes a U-Net-based architecture with a fused Distance-Intensity map, which is relatively robust to exudates and other clinical pathological regions in the retinal image (with Tukey's biweight loss function). The final estimate of the OD uses a Generalized Laplacian-of-Gaussian (gLoG) blob detector.

The literature review is split into approaches primarily involving hand-crafted features-based image processing methods, and those based on machine learning.

Hand-crafted Features-based Image Processing Techniques

Early methods use geometric models of vessel structures to locate the OD region (Nergiz et al. 2018), (Pathan et al. 2019). A circular transformation is proposed in (Lu 2011) to incorporate both the circular shape and intensity variation property of an OD. The authors in (Lalonde et al. 2001) use the convergence of retinal blood vessels at the OD. To locate the OD, the authors use a Hausdorff distance-based template matching technique on the edge map. Another method (Mendonça et al. 2013) combines vascular and intensity information. The authors use the entropy of vascular directions as a new measure for detecting the convergence point. An automated method for OD detection (Panda et al. 2017)

uses various vessel cues like global symmetry, component count, and local vessel symmetry inside the OD region.

(Foracchia et al. 2004) propose a method to identify the OD location based on preliminary detection of main retinal vessels. It relies on the assumption that all retinal vessels originate from OD and diverge towards the periphery of the retinal image. The work proposes a geometrical parabolic model to describe the general direction of retinal vessels (the model parameters are the coordinates of the OD center). The resulting algorithm is not just based on the detection of vessel convergent region, it focuses on fitting of a model with respect to the entire vascular structure. Additionally, (Soares et al. 2016) propose a vessel enhancement and convergence algorithm to locate OD coordinates with vessel density and high intensity features. The authors in (Yu et al. 2012) propose a template matching-based approach for fast localization. The OD location is identified using vessel properties on the OD surface. Apart from this, another template matching-based approach is that of (Wankhede and Khanchandani 2016), where the color histogram of the OD region is used for similarity matching.

The authors in (Avinash Ramakanth and Venkatesh Babu 2014) use an Approximate Nearest Neighbour Field to find the correspondence between a chosen OD reference image and an input image. Some methods such as (Xiong and Li 2016) and (Sharma et al. 2017) use as set of vessel and disc features such as vessel direction, density, connected components, angle of intersection and OD intensity, edges, size of bright regions etc. for OD localization. Saliency map-based visual features are employed in (Liang et al. 2020) along with morphological operations to detect the OD region. Another visual saliency thresholding method for the extraction of the OD along with ROI generation is seen in (Subha and Rayen 2022).

Machine Learning-based Techniques

The work in (Wei et al. 2018) uses local feature spectrum analysis for OD location. The authors create a dictionary of local features corresponding to the cropped OD, to train k-Nearest Neighbor and SVM classifiers. The local features used are pixel intensities present in non-overlapping patches in sub-images. In the classification step, the local feature spectra are used as a feature matrix. The technique does not perform well due to the high variability of non-OD features. A 24-layered deep convolutional neural network regression in (Mitra et al. 2018) attempts to estimate the coordinates of bounding box around the OD. The authors test the method on five public datasets: Messidor, Drishti-GS, Kaggle, DRIVE, and STARE. In a similar approach, a region proposal network is used in (Huang et al. 2020) to generate multiple bounding boxes around the OD. The proposal with maximum probability is recognized as the OD region. The authors in (Niemeijer et al. 2009) use an assumption about the location of the OD and the fovea with respect to each other. A kNN regressor predicts the position of these anatomical landmarks using a set of features (number of vessels, width of vessels, standard deviation of vessel width and orientation etc.) extracted from the vessel map. Another work (Al-Bander et al. 2018) performs OD localization using a deep CNN: it uses a 10-layer convolutional network to predict the OD center coordinates, trained and tested on Messidor and Kaggle public datasets. A simultaneous detection of the OD and fovea anatomical landmarks in (Meyer et al. 2018), strives to learn the distance of

the anatomical structure from each pixel location instead of performing pixel-wise classification. This suggest a pixel-wise regression task.

A semi-supervised machine learning algorithm in (Jana et al. 2021) initially generates an estimate of the OD with a Circular Hough Transform (CHT). The exact location is estimated using a supervised convolutional network. A supervised OD detection method (Calimeri et al. 2016) uses a pre-trained convolutional neural network (CNN). In (Niu et al. 2017), a saliency map is generated from retinal images using standard frequency-tuned approach, and classified into OD or non-OD region using a CNN. (Serte and Serener 2021) and (Latif et al. 2022) propose an ensemble of graph-based saliency and CNN for cropping of the OD. The focus of (Fu et al. 2021) lies not on OD detection, but rather on a task that is closely associated yet distinct i.e., OD segmentation (which involves accurate estimation of the OD boundary). The authors use a line-fitting model on the vessel segments generated by a U-Net model. The probable region with maximum intersection points between line segments is defined as the OD region. This method is highly sensitive to the vessel extraction step. The work of (Tulsani et al. 2021) uses the complete retinal image for feature extraction using a UNet++ architecture and predicts glaucoma using the segmented region properties. Segmentation tasks have performance measures quite different from detection, such as the use of Dice and Jaccard coefficients. The approach of (Nawaz et al. 2022) is also primarily geared towards segmentation. The work uses the LabelImg software to annotate the bounding box for the optic disc region.

The above systems either deal with good resolution images or are limited in generalization, since they work with similar categories of retinal images for OD localization. Moreover, most algorithms perform badly when validated on large volume of retinal images containing either imaging or pathological artefacts (Trucco et al. 2019). In most practical cases the algorithm used to classify images may not be equivalent to those being used for training purpose. This paper proposes a deep learning-based network for OD localization. Sec. 2.1 proposes our U-Net-based Distance-Intensity Map regression. The distance-and-intensity-based representation, along with a method for mitigating the effect of exudates and other clinical pathologies. Sec. 3 shows the results of extensive experiments with our system, across a wide range of challenging datasets. A summary of the contributions of our work is as follows:

- The paper proposes a normalized Distance-Intensity Map to model the intensity changes with distance from the OD center, to obliterate bright exudate pathologies.
- A U-Net-based regressor with Tukey's biweight loss function imparts robustness to exudates, lesions and haemorrhages, and lighting artifacts, while estimating the OD location. The final OD estimate comes from a generalized Laplacian-of-Gaussian (gLoG) operator.
- The paper shows both qualitative and quantitative results of extensive experimentation with both inter- and intra-dataset performance, across a wide range of publicly available large datasets: Messidor (Messidor Retinal Database 2016), Kaggle (Kaggle Diabetic Retinopathy Detection Competition 2015), DRIVE, DRIONS, STARE (Maninis and Pont-Tuset 2010) and
  Drishti-GS (Sivaswamy et al. 2014). The proposed method also shows extremely encouraging qualitative and quantitative results on challenging AIIMS

Community Camp dataset. This has images taken with poor resolution hand-held ophthalmoscopes, with blurred and noisy images, taken under non-uniform illumination conditions.

2 Methods

2.1 A U-Net-based Distance-Intensity Map for OD Localization

The Distance-Intensity map is a gray-scale two-dimension image that is used to locate the region of interest (the OD). The problem is formulated as regressing the distance of each pixel from the OD center (x_{od}, y_{od}), while taking into account intensity distributions corresponding to the OD. Fig. 3 shows a diagrammatic representation of the main building blocks of the proposed method. In what follows,

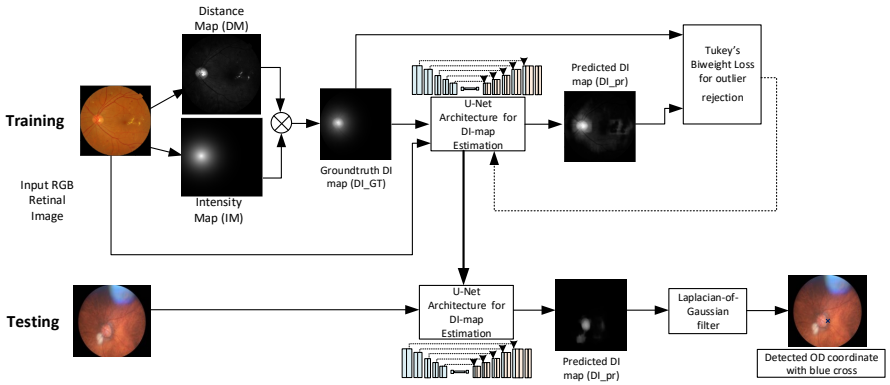


Fig. 3: An overall block-level representation of the proposed system

Sec. 2.1.1 describes our Distance-Intensity map. Sec. 2.1.2 proposes a U-Net-based method to estimate the Distance-Intensity map. Sec. 2.1.3 proposes a robust loss function (Tukey's biweight loss function) to suppress exudates and other clinical pathologies. Finally, a gLoG-based method is used to localize the OD center (Sec. 2.1.4).

2.1.1 ***Distance-Intensity Maps from Normalized Distance and Intensity Maps***

The motivation behind the Distance-Intensity Map is to model the intensity distribution in the OD region, with regard to both the distance from the OD center, and spatial intensity variations of the OD and obliterate the exudate bright pathology and light artefacts. The ground-truth Distance-Intensity map $DI_{GT}(x, y)$ or simply $DI(x, y)$ is the pixel-wise multiplication of normalized distance $D(x, y)^N_{GT}$ and intensity map $I(x, y)^N$ images.

$$DI_{GT}(x, y) = D(x, y)^N_{GT} \times I(x, y)^N \tag{1}$$

The ground-truth distance map $D(x,y)_{GT}$ can be calculated by finding the squared Euclidean distance of each pixel (x,y) from the OD center (x_{od}, y_{od}).

$$D(x,y)_{GT} = (x - x_{od})^2 + (y - y_{od})^2 \tag{2}$$

The normalized distance map $D(x,y)^N_{GT}$ is defined as follows.

$$D(x,y)^N_{GT} = \left(1 - \frac{D(x,y)}{\max_{(x,y)} D(x,y)}\right)^{\gamma} \tag{3}$$

Here GT represents the ground-truth and γ is a parameter that controls the decay of distance map $D(x,y)^N_{GT}$ over the image domain. γ is an integer greater than zero. The basic motivation behind the normalized distance map is to give a higher weightage of points closer to the expected OD center, and less weightage to exudates and artefacts, which usually occur farther away.

The motivation for the intensity map is the following. The OD region typically occupies 2-3% of the pixels in a typical retinal image, but accounts for about 60% of the total energy in a healthy retinal image (Niemeijer et al. 2009). The paper goes with a common convention of using the red $I(x,y)$ channel of the input RGB input image $X(x,y)$. The normalized Intensity map $I(x,y)^N$ lies in the range $[0,1]$:

$$I(x,y)^N = \left(\frac{I(x,y)}{\max_{(x,y)} I(x,y)}\right)^{\beta} \tag{4}$$

Similar to γ of the distance map, the intensity map uses a decay parameter β. Fig. 4 illustrates the relative effect of the decay parameters γ and β on the normalized Distance and Intensity maps, respectively. The top row shows the effect

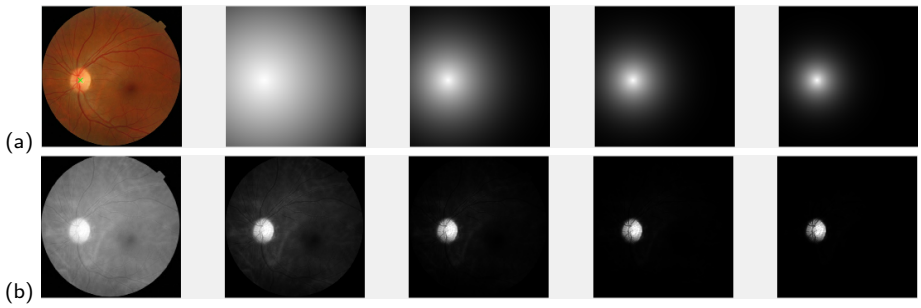


Fig. 4: The effect of decay parameters γ and β for sample values $\in \{1,3,5,7\}$, for (a) the normalized Distance Map, and (b) the Intensity Map. This also motivates the use of the fused Distance-Intensity Map.

of γ parameter variation for the distance map and bottom row shows the effect of β parameter variation for the intensity map. As γ values increase, the distance map values becomes more pronounced at the center of the OD. Similarly, for the intensity map, gray scale intensity values become more pronounced with high β values. The idea is to fuse the two maps namely distance and intensity in such a way that high values correspond to the OD region, and as one moves far from OD

center, the value decreases. The reason behind this approach is to downplay bright regions for retinal pathologies such as exudates or illumination artefacts. For such situations, the corresponding distance map coordinates are very low hence, the resultant product value will reduce. The goal is to look for high fused Distance-Intensity map areas i.e., coordinates with high normalized distance and intensity map values. This calculated ground-truth Distance-Intensity map $DI_{GT}(x, y)$ is further used for training the U-Net model, as explained in the following section.

2.1.2 ***Distance-Intensity Map Estimation using a U-Net Architecture***

A standard image generation (synthesis) procedure/task consists of generating new images from an existing dataset. Existing methods are based on architectures such as Generative Adversarial Networks (GANs) (Karras et al. 2018) and U-Nets (Ronneberger et al. 2015). The choice a U-Net-based architecture is for its proven track record in fast and precise generation of medical images (Ronneberger et al. 2015) (especially in localization and segmentation-related tasks). The paper proposes a U-Net-based scheme in an image-to-image translation (Isola et al. 2016) modality. It takes an RGB $X(x, y)$ image as input and generates the estimated gray scale Distance-Intensity map $\widehat{DI}(x, y)$. Given the ground-truth Distance-Intensity map $DI_{GT}(x, y)$(Sec. 2.1.1) and the estimated Distance-Intensity map $\widehat{DI}(x, y)$, the following section explains our loss function calculation. Fig. 5 shows the pipeline

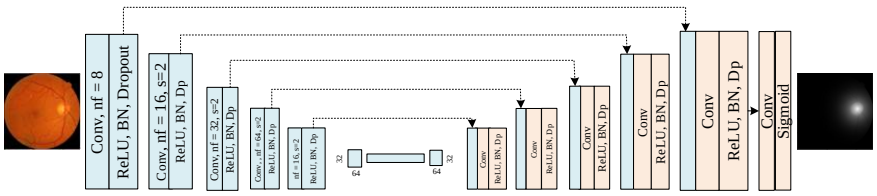


Fig. 5: The proposed U-Net-based structure for Distance-Intensity map generation. Each of the 5 blocks has convolution, activation, batch normalization (BN), and dropout. The output map gives the probability of association of each pixel to the OD region.

of our U-Net-based architecture. The input is an RGB retinal image of dimension 512×512. The output is a 2-D grayscale Distance-Intensity map. The model consists of five convolutional blocks for the encoder and decoder. Each convolutional block has a stride-two convolution operation followed by an activation function, batch normalization (BN) and dropout layer. The decoder (up-sampling) side imitates the same structure with five convolutional blocks: a convolution operation (with up-sampling) followed by an activation function, batch normalization and dropout layer. The skip connections concatenate the output at each level from the contracting side to the expanding one.

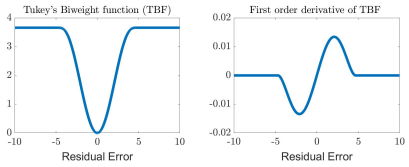
2.1.3 ***A Robust Loss Function to handle Artefacts, Exudates and other Clinical Pathological Regions***

For the OD detection problem, the regression model predicts the fused Distance-Intensity map values by minimizing the squared error between true(DI) and predicted($\widehat{DI}$) values summed over the entire image(X, with M rows and N columns). Retinal images often contain artefacts due to the motion blur, dust particles or insufficient and non-uniform illumination (Fig. 7 shows some representative examples of the same). Poor lighting (such as in a bright room) often bleeds excessive light in a retinal image. The green/blue lenses in a camera also create imaging artefacts. Additionally, there could be exudates, lesions, haemorrhages and other clinical pathological artefacts in retinal images (as in Fig. 14(c)). For such cases, a conventional loss function based on the Mean Square Error (MSE), or the Mean Average Error (MAE) may not suffice:

$$MSE = \frac{1}{MN} \sum_{x=1}^{x=M} \sum_{y=1}^{y=N} [DI(x,y) - \widehat{DI}(x,y)]^2 \tag{5}$$

$$MAE = \frac{1}{MN} \sum_{x=1}^{x=M} \sum_{y=1}^{y=N} |DI(x,y) - \widehat{DI}(x,y)| \tag{6}$$

Any loss function based on these errors would make a learning algorithm such



(a) (b)

Fig. 6: Tukey's biweight loss function (a) and its derivative (b). This robust M-estimator in the loss function of the U-Net architecture makes it robust to outliers.

as back-propagation dominantly influenced by large error values caused by imaging and clinical artefacts (outliers). The Huber loss function is a hybrid of MSE and MAE (Belagiannis et al. 2015). The Huber loss is only resistant to outliers present in the response, not to the outliers present in the input (Belagiannis et al. 2015). Tukey's biweight function imparts outlier resilience in the U-Net learning procedure. Fig. 6 shows a plot of Tukey's biweight loss function and its derivative, to the right. Tukey's biweight function down-weights the training pixel samples which have large residual value (generated by outliers) and use other samples with a small error in the back-propagation training.

$$\rho(r) = \begin{cases} c^2 \left[1 - \left(1 - \frac{r^2}{c^2} \right)^3 \right], & |r| \leq c \\ \frac{c^2}{6}, & \text{otherwise} \end{cases} \tag{7}$$

here c is a tuning constant that is usually (empirically) set to 4.6851, and r is the residual error between ground-truth (DI) and estimated map value ($\widehat{DI}$).

$$r = \widehat{DI}(x, y) - DI(x, y) \tag{8}$$

The loss function can be written as follows.

$$Loss = \frac{1}{MN} \sum_{x=1}^{x=M} \sum_{y=1}^{y=N} \rho(DI(x, y) - \widehat{DI}(x, y)) \tag{9}$$

A loss function based on Tukey's biweight function shows faster convergence and better generalization while training (Belagiannis et al. 2015).

2.1.4 ***gLoG-based OD Localization on the Estimated Distance-Intensity map***

The motivation behind using Tukey's biweight loss function in the U-Net-based Distance-Intensity estimation structure (Sec. 2.1.2) is to impart robustness to the presence of outliers, such as exudates, lesions and haemorrhages, and lighting artefacts. The proposed system uses a Generalized Laplacian-of-Gaussian (gLoG) operator (Kong et al. 2013) to identify the OD around the predicted location from the Distance-Intensity map ($\widehat{DI}(x, y)$). The OD typically appears as a whitish blob with a diffused boundary. In the RGB image, the red channel is used to extract the most probable region with the $\widehat{DI}(x, y)$ value at each pixel. The clinical reason behind the choice of the channel is that the OD typically appears highly saturated in the red channel as compared to the others. As the scale of gLoG operator increases, blob-like structures converge to local extrema at some scale (Kong et al. 2013).

3 Results

3.1 **Datasets**

To check for the robustness and efficacy of the proposed work, our experiments span six public and one private dataset. The six public datasets are the Kaggle Diabetic Retinopathy dataset, Messidor, Drishti-GS, DRIONS, STARE and DRIVE. Experiments on the challenging AIIMS Community Camp dataset also yield encouraging results. This has 2192 community camp-based images collected from the R. P. Centre, AIIMS New Delhi. This dataset has images taken in poor lighting conditions, using hand-held ophthalmoscopes. These result in various types of imaging artefacts (such as non-uniform illumination and blur), as in Fig. 7. Only the Kaggle (Kaggle Diabetic Retinopathy Detection Competition 2015) and the AIIMS (private) Community Camp dataset contain community camp or local health care-collected retinal images. Table 1 gives details about the datasets used in our experimentation. Of the 35,000 images in the Kaggle dataset (Kaggle Diabetic Retinopathy Detection Competition 2015) 30,337 have been manually chosen by an experienced clinical practitioner (The rest of the images have either no OD region, or a small part of the OD region present in them). A similar filtering by a trained clinical practitioner discards 64 out of 1200 images of the Messidor dataset (Messidor Retinal Database 2016).

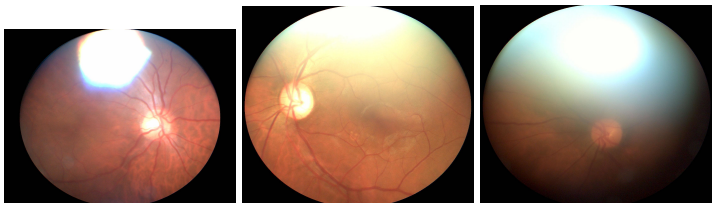
Table 1: Details of retinal image datasets used in this paper

S.No	Dataset Name	No of images	Image Dimensions
1	Drishti-GS (Sivaswamy et al. 2014)	101	2896×1944
2	Messidor (Messidor Retinal Database 2016)	1200	768×584
3	DRIVE (Maninis and Pont-Tuset 2010)	40	768×584
4	STARE (Maninis and Pont-Tuset 2010)	81	605×700
5	DRIONS (Maninis and Pont-Tuset 2010)	110	600×400
6	AIIMS Community Camp	2192	1536×1152
7	Kaggle (Kaggle Diabetic Retinopathy Detection Competition 2015)	30337	4752×3168

3.2 **Pre-processing and Augmentation**

This section describes the pre-processing steps and data augmentation in our system. Data augmentation is essential for a deep learning-based method, which typically need a very large number of training samples, which are in general difficult to obtain for medical datasets. All retinal images are down-scaled to a size of 512×512. Prior work in the area (Czarnowski and Jedrzejowicz 2018) shows that a size of 512×512 strikes a good balance between acceptable computational complexity and retaining sufficient amount of important information in retinal images.

The data augmentation starts with 3 geometric transformations for each sample image: a zoom of 0.2, rotation of 30 degrees, and a vertical flip. As described in Sec. 2.1.3 and the previous section (Fig. 7 shows a few representative examples), the Kaggle and AIIMS Community Camp dataset have numerous representative challenging cases of artefacts. In order to enhance the diversity of training dataset, training samples are augmented by performing various geometrical and intensity operations. Further, the augmentation is done so as to have a proper balanced training set between healthy controls and cases of retinal disorders. Hence, it is important to have a representation from the set of artefact images as well, in the augmented training set. In order to generate a new set of training images from the existing ones with poor lightning, a CLAHE-based adaptive histogram equalization is used to improve the contrast in local regions. Small saturated regions are in-painted with information from neighbouring regions. For large saturated regions (which tend to show up in the green channel more than the other two, for RGB images), the intensity artefact region is tuned down, from information in the green channel. The set of blurry retinal images is augmented with sharpened versions. To augment images with lens artefacts, a Circular Hough Transform (CHT) is used to find circular blobs with radii between 10 and 20 pixels on the red channel of an RGB image. These are in-painted over, with information from neighbouring regions. For each image therefore, our augmentation process generates between 3 to 6 augmented images.



(a) light artefacts: bright light saturation on a portion of the retina

(b) motion artefacts and image acquisition noise

(c) lens artefacts: dust on the camera lens

Fig. 7: Examples of retinal artefacts: representative images from the challenging AIIMS Community Camp dataset. Such images have been taken in poor lighting conditions, using hand-held ophthalmoscopes.

3.3 **Experimentation and Results**

As mentioned before (Table 1, Sec. 3.1), our experiments span six public datasets, and one private one. For each dataset, the training-validation-testing ratios are 80%-10%-10%. For the proposed U-Net-based architecture (Sec. 2.1.2), the training uses an ADAM optimizer with step learning rate of 0.05 which is reduced by half in every tenth epoch. A dropout value of 0.5 is used with an optimal batch size of 16. The experimental test-bed is an Intel i7 Windows 10 system with an NVIDIA Quadro P5000 card and 128 GB main memory. A description of some ablation studies follows, with different parameters of the system.

3.3.1 ***Motivation for the Distance-Intensity Map Formulation***

This section shows results of experimental validation of the motivation for using fused Distance-Intensity maps of Sec. 2.1.1. Fig. 8 shows some representative results with a distance map alone. The first image in each row corresponds to input RGB image, middle one is the ground-truth distance map and the rightmost, the predicted distance map using the proposed architecture. The distance map alone clearly mis-classifies the bright exudates around the OD margin as a probable region for OD detection using the subsequent gLoG-based localizer

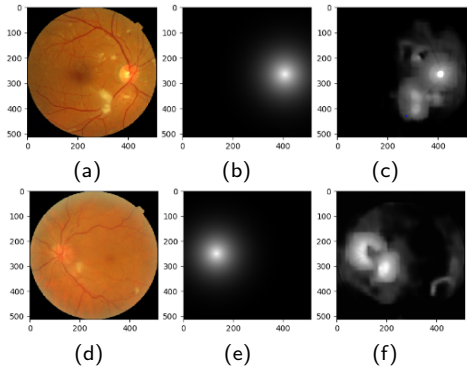


Fig. 8: Motivation for using a Distance-Intensity map, as opposed to a Distance map alone: two representative examples, (a) and (d) show the input RGB images, (b) and (e) show the respective ground-truth Distance map, and (c) and (f) show the predicted Distance map, with many false positives.

(Sec. 2.1.4). Further, the last row shows an example of uneven illumination. The possible candidate regions for OD localization have many false positives. Results with an intensity map alone are worse, and these are not presented here.

The Distance-Intensity map of Sec. 2.1.1 combines distance and intensity information in such a way that bright intensity non-OD candidate regions are suppressed by low distance map values. The two representative examples in Fig. 9 show results with worse examples that those of Fig. 8, in terms of the blur and poor quality of illumination. For each row again, the first image is the input RGB

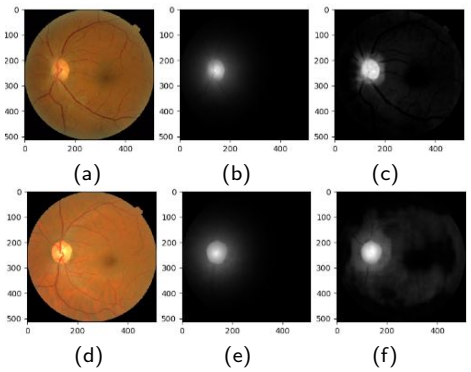


Fig. 9: A Distance-Intensity map (Sec. 2.1.1) shows good OD prediction results even for poor quality blurred and low illumination images. The two representative examples here are much worse than those of Fig. 8 with regard to their blur and low illumination. (a) and (d) show the input images, (b) and (e) the respective ground-truth Distance-Intensity maps, and (c) and (f), the OD prediction with the Distance-Intensity map.

image, the second is the ground truth distance map, and the one to the right

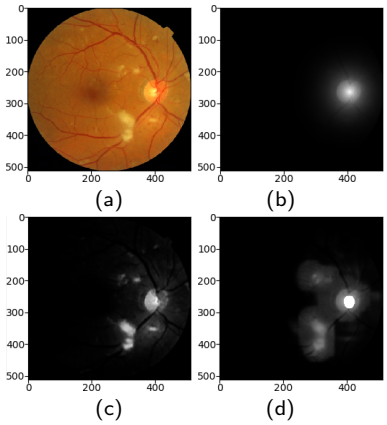


Fig. 10: The use of Tukey's biweight loss function (as opposed to the MSE) suppresses outlier exudate regions significantly, leading to better OD localization: (a) the input, (b) the ground-truthed Distance-Intensity map. (c) the OD prediction with the Distance-Intensity map with an MSE-based loss function, and (d) the OD prediction with the Distance-Intensity map for Tukey's biweight loss function.

is the predicted Distance-Intensity map corresponding to the input image. This alone does not suffice for clinical pathologies such as exudates and other clinical pathologies. Our formulation with Tukey's biweight loss function (in place of a simple MAE/MSE-based one) handles such cases, as the next section shows.

3.3.2 ***Distance-Intensity Map Prediction using Tukey's Biweight Loss Function***

Sec. 2.1.3 describes the use of Tukey's biweight loss function (Belagiannis et al. 2015) (in place of an MSE/MAE-based one). This enables proper training of the U-Net (Sec. 2.1.2) by down-weighting exudates and other clinical pathologies as outliers. Fig. 10 shows a representative example. The first row shows the original input image, and the ground-truthed Distance-Intensity map. The left image in the second row shows results with MSE-based loss function, which highlights exudates. The right image in the second row suppresses the exudates significantly.

Fig. 11 compares the mean pixel error curves of a MSE-based loss function and a loss function based on Tukey's biweight loss function, across epochs. The curves for Tukey's biweight loss function lie well below those for the MSE-based loss, for both the training and validation sets. It shows better generalization as well, as the validation errors lie well below those of the MSE-based case. Apart from better generalization, it also offers advantages in terms of computational complexity. Fig. 12 illustrates the faster convergence with Tukey's biweight loss function within approximately 200 epochs, compared to the MSE case (about 600 epochs). The above two plots are for a representative case for the Messidor dataset (Messidor Retinal Database 2016). We have chosen this particular dataset for the plots in Fig. 11 and Fig. 12 since all state-of-the-art OD detection papers show sample results for this dataset. Later in the paper, we compare our OD

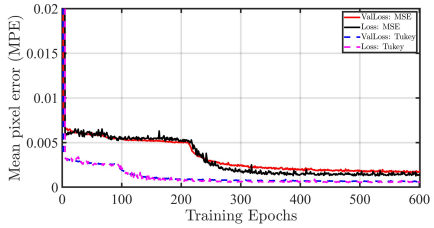


Fig. 11: Tukey's biweight loss curves for training and validation sets ('Loss: Tukey' and 'ValLoss: Tukey', respectively) lies well below the MSE case ('Loss: MSE' and 'ValLoss: Tukey', respectively), and thus leads to better generalization.

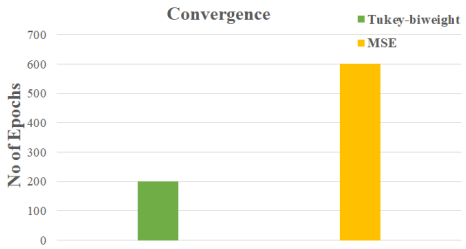


Fig. 12: Convergence rate comparison with MSE and Tukey's biweight loss functions. An MSE-based loss function requires about three times more iterations (epochs) to converge, as compared to one with Tukey's biweight loss function

detection results with the state-of-the-art methods for this particular dataset for the same reason (Table 3).

3.4 **Estimation of Decay Parameters γ and β**

The parameters γ and β in the Distance-Intensity Map (Sec. 2.1.1 and Fig. 4) measure the decay of the normalized Distance-Intensity values with distance. Fig. 13 shows a representative Euclidean Distance Error (EDE)-based empirical selection of the decay parameter γ. For the challenging AIIMS Community Camp dataset, which have been taken with hand-held ophthalmoscopes, often in poor illumination conditions. Our system estimates $\gamma = 7$ and $\beta = 7$ for this challenging dataset: values which give the minimum EDE (Euclidean Distance Error) with an acceptable spread. The EDE accuracy is measured with respect to the manual ground truth marking of the ODs by expert medical personnel.

3.5 **Robust OD Detection for a Wide Variety of Query Images: Blur, Poor Illumination, Pathology Artefacts**

As mentioned before (and illustrated in Fig. 7 for instance), the AIIMS Community Camp Dataset has images taken under poor illumination conditions with hand-held ophthalmoscopes. The challenges in correct OD detection in such cases include

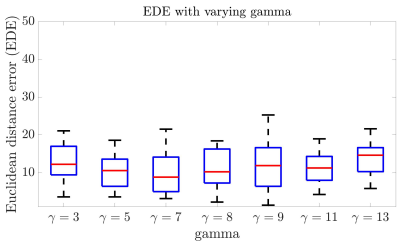


Fig. 13: Selecting Decay Parameter γ: a representative example. The Euclidean Distance Error (EDE) is plotted for each γ value in this example, for the challenging community camp-based AIIMS dataset. The choice of $\gamma = 7$ is empirical, since this gives the least mean EDE for a reasonable spread. $\gamma = 7$ and $\beta = 7$ give the minimum EDE (with an acceptable spread) for this challenging dataset.

non-uniform illumination, blurred retinal images (motion and sensor blur), clinical artefacts and other miscellaneous challenging cases (which could also include a combination of the above factors, and more). Fig. 14 shows representative examples of correct OD estimation in spite of such challenges. The figure shows results of inter-dataset training and testing. The training is with the Kaggle (Kaggle Diabetic Retinopathy Detection Competition 2015) which in itself, is quite a challenging dataset in terms of difficult cases. The testing is on the AIIMS Community Camp dataset. Each row of Fig. 14 shows four representative examples of a particular challenging scenario. Each of the four triads shows (respectively), the original image, the predicted Distance-Intensity map, and the OD detection marked with a small blue cross. The first set (a) shows examples of non-uniform illumination: bright saturation at some places, and darkness at others. The second set (b) shows four representative triads of blurred retinal samples due to motion and sensor blur. The third set (c) shows results of images with clinical artefacts. The fourth set (d) shows miscellaneous challenging cases, which include a combination of one or more of the above cases.

4 Discussion

This section considers the performance of the proposed method on inter- and intra-dataset calculations. It is well-known that the $1R$ criterion is the most common performance evaluation method for OD detection (Al-Bander et al. 2018). Thus, for the first set of experiments, the $1R$ criterion is considered. This represents the number of cases of the OD being within $1R$ of the ground-truthed OD radius, to the total number of cases in the dataset. As mentioned before, the ground truth has been marked by expert medical personnel. Table 2 shows the evaluated test accuracy for each dataset with different train and test combinations. In the second set, we see that the take-home points with the $0.5R$ condition are no different from those with the $1R$ condition. This paper shows the superiority of Tukey's biweight loss function on various counts: lower training and validation error with no overfitting (as shown in Fig. 11), better convergence (as shown in Fig. 12), and as specifically shown in Table 2, much better OD location accuracy over the MSE loss function.

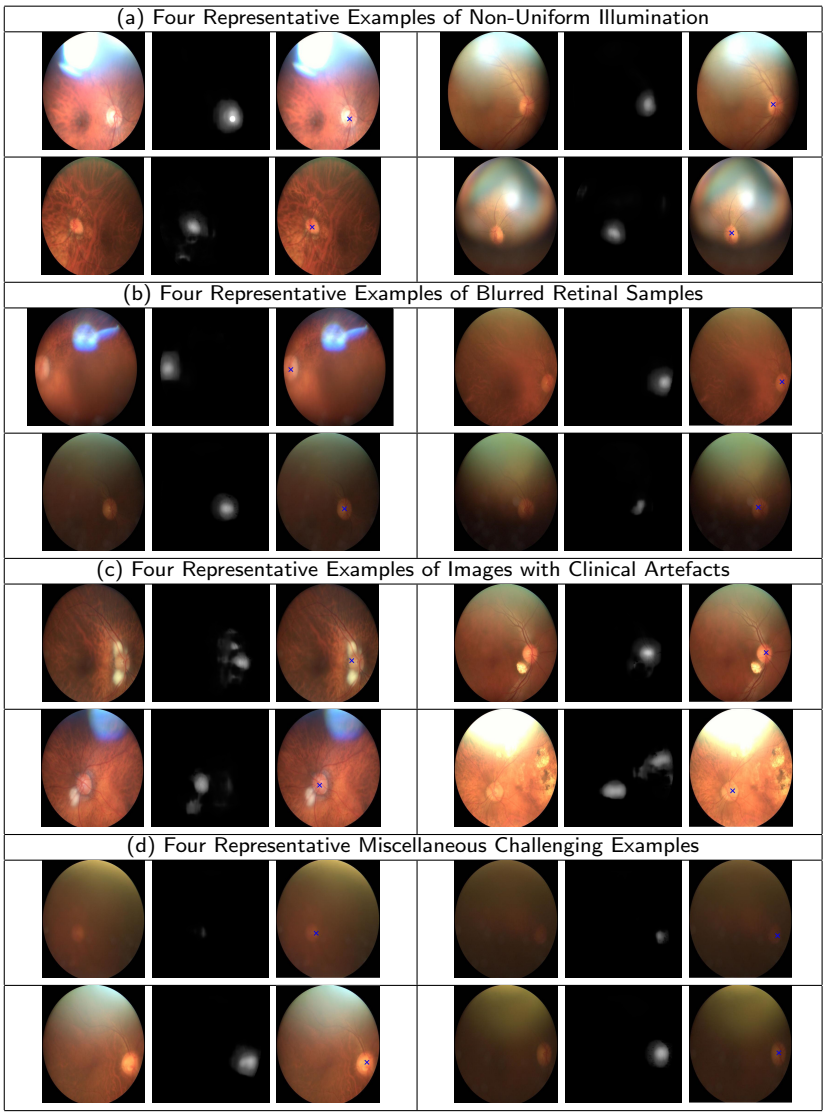


Fig. 14: Representative examples of successful OD estimation in challenging cases, in spite of inter-dataset training and testing (Kaggle and AIIMS, respectively). For each class of challenging conditions, we show four sample representative results. Each triplet of images in a row shows the original image the predicted Distance-Intensity map, and the estimated position of the OD marked with a blue cross, respectively.

The same kind of dataset (i.e., training and testing on the same dataset) expectedly performs well with 99.67%, 95.23% and 96.44% testing accuracy on Messidor (Messidor Retinal Database 2016), Kaggle (Kaggle Diabetic Retinopathy Detection Competition 2015) and the AIIMS Community Camp datasets respectively, using Tukey's biweight loss function. However, the aim of our proposed

Table 2: The evaluated test accuracy for each dataset with different train and test combinations.

S#	Training Dataset	#Train-Validate-Test Images	Testing Dataset	Test Accuracy %, for loss functions: Tukey ($1R$)	Tukey ($0.5R$)	MSE ($1R$)	MSE ($0.5R$)
1	Messidor	910-113-113	Messidor	99.67	99.11	98.23	96.26
2	Kaggle	24270-3034-3034	Kaggle	95.23	86.32	92.88	82.14
3	AIIMS	1672-250-272	AIIMS	96.44	86.36	95.69	81.88
4	Messidor+ AIIMS	2582-363-385	Messidor+ AIIMS	**99.83**	96.29	95.55	88.88
5	Messidor	910-113-333	DRIVE,DRIONS, STARE,Drishti-GS	94.85	92.21	92.43	90.11
6	Kaggle	24270-6068-333	DRIVE,DRIONS, STARE,Drishti-GS	83.1	76.31	85.26	75.15
7	Kaggle	24270-6068-1136	Messidor	93.75	91.02	94.81	89.38
8	Kaggle	24270-6068-2194	AIIMS	**98.83**	88.29	96.66	87.78

work is to evaluate the robustness of algorithm for inter-dataset scenarios (i.e., training and testing on different datasets). The Messidor dataset contains the maximum number of good quality retinal images. Our experimentation had training on the Messidor dataset (1136 images) and combined testing on the DRIVE, DRIONS, STARE and the Drishti-GS dataset (332 images). The obtained average accuracy for the case with Tukey's biweight loss function and MSE is 94.85% and 92.43% respectively. For the more difficult case of retinal images with artefacts, we consider the training on the Kaggle dataset. This has maximum number of artefact retinal images among all the competing datasets (approximately 30,000). In the experiment the network is trained on Kaggle and tested on the combined DRIVE, DRIONS, STARE and Drishti-GS dataset which gives 83.10% accuracy with Tukey's biweight loss function, and 85.26% with an MSE-based loss function. (This is a bit of an aberration, along with the seventh row: a possible reason is the different classes of datasets used for training and testing: the Messidor and DRIVE, DRIONS, STARE and Drishti-GS datasets have clean images, whereas the Kaggle and AIIMS Community Camp dataset have challenging cases of retinal images.)

On the 2194 images of the challenging AIIMS Community Camp dataset, the accuracies obtained are 98.83% and 96.66% with Tukey's biweight loss function and an MSE loss function, respectively. For the Messidor dataset, the accuracy figures are 93.75% and 94.81% for Tukey's biweight loss function and an MSE-based loss function, respectively. One of the aims of this work is to get good inter-dataset performance. As mentioned above, one gets good results when the training and test sets are similar in their nature: the Kaggle and AIIMS Community Camp datasets have challenging cases. The training on the large Kaggle dataset gives good accuracy values with the AIIMS Community Camp dataset. Training on the large (and clean) Messidor dataset gives good accuracies on the other (clean) datasets, the

combined DRIVE, DRIONS, STARE and Drishti-GS datasets. Table 3 compares OD detection accuracies for other competing state-of-the-art systems, all of which benchmark their results on the relatively clean Messidor dataset.

Table 2 also shows the accuracy figures for the $0.5R$ condition as well. As one expects, the relative accuracy figures for the more stringent $0.5R$ case are less than those with the $1R$ case, for each pair of corresponding results. The table shows that the general take-home points for the $0.5R$ case are no different from those of the $1R$ case.

It is worth mentioning here that all OD detection algorithms in Table 3 use different train-test splits, and thus comparing them accurately is a highlighted issue. For instance, (Al-Bander et al. 2018) train their system on the Kaggle dataset, but use the 1200 image Messidor dataset for testing. The systems of (Yu et al. 2012) and (Meyer et al. 2018) perform the training and testing on the whole Messidor dataset alone. On the other hand, the authors in (Huang et al. 2020) report a combination of a local dataset and the iDRID, and Messidor datasets for both training and testing. For this part, both the $1R$ and $0.5R$ criteria are considered

Table 3: A comparison of OD detection accuracies with other state-of-art papers, all of which have benchmarked their results on the relatively clean Messidor dataset under the $1R$ and $0.5R$ conditions.

S#	Paper	Approach	No. of Images	Test Accuracy % ($1R$)	($0.5R$)	Computation Time(sec)
1	(Al-Bander et al. 2018)	Deep Neural Network	1200	97.00	95.00	0.007
2	(Yu et al. 2012)	Template matching technique	1200	98.24	99.08	4.7
3	(Meyer et al. 2018)	U-Net	1136	98.94	97.10	-
4	(Huang et al. 2020)	Region proposal network (CNN)	1200	100	-	-
5	(Liang et al. 2020)	Saliency map of visual features	1200	99.58	-	5
6	(Jana et al. 2021)	CHT with CNN	1200	98.88	-	-
7	Proposed Approach	CNN regression network	1136	**99.67**	**99.11**	3

for OD localization accuracy. As seen in Table 3, our work compares favourably with the state-of-the-art for $1R$ accuracy, and outperforms the other systems, in the $0.5R$ case, with an average accuracy of 99.11%.

There is an interesting point about the use of both $1R$ and the $0.5R$ condition in Table 3 and showing results only for $1R$ in Table 2. It is well-known that the $1R$ criterion is the most common performance evaluation method for OD detection (Al-Bander et al. 2018). In Table 2, our aim is quite different: we wish to compare the performance of our system alone (and not with other systems), with respect to the choice of the loss function (our use of Tukey's biweight loss function, versus the more common MSE loss function). In this paper, we show the superiority of the use of Tukey's biweight loss function on various counts: lower training and validation error with no overfitting (as shown in Fig. 11), better convergence (as shown in Fig. 12), and as specifically shown in Table 2, much better OD location

accuracy over the MSE loss function. We wish to show this for various training and testing set combinations from different datasets, and draw conclusions about the possible suitability of specific types of combinations of training and testing sets. To avoid the take-home points from this specific study from getting missed in the details of $1R$ and $0.5R$ tests, we simply go ahead with the choice of the *de facto* standard $1R$ criterion, and just state that the conclusions and trends are no different for the $0.5R$ case as well.

5 Conclusion

This paper proposes a novel OD detection method which works well in case of difficult retinal images such as non-uniform illumination, blur, and exudates, artefacts and instance of clinical pathologies. The U-Net-based deep network is able to effectively learn and estimate the OD position well, with the aid of Tukey's biweight loss function. Results of extensive experimentation on inter-dataset performance achieve state-of-the-art accuracy figures. Qualitative results of impressive performance are obtained even on the difficult AIIMS Community Camp dataset, in addition to results on the popular Messidor, Kaggle and DRIVE, DRIONS, STARE and Drishti-GS datasets.

6 Declaration

Conflict of Interest

The authors declare no conflict of interest in the publication of this paper.

References

Al-Bander B, Al-Nuaimy W, Williams BM, Zheng Y (2018) Multiscale Sequential Convolutional Neural Networks for Simultaneous Detection of Fovea and Optic Disc. Biomedical Signal Processing and Control 40:91 – 101

Avinash Ramakanth S, Venkatesh Babu R (2014) Approximate Nearest Neighbour Field based Optic Disk Detection. Computerized Medical Imaging and Graphics 38(1):49 – 56

Belagiannis V, Rupprecht C, Carneiro G, Navab N (2015) Robust Optimization for Deep Regression. In: Proc. IEEE International Conference on Computer Vision (ICCV), pp 2830 – 2838

Calimeri F, Marzullo A, Stamile C, Terracina G (2016) Optic Disc Detection Using Fine Tuned Convolutional Neural Networks. In: Proc. International Conference on Signal-Image Technology & Internet-Based Systems, pp 69 – 75

Czarnowski I, Jedrzejowicz P (2018) An Approach to Data Reduction for Learning from Big Datasets: Integrating Stacking, Rotation, and Agent Population Learning Techniques. Complexity 2018:1 – 13

Dada T, Coote M (2010) Clinical Evaluation of Optic Nerve Head. International Society of Glaucoma Surgery

Dietter J, Haq W, Ivanov IV, Norrenberg LA, Völker M, Dynowski M, Röck D, Ziemssen F, Leitritz MA, Ueffing M (2019) Optic Disc Detection in the Presence of Strong Technical Artifacts. Biomedical Signal Processing and Control 53:1 – 11

Foracchia M, Grisan E, Ruggeri A (2004) Detection of Optic Disc in Retinal Images by Means of a Geometrical Model of Vessel Structure. IEEE Transactions on Medical Imaging 23:1189 – 1195

Fu Y, Chen J, Li J, Pan D, Yue X, Zhu Y (2021) Optic Disc Segmentation by u-net and Probability Bubble in Abnormal Fundus Images. Pattern Recognition 117:107971

Huang Y, Zhong Z, Yuan J, Tang X (2020) Efficient and Robust Optic Disc Detection and Fovea Localization using Region Proposal Network and Cascaded Network. Biomedical Signal Processing and Control 60:1 10

Isola P, Zhu JY, Zhou T, Efros AA (2016) Image-to-image Translation with Conditional Adversarial Networks. arXiv:1611.07004 [cs]

Jana S, Parekh R, Sarkar B (2021) A Semi-Supervised Approach for Automatic Detection and Segmentation of Optic Disc from Retinal Fundus Image. In: Nayak J, Naik B, Pelusi D, Das AK (eds) Handbook of Computational Intelligence in Biomedical Engineering and Healthcare, Academic Press, pp 65 – 91

Kaggle Diabetic Retinopathy Detection Competition (2015) https://www.kaggle.com/c/diabetic-retinopathy-detection/data, Accessed: 20 July, 2015

Karras T, Laine S, Aila T (2018) A Style-based Generator Architecture for Generative Adversarial Networks. arXiv:1812.04948 [cs]

Kong H, Akakin HC, Sarma SE (2013) A Generalized Laplacian of Gaussian Filter for Blob Detection and its Applications. IEEE Transactions on Cybernetics 43(6):1719 – 1733

Lalonde M, Beaulieu M, Gagnon L (2001) Fast and Robust Optic Disc Detection using Pyramidal Decomposition and Hausdorff-based Template Matching. IEEE Transactions on Medical Imaging 20(11):1193 – 1200

Latif J, Tu S, Xiao C, Rehman SU, Imran A, Latif Y (2022) Odgnet: A Deep Learning Model for Automated Optic Disc Localization and Glaucoma Classification using Fundus Images. SN Applied Sciences 4:98

Liang M, Zhang Y, Wang H, Li J (2020) Location of Optic Disk in the Fundus Image based on Visual Attention. In: Proc. International Conference on Computer Information and Big Data Applications (CIBDA), pp 446 – 449

Lu S (2011) Accurate and Efficient Optic Disc Detection and Segmentation by a Circular Transformation. IEEE Transactions on Medical Imaging 30:2126 – 2133

Maninis KK, Pont-Tuset J (2010) Retinal Databases. http://www.vision.ee.ethz.ch/ cvlsegmentation/driu/downloads.html

Mendonça AM, Sousa A, Mendonça L, Campilho A (2013) Automatic Localization of the Optic Disc by Combining Vascular and Intensity Information. Computerized Medical Imaging and Graphics 37(5-6):409 – 417

Messidor Retinal Database (2016) http://www.adcis.net/en/third-party/messidor/, Accessed: 24 September, 2016

Meyer MI, Galdran A, Medonça AM, Campilho A (2018) A Pixel-Wise Distance Regression Approach for Joint Retinal Optical Disc and Fovea Detection. In: Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp 39 – 47

Mitra A, Banerjee PS, Roy S, Roy S, Setua SK (2018) The Region of Interest Localization for Glaucoma Analysis from Retinal Fundus Image using Deep Learning. Computer Methods and Programs in Biomedicine 165:25 – 35

Nawaz M, Nazir T, Javed A, Tariq U, Yong HS, Khan MA, Cha J (2022) An Efficient Deep Learning Approach to Automatic Glaucoma Detection using Optic Disc and Optic Cup Localization. Sensors 22(2):434

Nergiz M, Akın M, Yıldız A, Ömer Takeş (2018) Automated Fuzzy Optic Disc Detection Algorithm using Branching of Vessels and Color Properties in Fundus Images. Biocybernetics and Biomedical Engineering 38(4):850 – 867

Niemeijer M, Abràmoff MD, van Ginneken B (2009) Fast Detection of the Optic Disc and Fovea in Color Fundus Photographs. Medical Image Analysis 13(6):859 – 870

Niu D, Xu P, Wan C, Cheng J, Liu J (2017) Automatic Localization of Optic Disc based on Deep Learning in Fundus Images. In: Proc. International Conference on Signal and Image Processing, pp 208 – 212

Panda R, Puhan NB, Panda G (2017) Robust and Accurate Optic Disk Localization using Vessel Symmetry Line Measure in Fundus Images. Biocybernetics and Biomedical Engineering 37(3):466 – 476

Pathan S, Kumar P, Pai R, Bhandary S (2019) Automated Detection of Optic Disc Contours in Fundus Images using Decision Tree Classifier. Biocybernetics and Biomedical Engineering 40:1 – 13

Pathan S, Kumar P, Pai RM, Bhandary SV (2021) Automated Segmentation and Classifcation of Retinal Features for Glaucoma Diagnosis. Biomedical Signal Processing and Control 63:1 –19

Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), vol 9351, pp 234 – 241

Serte S, Serener A (2021) Graph-based Saliency and Ensembles of Convolutional Neural Networks for Glaucoma Detection. IET Image Processing 15:1 – 8

Sharma A, Agrawal M, Lall B (2017) Optic Disc Detection Using Vessel Characteristics and Disc Features. In: Proc. National Conference on Communications (NCC), pp 1 – 6

Shukla AK, Pandey RK, Pachori RB (2020) A Fractional Filter based Efficient Algorithm for Retinal Blood Vessel Segmentation. Biomedical Signal Processing and Control 59:1 – 16

Sivaswamy J, Krishnadas SR, Joshi GD, Jain M, Ujjwal, Syed Tabish A (2014) Drishti-GS: Retinal Image Dataset for Optic Nerve Head (ONH) Segmentation. In: Proc. IEEE International Symposium on Biomedical Imaging (ISBI)

Soares I, Castelo-Branco M, Pinheiro AM (2016) Optic Disc Localization in Retinal Images Based on Cumulative Sum Fields. IEEE Journal of Biomedical and Health Informatics 20:574 – 85

Subha V, Rayen S (2022) A Novel Visual Saliency Thresholding Method for Accurate ROI Segmentation in Optic Disc from FEIs for Glaucoma Detection. International Journal of Computing and Digital Systems 11:477 – 489

Trucco E, MacGillivray T, Xu Y (2019) Computational Retinal Image Analysis: Tools, Applications and Perspectives. Mara Conner

Tulsani A, Kumar P, Pathan S (2021) Automated Segmentation of Optic Disc and Optic Cup for Glaucoma Assessment using Improved unet++ Architecture. Biocybernetics and Biomedical Engineering 41(2):819 – 832

Wankhede PR, Khanchandani KB (2016) Optic Disc Detection using Template Matching based on Color Plane Histograms. In: Proc. International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), pp 278 – 281

Wei Z, Hao W, Chengdong W, Xiaosheng Y, Yugen Y (2018) Automatic Optic Disc Detection in Color Retinal Images by Local Feature Spectrum Analysis

Xiong L, Li H (2016) An Approach to Locate Optic Disc in Retinal Images with Pathological Changes. Computerized Medical Imaging and Graphics 47:40 – 50

Yu H, Barriga ES, Agurto C, Echegaray S, Pattichis MS, Bauman W, Soliz P (2012) Fast Localization and Segmentation of Optic Disk in Retinal Images Using Directional Matched Filtering and Level Sets. IEEE Transactions on Information Technology in Biomedicine 16(4):644 – 657