Received: November 21, 2022. Revised: January 24, 2023. 332

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

# Optic Disc Localization in Retinal Fundus Images Based on You Only Look Once Network (YOLO)

**Hussein Mahdi Ali**[1] **Nidhal K. El Abbadi**[2]*

[1]*Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq*
[2]*Computer Techniques Engineering, Al-Mustaqbal College, Babylon, Iraq*
* Corresponding author's Email: nidhal.abass@fulbrightmail.org

**Abstract:** The retina is the visible component of the nervous system that links directly to the brain; hence, analyzing and extracting features of the retinal fundus image is essential in ophthalmology. Unfortunately, most retinal fundus images suffer from degradation or distortions, resulting from different imaging qualities during capture and infection by a disease of the eye, such as Glaucoma, Retinoblastoma, Diabetic Retinopathy (DR), Myopia, and Macular Edema. As a result, these changes make it difficult to identify regions of the retina that assist in diagnosing the disease. For instance, many retina diseases diagnosis is impossible to detect without detecting the optic disc (OD) correctly. All these issues motivate us to suggest an algorithm for the localization of the OD that is important for assisting in diagnosing some eye diseases. The proposed algorithm includes several stages. Firstly, we focus on gathering seventeen datasets that make the suggested algorithm more popular, where a group of images belongs to four datasets used in training and the rest in testing. Secondly, the quality of the image improves in three steps, crops the Field of View (FOV), which carries essential information in the retinal image, removing pixels at the top and bottom of the FOV because they represent redundant information, and we suggested a new method for enhancing the illumination and contrast of the retinal fundus image. Thirdly, we proposed using a You only look once (YOLO) for the localization of OD. The method locates the OD exactly (drawing box around the OD and each side of the box touches the OD), not the region around it as in the previous methods. The current method works excellently with all the challenges that may face in the detection and localization of the OD. The proposed model tested with about 6000 different images from various datasets and challenges The accuracy obtained was 100% with an average time of 0.19 seconds. These results were very good compared with previous works in terms of accuracy, average time, and the number of tested images.

**Keywords:** Fovea, Medical image processing, Object detection, Optic disc, Optic nerve head, YOLO.

## 1. Introduction

In recent years the pathologies are proven to relate to retinal features. This leads many researchers to focus on the importance of auto-retinal image analysis. One important feature that has significant information on the diagnosis of many diseases such as Diabetes, Glaucoma, Hypertension, and others is the optic disc (OD) or the optic nerve head (ONH) [1].

In ophthalmology, the most common way to examine the human eye is to take an eye-fundus photograph and analyze it. During this eye examination, a medical expert acquires a photo of the eye background through the pupil with a fundus camera. Note that the analysis of these images is commonly done by visual inspection. However, this process can require hours in front of a computer screen, in particular in the case of medical screening. Apart from that, the percentage of diagnosis errors is high due to the difficulty in determining all changes in the eye, specifically in the early diagnosis [2]. Furthermore, it is very hard to exploit efficiently the enormous amount of retinal data that contains a huge amount of clinical knowledge obtained from fundus images [3]. An example of a fundus image and the main important structures are shown in Fig. 1.

OD is the origin of the optic nerves and the entry and leave (point) of the main blood vessels that

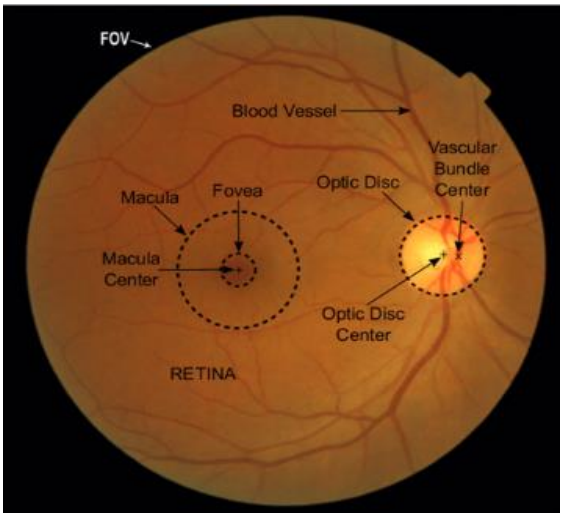

Figure. 1 The important structures in the fundus image

supply the retina. Usually, the OD looks like a bright region with orange or yellowish colors and has a circular or oval shape. The size of the OD is about one-sixth of the retina diameter of the retinal image (changes depending on the camera view) [4].

Many benefits can be achieved from OD. Some of them are: (1) Since the OD considers the entry point of blood vessels, it can be used as the starting point to track vessels. (2) OD features can significantly prevent some blinding eye diseases (such as Glaucoma and Cataracts). (3) Measuring the cup-to-disc ratio is important in evaluating Glaucoma. (4) It is used in image processing as a landmark to detect important regions, such as the macular region and fovea, in addition to retina abnormalities. (5) It can be used to detect lesions (due to equalization of luminance) like hemangioma and hemorrhage [5].

The scarcity of the fundus images dataset that can be used for automatic detection and segmentation of the OD features and for diagnosing retinal diseases has been a bottleneck for the successful application of computer-aided. Furthermore, the few datasets available for researchers suffered from the artifacts presented by cameras (such as noise, blur, contrast, and illumination), which are regarded as challenging [6].

However, most retinal fundus images suffer from degradation and distortions for two reasons, external which are associated with the environment such as the lighting variation, and camera efficiency, and internal reasons due to retina diseases, such as blood bleeding, the emergence of spots or lesions, swelling and erosion of some blood vessels, expansion of the parts of the optic nerve, etc.

The main purpose of this paper is to propose a method that can localize the optic disc in any retinal fundus image. The advantage of this proposal over the previous works is the ability to successfully localize the optic disc, regardless of how distorted the image is or whether the eye is healthy or injured, faster than the currents methods, not limited to a specific dataset, and bounded Precise the OD, not the region contained the OD as previous works.

The rest of the paper is organized as follows: section two focuses on the related works, while section three introduces a quick review of the YOLO. The proposed algorithm is presented in section four and the results of this algorithm introduce in section five. Finally, section six concludes the methodology and results.

## 2. Related works

The problem of disc localization has become inevitable and important since a correct disc location is crucial for identifying eye diseases. For this reason, in recent years, many and varied methods have been developed for this purpose. Some of them are present in this section.

Luangruangrong and Chinnasarn (2019) [7] proposed new hybrid algorithms for automatic disc localization. A modified smoothed gradient edge generation and a Hough transform are combined for this purpose. A voting method is used for combined all candidate results to locate the final location. This proposal concerns some challenges such as unbalanced shade in an image, image low contrast, unclear OD boundaries, and various bright lesions are found. the accuracy of this method was determined with seven various datasets and it was ranging between (91-100%).

Athab and Selman (2020) [8] suggested a method relying on the value of high density and affinity for blood vessels. The preprocessing step resized, normalized, and enhanced the contrast of the Green channel. OD regions are determined by using threshold and morphology operations. Then, it segmented the blood vessels for the initial OD regions using Gaussian and canny filters. Finally, it determined the OD by choosing the region with the maximum variance. The method achieved an average accuracy rate of 98.87% with an average speed of 0.5 seconds for six datasets. Other than that, it failed in low-contrast images.

Liang et al. (2020) [5] presented a method using visual attention. Firstly, it extracted fundamental visual characteristics from the image and generated a saliency map using region covariance. Secondly, it added a boundary bias based on the spatial characteristics of the OD and identified the OD using a saliency map that has been refined using morphological techniques. The method tested two datasets with an average accuracy rate of 99.79%

and an average speed of 4.75 seconds. However, the method was somewhat slow and failed in images with enlarged deformation of the lesion OD.

Khaing et al. (2022) [4] suggested a method for localizing OD based on blood vessels. It extracted the two main blood vessels by threshold and then found candidates for OD either by the Exclusion Method (EM) if vessels fit a parabola or the line detection method (LEM). Finally, the decision tree based on features determined the OD. The EM merges adjacent regions to form the horizontal regions of interest (HROI). The LEM extracted blood vessels using the line Hough transform and determined multiple OD candidates by a window that finds the highest average intensity value. Note that the method achieved an average accuracy rate of 94.673% with an average speed of 5 seconds for three available standard datasets. However, the method was slow and failed to localize the OD in blurred images.

## 3. You only look once (YOLO) model: a quick overview

You only look once (YOLO) is a state-of-the-art family dealing with object detection, originated by Joseph Redmon in 2016. It handles object detection as a classification and regression problem. YOLO one-stage object detection relied on deep learning [9].

YOLO model separates the input image into square blocks (S × S) as a grid of cells, each cell grid accountable for detecting the object presented. Note that each cell grid predicts bounding boxes (Bb) with an objectness score for each Bb and one set of conditional class probabilities (Cp). The objectness score is an indicator that shows the model's confidence that the box contains an object. Other than that, the YOLO algorithm also finds the class-specific confidence scores (Conf) for each box using Cp and the objectness score. Conf represents a class probability of appearing in the box and measures how closely the predicted box fits the object. For example, each Bb has five predicted attributes: the center of the Bb (x, y), width, height, and confidence score (Conf.) [9].

The architecture of YOLOv5s has three main components: Backbone, Neck, and Head. The architecture is shown in Fig. 2 [10,11].

● **Backbone** is utilized for extraction features a map of the image; it includes the focus layer, cross stage partial Darknet53 network (CSPDarknet53), and spatial pyramid pooling (SPP) layer.

● **Neck** is utilized for features aggregation by combining and mixing image features. It included the path aggregation network (PANet) and features pyramid network (FPN).

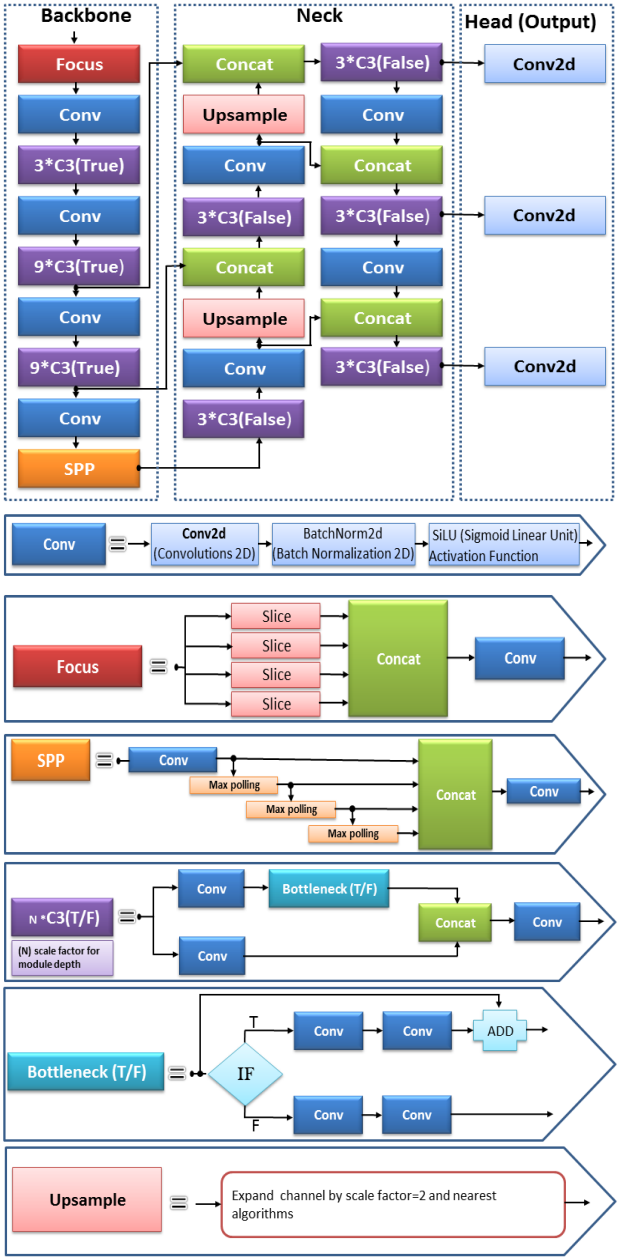


Figure. 2 YOLOv5s Architecture

● **Head** is utilized for the prediction of the boxes and classes. It depends on the anchor-based YOLO algorithm.

## 4. Proposed method

This paper aims to accurately create an effective model that can locate the optic nerve head (ONH) or optic disc (OD) in the retinal fundus images. The proposed method is based on the YOLOv5s model. Fig. 3 shows the block diagram of the proposed method. The proposed method consists of five main stages as follows:

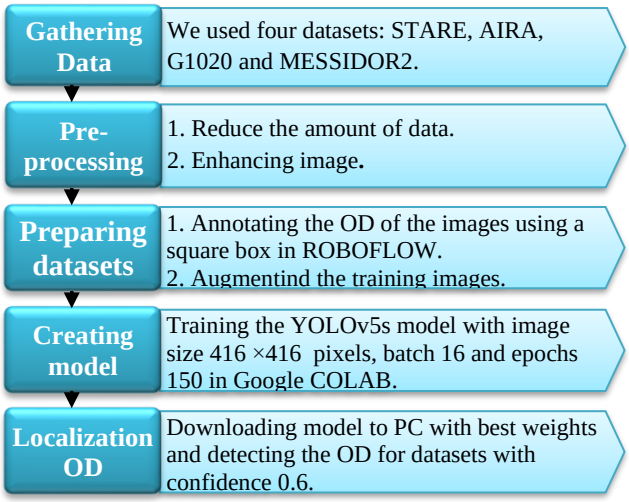


Figure. 3 Block diagram of the proposed method

Table 1. Datasets were collected for localizing the optic disc

| Datasets Name | Ref. | Total no. of images | No. of images used for | | |
|---|---|---|---|---|---|
| | | | Training | Validation | Testing |
| G1020 | [6] | 1020 | 612 | 306 | 102 |
| STARE | [12] | 329 | 198 | 98 | 33 |
| MESSIDOR-2 | [13] | 1748 | 1045 | 524 | 179 |
| ARIA | [14] | 143 | 86 | 42 | 15 |
| HRF | [2] | 45 | - | - | 45 |
| ORIGA | [3] | 650 | - | - | 650 |
| DRIVE | [15] | 40 | - | - | 40 |
| CHASE_DB1 | [16] | 28 | - | - | 28 |
| MESSIDOR-1 | [17] | 1200 | - | - | 1200 |
| DIARETDB0 | [18] | 130 | - | - | 130 |
| DIARETDB1 | [19] | 89 | - | - | 89 |
| ROC | [20] | 100 | | | 100 |
| DRIONS | [21] | 110 | - | - | 110 |
| RIGA | [22] | 750 | - | - | 750 |
| DRISHTI-GS1 | [23] | 101 | - | - | 101 |
| REFUGE-1 | [24] | 1200 | - | - | 1200 |
| REFUGE-2 | [25] | 1600 | - | - | 1600 |
| Total | | 9283 | 1941 | 970 | 6372 |

### 4.1 Gathering data

Generally, an object detection model needs a lot of images for training. Therefore, the authors gather seventeen well-known datasets, some of them were used for training while the others were used for testing as shown in Table 1.

### 4.2 Preprocessing image

Preprocessing is a common step in most computer vision applications and image processing. The preprocessing step in the current paper includes de-noise, de-blur, and de-haze. Hence, preprocessing the image is done by implementing two processes:

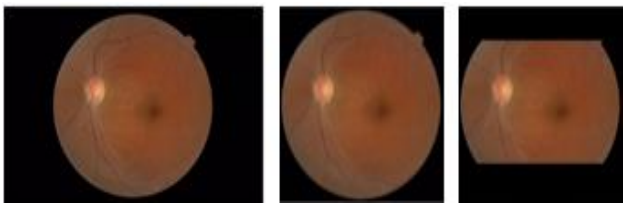

Figure. 4 Resize an image: (from left to right) input image, field of view, and resized image (RGB image)

#### 4.2.1. Reduce the amount of data

The first process aims to remove unimportant areas and focus on the parts of the image we need to find the disc. The results of reducing the amount of data are illustrated in Fig. 4, and the steps of this process are outlined below:

- In this work we focus on the region that carries significant information in the retina, which is the largest circular part of the image and contains the retina data called the Field of View (FOV) (Fig. 1). The data outside FOV is considered redundant and not useful in the diagnosis, so cutting this area.
- Scale the image size to 416x416. Almost the disc is near the center of the FOV height, so part of the top and bottom of the FOV is unimportant, we proposed to change the color of a stripe with size 20x416 pixels from the top and bottom of the image into the black color.

#### 4.2.2. Enhancing image

The second process enhances the image. It is one of the important contributions to this study. As mentioned previously, most images suffer from degradation. Therefore, the contrast in an image is poor, and parts of the image are not visible. Nevertheless, the process has proven its effectiveness in enhancing an image. The results of enhancing the image are illustrated in Fig. 5, and the steps of this process are outlined below:

**First**: Remove the noise by using the Gaussian filter (a low-pass filter) using Eq. (1):

$$G = \frac{1}{2\pi s^2} \exp\left[\frac{-(x^2+y^2)}{2s^2}\right] \quad (1)$$

Where *G* is the Gaussian smoothing filter and *s* is the standard deviation which is selected to be 13.

Each channel of the input RGB image is smoothed separately. And later concatenate the three channels.

**Third**: Calculate the light parameter in the image for each channel of the RGB image, based on mean and standard deviation. Eq. (2) is suggested by the authors for this purpose.

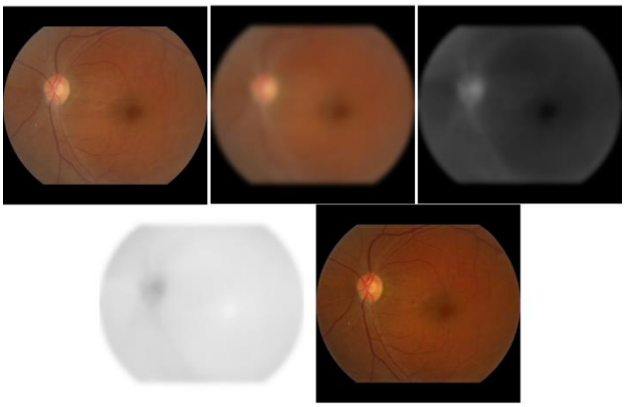

Figure. 5 Enhancing image: (from left to right and from top to bottom) the RGB input image, the smoothed RGB image, MC image, Depth, and enhanced image

$$\underset{c\in\{R,G,B\}}{\mathrm{SPLight}_C} = (M_c)^{std_c} \quad (2)$$

Where $\mathrm{SPLight}_C$ is a single value for each color channel c where $c \in \{R, G, B\}$, M is the mean of the non-zero value in channel c, and Std is the standard deviation of the non-zero value in channel c.

**Fourth**: In this step, we apply three operations:

A. Normalizing each channel of the smoothed RGB image by using Eq.(3) [26]:

$$N_C = \frac{\text{smoothed image}_c}{\mathrm{SPLight}_c} \quad \text{where } c \in \{R, G, B\} \quad (3)$$

Where Nc is the normalized RGB image channel.

B. Construct the new image (MC) from the normalized image Nc by selecting the minimum value for each pixel among the three values of (R, G, and B).

C. Calculating the depth using Eq. (4) [26]:

$$\mathrm{Depth}(x, y) = 1 - 0.63 \times MC \quad (4)$$

**Five**: Enhancing each channel of the RGB input image by applying Eq. (5) [26]:

$$EI_c = \frac{RGB_c - \mathrm{SPLight}_c}{\max(\text{depth}, \text{epsilon})} + \mathrm{SPLight}_c \quad (5)$$

Where $c \in \{R, G, B\}$, EI is the enhanced image, and epsilon is a very small value to prevent dividing by zero.

### 4.3 Preparing datasets

There are seventeen datasets used in this paper, four of them selected for training as mentioned previously. The four datasets were divided as 90% for training (60% training and 30% for validation), while 10% of these four datasets in addition to the thirteen datasets used for testing. The OD is annotated and bounded by a square box in training images. The number of training images after augmentation was 5823 images.

### 4.4 Creating model

In the current paper, we proposed to use the YOLOv5s model for detection and localization OD. YOLOv5s model with sixteen batches and 150 epochs trained on the described datasets in Table 1 from scratch. YOLOv5s model consists of three components: the backbone, the neck network, and the head, as shown in Fig. 2.

The backbone convolutional network extracts deep features from the input image. The focus module in the backbone is the first layer. The focus module is created specifically to speed up training and decrease model calculation. The input image with a size of 416×416×3 pixels is split into four slices and concatenated in the focus module, and the following layer received the output of the results to better extract the deep features.

The second component in the YOLOv5s model is the neck network. The neck is a sequence of feature aggregation layers of mixing and combining features of the image that are mostly used to produce feature pyramid networks. The feature pyramids in the neck help model generalizing and facilitates the detection of the same object in different sizes.

The final component in the model is the head (detection network). The head applies anchor boxes to the feature map produced from the preceding layer. The head produces a vector with the category probability of the target object (in our model disc) and bounding boxes surrounding the objects. The head is responsible for implementing the traditional concept of YOLO, which involves dividing image features into grids, and searching for objects. The output is an image containing a square box bounding the OD, with a notation indicating the probability value that this part is OD.

## 5. Experimental results

The performance of the YOLOv5s model is assessed using a variety of metrics, including recall, precision, and Mean Average Precision (mAP) when Intersection Over Union (IOU) is at 50% and 0.95%. After training the model, the evaluation of the model achieved a recall score of 0.996, a precision score of 0.996, a mAP@.5 score of 0.996, and a mAP@.95 score of 0.826. These results confirm the efficiency of the proposed method for accurately detecting the optic disc (OD).

Table 2. The results of detecting optic disc

| Datasets | Number of images | Accuracy |
|---|---|---|
| DRIVE | 40 | 100 |
| MESSIDOR-1 | 1200 | 100 |
| DIARETDB0 | 130 | 100 |
| DIARETDB1 | 89 | 100 |
| HRF | 45 | 100 |
| DRISHTI | 101 | 100 |
| DRIONS | 110 | 100 |
| REFUGE-1 | 1200 | 100 |
| REFUGE-2 | 1600 | 100 |
| CHASE_DB1 | 28 | 100 |
| ROC | 100 | 100 |
| RIGA | 750 | 100 |
| ORIGA | 650 | 100 |
| G1020 | 102 | 100 |
| STARE | 33 | 100 |
| MESSIDOR-2 | 179 | 100 |
| ARIA | 15 | 100 |

Table 3. Comparison of optic disc localization methods

| Datasets | Accuracy % for several methods | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Ref. [7] | Ref. [29] | Ref. [30] | Ref. [5] | Ref. [8] | Ref. [4] | Ref. [31] | Ref. [32] | Our |
| STARE | | | 97.78 | | | 90.12 | | 90 | 100 |
| DRIVE | | 100 | 100 | 100 | 100 | | 55 | 100 | 100 |
| MESSIDOR-1 | 98.58 | | 97.78 | 99.58 | 99.33 | | 89 | | 100 |
| DIARETDB0 | 99.23 | 96.92 | 99.49 | | 96.15 | 96.15 | | | 100 |
| DIARETDB1 | 100 | 98.8 | 99.49 | | 97.75 | 97.75 | | | 100 |
| HRF | 100 | | 100 | | 100 | | | | 100 |
| ARIA | 91.61 | | | | | | | | 100 |
| ROC | 99 | | 100 | | | | | | 100 |
| DRIONS | | | | | 100 | | | | 100 |
| Average Time(s) | 26.01 | | | 4.75 | 0.5 | 5 | | | 0.19 |

The model tests with 6372 images from seventeen well-known datasets (have been described in Table 1). The results of testing the proposed method are shown in Table 2. According to the results in Table 2, it can be seen that the model can determine the OD at 100% accuracy for all datasets.

The accuracy of the proposed method is compared with the accuracy of some of the previous methods, with a comparison of the time required to determine the OD for each method, as shown in Table 3.

The results show that the proposed method determines the OD with 100% accuracy for all datasets and an average speed of 0.19 seconds. It gives the best results compared to the other methods.

From Table 3 we conclude that all the previous method has various accuracy for each type of dataset, while the proposed method has 100% accuracy for all datasets.


Figure. 6 Comparison for crop OD: first column crop OD by the proposed method, second column crop OD by other methods [29]

The approaches presented in Table 3 can be categorized into five main classes: approaches based on OD features such as brightness and circularity (article [29]), approaches that relied on just vascular extraction (article [4]) or combined with OD characteristics (article [8]), approaches that employed some transformation (articles [7, 27]) or analysis image features for extraction OD (article [5]) and companies with first class, approaches that utilized the template matching technique for OD (article [28]), and approaches that utilized the deep learning for localization OD (article [30]).

Most papers localized the center of OD and then cropped the region around it. While in the proposed method crop just the OD with a minimum area around (just the area in the corner of the box). This is one of the strong points of this proposal. Fig. 6 shows a sample of the OD detection and cropping by [29] for the image in the Messidor1 dataset, compared with the proposed method.

From Fig. 6, we can notice that cropping OD by the proposed method is more accurate because there is a smaller area around the OD compared with another method.

Many of the previous methods failed to localize the OD because the OD is not always with the highest intensity in the image. Fig. 7 shows the comparison between previous methods that detect and localize OD from some challenge images, almost previous methods failed to detect and localize the OD, opposite of the proposed method in the same images.

Also, the other strong point of this proposal is the ability to detect and localize the OD even when the retina is infected and there are some lesions such as hemorrhage, and exudates. Most previous methods failed to detect and localize OD with a lesion, while

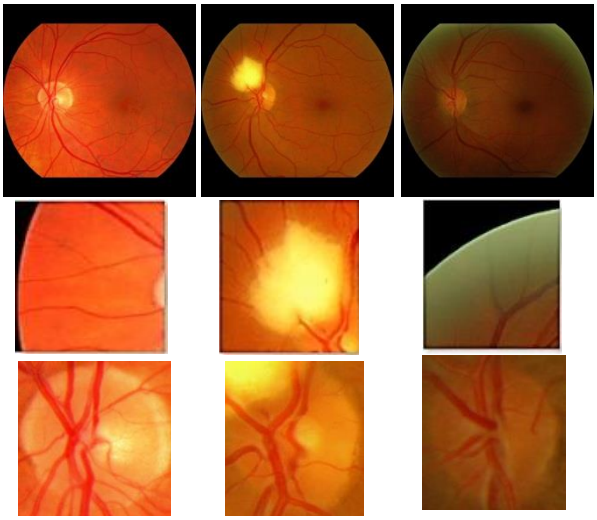

Figure. 7 Comparison with Triwijoyo [29]: first-row original images from the MESSIDOR-1 dataset, second-row crop OD by Triwijoyo, and third-row crop OD by the proposed method

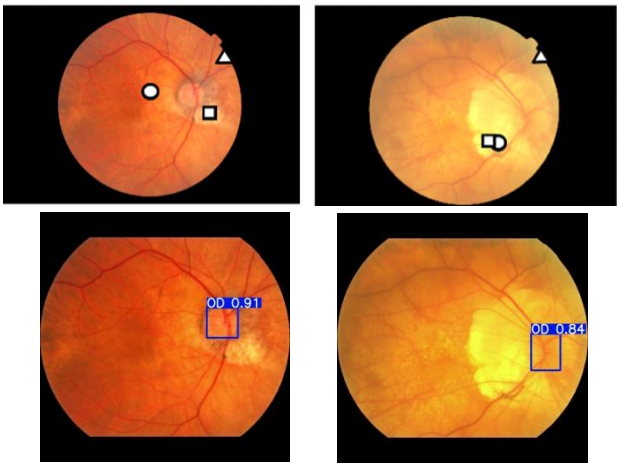


Figure. 8 Comparison of detecting and localizing OD when the retina is infected. First-row, failed OD detection images by Luangruangrong [7], and second-row, successful OD detection OD by the proposed method for the same image

the proposed method successfully detects and localizes the OD as shown in Fig. 8.

It is worth noting that the image in the second column of Fig. 8, is one of the difficult images when detecting the OD. Among the methods that failed to detect the OD is Liang [5] (technical details are mentioned in section 2 of the related work), and another image that handles almost the same characteristics is shown in Fig. 9.

As mentioned earlier, the retinal fundus images often suffer from degradation and distortion (especially as the disease progresses). The distortions or lesions make the regions of the fundus image abnormal. Therefore, automated detection of any part in a fundus image with an abnormal region is

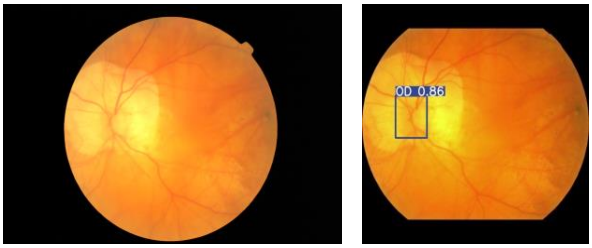


Figure. 9 Comparison with Liang: first-column, failed OD detection image by Liang and second-column, successful OD detection images by the proposed method for the same mage

commonly a challenging study. Due to this, a single image probably contains various types of retinal lesions with varying colors, shapes, sizes, locations, and textures.

Moreover, the kind and number of retinal lesions are usually not known in advance. Thus, the images often get different qualities during capturing.

These distortions are challenges for the automated detection of the OD correctly. We briefly highlight some challenges that any method often faces during OD localization, as well as the results of the proposed method for localizing the OD with challenges.

**The first challenge**: The most difficult challenge facing analyzing the retina is the presence of spots or lesions. These pests have different shapes, sizes, and locations, according to the disease and its level. Hemorrhage, exudates, and Retinal Pigment Epithelium (RPE) are examples of these lesions. The proposed method detects all images with these challenges successfully.

The results of the proposed method for localizing the OD with challenging hemorrhage, exudates, and RPE lesions are shown in Fig. 10, Fig. 11, and Fig. 12, respectively.

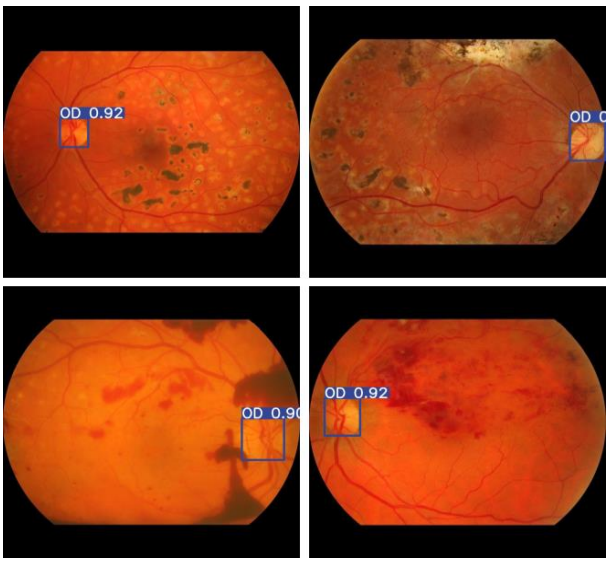


Figure. 10 OD localization with hemorrhage challenge

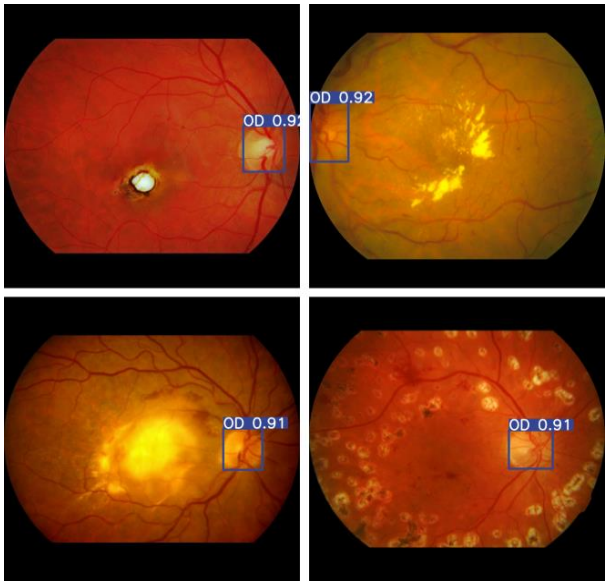


Figure. 11 OD localization with exudates challenge

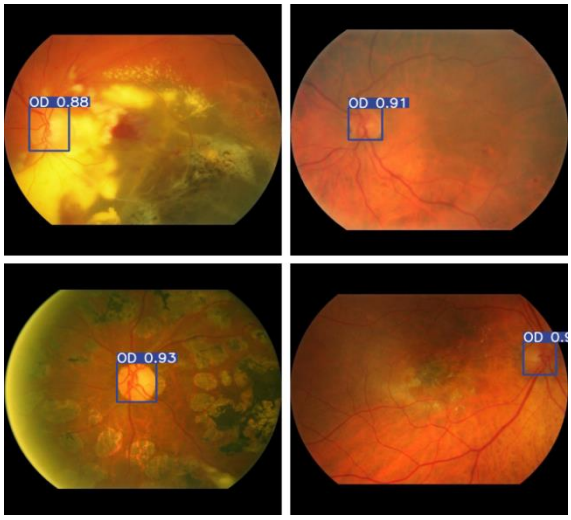


Figure. 12 OD localization with retinal pigment epithelium (RPE) challenge

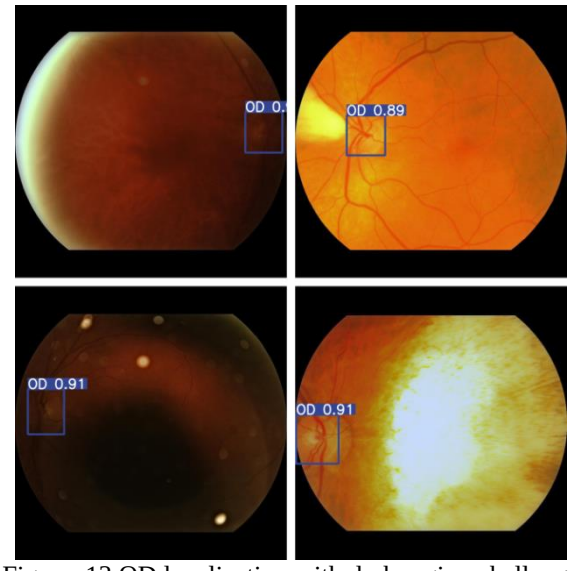


Figure. 13 OD localization with dark region challenge

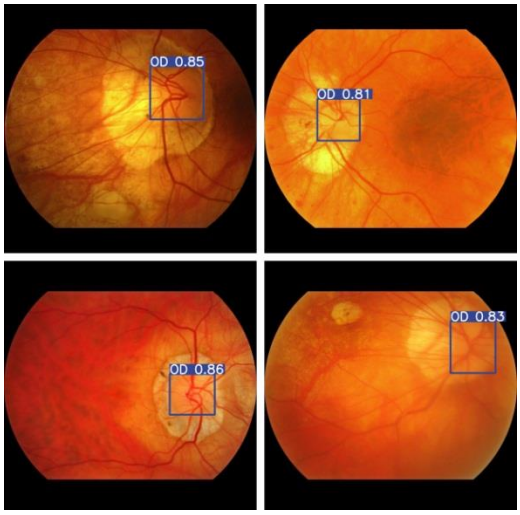


Figure. 14 OD localization with peripapillary atrophy challenge

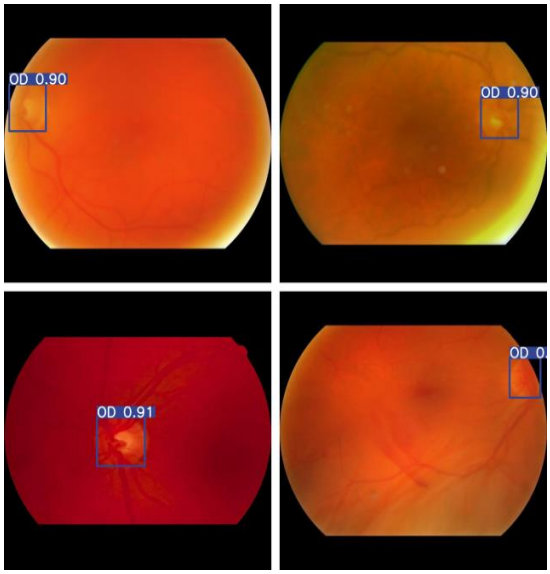


Figure. 15 OD localization with blur challenge

**The second challenge**: Most methods depend on the well-known principle that the OD is the brightest region in the image. Fig. 13 shows that the OD does not contain the highest values in the image and may even be in a dark region. Therefore, the previous methods of determining the OD do not work correctly with this type of image, while, the proposed method has determined the OD correctly.

**The third challenge**: Some retinal images have a region surrounding the OD called peripapillary atrophy (PPA). This region prevents determining the OD correctly because it has a wide boundary surrounding the OD and interfering with it. Subsequently, the OD boundary decays within this region. As a result, many methods fail to determine OD because it is difficult to distinguish between the boundary of OD and PPA. The results of the proposed

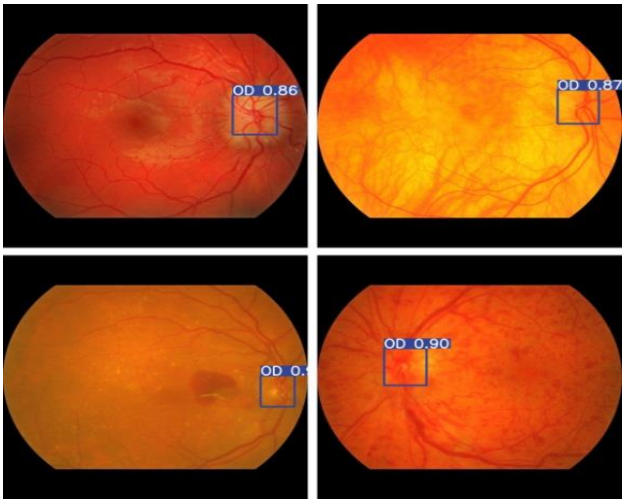


Figure. 16 OD localization with invisible edges challenge

method for this type of image are shown in Fig. 14, where it detected all the OD of images.

**The fourth challenge:** It is a general challenge in almost all computer vision and image processing applications that concerns image quality. The challenge arises when an image is blurred or has poor contrast. Methods that depend on segmentation of the OD or blood vessels, or both, fail in this type of image. The reason is that the boundaries of the regions are unclear, making it difficult to segment, consequently failing to identify the OD. Fig. 15 shows that the proposed method detected OD in images of this type.

**The fifth challenge**: This challenge resembles the blur and PPA challenges. The image in this challenge is generally clear, but the OD edges are invisible. Therefore, it is difficult to determine the OD, especially if the method depends on boundary detection. The results of the proposed method for this type of image are shown in Fig. 16, which detects and localized the OD.

## 6. Conclusion

In this paper, a new automated method for fast optic disc (OD) localization has been proposed. The main idea is to localize the OD using the Yolov5 model for object detection.

The main contribution of this paper is to use the object detection principle as a new principle in OD localization methods. The accuracy of detecting and localizing the OD from 13 different datasets was 100%, with relatively little time.

According to the results presented in the paper, we can conclude that this principle of object detection by using the Yolov5 model is successful and efficient in localizing OD.

To the best of our knowledge, there is no method to localize the optic disc (OD) acquired 100% accuracy with an average time of 0.19 seconds for thirteen well-known datasets in addition to 10% of the four datasets used for training. This paper proves that the proposed method is efficient in terms of accuracy and time.

In both normal and diseased images, our method achieved a high success rate of accuracy in OD detection and localization. Furthermore, in the current proposal, we noticed that the proposed method was the most accurate compared with the state-of-art methods. Apart from that, the approach precisely bounded the OD (the OD touches each side of the box) rather than determining a region around it.

Using the YOLO algorithm is new for OD detection and localization.

The current method works excellently with all challenges that reduce the detection accuracy in the previous works.

Depending on the results, the suggested method might be used to identify and segment parts of medical images to diagnose diseases (for example, Diabetic Retinopathy (DR) and Glaucoma) because it has high accuracy and speed that enables it to work in real-time.

## Conflicts of interest

The authors declare no conflict of interest.

## Author contributions

Conceptualization, Nidhal, and Hussien; methodology, Nidhal; software, Hussien; validation, Nidhal, Hussien; formal analysis, Hussien; investigation, Hussien; resources, Hussien; data curation, Hussien; writing—original draft preparation, Hussien; writing—review and editing, Nidhal; visualization, Nidhal; supervision, Nidhal; project administration, Nidhal.

## References

[1] R. J. Chalakkal, W. H. Abdulla, and S. S. Thulaseedharan, “Automatic detection and segmentation of optic disc and fovea in retinal images”, *IET Image Process.*, Vol. 12, No. 11, pp. 2100–2110, 2018.

[2] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, “Robust vessel segmentation in fundus images”, *Int. J. Biomed. Imaging*, Vol. 2013, 2013.

[3] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, “Origa-light: An online retinal fundus image database for glaucoma analysis and research”, In: *Proc. of 2010 Annual International Conference*

*of the IEEE Engineering in Medicine and Biology*, pp. 3065–3068, 2010.

[4] T. T. Khaing, P. Aimmanee, S. Makhanov, and H. Haneishi, "Vessel-based hybrid optic disk segmentation applied to mobile phone camera retinal images", *Med. Biol. Eng. Comput.*, Vol. 60, No. 2, pp. 421–437, 2022.

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

[6] M. N. Bajwa, G. A. P. Singh, W. Neumeier, M. I. Malik, A. Dengel, and S. Ahmed, "G1020: A benchmark retinal fundus image dataset for computer-aided glaucoma detection", In: *Proc. of 2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, 2020.

[7] W. Luangruangrong and K. Chinnasarn, "Optic disc localization in the complicated environment of a retinal image using circular-like estimation", *Arab. J. Sci. Eng.*, Vol. 44, No. 4, pp. 4009–026, 2019.

[8] S. D. Athab and N. H. Selman, "Localization of the Optic Disc in Retinal Fundus Image using Appearance Based Method andVasculature Convergence", *Iraqi J. Sci.*, pp. 164–170, 2020.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", In: *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, 2016.

[10] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement", *arXiv Prep*r. arXiv1804.02767, 2018.

[11] G. Jocher, L. Changyu, A. Hogan, L. Yu, P. Rai, and T. Sullivan, "ultralytics/yolov5: Initial Release (v1.0)", *Github Repository*, YOLOv5, 25-Jun-2020. [Online]. Available: https://doi.org/10.5281/zenodo.3908560#.YtyvzJ9PBf0.mendeley. [Accessed: 24-Jul-2022].

[12] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels", *IEEE Trans. Med. Imaging*, Vol. 22, No. 8, pp. 951–958, 2003.

[13] J. Krause, V. Gulshan, E. Rahimy, P. Karth, K. Widner, G. S. Corrado, L. Peng, and D. R. Webster, "Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy", *Ophthalmology*, Vol. 125, No. 8, pp. 1264–1272, 2018.

[14] D. J. J. Farnell, F. N. Hatfield, P. Knox, M. Reakes, S. Spencer, D. Parry, and S. P. Harding, "Enhancement of blood vessels in digital fundus photographs via the application of multiscale line operators", *J. Franklin Inst.*, Vol. 345, No. 7, pp. 748–765, 2008.

[15] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. V. Ginneken, "Ridge-based vessel segmentation in color images of the retina", *IEEE Trans. Med. Imaging*, Vol. 23, No. 4, pp. 501–509, 2004.

[16] M. M. Fraz, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification", *Int. J. Comput. Assist. Radiol. Surg.*, Vol. 9, No. 5, pp. 795–11, 2014.

[17] E. Decenciere, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, and A. Erginay, "Feedback on a publicly distributed image database: the Messidor database", *Image Anal. Stereol.*, Vol. 33, No. 3, pp. 231–234, 2014.

[18] T. Kauppi, V. Kalesnykiene, J. K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kalviainen, and J. Pietila, "DIARETDB0: Evaluation database and methodology for diabetic retinopathy algorithms", *Mach. Vis. Pattern Recognit. Res. Group, Lappeenranta Univ. Technol. Finl.*, Vol. 73, pp. 1–17, 2006.

[19] R. Kalviainen and H. Uusitalo, "DIARETDB1 diabetic retinopathy database and evaluation protocol", In: *Proc. of Medical Image Understanding and Analysis*, Vol. 1, pp. 61-65, 2007.

[20] M. Niemeijer, B. V. Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sánchez, B. Zhang, R. Hornero, M. Lamard, and C. Muramatsu, "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs", *IEEE Trans. Med. Imaging*, Vol. 29, No. 1, pp. 185–195, 2009.

[21] E. J. Carmona, M. Rincon, J. G. Feijoo, and J. M. M. D. L. Casa, "Identification of the optic nerve head with genetic algorithms", *Artif. Intell. Med.*, Vol. 43, No. 3, pp. 243–259, 2008.

[22] A. Almazroa, S. Alodhayb, E. Osman, E. Ramadan, M. Hummadi, M. Dlaim, M. Alkatee, K. Raahemifar, and V. Lakshminarayanan, "Retinal fundus images for glaucoma analysis: the RIGA dataset", In: *Proc. of Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, Vol. 10579, pp. 55–62, 2018.

[23] J. Sivaswamy, S. R. Krishnadas, G. D. Joshi, M. Jain, and A. U. S. Tabish, "Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation", In: *Proc. of 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pp. 53–56, 2014.

[24] J. I. Orlando, H. Fu, J. B. Breda, K. V. Keer, D. R. Bathula, A. D. Pinto, R. Fang, P. A. Heng, J. Kim, and J. Lee, "Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs", *Med. Image Anal.*, Vol. 59, p. 101570, 2020.

[25] H. Fang, F. Li, H. Fu, X. Sun, X. Cao, J. Son, S. Yu, M. Zhang, C. Yuan, and C. Bian, "REFUGE2 Challenge: Treasure for Multi-Domain Learning in Glaucoma Assessment", *arXiv Prepr*. arXiv2202.08994, 2022.

[26] N. K. E. Abbadi, H. Mahdi, and H. Rustum, "Single image haze removal via accurate atmosphere light", *Int. J. Appl. Eng. Res.*, Vol. 12, No. 19, pp. 9149–9158, 2017.

[27] H. M. Unver, Y. Kokver, E. Duman, and O. A. Erdem, "Statistical edge detection and circular hough transform for optic disk localization", *Appl. Sci.*, Vol. 9, No. 2, p. 350, 2019.

[28] A. Mvoulana, R. Kachouri, and M. Akil, "Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images", *Comput. Med. Imaging Graph.*, Vol. 77, p. 101643, 2019.

[29] B. K. Triwijoyo, "Optic Disk Segmentation Using Histogram Analysis", *Int. J. Eng. Comput. Sci. Appl.*, Vol. 1, No. 1, pp. 27–34, 2022.

[30] M. Tavakoli, "Automated optic disk detection in fundus images using a combination of deep learning and local histogram matching", In: *Proc. of SPIE Vol*, Vol. 12036, pp. 120360I–1, 2022.