

RESEARCH

Open Access



Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network

Ayoub Skouta^{1*}, Abdelali Elmoufidi², Said Jai-Andaloussi¹ and Ouail Ouchetto¹

*Correspondence:
ayskouta@gmail.com

¹ Computer Science and Systems Laboratory, Department of Mathematics and Computer Sciences, Faculty of Sciences Ain Chock, Hassan II University of Casablanca, Casablanca, Morocco

² Data4Earth Laboratory, Sultan Moulay Slimane University, Beni Mellal, Morocco

Abstract

Because retinal hemorrhage is one of the earliest symptoms of diabetic retinopathy, its accurate identification is essential for early diagnosis. One of the major obstacles ophthalmologists face in making a quick and effective diagnosis is viewing too many images to manually identify lesions of different shapes and sizes. To this end, researchers are working to develop an automated method for screening for diabetic retinopathy. This paper presents a modified CNN UNet architecture for identifying retinal hemorrhages in fundus images. Using the graphics processing unit (GPU) and the IDRiD dataset, the proposed UNet was trained to segment and detect potential areas that may harbor retinal hemorrhages. The experiment was also tested using the IDRiD and DIARETDB1 datasets, both freely available on the Internet. We applied preprocessing to improve the image quality and increase the data, which play an important role in defining the complex features involved in the segmentation task. A significant improvement was then observed in the learning neural network that was able to effectively segment the bleeding and achieve sensitivity, specificity and accuracy of 80.49%, 99.68%, and 98.68%, respectively. The experimental results also yielded an IoU of 76.61% and a Dice value of 86.51%, showing that the predictions obtained by the network are effective and can significantly reduce the efforts of ophthalmologists. The results revealed a significant increase in the diagnostic performance of one of the most important retinal disorders caused by diabetes.

Keywords: Diabetic retinopathy, Fundus images, Segmentation, Detection, Deep learning, Convolutional neural networks, Artificial intelligence, CAD system

Introduction

In recent years, Big Data has ushered in a new era of data. It recognizes highly varied and complicated data that is defined by the amount of data recorded, distinct data formats, speed of data generation, and veracity. The latter suggests a decline in confidence as the volume of data increases. Computer vision remains a critical research topic in the era of Big Data to enable a machine to analyze a large amount of data that represents distinct features of each image to predict with very good accuracy whether or not an area

belongs to the lesion. The challenges that Big Data can solve are two. The first is storage, as we are working with very large datasets containing very high resolution images. This has been quite well solved in terms of storage thanks to the availability of powerful machines. On the other hand, our main objective will be the challenge of analyzing the different images in real time. This can be done by an intelligent system that works in the same way as humans in order to extract value from the data by interpreting a large number of pixels of each image. With the help of artificial intelligence a machine can be enabled to mimic this type of behavior that we see in humans. So, finally the AI with the help of big data should be able to store, compute and then learn from the data.

Diabetes is a chronic disease caused by high blood sugar levels. In the short or long term, diabetes can damage nerves and blood vessels in many organs, such as the eyes and kidneys. There are two main types of diabetes: type 1 and type 2. Type 1 diabetes is caused by lack of insulin secretion by the pancreas [1]. The second type is caused by an increase in insulin consumption by the body's cells [2]. According to statistics, the number of diabetic patients has increased significantly for people over 18 years of age; their percentage has increased from 4.7 to 8.5% of patients [3]. The World Health Organization (WHO) predicts that diabetes ranks seventh in the world in terms of mortality rate. People with diabetes are more likely to develop diabetic retinopathy, as they are at greater risk of permanent vision loss if not treated in time. The retina, a sensitive organ of vision, is a thin membrane that lines the back of the eyeball and contains cells that receive light signals and then the information is transmitted by the optic nerve to the brain, which reconstructs the image. Figure 1 shows a normal fundus image.

Diabetic retinopathy is a serious complication of diabetes that affects 50% of patients with type 2 diabetes. High blood sugar levels weaken the capillary walls of the retina and make certain areas of the retina insufficiently oxygenated, causing blood vessels to burst, which can lead to permanent vision loss. At first, symptoms are not apparent, this is why regular check-ups and early detection by specialists are so important. Furthermore, retinopathy accelerates the onset of other eye diseases such as glaucoma [4] or cataracts. Most epidemiological studies have shown that diabetic retinopathy is the leading cause of blindness in people aged 20 to 74 years [5]. For this reason, the patient's retina should

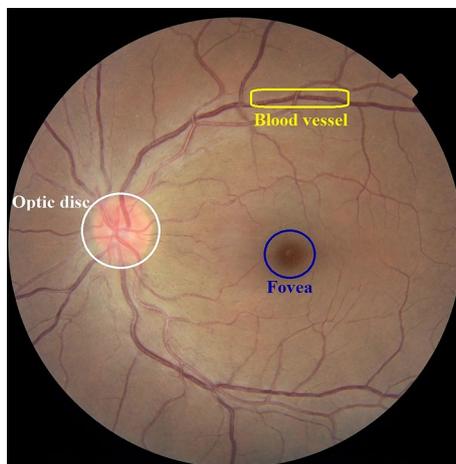


Fig. 1 Normal fundus image

be checked regularly by an ophthalmologist. The ophthalmologist performs several tests to determine the patency of the retinal vessels: measurement of visual acuity, measurement of eye tension, measurement by dilation of the pupil and the practice of retinal angiography.

These numerous maneuvers require a trip by the patient or the attending physician, a considerable amount of time and the intervention of a large number of human resources (optometrists, nurses, technicians, ophthalmologists, orthoptists...). In addition, this is only the beginning of a long series of analyzes leading to the morphological and physiological study of the eye and finally to the detection and estimation of lesions. Despite all this work, the results and diagnoses lack clarity and precision due to the very minute nature of the lesions and the complexity of their aspects (variations in size, color, morphology and shape). Today, artificial intelligence (AI) methods have already largely penetrated the fields of medical research, diabetes is no exception. The methods and algorithms that are particularly adapted to imaging data are very promising. It is the implementation of a set of techniques and computer theories aimed at simulating or reproducing human reasoning and learning in machines: (Deep Learning, convolutional neural networks). This alternative is beneficial for diabetic patients because computer assistance significantly improves the early diagnosis of advanced eye diseases, with reliability beyond the expert consensus. Over the past two decades, computer-aided diagnosis (CAD) has fueled the development of medicine in general [6, 7]. The innovative and effective uses of artificial intelligence have one goal: to improve the quality of care and provide benefits and facilities that have enabled: predict advanced eye diseases, automate disease detection in less time, measure their growth rates, reach a larger population, reduce costs and treat patients on time.

Medical imaging is a modern diagnostic tool in medicine that is undergoing a great evolution, allowing for better diagnosis and new hope for the treatment of many eye diseases. Thanks to advances in computer technology, it is possible to acquire images of the retina, to identify and visualize lesions indirectly. Moreover, these representations are subject to interpretation and facilitate the use of surgery. The eye is visualized using various retinal imaging screening techniques to detect and diagnose the presence of lesions indicative of early diabetic retinopathy, namely hemorrhages, microaneurysms, soft exudates and hard exudates [8]. There are several cross-sectional imaging techniques such as: Optical Coherence Tomography (OCT) [9], Fundus Fluorescein Angiography (FFA) and Color photography of the fundus of the eye (CF) [10].

In this article, we focus on diagnosis using fundus images. The identification of bleeding is a common telltale symptom of diabetic retinopathy, as it corresponds to the leakage of blood from capillary networks or venous branches. Weakening of the vascular walls leads to disruption of the inner blood-retinal barrier during sudden changes in perfusion pressure [3]. The different retinal abnormalities of diabetic retinopathy are illustrated in Fig. 2.

Segmentation of retinal images is a very complex process due to low contrast, uneven shape and blurred edges of the lesions. In recent years, deep learning (DL) methods have advanced considerably and produced better results through a set of specific techniques based on the careful application of fine engineering research that uses the computational power of machines to understand, analyze and process one or more images. These high

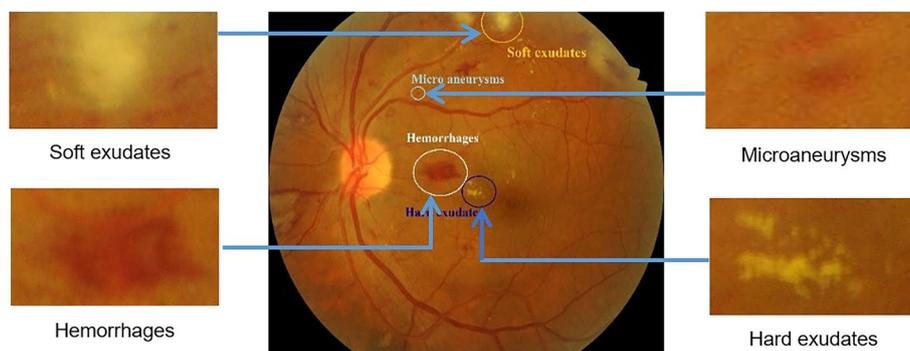


Fig. 2 A fundus image of the affected eye

definition techniques are used in various application areas such as: image classification [3, 11, 12], road mining [13], segmentation of brain tumors [14] and object detection [15].

Convolutional neural networks (CNNs), which are a branch of deep learning, have achieved the highest efficiency over traditional machine learning-based methods [10] for medical imaging segmentation using pixel-by-pixel labeling, due to their ability to extract and learn the most discriminative features at the pixel level. In this work, we use a convolutional neural network based on the U-Net architecture to develop a method for automatic segmentation of retinal hemorrhages in the fundus image. The neural networks we train on data sets that contain very high resolution images labeled pixel by pixel. The first step is to search the data to train the network to make decisions by providing correct information. Then it trains on this data set, we will give it a separate data set called the test data set. The validation data set to see how well it actually works after training.

The structure of our paper is as follows. The second part is a related work that mainly presents previous studies on bleeding segmentation and their obtained results. The third part describes semantic segmentation, deep learning and convolutional neural networks respectively and mainly explains our proposed method. The fourth part presents the datasets that were used to perform this experiment and the results of our experiment. Finally, the last part concludes and presents future work.

Related previous works

In recent years, machine vision has made dramatic progress, thanks in part to recent advances in optimization and the explosion of computer computing power. For example, advances that have been made in facial recognition are being replicated by the research community in the medical field. AI can process thousands of images in seconds and detect important information with great accuracy that would have taken radiologist months to find. For this reason, many computer-aided diagnostics have been developed to help physicians improve diagnostic results in many medical fields, such as breast cancer [16–20], brain cancer [21, 22], diabetic retinopathies [23, 24], etc. This section presents the different techniques that have been applied to segment retinal hemorrhages in fundus images of patients with diabetic retinopathy. There is an extensive literature on the detection of retinal hemorrhages. For example, the

regional growth-based technique described by The methodology published by Gardner et al. [25] applies an artificial neural network to indicate the presence or absence of a hemorrhage, exudate or blood vessel from the squares obtained by non-overlapping slicing of a retinal image of size 700×700 pixels to several small squares of 20×20 pixels. The network has a sensitivity of 73.8% for the detection of hemorrhages, which shows a high accuracy of detection compared with the results of the examination performed by an ophthalmologist. Another method presented by Sinthanayothin et al. [26] that automatically segments hemorrhages and exudates using the “Moat Operator” and increasing recursive segmentation of regions. Ophthalmologists noted hemorrhages and exudates in non-overlapping cut images of several small 10×10 squares. The authors did not use pixel segmentation to evaluate their segmentation, but instead evaluated their segmentation using segments. The non-proliferative diabetic retinopathy feature detection (NPDR) technique achieved a sensitivity of 77.5% and a specificity of 88.7% for the group containing hemorrhages and micro-aneurysms. To classify red lesions, Kande et al. [27] detected microaneurysms and hemorrhages by pixel classification and mathematical morphology. They used the red and green channels of the image to assess whether or not the image had red lesions. Then, the support vector machine (SVM) algorithm is used to classify candidate areas for red lesion containment. The proposed approach has a specificity of 91% and a sensitivity of 100%. Tang et al. [28] They detected hemorrhages using the k-Nearest Neighbors algorithm as a splat-based characteristic classifier selected by an envelope and filter method. This experiment yielded a receiver operating feature (ROC) curve score of 0.96 with the MESSIDOR data set. Grinsven et al. [29] created a CNN architecture to detect hemorrhage with nine layers formed by selective samples and 41×41 size patches labeled with or without evidence of hemorrhage. The results obtained are 84.8 and 90.4, respectively, for sensitivity and specificity to identify hemorrhages on images from the MESSIDOR and KAGGLE data sets. Tan et al. [30] developed a 10-layer multiclass neural network to segment hemorrhages, microaneurysms and exudates in images of the retinal fundus. Hemorrhage segmentation had a sensitivity of 62.57% and a specificity of 98.93%. Quellec et al. [31] developed a CNN model have using the ConvNets network structure that generates heat maps to simultaneously detect four forms of diabetic retinopathy lesions: microaneurysms, hemorrhages, exudates and absorbent cotton spots. For hemorrhages detection, the model showed an AUC of 0.614. Karkuzhali and Manimegalai [32] includes two preprocessing steps. The median filter is the initial preprocessing step, followed by the Sobel operator. After that, a slide segmentation is performed, which involves moving a kernel window over the entire retinal image. This approach has an accuracy of 93.21%, a sensitivity of 90.02% and a specificity of 88.43%. Lam et al. [33] tested five CNN models, including AlexNet, VGG16, GoogLeNet, ResNet and Inception-v3 to locate different types of lesions in retinal images. The ophthalmologist examines the image of the fundus to produce patches containing hemorrhages, microaneurysms, exudates, retinal neovascularization, or normal-looking structures. These image patches are used to train convolutional neural networks to predict the existence of these five categories. To create a probability map for the entire image, the sliding window method is used. Badar et al. [34] proposed an encoder-decoder for simultaneous segmentation

of hemorrhages, soft exudates and exudates based on a CNN inspired by the semantic segmentation network Segnet. Using the Messidor dataset for training and testing, the proposal achieved 97.86% accuracy for semantic hemorrhages segmentation. Orlando et al. [35] constructed a CNN and combined it with a random forest to segment hemorrhages and microaneurysms. Probability maps of hemorrhages and microaneurysms located at the image level are generated using the random forest algorithm that receives the green layer features from the patches extracted by the CNN architecture. In this experiment using the DIARETDB1 dataset, the approach has a sensitivity of 48.83 for detecting microaneurysms and hemorrhages. Saha et al. [36] used a method based on a fully convolutional deep neural network trained end-to-end for automatic segmentation of multiple lesions at once, including microaneurysms, hemorrhages, hard exudates, soft exudates and optic disc. The network, called “SegNet”, includes an encoder in the form of a 13-layer convolutional VGGNet and a decoder that manages the classification on a pixel-by-pixel basis. Ananda et al. [37] modified U-Net deep neural network by reducing the number of filters and coding layers to segment different types of retinal diseases; each of the U-Net models is used to segment one of the disease types, such as hemorrhages, microaneurysms, hard exudates or soft exudates and optic disc. The modified version of the U-Net model yielded a dice coefficient of 0.86 for bleeding segmentation. Guo et al. [38] wrote a paper using the L-Seg model, which is a modified version of the VGG16 architecture, to detect the presence of various diseases such as hemorrhages, microaneurysms, soft exudates and hard exudates. It is a method that combines multiscale properties to solve the problem of segmenting small areas as efficiently as possible. The results obtained by the L-Seg architecture are superior when the IDRiD dataset is utilised; this technique provided an AUC of 67.34 for hemorrhage. Yan et al. [39] proposed a study consisting of three parts: the first two are U-net models called GlobalNet and LocalNet. The third part is a fusion unit that integrates the outputs of two U-net models to segment the four types of retinal lesions present in each image, including hemorrhage, microaneurysm, hard exudate or soft exudate. The combined model was successful in segmenting hard exudate and microaneurysms. The GlobalNet model, on the other hand, exhibited the greatest hemorrhage and soft exudate segmentation rates. For hemorrhage segmentation, this approach yielded a score of 0.711. Huang et al. [40] used a convolutional neural network that combines BBR-Net, which can improve the accuracy of the annotations of training data and RetinaNet to identify the presence of hemorrhages. The system in question starts with a preprocessing step of the fundus image that uses CLAHE correction and adaptive gamma correction to adjust for irregular illumination and low contrast. The method outperformed the standard RetinaNet system, demonstrating its ability to refine manually traced hemorrhage annotations. A mean IoU value of 0.8715 was recorded on the IDRiD data set.

The literature discussed above presents different methods of retinal image analysis for computer-aided diagnosis of hemorrhages. The proposed work has resulted in considerable progress that is increasing daily. The types of methodologies discussed in this article have focused on: basic image processing methods and methods using mathematical approaches to morphology and region growth. These methods were effective at one time, but have been replaced by computer vision and machine

learning approaches that are characterized by the strength of feature extraction, selection and classification. This type of research has produced high performance measures. In addition, advances in deep learning and the application of CNN are leading to good results, especially in medical imaging which allows for more accurate localization compared to previous methods.

The related previous works shows that deep learning-based systems have become a popular research area due to their greater strength and ability to automatically extract features compared to machine learning-based methods. In addition, deep learning allows for accurate localization of the retinal boundaries. The only limitation is that its training is time consuming and remains difficult.

Materials and methods

Semantic segmentation

Segmentation is an important procedure applied to medical images of the retina, as it can greatly assist in the diagnosis of diabetic retinopathy and thus in the identification of areas of interest in retinal images that are often difficult to detect. It consists of dividing the medical image into a set of pixel groups or areas representing an anatomical structure such as the fovea, optic disc, retinal blood vessels or lesions like hemorrhages, microaneurysms, soft and hard exudates [41]. In this work, we will use a semantic segmentation algorithm, also known as pixel segmentation. It associates each pixel with its class so that the result of semantic segmentation is an image in which each pixel belongs to a certain group. To distinguish these groups, they will be assigned a different color than the others. Classical machine learning algorithms quickly replaced the first semantic segmentation methods. They were then overtaken by the development of deep learning, which proved to be much more efficient, comparable to that of humans [42]. Unlike classifiers such as AlexNet and VGGNet, where there is only one classifier that generates an output vector, the semantic segmentation architecture essentially consists of the contracting path (encoder) and the expansive path (decoder). An encoder is a network that takes an input image and generates a feature vector. The decoder accepts an input feature vector and produces a semantic segmentation mask [43]. Semantic segmentation has several classifiers that operate simultaneously. The number of classifiers is equal to the number of pixels in the input image. Each classifier generates its own prediction vector and each pixel is classified with a vector whose size is equal to the number of categories [44]. Figure 3 illustrates semantic segmentation in fundus image.

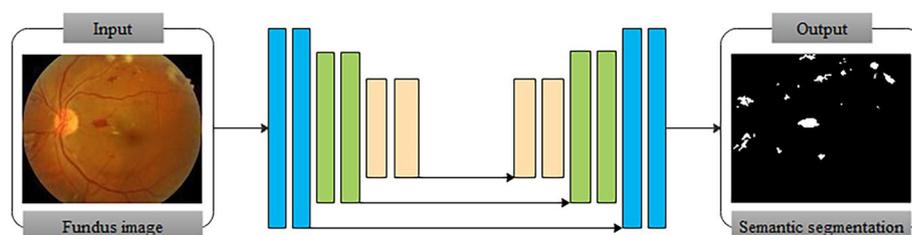


Fig. 3 Semantic segmentation in fundus image

Convolutional neural network

In this section, we discuss convolutional neural networks, also known as CNNs. They are a type of deep learning that has a great ability to automatically recognize the image content. It has proved that are very effective and more suitable for all computer vision applications such as autonomous driving, pattern recognition, facial recognition, robotic vision, satellite image segmentation and medical image diagnosis. These are probably the most successful AI models, which are inspired by biology and more precisely by the functioning of the brain [45], through these models machines can see and thus replace time-consuming and effort-consuming human work. To construct a CNN, primary and secondary layers are used. Among the most important primary layers are the convolutional layers, which are the backbone of the construction of a CNN and play a major role in extracting features from the image using the convolution process [46]. Pooling layers are the second type and they are used to reduce the size of the input sample. The third type is the flat and dense layers. These layers are fully connected layers, which are the last layer of the network in which classification takes place using the flattening operation to convert the output of the convolutional part of the CNN into 1D feature vector. Each neuron must perform a non-linear transformation of its input. To do this, there are several activation functions, which allow each neuron to perform a non-linear transformation of its input [46]. Secondary, layers play an important role in improving network performance, for example, dropout layers that significantly improve model generalization while preventing overfitting, batch normalization layers and regularization layers [44]. As shown in Fig. 4, the combinations of convolution and pooling layers are performed iteratively to extract features, so that each time this process is applied, the depth of the content acquired by the CNN is increased so that it has an accurate understanding, facilitates the creation of sophisticated descriptors. Regarding the classification process, fully connected layers are used to obtain an accurate classification of these descriptors [47].

CNN models such as VGGNet, ResNet and AlexNet have been very successful in image classification [48]. But they have failed to associate each pixel of an image with its class due to the propagation of feature maps through the different CNN stages, leading to relatively fuzzy boundaries due to the loss of information from the original image. To overcome these drawbacks, semantic segmentation is an interesting alternative that has achieved better results by dividing the retinal image into several distinct classes, each

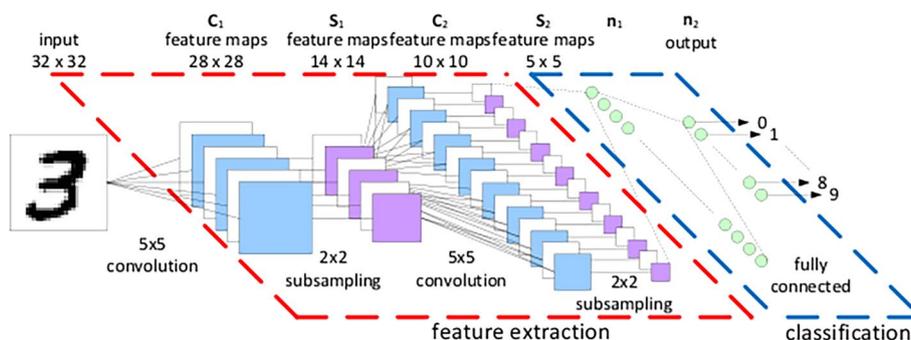


Fig. 4 Structure of a convolutional neural network

corresponding to a particular retinal lesion or anatomical structure. There are many deep learning-based architectures that can solve the semantic segmentation task. Such as fully convolutional neural network (FCN), SegNet architecture, DeepLab architecture, RefiNet architecture, PSPNet architecture, UNet architecture, etc. Since the UNet model is applicable to many biomedical segmentation problems and offers better performance, we chose to use it in this work to create a segmentation-capable program to handle the problem of segmentation of retinal hemorrhage from fundus images. After observing the structure of UNet, it is easy to understand where the name comes from, because this architecture is more or less symmetrical and it organizes the layers of neurons in the form of the letter “U”. The input to the neural network, also called the contraction path, is located on the left. It follows the architecture of a typical convolutional classification network but does not have fully connected layers; its role is to subsample the incoming image and create a feature vector. The output of the neural network, also known as the expansive path, is located on the right. Its function is to expand the size of the characterization maps and generate a mask of the same size as the incoming image, representing a semantic segmentation of retinal hemorrhages.

Proposed UNet architecture

To address the problem of segmentation of retinal hemorrhage, which is one of the most important indicators of the development of diabetic retinopathy produced by a vascular explosion due to the pressure given to it, we will use one of the most effective structures created by Ronneberger et al. [49]. We will modify the UNet design and write a special one to solve the retinal hemorrhage segmentation problem, because after testing, the architecture did not work with the original UNet network settings. Figure 5 shows a schematic representation of the main workflow used, which is divided into two steps. The first step is the training of the network using color images with three-dimensional

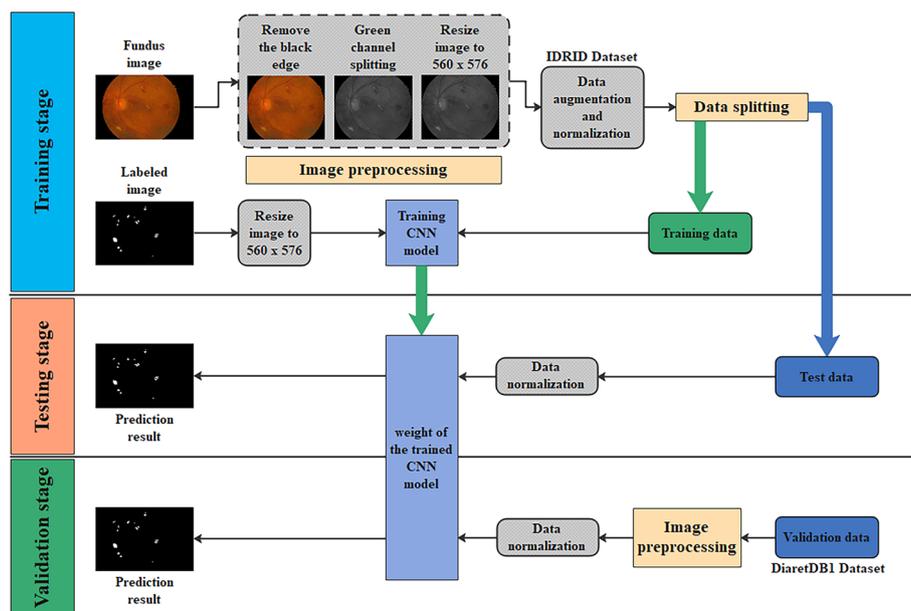


Fig. 5 An overview of the main process of our method

values, i.e., image width, length and depth, which refers to the number of RGB color channels. The second step is to examine the generalizability of the method by testing the network on two data sets.

A preprocessing step is necessary to obtain better data to facilitate subsequent operations. This step begins with the removal of the black border from the original data, followed by separation of the RGB color channels to extract the green channel, which has a strong contrast between hemorrhages and background in the fundus image compared to the red or blue channel [50, 51] and ends with resizing of the image. The results of the preprocessing step are shown in Fig. 6.

To be able to solve the data scarcity problem that deep learning requires training a convolutional neural network, we use data augmentation techniques to increase the amount of data available in the dataset. The data obtained through the data augmentation process is then normalized and separated into two parts, one for training and one for testing. We will feed the neural network with the training portion containing the green channel of the fundus images that provide finer and greater detail on the retinal hemorrhage in the input image. We will also need to provide the neural network with labeled data that is in the form of a training mask. This consists of black images with white spots that will mark the regions of interest to teach the neural network to recognize and identify retinal hemorrhages.

The encoder path on the left is composed of four blocks; the first one receives the green channel of the background image, while the following blocks receive the output of the previous block as subsampled images with a lower resolution than the previous layer. The first block consists of three convolutional layers, with 32 feature maps on the first and second layers and 64 feature maps on the third. The second block also consists of three convolutional layers, with 64 feature maps in the first two layers and 128 in the third. Only the third and fourth blocks comprise two convolutional layers and the number of feature maps is doubled at each subsampling step to 128 and 256, respectively. The last feature map is divided by four at the end of each block, using the most common form of pooling operation, Maxpooling. This layer is used to reduce the size of the input sample. While retaining the most critical information from the input layer. Then, an intermediate copy of the feature map is kept before each Maxpooling layer, which will be used to connect the contraction path to the expansion path. This duplicate of the feature map will help produce a more accurate segmentation result.

The decoder path on the right side of the design consists of four blocks that are symmetric in terms of the feature map layer configuration with the encoder path. Each block

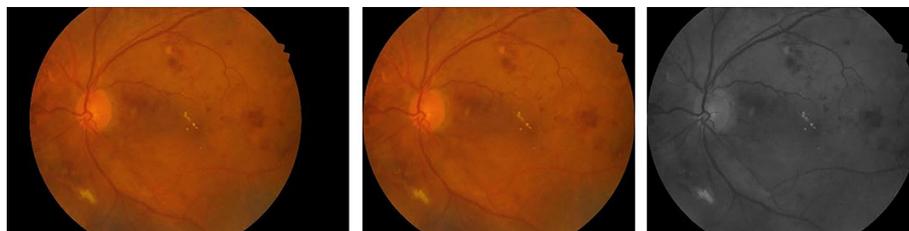


Fig. 6 (left) Input RGB fundus image with black border (middle) Input RGB fundus image after extracting the black border and (right) RGB color space green channel

ends with an operation that oversamples the feature map using the “UpSampling2D” layer with Step 2, which takes the number of feature maps from the previous block and divides it by two to reduce the number of feature maps in each sampling phase. In addition, the important procedure indicated by the blue horizontal arrows in Fig. 7 recovers the mirrored feature maps that were previously held before each maximum clustering layer in each of the four encoder path blocks and correlates them with the oversampling results drawn using the “UpSampling2D” layer of the decoder path. We added an additional subsampling step after the last block of the decoder path and then concatenation between the network result and the initial input image to achieve better segmentation of retinal hemorrhages. Finally, we need to perform a last convolution processing using a 1×1 kernel followed by a sigmoid activation function to compute the results of both maps. This function offers the final result of the pixel segmentation and decides whether the pixel belongs to the lesion or not.

$$g(z) = \frac{1}{1 + e^{-z}} \tag{1}$$

The bottleneck is located in the middle of the contraction and expansion paths, allowing them to connect. It has two convolution layers, the first of which enhances the feature maps and the second of which reduces them. Each convolution layer has a 3×3 kernel and all Maxpooling operations have a 2×2 kernel with a stride of 2. The size of the input image is slightly reduced when the convolution processing is performed. To avoid this reduction, generate an image with the same resolution as the input image. We surround

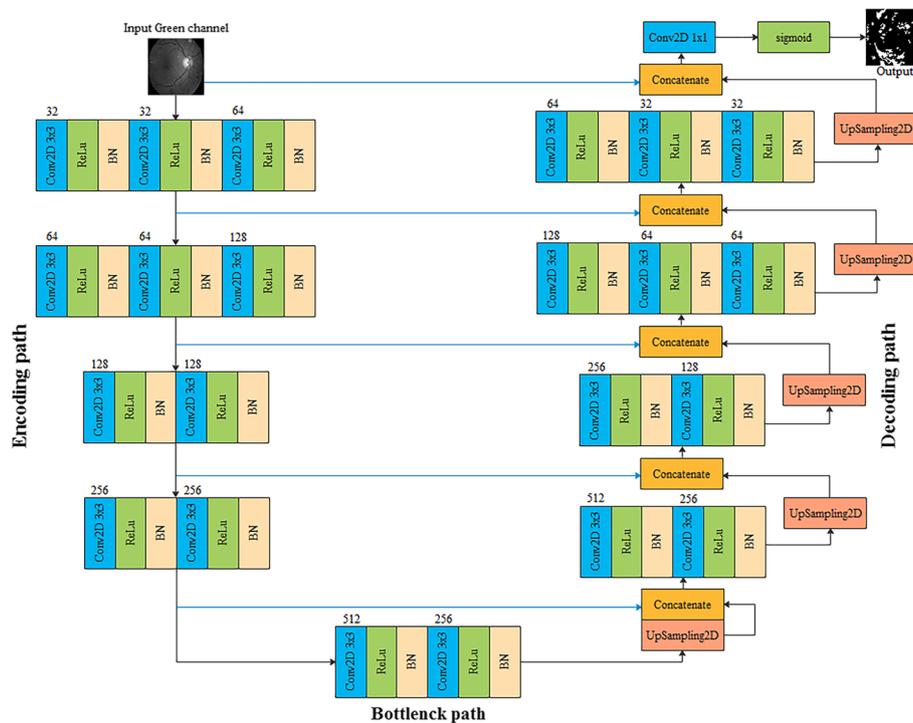


Fig. 7 Modified UNet architecture for hemorrhage segmentation in fundus images for the diagnosis of diabetic retinopathy

the original image with one or more rows and columns of zero. As a result, we will have a feature map with the same input and output layers. To put it another way, the padding is equal to the same. The convolutional layer is followed by the RELU activation function and the batch normalization layer. The rectified linear unit (RELU) activation function replaces all negative activations with 0 by applying the following function:

$$g(z) = \max\{0, z\} \quad (2)$$

The batch normalization layer normalizes the output of the activation function and scales it down to a zero-to-one scale. This prevents the network weights from becoming unbalanced due to very high or low values. Since the batch normalization layer is included in the gradient process, gradient deterioration is avoided. The inclusion of batch normalization in our model can significantly improve training speeds while reducing the possibility of exceeding high weights that would significantly impact the training process.

Experience and results

In this section, we detail the platform used to simulate the experimentation of this method, the datasets used, the approach taken to extend the amount of existing data and finally, we discuss the various results recorded. GPU (graphics processing unit) provided by Google's collaborative service (googlecolaboratory), as well as the Python Keras deep learning library and the TensorFlow backend, were adopted to simulate the experiment.

Dataset description

The data for this experiment are from the IDRiD [52] and DiaretDB1 [53] datasets that are freely available to the public [52, 53]. Currently, these are the only datasets with manual hemorrhage annotation [40]. The IDRiD dataset consists of 516 retinal fundus images with a resolution of 4288×2848 pixels, 80% of which are used for training and 20% for testing. To measure the degree of diabetic retinopathy and macular edema, IDRiD provides image-level annotations. It is the only dataset that also provides ground truths manually obtained from pixel-level annotations provided by retinal specialists for diabetic retinopathy-related lesions such as microaneurysms, hemorrhages, soft exudates and hard exudates. Ground truth is used to evaluate the performance of lesion segmentation techniques. The final function of the IDRiD database is to provide information on the location of the central pixel of the optic discs and the center of the fovea. Figure 8a shows a fundus image with hemorrhages from the IDRiD database, whereas Fig. 8b shows the accompanying labels.

The DIARETDB1 dataset contains 89 retinal fundus images with a resolution of 1500×1152 pixels, including 5 normal retina images and 84 abnormal retina images that were manually classified by experts. Microaneurysms, hemorrhages, hard and soft exudates are the four types of lesions identified by DIARETDB1. The DiaretDB1 dataset is an effective standard in screening for diabetic retinopathy as accurately as possible. It offers annotations to validate the results obtained by an encircling or delineating approach. Pixels are considered as ground truth only if the confidence level of the labeling is higher than the average of 75% of experts. Figure 9a shows a fundus image with hemorrhages from the DIARETDB1 database, whereas Fig. 9b shows the accompanying ground truth.

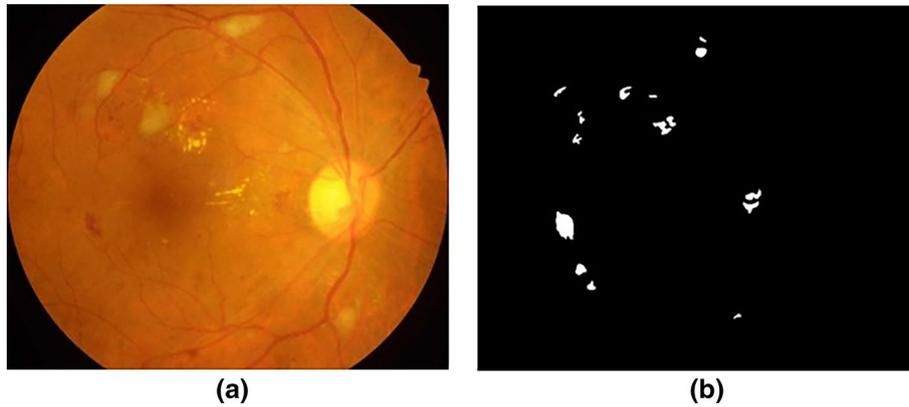


Fig. 8 **a** Examples of a fundus image with retinal hemorrhages from the IDRiD database and **b** ground-truth examples of retinal hemorrhages from the IDRiD database

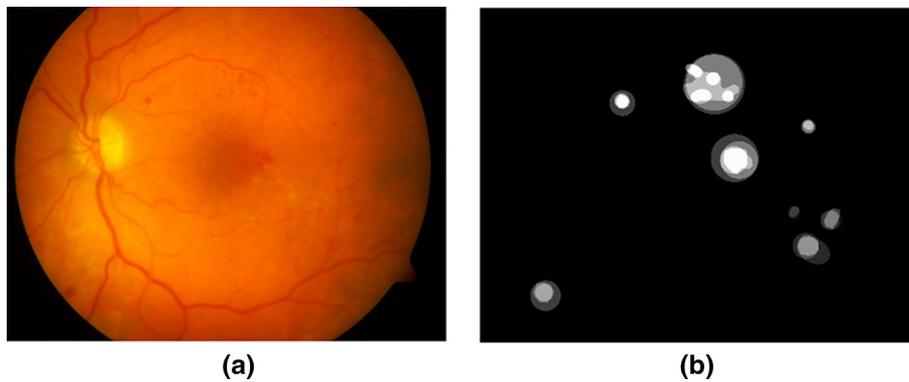


Fig. 9 **a** Examples of a fundus image with retinal hemorrhages from the DIARETDB1 database and **b** ground truth examples of retinal hemorrhages from the DIARETDB1 database

IDRID subset of data, that provides pixel-level annotations for segmentation, is used for training testing and validation in this study. We took 80 fundus images with bleeding symptoms and separated them into two parts: 70 images for training and testing and 10 images for validation. To facilitate the calculations, the images have been resized to a size of 560×576 pixels. The chosen geometry will allow us to avoid falling into odd numbers when dividing and multiplying by two when encoding and decoding the images. The size of the dedicated training and test dataset is relatively small for our convolutional network to work successfully. As a result, we will need to create more images using data augmentation technology, as 70 images is a relatively small number.

The DiaretDB1 dataset is only used in the validation stage and not in the learning phase. The bounding circles or frames are often incorrect, which could directly influence the performance of the CNN.

Data augmentation

A major difficulty in using deep learning is the amount of data, which must be huge. Increasing the amount of data is a critical component for the network to acquire the desired invariance and robustness characteristics to improve segmentation accuracy. It also avoids overfitting the network and provides a more representative image base, which

greatly improves the learning of the neural network. Convolutional neural networks are robust to changes in the original images due to the data augmentation approach, such as changes in image shape, brightness and other factors that generate more data. Both the original image and the generated images are used as training images. In the original images, the degree of similarity between the hemorrhages, blood vessels and the background is very high [7]. Through the transformations that occur in the images as a result of the data augmentation process, it is possible to provide the CNN with a set of features that allows it to distinguish between the hemorrhages, the blood vessels and the background in the original image. We will multiply the number of images by performing transformations on each original image. First, we will perform a 90° rotation from 0° to 360°. For each image generated by the rotation process, we will apply random manipulations. These manipulations are adding noise, changing the gamma (for brightness changes) and changing the colorimetry. We also perform vertical and horizontal flipping, as well as horizontal and vertical flipping of the image. The example of data augmentation is shown in Fig. 10. The final dataset consists of 1190 images, which will be divided into two subsets: training and test. We will randomly extract 90% of the total images for training and the remaining 10% will be reserved for testing. This gives 1071 training images and 119 test images. Not counting the 10 images that were recorded before the data augmentation process and will be used in the validation step. Finally, using the training dataset, we can begin teaching our network. We can use the network to generate predictions and create images once the training phase is complete.

Algorithm 1 Data augmentation

```

Input: Fundus original image;
 $\theta = 0$ 
for  $i=1$  To 4 do
  Apply:  $\text{Img}(\theta_i) \leftarrow \text{Rotation}(\text{Img}, \theta)$ ; // The function Rotation applied a rotation with 90 degrees
  in each iteration.
   $\theta = \theta + 90^\circ$ ;
  add  $\text{Img}(\theta_i)$  to table T;
   $T[i] = \text{Img}(\theta_i)$ ; // Save the rotate image in the table.
end for
for  $i=1$  To 4 do
  Apply:  $\text{Img}(i) \leftarrow \text{Random}(T[i])$ ; // The function Random applied a noise; a luminosity and
  varying the color.
  add  $\text{Img}(i)$  to table T // The instruction to add the new images in the same table.
end for
for  $i=-1$  To 1 do // for flipping the image either horizontally or vertically.
  Apply:  $\text{Img}(i) \leftarrow \text{flip}(T[i])$ ; // This function flip for flipping the images either horizontally or
  vertically.
  for  $i=1$  To 3 do
    Apply:  $\text{Img}(i,j) \leftarrow \text{Random}(\text{Img}(i))$ ;
    add  $\text{Img}(i,j)$  to table T // This instruction to add the new images in the same table with
    are generated after applied a rotation, a flipping and the random function.
  end for
end for
return T; // The table T contains all the image after applying the augmentation step.

```

Results

Experimental settings

Before the learning process begins, the next step is to experimentally adjust the hyper-parameter values by performing different permutations to find the parameters

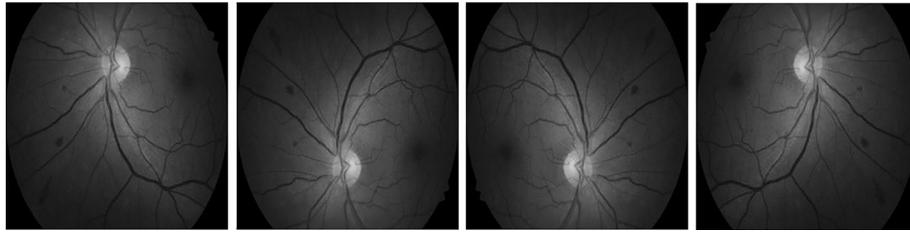


Fig. 10 Example of data augmentation

that provide the best segmentation results. We used the Adam optimizer [54] with default parameters to train the classifier on the training data. Since our pixel classification problem is a binary classification, we will use the binary cross-entropy function as the loss function. The batch size that we adopted in our experiment is four images per batch and our model training was adapted for 100 epochs.

The metrics that were used to evaluate the reliability of the suggested technique are presented in this section; these measures are often used in the process of evaluating the effectiveness of semantic segmentation. We describe their computational formulas for determining the success rate and demonstrate the level of similarity between the segmentation performed by the algorithm and the ground truth. The calculations are performed in terms of accuracy, precision, dice similarity coefficient, Intersection over Union (IoU), sensitivity and specificity.

The scoring values of the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) pixels were used to generate these metrics. True positives (TP) correspond to pixels correctly identified as retinal hemorrhages. Pixels correctly classified as non-retinal hemorrhage are true negatives (TN). False positives (FP) refer to pixels that do not belong to the retinal hemorrhage and are classified as retinal hemorrhages, whereas false negatives (FN) refer to pixels of retinal hemorrhages classified as not belonging to retinal hemorrhages.

- Accuracy: First, we use the accuracy, which is defined by the number of correctly predicted predictions divided by the total number of predictions. In semantic segmentation, the accuracy is formally computed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

- Precision: Corresponds to the percentage of positive samples correctly classified among all positive instances. Determined by dividing the number of true positives by the number of positive predictions. The precision is calculated using the following formula:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

- Similarity coefficient (Dice): Third, we will use the dice coefficient, or overlap index, to evaluate the similarity of two data sets. This is equivalent to the intersection of two data sets divided by the total size of the first set added to the second set. In semantic segmentation, the dice index is formally defined as follows:

$$Dice(X, Y) = \frac{2 |X \cap Y|}{|X| + |Y|} \quad (5)$$

We can use the Dice index as a metric to evaluate the accuracy of our neural network once we have discovered it. The Dice index is a number between 0 and 1. The two data sets are not identical if the Dice index is close to zero; otherwise, the two data sets are identical if the Dice index is close to one.

- Intersection over Union (IoU): Is a measure frequently used in most image segmentation algorithms to make a quantitative evaluation of performance. It is determined by dividing the intersection of the two data sets, which are the ground truth and the result of the algorithm, by the union of these same two sets. The following equation is used to calculate the IoU.

$$IoU = \frac{|X \cap Y|}{|X \cup Y|} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \quad (6)$$

- Sensitivity: Is the percentage of predictions correctly identified as retinal hemorrhages. It is determined by dividing the sum of correctly segmented positive cases by the sum of all positive cases in the ground truth. It is calculated with the following equation:

$$Sensitivity = \frac{TP}{TP + FN} \quad (7)$$

- Specificity: Is the percentage of predictors correctly identified as not belonging to retinal hemorrhages. It is determined by dividing the sum of correctly classified negative cases by the sum of all ground truth negative cases. It is calculated with the following equation:

$$Specificity = \frac{TN}{TN + FP} \quad (8)$$

Results and discussion

In this paper, a modified U-Net architecture proposed for the purpose of segmentation of suspicious hemorrhages in retinal images without first removing the blood vessels, optic disc, macula and fovea. U-Net architecture allows the reuse of features extracted

by the encoder network and correlates them with the results of the decoder path. This technique alleviates the degradation problem, allowing a faster convergence of the network and producing high resolution images.

The objective of this section is to discuss the obtained results and compared it with the similar published method in the literature of recent publications, which have used retinal hemorrhage segmentation approaches. We have trained and tested the model using the IDRiD dataset and the DIARETDB1 dataset. We have built a model and testing it with two different datasets to confirm the stability of our proposed approach. In the testing phase, we have used different datasets, to ensure that the result of the hemorrhages assessment is not biased due to the change of the dataset in the training and testing phases. Many similar studies on CAD system for hemorrhage diagnosis has recently been published [30, 34, 35, 37] these studies allow us make a comparison in order to assess the performance of the proposed approach.

To do so, we use a variety of performance measures that simplify the determination of the degree of effectiveness. Table 1 shows the performance obtained and Tables 2 and 3 show the results of the quantitative comparison of the different algorithms. Compared to previous techniques, we think that the results of our method are quite intriguing. The suggested model had an accuracy of 98.68%, a precision of 99.98%, a similarity coefficient (Dice) of 86.51%, an intersection over union (IoU) of 76.61%, a sensitivity of 80.49%, a specificity of 99.68%, and a loss to training data not exceeding 0.0038. Tan et al. [30] published a study using the CLEOPATRA database on a strategy for simultaneous segmentation of retinal diseases using a 10-layer convolutional neural network. For retinal hemorrhage segmentation, the proposed solution has a sensitivity of 62.57% and a specificity of 98.93%. Compared to our method, the approach has a lower sensitivity and a higher specificity. A technique for semantic segmentation of various retinal disorders was developed by [34]. The Messidor dataset is used in their technique, which is based on the Segnet architecture. Their technique had higher sensitivity and specificity than ours, with a sensitivity of 80.93% and specificity of 98.54%, respectively. Regarding accuracy, their method achieved a lower value than ours with a value of 97.86%. Orlando et al. [35] developed a method for detecting red lesions (micro-aneurysms or hemorrhages). Features are retrieved using a six-layer CNN architecture and manually developed procedures. Hemorrhage features collected from the DIARETDB1 database are fed into the random forest method, which creates a probability map to identify hemorrhages. The sensitivity obtained was lower than the sensitivity of the proposed system with a value of 48.83. Ananda et al. [37] suggested a methodology based on CNN by segmentation. They used the IDRiD and MESSIDOR databases, as well as a modified U-Net and a modified SegNet. They had a gain of 0.02% in the similarity index (Dice) compared to our approach, which had a value of 86.53%, and a higher training data loss, with a value of 0.1461, than our method.

Figure 11 shows the learning curve of the retinal hemorrhage segmentation algorithm as a function of the number of periods using the IDRiD dataset. We can observe that the accuracy value increases rapidly in the first few rounds and continues to increase with each subsequent iteration, showing that the learning process produces interesting results. The learning loss rate curve for the retinal hemorrhage segmentation algorithm as a function of the number of periods using the IDRiD dataset is shown in Fig. 12. The

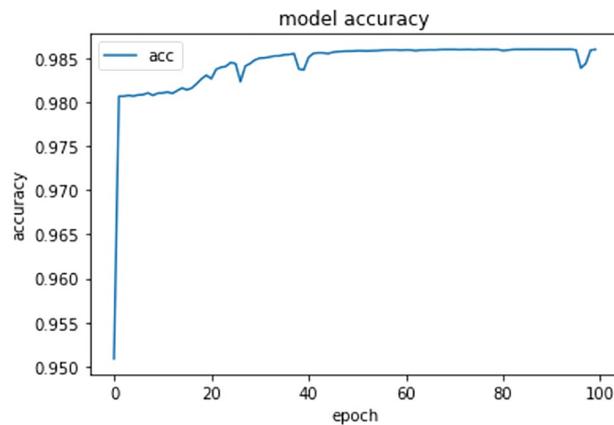


Fig. 11 Plots of the accuracy curve for the proposed method with 100 epochs

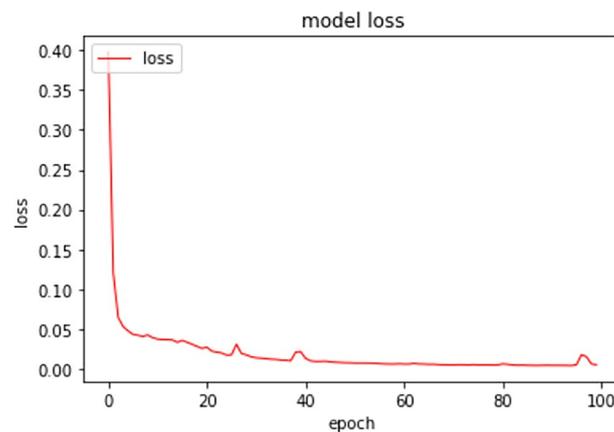


Fig. 12 Plots of the loss curve for the proposed method with 100 epochs

curve shows a substantial reduction in the learning loss rate in the first few iterations, which continues as the remaining iterations progress, giving us a modest loss value for the training data. Figure 13 represents the last part of the proposed model training procedure with 100 epochs.

We verified the proposed model using a portion of the images in the IDRiD and DIARETDB1 datasets for the model validation step. 51 of 89 fundus images in the DIARETDB1 dataset showed evidence of retinal hemorrhages according to independent labeling by four medical experts. The first column in Fig. 14 represents the original image from the validation set, the second column represents the hemorrhage label associated with the original image and the third column represents the retinal hemorrhage segmentation result obtained by the method using IDRiD as validation and test training data set.

In the validation step, we use the second dataset to demonstrate the generalizability of the model. The original DIARETDB1 dataset image is displayed in the first column of Fig. 15, the associated hemorrhage mask is displayed in the second column and the retinal hemorrhage segmentation result is displayed in the third column using the DIARETDB1 dataset as the validation set and IDRiD as the training and test set.

```

Epoch 97/100|
171/171 [=====] - 72s 422ms/step - loss: 0.0038 - acc: 0.9868
- recall: 0.7128 - precision: 0.9999 - Dice: 0.8640 - Jaccard: 0.7645

Epoch 98/100
171/171 [=====] - 72s 422ms/step - loss: 0.0038 - acc: 0.9868
- recall: 0.7127 - precision: 0.9998 - Dice: 0.8651 - Jaccard: 0.7661

Epoch 99/100
171/171 [=====] - 72s 422ms/step - loss: 0.0038 - acc: 0.9868
- recall: 0.7120 - precision: 0.9998 - Dice: 0.8639 - Jaccard: 0.7643

Epoch 100/100
171/171 [=====] - 72s 422ms/step - loss: 0.0038 - acc: 0.9868
- recall: 0.7126 - precision: 0.9998 - Dice: 0.8643 - Jaccard: 0.7649
    
```

Fig. 13 Illustration of the last part of the training process using the suggested technique with 100 epochs

In summary, the proposed method allows efficient reuse of network features, removes overfitting and mitigates the gradient phenomenon. The method will have a huge impact in avoiding future complexity of some patients. In the end, the experimental results illustrate a good advance over the state-of-the-art methods.

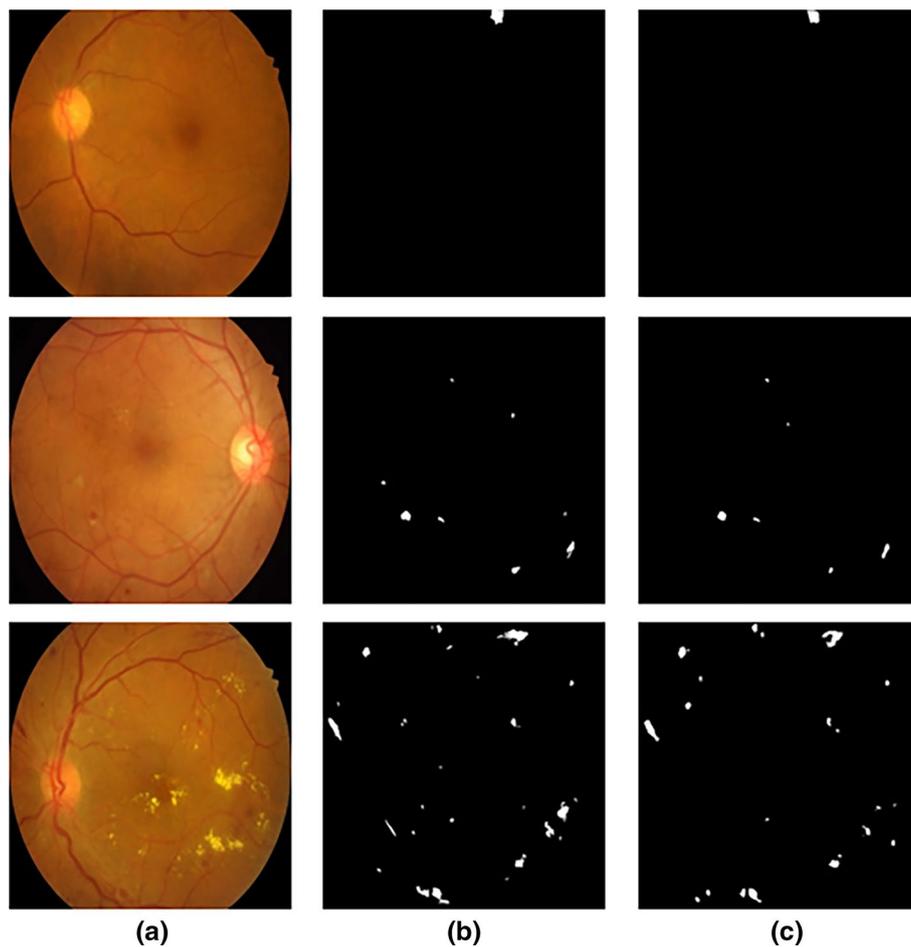


Fig. 14 Results of hemorrhages segmentation: **a** original image, **b** hemorrhages label and **c** segmented image

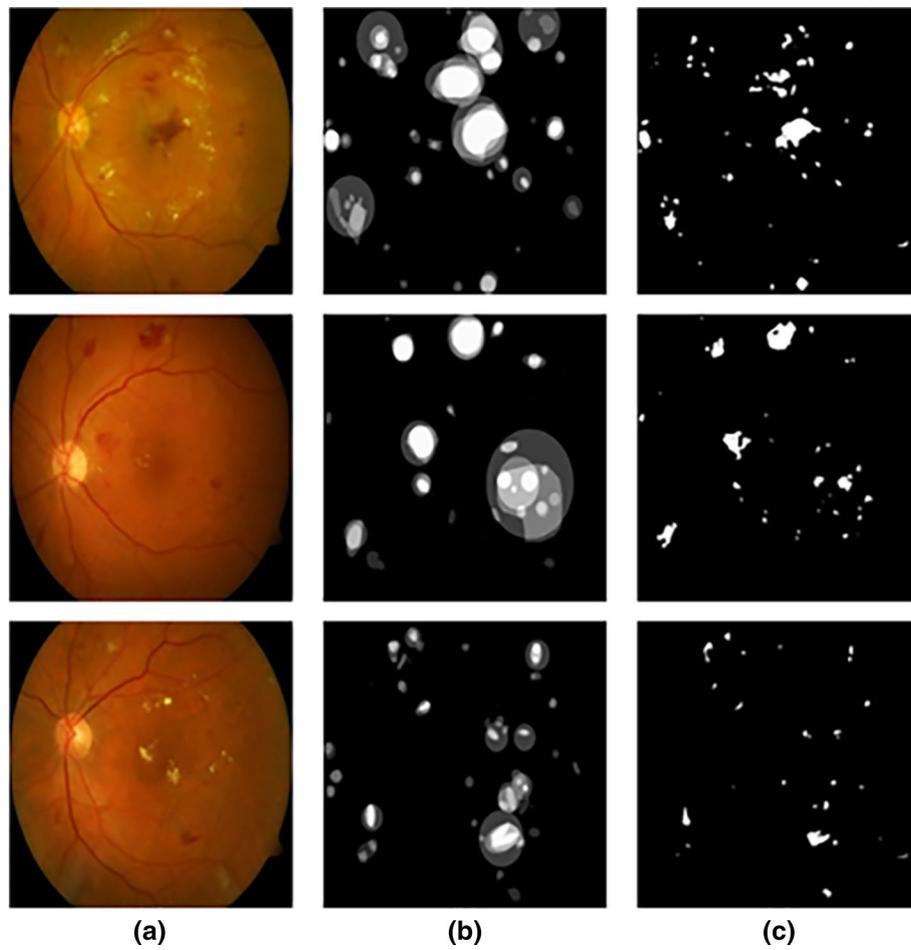


Fig. 15 Results of hemorrhages segmentation **a** original image, **b** hemorrhages label and **c** segmented image

Table 1 Performance obtained using the proposed method

Performance	Sen	Spe	Acc	Dice	IoU	Prec	LT
Avg (%)	80.49	99.68	98.68	86.51	76.61	99.98	0.0038

Sen sensitivity, Spe specificity, Acc accuracy, IoU Intersection over Union, Prec precision, LT loss for training data, Avg average

Table 2 Comparison of the results of the proposed method for retinal hemorrhage segmentation to the latest current methods in terms of accuracy, sensitivity and specificity

Authors and Refs.	Dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)
Ten et al. [30]	CLEOPATRA	62.57	98.93	—
Badar et al. [34]	MESSIDOR	80.93	98.54	97.86
Orlando et al. [35]	DIARETDB1 and E-Ophtha	48.83	—	—
Proposed	IDRID and DIARETDB1	80.49	99.68	98.68

Table 3 Comparison of the results of the proposed method for retinal hemorrhage segmentation to the latest current methods in terms of similarity coefficient (Dice), intersection over union (IoU), precision and loss for training data

Authors and Refs.	Dataset	Dice (%)	IOU (%)	Precision (%)	Loss for training data
Ananda et al. [37]	IDRID and MESSIDOR	86.53	–	–	0.1461
Proposed	IDRID and DIARETDB1	86.51	76.61	99.98	0.0038

Conclusions and future work

Due to the high degree of similarity between blood vessels, retinal hemorrhage and background in the original images, most of the methods presented in the literature do not separately address the problem of segmentation of the retinal hemorrhage. The difficulty of detecting diabetic retinopathy was addressed in this concept by segmenting the hemorrhages using digital fundus images. The goal is to automate diagnosis using the suggested approach and only digital fundus images, thus eliminating the need for additional tests. The framework of our method is to implement a modified UNet algorithm to segment the suspect region of interest. The proposal showed robustness despite the diversity of the dataset used for either training or validation. The results obtained were also compared to other results recently published in the literature and showed a significant improvement in diagnostic performance. The suggested architecture for retinal hemorrhage segmentation yields a sensitivity of 80.49%, specificity of 99.68%, accuracy of 98.68%, IoU of 76.61%, and Dice score of 86.51% when trained on the IDRiD data set and validated and tested on IDRiD and DIARETDB1. In future work, we will use the result of retinal hemorrhage segmentation to grade diabetic retinopathy. The appearance of retinal hemorrhages indicates a transition in the severity of retinopathy from minimal (grade 1) to moderate (grade 2) nonproliferative diabetic retinopathy. When retinal hemorrhages increase in all four quadrants of the retina, the classification of diabetic retinopathy becomes severe nonproliferative (grade 3). If the hemorrhages become more complex, it means that diabetic retinopathy becomes proliferative (grade 4). We also wish to extend our architecture or create a new convolutional neural network in the future, with the aim of obtaining more accurate results in automated and simultaneous segmentation tasks for various retinal disorders, such as hard and soft exudates, hemorrhages and micro-aneurysms.

Abbreviations

AI	Artificial intelligence
CAD	Computer-aided diagnosis
CLAHE	Contrast limited adaptive histogram equalization
CNN	Convolutional neural network
DL	Deep learning
FCNN	Fully convolutional neural network
FN	False negative
FP	False positive
GPU	Graphics processing unit
IDRiD	Indian diabetic retinopathy image dataset
IoU	Intersection over Union
NPDR	Non-proliferative diabetic retinopathy
ROC	Receiver operating characteristic
ReLU	Rectified linear unit

RGB	Red Green Blue
SVM	Support vector machine
TN	True negative
TP	True positive
VGG	Visual Geometry Group
WHO	World Health Organization
OCT	Optical Coherence Tomography
FFA	Fundus Fluorescein Angiography
CF	Color photography of the fundus

Acknowledgements

Not applicable.

Author contributions

AS designed and implemented the proposed model; OO, SA and AE supervised the methodology and tests. All authors read and approved the final manuscript.

Authors' information

Ayoub Skouta received his BE degree in systems, database and networks from Faculty of Science Ain Chock, Casablanca, Morocco in 2009 and ME degree in computer and internet engineering from Faculty of Science Ain Chock, Casablanca, Morocco in 2012. He is currently pursuing Ph.D. at the Department of Mathematics and Computer Science, Faculty of Science, Casablanca, Morocco. His areas of interest are segmentation and classification of fundus images for diabetic retinopathy diagnosis support by machine learning and deep learning.

Abdelali Elmoufidi (M'2012) received the Master and Ph.D. degrees, both in mathematics and computer science, from Faculty of Sciences Ain Chock, Hassan II University, Casablanca, Morocco, in 2012 and 2018, respectively. From 2019 to now, he is at the Faculty of Sciences and Techniques, Sultan Moulay Slimane University as an Associate Professor. He is currently in the Department of Computer Science, Faculty of Sciences and Techniques, Sultan Moulay Slimane University, Beni Mellal, Morocco. His research interests are related to image processing artificial intelligence and machine learning.

Said Jai Andaloussi was born in Fez (Morocco) on October 21, 1981. In 2010 he received a PhD degree in computer science jointly supervised by the Faculty of Science - University of Sidi Mohamed Ben-Abedillah, Fez, Morocco and Telecom Bretagne, Brest, France. He is a Professor of computer science in mathematics & computer science department, Faculty of Sciences Ain-chok, Casablanca, Morocco. He is the author of more than 40 papers. His current research interests include medical information processing, information modeling and analysis in medical images.

Ouail Ouchetto received a Master's degree in applied mathematics from Pierre et Marie Curie University-Polytechnic School of Paris, Paris, France, and a Master's degree in computer science from IMT Atlantique, Brest, France. He received his PhD from the Faculty of Sciences Paris-Sud (Paris-Saclay). He was a Postdoctoral Researcher at the University Blaise Pascal, Clermont-Ferrand, France, and an Assistant Professor (ATER) at the University of Paris-Sud, France. In 2010, he joined Ain Chock Hassan II University, Casablanca, Morocco, as a Professor of mathematics and computer science. His research interests include computer science, scientific computing, numerical computation, and electrical and electronic engineering.

Funding

Not applicable. This research received no specific grant from any funding agency.

Availability of data and materials

The data for this experiment are from the IDRID and DiaretDB1 datasets that are freely available to the public via links <https://iee-dataport.org/openaccess/indian-diabetic-retinopathy-image-dataset-idrid> and <http://www.it.lut.fi/project/imageret/diaretdb1/>.

Declarations

Ethics approval and consent to participate

The authors accept the Journal of Big Data's ethics approval and agree to share their work for scientific advancement.

Consent for publication

The authors hereby consent to the publication of the work in the Journal of Big Data.

Competing interests

The authors disclose that they do not have any competing interests.

Received: 4 March 2022 Accepted: 30 May 2022

Published online: 13 June 2022

References

1. Nagaraj P, Deepalakshmi P, Romany FM. Artificial flora algorithm-based feature selection with gradient boosted tree model for diabetes classification. *Diabetes Metab Syndrome Obes Targets Ther.* 2021;14:2789.
2. Zhang Y-H, Guo W, Zeng T, Zhang S, Chen L, Gamarra M, Mansour RF, Escorcia-Gutierrez J, Huang T, Cai Y-D. Identification of microbiota biomarkers with orthologous gene annotation for type 2 diabetes. *Front Microbiol.* 2021;12:1927.

3. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst.* 2012;25:1097–105.
4. Elmoufidi A, Skouta A, Jai-Andaloussi S, Ouchetto O. CNN with multiple inputs for automatic glaucoma assessment using fundus images. *Int J Image Graph.* 2022. <https://doi.org/10.1142/S0219467823500122>.
5. Pournaras CJ. *Pathologies vasculaires oculaires.* Issy-les-Moulineaux: Elsevier Masson; 2008.
6. Amin J, Sharif M, Yasmin M. A review on recent developments for detection of diabetic retinopathy. *Scientifica.* 2016. <https://doi.org/10.1155/2016/6838976>.
7. Mansour RF. Evolutionary computing enriched computer-aided diagnosis system for diabetic retinopathy: a survey. *IEEE Rev Biomed Eng.* 2017;10:334–49.
8. Devaraj D, Suma R, Kumar SP. A survey on segmentation of exudates and microaneurysms for early detection of diabetic retinopathy. *Mater Today Proc.* 2018;5(4):10845–50.
9. He Y, Carass A, Liu Y, Jedynek BM, Solomon SD, Saidha S, Calabresi PA, Prince JL. Deep learning based topology guaranteed surface and MME segmentation of multiple sclerosis subjects from retinal OCT. *Biomed Opt Express.* 2019;10(10):5042–58.
10. Shanthi T, Sabeenian R. Modified Alexnet architecture for classification of diabetic retinopathy images. *Comput Electr Eng.* 2019;76:56–64.
11. Skouta A, Elmoufidi A, Jai-Andaloussi S, Ouchetto O. Automated binary classification of diabetic retinopathy by convolutional neural networks. In: *Advances on smart and soft computing.* Singapore: Springer; 2021. p. 177–87.
12. El Hossi A, Skouta A, Elmoufidi A, Nachaoui M. Applied CNN for automatic diabetic retinopathy assessment using fundus images. In: *International conference on business intelligence.* Springer; 2021. p. 425–33.
13. Zhang Z, Liu Q, Wang Y. Road extraction by deep residual U-Net. *IEEE Geosci Remote Sens Lett.* 2018;15(5):749–53.
14. Noori M, Bahri A, Mohammadi K. Attention-guided version of 2D UNet for automatic brain tumor segmentation. In: *2019 9th international conference on computer and knowledge engineering (ICCKE).* IEEE; 2019. p. 269–75.
15. Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2014. p. 580–7.
16. Elmoufidi A, El Fahssi K, Jai-Andaloussi S, Sekkaki A, Gwenole Q, Lamard M. Anomaly classification in digital mammography based on multiple-instance learning. *IET Image Process.* 2018;12(3):320–8.
17. Elmoufidi A, El Fahssi K, Jai-Andaloussi S, Madrane N, Sekkaki A. Detection of regions of interest's in mammograms by using local binary pattern, dynamic k-means algorithm and gray level co-occurrence matrix. In: *2014 international conference on next generation networks and services (NGNS).* IEEE; 2014. p. 118–23.
18. Jai-Andaloussi S, Sekkaki A, Quellec G, Lamard M, Cazuguel G, Roux C. Mass segmentation in mammograms by using bidimensional empirical mode decomposition BEMD. In: *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC).* IEEE; 2013. p. 5441–4.
19. Elmoufidi A, El Fahssi K, Jai-Andaloussi S, Sekkaki A. Detection of regions of interest in mammograms by using local binary pattern and dynamic K-means algorithm. *Int J Image Video Process Theory Appl.* 2014;1(1):2336–0992.
20. Jafarzadeh Ghouschi S, Ranjbarzadeh R, Najafabadi SA, Osgooei E, Tirkolaei EB. An extended approach to the diagnosis of tumour location in breast cancer using deep learning. *J Ambient Intell Hum Comput.* 2021. <https://doi.org/10.1007/s12652-021-03613-y>.
21. El-Dahshan E-SA, Mohsen HM, Revett K, Salem A-BM. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. *Expert Syst Appl.* 2014;41(11):5526–45.
22. Ranjbarzadeh R, Bagherian Kasgari A, Jafarzadeh Ghouschi S, Anari S, Naseri M, Bendeche M. Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Sci Rep.* 2021;11(1):1–17.
23. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA.* 2016;316(22):2402–10.
24. Ghouschi SJ, Ranjbarzadeh R, Dadkhah AH, Pourasad Y, Bendeche M. An extended approach to predict retinopathy in diabetic patients using the genetic algorithm and fuzzy C-means. *BioMed Res Int.* 2021. <https://doi.org/10.1155/2021/5597222>.
25. Gardner GG, Keating D, Williamson TH, Elliott AT. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *Br J Ophthalmol.* 1996;80(11):940–4.
26. Sinthanayothin C, Boyce JF, Williamson TH, Cook HL, Mensah E, Lal S, Usher D. Automated detection of diabetic retinopathy on digital fundus images. *Diabet Med.* 2002;19(2):105–12.
27. Kande GB, Savithri TS, Subbaiah PV. Automatic detection of microaneurysms and hemorrhages in digital fundus images. *J Digit Imaging.* 2010;23(4):430–7.
28. Tang L, Niemeijer M, Reinhardt JM, Garvin MK, Abramoff MD. Splat feature classification with application to retinal hemorrhage detection in fundus images. *IEEE Trans Med Imaging.* 2012;32(2):364–75.
29. Van Grinsven MJ, van Ginneken B, Hoyng CB, Theelen T, Sánchez CI. Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. *IEEE Trans Med Imaging.* 2016;35(5):1273–84.
30. Tan JH, Fujita H, Sivaprasad S, Bhandary SV, Rao AK, Chua KC, Acharya UR. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Inf Sci.* 2017;420:66–76.
31. Quellec G, Charrière K, Boudi Y, Cochener B, Lamard M. Deep image mining for diabetic retinopathy screening. *Med Image Anal.* 2017;39:178–93.
32. Karkuzhali S, Manimegalai D. Retinal haemorrhages segmentation using improved toboggan segmentation algorithm in diabetic retinopathy images. *Biomed Res.* 2018. <https://doi.org/10.4066/biomedicalresearch.29-17-820>.
33. Lam C, Yu C, Huang L, Rubin D. Retinal lesion detection with deep learning using image patches. *Investig Ophthalmol Vis Sci.* 2018;59(1):590–6.
34. Badar M, Shahzad M, Fraz M. Simultaneous segmentation of multiple retinal pathologies using fully convolutional deep neural network. In: *Annual conference on medical image understanding and analysis.* Springer; 2018. p. 313–24.

35. Orlando JI, Prokofyeva E, Del Fresno M, Blaschko MB. An ensemble deep learning based approach for red lesion detection in fundus images. *Comput Methods Programs Biomed.* 2018;153:115–27.
36. Saha O, Sathish R, Sheet D. Fully convolutional neural network for semantic segmentation of anatomical structure and pathologies in colour fundus images associated with diabetic retinopathy. arXiv preprint. 2019. [arXiv:1902.03122](https://arxiv.org/abs/1902.03122).
37. Ananda S, Kitahara D, Hirabayashi A, Reddy KUK. Automatic fundus image segmentation for diabetic retinopathy diagnosis by multiple modified U-Nets and SegNets. In: 2019 Asia-Pacific signal and information processing association annual summit and conference (APSIPA ASC). IEEE; 2019. p. 1582–8.
38. Guo S, Li T, Kang H, Li N, Zhang Y, Wang K. L-Seg: an end-to-end unified framework for multi-lesion segmentation of fundus images. *Neurocomputing.* 2019;349:52–63.
39. Yan Z, Han X, Wang C, Qiu Y, Xiong Z, Cui S. Learning mutually local-global u-nets for high-resolution retinal lesion segmentation in fundus images. In: 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019). IEEE; 2019. p. 597–600.
40. Huang Y, Lin L, Li M, Wu J, Cheng P, Wang K, Yuan J, Tang X. Automated hemorrhage detection from coarsely annotated fundus images in diabetic retinopathy. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI). IEEE; 2020. p. 1369–72.
41. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 1–9.
42. Sambal N, Saini P, Syal R, Gupta V. Modified U-Net architecture for semantic segmentation of diabetic retinopathy images. *Biocybern Biomed Eng.* 2020;40(3):1094–109.
43. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina. *Prog Retinal Eye Res.* 2018;67:1–29.
44. Atienza R. Advanced deep learning with TensorFlow 2 and Keras: apply DL, GANs, VAEs, deep RL, unsupervised learning, object detection and segmentation, and more. Birmingham: Packt Publishing Ltd; 2020.
45. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol.* 1962;160(1):106–54.
46. Ranjbarzadeh R, Jafarzadeh Ghoushchi S, Bendeche M, Amirabadi A, Ab Rahman MN, Baseri Saadi S, Aghamohammadi A, Kooshki Forooshani M. Lung infection segmentation for COVID-19 pneumonia based on a cascade convolutional network from CT images. *BioMed Res Int.* 2021. <https://doi.org/10.1155/2021/5544742>.
47. Badar M, Haris M, Fatima A. Application of deep learning for retinal image analysis: a review. *Comput Sci Rev.* 2020;35: 100203.
48. Mansour RF. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Biomed Eng Lett.* 2018;8(1):41–57.
49. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer: Cham; 2015. p. 234–41.
50. Xiuqin P, Zhang Q, Zhang H, Li S. A fundus retinal vessels segmentation scheme based on the improved deep learning U-Net model. *IEEE Access.* 2019;7:122634–43.
51. Skouta A, Elmoufidi A, Jai-Andaloussi S, Ouchetto, O. Semantic segmentation of retinal blood vessels from fundus images by using CNN and the random forest algorithm; 2022.
52. Porwal P, Pachade S, Kamble R, Kokare M, Deshmukh G, Sahasrabudhe V, Meriaudeau F. Indian diabetic retinopathy image dataset (IDRIID): a database for diabetic retinopathy screening research. *Data.* 2018;3(3):25.
53. Kauppi T, Kalesnykiene V, Kamarainen J-K, Lensu L, Sorri I, Raninen A, Voutilainen R, Uusitalo H, Kälviäinen H, Pietilä J. The diaretdb1 diabetic retinopathy database and evaluation protocol. In: *BMVC*, vol. 1. 2007. p. 1–10.
54. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint. 2014. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.