Retinal blood vessel segmentation using a deep learning method based on modified U-NET model

Sanjeewani¹, Arun Kumar Yadav², Mohd Akbar³, Mohit Kumar⁴, Divakar Yadav⁵

¹*sanjeewani@gmail.com*, ²*ayadav@nith.ac.in*, ³*mohdakbar1971997@gmail.com*,

⁴*mohit@nith.ac.in*, ⁵*dsy99@rediffmail.com*,

^{1,2,4,5}Dept of Computer Science & Engineering, NIT Hamirpur (HP), India

³Dept of Computer Science & Engineering, AKGEC, Ghaziabad(UP), India

Abstract

Automatic retinal blood vessel segmentation is very crucial to ophthalmology. It plays a vital role in the early detection of several retinal diseases such as Diabetic Retinopathy, hypertension, etc. In recent times, deep learning based methods have attained great success in automatic segmentation of retinal blood vessels from images. In this paper, a U-NET based architecture is proposed to segment the retinal blood vessels from fundus images of the eye. Furthermore, 3 pre-processing algorithms are also proposed to enhance the performance of the system. The proposed architecture has provided significant results. On the basis of experimental evaluation on the publicly available DRIVE data set, it has been observed that the average accuracy (Acc) is .9577, sensitivity (Se) is .7436, specificity (Sp) is .9838 and F1-score is .7931. The proposed system outperforms all recent state of art approaches mentioned in the literature.

Segmentation Deep learning DRIVE CNN U-NET

1 INTRODUCTION

Retinal blood vessel network plays an important role in identification of many diseases like diabetic retinopathy, occlusion, glaucoma, hypertensive retinopathy and stroke, etc. The images of macular area of eyes provides necessary information to detect the discussed diseases [9]. In general, change of pattern of blood vessels are observed to detect the said diseases. Diabetic Retinopathy (DR) is Type-II disease that when diagnosed early, can save a patient's vision. In DR, new blood vessels are developed in the eye that causes damage to the blood circulation in the retina. The advanced stages of DR are characterized by damaged blood vessels, growth of new vessels and abnormal blood vessels in the retina [25, 7]. It makes outer layer of the retina thick and blood streams into retina that causes blood clotting in vessels and hemorrhage in the eye [45]. DR is classified into four stages :- Mild Non-proliferative Retinopathy, Moderate Non-proliferative Retinopathy, Severe Non-proliferative Retinopathy, Proliferative Retinopathy. Major complications that may result from DR include microaneurysms, cotton wool spots, bright and dark lesions, exudates, glaucoma and retinal detachment [30, 37]. In the paper[15], a detailed study has been carried out and commented that in the next 2 decades, more than 700 million diabetic patients will be diagnosed. About one third of the patients have signs of DR and one third of them have visual issues [39].

Blood vessel segmentation of retinal images has attracted the researchers from last one decade due to its marvelous success in the context of detecting DR related disease at an early stage. A small change in vascular structure generates a lot of abnormal states in blood vessels. The abnormal retinal images are very useful for diagnosing, but for better comparison healthy vessel images are also required [24]. The basic steps in diagnosing the DR is the extraction of blood vessels from fundus images that requires segmentation of retinal blood vessel. Segmentation process involves many issues during its execution for separation of infected blood vessel objects. For instance, a retinal image may contain

information about other objects that may relate to other diseases like microaneurysm and retinal detachment. To extract blood vessel manually requires the help of experts and it also consumes a lot of time. Object detection cannot be done without separating foreground & background and illumination factor also leads to increase in detection error [52].

Computerized methods are the promising alternative for segmentation of retinal fundus images for diagnosis of the retinal diseases. Further, computerized methods also reduce the challenges discussed previously such as improved diagnosis efficiency and reduced time of diagnosis as compared to manual diagnosis [4]. In the past few years, a lot of research has been carried out for segmentation of retinal images and validated in various studies in literature [51]. For the reasons here stated, the segmentation of retinal images can be a valuable aid for the detection and follow-up of several pathological states as well as a means of detecting and characterizing over time any of the changes in the blood vessels. In the current era, artificial intelligence is being employed in multiple disciplines of research such as healthcare, medicine, agriculture, etc. In recent times, machine learning and deep learning approaches have consistently performed remarkably well in the field of medical image analysis.

This paper introduces a new deep learning based method for retinal image segmentation that uses the convolutional neural network architecture. The proposed method works based on the popular U-NET architecture. The detailed description of the proposed model is explained in section 3.3. To increase the segmentation accuracy, the original U-NET architecture has been modified.

The proposed modifications revamp the evaluation parameters of proposed architecture that help the segmentation of the fundus images to detect the diseases. The model is evaluated and compared with state-of-art on DRIVE dataset, that is commonly used for this type of research work. The major contributions of proposed work is listed as follows:

- A masking algorithm is proposed to extract regions of interest relating to retinal blood vessel during preprocessing phase (Algorithm-1).
- An image enhancement algorithm is proposed to enhance the quality of image for better classification (Algorithm-2).
- An algorithm to crop images in 64X64 pixels is proposed for obtaining better accuracy and lower training/testing time in the proposed architecture (Algorithm-3).
- It provides more accurate results in lower time since input images used are of resolution 64x64.
- The activation function used in this architecture (ELU) works well for both positive and negative values, that increases its accuracy level.
- The proposed architecture provides accurate results for the segmentation of retinal fundus images for the mentioned database. This detection functionality can also be implemented for real dataset/images for automatic disease detection.

The rest of the paper is organised into following sections. Section 2 discusses the recent contributions in the literature with focus on image processing, machine learning and deep learning methods. Section 3 delineates the proposed pre-processing algorithms and model. Section 4 lists the performance evaluation metrics. Section 5 discusses analysis the results obtained. Section 6 compares the results of the proposed approach with recent state of art methods. Section 7 presents the conclusion and future scope of the work.

2 Literature survey

Medical image segmentation and medical image registration are ongoing research areas. Computational approaches are increasingly becoming the focus for automatic retinal blood vessel segmentation. In the past, a lot of research has been carried out for automatically segmenting the retinal

blood vessels[12, 29, 23, 36]. This section discussed the examined works from literature on retinal blood vessel segmentation using image processing (matched filter, morphological and vessel tracking methods), machine learning and deep learning methods as shown in Figure 1.

Figure 1: Classification of various retinal image segmentation approaches

In the paper [44], the authors use a region-based approach with distance regularisation term for segmenting of retinal blood vessels of fundus images. This method yielded an accuracy of 95.48% on DRIVE dataset. Again, in the paper [23], the authors examined the vision of people more than 49 years old. In this study, the authors evaluate the relationship between open-angle glaucoma (OAG) and retinal vessel diameter among baseline participants in the Blue Mountains Eye. Again, in the study [55], uses conditional random field based model with and got 94.69% accuracy, 96.74% specificity, 80.78% sensitivity and 79.42% F1-score on DRIVE dataset. The Gabor filter and Gaussian fractional derivative have been proposed in the paper [1] to significantly enhance performance for both blood vessel structure as well as its outlines detection. This approach obtains accuracy 95.03% and sensitivity 78.54% on DRIVE dataset. The matched filtering approach with multi-wavelet kernels (MFMK) is proposed in the paper [46]. This study evaluates the experiment on DRIVE dataset with accuracy 94.61%. In the paper [33], the authors proposed a Frechet PDF based Matched Filter Approach for Retinal Blood Vessels Segmentation for blood vessel segmentation. This study obtained an accuracy of 95.44%. In the paper [22], the authors proposed the detection of center lines or Morphological Reconstruction for retinal vessels segmentation. It attained accuracy of 94.63% on DRIVE dataset. In the paper [36], the authors used particle swarm optimization technique to determine the best filter parameters for multi-scale Gaussian matched filter, resulting in enhanced retinal vascular segmentation accuracy. Table 1 presents a summary of image processing based methods for retinal image segmentation.

Machine learning methods based segmentation include techniques based on support vector machines, K-Nearest Neighbour, Random Forests and Gaussian Mixture Models. In the paper [42], the authors introduced an overlapping-block-based method in retinal blood vessel segmentation with classification through support vector machines. This model achieved accuracy 95.38%, sensitivity 67.49% and specificity 97.73% on DRIVE dataset. A Hidden Markov model (HMM) based approach is proposed in [53] to quickly extract the primary features of vessels. The method is tested on DRIVE dataset with accuracy 94.75%, sensitivity 72.62% and specificity 98.03%. In the paper [50], the authors proposed a unique technique for identifying retinal vessels depending on radial projection using a semi-supervised algorithm. The proposed method is tested on DRIVE dataset with accuracy 94.34%, sensitivity 74.10% and specificity 94.51%. In the paper [18], the authors proposed presents a method based on distance comparison between the test patch and cluster centres to determine the minimal distance using unsupervised method and Frangi filter. This model is tested on DRIVE dataset with accuracy 95.4%, sensitivity 77.4% and specificity 98%. In the paper [11], the authors discuss the various Ensemble Classification-Based Approach used for segmentation in

Table 1: Summary of Image Processing based methods for retinal blood vessel segmentation

S.No.	Year/ citation	Model/Architecture
1	2005[23]	Retinal Vessel Diameter and Open-Angle Glaucoma
2	2006[22]	Detection of Centerlines and Morphological Reconstruction
3	2013[46]	Retinal vessel segmentation using multiwavelet kernels
4	2015[36]	Improved multiscale matched filter using PSO algorithm
5	2017[55]	Segmentation with feature learning & in-vessel enhancement
6	2018[47]	Saliency detection model and region optimization
7	2018[2]	Segmentation using Gabor filters, fractional derivatives
8	2020[44]	Segmentation using level set method
9	2020[33]	Fréchet PDF based Matched Filter

Table 2: Summary of Machine Learning based methods for retinal blood vessel segmentation

S.No.	Year/ citation	Model/Architecture
1	2006[35]	2-D Gabor Wavelet and Supervised Classification
2	2011 [50]	Radial projection and semi-supervised approach
3	2012[11]	Ensemble Classification-Based Approach
4	2015[17]	Morphological Component Analysis
5	2015[31]	Iterative Vessel Segmentation of Fundus Images
6	2016[28]	Fully Connected Conditional Random Field Model
7	2017[42]	Segmentation using Support Vector Machine Classification
8	2017[6]	Gradient approximation
9	2018[18]	Segmentation using mapping Functions
10	2020[53]	weighted line detector and hidden Markov model
11	2021[5]	Support vector machine with CNN

different categories. Also, they compare the results of algorithms on publicly available dataset. The MCA algorithm is proposed in the paper [17] for extracting retinal blood vessel segmentation and tested on DRIVE dataset with accuracy 95.25%, sensitivity 75.24% and specificity 97.53%. Again, in the paper [28], the authors proposed a model based on discriminatively trained and fully connected conditional random field yielding sensitivity 78.50%, specificity 96.70%, precision 77.70% and F1 score is 78.10%. In the study [31], the authors proposed an iterative vessel segmentation method of fundus images. The model was evaluated on DRIVE dataset with accuracy of 94.9%, sensitivity of 73.9% and specificity is 97.8%. Table 2 presents a summary of machine learning based methods for retinal image segmentation.

Unlike conventional supervised methods, deep learning based methods internally generate mathematical representation of the vascular structure. In the paper [14], the authors proposed a technique for segmenting vessels from fundus images based on deep learning. The method uses a deep neural network based on a reduced form of the U-NET architecture that integrates residual blocks with batch normalisation and provides an accuracy of 95.75%, sensitivity of 85.94% and specificity of 97.06% on DRIVE dataset. Again in the paper[48], the authors proposed a three stage deep learning model to separate thick and thin vessels differently. A U-NET featuring atrous convolution (AA-UNet) is proposed in [20] that leads a model that distinguishes vessel and non-vessel pixels and allows reuse of the deep features. This model was tested on DRIVE dataset and got an accuracy of 95.58%. A cross-connected convolutional neural network (CCNet) is presented in the paper[10] towards automated retinal vascular tree segmentation. Convolutional layers in CCNet extract features and forecast pixel classes based on learned characteristics. The model is tested on DRIVE dataset with accuracy 95.28% and sensitivity is 76.25%. To deal with the changing width and orientation of the vascular structure within retinal images, the authors of the paper [27] proposed a solution

Table 3: Summary of deep Learning based methods for retinal blood vessel segmentation

S.No.	Year/ citation	Model/Architecture
1	2016[19]	Cross-Modality Learning Approach
2	2016[13]	fully-connected conditional random fields
3	2018[48]	A Three-stage Deep Learning Model
4	2018[27]	Fully Convolutional Neural Networks
5	2019[26]	Scale-space approximated convolutional neural networks
6	2019[16]	Round-wise features aggregation
7	2019[43]	Dense U-net Based on Patch-Based Learning
8	2020[20]	Attention Guided U-Net with Atrous Convolution
9	2020[10]	A cross-connected convolutional network
10	2020[41]	RESwNet for retinal small vessel segmentation
11	2021[14]	convolutional kernels and modified U-Net model
12	2021[3]	Fully convolutional deep learning architecture
13	2021[32]	Vessel Specific Skip chain Convolutional Network
14	2021[34]	Encoder Enhanced Atrous (EEA) Unet

that uses existing multi-scale analysis given by a Stationary Wavelet Transform with a multi-scale Fully Convolutional Neural Network. The accuracy of this model is 95.76% as tested on DRIVE dataset. A line detector and hidden Markov model based blood vessel segmentation have been proposed in the paper [54]. The model is evaluate in DRIVE dataset with accuracy of 94.75%. In the research paper [56], the authors presented an innovative index called as Fusion Score (FS) that offers a holistic evaluation for binary images. It also uses a combination of size-invariant extracted features and dense connectivity to boost the standard CNN's learning abilities. This approach yields an of accuracy 95.2% on DRIVE dataset. Regarding retinal vascular segmentation, the authors in [26] propose the multi-scale convolutional neural network (CNN) architecture called as scale-space approximated CNNs (SSANet). The model is evaluated on DRIVE dataset and accuracy 95.69%, sensitivity 83.54% and specificity 97.46% is observed. A cross modality learning method has been proposed in the paper[19]. The model was tested on DRIVE dataset and accuracy 95.2%, sensitivity 75.69% and specificity 98.16% was attained on DRIVE dataset. In the paper [43], the authors proposed a dense U-NET architecture with patch based learning method that yielded an accuracy of 95.11%, sensitivity 79.86% and specificity 97.36% on DRIVE dataset. In the paper [41], the authors utilize U-NET with residual block and densely connected convolution and got an accuracy of 95.54% on DRIVE dataset. The authors of the paper [13] proposed Retinal vessel segmentation via deep learning and conditional random field. This approach provided accuracy 94.70% and sensitivity 72.94% on DRIVE dataset. Table 3 presents a summary of deep learning based methods for retinal image segmentation.

On the basis of literature review, it is observed that a lot of research has been carried out using image processing, machine learning and deep learning based methods for retinal image segmentation. However, a lot of gap is present for improvement in image segmentation performance. This survey has identified the following significant points. Firstly, most of the researchers used public datasets for image segmentation and achieved reasonable accuracy. It is required to try to further improve the accuracy. Secondly, some of the researchers used only limited number of evaluation metrics to report the results of segmentation but they have not commented on all of the available evaluation parameters. Thirdly, it has been observed that deep learning methods perform better than other segmentation methods. On the basis of these points, the following research points (RPs) have been defined:

- Is there any possibility to improve current architectures to achieve better results?
- Is there any possibility to enhance the quality of input images to obtain better results?

- Is there any possibility to reduce training and testing time of modified model?

3 Proposed Work

On the basis of the literature review and subsequent findings, a deep learning based model/method for retinal image segmentation is proposed. The model employs a modified version of the popular U-NET architecture along with proposed pre-processing algorithms in order to segment the input images. Figure 2 depicts the schematic diagram of the proposed method.

Figure 2: Schematic illustration of the proposed architecture for retinal image segmentation

The outline of the proposed method is as follows. Firstly, the dataset is pre-processed so that it is ready to be fed into the model. The pre-processing steps include colour conversion of the images followed by the image enhancement step. Finally, the last step of pre-processing is the cropping of the images into the shape of 512x512 pixels. After this last step, the images can be input into the model. However, U-NET or any CNN-based model does not perform best directly on large sized images. It may require a huge amount of training time along with a large number of hidden layers in the model. Therefore, in order to train the model efficiently and utilize the dataset effectively, every image in the dataset is divided into 64 parts with size 64x64. The training of the model is done with these images, followed by testing in the same manner for model evaluation. These steps are explained in detail in the following sections, beginning with the dataset description.

3.1 Dataset

There are many publicly available datasets that are relevant to retinal image segmentation task. However, most of them suffer from the problem that they do not have manually annotated ground truth labels. In this work, DRIVE [38] (Digital Retinal Images for Vessel Extraction) is used to train as well as evaluate the performance of the proposed model. The DRIVE dataset is a publicly available dataset that has been widely used by researchers, experts and doctors in regard to diagnosing and detecting eye-related diseases. The images in the dataset were obtained by a Canon CR5 non-mydratic 3-CCD camera with a 45-degree field of view (FOV). Each image is taken using 8 bits per colour plane at 565x584 pixel resolution. The FOV of each image is circular with a diameter of approximately 540 pixels. The dataset contains manually segmented ground truth labels by a vascular structure expert for every image. The dataset has a total of 40 colour fundus images, that is, the images were obtained by using a fundus camera to capture the image of inner surface of eye.

For the experiments in this study, the dataset is partitioned into two parts: training and testing. The training as well as testing portion contains 20 images each. Out of the total 40 images, 7 images correspond to unhealthy eyes. Such images contain exudates, haemorrhages, pigment, abnormal blood vessels, etc.

Figure 3: Sample images from DRIVE Dataset (a)Color image (b)Ground truth image (c)Mask image

3.2 Preprocessing

Pre-processing is an initial step that is followed in order to ready the input data to be fed into the model. This step also helps to clean the input data, remove irregularities present in the data, etc. such that the model can learn better [8]. Image processing [21] techniques are employed to reduce noise, resize images, contrast enhancement and image sharpening. The next subsections discuss the specific pre-processing steps that have been deployed in the current work.

Figure 4: Workflow for image pre-processing & enhancement

3.2.1 RGB to Gray Conversion

Firstly, the images are converted from colour to grayscale images. Using colour images directly into a learning algorithm or model may create unnecessary complexities of handling 3 channels of red, green and blue components while not helping much in the overall segmentation process. They also take more memory while processing. Therefore, grayscale images are obtained from colour images

as a first pre-processing step. For this step, the function `image.convert()` of the python PILLOW library has been utilized. The resulting grayscale images consist of only a single channel instead of the original three channels in the colour images. Such grayscale images may be represented as 2D matrices containing values from 0 to 255 where 0 represents absolute black and 255 corresponds to white.

Figure 5: Conversion of color image to grayscale (a) Color image (b) grayscale image

3.2.2 Image Enhancement

To improve the visual quality [49] or appearance of images, enhancement techniques are employed. This step is incorporated because it is expected that it will help to distinguish and therefore, segment the blood vessels better. Furthermore, enhancement techniques are used in medical images to reduce noise and improve sharpening details for better representation. In this study, three algorithms (Algorithms 1, 2 & 3) have been developed for image enhancement that is suited to the current task. The result of application of the algorithms is depicted in figure 6.

Algorithm 1 is used to extract the region of interest from the input image using a mask image. The basic idea behind this procedure is that while scanning the input image in a pixel-wise manner, if a pixel value is 0 (black) in mask, the corresponding pixel value in the input image is also set to 0. The other pixels of the input image retain their original values. In this way, the region of interest is extracted.

Algorithm 2 is used to enhance the contrast of the image. The output image from Algorithm 1 is the input to this algorithm. First, the minimum (min) and maximum (max) pixel values are obtained in the input image. Next, every pixel value is reduced by (min-1). Then the contrast factor is obtained as $255/\text{max}$. Finally, every pixel value is multiplied by the contrast factor. This enhanced image is the output of Algorithm 2.

In Algorithm 3, first the input image, which is the output from Algorithm 2, is cropped to 512x512 dimensions. Next, in order to partition the input image, boundary points are calculated, i.e., (0,0), (0,64), (0,128), ... and so on. Finally, from a single image of 512x512, 64 images of dimensions 64x64 are created by partitioning.

It is important to mention here that Algorithm 3 is also applied to the ground-truth (labels) images whereas Algorithm 1 & 2 are not applied because the ground-truth images are binary images and need no enhancement or colour conversion.

Algorithm 1 - Extract Region of Interest through Mask Image

```
1: image = Original Image, mask = Mask Image
2: for i in (0,565):
3:     for j in (0,584):
4:         if mask[i][j] == 0:
5:             image[i][j] = 0
6:         end if
7:     end for
8: end for
9: return image
```

Algorithm 2 - Image Enhancement

```
1: image = Output Image of Algorithm 1
2: min = Minimum intensity of image //Between 0-255
3: max = Maximum intensity of image //Between 0-255
4: for i in (0,565):
5:     for j in (0,584):
6:         image[i][j] = image[i][j] - (min-1)
7:     end for
8: end for
9: contrastFactor = 255 / max
10: for i in (0,565):
11:     for j in (0,584):
12:         image[i][j] = image[i][j] *contrastFactor
13:     end for
14: end for
15: return image
```

Algorithm 3 - Image Cropping and Partitioning

```
1: image = Image after applying Algorithm – 2
2: width = 565, height = 584, newWidth = 512, newHeight = 512
3: left = (width-newWidth) / 2, top = (height-newHeight) / 2
4: right = (width+newWidth) / 2, bottom= (height+newHeight) / 2
5: image = image.crop(left, top, right, bottom)
6: C_Img = [ ]
7: for i in range (0, 512, 64):
8:     for j in range (0,512,64):
9:         dim = [ ]
10:        dim.append(i), dim.append(j), C_Img.append(dim)
11:    end for
12: end for
13: List (Image) partitionedImages = [ ]
14: for i in range (0,64):
15:     partitionedImage= image.crop(C_Img[i][0], C_Img[i][1], C_Img[i][0] + 64, C_Img[i][1] + 64)
16:     partitionedImages.append(partitionedImage)
17: end for
18: return partitionedImages
```

Figure 6: Enhanced image (a) Grayscale image (b) Enhanced image

Figure 7: Partitioning of each image into 64 patches of dimension 64X64

3.3 Segmentation

It has been observed by researchers that early symptoms of numerous diseases such as diabetes, hypertension, etc. may be correlated with the health of retinal blood vessels. Manual segmentation of retinal blood vessels is a time taking task while also requiring the help of domain experts. This has prompted the need for automatic segmentation of retinal blood vessels. Consequently, segmentation has attracted the attention of many researchers experts. It has become one of the most researched topic in medical image analysis since the results of segmentation process are useful in almost all other areas in the medical domain.

The primary focus in this work is to extract the blood vessels from retinal images obtained from fundus camera images. Retinal vasculature is made up of hollow pipes of different sizes. It comprises of arteries, capillaries, veins and venules. The proposed model in this study (Figure 8) is a modified version of the original U-NET architecture that has become a benchmark in biomedical image segmentation. The U-NET model classifies the pixels of an input image into either a vessel pixel or a non-vessel pixel and therefore extracts the retinal blood vessels from the input image. The original U-NET architecture consists of two main components. The left half is the first component of the model that is called the contraction path or the encoder path while the right half is the second component of the model that is called the expansion path or the decoder path. The encoder comprises of convolution layers and max pooling layers. The decoder comprises of up-sampling layers along with convolution layers.

In the proposed architecture, the input images are firstly cropped to a size of 512X512. Next,

Figure 8: Proposed network architecture for retinal image segmentation

each image is divided into 64 patches of resolution 64X64 before feeding in to the main architecture. The depth of the proposed architecture is 4. As the input image moves through the network, it is reduced from 64X64 to 4X4 before the up-sampling phase begins. The proposed architecture produces [16, 32, 64, 128, 256]-sized feature maps at successive layers that is significantly lower than original U-NET architecture. This reduces the number of trainable parameters remarkably while also providing improvements in the performance. A succession of 2 convolutions are performed at every layer. Throughout this network, 3X3 convolutions are performed. Therefore, we have an initial image of 64X64. From this image, 16 filters are extracted twice, according to the convolutional step. Let us call this output C1 (output of convolution layer at depth 1). Since we use padding in the input image, the size of the image after the convolution steps does not change. Next, max pooling is performed and the size of the image is reduced from 64X64 to 32X32. Let us call this output P1. C1 gets passed from this layer to the decoder part at the same depth. P1 is passed to the next layer and the same process of successive convolutions is repeated at each level of depth. Now, C1, C2, C3, C4 & C5 are the outputs after the successive convolution operation and P1, P2, P3, P4 & P5 are the outputs after the max pooling operation.

For up-sampling, the image size gets double at every step, going from 4X4 to 64X64. Up-sampling performs 2X2 transposed convolutions with stride value 2 and padding setting ‘same’ with number of feature maps being halved each time. This operation doubles the input map dimensions. C1, C2, C3, C4 gets concatenated from encoder part to the symmetrically equivalent layer on the decoder part. Skip connections helps the network to extract features in various levels of detail and provides output of higher quality with localization information. Skip connections combine low level high resolution features and high level low resolution features. Finally, an output is produced according to binary classification of pixels: a pixel being either a vessel pixel or a non-vessel pixel. Similarly, predictions are generated for all the pixels in the image and thereby creating the retinal blood vessel structure.

The choice of activation function to be used in the network may have serious implications on the performance of the network. In this proposed model, ‘ELU’ activation function has been employed. It stands for Exponential Linear Unit. In contrast to the ‘RELU’ activation function, that is a popular choice among the community, the ‘ELU’ also takes negative values. Another activation function that was applied in the experiments is sigmoid. It takes real values as input and generates an output value between 0 and 1. It fulfils all the properties desirable of an activation function while also providing output in form of a probability.

Another important parameter that has been used in the proposed architecture is the padding.

Table 4: Confusion matrix

	Vessel pixel	Non-Vessel pixel
Predicted Vessel pixel	True Positive	False Positive
Predicted Non-Vessel pixel	False Negative	True Negative

As opposed to the original U-NET architecture, an extra layer of zero pixels were added in the experiments. The setting used was `padding = 'same'` that means that extra pixels on the edges are to be added so the output image dimensions are same as the input image. The choice of optimizer is another important variable. In the proposed work, 'Adam' optimizer was employed. A fundamental element in the success of a model is the loss function. It can impact the results to a significant degree. In this study, `binary cross entropy` loss function was deployed.

4 Performance Measure

The evaluation of performance of the proposed model is based on the following metrics: accuracy, sensitivity (recall), specificity (precision) and F1 score [40]. Each input image is classified, on pixel level, into one of the two categories: a vessel pixel or a non-vessel pixel. Each pixel may be classified as follows:

- **True positive(TP)** – When the pixel is recognized as a vessel pixel in ground truth image as well as segmented image.
- **True negative(TN)** – When the pixel is recognized as a non vessel pixel in both segmented image and ground truth image.
- **False positive(FP)** – When the pixel is recognized as vessel pixel in segmented image but is actually a non-vessel pixel in ground truth image.
- **False negative(FN)** – When a pixel is recognized as a non-vessel pixel in segmented image but is actually a vessel pixel in ground truth image.

4.1 Confusion matrix

A confusion matrix is a performance measurement technique that helps to understand how well a classification model performs, as show in Table 4. The objective is to explain the performance of the model from a variety of perspectives and not just using the accuracy prediction.

4.2 Accuracy

Accuracy can be calculated by measuring how often the model predicts a pixel correctly, that is, a vessel pixel is predicted as a vessel pixel and a non-vessel pixel is predicted as a non-vessel pixel. It is a function of the total correctly classified pixels of an input image.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

4.3 Sensitivity

Sensitivity or Recall is calculated by how often the model misdiagnosed a vessel pixel as a non-vessel pixel. It is a function of the ability of a method to recognize vessel pixels correctly.

$$Sensitivity = \frac{TP}{TP + FN}$$

4.4 Specificity

Specificity or Precision is calculated by measuring the number of times the model wrongly classifies a non-vessel pixel as a vessel pixel. It refers to the ability of a model to recognize the non-vessel pixels.

$$Specificity = \frac{TN}{TN + FP}$$

4.5 F-score

The F-score provides a decent overall estimate of how well a particular model is performing. It sums up a classifier's performance using a single score. It is calculated as depicted in the following expression.

$$F - score = \frac{2TP}{2TP + FP + FN}$$

5 Results and Analysis

To achieve as accurate results as possible, a series of experiments were performed with various hyper-parameter values. Initially, six (6) different models were tried. The description of these initial models along with their accuracy, precision, recall and F-score is shown in Table 5.

Table 5: Performance of initial models on 128X128 images taken from DRIVE dataset

S. No.	Model	Validation split	Acc	Se	Sp	F1-Score
1	Two Times Contrast(M 1)	0.3	.9352	.6076	.9841	.7092
2	Two Times Contrast (M 2)	0.2	.9414	.6717	.9818	.7490
3	Two Times Contrast (M 3)	0.1	.9438	.7262	.9763	.7706
4	Contrast images (M 4)	0.3	.9405	.6772	.9799	.7476
5	Contrast Images (M 5)	0.2	.9417	.6844	.9801	.7533
6	Contrast Images (M 6)	0.1	.9435	.7200	.9770	.7685

In these experiments, the Python library PILLOW is utilized for image colour conversion and image enhancement. For the models M1, M2 & M3, the images are contrasted 2 times using a contrast filter whereas they are contrasted only a single time for the models M4, M5 & M6. The input image sizes for the models were 384*384. However, before feeding the images into the network, they were resized into 128*128 resolution. Further, the validation split for the models M1 & M4 was 0.3, M2 & M5 was 0.2 and M3 & M6 was 0.1. The decision score for all the mentioned models was 0.4. With the above mentioned settings, encouraging results were obtained that provided useful insights based on which 3 subsequent models were developed.

- **Model 1:** In this model, the input image is converted to grayscale by using the `image.convert()` function of the PILLOW library available in Python. Next, images are enhanced and cropped using the proposed Algorithms 1, 2 & 3. Finally, before passing to the model, the images are divided into patches of size 64*64. The decision score is set to 0.4. With this model, the best specificity results were obtained. This indicates that model can efficiently acknowledge the non-vessel pixels in the input image.
- **Model 2:** In this model, the input image is converted to grayscale by using the `image.convert()` function similar to Model 1. Next, images are enhanced by applying histogram equalization. Finally, before passing to the model, the images are divided into patches of size 64*64. The decision score is set to 0.5.

- **Model 3:** This model is similar to Model 1 except that the decision score is set to 0.5. With this model, the best overall results were obtained.

Table 6: Performance of final models on 64X64 images with various decision scores on DRIVE dataset

S. No.	Model	Acc	Se	Sp	F1-Score
1	Enhanced cropped Patches with D-score - 0.4 on improved U-Net(Model 1)	.9391	.4948	.9935	.6396
2	Equalised cropped Patches with D-score - 0.5 on improved U-Net(Model 2)	.9550	.7260	.9831	.7791
3	Enhanced cropped Patches with D-score - 0.5 improved U-Net(Model 3)	.9577	.7436	.9838	.7931

Table 6 illustrates the details of the above discussed models along with the results obtained in terms of accuracy, sensitivity, specificity and F-score. It may be noted that out of the three models, two (M1 & M3) provide the best results in at least one metric but Model 3 has the best overall performance.

6 Comparison with state of the art

In order to emphasize the effectiveness of the proposed model, this section compares the present study and its results with the state of the art approaches mentioned in literature.

Table 7 depicts the comparison with the state of art approaches. It may be noted from the comparison that the proposed model outperforms all the recent approaches discussed in literature in terms of accuracy and specificity. Furthermore, the comparison in terms of accuracy, specificity and sensitivity are shown in figures 9, 10 & 11 respectively.

Figure 9: Comparison of Accuracy of proposed method with state of art on DRIVE dataset

Table 7: Comparison of the proposed approach with state of art on the DRIVE dataset (- indicates that result does not computed)

S.No.	YEAR	MODEL	Acc	Se	Sp	F1-Score
1	2011[50]	Radial projection	.9434	.7309	.9761	-
2	2013[46]	Multiwavelet kernels	.9461	-	-	-
3	2017[42]	Support vector machine	.9538	.6749	.9773	-
4	2018[18]	Mapping functions via training patches	.9540	.7740	.9800	-
5	2018[1]	Gabor filter fractional derivative	.9503	-	.7854	-
6	2019[26]	Scale-space approximated CNN	.9569	.8354	.9746	-
7	2019[43]	Dense U-NET	.9511	.7986	.9736	-
8	2020[41]	U-NET with residual blocks	.9554	-	-	-
9	2020[20]	Attention guided U-NET	.9558	-	-	-
10	2020[56]	Convolutional network with dense connectivity	.9569	-	-	.7866
11	2020[10]	Cross connected convolutional network	.9528	.7625	.9809	-
12	2020[53]	Robust method based on weighted line detector	.9535	.6749	.9773	-
13	2020[33]	Frechet PDF based matched filter	.9544	.7307	.9761	-
14	2021[14]	U-Net model	.9547	-	-	-
15	This study	Enhanced cropped Patches with D-score - 0.5 on improved U-Net	.9577	.7436	.9838	.7931

Figure 10: Comparison of Specificity of proposed method with state of art on DRIVE dataset

7 Conclusion and future work

The main focus of this paper is segmentation of the vascular tree structure from the fundus images of eyes. The objective of segmentation is to divide an image into areas that have a relatively uniform visual appearance or correspond to objects or portions of things. In this paper, a model based on modified version of U-NET architecture is introduced that achieves state of art performance on the stated task. The model works on extracted patches of images obtained after the pre-processing stage. The proposed model is utilized for segmentation and classifying each image pixel into vessel pixel

Figure 11: Comparison of Sensitivity of proposed method with state of art on DRIVE dataset

and non-vessel pixel. Using the proposed model, 95.77% accuracy, 98.38% specificity and 79.31% F-score is obtained on the DRIVE dataset. In future researchers can use improved optimization techniques, CNN models to improve accuracy or parameter selection.

References

- [1] Hugo Aguirre-Ramos, Juan Gabriel Avina-Cervantes, Ivan Cruz-Aceves, José Ruiz-Pinales, and Sergio Ledesma. Blood vessel segmentation in retinal fundus images using Gabor filters, fractional derivatives, and Expectation Maximization. *Applied Mathematics and Computation*, 339:568–587, 2018.
- [2] Hugo Aguirre-Ramos, Juan Gabriel Avina-Cervantes, Ivan Cruz-Aceves, José Ruiz-Pinales, and Sergio Ledesma. Blood vessel segmentation in retinal fundus images using gabor filters, fractional derivatives, and expectation maximization. *Applied Mathematics and Computation*, 339:568–587, 2018.
- [3] İbrahim Atli and Osman Serdar Gedik. Sine-net: A fully convolutional deep learning architecture for retinal blood vessel segmentation. *Engineering Science and Technology, an International Journal*, 24(2):271–283, 2021.
- [4] George Azzopardi, Nicola Strisciuglio, Mario Vento, and Nicolai Petkov. Trainable cosfire filters for vessel delineation with application to retinal images. *Medical image analysis*, 19(1):46–57, 2015.
- [5] Kishore Balasubramanian and NP Ananthamoorthy. Robust retinal blood vessel segmentation using convolutional neural network and support vector machine. *Journal of Ambient Intelligence and Humanized Computing*, 12(3):3559–3569, 2021.
- [6] Shouvik Chakraborty, Sankhadeep Chatterjee, Nilanjan Dey, Amira S Ashour, and Fuqian Shi. Gradient approximation in retinal blood vessel segmentation. In *2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)*, pages 618–623. IEEE, 2017.
- [7] Argyrios Christodoulidis, Thomas Hurtut, Houssem Ben Tahar, and Farida Cheriet. A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images. *Computerized Medical Imaging and Graphics*, 52:28–43, 2016.
- [8] Muhammed Gökhan Cinsdikici and Doğan Aydin. Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm. *Computer Methods and Programs in Biomedicine*, 96(2):85–95, 2009.
- [9] Abdolhossein Fathi and Ahmad Reza Naghsh-Nilchi. Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation. *Biomedical Signal Processing and Control*, 8(1):71–80, 2013.
- [10] Shouting Feng, Zhongshuo Zhuo, Daru Pan, and Qi Tian. CcNet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features. *Neurocomputing*, 392:268–276, 2020.
- [11] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman. Blood vessel segmentation methodologies in retinal images - A survey. *Computer Methods and Programs in Biomedicine*, 108(1):407–433, 2012.
- [12] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R Rudnicka, Christopher G Owen, and Sarah A Barman. Blood vessel segmentation methodologies in retinal images—a survey. *Computer methods and programs in biomedicine*, 108(1):407–433, 2012.
- [13] Huazhu Fu, Yanwu Xu, Stephen Lin, Damon Wing Kee Wong, and Jiang Liu. Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9901 LNCS(DI):132–139, 2016.

- [14] Manuel E. Gegundez-Arias, Diego Marin-Santos, Isaac Perez-Borrero, and Manuel J. Vasallo-Vazquez. A new deep learning method for blood vessel segmentation in retinal images based on convolutional kernels and modified U-Net model. *Computer Methods and Programs in Biomedicine*, 205:106081, 2021.
- [15] Stewart B Harris, Jordan W Tompkins, and Braden TeHiwi. Call to action: a new path for improving diabetes care for indigenous peoples, a global review. *Diabetes research and clinical practice*, 123:120–133, 2017.
- [16] Cam-Hao Hua, Thien Huynh-The, and Sungyoung Lee. Retinal vessel segmentation using round-wise features aggregation on bracket-shaped convolutional neural networks. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 36–39. IEEE, 2019.
- [17] Elaheh Imani, Malihe Javidi, and Hamid Reza Pourreza. Improvement of retinal blood vessel detection using morphological component analysis. *Computer Methods and Programs in Biomedicine*, 118(3):263–279, 2015.
- [18] Ahsan Khawaja, Tariq M. Khan, Khuram Naveed, Syed Saud Naqvi, Naveed Ur Rehman, and Syed Junaid Nawaz. An Improved Retinal Vessel Segmentation Framework Using Frangi Filter Coupled with the Probabilistic Patch Based Denoiser. *IEEE Access*, 7:164344–164361, 2019.
- [19] Qiaoliang Li, Bowei Feng, Linpei Xie, Ping Liang, Huisheng Zhang, and Tianfu Wang. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Transactions on Medical Imaging*, 35(1):109–118, 2016.
- [20] Yan Lv, Hui Ma, Jianian Li, and Shuangcai Liu. Attention Guided U-Net with Atrous Convolution for Accurate Retinal Vessels Segmentation. *IEEE Access*, 8:32826–32839, 2020.
- [21] M. Vijaya Maheswari and G. Murugeswari. A Survey on Computer Algorithms for Retinal image Preprocessing and Vessel Segmentation. *Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020*, pages 403–408, 2020.
- [22] Ana Maria Mendonça and Aurélio Campilho. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Transactions on Medical Imaging*, 25(9):1200–1213, 2006.
- [23] Paul Mitchell, Harry Leung, Jie Jin Wang, Elena Rochtchina, Anne J Lee, Tien Y Wong, and Ronald Klein. Retinal vessel diameter and open-angle glaucoma: the blue mountains eye study. *Ophthalmology*, 112(2):245–250, 2005.
- [24] Asloob Ahmad Mudassar and Saira Butt. Extraction of blood vessels in retinal images using four different techniques. *Journal of Medical Engineering*, 2013, 2013.
- [25] Uyen TV Nguyen, Alauddin Bhuiyan, Laurence AF Park, and Kotagiri Ramamohanarao. An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern recognition*, 46(3):703–715, 2013.
- [26] Kyoung Jin Noh, Sang Jun Park, and Soochahn Lee. Scale-space approximated convolutional neural networks for retinal vessel segmentation. *Computer Methods and Programs in Biomedicine*, 178:237–246, 2019.
- [27] Américo Oliveira, Sérgio Pereira, and Carlos A. Silva. Retinal vessel segmentation based on Fully Convolutional Neural Networks. *Expert Systems with Applications*, 112:229–242, 2018.
- [28] José Ignacio Orlando, Elena Prokofyeva, and Matthew B Blaschko. A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. *IEEE transactions on Biomedical Engineering*, 64(1):16–27, 2016.

- [29] John Alan Richards and JA Richards. *Remote sensing digital image analysis*, volume 3. Springer, 1999.
- [30] Somasis Roy, Anirban Mitra, Sudipta Roy, and Sanjit Kumar Setua. Blood vessel segmentation of retinal image using clifford matched filter and clifford convolution. *Multimedia Tools and Applications*, 78(24):34839–34865, 2019.
- [31] Sohini Roychowdhury, Dara D Koozekanani, and Keshab K Parhi. Iterative vessel segmentation of fundus images. *IEEE Transactions on Biomedical Engineering*, 62(7):1738–1749, 2015.
- [32] Pearl Mary Samuel and Thanikaiselvan Veeramalai. Vssc net: vessel specific skip chain convolutional network for blood vessel segmentation. *Computer Methods and Programs in Biomedicine*, 198:105769, 2021.
- [33] Sushil Kumar Saroj, Rakesh Kumar, and Nagendra Pratap Singh. Fréchet PDF based Matched Filter Approach for Retinal Blood Vessels Segmentation. *Computer Methods and Programs in Biomedicine*, 194, 2020.
- [34] Vallikutti Sathananthavathi and G Indumathi. Encoder enhanced atrous (eea) unet architecture for retinal blood vessel segmentation. *Cognitive Systems Research*, 67:84–95, 2021.
- [35] João VB Soares, Jorge JG Leandro, Roberto M Cesar, Herbert F Jelinek, and Michael J Cree. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Transactions on medical Imaging*, 25(9):1214–1222, 2006.
- [36] KS Sreejini and VK Govindan. Improved multiscale matched filter for retina vessel segmentation using pso algorithm. *Egyptian Informatics Journal*, 16(3):253–260, 2015.
- [37] Joes Staal, Michael D Abràmoff, Meindert Niemeijer, Max A Viergever, and Bram Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging*, 23(4):501–509, 2004.
- [38] Joes Staal, Michael D. Abràmoff, Meindert Niemeijer, Max A. Viergever, and Bram Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.
- [39] Alan W Stitt, Timothy M Curtis, Mei Chen, Reinhold J Medina, Gareth J McKay, Alicia Jenkins, Thomas A Gardiner, Timothy J Lyons, Hans-Peter Hammes, Rafael Simo, et al. The progress in understanding and treatment of diabetic retinopathy. *Progress in retinal and eye research*, 51:156–186, 2016.
- [40] Nasser Tamim, M. Elshrkawey, Gamil Abdel Azim, and Hamed Nassar. Retinal blood vessel segmentation using hybrid features and multi-layer perceptron neural networks. *Symmetry*, 12(6), 2020.
- [41] Yu Tang, Zhiyuan Rui, Changfeng Yan, Jingjun Li, and Jingpeng Hu. Reswnet for retinal small vessel segmentation. *IEEE Access*, 8:198265–198274, 2020.
- [42] Eva Tuba, Lazar Mrkela, and Milan Tuba. Retinal blood vessel segmentation by support vector machine classification. *2017 27th International Conference Radioelektronika, RADIOELEKTRONIKA 2017*, 2017.
- [43] Chang Wang, Zongya Zhao, Qiongqiong Ren, Yongtao Xu, and Yi Yu. Dense u-net based on patch-based learning for retinal vessel segmentation. *Entropy*, 21(2):168, 2019.
- [44] Chuang Wang and Yongmin Li. Blood Vessel Segmentation from Retinal Images. *Proceedings - IEEE 20th International Conference on Bioinformatics and Bioengineering, BIBE 2020*, (3):759–766, 2020.

- [45] Shuangling Wang, Yilong Yin, Guibao Cao, Benzhenh Wei, Yuanjie Zheng, and Gongping Yang. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, 149:708–717, 2015.
- [46] Yangfan Wang, Guangrong Ji, Ping Lin, and Emanuele Trucco. Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition. *Pattern Recognition*, 46(8):2117–2133, 2013.
- [47] Lan-Yan Xue, Jia-Wen Lin, Xin-Rong Cao, and Lun Yu. Retinal blood vessel segmentation using saliency detection model and region optimization. *Journal of Algorithms & Computational Technology*, 12(1):3–12, 2018.
- [48] Zengqiang Yan, Xin Yang, and Kwang-Ting Cheng. A three-stage deep learning model for accurate retinal vessel segmentation. *IEEE journal of Biomedical and Health Informatics*, 23(4):1427–1436, 2018.
- [49] Dan Yang, Guoru Liu, Mengcheng Ren, Bin Xu, and Jiao Wang. A multi-scale feature fusion method based on U-net for retinal vessel segmentation. *Entropy*, 22(8), 2020.
- [50] Xinge You, Qinmu Peng, Yuan Yuan, Yiu Ming Cheung, and Jiajia Lei. Segmentation of retinal blood vessels using the radial projection and semi-supervised approach. *Pattern Recognition*, 44(10-11):2314–2324, 2011.
- [51] Bob Zhang, Lin Zhang, Lei Zhang, and Fakhri Karray. Retinal vessel extraction by matched filter with first-order derivative of gaussian. *Computers in biology and medicine*, 40(4):438–445, 2010.
- [52] Yu Qian Zhao, Xiao Hong Wang, Xiao Fang Wang, and Frank Y Shih. Retinal vessels segmentation based on level set and region growing. *Pattern Recognition*, 47(7):2437–2446, 2014.
- [53] Chao Zhou, Xiaogang Zhang, and Hua Chen. A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden Markov model. *Computer Methods and Programs in Biomedicine*, 187:105231, 2020.
- [54] Chao Zhou, Xiaogang Zhang, and Hua Chen. A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden markov model. *Computer methods and programs in biomedicine*, 187:105231, 2020.
- [55] Lei Zhou, Qi Yu, Xun Xu, Yun Gu, and Jie Yang. Improving dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin-vessel enhancement. *Computer methods and programs in biomedicine*, 148:13–25, 2017.
- [56] Zhongshuo Zhuo, Jianping Huang, Ke Lu, Daru Pan, and Shouting Feng. A size-invariant convolutional network with dense connectivity applied to retinal vessel segmentation measured by a unique index. *Computer Methods and Programs in Biomedicine*, 196, 2020.