Dilated Deep Neural Architectures for Improving Retinal Vessel Extraction

V. Sathananthavathi[1] · **G. Indumathi[1]**

Abstract

Retinal vascular region is recognized as the promising anatomical region for the diagnosis of several commonly seen diseases including cardiovascular related and diabetes. In this paper we propose two novel deep neural architectures named as Dilated fully convolved convolutional neural network (FCNN) and dilated depth concatenated neural network (DCNN) to segment the retinal blood vessels. The proposed work is evaluated for both the proposed architectures with and without dilation. It is observed from the obtained results that dilation enhances the network performance. To eliminate the non-uniform illumination and low contrast differences effect the preprocessed images are used for training the architectures. The proposed methodologies are experimented on the two publicly available databases DRIVE and STARE database. The proposed dilated FCNN architecture can able to obtain high accuracy of about 95.39% which is high compared to the FCNN architecture. For the dilated DCNN architecture also, accuracy obtained is about 96.16% which is high compared to DCNN. The experimental results reveal the significance of dilation operation in improving the semantic segmentation of retinal blood vessels.

Keywords Retinal color fundus images · Deep neural network · Dilation · Fully convolved neural network · Depth concatenation · Entropy loss function

1 Introduction

Human eye is an absolute sensory organ for vision. Eye sight is entirely accomplished by the blood flow in retinal vessels in eye. Diseases such as diabetes retinopathy, hypertension and arteriosclerosis cause change in branching pattern and retinal blood vessel thickness leading to blindness. Hence primary requirement for screening and diagnosis of retinal diseases involves segmentation of retinal blood vessels. Two dimensional or three dimensional colour fundus images are used for segmenting retinal vessel. In early days analyst or experts manually segment the retinal vessel from the fundus images which is a time consuming progress and it leads to error. Automatic segmentation of retinal blood

✉ V. Sathananthavathi
  sathananthavathi@gmail.com

1 Department of Electronics & Communication Engineering, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu 626005, India

vessels reduces time and it helps in early detection and diagnosis of diseases. Improvement in computer aided system leads to development of various methods for retinal vessel segmentation. Several supervised or unsupervised methods have already been proposed for retinal vessel segmentation. Supervised methods rely on manual annotation groundtruth for segmentation of retinal vessel whereas unsupervised method does not depend on manual annotation.

Salazar Gonzalez et al. proposed an unsupervised graph cut based retinal vessel segmentation method [1] in which the input images are preprocessed by adaptive histogram equalization and distance transform. Recently, retinal vessel segmentation by unsupervised methodologies based on Frangi enhancement filter [2], ICA enhancement technique based method [3], line detector based methodology [4], Gaussian filtering with morphological processing [5] and also probabilistic tracking based method [6] are proposed. The retinal vascular region segmentation are chronically with similar intensity range for both background and vascular region and non-uniform brightness and contrast. These issues lead to detect the background region as vascular regions by the unsupervised method. Supervised methods can able to overcome the drawback of unsupervised methods by handcrafting the features describe vessel regions. Zhu et al. proposed a supervised method [7] by taking thirty nine discrete features for each fundus image comprising morphological feature, hessian feature and so on. The extracted feature together with ground truth is given as input to the extreme machine learning classifier and the image is classified as vessel and non-vessel region. High efficiency of neural network based methods lead to its implication for medical image processing. Marin et al, applied feed forward neural network as a classifier for retinal vessel extraction with gray and moment invariant features [8]. In many approaches the feature selection for segmentation of image is strenuous task as segmented result depends on features selected. Influence of significant features need to be considered for the vessel extraction are mentioned in the optimization based feature selection method [9] for the retinal vessel segmentation.

In [10] sumathi et al. proposed a robust algorithm for segmentation of retinal vessels. The input images are preprocessed to normalize the illumination and contrast problem and thirteen dimension feature vectors are extracted. Probability map are formed using neural network classifier and the images are classified as vessel region and non-vessel region. In [11] Shelhamer et al. proposed a transfer learning approach for retinal vessel segmentation using Alexnet. The conventional fully connected layer and classification layer on the output layer are modified to form a fully convolutional neural network.

The significance of both supervised and unsupervised methods are incorporated in some works [11, 12] to obtain the improved performance at the cost of increasing computational burden. In [12], radius based clustering algorithm is enhanced with a partial supervision strategy to segment blood vessels of small diameter and thickness. The thin vessel extraction is attempted through mean matting [13] method comprising both supervised and unsupervised methods.

Convolutional neural network process the image similar to that of human by applying feature engineering. Hence convolutional neural network was applied as an eminent feature extractor in some works [14–19]. Orlando et al. [14] segmented the retinal colour fundus images using convolutional neural network for feature extraction and classified the image using Support vector machine classifier. The input images are preprocessed to eliminate the contrast and illumination problem. Wang et al proposed convolutional neural network as a hierarchical feature extractor [15] and Random forest as a classifier for retinal vessel extraction. In [16] Liskowski et al. proposed a deep neural network based approach in which the input images are preprocessed with global contrast normalization. The neural

network is trained with four thousand samples which are data augmented using geometric transformation and gamma correction. In addition to the above surveys, several deep learning based methods [17–19] are also proposed for retinal blood vessel extraction.

Fu et al. proposed neural network based retinal vessel segmentation method [20] in which Convolutional neural network learns the features and generates probability map and segment the vessel by pixel concurrence using conditional random field. Zhexin et al. proposed transfer learning based approach [21] for retinal vessel segmentation and then the obtained result are post processed to improve the accuracy of segmented result. In [22] Olaf et al. proposed neural network architecture for segmenting medical images. The architecture consists of multiple hidden layers; the output of each hidden stage is up sampled and concatenated to previous layer output before upsampling. Dasgupta et al. [23] proposed a method for segmentation of retinal vessel in which input images are preprocessed by extracting green channel of fundus image to remove non uniform illumination and then the image is divided into patches and trained using fully convolved neural network. The images are then classified using probability map by setting threshold.

In [24], Qiaoliang et al. proposed a cross modality approach for retinal vessel segmentation. Mapping function is used to obtain the corresponding vessel map. The neural network consists of five layers viz, one input layer, output layer and three hidden layers with a sigmoid activation function consisting of 756 units. The neural network is trained with green channel image and it is transformed to vessel label.

Olivera et al. [25] proposed a method for retinal vessel segmentation in which multiscale analysis is performed by combining Stationary wavelet transform and fully convolutional neural network. The green channel of the images is taken and is its mean and variance is normalized. The remaining channels are stationary wavelet transformed and are then trained using neural network.

The fully convolved neural networks are successful in the semantic segmentation of medical images. As per retinal vessel segmentation is concerned, extraction of thin infinitesimal vessel is highly challenging. Several papers [26–30] are focused on modification of the fully convolved networks for enhancing the thin vessel extraction. The modification attempted on the fully convolved networks include adding residual blocks, multi paths, drop out or varying the number of layers. The cost for achieving the thin vessel extraction is done by increasing the computational burden. Hence it is essential to improve the thin vessel extraction without increasing the computational burden is to be focused.

The background intensity range coincides with the thin vessel region, therefore it is highly challenging to extract them. Dilation operation improves the field of view in the convolution operation [31]. Increasing the convolutional filter field of view helps to distinguish vessel region with the background. Even the field of view in the convolutional filter is increased, the computational complexity remain the same. In the proposed work, retinal vessels are segmented by two dilated convolutional neural network architectures derived from [22] and [32] respectively. The contribution of the proposed work includes the development of dilated fully convolved neural network and dilated Depth concatenated neural network architectures for the retinal vessel extraction. In order to verify the performance improvement by dilation operation, retinal vessel extraction with dilation factor of 1 is also experimented.

The proposed method has the following contribution and novelty:

- Development of dilated neural network architectures for retinal vessel segmentation.
- Enhancing the thin vessel extraction by the dilated convolution.
- Performance improvement is achieved without increasing the computational burden.

- Performance improvement is verified through comparing the dilated convolutional network result with non-dilated convolutional network.

Paper is organized as: Sect. 2 describes the proposed methodology. It describes the outline, preprocessing, dilation convolution, proposed architectures and class imbalance problem. Section 3 describes the results and discussion. It includes the database description, evaluation metric and experimental results. Section 4 describes the conclusion and future work.

2 Proposed Methodology

2.1 Outline

In the proposed work two different architectures of neural networks are proposed one is derived from segnet architecture [32] and it is named as fully convolved convolutional neural network (FCNN), the other is derived from unet architecture [22] and it is named as depth concatenated neural network (DCNN). The neural network consists of input layer, hidden layer for feature extraction and output layer for classification. Both the proposed architecture consists of five stages of encoder and decoder layers. All the convolutional layers in the proposed architecture are replaced by dilated convolutional layer. The ultimate significance of the dilated convolution is the large field of observation. It is inferred from our experimentation that convolution computed using the non-local neighbor pixels improves the classification performance.

In this work, the input images are preprocessed to eliminate illumination and low contrast between vessels and background. Features are extracted using multiscale CNN and are classified as vessel and non-vessel region using pixel classification. The vessel region is less compared to non-vessel region this leads to imbalance of pixel count between vessel and non-vessel region. The imbalance of pixel count is eliminated by cross entropy loss function thereby reducing misclassification of classes.

2.2 Preprocessing

Non uniform illumination and low contrast difference between vessels and non-vessels lead to improper segmentation of vessel. Retinal color fundus images are preprocessed to eliminate contrast and illumination problem. The green channel of the fundus image provides better contrast and can reduce illumination problem hence it is considered as the input image. The extracted green channel fundus images are further preprocessed for vessel enhancement. In the proposed work Contrast limited adaptive histogram equalization (CLAHE) is applied for vessel enhancement. CLAHE is a local enhancement technique that limits the amplification of histogram by clipping at a predefined value. To enable the vascular region more visible, CLAHE is the most commonly used technique in medical images. The preprocessed fundus image is shown in Fig. 1.

2.3 Dilated Convolution

Dilated convolution differs from conventional convolution layer by providing space between variables in kernel during convolution operation [31]. Dilation enlarges the
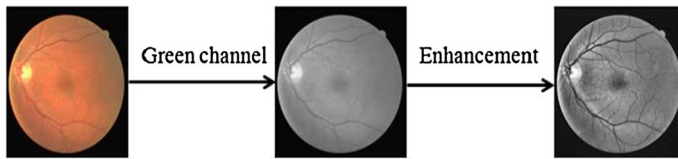
Fig. 1 Preprocessed image

receptive field without increasing the quantity of parameters or computation. The dilated convolution operation is given in Eq. 1.

$$(F *_l k)(p) = \sum_{s+lt=p} F(s)K(t) \tag{1}$$

In dilated convolution, a kernel of size k × k filter is enlarged to a filter of size k + (k−1) (r−1), where r is the dilated rate. It keeps the same resolution by allowing adaptable aggregation of the multiscale information. The pictorial representation of dilation working principle with 3 × 3 filter is shown in Fig. 2.

2.4 Fully Convolutional Neural Network

The proposed FCNN architecture is extracted from segnet architecture [32], which consists of five stages of encoder and decoder containing twenty six layers of convolution with two convolution layers in the first two stages of encoder and decoder and three in the next three stages with constant kernel size 7 × 7. The proposed FCCN architecture consists of five stages of encoder and decoder and it is shown in Fig. 3. All the stages of encoder consist of two dilated Convolution layer, each convolution layer is followed by batch normalization and rectified linear unit. Convolution operation performed in the proposed FCCN involves dilated filter of 3 × 3 with striding and padding of one. Batch normalization layer are included to increase the speed of training and rectified linear unit is used as the activation function. The output of each encoder hidden stage is unpooled and is convolved twice.
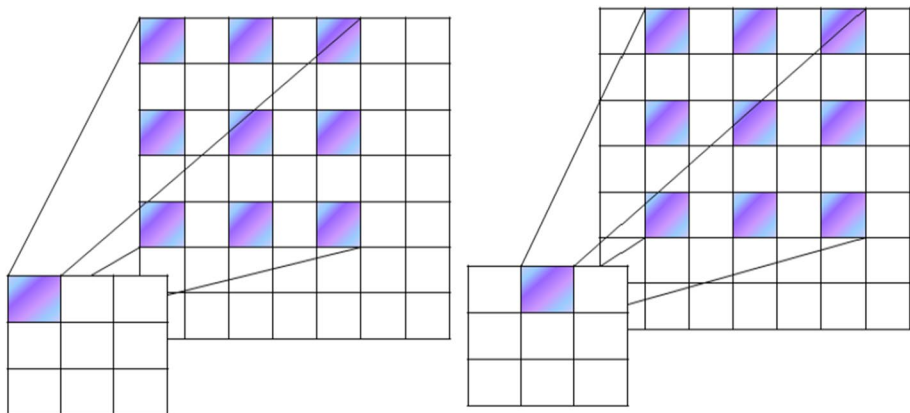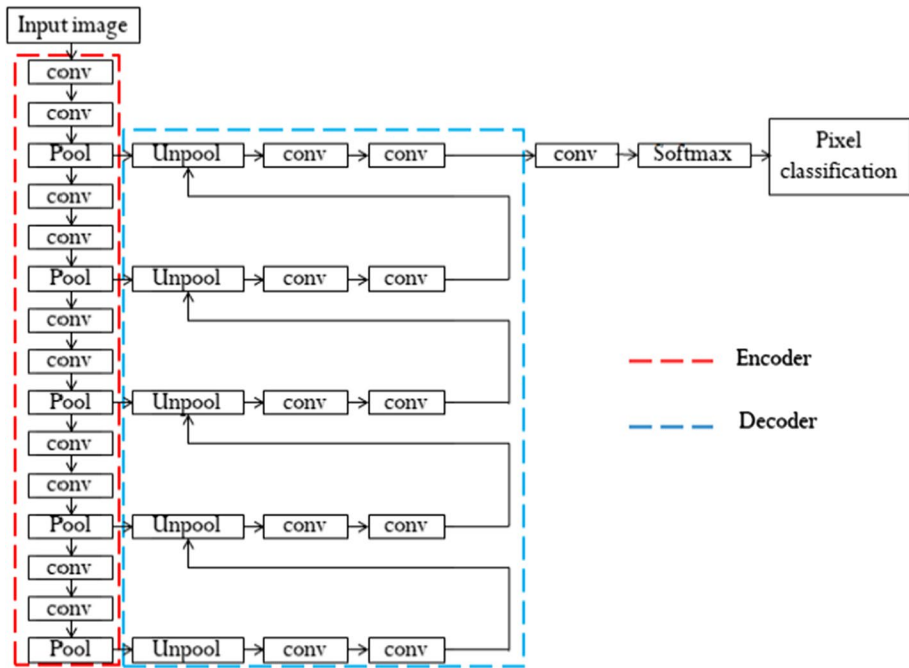

Fig. 2 Dilated convolution

Fig. 3 Fully convolved neural network

Pooling indices of unpooling layer is extracted from the encoder stage and using those indices the unpooling operation is done on the decoder stage. The number of filters in each layer remains the same throughout the stage. The maximum pooling operation of 2×2 is performed in the proposed FCCN. The output layer consists of softmax and pixel classification layer. The softmax layer assigns probability value for each class in an image. The pixel classification layer classifies the image as vessel and non-vessel region. The proposed architecture contains totally 73 layers comprising batch normalization and activation function layer; Increase of network layer enhances the architecture to learn significant features to segment the retinal vessels. Each layer provides useful information required for segmentation at various levels of features. The last convolution is a 1×1 convolution to map the channel to be the same as the input number of channels. The layer description of encoder and decoder for the proposed FCCN architecture are given in Tables 1 and 2 respectively.

2.5 Depth Concatenated Neural Network

The proposed DCNN architecture is modified from Unet architecture [22]. The hidden layer of DCN architecture consists of dilated convolution layer, maxpooling layer, ReLu layer, batch normalization layer, depth concatenation layer and upconvolutional layer. The proposed architecture is shown in Fig. 4.

The hidden layers form five encoder stages and five decoder stages. Each encoder stage consists of two convolutional layer and maxpooling layer. The depth concatenation layer in the proposed work concatenates the output from the encoder stage to the corresponding decoder stage. The depth concatenation layer concatenates the two inputs

Table 1 Encoder layer specification and output image size for FCNN

Encoder layer	Filter size and channel	Input image size	Output image size
1-CONV 1	3×3, 64	512×512	512×512
1-CONV 2	3×3, 64	512×512	512×512
Max pooling	2×2, 64	512×512	256×256
2-CONV 1	3×3, 64	256×256	256×256
2-CONV 2	3×3, 64	256×256	256×256
Max Pooling	2×2, 64	256×256	128×128
3-CONV 1	3×3, 64	128×128	128×128
3-CONV 2	3×3, 64	128×128	128×128
Max Pooling	2×2, 64	128×128	64×64
4-CONV 1	3×3, 64	64×64	64×64
4-CONV 2	3×3, 64	64×64	64×64
Max Pooling	2×2, 64	64×64	32×32
5-CONV 1	3×3, 64	32×32	32×32
5-CONV 2	3×3, 64	32×32	32×32
Max Pooling	2×2, 64	32×32	16×16

Table 2 Decoder layer specification and output image size for FCNN

Decoder layer	Filter size and channel	Input image size	Output image size
Unpool	–	16×16	32×32
5-CONV 1	3×3,64	32×32	32×32
5-CONV 2	3×3, 64	32×32	32×32
Unpool	–	32×32	64×64
4-CONV 1	3×3, 64	64×64	64×64
4-CONV 2	3×3, 64	64×64	64×64
Unpool	–	64×64	128×128
3-CONV 1	3×3,64	128×128	128×128
3-CONV 2	3×3,64	128×128	128×128
Unpool	–	128×128	256×256
2-CONV 1	3×3,64	256×256	256×256
2-CONV 2	3×3,64	256×256	256×256
Unpool	–	256×256	512×512
1-CONV 1	3×3, 64	512×512	512×512
1-CONV 2	3×3,64	512×512	512×512

that are of same height and width. The difference between the proposed architecture and the UNET architecture is that all the dropout layers are removed and batch normalization after each convolution layer is included and convolution layer used in each stage is dilated convolution. The upconvolution layer performs transpose convolution operation. The decoder stage consists of depth concatenation layer, two convolution layer and one upconvolution layer. The number of filters with which the images convolved in each stage doubles that of previous stage. The number of filter in the first stage is sixteen
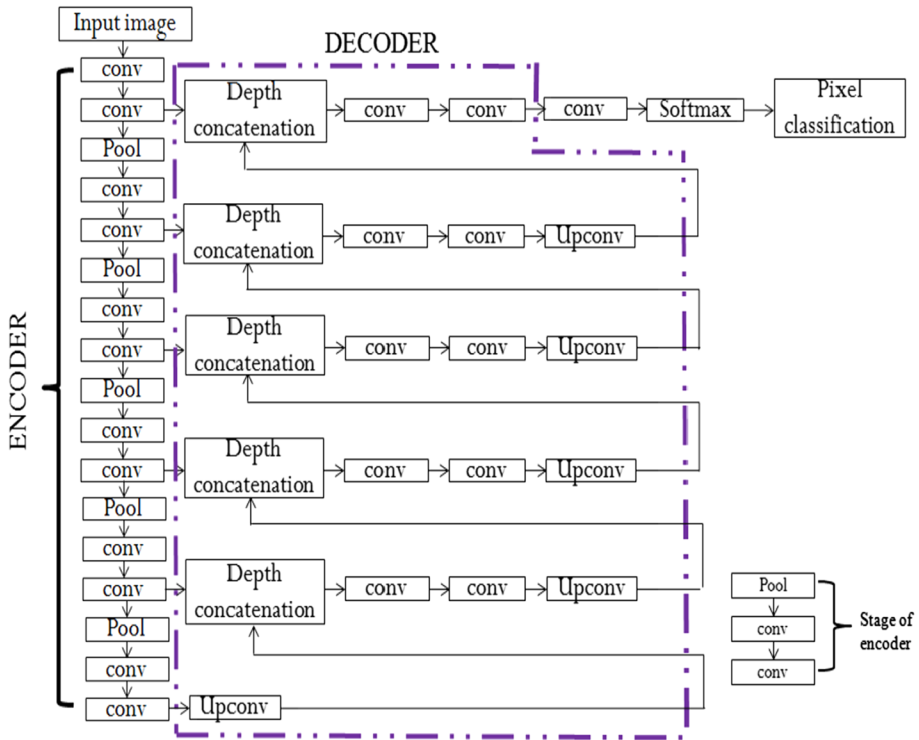
Fig. 4 Depth concatenated neural network

and hence for the fifth stage it is 512. The size of image is reduced to half the size of image for each encoder stage. The images are then upsampled by decoder. Convolution operation performed in the proposed work involves dilation factor of 2. The pooling operation performed in the proposed work is 2x2 maximum pooling. The output layer in DCN is same as that of FCCN. The encoder and decoder layer specifications are listed in Tables 3 and 4 respectively.

2.6 Class Imbalance problem

In pixel classification layer, all the pixels in the input images are classified as vessels and non-vessels. Vessel pixel count occupies less than one fourth of the total image pixel count which leads to misclassification of vessel as non-vessel. In order to avoid this class imbalance problem class weightage of vessel is increased. The images are classified as vessel and non-vessel, whose corresponding weights are denoted as $CW = CW^{(1)}, CW^{(2)}$. The total class weight (T(CW)) is given by Eq. (2)

$$T(CW) = CW^{(1)} + CW^{(2)} \qquad (2)$$

Class frequency are obtained by dividing each class weight by total class weight. Class frequency F is presented by Eq. (3)

Table 3 Encoder layer specification and output image size for DCNN

Encoder layer	Filter size and channel	Input image size	Output image size
1-CONV 1	$3 \times 3,16$	512×512	512×512
1-CONV 2	$3 \times 3,16$	512×512	512×512
Max pooling	$2 \times 2,16$	512×512	256×256
2-CONV 1	$3 \times 3,32$	256×256	256×256
2-CONV 2	$3 \times 3,32$	256×256	256×256
Max pooling	$2 \times 2,32$	256×256	128×128
3-CONV 1	$3 \times 3,64$	128×128	128×128
3-CONV 2	$3 \times 3,64$	128×128	128×128
Max pooling	$2 \times 2,64$	128×128	64×64
4-CONV 1	$3 \times 3,128$	64×64	64×64
4-CONV 2	$3 \times 3,128$	64×64	64×64
Max pooling	$2 \times 2,128$	64×64	32×32
5-CONV 1	$3 \times 3,256$	32×32	32×32
5-CONV 2	$3 \times 3,256$	32×32	32×32
Max Pooling	$2 \times 2,256$	32×32	16×16

Table 4 Decoder layer specification and output image size for DCNN

Decoder layer	Filter size and channel	Input image size	Output image size
TCONV	$2 \times 2,256$	16×16	32×32
5-CONV 1	$3 \times 3,256$	32×32	32×32
5-CONV 2	$3 \times 3,256$	32×32	32×32
5-TCONV	$2 \times 2,128$	32×32	64×64
4-CONV 1	$3 \times 3,128$	64×64	64×64
4-CONV 2	$3 \times 3,128$	64×64	64×64
4 -TCONV	$2 \times 2,64$	64×64	128×128
3-CONV 1	$3 \times 3,64$	128×128	128×128
3-CONV 2	$3 \times 3,64$	128×128	128×128
3 -TCONV	$2 \times 2,32$	128×128	256×256
2-CONV 1	$3 \times 3,32$	256×256	256×256
2-CONV 2	$3 \times 3,32$	256×256	256×256
2 -TCONV	$2 \times 2,16$	256×256	512×512
1-CONV 1	$3 \times 3,16$	512×512	512×512
1-CONV 2	$3 \times 3,16$	512×512	512×512

$$F^i = \frac{CW^i}{T(CW)}; \quad i = 1,2 \tag{3}$$

Inverse class weight $(CW_I^{(i)})$ is obtained by taking inverse of class frequency and is given by Eq. (4).

$$CW_I^{(i)} = \frac{1}{F^i}; \quad i = 1, 2 \tag{4}$$

Pixel classification layer is updated with class weight so that more weightage is given to the vessel region compared to the non vessel region. The network learns well by avoiding the class imbalance problem with the inclusion of weightage to the classes and it enhances the segmentation result.

3 Results

3.1 Database

The proposed methodology is experimented using two popular databases namely DRIVE (Digital Retinal Images for Vessel Extraction) [33] and STructured Analysis of Retina (STARE) [34]. Canon CR5 with 45° field of view is used for capturing DRIVE images. The database comprises totally 40 images; half the images are for training images and the half is for testing purpose. Each image is of size 565x584, and available with corresponding groundtruth and its mask image. Two ground truths are provided for testing; two experts were developed the groundtruth. Testing images are also available with one mask image. The STARE database are obtained using TopCon TRV-50 fundus camera at 35° Field of view. Images are of size 605×700 and for each image two independent manual segmentations are offered as groundtruth.

3.2 Evaluation Metric

The proposed methodology is evaluated using sensitivity, specificity and accuracy as performance metrics. The classified pixels may be under any of the following cases; True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN). TP is the outcome of vessel predicted correctly as vessel and those are wrongly predicted as non-vessel pixels instead of vessels are counted as FN. TN is the outcome of non-vessel pixel correctly predicted as non-vessel and the non-vessel pixel wrongly predicted as vessel are defined as FP. The sensitivity (Se), specificity (Sp) and accuracy (Acc) are calculated as follows.

$$Se = \frac{TP}{TP + FN} \tag{5}$$

$$Sp = \frac{TN}{TN + FP} \tag{6}$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

3.3 Experimental Results

The proposed Dilated FCNN and DCNN architectures are trained using twenty training images taken from DRIVE database and is evaluated for the remaining twenty test images.

In addition in order to validate the performance of the proposed trained architecture we tested twenty abnormal images from STARE database. In order to show the improvement in accuracy for the proposed dilated architectures, we do experimentation on both the architectures with and without dilation layers. The entire experimentation is carried out on Intel i5 core, 32 GB CPU system using Matlab 2018b. The number of epoch is a parameter that defines the number of times the training algorithm learns the entire training dataset. Both the proposed architectures are trained until the training accuracy improvement and loss reaches an optimum value; that is further improvement is very negligible. It is experiential that the training accuracy and training loss for FCCN architecture is 95.71% and 0.001 respectively. The DCNN architecture reaches the training accuracy of about 92.35% and loss of 0.131. Even though training accuracy is more for FCNN architecture, overall accuracy computation is better for DCNN architecture. Also it is observed that when all the convolutional layers are replaced by dilated convolutional layers, the overall accuracy gets increased for both FCNN and DCNN architectures. The improvement in overall accuracy for the dilated architectures is due to the decline of false positives more accurately by the proposed architectures. The segmented results for DRIVE database using FCCN and DCNN with and without dilation are shown in Fig. 5.

In order to verify the data independancy of the proposed architectures, the vessel extracted from the image with different levels of contrast and illumination are listed. In Fig 5, the first row shows the result of low contrast image, the result of high illuminated image in second row and the result of high illuminated and low contrast image in third row using the proposed architectures with and without dilation. The evaluation metric for twenty test images of DRIVE such as sensitivity, specificity and accuracy are calculated and tabulated in Table 5 and for dilation it is tabulated in Table 6. First human observer groundtruth is taken to perform the metric calculation.

The evaluation metric obtained for DCNN architecture is better compared to FCNN architecture. The accuracy value of DCNN is increased to nearly one percentage compared to FCCN. From Tables 5 and 6 it is observed that both architectures with dilated
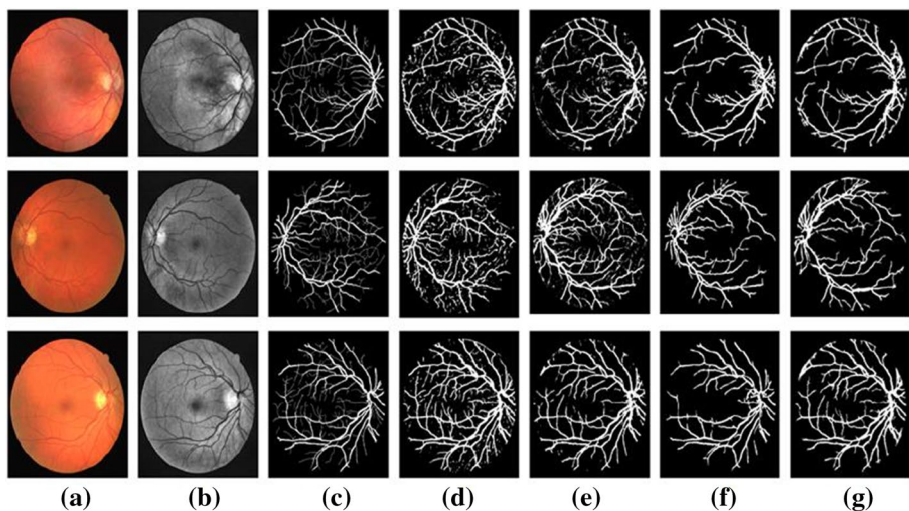

Fig. 5 Segmentation result of DRIVE dataset with dilation. **a** Color fundus image. **b** Enhanced green channel fundus image. **c** Ground truth. **d** FCCN. **e** DCNN. **f** FCCN with dilation. **g** DCNN with dilation

Table 5 Evaluation metrics of DRIVE database for FCNN and DCNN

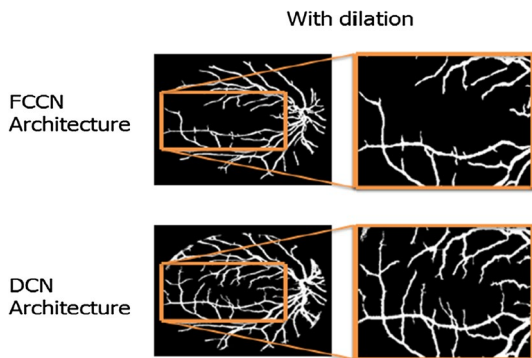| Image	FCNN architecture			DCNN architecture		
	Se(%)	Sp(%)	Acc(%)	Se(%)	Sp(%)	Acc(%)
1	77.67	96.44	93.97	80.86	96.99	95.61
2	76.78	96.81	93.80	80.92	96.93	95.31
3	82.86	96.73	94.81	86.89	95.70	95.02
4	72.23	97.11	93.74	83.74	96.44	95.44
5	66.84	97.54	93.31	84.40	95.91	95.04
6	68.51	98.13	93.89	84.32	95.54	94.68
7	69.62	97.17	93.47	78.61	96.42	94.93
8	61.05	98.02	93.34	83.85	96.13	95.29
9	57.87	98.64	93.78	85.29	96.23	95.56
10	71.48	96.91	93.83	82.20	96.76	95.71
11	89.85	95.93	95.13	73.42	96.58	94.48
12	67.87	97.54	93.77	82.88	96.40	95.43
13	72.19	97.24	93.64	86.96	95.73	95.06
14	73.35	96.34	93.61	80.28	96.98	95.72
15	76.72	96.34	94.01	70.99	97.96	95.74
16	70.04	97.67	94.00	81.51	96.82	95.56
17	74.02	97.93	94.24	83.30	96.47	95.56
18	70.68	97.26	94.16	78.22	97.49	95.97
19	84.55	96.96	95.44	79.03	97.80	96.14
20	73.12	97.31	94.70	77.25	97.71	96.18
Average	72.87	97.20	94.03	81.25	96.65	95.42

convolution layers have improved accuracy and sensitivity value. The visibility of the vessel extracted image is good for the DCNN dilated architecture due to the avoidance of artifacts arised due to retinal image capturing effects. The maximum sensitivity, specificity and accuracy for twenty images of DRIVE database for Dilated FCNN are 87.57%, 97.97% and 95.92% respectively. The maximum sensitivity, specificity and accuracy for twenty images of DRIVE database for Dilated DCNN is 89.44%, 98.31% and 97.24%. The comparison of segmented image of two Dilated architectures are shown in Fig. 6. From the figure shown, it is clearly identified that the segmented result obtained using DCNN with dilation are better compared to FCNN with dilation. Thin vessels are correctly distinguished from the background by the dilated convolutional layers of DCNN.

The cross validation of the proposed FCNN and DCNN is done using STARE database (i.e) the network is trained using DRIVE database and is tested on STARE database. The abnormal images form STARE database are purposely selected for cross validation. The segmented result for STARE database are shown in Fig. 7. The segmented result for the abnormal images also found good for Dilated DCNN compared to other architectures.

From Fig. 7, it is observed that the abnormalities like red and bright lesions present in the input images are eliminated in a better way by the proposed architectures. The evaluation metric for twenty test images of STARE using both architecture with dilation such as sensitivity, specificity and accuracy are calculated and tabulated in Tables 7 and 8. The maximum sensitivity, specificity and accuracy of STARE database for FCNN with dilation are 95.58%, 98.55% and 96.31% respectively. The maximum sensitivity, specificity

Table 6 Evaluation metrics of DRIVE database for FCNN and DCNN with dilation

Image	FCNN architecture			DCNN architecture		
	Se (%)	Sp (%)	Acc (%)	S e(%)	Sp (%)	Acc (%)
1	83.96	96.08	95.00	77.41	97.15	95.71
2	85.18	96.62	95.45	84.09	96.77	95.68
3	70.75	97.10	94.47	81.66	96.75	95.71
4	77.62	97.54	95.70	86.38	96.97	96.25
5	75.77	97.62	95.57	79.02	97.91	96.47
6	72.32	97.58	95.12	81.48	97.45	96.26
7	78.24	97.20	95.47	77.97	97.18	95.76
8	69.77	97.97	95.54	86.44	97.25	96.58
9	75.88	97.69	95.92	80.75	98.31	97.24
10	79.55	97.08	95.64	77.21	97.71	96.36
11	77.71	96.63	94.94	79.13	97.53	96.22
12	77.49	97.22	95.52	83.07	97.41	96.46
13	80.20	96.81	95.19	74.07	98.06	96.16
14	77.44	97.16	95.57	78.14	97.43	96.17
15	85.37	96.45	95.65	84.14	96.38	95.68
16	82.27	96.75	95.44	78.65	97.35	95.97
17	77.61	97.01	95.38	72.51	98.01	96.24
18	81.48	96.51	95.32	80.33	97.08	96.00
19	87.57	96.17	95.45	85.88	96.69	95.97
20	82.12	96.59	95.53	81.19	97.19	96.23
Average	78.91	96.99	95.39	80.48	97.33	96.16

Fig. 6 Comparison of segmented result between FCCN and DCNN architectures

and accuracy of STARE database for DCNN with dilation is about 95.16%, 98.98% and 98.07%. From the tables, it is observed that overall performance of DCNN architecture is better than FCNN architecture. It is also observed that sensitivity value get improved for the dilated covolution architecture compared to the conventional architecture. Fig. 8, shows the comparison of performance metrics for the FCNN and DCNN architectures with and without dilation. It is inferred that accuracy and specificity value improved for the Dilated DCNN architecture. Even the sensitivity value is reduced to 80% for the Dilated DCNN, it is still high among most of the existing methodologies.
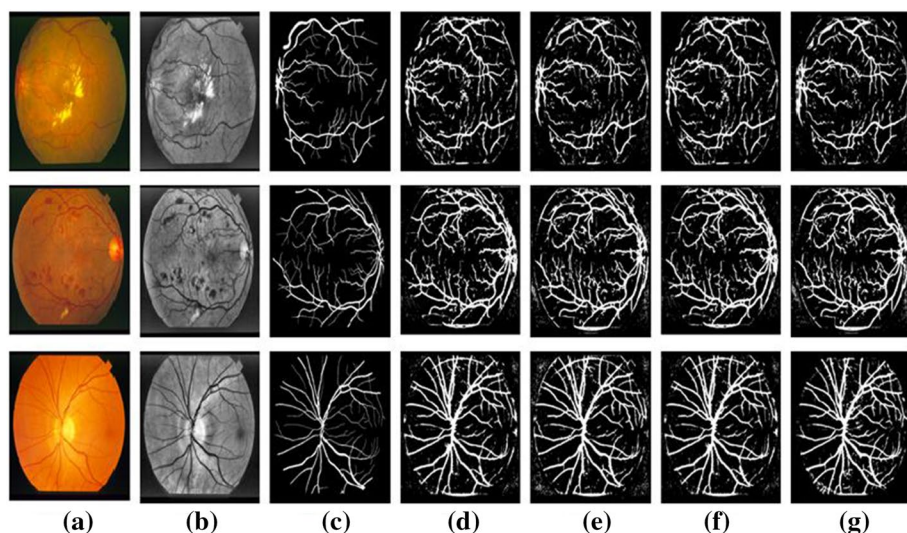
Fig. 7 Segmentation result of STARE database **a** Color fundus image. **b** Enhanced green channel fundus image. **c** Ground truth. **d** FCNN. **e** DCNN. **f** FCNN with dilation. **g** DCNN dilation

Table 7 Evaluation metrics of STARE database for FCNN and DCNN

Image	FCNN architecture			DCNN architecture		
	Se (%)	Sp (%)	Acc (%)	Se (%)	Sp (%)	Acc (%)
1	66.15	92.91	91.84	72.92	96.56	94.83
2	72.36	94.52	93.02	89.90	93.22	92.88
3	79.02	94.1	93.37	89.47	94.07	93.60
4	46.4	96.22	93.42	90.53	92.41	92.21
5	85.1	90.85	90.6	86.85	90.85	90.44
6	86.76	92.53	92.5	58.26	98.78	92.56
7	90.77	89.15	89.51	87.60	94.26	93.69
8	85.27	90.38	90.32	87.54	92.14	91.73
9	85.86	89.84	90.08	89.90	91.68	91.50
10	88.3	89.47	89.74	64.84	99.50	95.11
11	95.23	89.42	89.66	95.45	96.92	96.80
12	95.16	89.12	89.52	92.21	94.02	93.87
13	85.16	91.04	90.82	96.47	97.31	97.23
14	84.37	91.33	91.23	89.65	94.27	93.83
15	65.51	92.64	91.98	96.01	95.74	95.76
16	65.49	93.14	91.23	85.56	90.35	89.83
17	92.24	91.5	91.45	90.84	93.97	93.66
18	73.7	95.36	94.81	90.57	95.04	94.61
19	57.03	94.34	93.84	91.72	95.71	95.38
20	53.6	95.26	93.36	90.62	93.19	92.96
Average	77.67	92.16	91.62	86.84	94.50	93.62

Table 8 Evaluation metrics of STARE database for FCNN and DCNN with dilation

Image	FCNN architecture			DCNN architecture		
	Se (%)	Sp (%)	Acc (%)	Se (%)	Sp (%)	Acc (%)
1	76.88	94.33	91.93	87.96	97.85	97.01
2	67	95.76	94.09	90.1	97.48	96.67
3	79.26	94.44	93.45	89.7	98.25	97.31
4	52.43	98.55	94.49	90.62	96.71	96.02
5	87.55	92.87	92.12	87.82	99.25	98.07
6	92.01	92.87	92.41	90.08	97.1	96.38
7	94.76	91.42	91.36	89.45	95.92	95.18
8	89.34	92.85	92.24	92.77	96.17	95.76
9	93.65	93.14	92.52	89.63	95.98	95.31
10	93.63	93.5	93.05	94.57	95.2	95.15
11	94.1	91.13	91.45	93.48	95.26	95.12
12	95.58	92.19	92.44	90.99	97.35	96.82
13	88.15	93.85	93.03	94.04	95.62	95.47
14	90.07	94.36	93.4	86.24	96.24	95.27
15	83.22	94.91	92.19	95.16	94.81	94.84
16	71.63	96.81	93.45	85.68	98.98	97.55
17	90.74	93.11	93.02	89.43	92.9	92.57
18	82.19	97.6	96.31	90.84	95.06	94.61
19	79.95	97.74	95.86	90.94	95.9	95.38
20	62.31	94.84	93.24	91.63	94.51	94.2
Average	83.22	94.31	93.1	89.25	96.34	95.64
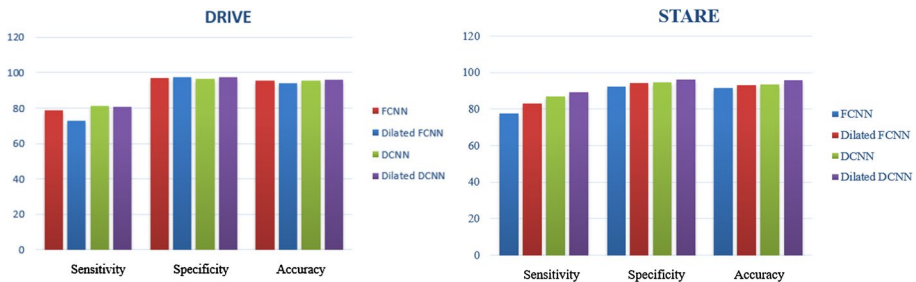
Fig. 8 Bar graph comparison of evaluation metric

In order to verify the significance of dilation in convolution layers, the proposed architectures are compared with some of the existing deep neural network (DNN) architectures and other recent methods. The performance comparison of DRIVE database with existing methodologies is listed in Table 7 and it is inferred that the proposed depth concatenated dilated architecture results are better compared to most of the existing methods. We also do experimentation on the data independency of the proposed architecture. The dilated architectures which are trained on normal images from DRIVE database is tested over selected abnormal images from STARE database. Sensitivity achieved by the depth concatenated

Table 9 Performance comparison with existing methodologies for DRIVE

Publication year	Method	Accuracy	Sensitivity	Specificity
2018	Shahid [2]	0.958	0.730	0.9793
2018	Soomro [3]	0.953	0.752	0.976
2018	Biswal [4]	0.9541	0.71	0.97
	DNN based Methods:			
2016	Li [24]	0.9527	0.7569	0.9816
2016	Liskowski [16]	0.9515	0.752	0.9806
2018	Oliveira [25]	0.9576	0.8039	0.9804
2018	Jiang [21]	0.9624	0.754	0.9825
2018	Yan [18]	0.9542	0.7653	0.9818
2018	Hu [19]	0.9533	0.7772	0.9793
	Proposed DNN			
	Dilated FCNN	0.9539	0.7891	0.9699
	Dilated DCNN	0.9616	0. 8048	0.9733

Table 10 Performance comparison with existing methodologies for STARE

Publication year	Method	Accuracy	Sensitivity	Specificity
2011	Marin[35]	0.9526	0.6944	0.9819
2012	Fraz [36]	0.9495	0.7010	0.9770
2016	Li [24]	0.9545	0.7027	0.9828
2018	Oliveira [25]	0.9597	0.8453	0.9726
	Proposed DNN			
	Dilated FCNN	0.931	0.8322	0.9431
	Dilated DCNN	0.9564	0.8925	0.9634

dilated architecture is about 89.25% and as far as we know, it is the maximum sensitivity obtained for the retinal vessel extraction for STARE database. Crossed trained STARE database also shows better compared to the existing methodologies and they are listed in Table 8. Accuracy and sensitivity of the Dilated DCNN architecture shows better performance compared to other methods (Tables 9 and 10).

4 Conclusion

Extraction of fine vessels from the retinal fundus images are still a challenging task. *I*n this work, we proposed two novel deep neural neural netwoks based on dilated convolution layers. Dilated convolution performs well in the vessel detection task irrespective of vessel thickness. *I*t is also inferred that for the pixel level classification, depth concatenation with dilated convolution improves the result in a better way. The proposed architectures are observed to be the optimum performing architctures due to the inclusion of class balanced cross entropy loss function. The proposed dilated architectures optimized the segmentation of retinal blood vessels. The proposed methodology is evaluated on DRIVE and STARE

database using the two different proposed architectures. The fine vessels are detected well in the DCNN with dilation architecture compared to FCNN architecture. In future the work can be further improved by enhancing encoder and decoder layers. The proposed dilated architectures show superior performance in the fine vessel detection even in the presence of abnormalities. *I*n conclusion, based on the performance measures calculated and also by the naked eye observation proposed architectures outperform the state of art methods.

Data Availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest Authors V. Sathananthavathi & G. Indumathi declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants performed by any of the authors.

References

1. Salazar-Gonzalez, A., Kaba, D., Li, Y., & Liu, X. (2014). Segmentation of the blood vessels and optic disk in retinal images. *IEEE Journal of Biomedical and Health Informatics, 18*(6), 1874–1886.
2. Shahid, M., & Taj, I. A. (2018). Robust retinal vessel segmentation using vessel's location map and Frangi enhancement filter. *IET Image Processing, 12*(4), 494–501.
3. Soomro, T. A., Khan, T. M., Khan, M. A. U., et al. (2018). Impact of ICA-based image enhancement technique on retinal blood vessels segmentation. *IET Image Processing, 6,* 3524–3533.
4. Biswal, B., Pooja, T., & Bala Subrahmanyam, N. (2018). Robust retinal blood vessel segmentation using line detectors with multiple masks. *IET Image Processing, 12*(3), 389–399.
5. Fraz, M. M., et al. (2011). Retinal vessel extraction using first-order derivativeof gaussian and morphological processing. In Advances in Visual Computing. Springer, pp. 410–420.
6. Yin, Y., et al. (2013). Automatic segmentation and measurement of vasculature in retinal fundus images using probabilistic formulation. *Computational and mathematical methods in medicine*. https://doi.org/10.1155/2013/260410.
7. Zhu, C., Zou, B., Zhao, R., Cui, J., Duan, X., Chen, Z., & Liang, Y. (2017). Retinal vessel segmentation in colour fundus images using extreme learning machine. *Computerized Medical Imaging and Graphics, 55,* 68–77.
8. Marín, D., Aquino, A., Gegúndez-Arias, M. E., & Bravo, J. M. (2011). A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants based features. *IEEE Transactions on Medical Imaging, 30*(1), 146–158.
9. Sathananthavathi, V., & Indumathi, G. (2018). BAT algorithm inspired retinal blood vessel segmentation. *IET Image Processing, 12*(11), 2075–2083.
10. Sumathi, T., Vivekanandan, P., & Balaji, R. (2018). Retinal vessel segmentation using neural network (RVSNN). *Image Processing, 12*(5), 669–678.
11. Shelhamer, E., Long, J., & Darrell, T. (2017). Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(4), 640–651.
12. Salem, S. A., Salem, N. M., & Nandi, A. K. (2007). Segmentation of retinal blood vessel using a novel clustering algorithm (RACAL) with a partial supervision strategy. *Medical & Biological Engineering & Computing, 45*(3), 261–273.
13. Sathananthavathi, V., Indumathi, G., Mahiya, R., & Priyadarshini, S. (2021). Improvement of thin retinal vessel extraction using mean matting method. *International Journal of Imaging Systems and Technology, 31,* 1455–1467.

14. Orlando, J. I., Prokofyeva, E., & Blaschko, M. B. (2017). 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.

15. Wang, S., Yin, Y., Cao, G., Wei, B., Zheng, Y., & Yang, G. (2015). Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing, 149,* 708–717.

16. Liskowski, Paweł, & Krawiec, Krzysztof. (2016). Segmenting retinal blood vessels with deepneural networks. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2016.2546227.

17. Wang, Xiaohong, Jiang, Xudong, & Ren, Jianfeng. (2019). Blood vessel segmentation from fundus image by a cascade classification framework. *Pattern Recognition, 88,* 331–341.

18. Yan, Zengqiang, Yang, Xin, & Cheng, Kwang-Ting. (2018). Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation. *IEEE Transactions on Biomedical Engineering, 65*(9), 1912–1923.

19. Kai, Hu., Zhang, Zhenzhen, Niu, Xiaorui, Zhang, Yuan, Cao, Chunhong, Xiao, Fen, & Gao, Xieping. (2018). Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function. *Neurocomputing, 309,* 179–191.

20. Fu, H., Xu, Y., Kee Wong, D. W., & Liu, J. (2016). Retinal vessel segmentation via deep learning and conditional random field, In Proceedings MICCAI, pp. 132-139.

21. Jiang, Z. Zhang, H., Wang, Y., & Ko, S.-B. (2018). Retinal blood vessel segmentation using fully convolutional network with transfer learning. *Computerized Medical Imaging and Graphics, 68,* 1–15.

22. Ronneberger, O., Fischer, P., Brox, T., (2015). U-net Convolutional neural networks for biomedical image segmentation, In Proceeding MICCAI, pp. 234-241

23. Dasgupta, A., Singh, S. (2017) A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation, In Proceeding ISBI, pp. 18-21.

24. Li, Qiaoliang, Feng, Bowei, Xie, LinPei, Liang, Ping, Zhang, Huisheng, & Wang, Tianfu. (2016). A cross-modality learning approach forvessel segmentation in retinal images. *IEEE Transactions on Medical Imaging, 35*(1), 109–118. https://doi.org/10.1109/TMI.2015.2457891.

25. Oliveira, Americo, Carlos, Sergio Pereira, & Silva, A. (2018). Retinal vessel segmentation based on fully convolutional neural networks. *Expert Systems with Applications 112*(1), 229–242.

26. Sathananthavathi, V., & Indumathi, G. (2020). Parallel architecture of fully convolved neural network for retinal vessel segmentation. *Journal of Digital Imaging, 33*(1), 168–180.

27. Tian, C., Fang, T., Fan, Y., & Wu, W. (2020). Multi-path convolutional neural network in fundus segmentation of blood vessels. *Biocybernetics and Biomedical Engineering, 40*(2), 583–595.

28. Atli, I., & Gedik, O. S. (2021). Sine-Net: A fully convolutional deep learning architecture for retinal blood vessel segmentation. *Engineering Science and Technology, an International Journal, 24*(2), 271–283.

29. Uysal, E., & Guraksin, G. E. (2021). Computer-aided retinal vessel segmentation in retinal images: Convolutional neural networks. *Multimedia Tools and Applications, 80,* 3505–3528.

30. Huang, L., & Liu, F. (2020). Retinal vessel segmentation using simple SPCNN model and line connector. *The Visual Computer.* https://doi.org/10.1007/s00371-020-02008-y.

31. Marios, et al. (2019). Semantic segmentation of pathological lung tissue with dilated fullyconvolutional networks. *IEEE Journal of Biomedical and Health Informatics, 23*(2), 714–722. https://doi.org/10.1109/JBHI.2018.2818620.

32. Badrinarayanan, Vijay. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transaction on Pattern Analysis and Machine Learning, 39*(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615.

33. Staal, J. J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., & van Ginneken, B. (2004). Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging, 23*, 501–509.

34. Hoover, A., Kouznetsova, V., & Goldbaum, M. (2000). Locating Blood Vessels in Retinal Images by Piece-wise Threshold Probing of a Matched Filter Response. *IEEE Transactions on Medical Imaging, 19*(3), 203–210.

35. Marin, Diego, Aquino, Arturo, Gegundez-Arias, Manuel Emilio, & Bravo, Jose Manuel. (2011). A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE Transactions On Medical Imaging, 30*(1), 146–158. https://doi.org/10.1109/TMI.2010.2064333.

36. Fraz, M. M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A. R., Owen, C. G., & Barman, S. (2012). A, An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering, 59*(9), 2538–2548.

V. Sathananthavathi is an Assistant professor (Sr.Grade) in ECE, Mepco Schlenk Engineering College. She obtained her B.E from Madurai Kamaraj University and M.E from Anna university in 2003 and 2010 respectively. She completed her Ph.D. on retinal image processing under Anna University. She is a lifetime member of ISTE and IETE. Her interests encompass medical image processing and optimization techniques.

G. Indumathi is a professor in ECE, Mepco Schlenk Engineering College. She obtained her B.E and M.E from Madurai Kamaraj University in 1990 and 1998 respectively. She received Ph.D. from Anna University. She is a lifetime member of ISTE and IETE. Her interests encompass several aspects of wireless communications such as Signal and Image processing, antenna design.