A Deep Convolutional Encoder–Decoder Architecture for Retinal Blood Vessels Segmentation

6 authors, including:

Adekanmi Adegun
Landmark University
25 PUBLICATIONS **19** CITATIONS

Marion O. Adebiyi
Covenant University Ota Ogun State, Nigeria
47 PUBLICATIONS **267** CITATIONS

AKANDE NOAH OLUWATOBI
Landmark University
40 PUBLICATIONS **27** CITATIONS

Roseline Ogundokun
Landmark University
35 PUBLICATIONS **27** CITATIONS

Some of the authors of this publication are also working on these related projects:

COMPUTATIONAL APPROACH TO MECHANISTIC AND PREDICTIVE TOXICOLOGY OF INSECTICIDE View project

Smart Automation View project

A Deep Convolutional Encoder-Decoder Architecture for Retinal Blood Vessels Segmentation

Adegun Adekanmi Adeyinka[1], Marion Olubunmi Adebiyi[2],
Noah Oluwatobi Akande[3(✉)], Roseline Oluwaseun Ogundokun[3],
Anthonia Aderonke Kayode[3], and Tinuke Omolewa Oladele[4]

[1] School of Computer Science, University of Kwazulu-Natal,
Durban, South Africa
[2] Department of Computer and Information Sciences,
Covenant University, Ota, Nigeria
[3] Bioninformatics Research Group, Computer Science Department,
Landmark University, Omu-Aran, Kwara, Nigeria
akande.noah@lmu.edu.ng
[4] Computer Science Department, University of Ilorin, Ilorin, Kwara, Nigeria

Abstract. Over the last decades, various methods have been employed in medical images analysis. Some state-of-the-arts techniques such as deep learning have been recently applied to medical images analysis. This research proposes the application of deep learning technique in performing segmentation of retinal blood vessels. Analyzing and segmentation of retina vessels has assisted in diagnosis and monitoring of some diseases. Diseases such as age-related fovea degeneration, diabetic retinopathy, glaucoma, hypertension, arteriosclerosis and choroidal neovascularization can be effectively managed by the analysis of retinal vessels images. In this work, a Deep Convolutional Encoder-Decoder Architecture for the segmentation of retinal vessels images is proposed. The proposed method is a deep learning system composed of an encoder and decoder mechanism allows a low resolution image set of retinal vessels to be analyzed by set of convolutional layers in the encoder unit before been sent into a decoder unit for final segmented output. The proposed system was evaluated using some evaluation metrics such as dice coefficient, jaccard index and mean of intersection. The review of the existing works was also carried out. It could be shown that the proposed system outperforms many existing methods in the segmentation of retinal vessels images.

Keywords: Retinal vessels · Deep learning · Convolutional layers · Encoder · Decoder · Images · Segmentation

1 Introduction

Over the years, retinal blood vessels have been revealed to contain rich features that could be used to diagnose many deadly diseases such as Diabetes Retinopathy (DR), Hypertensive Retinopathy (HR), glaucoma, cataract etc. Growth of false blood vessels

are pointer to the presence of DR while a measure of the Arteriovenous Ratio (AVR) could be used to detect and grade the presence of HR and glaucoma [1, 2]. AVR is a function of the Central Retinal Venular Equivalent (CRVE) and Central Retinal Artery Equivalent (CRAE) of the retinal blood vessels. Therefore, how to segment retinal blood vessels with a view to extract features that could facilitate the diagnosis of pathological signs has been a growing field of research. However, segmenting the components of retinal has been a difficult task especially when the retinal contains pathological signs [3]. Therefore, automated approach to retinal blood vessels segmentation is being explored by researchers towards segmenting these components. Various automated segmentation techniques such as Deep machine learning [4, 5]; supervised neural network [6, 7]; U-net based Convolutional neural network [8, 9]; Deep neural network [10–12] have been employed in the literature with an acceptable level of accuracy. However, a more accurate segmentation technique is still being craved for by researchers as well as medical practitioners. Image analysis and segmentation carried out based on deep learning methods has produced improved results with a very high accuracy [4]. This article presents a deep learning technique that uses encoder–decoder mechanism for retinal blood vessels segmentation.

2 Related Works

In the last decade, there have been a lot of research about the application of deep learning for medical image analysis. Some works have been particularly carried out in the segmentation process of medical image analysis most especially retina images using state-of-the arts techniques.

Authors in [13] carried out a review of existing methods in retinal blood vessels segmentation. They categorized the methods into: kernel-based techniques, vessel-tracking, mathematical morphology-based, multi-scale, model-based, adaptive local thresholding and machine learning. A Convolutional Neural Network (CNN) architecture for the segmentation of blood vessels fundus images was proposed by authors in [9]. The architecture was derived from U-Net architecture that implements an encoder-decoder architecture. A segmentation accuracy of 97.90% was achieved when the technique was evaluated with DRIVE dataset. Also, a trainable pre-weighted neural network termed Frangi-Net was employed for retinal blood vessel segmentation in [7]. The technique was tested with healthy, DR and glaucomatous retinal images. An accuracy of 84%, 82% and 84% were recorded with the healthy, DR and glaucomatous retinal images respectively. Authors in [14] also proposed a multi-scale convolutional neural network structure and label processing approach for the segmentation of retinal blood vessel. Their method used two different scale image segments to generate input for two deep convolutional networks. A sensitivity, specificity and accuracy of 78.90%, 98.83% and 97.10% were achieved with healthy retinal images available in STARE database while 78.43%, 98.37% and 96.90% was achieved with pathological retinal images available in the same STARE database.

Another model consisting of a combination of CNN and Long Short Term Memory (LSTM) was proposed to perform segmentation on retinal OCT images [5]. CNN was applied to extract layers of interest image and also extract the edges and LSTM was

applied to trace the layer boundary. Their model was trainable and applicable to minimal data. A pixel wise mean absolute error of 1.30 ± 0.48 was achieved when the technique was evaluated on three publicly available databases. Similarly, a multiscale and multi-level deeply supervised convolutional network with short connections was developed for retinal vessel segmentation in [6]. Short connections were used to transfer semantic information between the layers. Both forward and backward short connections were used for passing low level information and the proposed method was applied on DRIVE dataset.

CNN and Conditional Random Fields (CRFs) were combined into an integrated deep network to develop a model called Deep Vessel system to achieve state-of-the-art retinal vessel segmentation performance [16]. The model was experimented on DRIVE, STARE, and CHASE DB1 datasets. A segmentation accuracy and sensitivity of 95.23%, 76.03%; 95.85%, 74.12% and 94.89%, 71.30% were recorded with DRIVE, STARE, and CHASE DB1 datasets respectively.

A deep neural network (DNN) that uses max-pooling layers (MPCNNs) instead of subsampling or down sampling was also used for retina images segmentation [10]. The MPCNN mapped input samples into output class probabilities using hierarchical layers to extract features and fully connected layers to classify extracted features. An accuracy of 94.66% and an area under receiver operating characteristics curve of 97.43% were recorded.

Encoder-decoder based convolutional network architecture for semantic pixel-wise segmentation of retinal images was developed [11]. The network adapted the well-known SegNet model that learns high dimensional feature maps. This provided pixel wise segmentation of retinal abnormalities by assigning particular object class label. The technique was evaluated using MESSIDOR dataset. A mean accuracy of 99.24% was achieved for exudates segmentation, 88.65% for cotton wool spots segmentation and 97.86% for haemorrhages.

A fully convolutional network (FCN) for retinal layers' segmentation on retinal optical coherence tomography (OCT) images was developed [8]. Fan et al. [17] introduced a hierarchical image matting model for retinal blood vessel segmentation. The continuous and extendible characteristics of retinal blood vessels was incorporated into the image matting model. The performance of the technique was carried out on DRIVE datasets and an accuracy of 0.881 was obtained.

An intelligent blood vessel segmentation technique that could track vessels in the presence of occlusion was presented in [17]. The technique used hidden markov model for the vessel segmentation. When evaluated on DRIVE dataset, an AUC of 0.9 was achieved. In the same vein, convolutional neural network was employed for retinal blood vessel segmentation in [18]. When the performance was carried out on DRIVE dataset using ROC curve, an accuracy of 0.9802 was achieved. Condition random field was employed in [19] to reduce the number of encoders and decoders used by U-Net for a more improved retinal blood vessel segmentation. The improved segmentation technique was validated on DRIVE dataset while ROC curve was used for the performance evaluation. A segmentation accuracy of 0.9748 was achieved.

3   Methodology

3.1   Dataset

The proposed method was evaluated on publicly available retinal image dataset known as Digital Retinal Images for Vessel Extraction (DRIVE) [15], It contains 80 color fundus images. The images were further divided into a training set containing 57 samples and a test set, both containing 27 image samples. The training set and the testing set also come with their respective ground truth labels.

3.2   Model Implementation

The software used for the model implementation includes:

- Python Version 3.5
- Keras 2
- Tensor flow backend
- Scikit-image Version 0.14.1

The experimental platform is a personal computer equipped with an Intel Core i7 processor with ten (10) 3.4 GHZ cores, 16 GB memory and NVIDIA Tesla K40c GPU with Ubuntu 16.04 operating system.

3.3   General Architecture of the Proposed System

The proposed deep convolutional encoder-decoder model architecture adapts the popular U-Net model that implements a contemporary classification network into fully convolutional networks. It operates by learning from pixels to pixels in an end-to-end manner. Training image dataset and corresponding ground truth image dataset serves as the input into the system. These will be taken through the other sections of the system such as data preparation section and the learning and training section. The form of training employed is a supervised form of training whereby the target goal in the form of the ground truth images is supplied together with the input images. After a period of training, the validation image data set will now be fed into the system for the expected predictions. The output data from the prediction will finally be compared with our target goals for accuracy and the performance metrics evaluated.

The architecture is composed of the following major sections:

Data Preparation Section
The training image data set and the corresponding ground truth image data set were used to train the model simultaneously. The images were resized into 96×96 dimension, after which they were resampled to remove every form of noise. The ground truth serves as the expected output and the training set serves as the input to train the system. The input datasets (training and validations sets) and the ground truth are located in an online repository called DRIVE dataset. The sample training image data set and ground truth images are illustrated in Figs. 1 and 2 below.
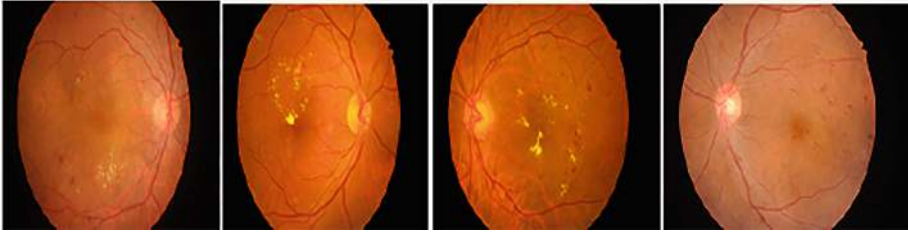
Fig. 1. A sample retina training image data set
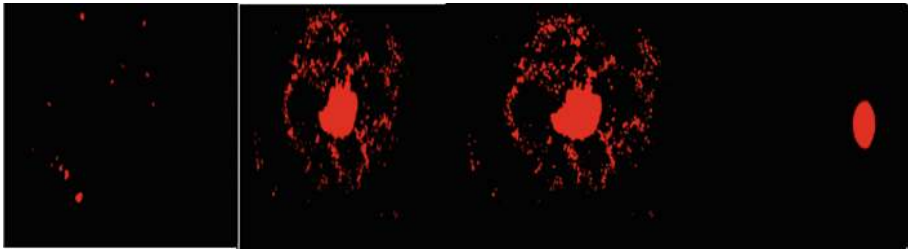

Fig. 2. The corresponding ground truth images

Learning and Training Section:

In this section, we have the encoder units and the decoder units.

Encoder:

The encoder is made up of set of convolution layers and pooling layers. Features extraction from the input image takes place at the convolution layer and the pooling layers reduce the resolution of the image feature maps. In this model, the encoder applies Rectified Linear Unit (RELU) activation function.

Each encoder uses the Max-pooling to translate invariance over small spatial shifts in the image and combine this with Subsampling to produce a large input image context in terms of spatial window. This method reduces the feature map size and this leads to image representation that is noisy with blurred boundaries. The restoration is done by decoder as the output image resolution must be the same as input image.

Decoder:

The decoder ensures that the image resolution of image set from the encoder units is increased to the initial resolution status. It is also made up of set of convolution layers. Each of the layers in the decoder stage corresponds to the layer in the encoder i.e. for each of the encoders there is a corresponding decoder which up samples the feature map using the already stored max-pooling indices. Sparse feature maps of higher resolutions are then produced. These are fed through the training section to produce dense feature maps.

Prediction:

The prediction section of this model is performed by predicting pixel-wise labels for an output which has the same resolution as the input image. The last part of the decoder is

connected to a softmax classifier which classifies each pixel. The sparse feature maps restored from the original resolution are then fed to the softmax classifier to produce the final segmentation.

Output:
The final segmented images are generated at this section. These are stored with preds as the output name. The performance of the model is also evaluated at this section. The model is described by the diagram in Fig. 3.

The layout of encoder-decoder architecture is illustrated with Fig. 3.
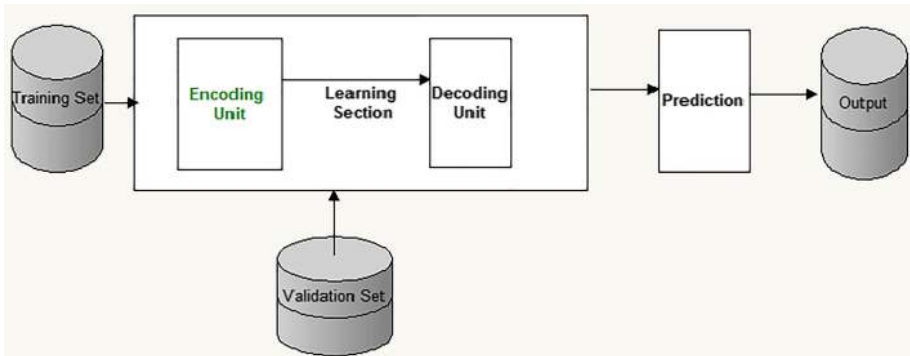

Fig. 3. Layout diagram of a proposed segmentation method for retina images analysis

3.4   Experimental Results and Analysis

The performance of the proposed model was evaluated and the results displayed below. The segmentation accuracy has been assessed by comparing the predicted results with the manual version. The final results of the validation image set used for testing segmentation are compared with the ground truth.

The metrics can be described as stated below:

- Positive Predictive Value (PPV): amount of prediction that turns positive. This was measured using Eq. 1:

$$\text{PPV}(A, B) = \frac{|A \cap B|}{|A|} \tag{1}$$

- Intersection of Union (IoU): this is also known as Jaccard similarity coefficient; it describes the similarity measured over two image sets. This can be calculated using Eq. 2:

$$\text{IOU}(A, B) = \frac{|A \cap B|}{|A| \cap |B|} \tag{2}$$

- Sensitivity: also known as true positive rate (TPR) Sensitivity: This is the proportion of actual positives which were predicted positive. It can be defined as the proportion of images which were predicted to belong to a class with respect to all the images that truly belong in the class.

$$\text{Sensitivity} = \frac{|A \cap B|}{|B|} \tag{3}$$

- Dice Coefficient: also shows similarity measure and it is related to Jaccard index. This was measured using Eq. 4:

$$\text{Dice} = \frac{|A \cap B|}{(|A| \cap |B|) \div 2} \tag{4}$$

Where A denotes the segmented region and B denotes the manually labelled region, $|A \cap B|$ denotes the overlap area between A and B, AND $|A|$ and $|B|$ represent the areas of A and B respectively.

The performance evaluation results as shown in Table 1 reveals that a high percentage of similarity and low level of diversity between the predicted results and ground truth results was achieved. The training accuracy curve on Fig. 5 shows that the accuracy is above 90%. This results can further be improved with the introduction of data augumentation.

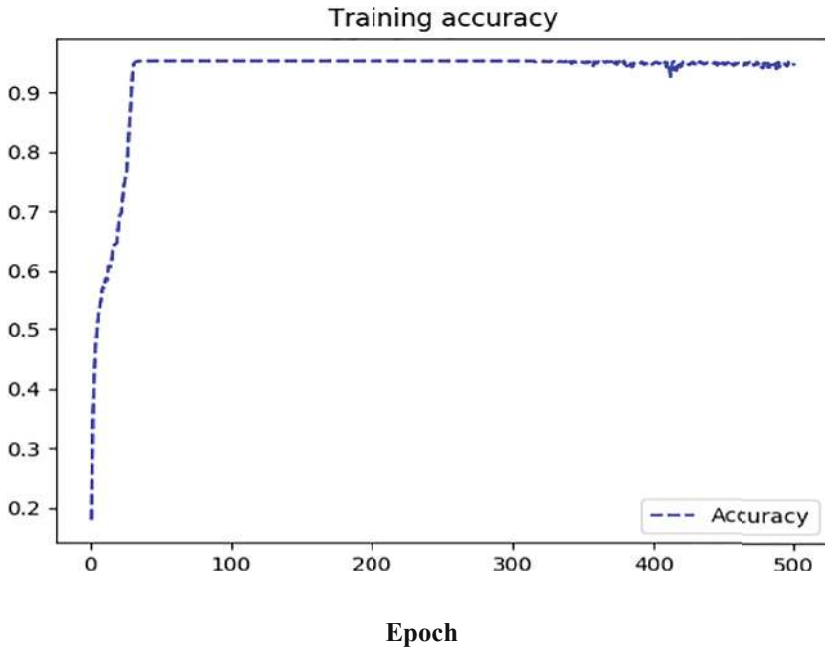**Table 1.** Performance evaluation of the proposed model

Performance metrics	Dice coefficient (%)	Intersection of union (%)	Positive predictive value (%)	Sensitivity (%)
Proposed model	82.23	72.23	96.5	96.5

The final segmented output is presented in the Fig. 4:
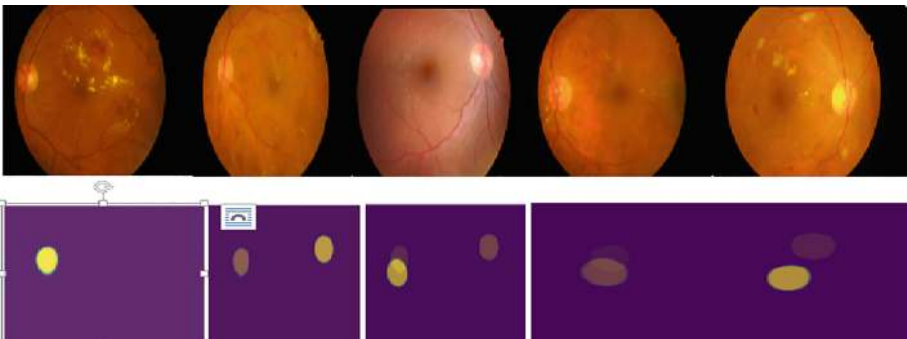
Fig. 4. The figure contains the validation image set and the predicted segmented image results from the proposed model.

Fig. 5. The figure shows the accuracy curve of the proposed model when trained on retina images dataset.

4   Conclusion

In this work, a novel deep learning method and its application on segmentation of retina images for detection and diagnosis of some diseases has been investigated. The review of the existing works was also carried out. The proposed model was evaluated and analyzed. This paper shows that the proposed model gave a better performance over the existing techniques. The encoder–decoder mechanism in this model can be improved on in the future works where ensemble methods with some other state-of-the arts techniques can also be explored. Application of data augmentation techniques will also improve the performance of the proposed system because limited data was used for the experiments. The detailed methodology which has been broken into learning, encoder, decoder, prediction was fully explored.

References

1. Akbar, S., Akram, M.U., Sharif, M., Tariq, A., Khan, S.A.: Arteriovenous ratio and papilledema based hybrid decision support system for detection and grading of hypertensive retinopathy. Artif. Intell. Med. **90**, 15–24 (2018). https://doi.org/10.1016/j.artmed.2018.06.004

2. Triwijoyo, B.K., Pradipto, Y.D.: Detection of hypertension retinopathy using deep learning and boltzmann machines detection of hypertension retinopathy using deep learning and boltzmann machines. J. Phys: Conf. Ser. **801**, 1–7 (2017). https://doi.org/10.1088/1742-6596/755/1/011001

3. Oluwatobi, A.N., et al.: Vascular networks segmented from retinal images of hypertensive retinopathy and glaucoma patients. J. Eng. Appl. Sci. (2019, in press)

4. Pound, M.P., et al.: Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. GigaScience **6**, gix083 (2017)

5. Gopinath, K., Rangrej, S.B., Sivaswamy, J.: A deep learning framework for segmentation of retinal layers from OCT images. arXiv preprint arXiv:1806.08859 (2018)

6. Guo, S., Gao, Y., Wang, K., Li, T.: Deeply supervised neural network with short connections for retinal vessel segmentation. arXiv preprint arXiv:1803.03963 (2018)

7. Fu, W., Breininger, K., Würfl, T., Ravikumar, N., Schaffert, R., Maier, A.: Frangi-Net: a neural network approach to vessel segmentation. arXiv preprint arXiv:1711.03345 (2017)

8. Ben-Cohen, A., et al.: Retinal layers segmentation using fully convolutional network in OCT images. RSIP Vision (2017)

9. Wang, X., et al.: Retina blood vessel segmentation using a U-net based Convolutional neural network. In: Procedia Computer Science: International Conference on Data Science (ICDS 2018), Beijing, China, 8–9 June 2018 (2018)

10. Melinščak, M., Prentašić, P., Lončarić, S.: Retinal vessel segmentation using deep neural networks. In: VISAPP 2015 (10th International Conference on Computer Vision Theory and Applications) (2015)

11. Badar, M., Shahzad, M., Fraz, M.M.: Simultaneous segmentation of multiple retinal pathologies using fully convolutional deep neural network. In: Nixon, M., Mahmoodi, S., Zwiggelaar, R. (eds.) MIUA 2018. CCIS, vol. 894, pp. 313–324. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-95921-4_29

12. Kang, S.H., Park, H.S., Jang, J., Jeon, K.: Deep neural networks for the detection and segmentation of the retinal fluid in OCT images. National Institute for Mathematical Sciences, Daejeon, Korea, 34047

13. Almotiri, J., Elleithy, K., Elleithy, A.: Retinal vessels segmentation techniques and algorithms: a survey. Appl. Sci. **8**(2), 155 (2018)

14. Li, M., Ma, Z., Liu, C., Zhang, G., Han, Z.: Robust Retinal blood vessel segmentation based on reinforcement local descriptions. Biomed. Res. Int. **2017**, 9 (2017)

15. Staal, J., Abràmoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. IEEE Trans. Med. Imaging **23**(4), 501–509 (2004)

16. Fu, H., Xu, Y., Lin, S., Kee Wong, D.W., Liu, J.: DeepVessel: retinal vessel segmentation via deep learning and conditional random field. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 132–139. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_16

17. Fan, Z., et al.: A hierarchical image matting model for blood vessel segmentation in fundus images. In: Computer Vision and Pattern Recognition, pp. 1–10 (2017). http://arxiv.org/abs/1701.00892

18. Hassan, M., Amin, M., Murtza, I., Khan, A., Chaudhry, A.: Robust hidden Markov model based intelligent blood vessel detection of fundus images. Comput. Methods Programs Biomed., 193–201 (2017). http://doi.org/10.1016/j.cmpb.2017.08.023
19. Güleryüz, M.Ş., Ulusoy, İ.: Retinal vessel segmentation using convolutional neural networks. In: IEEE 26th Signal Processing and Communication Applications Conference, pp. 1–4 (2018)