Hybrid UNet Architecture based on Residual Learning of Fundus Images for Retinal Vessel Segmentation

Sushma Nagdeote¹ and Sapna Prabhu²

¹Fr. CRCE, Department of ECS, Bandra, Mumbai, Maharashtra, India
²Fr. CRCE, Department of ECS, Bandra, Mumbai, Maharashtra, India

Abstract. This paper deals with the new segmentation techniques for retinal blood vessels on fundus images. This technique aims at extracting thin vessels to reduce the intensity difference between thick and thin vessels. This paper proposes the modified UNet model by incorporating ResNet blocks into it which includes structured prediction. In this work we generate the visualization of blood vessels from retinal fundus image for two loss functions namely cross entropy loss and Dice loss where the network classifies several pixels simultaneously. The results shows higher accuracy by considering a much more expressive UNet algorithm and outperforms the past algorithms for Retinal Vessel Segmentation. The benefits of this approach will be demonstrated empirically.

Keywords: RVS; ResNet; UNet; ROI; retinal images; fundus

1. Introduction
The study and analysis of Retinal Vessel Segmentation i.e. RVS is the basis of medical application which can be related to prior scrutiny of retinal diseases which can bulge to vision loss. Early diagnosis can prevent vision loss. The main causes of vision loss is diabetic retinopathy and maculopathy. The computerized image analysis is applied to various medical applications. Advanced machine learning depended techniques is applied to different domains of medical imaging. These approaches outperforms the current practice dependent on the image processing procedures. The main dominance of practicing leading-edge machine learning for classification and segmentation is suitability of training of huge database which makes such approach effective from previously implemented techniques. A deep Convolutional Neural Network i.e. D-CNN framework was devised for performing diverse imaging task on ordinary images. In retinal imaging the D-CNN is used to detect vessels but the efficiency subsides as the arteries and veins broadness gets cramped. This experimentation targets on retinal fundus image for vessel extraction by using modified UNet architecture. There are openly accessible datasets for fundus images DRIVE, STARE, CHASEDB1, HRF dataset, which can be used for the interpretation of retinal vessel disunion algorithm. This work focuses on Digital Retinal Images for Vessel Segmentation.

2. Related Work
Small changes in thin retinal vessels which cannot be seen by bare eye can only be detected using various computer algorithms and we can translate these findings into something useful which can be beneficial to ophthalmologist. For this we have gone through research papers which tells us about the recent techniques and developments in the field of retinal imaging. Lei Zhou et al. [1] proposed methods to improve the functioning of the dense CRF model and outstrips rest of the technique when assessed with reference to F1-score, Mathews correlation coefficient and G-mean.
Md. Tahsin Mostafiz et al. [2] proposed modified methods for fuzzy classifier and UNet autoencoder along with residual blocks. For Fuzzy classifier they considered mean and median property of fundus images for feature extraction and then use a fuzzy interface to extract vessel. They achieved accuracy 96.75% on test data of DRIVE dataset. Wang Xiancheng et al. [3] applies deep learning techniques to the retinal blood vessel segmentations. They used shape, size and arteriovenous crossing types of fundus images for determining numerous eye diseases. The accuracy for the proposed UNet was 97.90% for DRIVE dataset. Chen X. Qi et al. [4] devised deep contour network in accordance with FCN to speculate gland and contour unifying multilevel peculiarities.

Kai Hu et al. [5] proposed a novel retinal vessel segmentation scheme based on fully convolution network and Conditional random fields. They evaluated their scheme on DRIVE and STARE datasets. Zitao Zeng et al. [7] suggested altered network framework to integrate more discriminative as well as multiscale feature. Revised UNet, downsampling module to contain both residual[6] and inception segment. To derive more robust peculiarities residual connections turns out superior to interpret image features. Inception segment is recognized for its computational efficiency at the same time embodying multiscale peculiarities with dissimilar kernel dimensions. They included inception segment together with residual segment as RI segment which involves broader reception field and finer feature representation. Recently channel attention scheme turned out to be very successful in retrieving additional representative features for image classification in addition to semantic segmentation. Dhimas Arief Dharmawan et al. [8] proposes novel amalgamated approach for retinal vessel segmentation. This techniques highlights new enhancement techniques relying on modified Dolph-Chebyshev type 1 function matched filter. They modified matched filters to ensure that their approach detect all blood vessels with all calibers. They have used this enhanced image as input to modified UNet framework. The idea behind using this enhancement technique was to reinforce the blood vessel feature to ensure that the model is capable of learning from exceedingly imbalanced proportion among vessels and non vessels pels and narrow and broad vessels. Jose Ignacio Orlando et al. [9] suggested a novel approach for blood vessel segmentation relying on fully connected conditional random field (CRF) that is widely used to solve various segmentation problems in various application. The researcher proposed the segmentation procedure as an energy minimization challenge. To improve the accuracy in CRF every single pel corresponds to a node and every single node is bonded by an edge to its immediate neighbor i.e every single node is expected to serve as neighbor of one another. In this technique there is a long term correlations among pels of Uniary potentials and pair wise potentials add together to model energy function. Uninary potentials defines log likelihood over labeled data. Pair wise potentials defines interaction between pixels features and their labels. Rezty Amalia et al. [10] discusses about various segmentation technique by various researchers for retinal blood vessel segmentation. The author compared all these techniques and came up with conclusion that the best method for retinal vessel segmentation is morphological and multiscale method. This conclusion was based on the accuracy of the segmentation method. According to the authors review Dice, F1-score and Jaccard index are majorly used to evaluate the segmentation techniques. Wang et al.[12] used CNN as hierarchical feature extractor. These methods usually fails to capture some of good vessels. As per author these vessels are not important for clinical purpose. Xiuqin Pan et al. [13] proposed residual UNet which is the modified network of UNet. They have combined UNet with residual segment. For improving segmentation result they used Nesterov Adaptive Moment Estimation to train the network (Nadam). It has stronger constraints on learning rate and update of gradient. Deep Residual UNet may boost the execution of training process and minimizes issues related to vanishing gradient. To design a deeper network the hopping connection of residual segment can assure the upgradation of information. O. Ronneberger et al. [14] proposed guidelines to researchers to preserve some details while retaining most areas of interest. They suggested a UNet framework in accordance with residual learning to resolve the problems related to vascular segmentation. C. Y. Wu et al. [16] associates the residual framework with densely connected network (DenseNet) to significantly use again the vascular peculiarities which leads the framework to discover more robust morphological structure. H. Z. Fu et al. [15] transformed the vascular segmentation problem into a boundary detection problem, identified the features using a fully connected CNN and through the Fully Connected Conditional Random Fields (CRF) they generated a vessel probability map and then generated a dense global pixel correlation. L. L. Ming et al. [17] combined UNet and DenseNet
for fundus retinal vessels segmentation and achieved higher network performance. X. R. Gao et al.[18] addressed the issues regarding the robustness of segmented micro-vessels. The capability of geometric transformation modeling still comes from artificial features and data expansion, and up-sampling layer of the decoding part cannot recover the details of the coding loss. C. Qingcui et al. [19] proposed an approach of Local Adaptive Gamma Correction. Authors used to conduct Gamma matching according to different pixel features of blood vessels and background, so as to suppress the uneven light factor and the reflection of the center line and corrects the contrast of different regions in retinal images. J. Staal et al. [20] presented automatic ridge based vessel segmentation of retinal images for screening of diabetic retinopathy. The authors partitioned images into patches and computed its feature vector and used kNN classifier for classification. For this purpose they used Utrecht database and the Hoover database. The authors goal in their research was to classify each and every image pixel as vessel or non-vessel.

3. Challenges in Retinal Imaging

Retinal vessel segmentation i.e. RVS is the primary step preceding processing and identifying eye diseases. Many research work has been conducted for RVS but automatic classification of the segmented vessels have received less attention. There are some challenges which makes this RVS difficult. First factor is low contrast of fundus images as different blood vessels have different contrast with background i.e. thicker vessels have high contrast than thinner. Second factor is Inhomogeneous lighting of the background which is caused by imaging process itself. Generally veins are thicker and are darker than arteries, in many cases it becomes difficult to distinguish arteries from veins. In the low contrast retinal images the vessels in the peripheral areas are blackish because of overshadowing effect resulted from heterogeneous illumination of fundus images. In such lighting conditions arteries and veins seems undoubtedly identical that actuate misclassification.

Mohammad et al. [11] discussed constraints and conventional proposals of deep learning for various organ segmentation. To improve segmentation accuracy deep learning models, it should be capable to tackle sophisticated structures. To procure this accuracy the network requires generous amount of labelled instances for carrying out training. To collect massive amount of annotated dataset for medical images and annotating them is very expensive and tedious task. Several solution has been addressed for the limited annotated data problem. The most commonly used approach to address this problem is data augmentation. Transfer learning is alternative approach to tackle this problem and in contrast to the first approach transfer learning imparts explicit solution and requires more parameters. Another approach is patch-wise training in which image is decomposed into several overlapping or random patches. Higher accuracy is witnessed in overlapping patches and lower accuracy in random patching. Class imbalance is seen in Random patching and is not advisable in small organ segmentation. The hybrid approach to address this issues is to use unsupervised learning methods to derive additional precise information from weakly labeled data. These extracted annotated data can be used to train the network. In 3 D cases since fully annotated data is not always available sparsely labeled data must be used. In sparsely annotated dataset the weights of unlabeled data are set to zero to learn from labeled data.

In medical imaging the anatomy of region of interest i.e. ROI (occupying small portion) of images is of great importance in anomaly or lesion detection. Hence most of the patches extracted belongs to the background region. Training with such dataset results the trained network biased towards the background and gets trapped in local minima. Sample reweighting is used as a solution to this problem where higher weight is applied to foreground patches during training. Automatic sample reweighting is developed by Dice loss layer and Dice Coefficient. Still the effectiveness is limited in dealing with extreme class imbalance. A mechanism to have approximately balanced number of patches from background and foreground should be in place in order to deal with extreme class imbalance. It is also noted that the presence of hemorrhages in retina makes the blood vessel segmentation more challenging. The entire process is treated separately into dual segments first is training and second is testing. We applied thresholding on labeled images by replacing entire pixel levels equivalent to and in range of 127 to 255 to 1 and levels inferior to 127 is assigned 0 and converted them into one channel gray scale image. Normalize all the images along with their labeled images and rescale to 512 ×512. The Residual UNet model has encoder decoder architecture. Input images are processed by conv2D layer followed by two residual block as shown in Figure. 1. uses conv layer as hierarchical feature extractor to produce more
A ResNet section subsists of batch normalization in conjunction along a pair of subsequent convolutional layers. Batch normalization play the role of regularizer to accelerate discriminative feature learning and helps in achieving training speed up from higher learning rates. Activation layers is applied after each Conv2D module in order to add non linearity. Batch normalization layer throughput is channeled to the concluding convolutional layer of the ResNet section throughput. This results ResNet section acting as a feedback nexus utilize residues through preceding layer. The drop-out layer after each max-pooling layer is added.

4. Results
We used publicly available retinal image database DRIVE for training and evaluating our model on DRIVE, HRF and CHASEDB1 data set. Training of the model was done with binary cross-entropy and dice coefficient loss functions and the training process was carried for 100 epochs. We also experimented our model on HRF and CHASEDB1 dataset as shown in Figure 3 and 4 respectively. Also the images from HRF and CHASEDB1 data set are tested with the trained DRIVE dataset. We evaluated the Sensitivity and Specificity of our method for DRIVE, HRF and CHASEDB1 data set.. It is observed that as training progresses the training accuracy increases and training loss decreases. The training accuracy curve and training loss curve for CHASEDB is shown in Figure. 5(a) and (b) respectively. The training accuracy and loss curve for HRF dataset is shown in Figure. 5(c) and (d) respectively. We conducted the investigation of our scheme with respect to sensitivity(Se) and specificity(Sp) as shown in Table I. Se is true positive rate refers to the tests potential to characterize an individual along disease as positive. High Se signifies that there are meager faulty negative outcome. Sp is true negative rate refers to the tests potential to characterize an individual not having a disease as negative. High Sp signifies that there are meager faulty positive outcome. It is recommendable to possess a test especially
being highly sensitive and specific. This is intermittently impossible. Generally there exists tradeoff. After doing extensive testings it is observed that certain subjects are apparently healthy certain subjects are apparently unhealthy and certain subjects lie in the midst of them. So the criteria must be made for positive and negative results. Confusion matrix aka error matrix is frequently accustomed to depict the behavior of model based on group of test dataset for which the actual values are familiar.

The HRF and CHASEDB1 dataset are tested on trained DRIVE dataset. Their performance indicators are recorded and hence we can comment that this algorithm works well in unseen data. Secondly we also trained DRIVE, HRF and CHASEDB1 dataset separately and their performance indicators are recorded. The Figure 2. (a), (b) and (c) shows Test image for the Drive dataset, ground truth image and the result of segmentation map for Drive dataset respectively. The Figure 3. (a), (b) and (c) shows Test image for the HRF dataset, ground truth image and the result of segmentation map for HRF dataset respectively. The Figure 4. (a), (b) and (c) shows Test image for the CHASEDB1 dataset, ground truth image and the result segmentation map for CHASEDB1 dataset respectively.

5. Performance Evaluation
The intended computation is assessed with regard to sensitivity (Se), specificity (Sp), accuracy (Acc), f1 score, G and MCC as shown in Table I. The performance indicator is calculated in terms of Tp, Tn, Fp and Fn. Where Tp, Tn, Fp and Fn is True Positive, True Negative, False Positive and False Negative accordingly. Se is the potential of the model to accurately speculate the blood vessel and Sp is the
potential of the model to accurately speculate the non-vessel. G-mean is the root of the product of class wise sensitivity. This indicator maximizes the each class accuracies while keeping the accuracies balanced. Positive predicted value (PPV) is the probability of correctly discovered vessels which are actually vessels. Mathews correlation coefficient (MCC) is a correlation among the predicted throughput and ground truth.

It is recorded that intended computation is superior than all the other contemporary schemes. This is because performing retinal image segmentation in fundus images is more challenging. This algorithm has a capability of detecting micro vessels in low contrast images. As well as has capability to detect blood vessels with central reflex. The findings of different researchers are compared with our hybrid method and is shown in Table 1. It is seen that intended computation of our hybrid method is more superior than all the other methods.

Figure 5. (a) CHDB Training Accuracy Curve (b) CHDB Training Loss (c) HRF Training Accuracy Curve (d) HRF Training Loss Curve
Table 1. Comparison of Performance metrics on different datasets.

| Dataset         | Se    | Sp    | Acc   | PPV   | f1_score | G    | MCC  |
|-----------------|-------|-------|-------|-------|----------|------|------|
| Mostafiz et al. [2] | DRIVE 0.7690 | 0.9865 | 0.9675 | 0.8453 | -        | -    | -    |
| Kai Hu et al. [5]  | DRIVE 0.7772 | 0.9793 | 0.9533 | -      | -        | -    | -    |
|                 | STARE 0.7543 | 0.9814 | 0.9632 | -      | -        | -    | -    |
| Xurong Gao et al. [18] | DRIVE 0.7802 | 0.9876 | 0.9636 | 0.8948 | -        | -    | -    |
| Le Zhou et. al. [1] | DRIVE 0.8078 | 0.9674 | 0.9469 | 0.7861 | 0.7942   | 0.8835 | 0.7656 |
|                 | STARE 0.8065 | 0.9761 | 0.9585 | 0.8113 | 0.8107   | 0.8859 | 0.7830 |
|                 | CHDB1 0.7553 | 0.9751 | 0.9520 | 0.7804 | 0.7644   | 0.8579 | 0.7398 |
|                 | HRF    0.8015 | 0.9699 | 0.9544 | 0.7325 | 0.7627   | 0.8812 | 0.7402 |
| Dhimas Arief Dharmawan et. al. [8] | DRIVE 0.8314 | 0.9726 | -      | 0.8157 | 0.8235   | 0.8992 | 0.7991 |
|                 | STARE 0.7924 | 0.9827 | -      | 0.8412 | 0.8161   | 0.8824 | 0.7959 |
|                 | HRF    0.8136 | 0.9770 | -      | 0.7838 | 0.7984   | 0.8916 | 0.7775 |
| Proposed Method  | DRIVE 0.7543 | 0.9927 | 0.9778 | 0.8956 | 0.8285   | 0.8622 | 0.8685 |
|                 | HRF    0.7233 | 0.9871 | 0.9698 | 0.8266 | 0.7715   | 0.8450 | 0.8364 |
|                 | CHDB1  0.7403 | 0.9894 | 0.9839 | 0.8391 | 0.7867   | 0.8558 | 0.8445 |

The algorithm correctly detects vessels with Se 0.7543, 0.7233 and 0.7403 for DRIVE, HRF and CHDB dataset respectively. This algorithm is useful for practical applications. The algorithm predicts non vessels with Sp 0.9927, 0.9871 and 0.9894 for DRIVE, HRF and CHDB dataset respectively. The accuracy recorded for DRIVE, HRF and CHDB dataset is 0.9778, 0.9698 and 0.9839 respectively. The PPV recorded is 0.8956, 0.8266 and 0.8391 for DRIVE, HRF and CHDB dataset respectively. The f1_score is 0.8285, 0.7715 and 0.8558 for DRIVE, HRF and CHDB dataset respectively. G-mean score recorded is 0.8622, 0.8450 and 0.8558 for DRIVE, HRF and CHDB dataset respectively. The correlation among the segmentation throughput and ground truth is recorded as 0.8685, 0.8364 and 0.8445 for DRIVE, HRF and CHDB dataset respectively.

6. Conclusion and Future Scope

This research article introduces a modified computation for retinal blood vessel segmentation on fundus images originated from UNet. The findings are shown experimentally on openly accessible dataset as mentioned in section I. which outperforms state of art algorithms. This algorithm also works in challenging situation similar to low contrast imaging of microvasculature and vessels in presence of abnormality. We have explored this experiments on unseen retinal data. We have trained DRIVE dataset and tested HRF and CHASEDB1 on trained DRIVE dataset. We also have trained the above three mentioned dataset separately and recorded their performance indicators in Table I. The accuracy recorded with proposed method for DRIVE, HRF and CHDB dataset is 0.9778, 0.9698 and 0.9839 respectively.

We can further exploit the proposed algorithm in other medical applications for segmentation of vessel like structures. Effective and true candidate lesion detection depends on the accuracy belonging to retinal vessel segmentation which might be used for practical research or practical medical treatment. Ophthalmologist will be benefited as different deep learning algorithms can be used for automatic retinal vessel segmentation and classification so that they can take correct decision.

References

[1] Lei Zhou, Qi Yu, Xun Xu, Yun Gu, Jie Yang, “Improving Dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin vessel enhancement” in Computer Methos and Programs in Biomedicine, ELSEVIER, 2017, pp. 13-25.

[2] Md. Tahsin Mostafiz, Ismat Jarin, Shaikh A. Fattah, Celia Shahnaz, “Retinal Blood Vessel Segmentation using Residual Block Incorporated U-Net Architecture and Fuzzy Inference
[3] Wang Xiancheng, Li Wie, Miao Bingyi, Jing He, Zhangwei Jiang, Wen Xu, Zhenyan Ji, Gu Hong, Shen Zhanomneg, “Retinal Blood Vessel Segmentation using a U-Net Based Convolutional Neural Network”, International conference on Data Science(ICDS 2018), Published by Procedia Computer Science

[4] Chen, X. Qi, L.Yu and P.-A. Heng, “ DCAN:Deep counter-aware networks for accurate gland segmentation,” in proc. Int. Conf. Comput.Vis. Pattern Recognition(CVPR), Jul.2016, pp. 2487-2496

[5] K. Hu, Z. Zhang, X. Niu, Y. Zhang, C. Cao, F. Xiao, et al., "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function", Neurocomputing, vol. 309, pp. 179-191, Oct. 2018.

[6] K. He, X. Zhang, S. Ren, and J.Sun, “Deep Residual learning for image recognition,” in in proc. Int. Conf. Comput. Vis. Pattern Recognition(CVPR), June 2016, pp. 770-778

[7] Zitao Zeng, Weihao Xie, Yunzhe Zhang, Yao Lu, “RIC-Unet for Nuclei Segmentation in Histology images”, IEEE open access jurnal, special session on Deep Learning for Computer-Aided Medical Diagnosis., DOI:10.1109/ACCESS.2019.2896920, February 27, 2019

[8] Dhimas Arief Dharmawan, Di Li, Boon Poh NG, Susanto Rahardja“A New Hybrid Algorithm for Retinal Vessels Segmentation on Fundus Images” 2169-3536 © Volume 7 2019 IEEE. Translations and content mining, Digital Object Identifier 10.1109/ACCESS.2019.2906344

[9] José Ignacio Orlando, Matthew Blaschko “Learning fully-connected CRFs for blood vessel segmentation in retinal images” Medical Image Computing and Computer Assisted Intervention (MICCAI), Sep 2014, Boston, United States. 10.1007/978-3-319-10404-1_79, hal-01024226

[10] Rezty Amalia Aras, Tri Lestari, Hanung Adi Nugroho, Igi Ardiyanto, “Segmentation of retinal blood vessels for detection of diabetic retinopathy: A review” Communications in Science and Technology 1 (2016) 33-41

[11] Mohammad Hesam Hesamian, Wenjing Jia, Xiangjian He & Paul Kennedy “Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges” Journal of Digital Imaging (2019) 32:582-596, https://doi.org/10.1007/s10278-019-00227-x

[12] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, G. Yang, “Hierarchical retinal blood vessel segmentation based on feature and ensemble learning, Neurocomputing, 149(2015) 708-717

[13] Xiuqin Pan, Qinrui Zhang, And Sumin Li, “A Fundus Retinal Vessels Segmentation Scheme Based on the Improved Deep Learning U-Net Model” Special Section on Deep Learning Algorithms for Internet of Medical Things. Digital Object Identifier 10.1109/ACCESS.2019.2935138

[14] O. Ronneberger, P. Fischer, and T.Brox, “ U-Net: Convolutional Networks For Biomedical Image Segmentation”, in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent., 2015, pp. 234-241.

[15] H. Z. Fu, Y. W. Xu, and D. W. K. Wong, “Retinal Vessel Segmentation via Deep Learning Network and Fully Connected Conditional Random Fields,” in Proc. IEEE Int. Symp. Biomed. Imag., Jul 2016, pp. 698-701

[16] C. Y. Wu, B. S. Yi and Y. G. Zhang, “Retinal Vascular Image Segmentation Based On Improved Convolution Neural Network,” Acta Optica Sinica, vol. 38, no. 11, 2018, Art. no. 1111100

[17] L. L. Ming and X. S. Qi, “Improved U-Net Fundus Retinal Vessels Segmentation,” Appl. Res. Comput., vol. 6, pp. 1-6, April 2019

[18] X. R. Gao, Y. H. Cai, and C. Y. Qiu, “Retinal Blood Vessel Segmentation Based On The Gaussian Matched Filter And U-Net,” in Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Inform. (CISP-BMEI). Aug. 2017, pp. 1-5

[19] C. Qingcui, W. Huabin, and T. Liang, “Local Adaptive Gamma Correction of Images,” Comput. Eng. Appl. Vol. 51, no. 7, pp. 189-193, 2015

[20] J. Staal. M. D. Abramoff, and M. Niemeijer, “Ridge-Based Vessel Segmentation in Color Images Of The Retina,” IEEE Trans. Med. Imag., vol. 148, pp. 13-25, Apr. 2017.