IMPLEMENTATION OF ABBASI ET AL'S PAPER: "LMBIS-NET: A LIGHTWEIGHT MULTIPATH BIDIRECTIONAL SKIP CONNECTION BASED CNN FOR RETINAL BLOOD VESSEL SEGMENTATION"

Milad Farazian, Charlie Floeder, Rizq Khateeb, Harshit Shah, Yash Sharma
Department of Computer Science
University of Southern California
Los Angeles, CA 90089, USA

1 INTRODUCTION

The goal of this project is to implement the LMBiS-Net model that is presented in "LMBiS-Net: A Lightweight Multipath Bidirectional Skip Connection based CNN for Retinal Blood Vessel Segmentation." (1) In addition to implementing LMBiS-Net, we decided to apply the model to a different dataset that is not used in the paper. The dataset we chose was FIVES, a Fundus image dataset for AI-based vessel segmentation. The authors of Abbasi et al. did not make their implementation of LMBiS-Net publicly available, and thus we worked on creating our own implementation of LMBiS-Net from scratch to see if we were able to replicate the results achieved by Abbasi et al. To further test the robustness of our model, we further tested on the FIVES dataset to confirm that our model can be used with other datasets beyond just CHASE_DB1 and STARE.

1.1 IMPORTANCE

It's not uncommon for researchers to purposely choose datasets where their models perform particularly well. Applying the model to a new dataset will add more credibility to the generalizability of LMBiS-Net. LMBiS-Net's primary benefit is that it provides an accurate retinal blood vessel segmentation model that is more computationally efficient compared to current state-of-the-art models. This CNN model is applied to retinal imaging. The LMBiS-Net research paper claims that retinal blood vessel segmentation is beneficial for early detection and treatment of retinal diseases. This model can significantly reduce the amount of time ophthalmologists spend manually identifying retinal vessels. Additionally, the model can reduce human error in that task. Retinal diseases are a major cause of visual impairment and blindness (12). Studies show that 5%-20% of the global population ages 40+ have retinal disorders (12). Examining retinal vessels can provide important information regarding the underlying medical conditions leading to retinal diseases.

1.2 PROJECT CONTRIBUTIONS

Our team created the first publicly available implementation of LMBiS-Net along with code to augment retinal images to increase the size of training datasets. Our findings add credibility to the original paper's claims that LMBiS-Net is a computationally efficient and accurate state-of-the-art retinal blood vessel segmentation model.

2 RELATED WORK

Image segmentation is a well established research area in the medical field. As early as 2004, Staal et al. (11) implemented neural network classifiers and sequential forward feature selection to automate blood vessel segmentation from retinal images. They set a baseline for work regarding blood vessel image segmentation that has since been changed as many newer models use the U-Net model.

Over the years, continuous new advancements were made in blood vessel segmentation, particularly in implementing new methods. Fraz et al. (5) presents a new supervised technique for segmenting blood vessels in retinal photographs that uses ensemble learning using Decision Trees. Their feature extraction techniques are based on methods of traditional digital image transformations such as morphological transformations or Gabor filters. The datasets tested in the study included CHASE_DB1 and STARE, both of which we explored with our LMBiS-Net implementation.

Fathi and Naghsh-Nilchi's (4) work presented another new multi-scale vessel enhancement method based on complex continuous wavelet transform (CCWT). Their method is different from the other approaches described as they rely on digital transformations to segment the retinal images. Their method sets up good benchmarks on the same datasets used other papers as well as ourselves, such as CHASE_DB1 and STARE. In addition, since the method described by Fathi and Naghsh-Nilchi does not use convolutional neural network (CNNs), we could compare this method with CNN-based methods to consider various approaches for the same problem.

Despite all of the earlier methods that were designed, none of them had as strong an impact on blood vessel segmentation as U-Net. In 2015, Ronneberger et al. (10) introduced U-Net, a type of CNN, designed with intentions to be used for biomedical image segmentation and proven to be very accurate and efficient as opposed to preceding methods. The U-Net architecture introduces skip connections to preserve features. The skip connections between encoder and decoder layers prove to be a significant evolution in this field as U-Net is considered a state-of-the-art method and possibly an inspiration for future models to come, including LMBiS-Net.

Some studies sought to categorize the images based on features. In 2018, Yan et al. (14) published their work on blood vessel segmentation model that targets the different problems associated with segmenting thick and thin vessels. They propose to approach the segmentation with a three stage deep learner architecture, which separates learning into thin, thick and fusion elements. It is an innovative and practical idea to treat the imbalance of thick and thin elements in medical image segmentation. However, their model size and scalability are not discussed.

By building upon Full Convolutional Neural Networks (FCNs) that additionally included a new multi-scale input block, Jiang et al's (7) Multiscale Multimodal Multiobjective Fully Convolutional Network (M3FCN) initiated a new kind of automatic retinal vessel segmentation framework. M3FCN is an end-to-end framework that automatically and efficiently performs retinal vessel segmentation, and was assessed once again on the CHASE_DB1 and STARE datasets. Experimental results highlighted the framework's high vessel segmentation performance for all of their criteria. Abbasi et al.'s LMBiS-Net compares their own results with M3FCN's on the datasets to evaluate if they improve upon their predecessor's results. It is a useful benchmark to see if their advancement of the U-Net is headed in the right direction.

Like Yan et al., Wang et al. (13) also classified images into different categories before segmenting, yet with different criteria. The team created a multi-decoder model that first decides which parts of the image are 'hard' or 'easy' to analyze, and from there, images are classified as either 'hard' or 'easy' based on their features. Their implementation uses attention mechanism to focus on 'hard' segmentation parts in the image. Their implementation also does not mention model size and scalability issues.

Du et al. (3) provides a comprehensive overview of medical image segmentation using U-Net. Their review focuses on the application of U-Net in various medical imaging systems like CT, MRI, ultrasound, OCT, PET, and X-ray. It discusses the advantages of U-Net in accurately segmenting target features, efficiently processing medical images, and aiding in precise medical diagnoses. This overview illustrates various ways in which U-Net can be enhanced to fit specific use cases within the medical field, and since LMBiS-NET built upon the architecture of U-Net by introducing multipath feature extraction blocks and bi-directional skip connections, Du et al.'s paper presents practical cases of improvement by LMBiS-Net.

All of these papers contributed to LMBiS-Net, given its utilization of these aspects of the U-Net design. LMBiS-Net, however, also introduces modifications such as multipath feature extraction blocks and bidirectional skip connections, which enhance its performance. LMBiS-Net is optimized for segmentation of retinal blood vessels and designed to be lightweight with a lower number of parameters, making it more efficient, especially in terms of computational resources and training time.

U-Net was used as a base model by many research studies, and their resulting models often sow similarities to LMBiS-Net, which is itself based off of U-Net. Zhang et al. (16) presented Bridge-Net, which used U-Net as its base model, as a CNN that uses large-scale patches to give contextual information to smaller-scale patches, thus helping boost efficiency. When processing vascular features, Bridge-Net uses large-scale patches to add contextual information to small-scale patches.

All of the previously mentioned work suffers from the same problem: computational complexity caused by the multipath structure to segment thick and thin blood vessels. Additionally, these models can struggle with segmenting thin vessels because some of the vessel features are lost during the successive convolution and pooling operations. LMBiS-Net addresses both of those issues. This model uses few model parameters compared to the existing body of knowledge, which addresses the issue of high computational complexity. To reduce the spatial information loss of thin blood vessels, LMBiS-Net only uses two max-pooling layers.

The introduction of Dual-Encoder Fusion Network DEF-Net (9), which adds a dual-encoder and decoder to an already established U-Net model, differs from all of the previous alterations to U-Net. Li et al.'s DEF-Net uses multiple paths involving convolution and adds depth to the fusion feature process by having two encoders instead of one, which provides us with better features. These fundamental changes are what significantly influenced LMBiS-Net's architectural enhancements. The integration of both the multipath feature extraction blocks and bidirectional skip connections in our model stemmed from these concepts, boosting efficiency and precision.

Deshmukh and Roy's paper (2) focuses on diabetic retinopathy, a major cause of blindness, and presents an automatic detection method for blood vessel segmentation using a U-net architecture. Their method involves a multi-layer convolutional neural network for segmentation, employing a Gaussian filter for feature extraction after preprocessing and before passing it to the U-Net for segmentation. Abbasi et al.'s paper, also deals with retinal blood vessel segmentation, but it emphasizes a "lightweight multipath bidirectional skip connection based CNN" for this purpose. The multipath feature extraction block builds upon the previous work to generalize retinopathy and pave way for lighter and scalable segmentation models.

Despite the fact that all of these papers proposed new and potentially preferable methods, we stuck with LMBis-Net due to reliability and consistency. Our results depend on furthering the techniques used in Abbasi et al. (1). Furthermore, the methods used in the other papers were tested, but only on certain datasets or in certain situations. This does not guarantee their reliability in the cases we are using LMBiS-Net for. LMBiS-Net has been tested over a variety of datasets and situations, and thus is fairly well established. Thus, we know that it will not act as a confounding variable that may affect our results and conclusions.

3 PROJECT SCOPE/STRUCTURE

3.1 LMBIS-NET ARCHITECTURE

Figure 1 shows the model architecture used for implementation. We use convolution layers and ReLU (rectified linear unit) activation functions to find important patterns in the input layer. To make training more stable and speed up the learning process, we add a batch normalization layer. After that, we use max pooling to shrink the size of the feature maps while maintaining the important details.

In the multipath feature extraction block, we parallelize the input to 1x1, 3x3, and 5x5 convolution blocks, each followed by ReLU activation and batch normalization to grab important details. After each block, we concatenate the results and get a full feature map showing what the encoder has found. The bottleneck layer is a key part that helps the model learn well by connecting the encoder and decoder smoothly.

In the decoder stage, transpose convolution layers are used to enlarge the feature maps and consequently cover more space, the opposite of the usual convolution process. At the last stage of the decoder, a softmax layer is used to present the model's output as probabilities.

Dice pixel classification determines the similarities between the predicted and actual images. Skip connections are both a notable part of neural networks and helpful in aiding with the training of deep models. Our suggested LMBiS-Net utilizes skip connections that link its encoding and decoding layers in both directions. This helps mix low-level and high-level features using forward connections while also bringing the decoded features back to the encoder using reverse connections.

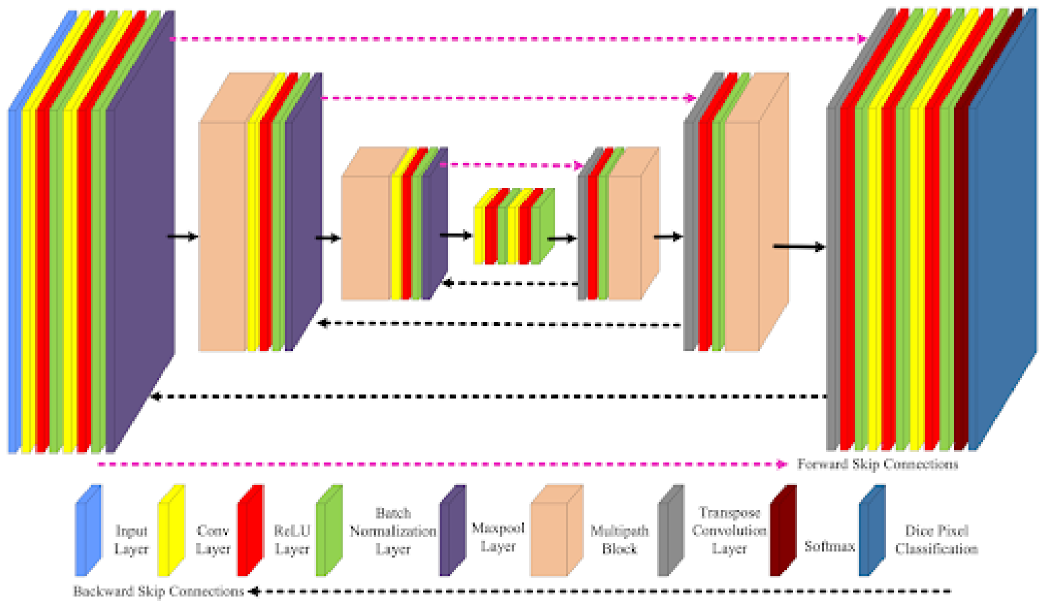

Figure 1: Block Diagram of LMBiS-Net (1)

3.2 Software Design

We implemented LMBiS-Net using Tensorflow on Google Colab. The implementation is divided into several sections: Dataset Preparation, Model Architecture, Model Training, and Model Evaluation. In Dataset Preparation, the retinal image datasets are loaded. Data augmented is used to increase the size of the datasets to match the sizes used in the original paper. The Model Architecture details all the layers of the LMBiS-Net. This CNN is composed of three encoder layers, a bottleneck layer, and three decoder layers. Our implementation uses a combination of self-built functions and functions from the tensorflow.keras library. The Model Training section simply takes the compiled model from the Model Architecture section and trains it for a specified number of iterations. The Model Evaluation displays the evaluation metrics (See Section Below) after feeding the test dataset to the trained model. All this code is packaged into an easy-to-use file where our implementation of LMBiS-Net can be executed by simply running all cells in the Google Colab.

4 Methodology and Data Description

Our first goal was implementing LMBiS-Net from scratch, and then using our LMBiS-Net on the Chase_DB1 (5) and STARE(6) datasets that were used in Abbasi et al. After this, we chose to further test our implementation on the FIVES (8) dataset. Although each dataset had different image resolution, each image was resized to 512x512 pixels before model training. Data augmentation techniques were used to increase the training images. Augmentation techniques used include random rotation, random cropping, random brightness, random contrast, and random flipping.

Chase_DB1 is a dataset commonly used in retinal blood vessel segmentation. It contains 28 color retina images collected from the left and right eyes of 14 children, and the images were annotated by two human experts independently.

STARE, or Structured Analysis of the Retina, is another well known dataset when it comes to blood vessel segmentation. It consists of 20 color fundus images, with two sets of annotations for each image.

The last dataset we chose to implement, and one that was not explored by Abbasi et al. was the FIVES dataset. FIVES consists of 800 multi-disease color fundus images. Annotations were crowd-sourced and are available for every image in the dataset.

4.1 ATTEMPTED IDEAS

We opted to retain LMBiS-Net rather than exploring alternative methods or models for several reasons. Most notably, the original paper by Abbasi et al. employed LMBiS-Net, thus influencing our decision to maintain consistency. We reconstructed the model from the ground up and applied it to both the CHASE and STARE datasets, enabling a direct comparison of results. This approach allowed us to assess the extent of dissimilarity, if any, between our implementation and the model used by Abbasi et al. Choosing a different model could introduce confounding variables, making it challenging to attribute changes in results to the correctness of our implementation rather than the choice of model.

5 RESULTS

5.1 EVALUATION METRICS

Sensitivity(S_e), Specificity(S_p), Accuracy(A_cc), F1-Score, and Area Under the Curve (AUC) were the metrics used in the original paper to evaluate the performance of the LMBiS-Net. The LMBiS-Net model is essentially a binary classification of every pixel in the image 1 indicates a blood vessel, 0 indicates not a blood vessel. The Formulas:

[T_P = true positive, T_N = true negative, F_P = false positive, F_N = false negative]

$$S_e = \frac{T_P}{T_P + F_N} \qquad S_E = \frac{T_P}{T_P + F_N}$$

$$A_{cc} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}$$

$$F_1 - Score = \frac{2 \times T_P}{(2 \times T_P) + F_P + F_N}$$

$$AUC = 1 - \frac{1}{2} \left(\frac{F_P}{F_P + T_N} + \frac{F_N}{F_N + T_P}\right) \qquad IoU = \frac{T_P}{T_P + F_P + F_N}$$

5.2 ANALYSIS OF RESULTS

Dataset	AUC	IoU	Specificity	Dice Coef.
CHASE _ DB1	0.8688	0.6562	0.9493	0.7931
STARE	0.6835	0.4300	0.9870	0.6100
FIVES	0.8321	0.7030	0.9959	0.8147

Table 1: Results Achieved with Our Team's Implementation of LMBiS-Net

Table 1 shows the results that we achieved by using our LMBiS-Net model implementation on the CHASE_DB1, STARE, and FIVES datasets. We were successful in implementing a LMBiS-Net model from scratch. As Fig. 2 shows, there are very minor differences between our predictions and the ground truth from Abbasi et al. Thus, we were able to replicate the original results achieved with the CHASE_DB1 and STARE datasets to a close degree.

Since FIVES was not used in Abbasi et al., we have no results to compare our results towards. The main idea behind including this dataset was confirming that our implementation would extend to previously unused datasets. Our results confirm that our implementation did indeed work on FIVES

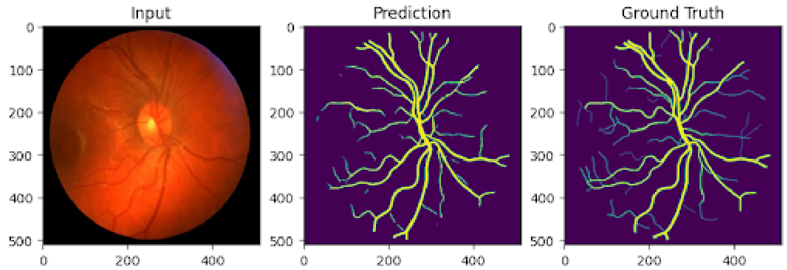

Figure 2: Retinal Blood Vessel Segmentation Example from CHASE$_D B1$

and we were able to extract results. The only comparisons we can draw are between the results achieved with FIVES and the results achieved with the other two datasets. As Table 1 shows, we achieve a high specificity with the FIVES dataset, and data in similar ranges for the other categories. This confirms that our model is well-suited to work with many datasets.

6 Conclusion / Future Work

Our team has developed the first publicly available implementation of LMBiS-Net, complete with code for enhancing retinal images to expand training datasets. Our research corroborates the assertions made in the original paper, Abbasi et al., affirming that LMBiS-Net stands as a computationally efficient and accurate state-of-the-art model for retinal blood vessel segmentation.

The limitations of our implementation were brought on by a multitude of factors. The hyper-parameters that Abbasi et al. used for their implementation were not specified in the paper, so we had to decide the weight of these parameters at our own discretion. Given more time, we could have better fine tuned these parameters to produce better results for all of our evaluation metrics. Another large limitation to our implementation was the lack of computational capabilities and processing power when working within Google Colab. Although we benefited greatly from the ease of working dynamically together on building this model due to Colab, we had to compromise on performance somewhat, which may have affected the result.

To improve our project, we could compare against papers with results beyond Abbasi et al. Doing so would further cement our results and the validity and accuracy of our LMBiS-Net model. Furthermore, we can also compare our results with studies that used different datasets than ours, and possibly even papers that used their models on the FIVES dataset. This would support our results if correct, or potentially point out any flaws in our implementation if not. Extending our implementation to more datasets and conditions would only add to the credibility of LMBiS-Net, which itself is already considered a accurate and reliable model. In addition, we could work off our limitations and work on fine-tuning parameters for better results and using a more suitable coding environment without the limitations of Colab. Although such issues did not harm our work significantly, they did make our tasks more complicated and time-consuming, and making changes on these areas would only streamline our work and contribute to accuracy and efficiency.

References

[1] Mufassir M. Abbasi, Shahzaib Iqbal, Asim Naveed, Tariq M. Khan, Syed S. Naqvi, and Wajeeha Khalid. Lmbis-net: A lightweight multipath bidirectional skip connection based cnn for retinal blood vessel segmentation, 2023.

[2] S. V. Deshmukh and A. Roy. Retinal blood vessel segmentation based on modified cnn and analyze the perceptional quality of segmented images. In *International Conference on Advanced Network Technologies and Intelligent Computing*, pages 609–625. Springer, 2022.

[3] G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan. Medical image segmentation based on u-net: A review. *Journal of Imaging Science & Technology*, 64(2), 2020.

[4] A. Fathi and A. R. Naghsh-Nilchi. Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation. *Biomedical Signal Processing and Control*, 8(1):71–80, 2013.

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

[6] A.D. Hoover, V. Kouznetsova, and M. Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, 2000.

[7] Y. Jiang, H. Zhang, N. Tan, and L. Chen. Automatic retinal blood vessel segmentation based on fully convolutional neural networks. *Symmetry*, 11(9):1112, 2019.

[8] Kai Jin, Xingru Huang, Jingxin Zhou, Yunxiang Li, Yibao Sun, Qianni Zhang, Yaqi Wang, and Juan Ye. Fives: A fundus image dataset for artificial intelligence based vessel segmentation. *Scientific Data*, 9, 2022.

[9] J. Li, G. Gao, L. Yang, Y. Liu, and H. Yu. Def-net: A dual-encoder fusion network for fundus retinal vessel segmentation. *Electronics*, 11(22):3810, 2022.

[10] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[11] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.

[12] R. Thapa, S. Khanal, H.S. Tan, S.S. Thapa, and G.H.M.B. van Rens. Prevalence, pattern and risk factors of retinal diseases among an elderly population in nepal: The bhaktapur retina study. *Clin Ophthalmol*, 14:2109–2118, 2020. PMID: 32801619; PMCID: PMC7399464.

[13] D. Wang, A. Haytham, J. Pottenburgh, O. Saeedi, and Y. Tao. Hard attention net for automatic retinal vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 24(12):3384–3396, 2020.

[14] Z. Yan, X. Yang, and K.-T. Cheng. A three-stage deep learning model for accurate retinal vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 23(4):1427–1436, 2018.

[15] XX. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang. U-net-based medical image segmentation. *J Healthc Eng*, 2022:4189781, 2022. Retraction in: J Healthc Eng. 2023 Oct 18;2023:9890389.

[16] Y. Zhang, M. He, Z. Chen, K. Hu, X. Li, and X. Gao. Bridge-net: Context-involved u-net with patch-based loss weight mapping for retinal blood vessel segmentation. *Expert Systems with Applications*, 195:116526, 2022.