



Segmentation of Retinal Blood Vessels using an Attention-Gated U-Net (AG-U-Net)

Carolyn YTW^{1,2,3*}, Timing L^{1,2}, Tin LW³ and Henry HWL⁴

¹Institute of Ophthalmology, University College London, United Kingdom

²Moorfields Eye Hospital NHS Foundation Trust, United Kingdom

³Faculty of Medicine, The Chinese University of Hong Kong, China

⁴Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, China

Research Article

Volume 9 Issue 1

Received Date: June 05, 2024

Published Date: June 19, 2024

DOI: 10.23880/oajo-16000317

***Corresponding author:** Carolyn Yu Tung Wong, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, New Territories, Hong Kong, Email: carolcarol.carolyn@gmail.com

Abstract

Retinal vessel segmentation plays a crucial role in the automated examination of fundus images for screening and diagnosing diabetic retinopathy, a common complication of diabetes leading to sudden vision loss. Automated segmentation of retinal vessels can detect these changes more efficiently and accurately compared to manual assessment by an ophthalmologist. The proposed method aims to precisely identify blood vessels in retinal images while simplifying the segmentation process and reducing computational complexity. This approach can enhance the accuracy and reliability of retinal image analysis, aiding in the diagnosis of various eye diseases. The Attention Gated U-Net architecture is a key component in retinal image segmentation for retinal pathologies like diabetic retinopathy, showing promising results in improving segmentation accuracy, especially in scenarios with limited training data and ground truth. This method involves incorporating an attention mechanism into the U-Net to focus on relevant regions of the input image, enhancing the performance of semantic segmentation models. Extensive experiments conducted on a retinal segmentation dataset demonstrated that the proposed approach outperformed existing methods in terms of performance.

Keywords: Artificial Intelligence; Image Segmentation; Retinal Blood Vessel; Colour Fundus Photograph

Introduction

The retinal vascular system contains valuable information regarding the eye's condition [1], and segmenting retinal vessels is a crucial task for analyzing eye fundus images and diagnosing fundus diseases [2,3]. This process is essential for various advanced applications, including evaluating artery/vein ratios [4], analyzing blood flow [5], assessing image quality [6], registering and synthesizing retinal images

[7,8], and aiding in the early detection of systemic vascular diseases [1]. Consequently, the automatic segmentation of retinal blood vessels from fundus images has emerged as a prominent research focus in the medical imaging domain [1].

Early methods for retinal vessel segmentation were unsupervised and utilized traditional image processing techniques such as mathematical morphology [9,10] or modified edge detection operations [11]. These approaches

Methodology

U-Net

Our proposed algorithm is based on the U-Net architecture, which is well-suited for biomedical image segmentation due to its efficiency and accuracy. The U-Net model features an encoder and decoder, as illustrated in Figure 1. The encoder, the initial component of the U-Net structure, involves downsampling and max-pooling following a convolution block to represent input images as featured images at multiple levels [21]. The decoder, situated in the latter part, utilizes convolution, upsampling, and concatenation to transform low-resolution feature images from the encoder into high-resolution pixel space, generating categorized dense blocks [22]. In this study, we employed the U-Net architecture shown in Figure 1 for retinal blood vessel segmentation.

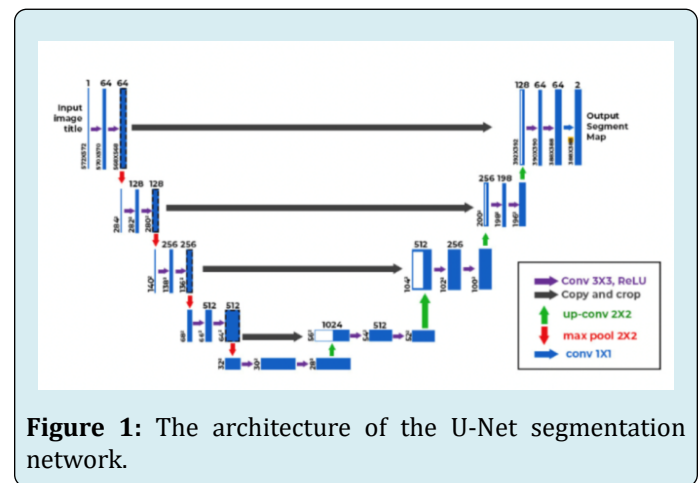


Figure 1: The architecture of the U-Net segmentation network.

Incorporating Attention in the Network

Drawing inspiration from TransNorm [23], a modified three-level gate module known as three-level attention (TLA) was developed to enhance feature focus and suppress irrelevant features [23]. The first component of TLA is the attention gate (AG), which targets essential features for a specific task while suppressing responses from irrelevant background areas [24]. During down sampling, where small objects may exhibit significant shape variations, the AG utilizes a 1×1 channel wise convolution for linear transformation. The resulting feature map outputs are combined through element wise addition, followed by ReLU activation and another 1×1 convolution with sigmoid activation. Trilinear interpolation is then applied to scale the output. In the TLA framework, the AG is strategically positioned after skip connections to optimize information fusion [25]. This approach aids in noise reduction from the background and addresses potential blurred boundaries

aimed to enhance vessel intensities in retinal images through pre-processing and subsequent thresholding for segmentation. Despite ongoing research into advanced filtering methods for retinal vessel segmentation in recent years [12,13], these techniques have struggled to achieve competitive performance on standard benchmarks. This limitation is likely attributed to their challenges in effectively handling images with pathological structures and adapting to various appearances and resolutions.

On the other hand, artificial intelligence (AI)-based methods, including machine learning (ML) and deep learning (DL), have demonstrated more promising outcomes and superior performance compared to traditional approaches [3]. Liskowski P, et al. [14] introduced a supervised segmentation technique utilizing a deep neural network trained on a large dataset preprocessed with global contrast normalization, zero-phase whitening, and augmented with geometric transformations and gamma corrections. Suryani E, et al. [15] employed a self-organizing graph artificial neural network for blood vessel segmentation, involving preprocessing, segmentation, and performance analysis stages. Wang S, et al. [16] proposed a supervised method based on feature and ensemble learning. Zhou L, et al. [17] suggested a discriminative feature learning approach using a convolutional neural network (CNN) for the dense conditional random field model. Fu H, et al. [18] treated retinal vascular segmentation as a boundary detection task, incorporating multi-scale context information, a side output layer for hierarchical structure learning, and conditional random fields for long-term pixel dependence modeling. Zhou Y, et al. [19] presented an end-to-end synthetic neural network with components like a symmetric equilibrium generative adversarial network, multi-scale features refine blocks, and an attention mechanism to enhance vessel segmentation capabilities. Xiuqin P, et al. [20] proposed a retinal vessel segmentation method based on an improved deep learning U-Net model to address performance degradation issues in residual networks with extreme depth.

To address the challenges of low segmentation accuracy and incomplete segmentation of small vessels, this study introduces an enhanced segmentation model called the attention-gated U-Net. This model incorporates an attention gate to improve segmentation performance. Our methodology integrates adaptive histogram equalization with contrast limitation (CLAHE), median filtering data normalization, and multi-scale morphological transformation to enhance vascular-feature information. Additionally, artifact correction is achieved through adaptive gamma correction. The preprocessed results are then segmented using the attention-gated U-Net (AG-U-Net) model to accurately segment fine vessels.

during upsampling [25]. The second attention stage in TLA involves channel attention, which computes average and maximum pooling operations on the input, combines them through element wise addition, and passes the result through a multilayer perceptron layer with sigmoid activation to dynamically adjust channel weights [23]. TLA further incorporates two consecutive convolutional layers with ReLU activation and concludes with a batch normalization layer, aligning with standard decoder blocks in UNet models. The final stage of TLA involves element wise multiplication between Transformer coefficients and the feature map.

Dataset

The Retinal blood vessel segmentation dataset [26] is an openly available dataset that includes high-resolution retinal fundus images and ground truth labels for retinal blood vessels. These fundus images were captured using state-of-the-art imaging equipment. Each image is accompanied by detailed pixel-level ground truth annotations that precisely identify the blood vessel locations. The dataset provides corresponding pixel-wise annotations in a binary mask format for each image, where blood vessel pixels are denoted as 1 and background pixels as 0.

For faster training, the images were resized to 224 x 224 pixels and stored in JPEG format during pre-processing. The 80 images were divided into three sets: training, validation, and test, with proportions of 70%, 10%, and 20%, respectively. Data augmentation was excluded to minimise computational costs and avoid introducing extra noise to the initial dataset. Our primary objective was to assess the enhanced AG-U-Net's performance and contrast it with the baseline U-Net segmentation model. Re-labelling was also sidestepped to prevent bias, considering the lack of external annotators and blinding [27].

Data Pre-Processing

The hue and saturation levels in individual retinal colour images exhibit significant variation. Each basic image must be converted into an intensity image, which is then normalised to achieve zero mean and unit variance. The normalised intensities are subsequently adjusted to fall within the range of 0 to 255. In human vision, a gamma correction algorithm is applied to process the images. The Contrast Limited Adaptive Histogram Equalization (CLAHE) technique is a commonly employed pre-processing method for retinal vessel segmentation. It enhances the quality of available information, thereby improving the performance of segmentation models. The CLAHE 27 technique is utilized to enhance the contrast of retinal images by controlling noise amplification in neighbouring regions and addressing low-intensity contrast issues. By strengthening the contrast

of retinal images, the CLAHE algorithm preserves overall brightness and colour balance, aiding in the differentiation of retinal vessels from the background. This enhancement can boost the accuracy of segmentation models, particularly in scenarios where the contrast between vessels and the background is minimal.

Model Training, Testing and Results Extraction

During the development of the AG-U-Net and baseline segmentation models, 80 retinal images and masks were utilized for training. The models were configured with a kernel size of 3x3, a learning rate of 10-5, a batch size of 8, and a dropout rate of 50%. The activation function employed was ReLU, and the loss function used was Dice, as studies have indicated that these choices can enhance accuracy and address pixel imbalance, respectively [28,29].

Training and validation were conducted using 64 pairs of retinal images and masks, following the initial dataset split. The objective of the model is to minimize the loss function during training, while the validation set aids in fine-tuning the parameters. At the end of each training epoch, the model is saved if its validation score improves, and the training data is shuffled to prevent bias from the presentation order. Subsequently, the segmentation model is evaluated using a test set consisting of 16 pairs of retinal images and masks. Each iteration aims to produce a model that surpasses its predecessor in performance on the test set.

Evaluation Metrics

In vessel segmentation, various evaluation metrics can be utilised to gauge the performance of the segmentation model, including accuracy, F1 score, Jaccard score, recall score, and precision. These metrics rely on true-positives (TPs), true-negatives (TNs), false-positives (FPs), and false-negatives (FNs) derived from comparing the predicted binary vessel map to the ground truth binary vessel map. This comparison helps evaluate the segmentation results for addressing the challenge of binary segmentation of retinal vessels. TPs represent the pixels correctly identified as vessels in both the ground truth and predicted maps, while TNs denote the pixels correctly identified as non-vessels in both maps. FPs indicate pixels identified as vessels in the predicted map but not in the ground truth map, and FNs refer to pixels classified as vessels in the ground truth map but not identified as vessels in the predicted map.

Accuracy refers to the percentage of correctly identified pixels in the predicted binary map, calculated as $[(TP + TN) / (TP + TN + FP + FN)]$. Precision represents the proportion of predicted vessels correctly identified by the model, calculated

Results

Retinal Blood Vessel Segmentation

Table 1 presents the performance metrics of the AG-U-Net and the baseline U-Net models for retinal blood vessel segmentation. The AG-U-Net demonstrated an accuracy of 0.9488, precision of 0.8186, recall of 0.7419, a dice score of 0.7765, and a Jaccard score of 0.6356. In contrast, the baseline U-Net model achieved an accuracy of 0.8865, precision of 0.7245, recall of 0.6432, a dice score of 0.5738, and a Jaccard score of 0.4189.

| | Accuracy | Precision | Recall | Dice score | Jaccard score (IoU) |
|----------|----------|-----------|--------|------------|---------------------|
| AG-U-Net | 0.9488 | 0.8186 | 0.742 | 0.7765 | 0.6356 |
| U-Net | 0.8865 | 0.7245 | 0.643 | 0.5738 | 0.4189 |

Table 1: The AG-U-Net and performance on retinal blood vessel segmentation in colour fundus photographs.

Discussion

The AG-U-Net model for retinal vascular segmentation has shown promising outcomes in accurately recognizing and segmenting vessels in retinal images. This enhanced approach focuses on specific image regions and extracts relevant features for segmentation through the utilization of global and local attention mechanisms. The AG-U-Net model has demonstrated an accuracy of 0.9488, precision of 0.8186, recall of 0.7419, a dice score of 0.7765, and a Jaccard score of 0.6356 in the Retinal blood vessel segmentation dataset. These results indicate that the AG-U-Net model outperforms the traditional U-Net model in terms of segmentation accuracy.

Additionally, a potential future improvement could involve incorporating a more extensive and diverse set of training data. The current training of the network is based on a limited sample of images, which may not fully capture the variability present in retinal scans. By expanding the training dataset to include a broader range of images, the network's ability to generalise to new data could be enhanced, leading to improved performance on previously unexplored content. Furthermore, enhancing the AG-U-Net model could involve incorporating different modalities beyond conventional grayscale images. For example, integrating channels like fluorescein angiography or optical coherence tomography could provide the network with additional context and information for more accurate vein segmentation. Utilising the TLA network for retinal vascular segmentation has the potential to significantly enhance the precision and effectiveness of retinal vessel analysis in the

future. Moreover, a sudden decrease in Intersection over Union (IoU) indicates a lack of a comprehensive ground truth training dataset. Moving forward, efforts can be directed towards leveraging the existing limited training data to enhance the model's generalisation capabilities. Lastly, improving the transparency of annotators' qualifications and the consensus process for addressing inter-observer variability in annotation can help mitigate uncertainties and doubts surrounding the accuracy of ground truth labels.

Conclusion

The results demonstrate that the enhanced AG-U-Net segmentation model outperforms the standard U-Net baseline model in accurately segmenting retinal blood vessels, achieving a high accuracy score of 0.9488 on the dataset. This improvement is notably superior to the baseline U-Net model's performance. The enhanced method excels in segmentation-specific metrics like dice score and IoU, as well as overall performance, showcasing its effectiveness in accurately delineating retinal blood vessels. The approach exhibits adaptability and performs well even in challenging scenarios. Moving forward, there is potential to develop a more intricate model that can identify vessels with greater precision and enhance the connectivity of vascular structures. Additionally, enhancing the interpretability of deep learning models is essential for understanding the model's detected regions of interest. Visualization techniques can aid in comprehending the model's inner workings, leading to informed decision-making and more efficient problem-solving.

Conflict of Interest: All authors have disclosed no conflict of interests.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability Statement: All data is publicly available and can be retrieved from open-source platforms like Google Dataset Search and Kaggle. The link to the dataset used was also cited in the reference.

References

- Li Z, Jia M, Yang X, Xu M (2021) Blood Vessel Segmentation of Retinal Image Based on Dense-U-Net Network. *Micromachines (Basel)* 12(12): 1478.
- Roychowdhury S, Koozekanani DD, Parhi KK (2015) Blood Vessel Segmentation of Fundus Images by Major Vessel Extraction and Subimage Classification. *IEEE J Biomed Health Inform* 19(3): 1118-1128.
- Galdran A, Anjos A, Dolz J, Chakor H, Lombaert H, et al. (2022) State-of-the-art retinal vessel segmentation with minimalistic models. *Sci Rep* 12(1): 6174.
- Niemeijer M, Xu X, Dumitrescu AV, Gupta P, Ginneken B, et al. (2011) Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs. *IEEE Trans Med Imaging* 30(11): 1941-1950.
- Orlando JI, Breda JB, van Keer K, Blaschko MB, Blanco PJ, et al. (2018) Towards a glaucoma risk index based on simulated hemodynamics from fundus images. *Computer Science* 4: 1-9.
- Welikala RA, Fraz MM, Foster PJ, Whincup PH, Rudnicka AR, et al. (2016) Automated retinal image quality assessment on the UK Biobank dataset for epidemiological studies. *Comput Biol Med* 71: 67-76.
- Kim GY, Kim JY, Lee SH, Kim SM (2022) Robust Detection Model of Vascular Landmarks for Retinal Image Registration: A Two-Stage Convolutional Neural Network. *Biomed Res Int* 2022: 1705338.
- Costa P, Galdran A, Meyer MI, Niemeijer M, Abramoff M, et al. (2018) End-to-End Adversarial Retinal Image Synthesis. *IEEE Trans Med Imaging* 37(3): 781-791.
- Zana F, Klein JC (2001) Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Trans Image Process* 10(7):1010-1019.
- Mendonça AM, Campilho A (2006) Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Trans Med Imaging* 25(9): 1200-1213.
- Frangi AF, Niessen WJ, Vincken KL, Viergever MA (1998) Multiscale vessel enhancement filtering. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*. Springer Berlin Heidelberg pp: 130-137.
- Azzopardi G, Strisciuglio N, Vento M, Petkov N (2015) Trainable COSFIRE filters for vessel delineation with application to retinal images. *Med Image Anal* 19(1): 46-57.
- Zhang J, Dashtbozorg B, Bekkers E, Pluim JPW, Duits R, et al. (2016) Robust Retinal Vessel Segmentation via Locally Adaptive Derivative Frames in Orientation Scores. *IEEE Trans Med Imaging* 35(12): 2631-2644.
- Liskowski P, Krawiec K (2016) Segmenting Retinal Blood Vessels With Deep Neural Networks. *IEEE Trans Med Imaging* 35(11): 2369-2380.
- Suryani E, Susilo M (2019) The hybrid method of SOM artificial neural network and median thresholding for segmentation of blood vessels in the retina image fundus. *Int J Fuzzy Log Intell Syst* 19(4): 323-331.
- 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.
- Zhou L, Yu Q, Xu X, Gu Y, Yang J (2017) Improving dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin-vessel enhancement. *Comput Methods Programs Biomed* 148: 13-25.
- Fu H, Xu Y, Lin S, Kee WDW, Liu J (2016) DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*. Springer International Publishing pp: 132-139.
- Zhou Y, Chen Z, Shen H, Zheng X, Zhao R, et al. (2021) A refined equilibrium generative adversarial network for retinal vessel segmentation. *Neurocomputing* 437: 118-130.
- Xiuqin P, Zhang Q, Zhang H, Li S (2019) A Fundus Retinal Vessels Segmentation Scheme Based on the Improved Deep Learning U-Net Model. *IEEE Access* 7: 122634-122643.
- Radha K, Yepuganti K, Saritha S, Kamireddy C, Bavirisetti DP (2023) Unfolded deep kernel estimation-attention

- UNet-based retinal image segmentation. *Sci Rep* 13(1): 20712.
22. Suri JS, Bhagawati M, Agarwal S, Paul S, Naidu S (2022) UNet Deep Learning Architecture for Segmentation of Vascular and Non-Vascular Images: A Microscopic Look at UNet Components Buffered With Pruning, Explainable Artificial Intelligence, and Bias. *IEEE Access* (11): 595-645.
 23. Azad R, Al-Antary MT, Heidari M, Merhof D (2022) TransNorm: Transformer Provides a Strong Spatial Normalization Mechanism for a Deep Segmentation Model. *IEEE Access* 10: 108205-108215.
 24. Li S, Dong M, Du G, Mu X (2019) Attention Dense-U-Net for Automatic Breast Mass Segmentation in Digital Mammogram. *IEEE Access* 7: 59037-59047.
 25. Zhang S, Fu H, Yan Y, Zhang Y, Wu Q, et al. (2019) Attention Guided Network for Retinal Image Segmentation. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Electrical Engineering and Systems Science* pp: 797-805.
 26. Ibrahim AW (2023) Retina Blood Vessel.
 27. Setiawan AW, Mengko T, Santoso OS, Suksmono A (2013) Color retinal image enhancement using CLAHE. In: *ICT for Smart Society (ICISS), 2013 International Conference on*. Unknown pp: 1-3.
 28. He K, Zhang X, Ren S, Sun J (2015) Delving Deep into Rectifiers: Surpassing Human-Level Performance on Image Net Classification. *Computer Science*: 1-11.
 29. Sudre CH, Li W, Vercauteren T, Ourselin S, Jorge CM (2017) Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support* 240-248.