



# **Gommunication G-Net Light:** A Lightweight Modified Google Net for Retinal Vessel Segmentation

Shahzaib Iqbal <sup>1</sup>, Syed S. Naqvi <sup>1</sup>, Haroon A. Khan <sup>1</sup>, Ahsan Saadat <sup>2</sup>, and Tariq M. Khan <sup>3,\*</sup>

- <sup>1</sup> Department of Electrical and Computer Engineering, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan
- <sup>2</sup> School of Electrical Engineering and Computer Science, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan
- <sup>3</sup> School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia
- \* Correspondence: tariq.khan@unsw.edu.au

**Abstract:** In recent years, convolutional neural network architectures have become increasingly complex to achieve improved performance on well-known benchmark datasets. In this research, we have introduced G-Net light, a lightweight modified GoogleNet with improved filter count per layer to reduce feature overlaps, hence reducing the complexity. Additionally, by limiting the amount of pooling layers in the proposed architecture, we have exploited the skip connections to minimize the spatial information loss. The suggested architecture is analysed using three publicly available datasets for retinal vessel segmentation, namely DRIVE, CHASE and STARE datasets. The proposed G-Net light achieves an average accuracy of 0.9686, 0.9726, 0.9730 and  $F_1$ -score of 0.8202, 0.8048, 0.8178 on DRIVE, CHASE, and STARE datasets, respectively. The proposed G-Net light achieves state-of-the-art performance and outperforms other lightweight vessel segmentation architectures with fewer trainable number of parameters.

Keywords: deep learning; convolutional neural networks; medical image segmentation



Citation: Iqbal, S.; Naqvi, S.S.; Khan, H.A.; Saadat, A.; Khan, T.M. G-Net Light: A Lightweight Modified Google Net for Retinal Vessel Segmentation. *Photonics* **2022**, *9*, 923. https://doi.org/10.3390/ photonics9120923

Received: 29 August 2022 Accepted: 25 November 2022 Published: 30 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

## 1. Introduction

Diabetic retinopathy (DR) has gained a great deal of attention recently due to its connection with long-standing diabetes, which is one of the most common causes of avoidable blindness in the world [1,2]. Additionally, diabetic retinopathy is one of the major contributors of vision loss, especially in those of working age [3,4]. Lesions are the first signs of diabetic retinopathy. They include exudates, microaneurysms, haemorrhages, vessel abnormalities and leakages [5–7]. The number and type of lesions that form on the surface of the retina affect the severity and diagnosis of the disease. Thus, the effectiveness of an automated system for extensive screening may depend on the precision of segmenting blood vessels, optical cup/disc and retinal lesions [8]. Along these lines, it has long been thought that detecting retinal blood vessels is the most difficult problem, and it is frequently thought that it is the most crucial part of an automated computer-aided diagnostic (CAD) system [1,9]. This is because the vessels in the retina are hard to see because of their tortuous shape, density, diameter and branching pattern [10]. Even more challenging to identify are the centerline reflex and the many components that make up the retina, including the macula, optic cup/disc, exudates and so on, all of which may have lesions or other flaws. Finally, the settings used for camera calibration and the acquisition method can also bring unpredictability into the imaging process.

For the purpose of blood vessels segmentation, when a machine learning or deep learning architecture is used, training is usually conducted using a dataset of manually labelled segmented images [11]. In order to diagnose serious disorders including retinal vascular occlusions [12], glaucoma [13], AMD [14], DR [15] and chronic systematic hypoxemia (CSH) [16], these techniques have been used to detect retinal vessels. Kadri et al. [17] introduced a multi-scale filter (MSMF) utilising the slime mould optimization technique. Furthermore, deep learning-based approaches have attained cutting-edge accuracy in applications including vessel detection and optic cup/disc detection [18]. Therefore, it is believed that the predominant technique for creating retinal diagnostics systems is now supervised machine learning models [19,20]. Despite considerable success of supervised ML models [21,22], it is still challenging to find blood vessels when there are noticeable differences and abnormalities. It becomes significantly more challenging when the vessels' diameter is small. Moreover, training these architectures is time-consuming, despite the fact that the results of supervised segmentation obtained by these methods are superior to those of unsupervised segmentation. The lack of comprehensively (labelled) data for a variety of ailments and imaging modalities makes this more challenging.

According to [23,24], the usage of intricate CNN architecture based models does not produce the optimal results for the majority of segmentation algorithms. Keep in mind that the quantity of hidden layers and the number of filters used in each layer have a significant impact on the number of trainable parameters. In these circumstances, shallow networks are frequently suggested as a deep network substitute [20]. In comparison to their deep counterparts, these shallow networks employ fewer filters per layer. Our network's layout is intended to utilize the most filters possible in each layer while minimizing the complexity of the system as a whole. If an image has less feature variation, performance does not rise with more filters in a convolution layer, but complexity does [25,26]. By recommending smallscale networks with fewer layers, convolution networks' complexity has been lowered in the literature [27–32]. Furthermore, the significance in terms of the performance and complexity is not addressed in [27]. Here, the characteristic complexity is used to determine the number of filters.

To the best of our knowledge, GoogleNet based encoder–decoder architecture for image segmentation is not proposed so far, hence one of the major contributions of the proposed work is to design a decoder of GoogleNet. Inspired from the GoogleNet [33], this study introduces G-Net lite, a simple yet effective small scale neural network architecture for retinal blood vessels' segmentation. This is because G-Net light only has a small number of parameters, which means that it requires relatively lesser memory and GPU resources than alternatives with significantly higher parameters. In addition, the encoder employs only two max-pooling layers to reduce the spatial information loss. Experiments are conducted on three different datasets of retinal blood vessels segmentation to demonstrate the efficacy of the proposed architecture for medical image segmentation.

#### 2. Related Work

Recent research has been presented where the U-Net structure is extended with changing module design and network building, demonstrating its potential on numerous visual analysis tasks. V-Net [34] extends U-Net to higher dimension pixels while retaining the vanilla internal structures. W-Net [35] adapts U-Net to address the problem of unsupervised segmentation by concatenating two U-Nets using an autoencoder style model. In contrast to U-Net, M-Net [36] appends different scales of input characteristics to different levels, allowing a sequence of downsampling and upsampling layers to capture multi-level visual information. U-Net++ [37] has recently adopted nested and dense skip connections to more efficiently depict fine-grained object information. Furthermore, attention U-Net [38] employs extra branches to apply the attention mechanism adaptively on the fusion of skipped and decoded data. However, these suggestions may include extra building pieces, resulting in a bigger number of network parameters and, as a result, more GPU RAM. It has been established that using recurrent convolution to repeatedly modify the features extracted at different periods is feasible and successful for many computer vision problems [39–42]. Guo et al. [39] advocated reusing residual blocks in ResNet to completely utilise available parameters and greatly reduce model size. Such a mechanism is also beneficial to the evolution of U-Net. As a result, Wang et al. [42] created R-U-Net, which recurrently connects multiple paired encoders and decoders of U-Net to improve its

discrimination strength for semantic segmentation; however, as a trade-off, extra learnable blocks are included.

#### 3. G-Net Light

This section presents and explains the the proposed network architecture. In Figure 1, overall architecture of G-Net light is presented. The proposed network starts with an input image layer, then a convolutional layer, and finally the essential final layers that create the pixel-wise segmentation map. We have performed nonlinear activations (ReLU) on the segmentation map. The feature maps are then fed into the max-pooling layer. The inception block is used after the max-pooling layer, followed by another max-pooling layer. There is an inception block which connects the encoder and decoder blocks. At the decoder side, the up-sampling layer (max-unpooling) is used followed by the same inception block, another up-sampling layer and another inception block. Once the spatial information is restored using up-sampling layers, a convolutional layer (CL) followed by nonlinear activations (ReLU), and the batch normalization layer (BN) is applied. After a soft-max layer, the final classification layer is a dice pixel classification layer. Note that the proposed architecture has four inception blocks, where the first block is used after the first downsampling. There is an intermediary inception block that connects the encoder and decoder blocks. There are two inception blocks at decoder followed by the convolutional layer, which is supplied with the necessary final layers required for constructing the pixel-wise segmentation map. Using the convolution layers in between the filter banks and input feature maps, each encoder block creates its own collection of features. We have performed nonlinear activations (ReLU) on these features. Depending on whether the block is upsampling or down-sampling, the produced feature maps are subsequently supplied to the max-pooling or unpooling layers. All max-pooling and unpooling layers are  $2 \times 2$ , non-overlapping, with a stride size of 2.



**Figure 1.** Block diagram of the proposed network.

It is worth noting that the proposed network design responds to multiple motivations. To begin, we wanted to use as few pooling layers as possible in the proposed architecture. This is due to the fact that pooling frequently reduces the size of the feature maps and can also result in a spatial information loss. Second, we have used a limited number of convolutional layers. Finally, within each layer, the total number of convolutional filters are minimized. Skip connections have been used between the encoder and the associated

decoder blocks to preserve structural information. Figure 1 depicts these as dotted lines with arrowheads. Another motivating force behind the choice to adopt skip connections as an alternative to dense skip paths is the assumption that feature retention within each convolutional layers may assist with reducing the semantic gap of the encoder side and decoder side while keeping computational overhead under control. In order to preserve fine-grained structures, which are frequently important in medical image segmentation, the number of pooling layers is reduced in the proposed network.

#### The Inception Block

The key idea of the inception block is to apply the dimension reductions wisely. These reductions are computed using the  $1 \times 1$  filter size for the convolution operations prior to the 3  $\times$  3 and 5  $\times$  5 filter size for the convolutions operations. They are dualpurpose because, in addition to being utilised as reductions, they also utilise rectified linear activation. Figure 2 depicts the ultimate design of the inception block. An Inception block generally is an architecture made up of the above-mentioned modules that are vertically stacked with intermittent max-pooling layers with stride 2 that result in the reduction of the grid's resolution. It appeared preferable to start using the inception blocks only at higher layers and leave the lower layers in typical convolutional form for technical reasons during the training. One of this architecture's key benefits is that it permits significant increases in the number of units at each step without increasing the complexity in terms of computations. The widespread use of reduction of the dimensions enables hiding the high volume of input filters from the preceding stage to the succeeding layer. This is achieved by initially lowering their dimension prior to convolving over them with a large patch size. This method also adheres to the idea that visual data should be processed at various scales before being aggregated, allowing the subsequent stage to simultaneously extract features from different scales. Because the processing resources are being used more efficiently, it is possible to increase both the total number of stages and the width of each stage without encountering computational challenges. Developing significantly less effective but computationally less expensive variations of the inception block is another way to use it. It can be seen that all of the available knobs and levers enable a controlled balancing of computational resources. This can lead to architectures that are twice as fast or three times as fast as similarly performing networks without the inception blocks, though this requires a careful manual design.



Figure 2. The inception block.

#### 4. Experimental Setup

#### 4.1. Datasets

For the segmentation of retinal vessels, we tested our proposed network using three public image data sets: DRIVE [43], CHASE [44] and STARE [45]. DRIVE [43] is made up of 20 colour images for testing and 20 colour images for training, both of which are saved  $584 \times 565$  image size in JPEG format and cover a wide range of age of DR patients. A field of view (FOV) binary mask is available for all images. Both the test and training images contain manually segmented ground truth vessels' labels.

The CHASE [44] dataset includes 28 colour images acquired with a 30° FOV centred at the optical disc and an image resolution of 999  $\times$  960 pixels. Two distinct manually segmentation ground truth maps are available. For the experiments, the first expert's segmentation map is used.

The STARE [45] dataset consists of 20 colour retinal fundus images with a size of  $700 \times 605$  pixels per image that were taken at a 35° FOV. Each of these images has two separate manual segmentation maps available. Here, we have used the initial ophthalmologist segmentation as the benchmark.

#### 4.2. Implementation and Training

All of our studies have been run using a GeForce GTX2080TI GPU and an Intel(R) Xeon(R) W-2133 3.6 GHz CPU with 96GB RAM. With a fixed learning rate, stochastic gradient descent was used in our RC-Net implementation. A weighted cross-entropy loss is employed as an objective function for training in all of our experiments. This decision was made after it was discovered that, in each retinal image's vessel segmentation, the *non*-*vessel* pixels outweighed the *vessel* pixels by a significant margin. Various techniques can be employed to assign the loss weights. Here, we use median frequency balancing to determine class association weights [46]. Note that STARE and CHASE datasets do not have a specified test set available. In the literature, a *"leave-one-out"* strategy is frequently utilised for STARE [47]. With 10 images for training and 10 for testing, we have employed *"leave-one-out"* data split in this paper. We have also employed data augmentation to generate sufficient images for training since the retinal vascular segmentation image datasets used are relatively small. Contrast enhancement and rotations were utilised for the data augmentation. Each image is rotated by 1° for the rotations at the training stage. The image brightness was randomly increased and decreased to enhance the contrast.

#### 4.3. Evaluation Criteria

Remember that pixel markings on blood vessels segmentation are binary, indicating whether a pixel is a vessel or the background. Publicly accessible datasets include *ground truth* that is manually annotated by experienced clinicians. As a result, each pixel is categorized as vessel pixel, if the area of interest is present in a image such as blood retinal vessels. There can be four possible outcomes for each output image: pixels that are correctly categorized as areas of interest ( $T_P$ : true positive), pixels that are correctly categorized as non-interest ( $T_N$ : true negative), pixels of non-interest that were incorrectly categorized as such (( $F_N$ ): false negative). Four commonly used performance parameters Accuracy, Sensitivity, Specificity, and  $F_1$ -score are frequently used in the literature to compare approaches using these components:

$$A_{cc} = \frac{T_P + T_N}{F_P + F_N + T_P + T_N} \tag{1}$$

The term accuracy  $(A_{cc})$  in Equation (1) denotes the proportion of successfully segmented pixels to all of the pixels in the expertly annotated (labelled) mask. The *Sp* and *Se* indicate the model's specificity and sensitivity, which demonstrate how the no-vessel and vessel pixels are correctly distinguished and given in Equations (2) and (3), respectively:

$$S_p = \frac{I_N}{T_N + F_P} \tag{2}$$

$$S_n = \frac{T_P}{T_P + F_N} \tag{3}$$

The  $F_1$ -score, which is the harmonic mean of  $S_n$  and precision, is another technique to assess the model's performance and can be calculated using Equation (4):

$$F_1 - score = \frac{2 \times T_P}{(2 \times T_P) + F_P + F_N} \tag{4}$$

#### 5. Analysis of the Results and Comparisons

The qualitative and quantitative analysis of the proposed architecture with a number of commonly used alternatives methods in retinal image segmentation is included in this section. Table 1 shows the summary of quantitative performance of the proposed G-Net light w.r.t the ground-truths marked by different observers on DRIVE, CHASE and STARE dataset. The average performance on each data set is also shown in Table 1.

Dataset	Ground-Truth –	Performance						
		S <sub>n</sub>	S <sub>p</sub>	A <sub>cc</sub>	F <sub>1</sub> -Score			
DRIVE	1st Manual 2nd Manual	0.8192 0.8714	0.9829 0.9734	0.9686 0.9807	0.8202 0.8724			
	Average	0.8453	0.9782	0.9747	0.8463			
CHASE	1st Observer 2nd Observer	0.8210 0.8932	0.9838 0.9823	0.9726 0.9847	0.8048 0.8770			
	Average	0.8571	0.9831	0.9787	0.8409			
STARE	Dr. Adam Hoover Dr. Valentina Kouznetsova	0.8170 0.8892	0.9853 0.9838	0.9730 0.9851	0.8178 0.8900			
	Average	0.8531	0.9846	0.9791	0.8539			

Table 1. Quantitative performance of the proposed G-Net light on DRIVE, CHASE and STARE dataset.

The qualitative segmentation findings for the retinal vessels on the DRIVE dataset are analysed and discussed first. In Figure 3, the analysis of the segmented output is illustrated. In Figure 3a, noisy test images 3, 4 and 19 from the DRIVE dataset are presented. Corresponding ground truth images of the 1st observer are given in Figure 3b. Figure 3c,d presents the output of the networks SegNet [48] and U-Net [49], respectively. The segmentation output of the proposed architecture is given in Figure 3e. The segmentation maps' black and green colours represent accurately predicted pixels, whereas the blue and red colours represent false negatives and false positives, respectively. It is apparent that the suggested G-Net Light outperforms the U-Net [49] and SegNet [48] in terms of visual performance. The segmentation maps' black and green colours represent accurately predicted pixels, whereas the blue and red colours represent false negatives and false positives, respectively. It can be clearly observed that the visual performance of the proposed G-Net Light is better than the SegNet [48] and U-Net [49].

The vessel segmented maps of the proposed architecture on CHASE dataset are given in Figure 4. The segmentation maps' black and green colours represent accurately predicted pixels, whereas the blue and red colours represent false negatives and false positives, respectively. In the 1st row, noisy images of CHASE dataset are illustrated, and the corresponding ground truth images marked by 1st observer are shown in the 2nd row of the Figure 4. The final vessels' segmented vessels map images of the proposed architecture are shown in the 3rd row of Figure 4.



**Figure 3.** Analysis of segmented output. The segmentation maps' black and green colours represent accurately predicted pixels, whereas the blue and red colours represent false negatives and false positives, respectively: (**a**) noisy test images 3, 4 and 19 of the DRIVE dataset; (**b**) corresponding ground truth images from the 1st manual; (**c**) the output of SegNet [48]; (**d**) the output of U-Net [49]; (**e**) the output of the proposed G-Net light architecture.



**Figure 4.** Analysis of segmented output. The segmentation maps' black and green colours represent accurately predicted pixels, whereas the blue and red colours represent false negatives and false positives, respectively: in row one, noisy test images of CHASE dataset; in row two, corresponding ground truth images marked by 1st observer. In row three, the output of the proposed network is presented.

Tables 2 and 3 compare the G-Net light network performance to some state-of-the-art supervised approaches. The proposed architecture obtains an average sensitivity of 81.92% for the DRIVE database and 82.10% for the CHASE database. In terms of the sensitivity parameter, the proposed G-Net light architecture outperforms all other techniques on the DRIVE dataset and is the 3rd highest on CHASE dataset. The average accuracies of

the proposed G-Net light are 96.86% and 97.26%, the highest on the DRIVE and CHASE datasets, respectively. The proposed architecture achieves an average specificity of 98.29% on DRIVE and 98.38% on CHASE, the 3rd and 2nd highest, respectively. Finally, the proposed network achieves 82.02% of  $F_1$ -score, the highest on DRIVE dataset and the 3rd highest value of 80.48% on CHASE dataset.

**Table 2.** Comparison results on DRIVE dataset. Red is the best, green is the 2nd best, and blue is the3rd best.

Method	Year	S <sub>n</sub>	Sp	A <sub>cc</sub>	F <sub>1</sub> -Score
SegNet [46]	2017	0.7949	0.9738	0.9579	0.8180
MŠ-NFN [50]	2018	0.7844	0.9819	0.9567	N.A
FCN [51]	2018	0.8039	0.9804	0.9576	N.A
BTS-DSN [52]	2019	0.7891	0.9804	N.A	N.A
Three-stage CNN [53]	2019	0.7631	0.9820	0.9538	N.A
DE U-Net [54]	2019	0.7986	0.9736	0.9511	N.A
EL Approach [55]	2019	0.7880	0.9819	0.9569	N.A
GGM [56]	2019	0.7820	0.9860	0.9600	N.A
VessNet [57]	2019	0.8022	0.9810	0.9655	N.A
Vessel-Net [58]	2019	0.8038	0.9802	0.9578	N.A
CcNet [59]	2020	0.7625	0.9809	0.9528	N.A
AWS FCM [60]	2022	0.7020	0.9844	0.9605	0.7531
Proposed Method	2022	0.8192	0.9829	0.9686	0.8202

**Table 3.** Comparison results on the CHASE dataset. Red is the best, green is the 2nd best, and blue is the 3rd best.

Method	Year	S <sub>n</sub>	$S_p$	$A_{cc}$	F <sub>1</sub> -Score
U-Net [61]	2016	0.7764	0.9865	0.9643	N.A
R2u-net [62]	2018	0.7756	0.9820	0.9634	N.A
Laddernet [63]	2018	0.7978	0.9818	0.9656	0.8031
Ce-net [64]	2019	0.8008	0.9723	0.9633	N.A
Iternet [65]	2020	0.7969	0.9820	0.9702	0.8073
SA-Unet [66]	2021	0.8151	0.9809	0.9708	0.7736
AACA-MLA-D-Unet [67]	2021	0.8302	0.9810	0.9673	0.8248
MC-UNet [68]	2022	0.8366	0.9829	0.9714	0.7741
Proposed Method	2022	0.8210	0.9838	0.9726	0.8048

In Figure 5, the analysis of the segmented output is illustrated. In Figure 5a, noisy test images from STARE dataset are presented. Corresponding ground truth images marked by Adam Hoover are given in Figure 5b. Figure 5c,d presents the output of the networks SegNet [48] and U-Net [49], respectively. The vessels' segmentation maps of the proposed architecture are shown in Figure 5e. The segmentation maps' black and green colours represent accurately predicted pixels, whereas the blue and red colours represent false negatives and false positives, respectively. It can be clearly observed that the visual performance of the proposed G-Net Light is better than the SegNet [48] and U-Net [49]. Table 4 compares the proposed network performance to state-of-the-art supervised approaches. The proposed architecture obtains an average sensitivity of 81.70%, which is 2nd highest among the all methods. The average accuracy of the proposed G-Net light architecture is 97.30%, which is 3rd highest. Finally, the proposed network achieves 81.78% of *F*<sub>1</sub>-*score*, the 2nd highest among the all methods on the STARE dataset.



**Figure 5.** Analysis of segmented output. The segmentation maps' black and green colours represent accurately predicted pixels, whereas the blue and red colours represent false negatives and false positives, respectively: (**a**) noisy test images of STARE dataset; (**b**) corresponding ground truth images marked by Adam Hoover; (**c**) the output of SegNet [48]; (**d**) the output of U-Net [49]; (**e**) the output of the proposed G-Net light architecture.

Method	Year	S <sub>n</sub>	$S_p$	$A_{cc}$	F <sub>1</sub> -Score
U-Net [61]	2016	0.7764	0.9865	0.9643	N.A
R2u-net [62]	2018	0.7756	0.9820	0.9634	N.A
Laddernet [63]	2018	0.7822	0.9804	0.9613	0.7994
BTS-DSN [52]	2019	0.8212	0.9843	N.A	N.A
Dual Encoding U-Net [54]	2019	0.7914	0.9722	0.9538	N.A
GGM [56]	2019	0.7960	0.9830	0.9610	N.A
Ce-net [64]	2019	0.7909	0.9721	0.9732	N.A
CcNet [59]	2020	0.7709	0.9848	0.9633	N.A
Iternet [65]	2020	0.7969	0.9823	0.9760	0.8073
SA-Unet [66]	2021	0.7120	0.9930	0.9521	0.7736
AACA-MLA-D-Unet [67]	2021	0.7914	0.9870	0.9665	0.8276
MC-UNet [68]	2022	0.7360	0.9947	0.9572	0.7865
Proposed Method	2022	0.8170	0.9853	0.9730	0.8178

**Table 4.** Comparison results on the STARE dataset. Red is the best, green is the 2nd best, and blue is the 3rd best.

Comparisons of the segmentation of retinal vessels using the proposed G-Net light and current lightweight networks in terms of learnable parameters and quantitative performance are also carried out in Table 5. The accuracy and  $F_1$ -score results are compared from G-Net light to current lightweight networks on the DRIVE, CHASE and STARE datasets. Table 5 shows that G-Net light outperforms the state-of-the-art alternatives in terms of accuracy and  $F_1$ -score with a minimal learnable parameters.

	Parameters (M)	Size in (MB)	DRIVE		CHASE		STARE	
Method			A <sub>cc</sub>	F <sub>1</sub> -Score	A <sub>cc</sub>	F <sub>1</sub> -Score	A <sub>cc</sub>	F <sub>1</sub> -Score
Image BTS-DSN [52]	7.80	N.A	0.9551	0.8201	0.9627	0.7983	0.9660	N.A
MobileNet-V3-Small [27]	2.50	11.00	0.9371	0.6575	0.9571	0.6837	N.A	N.A
ERFNet [69]	2.06	8.00	0.9598	0.7652	0.9716	0.7872	N.A	N.A
Sine-Net [70]	0.69	N.A	0.9685	N.A	0.9676	N.A	0.9711	N.A
M2U-Net [71]	0.55	2.20	0.9630	0.8091	0.9703	0.8006	N.A	0.7814
Proposed G-Net Light	0.39	1.52	0.9686	0.8202	0.9726	0.8048	0.9730	0.8178

**Table 5.** Comparison with recent light-weight networks in terms of accuracy and  $F_1$ -*score* on DRIVE, CHASE and STARE datasets. Best results are highlighted in bold font. The measures that are not available are represented by N.A.

In Figure 6, the analysis of the quantitative results is illustrated. In Figure 6a, comparison of the quantitative results of G-Net light with other methods on DRIVE dataset are given. Figure 6b,c, shows the comparison of the quantitative results of G-Net light with other methods on CHASE and STARE datasets, respectively. It can be observed from Figure 6 that the performance of the proposed G-Net light is clearly comparable with the other state-of-the-art methods.



Figure 6. Quality measurements for datasets: (a) DRIVE; (b) CHASE; (c) STARE.

## 6. Discussion

A sizeable portion of the global population is affected by a variety of retinal illnesses that can compromise one's vision. This significant worry has arisen in part as a result of the high cost of the necessary equipment that is required for the diagnosis of ophthalmological diseases, and in part as a result of the scarcity of ophthalmological specialists who are readily available. It is essential to make a prompt diagnosis of these retinal illnesses in order to avert vision loss and blindness. In this regard, accessible computer-aided diagnostic techniques have the potential to play a pivotal role. The majority of the deep learning models that have been developed for the diagnosis of retinal disorders function effectively, despite the fact that they computationally expensive. This constitutes a major obstacle in the way of the deployment of such models on portable edge devices. Therefore, the proposed lightweight model for the segmentation of retinal vessels can play a vital role in the development of computationally less expensive diagnostic systems. The proposed model uses significantly less trainable parameters without sacrificing performance.

## 7. Conclusions

In this research paper, we have introduced and analyzed G-Net light, a lightweight modified GoogleNet with improved filter count per layer to reduce feature overlaps and complexity. Additionally, by reducing the amount of pooling layers in the proposed architecture, we have exploited the skip connections to minimize the spatial information loss. Our investigations are examined on publicly available DRIVE, CHASE and STARE datasets. In the experiments, the proposed G-Net light achieves state-of-the-art performance

and outperforms other lightweight vessel segmentation architectures in terms of accuracy and  $F_1$ -score with fewer trainable number of parameters.

Author Contributions: Conceptualization, S.I. and T.M.K.; methodology, S.I. and S.S.N.; software, S.I.; validation, S.I., T.M.K. and S.S.N.; writing—original draft preparation, S.S.N.; writing—review and editing, T.M.K., A.S., H.A.K.; supervision, T.M.K. and S.S.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- Khawaja, A.; Khan, T.M.; Naveed, K.; Naqvi, S.S.; Rehman, N.U.; Junaid Nawaz, S. An Improved Retinal Vessel Segmentation Framework Using Frangi Filter Coupled With the Probabilistic Patch Based Denoiser. *IEEE Access* 2019, 7, 164344–164361. [CrossRef]
- Manan, M.A.; Khan, T.M.; Saadat, A.; Arsalan, M.; Naqvi, S.S. A Residual Encoder-Decoder Network for Segmentation of Retinal Image-Based Exudates in Diabetic Retinopathy Screening. arXiv 2022, arXiv:2201.05963.
- Khan, T.M.; Robles-Kelly, A.; Naqvi, S.S.; Muhammad, A. Residual Multiscale Full Convolutional Network (RM-FCN) for High Resolution Semantic Segmentation of Retinal Vasculature. In Proceedings of the Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshops, S+ SSPR 2020, Padua, Italy, 21–22 January 2021; Springer Nature: Berlin/Heidelberg, Germany, 2021; p. 324.
- Khan, T.M.; Khan, M.A.; Rehman, N.U.; Naveed, K.; Afridi, I.U.; Naqvi, S.S.; Raazak, I. Width-wise vessel bifurcation for improved retinal vessel segmentation. *Biomed. Signal Process. Control* 2022, 71, 103169. [CrossRef]
- Soomro, T.A.; Khan, T.M.; Khan, M.A.U.; Gao, J.; Paul, M.; Zheng, L. Impact of ICA-Based Image Enhancement Technique on Retinal Blood Vessels Segmentation. *IEEE Access* 2018, *6*, 3524–3538. [CrossRef]
- Abdullah, F.; Imtiaz, R.; Madni, H.A.; Khan, H.A.; Khan, T.M.; Khan, M.A.U.; Naqvi, S.S. A Review on Glaucoma Disease Detection Using Computerized Techniques. *IEEE Access* 2021, 9, 37311–37333. [CrossRef]
- Iqbal, S.; Khan, T.M.; Naveed, K.; Naqvi, S.S.; Nawaz, S.J. Recent trends and advances in fundus image analysis: A review. Comput. Biol. Med. 2022, 151, 106277. [CrossRef] [PubMed]
- 8. Naveed, K.; Abdullah, F.; Madni, H.A.; Khan, M.A.; Khan, T.M.; Naqvi, S.S. Towards Automated Eye Diagnosis: An Improved Retinal Vessel Segmentation Framework Using Ensemble Block Matching 3D Filter. *Diagnostics* **2021**, *11*, 114. [CrossRef]
- Imtiaz, R.; Khan, T.M.; Naqvi, S.S.; Arsalan, M.; Nawaz, S.J. Screening of Glaucoma disease from retinal vessel images using semantic segmentation. *Comput. Electr. Eng.* 2021, 91, 107036. [CrossRef]
- Khan, M.A.; Khan, T.M.; Aziz, K.I.; Ahmad, S.S.; Mir, N.; Elbakush, E. The use of fourier phase symmetry for thin vessel detection in retinal fundus images. In Proceedings of the 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, United Arab Emirates, 10–12 December 2019; pp. 1–6.
- Khan, T.M.; Robles-Kelly, A.; Naqvi, S.S. A Semantically Flexible Feature Fusion Network for Retinal Vessel Segmentation. In Proceedings of the Neural Information Processing; Yang, H., Pasupa, K., Leung, A.C.S., Kwok, J.T., Chan, J.H., King, I., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 159–167.
- Muraoka, Y.; Tsujikawa, A.; Murakami, T.; Ogino, K.; Kumagai, K.; Miyamoto, K.; Uji, A.; Yoshimura, N. Morphologic and Functional Changes in Retinal Vessels Associated with Branch Retinal Vein Occlusion. *Ophthalmology* 2013, 120, 91–99. [CrossRef] [PubMed]
- Thakoor, K.A.; Li, X.; Tsamis, E.; Sajda, P.; Hood, D.C. Enhancing the Accuracy of Glaucoma Detection from OCT Probability Maps using Convolutional Neural Networks. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 2036–2040.
- 14. Cicinelli, M.V.; Rabiolo, A.; Sacconi, R.; Carnevali, A.; Querques, L.; Bandello, F.; Querques, G. Optical coherence tomography angiography in dry age-related macular degeneration. *Surv. Ophthalmol.* **2018**, *63*, 236–244. [CrossRef]
- Zeng, X.; Chen, H.; Luo, Y.; Ye, W. Automated Diabetic Retinopathy Detection Based on Binocular Siamese-Like Convolutional Neural Network. *IEEE Access* 2019, 7, 30744–30753. [CrossRef]
- Traustason, S.; Jensen, A.S.; Arvidsson, H.S.; Munch, I.C.; Søndergaard, L.; Larsen, M. Retinal Oxygen Saturation in Patients with Systemic Hypoxemia. *Investig. Ophthalmol. Vis. Sci.* 2011, 52, 5064. [CrossRef] [PubMed]
- Kadry, S.; Rajinikanth, V.; Damaševičius, R.; Taniar, D. Retinal vessel segmentation with slime-Mould-optimization based multiscale-matched-filter. In Proceedings of the 2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 25–27 March 2021; pp. 1–5.

- Jiang, Y.; Zhang, H.; Tan, N.; Chen, L. Automatic Retinal Blood Vessel Segmentation Based on Fully Convolutional Neural Networks. *Symmetry* 2019, 11, 1112. [CrossRef]
- Khan, T.M.; Naqvi, S.S.; Arsalan, M.; Khan, M.A.; Khan, H.A.; Haider, A. Exploiting residual edge information in deep fully convolutional neural networks for retinal vessel segmentation. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–8.
- Khan, T.M.; Abdullah, F.; Naqvi, S.S.; Arsalan, M.; Khan, M.A. Shallow Vessel Segmentation Network for Automatic Retinal Vessel Segmentation. In Proceedings of the 2020 International Joint Conference on Neural Networks, Glasgow, UK, 19–24 July 2020; pp. 1–7. [CrossRef]
- 21. Khan, T.M.; Robles-Kelly, A. Machine learning: Quantum vs classical. IEEE Access 2020, 8, 219275–219294. [CrossRef]
- Khan, T.M.; Robles-Kelly, A. A derivative-free method for quantum perceptron training in multi-layered neural networks. In Proceedings of the International Conference on Neural Information Processing, Bangkok, Thailand, 18–22 November 2020; Springer: Cham, Switzerland, 2020; pp. 241–250.
- Chen, W.; Liu, Y.; Kira, Z.; Wang, Y.; Huang, J. A Closer Look at Few-shot Classification. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
- Kevin, M.; Serge, B.; Ser-Nam, L. A Metric Learning Reality Check. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020.
- Khan, T.M.; Naqvi, S.S.; Robles-Kelly, A.; Meijering, E. Neural Network Compression by Joint Sparsity Promotion and Redundancy Reduction. arXiv 2022, arXiv:2210.07451.
- Arsalan, M.; Khan, T.M.; Naqvi, S.S.; Nawaz, M.; Razzak, I. Prompt Deep Light-weight Vessel Segmentation Network (PLVS-Net). IEEE/ACM Trans. Comput. Biol. Bioinform. 2022. [CrossRef]
- Howard, A.; Sandler, M.; Chu, G.; Chen, L.C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for MobileNetV3. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019.
- Ma, N.; Zhang, X.; Zheng, H.T.; Sun, J. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.
- Khan, T.M.; Robles-Kelly, A.; Naqvi, S.S. T-Net: A resource-constrained tiny convolutional neural network for medical image segmentation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 4–8 January 2022; pp. 644–653.
- Khan, T.M.; Naqvi, S.S.; Meijering, E. Leveraging Image Complexity in Macro-Level Neural Network Design for Medical Image Segmentation. arXiv 2021, arXiv:2112.11065.
- Khan, T.M.; Robles-Kelly, A.; Naqvi, S.S. RC-Net: A Convolutional Neural Network for Retinal Vessel Segmentation. In Proceedings of the 2021 Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, Australia, 29 November–1 December 2021; pp. 1–7.
- 32. Khan, T.M.; Arsalan, M.; Robles-Kelly, A.; Meijering, E. MKIS-Net: A Light-Weight Multi-Kernel Network for Medical Image Segmentation. *arXiv* 2022, arXiv:2210.08168.
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
- Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
- 35. Xia, X.; Kulis, B. W-net: A deep model for fully unsupervised image segmentation. *arXiv* **2017**, arXiv:1711.08506.
- Mehta, R.; Sivaswamy, J. M-net: A convolutional neural network for deep brain structure segmentation. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 437–440.
- Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imaging* 2019, 39, 1856–1867. [CrossRef]
- 38. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention u-net: Learning where to look for the pancreas. *arXiv* **2018**, arXiv:1804.03999.
- Guo, Q.; Yu, Z.; Wu, Y.; Liang, D.; Qin, H.; Yan, J. Dynamic recursive neural network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 5147–5156.
- Han, W.; Chang, S.; Liu, D.; Yu, M.; Witbrock, M.; Huang, T.S. Image super-resolution via dual-state recurrent networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 1654–1663.
- Alom, M.Z.; Yakopcic, C.; Taha, T.M.; Asari, V.K. Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net). In Proceedings of the NAECON 2018-IEEE National Aerospace and Electronics Conference, Dayton, OH, USA, 23–26 July 2018; pp. 228–233.
- 42. Wang, W.; Yu, K.; Hugonot, J.; Fua, P.; Salzmann, M. Recurrent U-Net for resource-constrained segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 2142–2151.

- Staal, J.; Abramoff, M.D.; Niemeijer, M.; Viergever, M.A.; van Ginneken, B. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging* 2004, 23, 501–509. [CrossRef]
- The Child Heart and Health Study in England (CHASE). Available online: https://blogs.kingston.ac.uk/retinal/chasedb1/ (accessed on 14 June 2020).
- 45. Hoover, A.D.; Kouznetsova, V.; Goldbaum, M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. Med. Imaging* **2000**, *19*, 203–210. [CrossRef] [PubMed]
- Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 2481–2495. [CrossRef]
- 47. Soares, J.V.B.; Leandro, J.J.G.; Cesar, R.M.; Jelinek, H.F.; Cree, M.J. Retinal vessel segmentation using the 2D Gabor wavelet and supervised classification. *IEEE Trans. Med. Imaging* **2006**, *25*, 1214–1222. [CrossRef] [PubMed]
- Tariq, M.K.; Musaed, A.; Khursheed, A.; Muhammad, A.; Syed, S.N.; Junaid, N.S. Residual Connection Based Encoder Decoder Network (RCED-Net) For Retinal Vessel Segmentation. *IEEE Access* 2020, *8*, 131257–131272. [CrossRef]
- Song, G. DPN: Detail-Preserving Network with High Resolution Representation for Efficient Segmentation of Retinal Vessels. J. Ambient. Intell. Humaniz. Comput. 2020, 1–14. [CrossRef]
- Wu, Y.; Xia, Y.; Song, Y.; Zhang, Y.; Cai, W. Multiscale Network Followed Network Model for Retinal Vessel Segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention, Granada, Spain, 16–20 September 2018; pp. 119–126.
- 51. Oliveira, A.; Pereira, S.; Silva, C.A. Retinal vessel segmentation based on fully convolutional neural networks. *Expert Syst. Appl.* **2018**, *112*, 229–242. [CrossRef]
- Song, G.; Kai, W.; Hong, K.; Yujun, Z.; Yingqi, G.; Tao, L. BTS-DSN: Deeply supervised neural network with short connections for retinal vessel segmentation. *Int. J. Med. Inform.* 2019, 126, 105–113.
- Yan, Z.; Yang, X.; Cheng, K. A Three-Stage Deep Learning Model for Accurate Retinal Vessel Segmentation. *IEEE J. Biomed. Health Inform.* 2019, 23, 1427–1436. [CrossRef]
- Wang, B.; Qiu, S.; He, H. Dual Encoding U-Net for Retinal Vessel Segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention, Long Beach, CA, USA, 10–15 June 2019; pp. 84–92.
- 55. Ribeiro, A.; Lopes, A.P.; Silva, C.A. Ensemble learning approaches for retinal vessel segmentation. In Proceedings of the Portuguese Meeting on Bioengineering, Lisbon, Portugal, 22–23 February 2019; pp. 1–4.
- Khan, M.A.; Khan, T.M.; Naqvi, S.S.; Aurangzeb Khan, M. GGM classifier with multi-scale line detectors for retinal vessel segmentation. *Signal Image Video Process.* 2019, 13, 1667–1675. [CrossRef]
- 57. Muhammad, A.; Muhamamd, O.; Tahir, M.; Se Woon, C.; Kang Ryoung, P. Aiding the Diagnosis of Diabetic and Hypertensive Retinopathy Using Artificial Intelligence-Based Semantic Segmentation. *J. Clin. Med.* **2019**, *8*, 1446. [CrossRef]
- Wu, Y.; Xia, Y.; Song, Y.; Zhang, D.; Liu, D.; Zhang, C.; Cai, W. Vessel-Net: Retinal Vessel Segmentation Under Multi-path Supervision. In Proceedings of the Medical Image Computing and Computer Assisted Intervention, Long Beach, CA, USA, 10–15 June 2019; pp. 264–272.
- Feng, S.; Zhuo, Z.; Pan, D.; Tian, Q. CcNet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features. *Neurocomputing* 2020, 392, 268–276. [CrossRef]
- 60. Mahapatra, S.; Agrawal, S.; Mishro, P.K.; Pachori, R.B. A novel framework for retinal vessel segmentation using optimal improved frangi filter and adaptive weighted spatial FCM. *Comput. Biol. Med.* **2022**, *147*, 105770. [CrossRef]
- Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
- Alom, M.Z.; Hasan, M.; Yakopcic, C.; Taha, T.M.; Asari, V.K. Recurrent residual convolutional neural network based on U-net (R2U-net) for medical image segmentation. arXiv 2018, arXiv:1802.06955.
- 63. Zhuang, J. LadderNet: Multi-path networks based on U-Net for medical image segmentation. arXiv 2018, arXiv:1810.07810.
- 64. Gu, Z.; Cheng, J.; Fu, H.; Zhou, K.; Hao, H.; Zhao, Y.; Zhang, T.; Gao, S.; Liu, J. Ce-net: Context encoder network for 2d medical image segmentation. *IEEE Trans. Med. Imaging* 2019, *38*, 2281–2292. [CrossRef] [PubMed]
- Li, L.; Verma, M.; Nakashima, Y.; Nagahara, H.; Kawasaki, R. Iternet: Retinal image segmentation utilizing structural redundancy in vessel networks. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass Village, CO, USA, 1–5 March 2020; pp. 3656–3665.
- Zhang, Q.L.; Yang, Y.B. Sa-net: Shuffle attention for deep convolutional neural networks. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 2235–2239.
- Yuan, Y.; Zhang, L.; Wang, L.; Huang, H. Multi-level attention network for retinal vessel segmentation. *IEEE J. Biomed. Health Inform.* 2021, 26, 312–323. [CrossRef]
- 68. Zhang, T.; Li, J.; Zhao, Y.; Chen, N.; Zhou, H.; Xu, H.; Guan, Z.; Yang, C.; Xue, L.; Chen, R.; et al. MC-UNet Multi-module Concatenation based on U-shape Network for Retinal Blood Vessels Segmentation. *arXiv* 2022, arXiv:2204.03213.
- Romera, E.; Álvarez, J.M.; Bergasa, L.M.; Arroyo, R. ERFNet: Efficient Residual Factorized ConvNet for Real-Time Semantic Segmentation. *IEEE Trans. Intell. Transp. Syst.* 2018, 19, 263–272. [CrossRef]

- 70. Atli, I.; Gedik, O.S. Sine-Net: A fully convolutional deep learning architecture for retinal blood vessel segmentation. *Eng. Sci. Technol. Int. J.* **2021**, *24*, 271–283. [CrossRef]
- Laibacher, T.; Weyde, T.; Jalali, S. M2U-Net: Effective and Efficient Retinal Vessel Segmentation for Real-World Applications. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2019, Long Beach, CA, USA, 16–20 June 2019; pp. 115–124.