Segmentation of Retinal Blood Vessels Using Growing GANs

Christian Muth, Mark-Christian Bront, and Renata Georgia Raidou

TU Wien, Austria

Abstract

We investigate a progressively growing U-Net-based generative adversarial network (GAN) architecture for the segmentation of retinal blood vessels in fundus images. Employing growing GANs for segmentation tasks has yielded promising results, but it has not been applied to retinal vasculature segmentation yet. We start by training our GAN on downscaled versions of the training images. Subsequently, the learned weights of this core are transferred to the model, where additional convolution layers are appended at the front and back of the core and the model is incrementally trained with upsampled images. Our approach yields comparable results (i.e., similar metrics within a similar training time) to state-of-the-art approaches—opening interesting directions for employing growing GANs in retinal vasculature segmentation.

CCS Concepts
• Computing methodologies → Machine learning; • Applied computing → Life and medical sciences;

1. Introduction

Retinal imaging techniques serve ophthalmologists in detecting certain diseases and predispositions like increased blood pressure [SPJ19]. Yet, segmenting retinal images can often be time-consuming when done manually and automatic approaches are being investigated to streamline the diagnostic process. Basic segmentation approaches, though, work poorly as the images have several artifacts, such as bright and dark spots. Deep learning has the potential to yield better results but is computationally more difficult—especially for small vessels.

Following the advancements of neural networks, and specifically the generative adversarial network (GAN) architecture presented by Goodfellow et al. [GPAM*20], we investigate a progressively growing GAN for retinal imaging segmentation. Growing GANs employ a progressively growing training approach, where the generator and discriminator grow incrementally [KALL17]. Starting from a low resolution, new layers that model increasingly fine details are added as training progresses. Previously, Karras et al. [KALL17] proposed this idea for architectures using fully-connected layers. However, we wanted to apply this concept to a GAN architecture that consists of *convolutional layers*, as this design choice promotes flexible learning. To this end, we employ the base GAN architecture proposed by Son et al. [SPJ19] and adapt it to a growing GAN approach. Our work is the first known investigation of a growing GAN architecture for the segmentation of retinal blood vessels and a comparison to state-of-the-art approaches.

2. Our Growing GAN Approach

First, we train a model following the approach described by Son et al. [SPJ19]. They follow a typical GAN architecture with a generator and a discriminator. The discriminator receives either a real pair of a retinal image and its segmentation done by an expert or a fake pair that takes the output of the generator as segmentation. The discriminator receives the segmentation as a 4th color channel and shrinks the image size while increasing its depth to store the features it calculates. In the end, it outputs a "guess" of whether it thinks this is a real or fake pair—thus, calculating the loss for the generator. The generator follows a U-Net structure in which it receives a retinal image as input, shrinks the image, and increases the depth, calculating features. This can be understood as features describing small neighborhoods for the larger, but thinner layers and features describing larger areas for smaller and thicker layers. When the image size is increased again it takes in the upsampled features of the previous layer and the layer of the same size from before. For more details on the architecture, we refer to the original paper of Son et al. [SPJ19].

Karras et al. [KALL17] showed that starting a GAN training with a downsampled image and increasing its size over the training yields better results, more stable training, and faster convergence. The inner layers do not increase in depth—as opposed to the approach of Son et al. [SPJ19]. Therefore, to combine the two approaches we needed to only consider the trained weights of the layers after the first one. We also do not blend in the new layers. To achieve this growth we append additional convolutional layers at the beginning of both the generator and discriminator and at the end of the generator. The model is then trained with the less downsampled images, and the process is repeated until the full image size is reached, i.e., we grow from image sizes $(80, 80)$ to $(160, 160)$ to $(320, 320)$ and to $(640, 640)$. These were the input dimensions for the respective convolutions of the original ar-
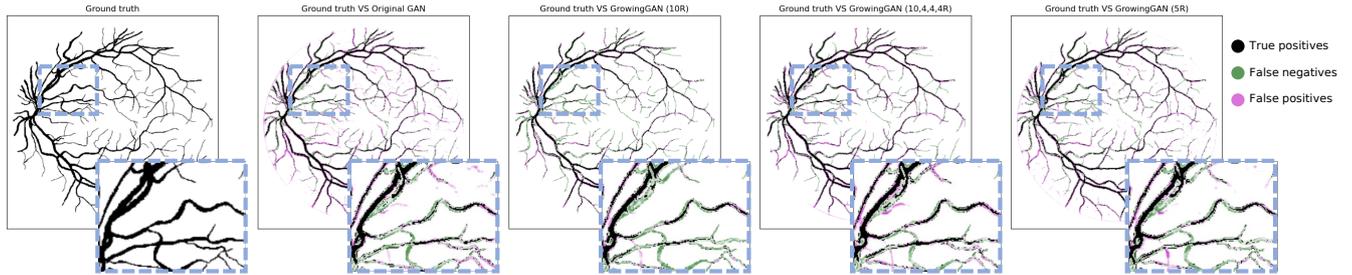
Figure 1: *Comparison of the applied approaches. The first column includes the ground truth only. The results of the second column were achieved with the implementation by Son et al. [SPJ19]. The third, fourth, and fifth columns depict the growing GAN approach with 10 rounds (i.e., 10,10,10,10), with a varying number of rounds (i.e., 10,4,4,4), and with 5 rounds (i.e., 5,5,5,5) respectively. The colors correspond to true positives (correct segmentations; black), false negatives (under-segmentations; green), and false positives (over-segmentations; magenta).*

chitecture. For the growing GAN, we investigate different growth rounds—namely, one approach with 10 rounds (i.e., 10,10,10,10), one with 5 rounds (i.e., 5,5,5,5), and one with a varying number of rounds (i.e., 10,4,4,4). This was intended to demonstrate the impact of different growth strategies on the architecture's performance.

Computational Resources. Our implementation is available on GitHub [git]. To train the models, we used a desktop with 64GB of RAM and an RTX 3090. It took about 3 hours for the model of Son et al., while the growing approaches with 10, 5, and a varying number of rounds took 4.7, 2.5, and 2.3 hours respectively.

3. Results and Conclusion

Our growing approaches were trained on the DRIVE dataset [SAN*04]. To evaluate our outcomes, we compare metrics of the area under the receiver operating characteristic (AU-ROC) and the area under the precision and recall curve (AU-PR), as these were also the metrics used by Son et al. [SPJ19]. For completion, we also include the Dice coefficient, the sensitivity (SE), and the specificity (SP). All metrics are included in Table 1.

In their original paper, Son et al. obtained an AU-ROC value of 0.9800 with their "image" discriminator and an AU-PR value of 0.9134 within 10 rounds of training. When re-training the same model ourselves, we obtained an AU-ROC of 0.9710 and AU-PR of 0.8892. The growing GAN model with the best accuracy score of 0.9502 was the one with the 5 rounds and yielded an AU-ROC value of 0.9600 and an AU-PR 0.8748. The calculated metrics are summarized in Table 1. Our results indicate that *our growing GAN approach yields comparable results to the original* by Son et al. and future work holds promise for better segmentation outcomes.

Figure 1 includes a visual comparison of the proposed approaches against the original approach and the ground truth. Here, the false negatives and false positives (i.e., the under- and over-segmentations respectively) are evident for the different growing GAN approaches. The approach with the varying growth generates more false positives than the other two approaches, while the approach with the 10 rounds yields more fragmented segmentations than the 5 rounds. In areas with thin vasculatures (see top right corner of the zoomed-in image), the original approach tends to generate more false positives than the growing approaches. Downsampling images and segmentations of vascular structures yield images resembling the full-size ones. For other tasks where this is not the case, it is questionable if the GAN training benefits from growing.

To conclude, we showed that growing GANs is a promising solution to retinal blood vessel segmentation. Growing GANs—once more mature—can enhance the accuracy of diagnosing retinal vascular diseases, enabling early detection and (tailored) intervention. They can also provide valuable data for tracking disease progression and treatment efficacy, leading to better-informed clinical decisions. Future work might combine aspects of new approaches, using noisy labels, alternative growing procedures, deeper convolutional neural networks, transformer layers, or different loss metrics.

References

[git] Retinal-Growing-GAN. URL: https://github.com/RedSweaterGuy/Retinal-Growing-GAN.git. 2

[GPAM*20] GOODFELLOW I., POUGET-ABADIE J., MIRZA M., XU B., WARDE-FARLEY D., OZAIR S., COURVILLE A., BENGIO Y.: Generative adversarial networks. *Communications of the ACM 63*, 11 (2020), 139–144. doi:https://doi.org/10.1145/3422622. 1

[KALL17] KARRAS T., AILA T., LAINE S., LEHTINEN J.: Progressive growing of GANs for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196* (2017). doi:https://doi.org/10.48550/arXiv.1710.10196. 1

[SAN*04] STAAL J., ABRÀMOFF M. D., NIEMEIJER M., VIERGEVER M. A., VAN GINNEKEN B.: Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging 23*, 4 (2004), 501–509. doi:10.1109/TMI.2004.825627. 2

[SPJ19] SON J., PARK S. J., JUNG K.-H.: Towards accurate segmentation of retinal vessels and the optic disc in fundoscopic images with generative adversarial networks. *Journal of digital imaging 32*, 3 (2019), 499–512. doi:https://doi.org/10.1007/s10278-018-0126-3. 1, 2

Model	AU-ROC	AU-PR	Dice	ACC	SE	SP
Original [SPJ19]	0.9710	0.8892	0.8032	0.9510	0.7855	0.9751
GrowingGAN (10R)	0.9664	0.8754	0.7604	0.9471	0.6584	0.9893
GrowingGAN (10,4,4,4R)	0.9600	0.8748	0.7912	0.9502	0.7414	0.9806
GrowingGAN (5R)	0.9670	0.8774	0.7897	0.9492	0.7487	0.9784

Table 1: *Resulting metrics for the tested models: area under the receiver operating characteristic (AU-ROC), area under the precision and recall curve (AU-PR), Dice coefficient, accuracy (ACC), sensitivity (SE), and specificity (SP).*