Exploring Generative Adversarial Networks for Diabetic Foot Ulcer Image Segmentation

David Roth, B.S. in Mathematics and Actuarial Science
North Dakota State University and University of Iowa; Funded by the NSF

Abstract

The early and accurate diagnosis of diabetic foot ulcers (DFUs) is crucial for effective patient care. However, the conventional approach of visual inspection and manual measurements by medical experts can be subject to human error, leading to limited precision in ulcer assessment. To overcome these limitations and enhance the diagnostic process, this research focuses on leveraging advanced image segmentation techniques. While certain convolutional neural network architectures, such as U-Net and SegNet, have been applied for image segmentation, this paper delves into exploring the untapped potential of Generative Adversarial Networks (GANs) in this domain. GANs have shown remarkable success in various computer vision tasks, including image generation and image-to-image translation. We aim to investigate the effectiveness of GAN-based image segmentation methods, particularly Pix2Pix and SegAN, in accurately identifying and segmenting DFU medical images. To accomplish this, we propose the use of performance measures such as the Dice Coefficient and the Jaccard Index, among others. By identifying the most effective GAN-based approach for DFU segmentation, this research seeks to contribute to the development of more reliable and automated diagnostic tools, leading to improved patient outcomes and reduced workload for healthcare professionals.

Introduction

Diabetes presents significant health challenges, with complications like neuropathy, retinopathy, and cardiovascular issues. Diabetic foot ulcers (DFUs) are critical, affecting about one in four diabetic patients, with many remaining unhealed, leading to potential amputations.

To improve DFU care and treatment outcomes, precise and frequent wound measurements are vital. Conventional methods relying on visual inspection and manual measurements have limitations due to human error and subjectivity.

Medical image segmentation, a computer vision technique, partitions an image into distinct regions of interest. For DFUs, image segmentation accurately identifies affected areas, aiding treatment planning and monitoring.

This research explores advanced segmentation techniques, leveraging Generative Adversarial Networks (GANs), a revolutionary machine learning breakthrough. GANs consist of a generator and a discriminator engaged in an adversarial process. The generator creates synthetic images resembling samples from a training set, while the discriminator distinguishes real from synthetic images. Through iterative adversarial training, GANs produce highly realistic synthetic images. In DFU diagnostics, GANs hold immense potential. Training on real ulcer images, the generator synthesizes realistic ulcer representations, while the discriminator learns to distinguish real and synthetic ulcers.

Integrating GANs into medical image segmentation could significantly enhance DFU diagnosis, providing better wound measurement accuracy and improved patient outcomes. This cutting-edge approach may reduce the need for invasive procedures and discomfort, thus improving the overall quality of life for patients suffering from DFUs.

Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs) are a powerful framework in machine learning created in 2014 by Ian Goodfellow and colleagues. They consist of two neural networks engaged in a competitive "minimax" game. The generator takes random noise as input and transforms it into synthetic data, closely resembling real samples from a dataset. The discriminator acts as a binary classifier, distinguishing real from synthetic data.

The training process involves random initialization of the generator and discriminator. The generator produces synthetic samples, and the discriminator tries to classify them along with real data. Both networks continuously adapt through iterative backpropagation, with the generator improving its ability to create realistic data and the discriminator becoming more discerning.

As training progresses, the generator learns to create data indistinguishable from real samples, fooling the discriminator. This adversarial interplay allows GANs to learn the underlying data distribution, generating highly realistic data.

GANs' remarkable capability has led to groundbreaking applications in image generation, image-to-image translation, and medical image segmentation, revolutionizing various fields.

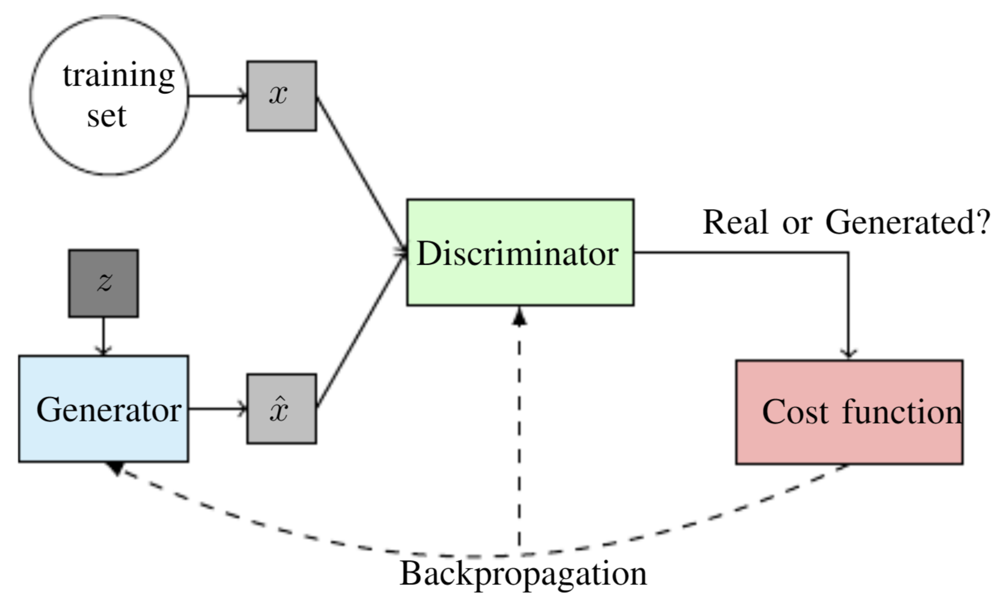

Figure 1: Diagram of the GAN general training process

The Dataset

Chaunbo Wang et al. created their dataset to address the scarcity of publicly available datasets for training deep-learning-based wound segmentation models. Partnering with the Advancing the Zenith of Healthcare (AZH) Wound and Vascular Center in Milwaukee, WI, they collected 1109 foot ulcer images from 889 patients over two years. Each image is paired with a manually created segmentation mask by a medical expert, making it a valuable resource for evaluating deep learning algorithms in wound segmentation.

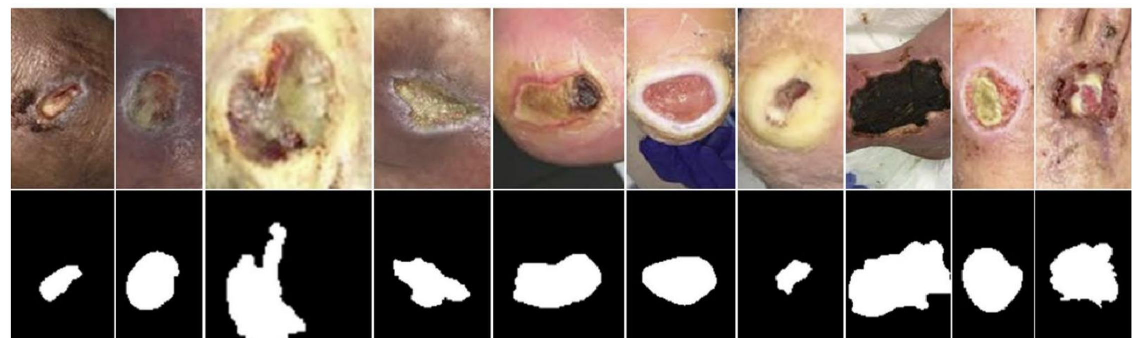

Figure 2: A few examples from the AZH dataset []. Each column shows the raw image and the corresponding ground truth segmentation.

GAN Architectures: Pix2Pix & SegAN

Pix2Pix

Pix2Pix is a powerful image-to-image translation model introduced by Phillip Isola and colleagues in 2017. Image-to-image translation involves transforming one representation of a scene into another (see Figure 3), and Pix2Pix achieves this by making pixel-by-pixel changes to an input image to produce a corresponding output image with desired properties. For instance, it can transform an image of a horse into a zebra, altering the mane to resemble the black and white stripes of a zebra while retaining other features.

Pix2Pix leverages the concept of conditional generative adversarial networks (cGANs), where the generator receives not only random noise but also additional conditional information, which, in this case, is the input image itself. This allows the generator to produce the desired output based on the input. Pix2Pix's versatility makes it suitable for various image translation tasks.

The Pix2Pix generator adopts an encoder-decoder structure with skip connections, similar to the U-Net architecture. The encoder extracts high-level features from the input image, while the decoder reconstructs the segmented output by gradually upsampling the encoded features. Skip connections maintain fine-grained details during the upsampling process, enhancing the quality of the output.

The Markovian Discriminator, also known as PatchGAN, is a crucial component that evaluates the realism and quality of the generated images. Unlike traditional discriminators, PatchGAN focuses on local image patches, enabling it to capture high-frequency structural details effectively.

The training process involves iteratively updating the discriminator and generator using minibatch stochastic gradient descent with the Adam solver. Measures are taken to control the relative training rate of the discriminator for better results.

Pix2Pix's combination of the generator's image-to-image translation capabilities and the PatchGAN discriminator's focus on local patches makes it suitable for tasks like diabetic foot ulcer image segmentation. The model can generate visually appealing results with fine-grained information, addressing the challenges of image synthesis tasks and producing high-quality segmented outputs.

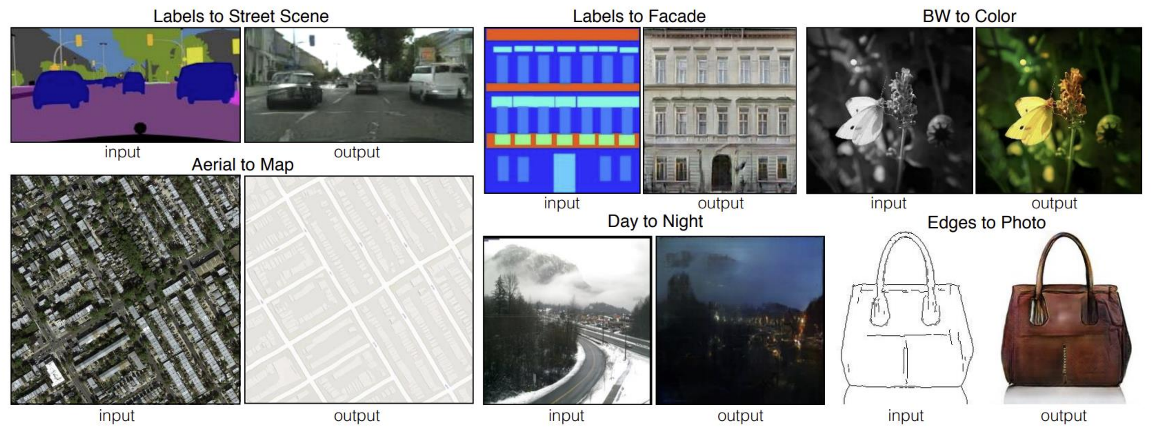

Figure 3: Examples of image-to-image translation that Pix2Pix is capable of []

SeGAN

SegAN, introduced by Xue and Xu in 2018, is a medical image segmentation framework inspired by GANs. Its unique contribution is the use of a critic network with a multi-scale L1 loss function for better spatial understanding. This loss function measures the absolute difference between features extracted from the critic network for the masked input image and the corresponding ground truth. By training the segmentor and critic in an adversarial manner, SegAN achieves effective medical image segmentation.

GAN Architectures: SegAN (Continued)

The segmentor's architecture resembles Pix2Pix with an encoder-decoder structure similar to U-Net. It efficiently learns and extracts features from input images, aided by skip connections to preserve fine-grained details during upsampling. The critic, structured like the decoder in the segmentor, extracts hierarchical features at different scales to capture long and short-range spatial relationships between pixels.

The training process involves alternately training the segmentor and critic in a min-max game. First, the critic maximizes the multi-scale L1 loss function to provide valuable feedback to the segmentor. Then, the segmentor minimizes the same loss function to generate accurate segmentation maps resembling the ground truth. This alternating training makes both networks more powerful over time. During training, RMSProp optimization with a batch size of 64 and a learning rate of 0.00002 is used. The architectures of both networks are selected through a grid search process. SegAN's bounded loss function ensures convergence during training.

SegAN outperforms the state-of-the-art U-Net method, providing accurate and smooth predicted label maps for medical image segmentation tasks. Its success lies in the adversarial training scheme and the multi-scale L1 loss function, enabling the model to capture global and local image elements effectively.

Results (Pending) & Conclusion

At this stage of our research, we are unable to present definitive results due to implementation challenges and hyperparameter issues. While the Pix2Pix model was functioning correctly, we encountered problems with the learning rate, resulting in the model generating predominantly black images. However, adjustments to the hyperparameters, such as reducing the learning rate and increasing the L1 lambda parameter, showed promising improvements in segmentation quality.

Despite not presenting complete results in this poster, the modified Pix2Pix model exhibits potential for generating more accurate DFU segmentations. Additionally, we are on the verge of conducting experiments with the SegAN architecture and are excited to report our findings soon. Stay tuned for forthcoming results and evaluations of SegAN's performance in DFU image segmentation.

Conclusion:

Our research explores the potential of GANs for DFU image segmentation, focusing on Pix2Pix and SegAN architectures. Though conclusive results are pending, our study contributes to the development of more accurate and automated DFU diagnostic tools, benefiting patients and healthcare professionals alike. GANs hold promise to revolutionize the DFU diagnostic process, leading to improved patient outcomes and reduced workload for medical practitioners.

Acknowledgments & References

The authors acknowledge the support of the NDSU Department of Computer Science sponsored by NSF Directorate for Engineering and EPSCoR Co-Funding through NSF grant EEC-2050175. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

[1] A. Mahbod, G. Schaefer, R. Ecker, and I. Ellinger, "Automatic Foot Ulcer Segmentation Using an Ensemble of Convolutional Neural Networks," 2022.
[2] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015.
[3] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," 2016.
[4] J. A. Rubio and et al., "Mortality in Patients with Diabetic Foot Ulcers: Causes, Risk Factors, and Their Association with Evolution and Severity of Ulcer," Journal of Clinical Medicine, vol. 9, no. 9, p. 3009, 2020.
[5] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 7, pp. 3523–3542, 2022.
[6] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," 2014.
[7] M. Ponti, L. Ribeiro, T. Nazare, T. Bui, and J. Collomosse, "Everything You Wanted to Know about Deep Learning for Computer Vision but Were Afraid to Ask," pp. 17–41, 2017.
[8] T. A. A. T. e. a. Jeong, J.J., "Systematic Review of Generative Adversarial Networks (GANs) for Medical Image Classification and Segmentation," Journal of Digital Imaging, vol. 35, pp. 137–152, 2022.
[9] Y. G¨und¨u¸c, "Vit-GAN: Image-to-image Translation with Vision Transformes and Conditional GANS," 2021.[10] M. Z. Hossain, F. Sohel, M. F. Shiratuddin, H. Laga, and M. Bennamoun, "Text to Image Synthesis for Improved Image Captioning," IEEE Access, vol. 9, pp. 64918–64928, 2021.
[11] A. D. W. V. e. a. Wang, C., "Fully Automatic Wound Segmentation with Deep Convolutional Neural Networks," Scientific Reports, vol. 10, p. 21897, 2020.
[12] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional 12 Adversarial Networks," 2018.
[13] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," 2014.
[14] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context Encoders: Feature Learning by Inpainting," 2016.
[15] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," 2016.
[16] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based Generative Adversarial Network," 2017.
[17] H. Noh, S. Hong, and B. Han, "Learning Deconvolution Network for Semantic Segmentation," in 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1520–1528, 2015.
[18] A. Andonian, T. Park, B. Russell, P. Isola, J.-Y. Zhu, and R. Zhang, "Contrastive Feature Loss for Image Prediction," 2021.
[19] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 2017.
[20] Y. Xue, T. Xu, H. Zhang, L. R. Long, and X. Huang, "SegAN: Adversarial Network with Multiscale L1 Loss for Medical Image Segmentation," Neuroinformatics, vol. 16, pp. 383–392, May 2018.