Multi-Modality Microscopy Image Style Augmentation for Nuclei Segmentation

Sophia J. Wagner[*1,2] SOPHIA.WAGNER@HELMHOLTZ-MUENCHEN.DE
[1] *Technical University Munich, Munich, Germany*
[2] *Helmholtz AI, Neuherberg, Germany*

Ye Liu[*1] YEEE.LIU@TUM.DE
Tingying Peng[2] TINGYING.PENG@HELMHOLTZ-MUENCHEN.DE

Editors: Under Review for MIDL 2022

Abstract

Annotating microscopy images for nuclei segmentation is laborious and time-consuming. To leverage the few existing annotations, also across multiple modalities, we propose a microscopy-style augmentation technique based on a generative adversarial network (GAN). Unlike other style transfer methods, it can not only deal with different cell assay types and lighting conditions but also with different imaging modalities, such as bright-field and fluorescence microscopy. Using disentangled representations for content and style, we can preserve the structure of the original image while altering its style during augmentation. We evaluate our data augmentation on the 2018 Data Science Bowl dataset, consisting of various cell assays, lighting conditions, and imaging modalities. With our style augmentation, the segmentation accuracy of the two top-ranked Mask R-CNN-based nuclei segmentation algorithms in the competition increases significantly. Thus, our augmentation technique renders the downstream task more robust to the test data heterogeneity and helps counteract class imbalance without resampling of minority classes[1].

Keywords: Data augmentation, style transfer, disentanglement, nuclei segmentation.

1. Introduction

The evaluation of cell-level features, such as nuclei shape and distribution, is a key task in the histopathological workflow. However, deep learning models require accurate and time-consuming segmentation masks. In this paper, we propose to facilitate network training by a GAN-based style transfer data augmentation technique as has been shown to be effective for histological images (Wagner et al., 2021). By synthesizing less-represented image types from well-represented ones, our style augmentation can increase the amount of images of minority types in the training set. We evaluate the augmentation technique on the dataset of the Kaggle competition 2018 Data Science Bowl (DSB'18)[2], which is highly imbalanced in the contained imaging modalities. For nuclei segmentation, we used two implementations from the top-5 ranked methods of the competition leaderboard that provided their code as baselines. Both implementations apply Mask R-CNN for instance segmentation and differ only in the image pre-processing and preparation of the network training. In the following, we focus on the training of our style transfer network.

* Contributed equally
1. For a detailed presentation of the method, please refer to Liu et al. (2022).
2. https://www.kaggle.com/c/data-science-bowl-2018

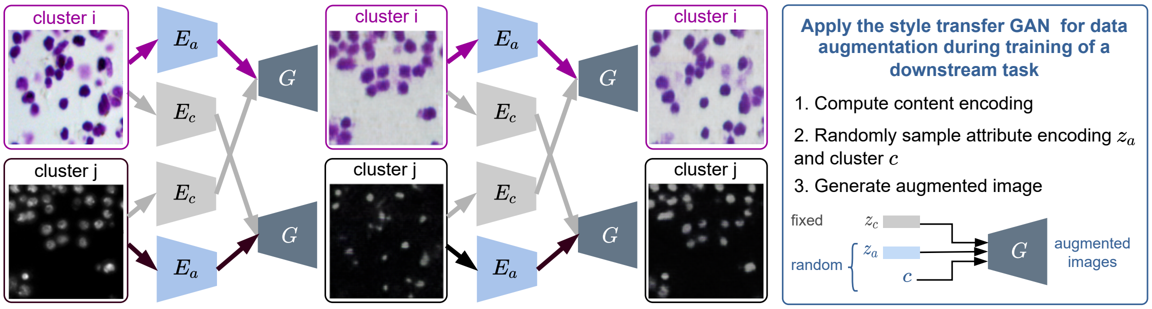

Figure 1: Left: our style transfer GAN is trained with a cross-cycle consistency loss. Right: how the GAN is applied for data augmentation during training the nuclei segmentation network.

2. Materials and Methods

Our nuclei segmentation workflow consists of three steps: data augmentation (clustering into modalities and training of the GAN), training of the instance segmentation network, and evaluating the segmentation network with test time augmentation.

Clustering. First, the dataset needs to be clustered into different modalities, assuming the content space is shared between the modalities. We divide the training images of the DSB'18 dataset into six clusters based on their HSV representation using the K-means algorithm. Each cluster should ideally correspond to one imaging modality (see Figure 2).

Multi-Modality Style Transfer. The style transfer GAN consists of two encoders (disentangling image style and content) and a generator that takes content, style, and domain encoding as input (Lee et al., 2018). The network is trained with a cross-cycle consistency loss, such that paired images are not required (see Figure 1). An adversarial loss on the content encoder additionally enforces domain independence of the content encodings. This ensures that only the appearance is changed from modality to modality.

To apply the style transfer GAN as augmentation technique, we sample attribute and domain vector randomly, while leaving the content-encoding fixed (see Figure 1). We randomly augment half of the training images additionally to standard augmentations. Since the content encoding is fixed during augmentation, the augmented image has the same nuclei location and shape as the input image and thus inherits the nuclei segmentation mask from the original image. This is a key difference between our approach and a common CycleGAN-based image style transfer (Zhu et al., 2017), where there is no guarantee of content invariance.

3. Results and Conclusion

Figure 2 shows samples created by our multi-modality style transfer GAN from one domain to the others. We quantify the add-on value of our proposed augmentation method by training models with and without our augmentation. The final evaluation score is based on the Intersection-over-Union (IoU) metric, determined by submitting our segmentation results of the second-stage test dataset to the Kaggle competition. As shown in Table 1, including our augmentation in the methods increased the nuclei segmentation accuracy

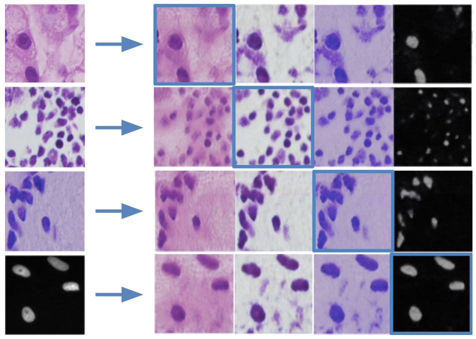

Figure 2: Results of style transfer network.

Method	IoU score
BIOMAGic	0.570
Deep Retina	0.532
Deep Retina + our aug.	**0.609**
Inom Mirzaev	0.599
Inom Mirzaev + our aug.	**0.613**

Table 1: Results on the DSB'18 dataset.

from 53.2% to 60.9% for Deep Retina[3] and 59.9% to 61.3% for Inom Mirzaev[4], respectively. Notably, a score around 61% was ranked among the top-5 and almost only achieved by using additional datasets. In addition to the above two baseline methods, we also quote the result from the team BIOMAGic (57.0%), the only method using style transfer among the top-25 submissions during the competition.

In summary, we developed an augmentation technique using a multi-modality style transfer GAN to transfer microscopy nuclei images between imaging modality. During training a Mask R-CNN for nuclei segmentation, this augmentation strategy facilitates the training by increasing the diversity of the training images, hence making it more robust to the test data heterogeneity and resulting in better segmentation accuracy.

Acknowledgments

The first author is supported by the Munich School for Data Science (MUDS).

References

Hsin-Ying Lee, Hung-Yu Tseng, Jia-Bin Huang, Maneesh Kumar Singh, and Ming-Hsuan Yang. Diverse image-to-image translation via disentangled representations. In *European Conference on Computer Vision*, 2018.

Ye Liu, Sophia J Wagner, and Tingying Peng. Multi-Modality Microscopy Image Style Augmentation for Nuclei Segmentation. *Journal of Imaging*, 8(3), 2022. ISSN 2313-433X. doi: 10.3390/jimaging8030071. URL https://www.mdpi.com/2313-433X/8/3/71.

Sophia J. Wagner et al. Structure-preserving multi-domain stain color augmentation using style-transfer with disentangled representations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 257–266. Springer, 2021.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

3. https://github.com/Lopezurrutia/DSB_2018
4. https://github.com/mirzaevinom/data_science_bowl_2018