

Neural Style Transfer to deal with the Domain Shift Problem on Glioblastoma Segmentation

Manuel García-Domínguez, César Domínguez, Jónathan Heras, Eloy Mata, and Vico Pascual

dept. of Mathematics and Computer Science

University of La Rioja

Logroño, Spain

{manuel.garciad, cesar.dominguez, jonathan.heras, eloy.mata, vico.pascual}@unirioja.es

Abstract—Domain shift is a generalisation problem of machine learning models that occurs when the data distribution of the training set is different to the data distribution encountered by the model when it is deployed. This is common in the context of biomedical image segmentation due to the variance of experimental conditions, equipment, and capturing settings. In this work, we address this challenge by studying neural style transfer algorithms in the context of the segmentation of tumour spheroids. We show that, using these algorithms, it is possible to recover the performance of a model that suffers from the domain shift problem. Finally, we provide an API to facilitate the use of neural style transfer techniques in the bioimaging context.

Index Terms—Domain adaption, Style Transfer, Semantic Segmentation, Glioblastoma, Spheroids

I. INTRODUCTION

Deep convolutional neural networks have become the state-of-the-art approach to tackle segmentation problems in medicine [19], [22]. However, there are several challenges that hinder the training and deployment of deep learning models in this context. First of all, a considerable amount of annotated images is needed to train a deep model, and annotating datasets for image segmentation is a tedious and time-consuming task that requires expert knowledge [13]. Moreover, there is an important generalisation challenge when using trained models that is known as *domain shift* (also known as distribution shift) [2], [4]. This problem arises when the data distribution of the training dataset employed for training a model is different to the data that the model encounters when deployed. This is common in biomedical datasets since images greatly vary due to experimental conditions, and the equipment (for instance, microscopes) and settings (for instance, focus and magnification) employed for capturing those images.

This generalisation problem can be tackled by combining datasets from multiple sources [7] or using techniques like data augmentation [21]; nevertheless, it is not possible to foresee every new and unknown distribution. A different approach consists in applying transfer learning [18], a technique that, instead of training a model from scratch, reuses a model pre-trained in a source dataset to train a new model in a target dataset. However, this requires the annotation of the target dataset, a time-consuming task that should be carried

out for every new dataset. A different approach to handle the domain shift problem is image-to-image translation [8], a set of techniques that aim to learn the mapping between an input image and an output image using a training set of aligned image pairs; however, this requires paired data from the source and target domains, a challenge that can be faced by using unpaired image-to-image translation [26].

Unpaired image-to-image translation methods translate an image from a domain A to a domain B , and vice versa, in the absence of paired examples. This approach has been already employed in several medical segmentation tasks; for instance, the segmentation of the left ventricle in magnetic resonance images [24], the segmentation of digitally reconstructed radiographs [25], and the segmentation of magnetic resonance imaging (MRI), abdominal CT and MRI, and mammography X-rays [9]. All these works are based on variants of CycleGAN [26], an unpaired image-to-image translation method based on Generative Adversarial Networks (GANs) that requires two datasets: one of them contains images from the distribution employed for training the segmentation model, and the other contains images acquired in a different setting. This approach poses two challenges. First, both datasets must be available, and this might be an issue due to privacy concerns [1]; and, secondly, CycleGAN variants must be trained, a process that demands the usage of GPUs and might be challenging for several users due to the difficulties of training GAN models [20]. The approach proposed in this paper to tackle these drawbacks consists in using style transfer methods [5]; that is, techniques that render the content of an image using the style of another. Those techniques do not require a training process, and it is enough with releasing one image of the dataset employed for training the model that suffers the domain shift problem.

In this work, we have studied 3 neural style transfer techniques to deal with the domain shift problem in the context of segmenting tumour spheroids [15]. In this task, we have observed, see Section II, that models that achieve a mean IoU over 97% when evaluating with data following the same distribution as the training set, fail when they are employed with data following a different distribution (the IoU is, in some cases, under 15%). We have faced this domain shift problem by using style transfer techniques. Namely, the contributions of our work are:

This work was partially supported by Ministerio de Economía y Competitividad [MTM2017-88804-P], and a FPI Grant of Community of La Rioja

Dataset	# Images	Image size	Microscope	Magnification	Format	Type
BL5S	50	1296 × 966	Leica	5x	TIFF	RGB
BN2S	154	1002 × 1004	Nikon	2x	ND2	Gray 16bits
BN10S	105	1002 × 1004	Nikon	10x	ND2	Gray 16bits
BO10S	64	3136 × 2152	Olympus	10x	JPG	RGB

TABLE I

FEATURES OF THE 4 DATASETS EMPLOYED IN THIS WORK. THE FORMER THREE DATASETS WERE EMPLOYED FOR TRAINING AND THE LAST DATASET WAS USED FOR TESTING

- We explore several state-of-the-art style transfer translation methods to tackle the domain shift problem in the context of tumour spheroid segmentation.
- We demonstrate the effectiveness of using style transfer to improve the performance of a variety of advanced deep segmentation networks.
- We provide an API to apply the studied methods not only in the context of spheroid segmentation but in general for medical imaging tasks. The API is available at <https://github.com/ManuGar/ImageStyleTransfer>

II. MATERIALS

Spheroids are the most widely used 3D models to study cancer since they can be used for studying the effects of different micro-environmental characteristics on tumour behaviour and for testing different preclinical and clinical treatments [15]. The images from tumour spheroids greatly vary depending on the experimental conditions, and also on the equipment (microscopes) and conditions (focus and magnification) employed to capture the images [11].

For our experiments, we have employed the 4 datasets presented in [11]; a description of those datasets is provided in Table I, and an image of each dataset is shown in Figure 1. As can be noticed from Table I and Figure 1, there are considerable differences among the images of each dataset. Three of those datasets (the BL5S, BN2S, and BN10S datasets) were employed for training 4 segmentation models (using the algorithms DeepLab v3 [3], HRNet Seg [23], U-net [19] and U²-Net [17]) and the last dataset (the BO10S dataset) was employed for testing. We have used this dataset split because the last dataset comes from a different laboratory; so, its style will not be the same as the others. The definition of the 4 segmentation architectures is available in the SemTorch package¹. All the architectures were trained with the libraries PyTorch [16] and FastAI [6] and using a GPU Nvidia RTX 2080 Ti. In order to set the learning rate for the different architectures, we employed the procedure presented in [6]; and, we applied early stopping when training all the architectures to avoid overfitting. The metric employed to measure the accuracy of the different methods is the IoU, also known as Jaccard index — this metric measures the area of intersection between the ground truth and the predicted region over the area of union between the ground truth and the predicted region.

When the models were evaluated using a test set formed by images following the same distribution than the training set,

¹The SemTorch package is available at <https://github.com/WaterKnight1998/SemTorch>

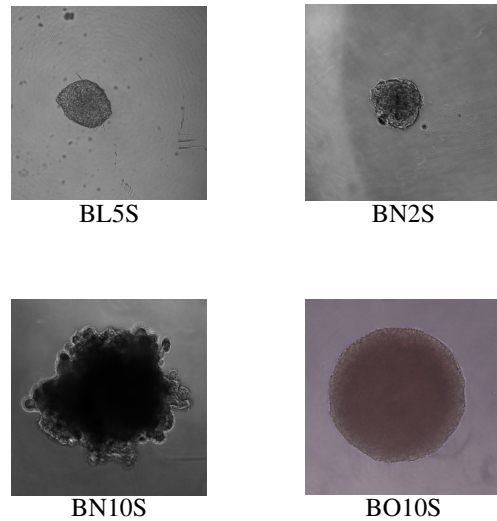


Fig. 1. Samples from the 4 datasets employed in this work

	DeepLab v3	HRNet-Seg	U-Net	U2-Net
BL5S-BN2S-BN10S	97.00	97.32	97.25	97.26
BO10S	83.61	92.65	13.64	95.65

TABLE II

PERFORMANCE OF THE 4 MODELS WHEN EVALUATING IN A TEST SET FORMED FROM IMAGES FOLLOWING THE SAME DISTRIBUTION THAN THE TRAINING SET (BL5S-BN2S-BN10S), AND WHEN EVALUATED USING A TEST SET FROM A DIFFERENT DISTRIBUTION (BO10S)

the 4 models achieved a performance over 97%, see Table II. On the contrary, when those models were employed with images captured under different conditions (namely, using the BO10S dataset), the performance of the models decreased by up to 84%. In the next section, we explore how style transfer methods can serve to deal with the domain shift problem in this context.

III. STYLE TRANSFER

This section is devoted to present how style transfer methods can handle the domain shift problem. In addition, we introduce the API that we have developed to facilitate the use of those methods.

We start by explaining the procedure to apply style transfer methods to deal with the domain shift problem of a model — such a procedure is summarised in Figure 2. We assume that a model has been trained using a source dataset of images, and we are interested in applying such a model to obtain the prediction associated with an image from a different distribution than the source dataset; we call this image, the target image. Instead of feeding the target image directly to the model, we first take an image from the source dataset and transfer the style of that image to the target image but preserving its content producing a transformed image. Finally, the transformed image is fed to the model to obtain the associated prediction.

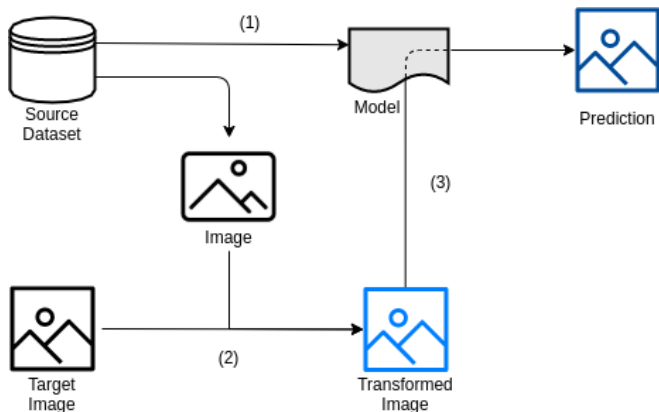


Fig. 2. Workflow of the style transfer approach. (1) A model is trained using a source dataset. (2) The target image is transformed using the style from an image of the source dataset. (3) The transformed image is fed to the model.

The key component of the aforementioned process is the algorithm that transfers the style from the source dataset but keeping the content of the target image. In the literature, there are several style transfer algorithms [14]; but, for our experiments, we have focused on three of them: neural style transfer (NST) [5], an optimisation technique that uses a Convolutional Neural Network (CNN) to decompose the content and style from images; deep image analogy [12], a method that finds semantically-meaningful correspondences between two input images by adapting the notion of image analogy with features extracted from a CNN; and STROTSS [10], a variant of the NST algorithm that changes the optimisation objective of NST.

It is worth noting that the style transfer approach presented here can be applied to deal with the domain shift problem not only for segmentation problems, as in our work, but also to other computer vision tasks. Hence, these methods can be helpful for a great variety of problems. However, it might be difficult to apply these techniques since they are implemented in different libraries and using different frameworks, and each of them has its own particularities. In this work, we have addressed this drawback by developing a high-level Python API that allows the integration of style transfer algorithms independently of their underlying library and framework. The API currently includes the aforementioned methods (the project webpage provides information about the library that implements each method) and can be easily extended with new techniques. In order to apply the previously introduced procedure using our API, users only have to provide the style image, the target image, and the name of the algorithm to apply; the rest of the transformation process is automatically conducted by the API.

IV. RESULTS AND DISCUSSION

In our running example of segmenting tumour spheroids, and using our API, we randomly picked an image from the combination of the datasets BL5S, BN2S, and BN10S, and used it to transform the images from the BO10S dataset. Subsequently, we fed those images to the segmentation models

	DeepLab v3	HRNet-Seg	U-Net	U ² -Net
Base	83.61	92.65	13.64	95.65
NST	95.64 \uparrow	94.91 \uparrow	89.21 \uparrow	95.89 \uparrow
Deep Image Analogy	0.00 \downarrow	45.13 \downarrow	0.66 \downarrow	0.84 \downarrow
STROTSS	94.86 \uparrow	92.38 \downarrow	78.08 \uparrow	94.14 \downarrow

TABLE III
PERFORMANCE FOR THE BO10S DATASET USING THE DIFFERENT STYLE-TRANSFER METHODS TO DEAL WITH THE DOMAIN SHIFT PROBLEM. A \uparrow INDICATES AN IMPROVEMENT WITH RESPECT TO THE BASE MODEL, WHEREAS A \downarrow INDICATES A DECLINATION IN THE PERFORMANCE.

presented in Section II, and evaluated their performance, see Table III. From the 3 studied style transfer algorithms, both the NST and STROTSS algorithms handle the domain shift problem; whereas, the images transformed with the deep image analogy algorithm produce even worse results than the original images from the BO10S dataset. Using the NST algorithm, all the segmentation models improve their IoU (the U-Net model improves its performance from 13.64% to 89.21%, and the other models have an IoU close to 95%). For the STROTSS algorithm, the results are also positive: two of the segmentation models improve (DeepLab and U-Net), and the other two achieve worse results, but still their IoU is over 92%.

We can also visually inspect the images produced by the different transformation algorithms to discover the difficulties faced by the segmentation models, see Figure 3. We can notice that the 2 successful models (NST and STROTSS) produce images that preserve the content of the image but with a style that is similar to the style of those used for training the segmentation models. On the contrary, the deep image analogy method does not keep the content of the image; and, thus the segmentation models are not able to properly segment the images.

From Figure 3, we can also appreciate the sensibility of the segmentation models to variations in the input image. The HRNet Seg and U²-net models are more robust than the DeepLab and U-net models — recall that all the models achieved an IoU over 97% when evaluating in data from the distribution of the training set. Hence, the style transfer methods can be employed not only to deal with the domain shift problem of computer vision models, but also to evaluate the robustness of such models.

V. CONCLUSIONS

In this paper, we have studied the benefits of applying style transfer techniques to deal with the domain shift problem in the context of tumour segmentation. The results show us that, using those translation methods, it is possible to recover the performance of a model that suffers from the domain shift problem. In contrast with other existing methods that deal with the domain shift problem, such as image-to-image translation models, style transfer algorithms have the advantage of not requiring a training step, and can be deployed by providing a single image from the source dataset.

In this work, we have only used style transfer techniques, but it remains as future work to compare these methods

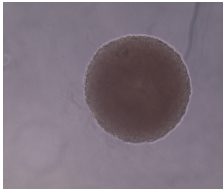




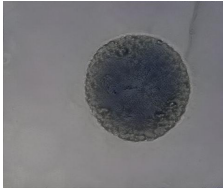




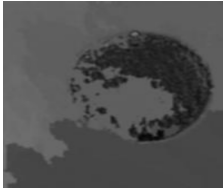

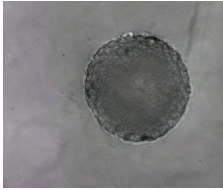
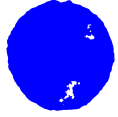



	Image	DeepLab v3	HRNet Seg	U-Net	U ² -Net
Base					
NST					
Deep Image Analogy					
STROTSS					

Fig. 3. An example showing the segmentation produced by the DeepLab, HRNet, U-Net and U²-Net models after applying a style transfer algorithm or an image-to-image translation model to a given image

with unpaired image-to-image translation methods and other techniques that deal with domain shift. Since, it is known that not all algorithms work equally well for all problems, we plan to extend our API to include several image-to-image methods using GANs. Finally, we will test the performance of both neural style transfer methods and unpaired image-to-image models with other kinds of images and computer vision tasks.

REFERENCES

- [1] Adler-Milstein, J., Jha, A.K.: Sharing clinical data electronically: a critical challenge for fixing the health care system. *JAMA* **307**(16), 1695–1696 (2012). <https://doi.org/10.1001/jama.2012.525>
- [2] Arvidsson, I., et al.: Generalization of prostate cancer classification for multiple sites using deep learning. In: *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. pp. 191–194. *IEEE* (2018). <https://doi.org/10.1109/ISBI.2018.8363552>
- [3] Chen, L.C., Zhu, Y., Papandreou, G., et al.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Computer Vision – ECCV 2018*. pp. 833–851. Springer International Publishing (2018). https://doi.org/10.1007/978-3-030-01234-2_49
- [4] Choudary, A., et al.: Advancing medical imaging informatics by deep learning-based domain adaptation. *Year book of medical informatics* **29**(1), 129–138 (2020). <https://doi.org/10.1055/s-0040-1702009>
- [5] Gatys, L., Ecker, A., Bethge, M.: A neural algorithm of artistic style. *Journal of Vision* **16**, 326 (2016). <https://doi.org/10.1167/16.12.326>
- [6] Howard, J., Gugger, S.: *Deep Learning for Coders with FastAI and Pytorch*. O'Reilly (2020)
- [7] Irvin, J., Rajpurkar, P., Ko, M., et al.: Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In: *Thirty-Third AAAI Conference on Artificial Intelligence. AAAI'19*, vol. 33, pp. 590–597 (2019)
- [8] Isola, P., et al.: Image-to-image translation with conditional adversarial networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 5967–5976. *IEEE* (2017). <https://doi.org/10.1109/CVPR.2017.632>
- [9] Jinzheng, C., et al.: Towards cross-modal organ translation and segmentation: A cycle- and shape-consistent generative adversarial network. *Medical Image Analysis* **52**, 174 – 184 (2019). <https://doi.org/10.1016/j.media.2018.12.002>
- [10] Kolkin, N., Salavon, J., Shakhnarovich, G.: Style transfer by relaxed optimal transport and self-similarity. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 10043–10052. *IEEE* (2019). <https://doi.org/10.1109/CVPR.2019.01029>
- [11] Lacalle, D., et al.: SpheroidJ: An Open-Source Set of Tools for Spheroid Segmentation. *Computer Methods and Programs in Biomedicine* (2021). <https://doi.org/10.1016/j.cmpb.2020.105837>
- [12] Liao, J., et al.: Visual attribute transfer through deep image analogy. *ACM Transactions on Graphics* **36**(4), 120:1–120:15 (2017). <https://doi.org/10.1145/3072959.3073683>
- [13] Litjens, G., et al.: A survey on deep learning in medical image analysis. *Medical Image Analysis* **42**, 60–88 (2017). <https://doi.org/10.1016/j.media.2017.07.005>
- [14] Liu, L., et al.: Advanced deep learning techniques for image style transfer: A survey. *Signal Processing: Image Communication* **78**, 465–470 (2019). <https://doi.org/10.1016/j.image.2019.08.006>

- [15] Nath, S., Devi, G.R.: Three-dimensional culture systems in cancer research: Focus on tumor spheroid model. *Pharmacology & Therapeutics* **163**, 94–108 (2016). <https://doi.org/10.1016/j.pharmthera.2016.03.013>
- [16] Paszke, A., Gross, S., Massa, F., et al.: Pytorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems* 32, pp. 8024–8035. Curran Associates, Inc. (2019)
- [17] Qin, X., Zhang, Z., Huang, C., et al.: U²-net: Going deeper with nested u-structure for salient object detection. *Pattern Recognition* **106**, 107404 (2020). <https://doi.org/10.1016/j.patcog.2020.107404>
- [18] Razavian, A.S., Azizpour, H., Sullivan, J., et al.: CNN features off-the-shelf: An astounding baseline for recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. pp. 512–519. CVPRW'14, IEEE (2014)
- [19] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Lecture Notes in Computer Science, vol. 9351, pp. 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28
- [20] Salimans, T., et al.: Improved techniques for training gans. In: *30th International Conference on Neural Information Processing Systems*. pp. 2234–2242. NIPS'16, Curran Associates Inc. (2016)
- [21] Simard, P., Steinkraus, D., Platt, J.C.: Best practices for convolutional neural networks applied to visual document analysis. In: *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR'03)*. vol. 2, pp. 958–964 (2003)
- [22] Sirinukunwattana, K., et al.: Gland segmentation in colon histology images: the glas challenge contest. *Medical Image Analysis* **35**, 489–502 (2017). <https://doi.org/10.1016/j.media.2016.08.008>
- [23] Wang, J., Sun, K., Cheng, T., et al.: Deep high-resolution representation learning for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020). <https://doi.org/10.1109/tpami.2020.2983686>
- [24] Yan, W., et al.: The domain shift problem of medical image segmentation and vendor-adaptation by unet-gan. In: *Medical Image Computing and Computer Assisted Intervention (MICCAI 2019)*. pp. 623–631. Springer International Publishing (2019). https://doi.org/10.1007/978-3-030-32245-8_69
- [25] Zhang, Y., et al.: Task driven generative modeling for unsupervised domain adaptation: Application to x-ray image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 599–607 (2018)
- [26] Zhu, J.Y., et al.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *IEEE International Conference on Computer Vision (ICCV 2017)*. pp. 2242–2251. IEEE (2017). <https://doi.org/10.1109/ICCV.2017.244>