

OTLesMix: Wasserstein Barycenter and Optimal Transport Map for Synthetic Lesion Generation with Diverse Shapes and Locations

Robin Trombetta^[0009–0001–7945–6953] and Carole Lartizien^[0000–0001–7594–4231]

Univ. Lyon, CNRS UMR 5220, Inserm U1294, INSA Lyon, UCBL, CREATIS, France
`robin.trombetta@creatis.insa-lyon.fr`

Abstract. The development of deep learning over the past decade has revolutionized medical imaging segmentation, allowing the extraction of precise descriptors from large volumes to characterize pathologies. Data augmentation is a technique widely regarded as a way to improve model training. It includes simple transformations like spatial operations or intensity modifications, but also more advanced synthesis techniques. Their goal is to generate new realistic samples from an existing dataset to diversify the images used during training. Among them, several propose different mixing strategies to combine real samples. However, one of their major shortcomings is to yield limited variability in terms of generated lesion shapes and locations. In this work, we introduce a novel image synthesis method, called OTLesMix, that leverages Wasserstein barycenter and optimal transport plan to generate realistic and diverse samples. We evaluated our method on three brain lesion segmentation tasks, on which it improves the Dice score compared to a model trained without synthetic data by 2.9 to 6.6 points, and outperforms state-of-the-art *mix-based* methods.

Keywords: Data Augmentation · Image Synthesis · Optimal Transport · Medical Image Segmentation

1 Introduction

Deep learning models for image segmentation have opened new perspectives for medical image analysis, allowing for a fine description of anatomical or functional structures. Supervised training remains the dominant and most performing training paradigm, but the cost of getting annotations can be prohibitive, especially for medical imaging applications. It requires expert knowledge and is time-consuming, especially for tasks like dense image segmentation which can take up to several hours for a single 3D imaging volume [4].

Data augmentation is a broadly used technique to improve the training of neural networks. It works by increasing the diversity of the training samples through the application of some transformations to the available data. Basic random image transformations are spatial and intensity transformations, or noise injection.

To improve further the diversity of generated samples, more advanced techniques leveraging generative models, like GANs or Diffusion Models, have also been widely studied over the past few years. Such methods yield state-of-the-art results in terms of realism of the generated samples. Conditioning the generation process on additional attributes, such as class, mask or text, also allows guiding the sampling process [5,14]. However, such generative models demand prior training based on diverse and numerous data, which can be a deceptive solution. If the available data is restricted, training a generative deep learning model may not be efficient for generating new realistic data.

To overcome the limitations of methods based on deep generative models, other approaches propose to mix available annotated data with simpler operations. Two early approaches falling into this category of *mix-based* methods are MixUp [12] and CutMix [11]. In the former, a new sample is produced by a random linear combination of two existing images and their associated ground truth. The latter consists in cutting a rectangle in a first image and pasting it in a second image to corrupt it. These two methods have been primarily designed for classification tasks.

Some works focus on designing data augmentation techniques more suited to segmentation tasks, and in particular for medical imaging applications. CarveMix [13] proposes to carve an image and its segmentation mask according to the location and geometry of the ground truth lesion, and inpaint the resulting region of interest onto another image. LesionMix [2] generates lesions using a marching method to augment existing images, allowing the creation of more diverse lesion shapes, with control over the synthetic lesion load. SelfMix [15] fuses several random lesion and non-lesion regions from the training dataset with computed weights to create new samples. These data augmentation techniques can produce unrealistic synthetic images, in particular when compared to samples generated with GANs or Diffusion Models, but recent works have shown that such *out-of-domain* images can still improve performance of deep segmentation models [3]. There is a tradeoff between the domain gap between real and synthetic images, and the diversity of generated samples.

As can be seen in the most recent research works in the field of medical imaging data augmentation, the main areas of improvement are the realism of generated images and the diversity of the shapes and locations of the masks. The new shapes are usually obtained via a simple random sampling process from existing shapes [2,15], which gives limited diversity. Similarly, the locations of the generated lesion masks are drawn from existing ones [13,2], sometimes with additional simple transformations, such as rotations or translations [15], which restrict their variety.

In this work, we propose to leverage Optimal Transport (OT) to produce synthetic samples with high diversity both in terms of lesion shapes and locations. Our method consists of two main steps. First, given two or more lesion images from the training dataset with their respective masks, referred to as base images in the following, we compute the Wasserstein barycenter of the masks to generate a new lesion shape and location. The obtained mask can

be located anywhere between the two original masks, and its shape is the interpolation between the base masks. Since the generated mask does not fully overlap any anomalous region in the base images, we then compute the optimal transport plan between the generated mask and one of the two base masks and use it to derive the intensities of the generated image inside the lesion mask. The synthesized lesion is inpainted onto one of the two base images, or onto any other image from the training dataset. We evaluate our approach, dubbed OTLesMix, on three brain segmentation tasks, on which it is shown to outperform state-of-the-art *mix-based* data augmentation techniques. The code that implements our method and reproduces the experiments of this study is available at <https://github.com/robintrmbtt/otlesmix>.

2 Method

Let us consider a dataset $\mathcal{D} = \{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i=1}^n$, composed of pairs of images $\mathbf{X}_i$ and associated pixel-level ground truth annotations $\mathbf{Y}_i$, where $\mathbf{Y}_i$ can take values in $\mathbb{N}$ in the case of a multi-class segmentation task. The goal is to generate a synthetic dataset $\mathcal{D}_{\text{synth}} = \{(\tilde{\mathbf{X}}_i, \tilde{\mathbf{Y}}_i)\}_{i=1}^m = f(\mathcal{D})$ from the base dataset $\mathcal{D}$. We hope that our data generation process f is such that a model trained on $\mathcal{D} \cup \mathcal{D}_{\text{synth}}$ achieves better performance on a test set than a baseline model trained only on $\mathcal{D}$.

2.1 Overview of the method

Our method, illustrated in Figure 1, is composed of two main steps, described in more detail in Sections 2.2 and 2.3. The first step consists in combining two ground truth masks to create a new lesion mask $\tilde{\mathbf{Y}}_{\text{les}}$. We randomly draw two base masks $\mathbf{Y}_i$ and $\mathbf{Y}_j$ from the training dataset and compute their Wasserstein barycenter to obtain $\tilde{\mathbf{Y}}_{\text{les}}$. As shown in Figure 1, this step allows interpolating between two existing shapes to generate original and realistic lesion shapes. Moreover, the location of the new lesion falls between the two base masks.

Once we have obtained a new lesion $\tilde{\mathbf{Y}}_{\text{les}}$, the second step consists in computing both the intensities and the labels of its constituting pixels. To do so, we *transport* every pixel of the lesion mask $\tilde{\mathbf{Y}}_{\text{les}}$ onto either the lesion of $\mathbf{Y}_i$ or the lesion of $\mathbf{Y}_j$. For each pixel of the lesion mask $\tilde{\mathbf{Y}}_{\text{les}}$, we look at where it is mapped by the transport plane, and interpolate the image intensity values and the label accordingly. Finally, we can inpaint the synthetic lesion onto $(\mathbf{X}_i, \mathbf{Y}_i)$, $(\mathbf{X}_j, \mathbf{Y}_j)$ or any other sample from the training dataset to generate a new sample $(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$.

2.2 Lesion mask generation with Wasserstein barycenter

Let us take two samples $(\mathbf{X}_i, \mathbf{Y}_i)$ and $(\mathbf{X}_j, \mathbf{Y}_j)$ from the training dataset. We denote by $\alpha_i \stackrel{\text{def}}{=} (x_{i,k}) \in \mathbb{R}^{N_i \times D}$ and $\alpha_j \stackrel{\text{def}}{=} (x_{j,k}) \in \mathbb{R}^{N_j \times D}$ the coordinates of

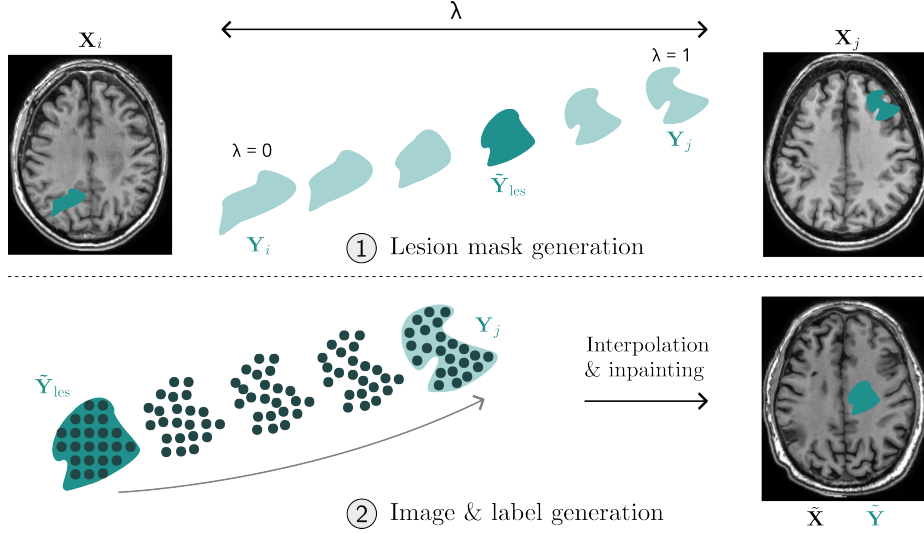


Fig. 1. Illustration of the sample generation process with OTLesMix.

the N_i (resp. N_j) points of dimension D where the respective label maps $\mathbf{Y}_i$ (resp. $\mathbf{Y}_j$) are not null. To interpolate between the two lesion shapes of $\mathbf{Y}_i$ and $\mathbf{Y}_j$, we make use of the theory of optimal transport, which provides a means to compare two probabilistic distributions. More formally, if we consider the distance, or *cost*, between two points $\mathbf{C} : (x, y) \in \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbf{C}(x, y) = \frac{1}{2} \|x - y\|^2 \in \mathbb{R}_{\geq 0}$, Kantorovich's formulation of the optimal transport problem consists in finding the transport plan $\pi \in \mathbb{R}_{\geq 0}^{N_i \times N_j}$ that minimizes the loss:

$$\text{OT}(\alpha, \beta) \stackrel{\text{def}}{=} \min_{\pi \in \mathbb{R}_{\geq 0}^{N_i \times N_j}} \sum_{m=1}^{N_i} \sum_{n=1}^{N_j} \pi_{m,n} \mathbf{C}(x_{i,m}, x_{j,n}) \quad (1)$$

under the marginal constraints $\forall i, j, \sum_{n=1}^{N_j} \pi_{i,n} = \frac{1}{N_i}$ and $\sum_{m=1}^{N_i} \pi_{m,j} = \frac{1}{N_j}$. π represents how the *mass* of the points α_i is distributed among the points α_j .

We denote by $\beta \stackrel{\text{def}}{=} (y_k) \in \mathbb{R}^{N \times D}$ the coordinates of the points of the new lesion mask, and consider a weighting parameter λ randomly drawn in the distribution $\mathcal{U}(0, 1)$. We optimize the coordinates of these points such that they are the Wasserstein barycenter of α_i and α_j , *i.e.* we minimize the loss:

$$\mathbf{B}(\lambda, \alpha_i, \alpha_j, \beta) \stackrel{\text{def}}{=} \lambda \text{OT}(\alpha_i, \beta) + (1 - \lambda) \text{OT}(\alpha_j, \beta) \quad (2)$$

When the optimal $\{y_k\}_{k=1}^N$ are found, we simply round their coordinates to the closest integer to obtain the final binary lesion mask $\hat{\mathbf{Y}}_{\text{les}}$.

In practice, solving Equation (1) does not scale well with the number of points $N_i + N_j$. We thus consider a regularized version of the OT cost, called *de-biased Sinkhorn divergence* [6]. This formulation can be solved fast using the

Sinkhorn algorithm, with a linear instead of quadratic memory footprint, and for measures with up to $N \sim 1e^6$ points in a matter of seconds on modern GPUs.

2.3 Image and label generation with optimal transport map

Given the coordinates (y_k) of the points of the generated lesion mask $\tilde{\mathbf{Y}}_{\text{les}}$, we now want to *transport* these points to either the lesion of $\mathbf{Y}_j$ – or similarly $\mathbf{Y}_i$. More formally, we aim to find the optimal transport map $T : x \in \mathbb{R}^D \rightarrow T(x) \in \mathbb{R}^D$ which minimizes $\sum_{k=1}^N \sum_{n=1}^{N_j} \mathbf{C}(T(y_k), x_{j,n})$.

To obtain the intensities of the pixel of the synthetic lesion, for each point of the mask $\tilde{\mathbf{Y}}_{\text{les}}$ with coordinates y_k , we compute the linear interpolation of the point $T(y_k)$ based on the intensities of the image $\mathbf{X}_j$. If the segmentation task is multi-class, we also compute the ground truth label map of the synthetic by nearest-neighbour interpolation of the point $T(y_k)$ based on the ground truth label map $\mathbf{Y}_j$. Finally, the obtained lesion can be inpainted onto $(\mathbf{X}_i, \mathbf{Y}_i)$, $(\mathbf{X}_j, \mathbf{Y}_j)$ or any other sample from the training dataset to generate a new sample $(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$.

3 Experiments

3.1 Experimental Setup

To evaluate the effectiveness of OTLesMix as a data augmentation technique to improve the training of a deep learning model, we apply the generation procedure described in Section 2 to increase the training datasets of several CNN-based segmentation tasks, both in 2D and 3D. For a given training dataset containing images and associated ground truth label maps, we repeatedly apply OTLesMix to pairs of training samples to generate synthetic data and increase the number of training samples. Then, two identical neural networks are trained, the first on real data only, the second on both real and synthetic data. Both models are evaluated solely on real-world data to assess the impact of using partially synthesized data during training and determine whether or not this improves the downstream segmentation performance.

Datasets We use three publicly available datasets for the segmentation of brain pathologies:

- **BraTS 2020** [10,1] is a multi-institutional database of pre-operative MRI scans of glioblastoma and low-grade glioma. It contains images from 369 patients with four modalities: native T1-weighted, post-contrast T1-weighted, T2-weighted and FLAIR, and annotated by expert radiologists with three labels: enhancing tumor, the peritumoral edema, and the necrotic and non-enhancing tumor core. The volumes are provided already pre-processed, namely with co-registration, skull-stripping and resampling to isotropic resolution of 1mm^3 . We extract 200 2D slices from 37 patients for the training, 100 with tumors and 100 without, and 10 000 slices with tumors from the

remaining patients for the evaluation. On this dataset, we augment the training dataset by generating 1000 synthetic images. We create synthetic lesions from pairs of slices with tumors and inpaint them on non-tumorous slices.

- **ATLAS v2.0** [9]. We leverage the 655 openly available T1w volumes of this manually annotated stroke neuroimaging dataset. The images are defaced, registered on the standard atlas MNI, and their intensities are normalized. In addition, we crop the volumes to keep only the 60 central slices, which gives us 3D images of size $197 \times 233 \times 60$ mm³. One-hundred 3D images are used for the training and the rest are reserved for evaluating the models. We synthesize 200 volumes on this dataset, by generating 3D synthetic lesions from pairs of 3D base volumes and inpainting them on either one or the other base volumes.
- **ISLES 2022** [7] is a challenge for Ischemic Stroke Lesion Segmentation, comprising 250 publicly available MRI cases of acute to sub-acute stroke lesions. Each patient has at least a 3D FLAIR and a Diffusion-Weighted Image (DWI) exam, but we only keep the latter modality in our experiments. The volumes are resampled to 2 mm³, skull-stripped and registered to MNI-152 space. We use 50 volumes for the training, 200 for the evaluation, and generate 100 additional synthetic images, following the same protocol as for the ATLAS dataset.

Compared methods OTLesMix is compared to related state-of-the-art methods, namely MixUp [12], CutMix [11] and CarveMix [13]. We generate the same number of synthetic images and use the same random seed for all methods. Each training dataset, endowed with the synthetic data, is used within the framework nnUNet [8] for fair comparison. nnUNet was chosen as one of the highest-performing state-of-the-art versatile backbone for medical image segmentation. In addition to synthetic data, all methods also use *traditional data augmentations* (TDA) – rotation, scaling, blurring, etc. –, which are included in the framework of nnUNet. As a baseline, we also report the performance of nnUNet trained without any synthetic data, *i.e.* only with TDA.

Implementation details The Python package GeomLoss¹ is used to implement our method, in particular to compute the Wasserstein barycenter and the optimal transport map described in Section 2. The entropic penalty ϵ in Equation 1 is set to 0.01. To smooth the contours of the synthetic lesions, we apply a Gaussian filter with an isotropic kernel size of 1 mm. Moreover, when a synthetic lesion is inpainted on an image, we remove the parts of the lesion that intersect existing lesions or that land on unrealistic locations, such as the skull or the cerebrospinal fluid. Those post-processing steps are applied to all methods. The nnUNet models are trained for 1000 epochs on BraTS and 500 on ATLAS and ISLES. By default, the batch size is capped by the number of training samples, so it is smaller for the model trained without synthetic data. For fair comparison, we manually set the same batch size for all methods.

¹ <https://github.com/jeanfeydy/geomloss>

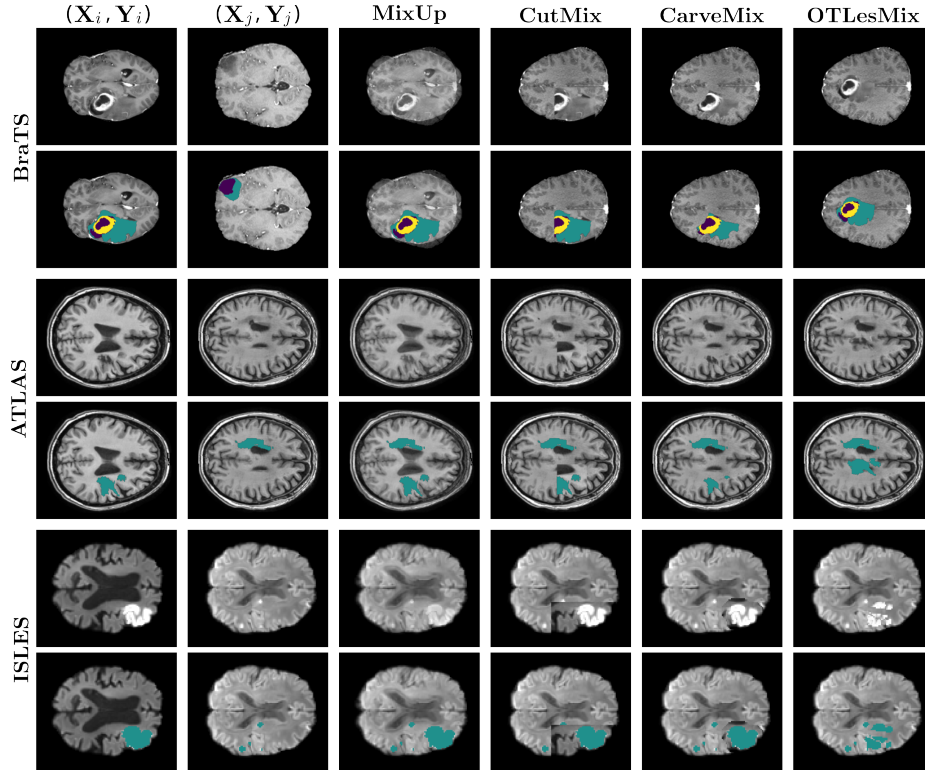


Fig. 2. Examples of generated samples from all methods. For each database, the first row shows the image and the second the corresponding ground truth label map. The first two columns represent the two base images with their corresponding original mask, while the other columns depict the resulting synthetic lesions inpainted on a non-tumorous 2D slice for BraTS and on either one or the other base 3D images for ATLAS and ISLES. Note that for BraTS, the synthetic tumors are inpainted on a third non-tumorous image, which is not shown here, and only OTLesMix uses $\mathbf{Y}_j$.

Evaluation To evaluate the compared methods, we report the Dice Similarity Coefficient (DSC) for all datasets. On BraTS 2020, we follow the standard practice and report the DSC of 3 classes: Whole Tumor (WT), Tumor Core (TC) and Enhanced Core (EC). The statistical significance of the difference between the best-performing method and the others is assessed with a pairwise Wilcoxon signed-rank test.

3.2 Results

Table 1 presents the quantitative results on the three datasets and Figure 2 shows examples of generated samples for all methods.

On every dataset, OTLesMix outperforms all concurrent methods, increasing DSC by 0.8 to 1.8 points over the second-best method. Notably, the difference

Table 1. Average and standard deviation of the DSC (%) on the three test datasets. The best model on each task is highlighted in **bold**. Asterisks indicate the level of statistical significance of the difference in DSC between a given method and the best performing one with a pairwise Wilcoxon signed-rank test (*: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$).

	BraTS			ATLAS	ISLES
	WT	TC	ET		
TDA	$77.0 \pm 30.1^{***}$	$72.5 \pm 29.6^{***}$	$66.4 \pm 30.9^{***}$	$45.8 \pm 33.8^{***}$	$68.8 \pm 22.4^{***}$
MixUp	$80.6 \pm 24.5^{***}$	$75.2 \pm 24.7^{***}$	$68.2 \pm 27.6^{***}$	$43.7 \pm 33.0^{***}$	$66.3 \pm 23.9^{***}$
CutMix	$81.9 \pm 22.5^{***}$	$76.7 \pm 23.1^{***}$	$70.0 \pm 26.7^{***}$	$48.1 \pm 33.2^{**}$	$69.8 \pm 22.6^{***}$
CarveMix	$82.5 \pm 22.3^{***}$	$77.2 \pm 23.1^{***}$	$69.6 \pm 26.9^{***}$	$48.5 \pm 32.5^*$	$70.8 \pm 21.4^{**}$
OTLesMix	83.3 ± 22.8	78.4 ± 23.2	71.4 ± 26.5	49.9 ± 32.9	71.7 ± 20.8

in DSC between OTLesmix and the other generative methods is statistically significant with a p-value of at least 0.05, the greatest p-value being of 0.012 for CarveMix on the ATLAS database. Compared to the standard nnUNet without data synthesis (TDA), our approach improves the DSC on the test set by between 2.9 points on ISLES and 6.6 points on the class WT of the BraTS dataset.

Qualitatively, Figure 2 illustrates how OTLesMix allows generating lesion masks with more diversity in shapes and locations than other state-of-the-art *mix-based* methods. The example on the BraTS dataset – first two rows – shows that our method works on a multi-class segmentation setup. On ATLAS and ISLES, we can see that the shape interpolation computed with Wasserstein barycenter is consistent when the original lesion masks $\mathbf{Y}_i$ and $\mathbf{Y}_j$ contain multiple connected components.

4 Discussion and Conclusion

In this work, we propose to leverage Wasserstein barycenter and optimal transport map to generate synthetic lesions and enhance the training of deep learning segmentation models. Compared to related methods that also mix real samples, our method yields more varied shapes and locations. Although it achieves better segmentation performance than the compared methods, OTLesMix still suffers from several limitations. Typically, generating one sample takes around 1 second for 2D images (BraTS) and 1-3 minutes on 3D datasets (ATLAS and ISLES), depending on the resolution and size of the images. Moreover, synthetic lesions are obtained simply by linear interpolation of the pixel intensities of real lesions. This approach can result in unrealistic inpainted lesions, in particular when there is a domain gap – different scanner or acquisition protocol for instance – between the two original base images $\mathbf{X}_i$ and $\mathbf{X}_j$. One perspective of this work would be to improve the quality of the synthetic samples by combining the mask

generation procedure of OTLesMix with mask-conditioned deep-learning-based generation models.

Acknowledgments. This work was funded by the *Agence Nationale de la Recherche* (ANR) under projects ANR-24-CE45-4399 (SEIZURE), ANR-11-INBS-0006 (FLI) and 11-LABX-0063 (Labex PRIMES). This work was granted access to the HPC resources of IDRIS under the allocation 2026-AD011014900R2 made by GENCI.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J.S., Freymann, J.B., Farahani, K., Davatzikos, C.: Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data* **4**(1), 170117 (2017)
2. Basaran, B.D., Zhang, W., Qiao, M., Kainz, B., Matthews, P.M., Bai, W.: Lesionmix: A lesion-level data augmentation method for medical image segmentation. In: 0002, Y.X., 0042, C.C., 0012, C.C., Zuo, L., 0003, Y.L. (eds.) *Data Augmentation, Labelling, and Imperfections - Third MICCAI Workshop, DALI 2023, Held in Conjunction with MICCAI 2023, Vancouver, BC, Canada, October 12, 2023, Proceedings. Lecture Notes in Computer Science*, vol. 14379, pp. 73–83. Springer (2023). https://doi.org/10.1007/978-3-031-58171-7_8, https://doi.org/10.1007/978-3-031-58171-7_8
3. Cubuk, E.D., Dyer, E.S., Lopes, R.G., Smullin, S.: Tradeoffs in data augmentation: An empirical study. In: *ICLR* (2021), <https://openreview.net/forum?id=ZcKPWuhG6wy>
4. Dang, V.N., Galati, F., Cortese, R., Di Giacomo, G., Marconetto, V., Mathur, P., Lekadir, K., Lorenzi, M., Prados, F., Zuluaga, M.A.: Vessel-captcha: An efficient learning framework for vessel annotation and segmentation. *Medical Image Analysis* **75**, 102263 (2022)
5. Dorjsembe, Z., Pao, H.K., Odonchimed, S., Xiao, F.: Conditional diffusion models for semantic 3d brain mri synthesis. *IEEE Journal of Biomedical and Health Informatics* **28**(7), 4084–4093 (2024)
6. Feydy, J., Séjourné, T., Vialard, F.X., Amari, S.i., Trounev, A., Peyré, G.: Interpolating between optimal transport and mmd using sinkhorn divergences. In: *The 22nd international conference on artificial intelligence and statistics*. pp. 2681–2690. PMLR (2019)
7. Hernandez Petzsche, M.R., De La Rosa, E., Hanning, U., Wiest, R., Valenzuela, W., Reyes, M., Meyer, M., Liew, S.L., Kofler, F., Ezhov, I., et al.: Isles 2022: A multi-center magnetic resonance imaging stroke lesion segmentation dataset. *Scientific data* **9**(1), 762 (2022)
8. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods* **18**(2), 203–211 (2021)
9. Liew, S.L., Lo, B.P., Donnelly, M.R., Zavaliangos-Petropulu, A., Jeong, J.N., Barisano, G., Hutton, A., Simon, J.P., Julian, J.M., Suri, A., et al.: A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms. *Scientific data* **9**(1), 320 (2022)

10. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging* **34**(10), 1993–2024 (2014)
11. Yun, S., Han, D., Chun, S., Oh, S.J., Yoo, Y., Choe, J.: Cutmix: Regularization strategy to train strong classifiers with localizable features. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 6022–6031 (2019). <https://doi.org/10.1109/ICCV.2019.00612>
12. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. In: International Conference on Learning Representations (2018)
13. Zhang, X., Liu, C., Ou, N., Zeng, X., Zhuo, Z., Duan, Y., Xiong, X., Yu, Y., Liu, Z., Liu, Y., Ye, C.: Carvemix: A simple data augmentation method for brain lesion segmentation. *NeuroImage* **271**, 120041 (2023). <https://doi.org/https://doi.org/10.1016/j.neuroimage.2023.120041>, <https://www.sciencedirect.com/science/article/pii/S1053811923001878>
14. Zhang, Z., Yao, L., Wang, B., Jha, D., Durak, G., Keles, E., Medetalibeyoglu, A., Bagci, U.: Diffboost: Enhancing medical image segmentation via text-guided diffusion model. *IEEE Transactions on Medical Imaging* **44**(9), 3670–3682 (2024)
15. Zhu, Q., Wang, Y., Yin, L., Yang, J., Liao, F., Li, S.: Selfmix: A self-adaptive data augmentation method for lesion segmentation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. pp. 683–692. Springer Nature Switzerland, Cham (2022)