UNIVERSITY OF COPENHAGEN

Københavns Universitet

Local Gamma Augmentation for Ischemic Stroke Lesion Segmentation on MRI

Middleton, Jon; Bauer, Marko; Johansen, Jacob; Perslev, Mathias; Sheng, Kaining; Ingala, Silvia; Nielsen, Mads; Pai, Akshay

Published in:
Proceedings of the 5th Northern Lights Deep Learning Conference ({NLDL})

Publication date:
2024

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Middleton, J., Bauer, M., Johansen, J., Perslev, M., Sheng, K., Ingala, S., Nielsen, M., & Pai, A. (2024). Local Gamma Augmentation for Ischemic Stroke Lesion Segmentation on MRI. In *Proceedings of the 5th Northern Lights Deep Learning Conference ({NLDL})* (pp. 158-164). PMLR. https://proceedings.mlr.press/v233/

Download date: 11. Oct. 2026

Local Gamma Augmentation for Ischemic Stroke Lesion Segmentation on MRI

Jon Middleton[†*1,3], Marko Bauer[†3], Jacob Johansen[1,3], Mathias Perslev[3], Kaining Sheng[1,3,4], Silvia Ingala[3,4], Mads Nielsen[1,2,3], and Akshay Pai[3,4]

[1]Department of Computer Science, University of Copenhagen
[2]Pioneer Centre for Artificial Intelligence, University of Copenhagen
[3]Cerebriu A/S
[4]Department of Diagnostic Radiology, Rigshospitalet, Copenhagen, Denmark

Abstract

The identification and localisation of pathological tissues in medical images continues to command much attention among deep learning practitioners. When trained on abundant datasets, deep neural networks can match or exceed human performance. However, the scarcity of annotated data complicates the training of these models. Data augmentation techniques can compensate for a lack of training samples. However, many commonly used augmentation methods can fail to provide meaningful samples during model fitting. We present local gamma augmentation, a technique for introducing new instances of intensities in pathological tissues. We leverage local gamma augmentation to compensate for a bias in intensities corresponding to ischemic stroke lesions in human brain MRIs. On three datasets, we show how local gamma augmentation can improve the image-level sensitivity of a deep neural network tasked with ischemic stroke lesion segmentation on magnetic resonance images.

1 Introduction

The potential of deep neural networks to identify and localise pathological tissues in medical images continues to motivate research into novel training methods. In theory, deep architectures can achieve arbitrarily high levels of performance. In practice, however, these algorithms are trained on insufficient quantities of data. Overfitting arises as a result, and the full potential of these algorithms remains unrealised.

Traditional approaches to the data scarcity problem use data augmentation to expand the size of the dataset. A typical setup selects a collection of data-agnostic operations which come from parameterized families of spatial and intensity transformations. Examples include affine transformations, elastic transformations, global gamma transformations, additive and multiplicative noise transformations, blurring transformations, and brightness transformations. An unlimited collection of novel samples can then be obtained by fixing probability distributions on the parameter spaces and sampling augmentation parameters from them.

Though a dataset can be enlarged in this fashion, the approach may not provide relevant samples for image models tasked with classifying or segmenting regions of interest. Moreover, the augmentations themselves are not guaranteed to have an appreciable impact on model training. For example, the literature often claims that elastic deformations improve deep neural network performance [1–4]. Nonetheless, these transformations do not appear in all well-known augmentation pipelines (for example, [5]), and some research indicates that these transformations can impair a network's performance [6, 7]. The latter examples fit into the claim in [8] that the effectiveness of a data augmentation pipeline depends crucially on the content of a dataset.

Thus, for models tasked with identifying and localising pathological tissues in medical images, practitioners are developing augmentation techniques that target these tissues. Recent work [4, 6, 9–11] has focused on local augmentations that expose deep networks to additional examples by isolating and altering regions of interest.

We contribute a simple, targeted method of local data augmentation for a pathology segmentation model and apply this method to the problem of ischemic stroke lesion segmentation on multi-modal magnetic resonance images. The method exploits both the single-parameter family of gamma transformations and ischemic stroke lesion segmentation maps to mitigate an intensity bias in the training data. We compare local gamma augmentation against a baseline without it, and we show an improvement in image-level sensitivity as a result.

2 Related Work

Gamma transformations frequently appear in probabilistic data augmentation pipelines for deep segmentation networks. The training pipeline for nnU-Net

*Corresponding Author: jami@di.ku.dk
†Equal contribution.

Proceedings of the 5th Northern Lights Deep Learning Conference (NLDL), PMLR 233, 2024.
© 2024 Jon Middleton[†], Marko Bauer, Jacob Johansen, Mathias Perslev, Kaining Sheng, Silvia Ingala, Mads Nielsen, & Akshay Pai. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).

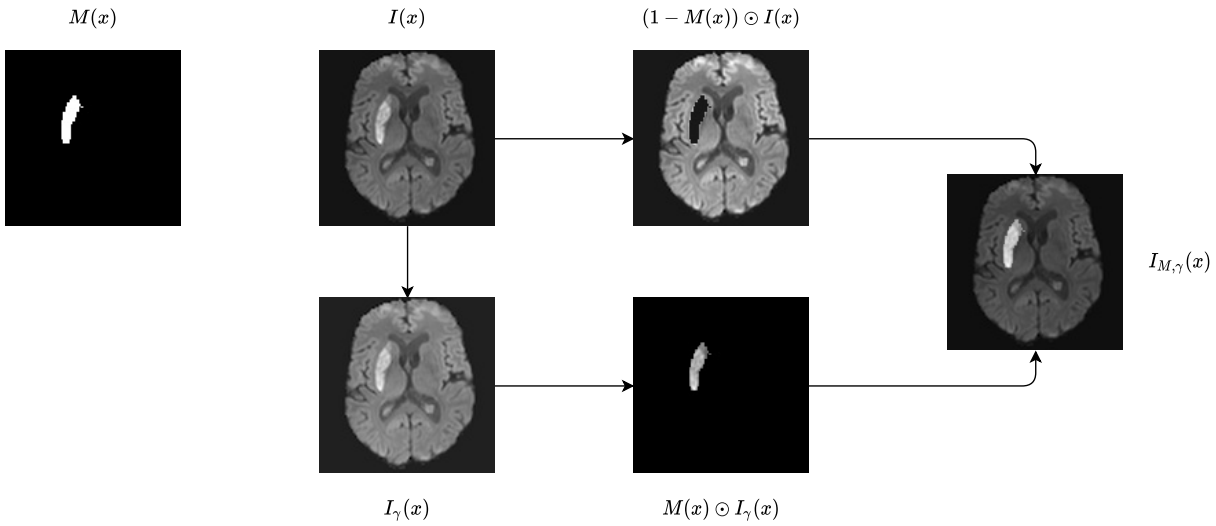

Figure 1. Depiction of local gamma augmentation. An image $I(x)$ and its gamma transformation $I_\gamma(x)$ are multiplied pointwise with a segmentation map and combined to produce an image $I_{M,\gamma}(x)$ featuring an augmentation of a pathology.

[5] uniformly samples parameters for both gamma compression and gamma expansion, with a bias toward gamma expansion parameters. A similar gamma augmentation scheme for a tumour segmentation network appears in [3], with equal weighting of gamma compression and gamma expansion. In [12], random channel-wise gamma augmentations are applied to RGB images of retinal vessels, with a strong bias toward gamma expansion. In [13], random gamma compressions augment the saturation and value channels of HSV images of retinal vessels. In all cases, the gamma transformations are applied to the whole image and are uninformed by any features of the dataset.

Other applications for global gamma augmentations include generative modeling, out-of-distribution benchmarking, and unsupervised domain adaptation. In [14], log-normally distributed gamma transformations assist in the construction of synthetic MRI training data for a segmentation network. In [15], gamma augmentations help to produce synthetic training data for a contrast-agnostic MRI-to-CT generation network. In [16], the intensity distribution shifts induced by gamma transformations are used to assess the robustness of image segmentation networks. Moreover, the authors observe that gamma compression impairs a segmentation network's ability to localise white-matter hyperintensities. In [17], a GAN framework is trained to simulate global intensity distribution shifts using instance-specific gamma transformations.

Local image augmentations arise in many training schemes. CarveMix [11] augments MRI training data by inserting pathologies from one image into a second image. Moreover, for brain tumours, the authors use deformable image registration to generate plausible intracranial mass effects. CarveMix is also used in [4] to augment MRIs featuring multiple sclerosis lesions. Mask-based augmentations which corrupt image data, such as random erasing [18] and CutOut [19], have been explored in medical image analysis scenarios and show mixed results. The authors in [6] show that random erasing, which replaces random image patches with independent and identically distributed point intensities, impairs the performance of a melanoma classification model. In [9], a white matter tract segmentation algorithm is fine-tuned on MRIs in which random subsets of white matter tracts are cut out. In [10], the authors observe that masked image modelling can improve the performance of a prostate lesion classifier.

3 Methods

3.1 Gamma transformations

Let $I(x)$ be an image with intensity values in the range $[0, 1]$. A gamma transformation is defined by a choice of $\gamma > 0$ via $I(x) \to I(x)^\gamma$. Thus a gamma transformation yields a pointwise, non-linear transformation of I.

Gamma transformations generalize to images with arbitrary values by conjugation with min-max normalization:

$$I(x) \longrightarrow (m_2 - m_1) \left(\frac{I(x) - m_1}{m_2 - m_1} \right)^\gamma + m_1, \quad (1)$$

where $m_1 = \min_x I(x)$ and $m_2 = \max_x I(x)$.

A choice of γ between 0 and 1 gives a gamma transformation which increases intensities and is called a gamma compression. For γ greater than 1,

the defined transformation decreases intensities and is called gamma expansion.

3.2 Local gamma augmentations

To adapt gamma transformations to a local region of interest, we alter the min-max normalization in equation 1 as follows. Given a binary function $M(x)$, let $S = \{x : M(x) = 1\}$ be the set of foreground points of M. Set $\tilde{m}_1 = \min_x S$ and $\tilde{m}_2 = \max_x S$. Then the map

$$I_\gamma(x) = (\tilde{m}_2 - \tilde{m}_1)\left(\frac{I(x) - \tilde{m}_1}{\tilde{m}_2 - \tilde{m}_1}\right)^\gamma + \tilde{m}_1 \quad (2)$$

applies a non-linear transformation which is normalized relative to the foreground S.

We then complete the local gamma transformation by using $M(x)$ as a mask to mix intensity-augmented pathological tissue in S with normal tissue outside S:

$$I_{M,\gamma}(x) = (1 - M(x)) \odot I(x) + M(x) \odot I_\gamma(x), \quad (3)$$

where $\odot$ denotes pointwise multiplication. Local gamma augmentation is depicted in Figure 1, with a segmentation map indicating diseased brain tissue.

When used to augment image data, gamma transformations are randomly chosen by sampling γ from a probability distribution. Typical choices include the log-normal distribution for a weakly constrained value of γ or a Beta distribution supported in a fixed interval. Rather than sampling gamma parameters as in [5], in which γ is drawn from the uniform distribution on the interval $[0.7, 1.5]$, we instead sample from a mixture of uniform distributions:

$$\gamma \sim \frac{1}{2}U(0.7, 1.0) + \frac{1}{2}U(1.0, 1.5). \quad (4)$$

Thus we effectively partition this interval into a half-open gamma compression interval $[0.7, 1.0)$ and a closed gamma expansion interval $[1.0, 1.5]$. During each choice of γ, one of the two intervals is selected with equal probability.

Figure 2 depicts various local and global gamma transformations applied to a diffusion-weighted image from the ISLES-2022 dataset. In this image, one can see a clear increase in contrast between healthy brain tissue and diseased brain tissue for $\gamma < 1$, whereas this contrast is decreased for $\gamma > 1$. This differs from global gamma transformations, in which the local contrast between healthy and pathological tissue remains.

4 Data

4.1 Training Data

The models were trained on a multi-site in-house brain MRI dataset consisting of 1308 studies, comprising of 420 ischemic strokes, 265 tumors and 234 hemorrhages, as well as 530 normal studies without abnormalities. Each study consisted of three MRI sequences: fluid-attenuated inversion recovery (FLAIR), diffusion-weighted imaging (b0, b1000, and ADC), and either susceptibility-weighted imaging (SWI) or T2* gradient echo (T2*GRE). These sequences were annotated by an in-house medical team and verified by a certified neuroradiologist.

Medical experts classify presentations of ischemic stroke on MRI according to their age [20]. The various classes include *hyperacute* ischemic stroke, which has the shortest time window from stroke onset. When imaged sufficiently early, hyperacute ischemic stroke lesions appear hyperintense in a DWI sequence and isointense in a FLAIR sequence [21]. Notably, the MRIs in our training set underrepresent hyperacute ischemic strokes, as shown in Table 1. This yields a bias in ischemic stroke signal intensities between the DWI and FLAIR sequences.

In addition, we include two publicly available datasets for training. These include the Multimodal Brain Tumor Segmentation (BraTS) [22–24] Challenge 2019 and Ischemic Stroke Lesion Segmentation (ISLES) Challenge 2022 datasets [25, 26].

Pathology	Type	Count
Ischemic Stroke	Acute/Subacute	342
	Hyperacute	15
	Hemorrhagic	63
Hemorrhage	Intra-Axial	137
	Extra-Axial	97
Tumor	Intra-Axial	95
	Extra-Axial	170
Infection	Intra-Axial	27
	Extra-Axial	2

Table 1. Representation of tissue pathologies in training data. Of note is the relative under-representation of hyperacute infarcts.

All sequences were acquired in different resolutions and acquisition planes. Accordingly, as part of the pre-processing step, DWI sequences were resampled to isotropic resolution, while FLAIR and SWI/T2* GRE sequences were resampled and co-registered to DWI space.

4.2 Test Data

Our methods were evaluated on three independent private datasets, yielding a train-test split that generally arises in a real-world scenario, with a mixture of public and private data for the training set and private data entirely sourced from external sites for testing. The test data consist of DWI (b0 and b1000), FLAIR, and ADC sequences for each patient. At training time, the test data consisted of image-level labels only, rather than the pointwise

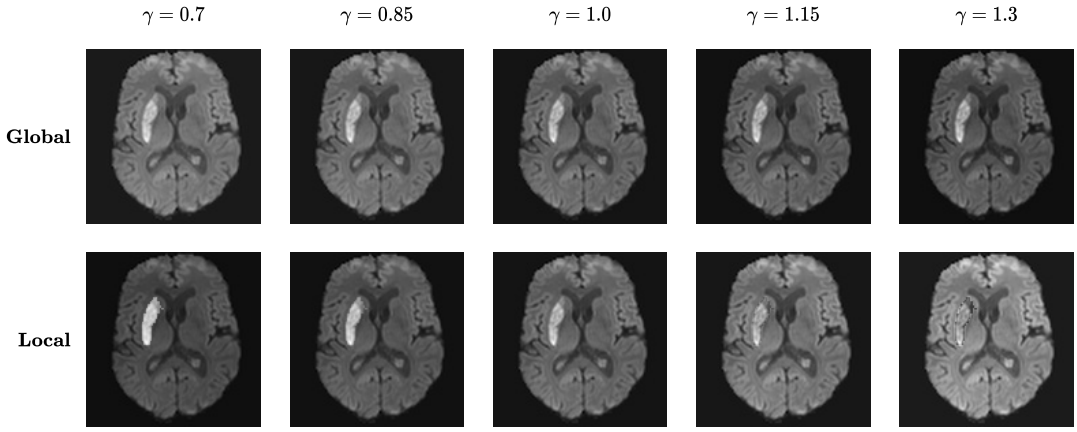

Figure 2. Gamma transformations of a diffusion weighted image from the ISLES-2022 dataset. The center column shows the original image. The top row contains globally gamma compressed images to the left and globally gamma expanded images to the right. The bottom row consists of the same gamma transformations restricted to the location of an acute ischemic stroke.

	MVS		**RH**		**WUS**	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
U-Net	0.781	0.928	0.757	0.899	0.566	N/A
U-Net+LG	0.844	0.910	0.814	0.895	0.736	N/A

Table 2. Image-level sensitivity and specificity for ischemic stroke detection across three MRI datasets for a baseline U-Net versus a U-Net trained with local gamma augmentation. Specificity for the WUS dataset is non-applicable since all samples in the dataset contain ischemic strokes.

annotations that are customarily supplied to a segmentation model. The three private datasets are elaborated below:

MVS. A multi-site dataset consisting of 387 studies obtained with Siemens scanners, comprising of 139 ischemic strokes, 93 tumors, 76 hemorrhages, as well as 98 studies with no abnormalities, which were annotated on image level by two neuroradiologists.

RH. A consecutively sampled dataset of 262 acquired brain MRI scans in a routine setting from Rigshospitalet (RH) in Copenhagen, Denmark. The dataset has further been consecutively enriched with positive findings of ischemic strokes (hyperacute, acute, and subacute), brain hemorrhage (intra-axial and extra-axial) and brain tumors (intra-axial and extra-axial) from three other hospital sites in the Capital Region of Denmark. This resulted in a multi-site cohort consisting of 487 scans including 96 ischemic strokes, 80 tumors, 76 hemorrhages, as well as 162 completely normal scans. Furthermore, 73 scans have significant findings that are not attributed to ischemic stroke, hemorrhage, or tumors, e.g., inflammatory disease, pineal cysts, aneurysms, and cavernous malformations. The pathology type of the scans in the dataset was classified by a medical doctor with 3 years' experience in brain imaging using original radiological report impressions as reference.

WUS. A collection of brain MRI scans featuring 51 patients who awoke with symptoms of ischemic stroke which were not present before falling asleep. This *wake-up stroke* dataset contains two distinct consecutively sampled cohorts, assembled using distinct stroke protocols, The first cohort contains 27 adult patients, and the second cohort consists of 24 adult patients. Both cohorts feature visible stroke lesions on the DWI sequence for all patients, as patients without visible stroke lesions were excluded. Of the 51 patients, 16 were determined to have a mismatch in ischemic stroke signal intensities between the DWI sequence and the FLAIR sequence. All studies were annotated on an image level according to patient reports.

5 Experiments and Results

5.1 Experimental Setup

All experiments used one 40GB Nvidia A100 Tensor Core GPU for model training. The model architecture largely follows the implementation in [27]. The architecture has a depth of five and incorporates padded convolution, instance normalization, and ReLU activations.

Models were trained for 500 epochs with a batch size of 2 and 4-step gradient accumulation on randomly extracted $128 \times 128 \times 128$ patches, with deep supervision, combining cross-entropy and general-

ized dice loss, and using an Adam optimizer with an initial learning rate of 0.0001 and a polynomial-rate schedule.

We trained two U-Net models: a baseline U-Net trained without local gamma augmentation, and a U-Net trained with local gamma augmentations applied to DWI sequences, which we call U-Net+LG. Augmentations for *both* models include: Rician noise, Gaussian noise, Gaussian blurring, contrast transformations, brightness transformations, global gamma correction, bias field correction, histogram equalization, elastic deformation, rotation, scaling, mirroring, random channel shifts and ghosting.

Of note is that a local gamma augmentation scheme as described in Section 3.2 does not account for images without a pathology. Such images are paired with a segmentation map $M(x)$ which is identically zero. Such an $M(x)$ reduces Equation 3 to the identity map. To avoid this reduction for such images, we set $M(x) = 1$, so that a global gamma transformation can be applied to the patch.

5.2 Evaluation

We used image-level sensitivity and image-level specificity to evaluate the performance of a baseline U-Net and a local gamma augmented U-Net+LG on an ischemic stroke classification task. As segmentation models, U-Nets are typically evaluated with pointwise metrics such as the Dice-Sørensen coefficient or the Jaccard index. Our method of model evaluation, namely classification instead of segmentation, is required due to the lack of segmentation maps as annotations for the data set. To treat a segmentation model as a classification model, we define a U-Net prediction to be positive if at least one voxel in the model's output is predicted as corresponding to ischemic stroke.

5.3 Results

Table 2 summarises our results. We observe an increase in image-level sensitivity from 0.781 to 0.844 for the MVS dataset and an increase in image-level sensitivity from 0.757 to 0.814 for the RH dataset. These increases are accompanied with small reductions in image-level specificity on both datasets.

For the WUS dataset, the improvement is even greater: an increase in image-level sensitivity from 0.566 to 0.736. We speculate that this larger increase in sensitivity is due to the influence of local gamma augmentation. In particular, the local gamma augmentation scheme exposes U-Net+LG to more examples of MRIs with signal differences between the DWI sequence and the remaining sequences. We also mention that specificity is not reported for WUS: a specificity score for this dataset is meaningless since every image contains an ischemic stroke.

5.4 Discussion

The simplicity of local gamma augmentation, namely its use of a one-parameter family of intensity transformations for data augmentation, means that it is very straightforward to implement and hence is amenable to further exploration. One potential research avenue would be to identify better choices of the gamma augmentation parameter. Our method uses a mixture distribution to select augmentation parameters, but this distribution is fixed. In this way, the chosen augmentation parameters are data-independent. Other research avenues could pursue the use of targeted gamma augmentations for other applications, such as adversarial training or pseudo-healthy synthesis.

A serious limitation of the method is its dependence on pointwise labels: without segmentation maps, only image-level gamma augmentations can be applied. Another limitation is the dependence on pathological tissues which are characterized by intensity differences relative to normal tissue. Not all brain diseases are characterized by intensity differences. Hence it is questionable that local gamma augmentation (or local intensity augmentation of any kind) would be effective.

6 Conclusion

We presented local gamma augmentation, an intensity-based method of data augmentation which, in a supervised setting, can be used to target pathological tissues in human brain MRIs. We applied this method to compensate for an intensity bias in the training set due to an over-representation of non-hyperacute ischemic strokes. This method was compared against a baseline data augmentation method without local gamma augmentation on three MRI data sets. The results suggest that restricting gamma transformations to ischemic stroke lesions during training of a segmentation model can enhance image-level sensitivity without impairing image-level specificity.

References

[1] E. Castro, J. S. Cardoso, and J. C. Pereira. "Elastic deformations for data augmentation in breast cancer mass detection". In: *2018 IEEE EMBS Int. Conf. Biomed. Heal. Informatics*. Vol. 2018-Janua. IEEE, Mar. 2018, pp. 230–234. ISBN: 978-1-5386-2405-0. DOI: 10.1109/BHI.2018.8333411. URL: http://ieeexplore.ieee.org/document/8333411/.

[2] P. Novosad, V. Fonov, and D. L. Collins. "Accurate and robust segmentation of neuroanatomy in T1-weighted MRI by combining spatial priors with deep convolutional neural networks". In: *Hum. Brain Mapp.* 41.2 (2020), pp. 309–327. ISSN: 10970193. DOI: 10.1002/hbm.24803. arXiv: 1902.01478.

[3] M. D. Cirillo, D. Abramian, and A. Eklund. "What Is the Best Data Augmentation for 3D Brain Tumor Segmentation?" In: *Proc. - Int. Conf. Image Process. ICIP* 2021-Septe (2021), pp. 36–40. ISSN: 15224880. DOI: 10.1109/ICIP42928.2021.9506328. arXiv: 2010.13372.

[4] B. D. Basaran, P. M. Matthews, and W. Bai. "New lesion segmentation for multiple sclerosis brain images with imaging and lesion-aware augmentation". In: *Front. Neurosci.* 16 (2022). ISSN: 1662453X. DOI: 10.3389/fnins.2022.1007453.

[5] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". In: *Nat. Methods* 18.2 (Dec. 2021), pp. 203–211. ISSN: 15487105. DOI: 10.1038/s41592-020-01008-z. URL: http://dx.doi.org/10.1038/s41592-020-01008-z%20http://www.nature.com/articles/s41592-020-01008-z.

[6] F. Perez, C. Vasconcelos, S. Avila, and E. Valle. "Data Augmentation for Skin Lesion Analysis". In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. Vol. 11041 LNCS. 2018, pp. 303–311. ISBN: 9783030012007. DOI: 10.1007/978-3-030-01201-4_33. arXiv: 1809.01442. URL: http://link.springer.com/10.1007/978-3-030-01201-4%7B%5C_%7D33.

[7] D. Karimi and A. Gholipour. "Improving Calibration and Out-of-Distribution Detection in Deep Models for Medical Image Segmentation". In: *IEEE Trans. Artif. Intell.* 4.2 (2023), pp. 383–397. ISSN: 26914581. DOI: 10.1109/TAI.2022.3159510.

[8] R. Balestriero, L. Bottou, and Y. LeCun. "The Effects of Regularization and Data Augmentation are Class Dependent". In: (2022), pp. 1–21. arXiv: 2204.03632. URL: http://arxiv.org/abs/2204.03632.

[9] W. Liu, Q. Lu, Z. Zhuo, Y. Liu, and C. Ye. "One-Shot Segmentation of Novel White Matter Tracts via Extensive Data Augmentation". In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 13431 LNCS (2022), pp. 133–142. ISSN: 16113349. DOI: 10.1007/978-3-031-16431-6_13.

[10] A. Fernandez-Quilez, C. Gabrielsen Andersen, T. Eftestøl, S. R. Kjosavik, and K. Oppedal. "3D Masked Modelling Advances Lesion Classification in Axial T2w Prostate MRI". In: *Proc. North. Light. Deep Learn. Work.* 4 (2023), pp. 1–8. DOI: 10.7557/18.6787.

[11] X. Zhang, C. Liu, N. Ou, X. Zeng, Z. Zhuo, Y. Duan, X. Xiong, Y. Yu, Z. Liu, Y. Liu, and C. Ye. "CarveMix: A simple data augmentation method for brain lesion segmentation". In: *Neuroimage* 271.March (2023). ISSN: 10959572. DOI: 10.1016/j.neuroimage.2023.120041.

[12] X. Sun, H. Fang, Y. Yang, D. Zhu, L. Wang, J. Liu, and Y. Xu. "Robust Retinal Vessel Segmentation from a Data Augmentation Perspective". In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 12970 LNCS (2021), pp. 189–198. ISSN: 16113349. DOI: 10.1007/978-3-030-87000-3_20. arXiv: 2007.15883.

[13] P. Liskowski and K. Krawiec. "Segmenting Retinal Blood Vessels with Deep Neural Networks". In: *IEEE Trans. Med. Imaging* 35.11 (2016), pp. 2369–2380. ISSN: 1558254X. DOI: 10.1109/TMI.2016.2546227.

[14] B. Billot, D. N. Greve, O. Puonti, A. Thielscher, K. Van Leemput, B. Fischl, A. V. Dalca, and J. E. Iglesias. "SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining". In: *Med. Image Anal.* 86 (May 2023), p. 102789. ISSN: 13618415. DOI: 10.1016/j.media.2023.102789. arXiv: 2107.09559. URL: http://arxiv.org/abs/2107.09559%20https://linkinghub.elsevier.com/retrieve/pii/S1361841523000506.

[15] L. Nijskens, Cornelis, A. van den Berg, J. J. Verhoeff, and M. Maspero. "Exploring contrast generalisation in deep learning-based brain MRI-to-CT synthesis". In: *Phys. Medica* 112.May (2023), p. 102642. ISSN: 1120-1797. DOI: 10.1016/j.ejmp.2023.102642. arXiv: 2303.10202. URL: https://arxiv.org/abs/2303.10202v1.

[16] L. Boone, M. Biparva, P. Mojiri Forooshani, J. Ramirez, M. Masellis, R. Bartha, S. Symons, S. Strother, S. E. Black, C. Heyn, A. L. Martel, R. H. Swartz, and M. Goubran. "ROOD-MRI: Benchmarking the robustness of deep learning segmentation models to out-of-distribution and corrupted data in MRI". In: *Neuroimage* 278.April (Sept. 2023), p. 120289. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2023.120289. arXiv: 2203.06060. URL: https://

linkinghub.elsevier.com/retrieve/pii/S1053811923004408.

[17] P. Diao, A. Pai, C. Igel, and C. H. Krag. "Histogram-Based Unsupervised Domain Adaptation for Medical Image Classification". In: vol. 3. Springer Nature Switzerland, 2022, pp. 755–764. ISBN: 9783031164484. DOI: 10.1007/978-3-031-16449-1_72. URL: http://dx.doi.org/10.1007/978-3-031-16449-1%7B%5C_%7D72%20https://link.springer.com/10.1007/978-3-031-16449-1%7B%5C_%7D72.

[18] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang. "Random Erasing Data Augmentation". In: *Proc. AAAI Conf. Artif. Intell.* 34.07 (Apr. 2020), pp. 13001–13008. ISSN: 2374-3468. DOI: 10.1609/aaai.v34i07.7000. URL: https://ojs.aaai.org/index.php/AAAI/article/view/7000.

[19] T. DeVries and G. W. Taylor. "Improved Regularization of Convolutional Neural Networks with Cutout". In: *arXiv* (Aug. 2017). arXiv: 1708.04552. URL: http://arxiv.org/abs/1708.04552.

[20] J. Bernhardt, K. S. Hayward, G. Kwakkel, N. S. Ward, S. L. Wolf, K. Borschmann, J. W. Krakauer, L. A. Boyd, S. T. Carmichael, D. Corbett, and S. C. Cramer. "Agreed definitions and a shared vision for new standards in stroke recovery research: The Stroke Recovery and Rehabilitation Roundtable taskforce". In: *Int. J. Stroke* 12.5 (2017), pp. 444–450. ISSN: 17474949. DOI: 10.1177/1747493017711816.

[21] G. Thomalla, B. Cheng, M. Ebinger, Q. Hao, T. Tourdias, O. Wu, J. S. Kim, L. Breuer, O. C. Singer, S. Warach, S. Christensen, A. Treszl, N. D. Forkert, I. Galinovic, M. Rosenkranz, T. Engelhorn, M. Köhrmann, M. Endres, D.-W. Kang, V. Dousset, A. G. Sorensen, D. S. Liebeskind, J. B. Fiebach, J. Fiehler, and C. Gerloff. "DWI-FLAIR mismatch for the identification of patients with acute ischaemic stroke within 4·5 h of symptom onset (PRE-FLAIR): a multicentre observational study". In: *Lancet Neurol.* 10.11 (Nov. 2011), pp. 978–986. ISSN: 14744422. DOI: 10.1016/S1474-4422(11)70192-2. URL: http://dx.doi.org/10.1016/S1474-4422(11)70192-2%20https://linkinghub.elsevier.com/retrieve/pii/S1474442211701922.

[22] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, and et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)". In: *IEEE Trans. Med. Imaging* 34.10 (2015), pp. 1993–2024. ISSN: 1558254X. DOI: 10.1109/TMI.2014.2377694.

[23] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, and et al. "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features". In: *Sci. Data* 4.July (2017), pp. 1–13. ISSN: 20524463. DOI: 10.1038/sdata.2017.117. URL: http://dx.doi.org/10.1038/sdata.2017.117.

[24] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, and et al. "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge". In: (2018). arXiv: 1811.02629. URL: http://arxiv.org/abs/1811.02629.

[25] C. W. Cereda, S. Christensen, B. C. Campbell, N. K. Mishra, M. Mlynash, C. Levi, M. Straka, M. Wintermark, R. Bammer, G. W. Albers, M. W. Parsons, and M. G. Lansberg. "A benchmarking tool to evaluate computer tomography perfusion infarct core predictions against a DWI standard". In: *J. Cereb. Blood Flow Metab.* 36.10 (2016), pp. 1780–1789. ISSN: 15597016. DOI: 10.1177/0271678X15610586.

[26] A. Hakim, S. Christensen, S. Winzeck, M. G. Lansberg, M. W. Parsons, C. Lucas, D. Robben, R. Wiest, M. Reyes, and G. Zaharchuk. "Predicting Infarct Core From Computed Tomography Perfusion in Acute Ischemia With Machine Learning: Lessons From the ISLES Challenge". In: *Stroke* 52.7 (2021), pp. 2328–2337. ISSN: 15244628. DOI: 10.1161/STROKEAHA.120.030696.

[27] O. Ronneberger, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Lect. Notes Comput. Sci.* Vol. 9351. Springer International Publishing, May 2015, pp. 234–241. ISBN: 9783319245737. DOI: 10.1007/978-3-319-24574-4_28. arXiv: 1505.04597. URL: http://link.springer.com/10.1007/978-3-319-24574-4%7B%5C_%7D28.