

Learning From Mouse CT-Scan Brain Images To Detect MRA-TOF Human Vasculatures*

Sara Chater¹, Nathan Lauzeral², Anass Nouri¹, Youssef El Merabet¹ and Florent Atrousseau²

Abstract—The earlier studies on brain vasculature semantic segmentation used classical image analysis methods to extract the vascular tree from images. Nowadays, deep learning methods are widely exploited for various image analysis tasks. One of the strong restrictions when dealing with neural networks in the framework of semantic segmentation is the need to dispose of a ground truth segmentation dataset, on which the task will be learned. It may be cumbersome to manually segment the arteries in a 3D volumes (MRA-TOF typically). In this work, we aim to tackle the vascular tree segmentation from a new perspective. Our objective is to build an image dataset from mouse vasculatures acquired using CT-Scans, and enhance these vasculatures in such a way to precisely mimic the statistical properties of the human brain. The segmentation of mouse images is easily automatized thanks to their specific acquisition modality. Thus, such a framework allows to generate the data necessary for the training of a Convolutional Neural Network - i.e. the enhanced mouse images and their corresponding ground truth segmentation - without requiring any manual segmentation procedure. However, in order to generate an image dataset having consistent properties (strong resemblance with MRA images), we have to ensure that the statistical properties of the enhanced mouse images do match correctly the human MRA acquisitions. In this work, we evaluate at length the similarities between the human arteries as acquired on MRA-TOF and the “humanized” mouse arteries produced by our model. Finally, once the model duly validated, we experiment its applicability with a Convolutional Neural Network.

I. INTRODUCTION

A. Motivation and context

This work takes place in a broader framework, in which we intend to detect and monitor intra-cranial aneurysms on human Time-Of-Flight (TOF) Magnetic Resonance Angiography (MRA) images (abbreviated as TOF in the remaining of the paper). Prior to detect or characterize cerebral aneurysms, an efficient image segmentation is mandatory. In this work, we are particularly interested by the segmentation of the Circle of Willis, as this is where most of the intra-cranial aneurysms occur [1]. A proper segmentation of the cerebral vascular tree would allow a precise characterization of the intra-cranial arteries. As a matter of fact, knowing precisely the geometrical arrangement of the arteries can be of tremendous help in order to determine the risk of developing an aneurysm [2]. In this work, we intend to optimize the image segmentation in order to detect aneurysms, but also to possibly study the geometric disposition of the bifurcations,

which might lead us to estimate the risk of occurrence of a new aneurysm. To reach state-of-the-art performance for semantic image segmentation, we follow the route of the most recent works which leverage the impressive capabilities of the deep learning algorithms.

In the literature, several architectures were proposed to perform the segmentation task. Generally, the Convolutional Neural Networks (CNN) provide the best results. For example, simple architectures were successfully used in [3] and [4] to segment human TOF. Furthermore, the particular U-net architecture [5] became a reference when it comes to the segmentation of biomedical images. Its encoder-decoder structure has been the basis of many new architectures. In the context of vascular tree segmentation it was used, for instance, in [6] for the segmentation of TOF in patients with cerebrovascular disease, or in [7] for the segmentation of digital subtraction angiography. However, all these approaches require a large amount of annotated data to train the networks. Indeed, in the context of medical image analysis, and in particular for semantic segmentation purpose, images must be associated with their corresponding ground-truth segmentation. The networks are then trained to reproduce the segmentation operation by learning a mapping function from the domain of the raw medical images to the segmented images domain. The initial task of data labelling is often tedious and prone to variability, especially in the segmentation case which is known to be highly observer-dependent [8]. Moreover, the inter-variability of the acquisition devices and procedures exacerbates the problem as the neural network training must be repeated for each new configuration and thus requires new ground-truth segmented images. To circumvent the lack of labelled data, techniques such as the few-shot semantic segmentation [9] have been developed. In [10], the authors even proposed a medical images data augmentation approach for the extreme case where only one labelled image is available. Alternatively, the authors in [11] proposed an approach based on a pre-segmentation based on a hidden Markov random field model coupled with a specifically tailored CNN. They showed that their approach was able to handle images from different devices with different resolution while being trained on a small amount of annotated data. In the deep learning framework, training a network on a specific class of images to be applied on a similar but different image class is known as *domain adaptation*. Such a procedure has been used to alleviate the domain shift induced by the multi-modal nature of medical images. In particular, the use of the Cycle-GAN structures [12] has proven to be efficient on X-ray to X-ray, CT-scan to X-ray, and MRI to

*This work was supported by the U-CAN project

¹S. Chater, A. Nouri and Y. El Merabet are with Ibn Tofail University, Morocco. <FirstName>.<LastName>@uit.ac.ma

²N. Lauzeral and F. Atrousseau are with the University of Nantes, France. <FirstName>.<LastName>@univ-nantes.fr

MRI datasets [13]. To circumvent the lack of high quality labelled medical images, some authors also proposed to use synthetic data [14]. Nonetheless, synthetic data may increase the domain shift between the source and target images.

B. Scope of the current work

In this paper, we propose to alleviate the labelling burden associated with the generation of a ground-truth segmentation by simplifying the generation of the segmented images while using biological data. To this end, we leverage the availability of images acquired on animal subject and, especially, the possibility to inject a contrast agent (baryum sulfate) before the acquisition. It allows to drastically enhance the intensity gap between the injected vessels and the non-essential background tissues. By coupling this procedure with the use of a microtomography (μ -CT) scanner, we were then able to generate high-quality images of mice brain vasculatures with a very high contrast, making the segmentation trivial. The injection of baryum sulfate was lethal for the mice and their brains were removed from the skulls before the μ -CT acquisitions. A 2D image of a mouse brain acquired through the μ -CT scanner is shown on Fig. 1. It appears clearly that only the vascular tree is visible, without any alteration from other brain structures such as grey matter, cerebro-spinal fluid, etc. The segmentation of such an image requires little to no supervision and there is no need for medical experts to validate it, a simple thresholding returns satisfying results. To insure the robustness of the method and the quality of the segmentations, the well-known maximum entropy thresholding was applied. As mentioned before, using a neural network trained on the mice μ -CT scans and tested on humans TOF images is not trivial because of the domain shift. Two major differences arise. First, the specie on which the data have been collected is not human, and second, the acquisition modality is not the same; we have μ -CT scans acquired on a unique machine for the mice and TOF acquired on a broad panel of devices for the humans. We thus need to perform domain adaptation. To this end, we propose to “humanize” the mice images in a supervised manner using a well-chosen set of representative features.

In Section II the “humanization” procedure is described. The latter is composed of two main steps: a noise generation detailed in Section II-B and a template matching step explained in Section II-C. Then, the humanization process is validated in Section III using the half-Unet structure proposed in [6]. First, the quality of the generated images is verified in Section III-A, then, the neural networks performance are evaluated in Section III-B. Eventually, we conclude the paper and propose some perspectives in Section IV.

II. TOWARDS HUMANIZED MICE VASCULATURES

A. Overview

As shown in Fig. 1, the μ -CT acquisition significantly differs from the TOF image. While on the mice acquisitions only the vascular tree appears, on human acquisitions, several brain tissues (hypothalamus, cerebro-spinal fluid, grey matter, etc.) interfere with the proper detection of the vascular

tree. However, we believe that using some image processing techniques might allow us to transform the mice acquisitions in such a way that they resemble the TOF volumes.

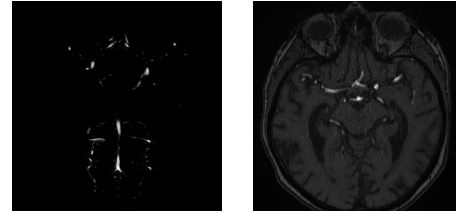


Fig. 1. 2D view (slice) of a non-segmented mouse brain micro-CT (left) and a non segmented human brain TOF (right).

Working on full-scale 3D images might lead to major difficulties because of the high dimensionality and complexity of the data that we should encompass in the humanization process. We thus plan to tune a neural network to carry the segmentation out on 2D patches of the images. Hence, we only need to reach a sufficient resemblance on a local viewpoint. The humanization process is decomposed into two stages. At first, 2D patches are extracted from the mouse μ -CT scans and a well-defined noise is added. This latter is computed so as to match the human TOF noise. Then, a template matching is used in order to insure that the newly generated patches of humanized data correctly match with realistic human TOF. To this end a Normalized Cross-Correlation (NCC) score is computed for each patch and only those with sufficiently high values are kept to form the training data set. The procedure is summarized in Fig. 2.

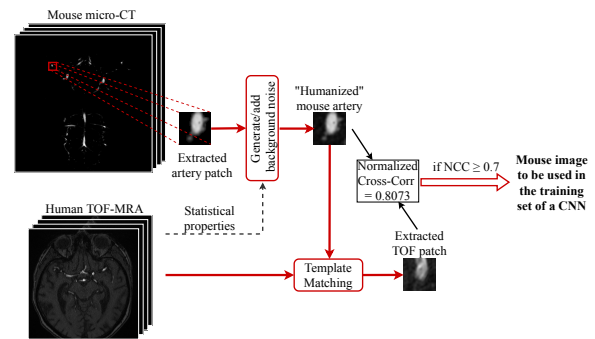


Fig. 2. Schematic overview of the proposed approach, background noise is added on mice arteries, so that they properly match human MRI arteries. The most resembling image patches are used to form the training dataset.

B. Noise generation

First, a noise model corresponding to the TOF is defined. This noise must not only follow the same kind of distribution as the one that stems from the TOF, but its statistical properties must also closely match those of the TOF acquisitions. We thus evaluate the similarities with regard to three different aspects: the distribution type, the basic statistical properties (mean, variance, and amplitude), and a blur factor. The choice of the blur factor is mandatory to insure that the spatial frequency of our model fits the one from the TOF

device. This means that not only we shall be able to produce some noise that will match various MRI scanners or various acquisition settings on a particular machine, but we also might be able to reproduce, within a given TOF volume, a large variety of background noises due to physiological properties, i.e. blood vessels surrounded by the grey or white matter or by the cerebrospinal fluid.

Noise on TOF signals follows a Gaussian distribution but manifests itself in magnitude images as a Rician distribution of pixel intensities [15]. In order to generate humanized mice arteries, we apply the following process:

- 1) Small TOF noise patches are extracted from the full 3D stacks.
- 2) For each extracted patch, a noise patch - of the same size - following a Rician distribution is generated so as to match its statistical properties.
- 3) A blur kernel is then iteratively applied on the noise patches in order to match the blur score of their corresponding TOF patches. The matching error ϵ_i for the i -th patch is defined as

$$\epsilon_i = |\mu_i^N - \mu_i^{TOF}| + |\sigma_i^N - \sigma_i^{TOF}| + |\mathcal{B}_i^N - \mathcal{B}_i^{TOF}|,$$

where μ^N (resp. μ^{TOF}) is the mean of the noise patch (resp. TOF patch), σ^N (resp. σ^{TOF}) is the standard deviation and \mathcal{B}^N (resp. \mathcal{B}^{TOF}), a blur score collected on the images. At each iteration the filter size is increased, a stopping criterion is defined such as iterations stop when ϵ_i starts increasing.

- 4) A correction of the noise patches' average values is applied so that each set of corresponding images are perfectly aligned.
- 5) Small patches of artery sections are extracted from the μ -CT mice acquisitions. To this end, a skeletonization of the vascular tree is first applied on the μ -CT scans. At regular intervals along the skeleton, patches are extracted across the x , y and z planes describing the 3D space. To enrich the dataset, a data augmentation step is performed. Rescaling, flips and various rotations are applied to generate new patches. Scaling ratios were chosen so that the modelled arteries are of similar diameters as the humans'.
- 6) A flattening of the previously constituted database of arteries is applied so that the amplitude of the mice arteries matches the humans' ones from the TOF.
- 7) For each artery image, a patch of noise generated according to steps 1, 2, and 3 is added.
- 8) A final Gaussian blur with a fixed kernel is applied onto the assembled noisy artery.

C. Template matching

Once the dataset of humanized mice images is constituted, a template matching [16] is used to insure the quality of the humanization. This procedure detects, among a slice of human TOF, the best patch matching a given humanized image. It then returns a NCC in $[0, 1]$ defining the goodness of the humanization process for the current image. Only images with a NCC above 0.7 (a threshold experimentally

fixed) were kept. Some of the results are shown in Fig. 3. Visually, mice images augmented with the noise model closely match the human patches.

It should be noted that thanks to the μ -CT scan and the MRI resolutions, both human and mice arteries exhibit similar diameters on the processed images. Indeed, mice brain acquisitions were obtained on a Bruker μ -CT with a 12 μ m per pixel resolution, the MRA acquisitions came from various MRI scanners, but mostly our resolution was 351 μ m per voxel (along the x and y axis). Thus, looking at the respective artery diameters in rodent and humans [17], [18], we can observe that in our gathered dataset, human arteries range from 7.12 to 11.36 pixels, whereas mice arteries fall in the range 4.16 to 12.07 pixels.

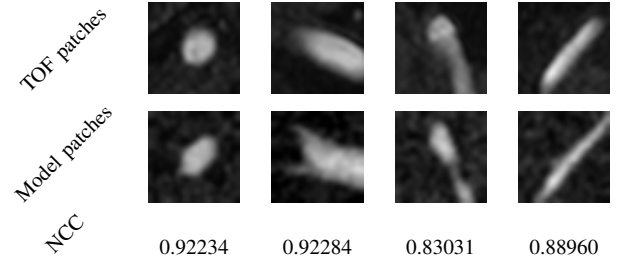


Fig. 3. Some examples of corresponding images obtained after template matching. The top row contains TOF patches and the bottom row shows patches from our model. For each pair, the NCC is given.

III. EXPERIMENTAL RESULTS

A. Humanized mice arteries

First, 5 172 modeled patches were generated from seven mice μ -CT scans. After data augmentation 20 000 images patches were available. Finally, thanks to the template matching filtering procedure, 17 121 patches of humanized mice images were retained for the neural network's training step.

In order to avoid feeding the neural network with images bearing several arteries, or even in which various tissues are present, we have opted for the generation of image patches of size 40×40 . Several patch sizes were tested (from 16×16 to 64×64), a patch size of 40 provided the best results.

Once the mice brain images were humanized, it was crucial to evaluate their resemblance with the TOF. As we have previously discussed, our model is tuned so that the background noise can be closely reproduced. Several properties of the said noise must be scrupulously studied. In the following, we will compare such image patches extracted from the TOF images (the target we aim to reproduce), and the image patches generated by our model.

Both the average and standard deviation were collected from some cropped portions of the TOF and from our model. The distribution of the gathered data is coherent, and very few modulations can be observed between the TOF and our model (see Fig. 4).

Besides the basic statistical properties, we have conducted two separate tests. We first compared the images from a texture analysis viewpoint. We used the framework proposed by Haralick [19], this latter provides 14 measures on a

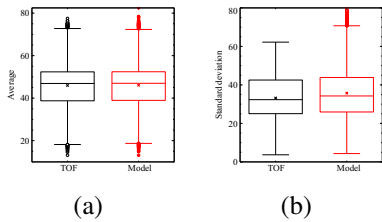


Fig. 4. Comparison of (a) the average noise and (b) the standard deviation for square patches on both a TOF volume (left boxplots) and our model (right boxplots).

given image. Combinations of these measures can lead to a total of 28 features. However, the first 13 are considered to present a good representation of the texture (the 14th is subject to controversies). Future works might consider using LBP for this texture assessment task. This set of features represents various aspects of the given texture. We show on Fig. 5(a) how, with respect to all 13 Haralick features, our model (represented by blue bars), matches quite well the TOF images (red bars). Apart from the texture analysis,

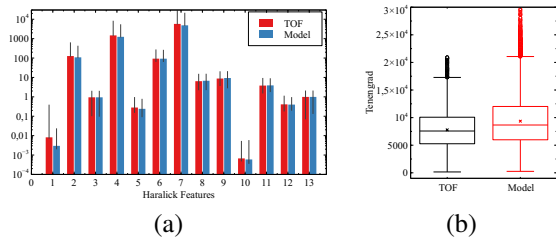


Fig. 5. (a) Repartition of the Haralick features and (b) the Tenengrad coefficients for both the TOF image patches and the measures from the model

we also decided to compute a blur measure of the modeled patches. To this end, we used the *Tenengrad* coefficient [20]. The Fig. 5(b) clearly shows that the distribution of the blur measures (Tenengrad) for the model closely matches those of the ground truth TOFs.

Once the modeled humanized patches evaluated, we can thus carry on with the next step, being the vaculature segmentation via a neural network.

B. Application to human brains : preliminary results

The half U-net convolutional neural network structure from [6] was employed to test the proposed approach. Slight modifications were brought. First, each convolutional layers was followed by a batch-normalization layer and a rectified linear unit, excepted for the last layer where no batch-normalization was used and a pixel-wise sigmoid activation function was applied. Also, no dropout was applied. The network was trained using the Adam optimizer with a learning rate of 0.01, the Dice coefficient loss function described in [6], and a batch-size of 32. Three training datasets were employed to assess the performance of the proposed approach. The first dataset - afterward referred to as \mathcal{D}_1 - was composed of the previously introduced 17 121 images of humanized mice arteries plus 1712 images of background noise generated as mentioned in Section II. The second dataset, \mathcal{D}_2 , was based on the first one, but the images

were augmented on-the-fly using the *ImageDataGenerator* function implemented in Keras. This function accepts the original data, randomly transforms it, and returns only the new, transformed data in a random order. Here, standard transformations were used like rotations, shifts, and scalings. 18 833 transformed images were used per epoch. The last dataset, \mathcal{D}_3 , was composed of 12 477 images of mice arteries without any noise enhancement. These data were also augmented using the *ImageDataGenerator* function. The network was trained on each dataset for 30 epochs. A validation dataset composed of 16 749 patches (40×40 pixels) extracted from 36 TOFs was used to monitor the performance. No overfitting was detected. Then, a test dataset of 43 TOFs was used to evaluate the performance of each network for each training dataset. The ground-truth segmentation of the test images was computed using a semi-automatic segmentation software 3DSlicer¹. Note that all human images were preprocessed with a brain extraction tool [21]. To obtain the predicted segmentation of each network a threshold of 0.5 was used on the probability maps.

The performance were evaluated with the Dice coefficient, the 95%-percentile and average Hausdorff distances. Results are shown in Fig. 6.

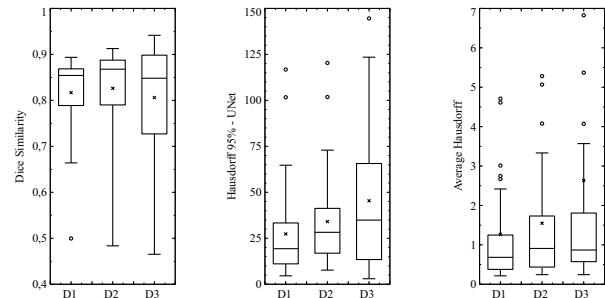


Fig. 6. Performance of the half-Unet evaluated using the Dice coefficient, the 95%-percentile and averaged Hausdorff distances on the three training datasets.

The choice of the training dataset shows a significant impact on the performance. First, it is interesting to notice that the performance using the dataset \mathcal{D}_3 are already acceptable. Median values of the Dice coefficient and of the Hausdorff distances are close to those found in the literature. The main flaw lies in the high variability of the results. The humanization in \mathcal{D}_1 and \mathcal{D}_2 allows to significantly reduce the latter for all the performance metrics. The neural networks are thus more robust and reliable, which is primordial for medical applications. Furthermore, when looking at the median values of the Dice coefficient, it appears that the humanization of the images also allows to get better results, that is all the more true if the data augmentation is used. This result proves the validity of our approach as the neural networks perform better in average when the image adaptation is done and with less variations. However, the gain in performance is not extremely significant and additional preprocessing might be beneficial to get a better humanization of the mouse

¹<https://www.slicer.org>

images. Finally, the Hausdorff distance is also generally lower - i.e. better - for the humanized training data. This metric is better suited to evaluate the quality of the vessels delineation. It means that the humanization process permits to better detect the borders of the vessels, which is very important for applications such as aneurysm characterization, that we intend to perform in future works. Fig. 7 shows the error map for a Unet trained on the three previously detailed datasets. The insets in the lower left corner presents a zoom in the central part of the slices (delimited in the white square). Brightness was increased in the insets for optimized visualization.

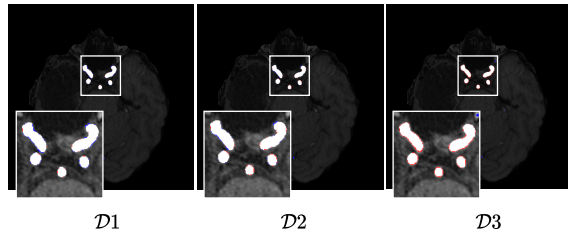


Fig. 7. Error map for one TOF slice when training from data set $\mathcal{D}1$, $\mathcal{D}2$, or $\mathcal{D}3$. True positives are shown in white, false negatives in red and false positives in blue.

IV. CONCLUSION AND PERSPECTIVES

We tackled in this work the problem of ground-truth data generation for the automatic segmentation of human brain vasculature on TOF acquisitions. The issue has been addressed from an original point of view as instead of asking experts to manually label the data, as commonly seen in the literature, we suggested to employ ready-to-use data in the form of mouse μ -CT scans (with prior sulfate baryum injection). These acquisitions can be automatically segmented in an unsupervised manner, hence avoiding the manual segmentation burden. To use these data to segment human TOF, we presented a procedure to humanize the mice μ -CT scan in order to train a neural network. The resemblance of the generated images with human TOF were positively evaluated regarding 3 aspects: the statistical correspondance, a full texture analysis, and the spatial-frequency similarity. Preliminary segmentation results showed a positive impact of this humanization on the training of a U-Net neural network architectures in terms of both Dice score and Hausdorff distance. However, the humanization process could still be improved in regard to the neural networks efficiency as the performance gap between humanized and non-humanized trained networks remains modest. For instance, the thresholding factor set to 0.7 in the template matching procedure might be too low and would allow poorly humanized images to be used as training data in the neural networks. Moreover, the study remained in 2D, further work should focus on 3D applications that might provide better performance. Eventually, the domain adaptation could be embedded within the neural networks. The humanization would then be done in an unsupervised manner and the difficult choices of the noise type and the blurring method could be bypassed.

REFERENCES

- [1] A. Keedy, "An overview of intracranial aneurysms," *McGill Journal of Medicine, MJM*, vol. 9, no. 2, pp. 141–146, 2006.
- [2] A. S. E. Bor, B. K. Velthuis, C. B. Majoie, and G. J. Rinkel, "Configuration of intracranial arteries and development of aneurysms," *Neurology*, vol. 70, no. 9, pp. 700–705, 2008.
- [3] R. Phellan, A. Peixinho, A. Falcão, and N. D. Forkert, "Vascular segmentation in tof mra images of the brain using a deep convolutional neural network," in *Intravascular Imaging and Computer Assisted Stenting, and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*. Springer, 2017, pp. 39–46.
- [4] G. Tetteh, V. Efremov, N. D. Forkert, M. Schneider, J. Kirschke, B. Weber, C. Zimmer, M. Piraud, and B. H. Menze, "Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes," *arXiv preprint arXiv:1803.09340*, 2018.
- [5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [6] M. Livne, J. Rieger, O. U. Aydin, A. A. Taha, E. M. Akay, T. Kossen, Sobesky, et al., "A u-net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease," *Frontiers in neuroscience*, vol. 13, p. 97, 2019.
- [7] C. Meng, C. Sun, S. Guan, Q. Wang, R. Zong, and L. Liud, "Multiscale dense convolutional neural network for DSA cerebrovascular segmentation," *Neurocomputing*, vol. 373, pp. 123–134, 2020.
- [8] L. Joskowicz, D. Cohen, N. Caplan, and J. Sosna, "Inter-observer variability of manual contour delineation of structures in ct," *European radiology*, vol. 29, no. 3, pp. 1391–1399, 2019.
- [9] N. Dong and E. P. Xing, "Few-shot semantic segmentation with prototype learning," in *BMVC*, vol. 3(4), 2018.
- [10] A. Zhao, G. Balakrishnan, F. Durand, J. V. Guttag, and A. V. Dalca, "Data augmentation using learned transformations for one-shot medical image segmentation," in *IEEE conf. on computer vision and pattern recognition*, 2019, pp. 8543–8553.
- [11] S. Fan, Y. Bian, H. Chen, Y. Kang, Q. Yang, and T. Tan, "Unsupervised cerebrovascular segmentation of tof-mra images based on deep neural network and hidden markov random field model," *Front. Neuroinform.*, vol. 13, p. 7, 2020.
- [12] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.
- [13] Y. Zhang, S. Miao, T. Mansi, and R. Liao, "Task driven generative modeling for unsupervised domain adaptation: Application to x-ray image segmentation," in *Intl Conf. on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 599–607.
- [14] F. Mahmood, R. Chen, and N. J. Durr, "Unsupervised reverse domain adaptation for synthetic medical images via adversarial training," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2572–2581, 2018.
- [15] A. Cardenas-Blanco, C. Tejos, P. Irazrazaval, and I. Cameron, "Noise in magnitude magnetic resonance images," *Concepts in Magnetic Resonance*, 2008.
- [16] J. Lewis, "Fast normalized cross-correlation," in *Vision interface*, vol. 95, 1995.
- [17] S. Mujagic, M. Moranjic, N. Mesanovic, and S. Osmanovic, "The inner diameter of arteries of the circle of willis regarding gender and age on magnetic resonance angiography," *Acta Medica Saliniana*, vol. 42(2), pp. 6–12, 2013.
- [18] B. Qian, R. Rudy, T. Cai, and R. Du, "Cerebral artery diameter in inbred mice varies as a function of strain," *Frontiers in Neuroanatomy*, Feb. 2018.
- [19] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979.
- [20] U. Ali and M. Mahmood, "Analysis of blur measure operators for single image blur segmentation," *Applied Sciences*, vol. 8, no. 5, pp. 807–839, 2018.
- [21] S. M. Smith, "Fast robust automated brain extraction," *Human brain mapping*, vol. 17, no. 3, pp. 143–155, 2002.