



OPEN ACCESS

EDITED BY

Ken Butcher,
University of Alberta, Canada

REVIEWED BY

Carole Frindel,
Université de Lyon, France
Freda Werdiger,
The University of Melbourne, Australia

*CORRESPONDENCE

Tabea Kossen
✉ tabea.kossen@charite.de

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION

This article was submitted to Stroke, a section of the journal Frontiers in Neurology

RECEIVED 22 September 2022

ACCEPTED 15 December 2022

PUBLISHED 10 January 2023

CITATION

Kossen T, Madai VI, Mutke MA, Hennemuth A, Hildebrand K, Behland J, Aslan C, Hilbert A, Sobesky J, Bendszus M and Frey D (2023) Image-to-image generative adversarial networks for synthesizing perfusion parameter maps from DSC-MR images in cerebrovascular disease. *Front. Neurol.* 13:1051397. doi: 10.3389/fneur.2022.1051397

COPYRIGHT

© 2023 Kossen, Madai, Mutke, Hennemuth, Hildebrand, Behland, Aslan, Hilbert, Sobesky, Bendszus and Frey. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Image-to-image generative adversarial networks for synthesizing perfusion parameter maps from DSC-MR images in cerebrovascular disease

Tabea Kossen^{1,2*†}, Vince I. Madai^{1,3,4†}, Matthias A. Mutke⁵, Anja Hennemuth^{2,6,7}, Kristian Hildebrand⁸, Jonas Behland¹, Cagdas Aslan⁸, Adam Hilbert¹, Jan Sobesky^{9,10}, Martin Bendszus⁵ and Dietmar Frey¹

¹Charité Lab for Artificial Intelligence in Medicine (CLAIM), Charité-Universitätsmedizin Berlin, Berlin, Germany, ²Department of Computer Engineering and Microelectronics, Computer Vision and Remote Sensing, Technical University Berlin, Berlin, Germany, ³QUEST Center for Responsible Research, Berlin Institute of Health (BIH), Charité-Universitätsmedizin Berlin, Berlin, Germany, ⁴Faculty of Computing, Engineering and the Built Environment, School of Computing and Digital Technology, Birmingham City University, Birmingham, United Kingdom, ⁵Department of Neuroradiology, Heidelberg University Hospital, Heidelberg, Germany, ⁶Institute for Imaging Science and Computational Modelling in Cardiovascular Medicine, Charité-Universitätsmedizin Berlin, Berlin, Germany, ⁷Fraunhofer MEVIS, Bremen, Germany, ⁸Department of Computer Science and Media, Berlin University of Applied Sciences and Technology, Berlin, Germany, ⁹Center for Stroke Research Berlin, Charité-Universitätsmedizin Berlin, Berlin, Germany, ¹⁰Johanna-Etienne-Hospital, Neuss, Germany

Stroke is a major cause of death or disability. As imaging-based patient stratification improves acute stroke therapy, dynamic susceptibility contrast magnetic resonance imaging (DSC-MRI) is of major interest in image brain perfusion. However, expert-level perfusion maps require a manual or semi-manual post-processing by a medical expert making the procedure time-consuming and less-standardized. Modern machine learning methods such as generative adversarial networks (GANs) have the potential to automate the perfusion map generation on an expert level without manual validation. We propose a modified pix2pix GAN with a temporal component (temp-pix2pix-GAN) that generates perfusion maps in an end-to-end fashion. We train our model on perfusion maps infused with expert knowledge to encode it into the GANs. The performance was trained and evaluated using the structural similarity index measure (SSIM) on two datasets including patients with acute stroke and the steno-occlusive disease. Our temp-pix2pix architecture showed high performance on the acute stroke dataset for all perfusion maps (mean SSIM 0.92–0.99) and good performance on data including patients with the steno-occlusive disease (mean SSIM 0.84–0.99).

While clinical validation is still necessary for future studies, our results mark an important step toward automated expert-level perfusion maps and thus fast patient stratification.

KEYWORDS

stroke, perfusion-weighted imaging, dynamic susceptibility contrast MR imaging (DSC-MR imaging), cerebrovascular disease, generative adversarial network (GAN)

1. Introduction

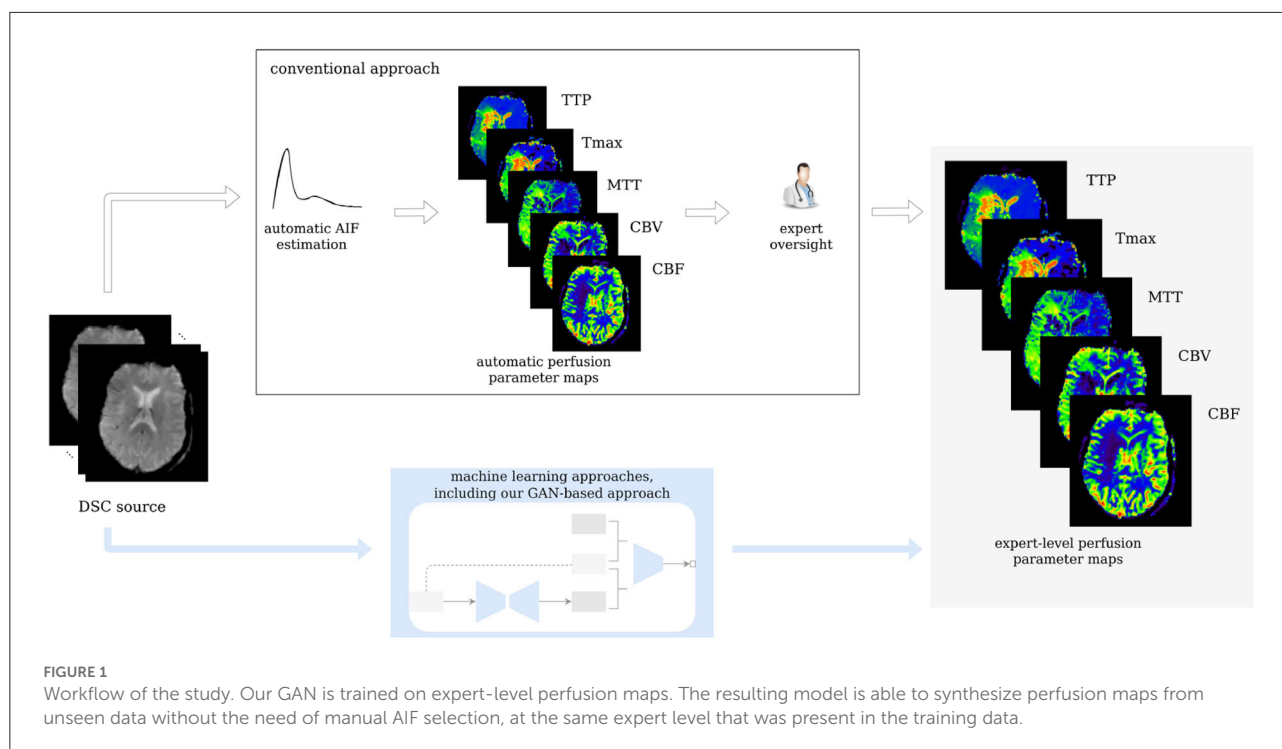
Ischemic stroke is a leading cause of death or disability worldwide.¹ In such a situation, time is brain (1). This requires rapid decision-making in the clinical setting to ensure an optimal outcome for an affected patient. Standard treatment strategies include recanalization by mechanical or pharmacological intervention, or a combination of both (2, 3). In this context, the eligibility of patients for treatment is mainly based on large cohorts of interventional trials that implement few imaging information (4, 5). However, this means that some patients will not receive treatment that would be beneficial for them, and conversely, some patients will be subjected to futile treatment attempts (6). An alternative approach to improve outcomes is individualized patient stratification based on specific patient characteristics (7, 8). One of the most important techniques for this approach is perfusion-weighted imaging, a special imaging technique used in both computed tomography (CT) and magnetic resonance imaging (MRI) (9). It provides highly relevant information about (patho)physiological blood flow in and around the ischemic brain tissue (10). In MRI, the most commonly used perfusion imaging technique is dynamic susceptibility contrast (DSC) MRI (11). It measures brain perfusion by injecting a gadolinium-based contrast agent into the patient's blood (11), followed by a series of T2- or T2*-weighted MRI sequences that record the flow of the contrast agent through the brain. The resulting four-dimensional image is deconvolved voxel-wise with an arterial input function (AIF) (12). The tissue concentration curve and the deconvolved curve result in interpretable perfusion parameter maps, such as the cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT), time-to-maximum (Tmax), and time-to-peak (TTP) (12). These maps are different representations of the information encoded in the time-intensity curve for each voxel. For all except TTP, to derive robust and valid parameter maps, the time-intensity curve must be deconvolved with an AIF (12). Ideally, the AIF is derived for each voxel separately, but in the clinical setting, the calculation of a global AIF is preferred (12).

The gold standard is the manual selection of several—usually 3 or 4—AIFs in the hemisphere contralateral to the stroke, from segments of the middle cerebral artery (12). The manual selection of AIFs is a tedious and time-consuming process that can only be performed after training (12). Therefore, automated methods whose results are subsequently reviewed by an expert are preferred in clinical practice (12).

However, the correct shape of the AIF has to be visually confirmed by an expert user. Otherwise, incorrect perfusion maps are generated, which can result in a serious diagnostic error (13). Therefore, there is a great clinical demand for novel automation approaches that provide validated, expert-level perfusion maps without the necessity for any expert oversight. This is especially important in smaller institutions where no expert is readily available and yet a quick decision must be made whether to treat a patient with an acute stroke or even transfer them for further endovascular therapy.

One possible solution is the application of modern artificial intelligence (AI) methods based on machine learning and in this study, particularly, deep learning approaches. These have shown great promise for solving medical imaging problems in the past few years (14, 15). This includes problems in stroke, such as stroke time onset prediction (16), lesion segmentation (17), and patient outcome prediction (18, 19). Among deep learning applications, generative adversarial networks (GANs) are particularly promising for the generation of expert-level perfusion maps. For example, GANs can be presented both with an original image and a processed image and learn to generate the processed image from the original. This is achieved by the special architecture of GANs, and they consist of two neural networks that try to fool each other (20). One network, the generator, synthesizes a data sample such as an image, whereas the other network, the discriminator, decides whether the sample looks like a real sample or not. At the end of the training, the generated sample should resemble the original as closely as possible. For image-to-image translations, GANs are considered to be state of the art in the medical field (21, 22), and a conditional GAN, such as the pix2pix GAN, can be applied (23). For example, pix2pix GANs have been successfully applied to transform MR images to CT images (cross-modal) or to transform 3T MR images to 7T MR images (intramodal) (24, 25).

1 WHO EMRO Stroke, Cerebrovascular Accident | Health Topics. Available online at: <http://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html>.



Given that the translation of a time series of perfusion information from source images to a single perfusion map can be seen as a highly similar medical image-to-image translation problem, GANs are a highly promising method for this use case. Preliminary work on GANs for the translation of time-series in dynamic cine applications has been published (26). Yet, to the best of our knowledge, no study has investigated the generation of DSC perfusion images from perfusion source data so far. The use of GANs would also present a new advantage: Since we use the final perfusion map for training, the GAN would not simply copy the map generation algorithm but would merge the map generation algorithm and the optimal AIF placement information into one algorithm. The GAN system would thus be able to generate perfusion maps even on images where manual AIF placement is not possible, such as due to motion artifacts, which are quite common in acute stroke (13).

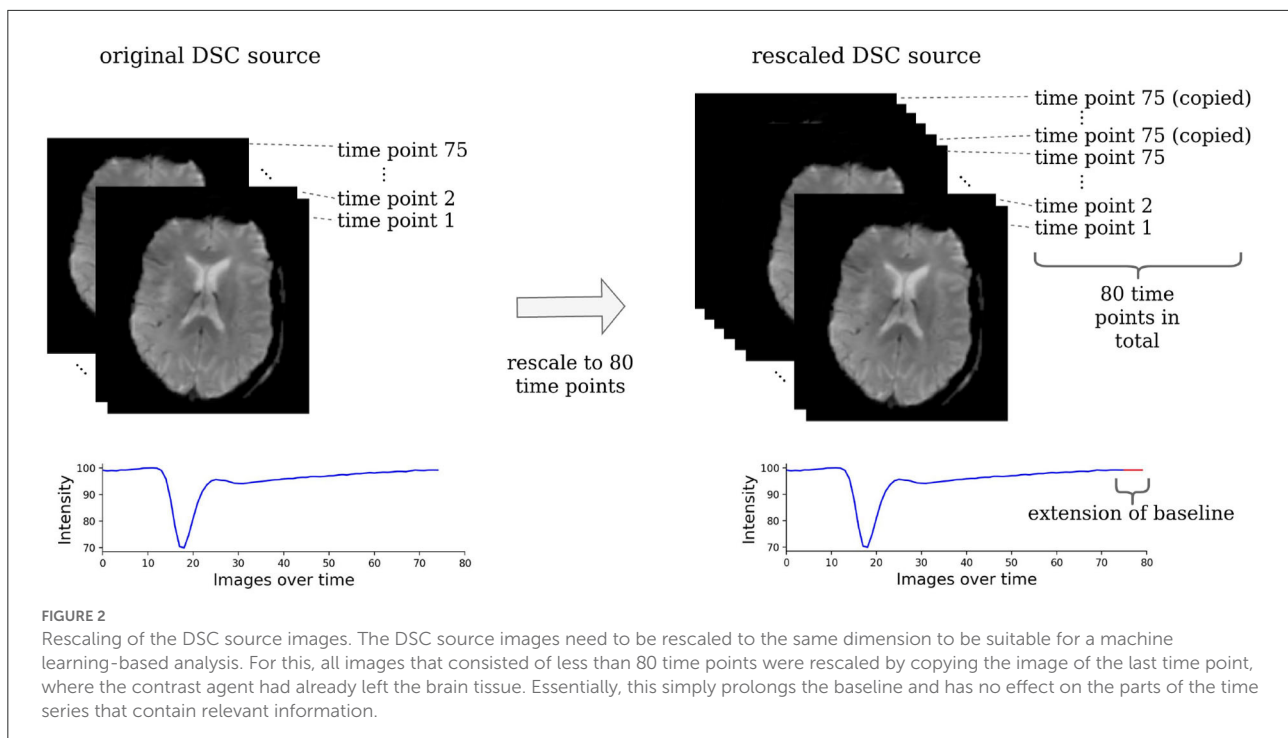
Thus, we propose a modified slice-wise pix2pix GAN with a temporal component (temp-pix2pix-GAN) to account for the time dimension in DSC source perfusion imaging. Our GAN model automatically generates perfusion parameter maps in an end-to-end fashion. We train our model on expert-level perfusion parameter maps (see Figure 1). The performance of our temp-pix2pix-GAN model is compared to a standard pix2pix GAN without a temporal component. We train and test our approach on two different datasets including patients with acute stroke and those chronic cerebrovascular disease, and perfusion data with motion artifacts.

2. Materials and methods

2.1. Data

In total, 276 patients were included in this study. Of which, 204 patients from a study performed at Heidelberg University Hospital suffered from acute stroke. Imaging was performed with a T2*-weighted gradient-echo EPI sequence with fat suppression TR = 2,220 ms, TE = 36 ms, flip angle 90°, field of view: 240 x 240 mm², image matrix: 128 x 128 mm, and 25–27 slices with ST of 5 mm and was started simultaneously with bolus injection of a standard dose (0.1 mmol/kg) of an intravenous gadolinium-based contrast agent on 3 Tesla MRI systems (Magnetom Verio, TIM Trio and Magnetom Prisma; Siemens Healthcare, Erlangen, Germany). In total, 50–75 dynamic measurements were performed (including at least eight prebolus measurements). Bolus and prebolus were injected with a pneumatically driven injection pump at an injection rate of 5 ml/s. The study protocol for this retrospective analysis of our prospectively established stroke database was approved by the Ethics Committee of Heidelberg University, and patient informed consent was waived.

A total of 72 patients with steno-occlusive disease were included in the PEGASUS study (27). A total of 80 whole-brain images were recorded using a single-shot FID-EPI sequence (TR = 1,390 ms, TE = 29 ms, and voxel size: 1.8 x 1.8 x 5 mm³) after automatic and synchronized injection of 5 ml Gadovist (Gadobutrol, 1 M, Bayer Schering Pharma AG, Berlin) followed



by 25 ml saline flush by a power injector (Spectris, Medrad Inc., Warrendale PA, USA) at a rate of 5 ml/s. The acquisition time was 1:54 min. All patients gave their written informed consent, and the study has been authorized by the Ethical Review Committee of Charité - Universitätsmedizin Berlin. DSC post-processing was performed blinded to clinical outcome.

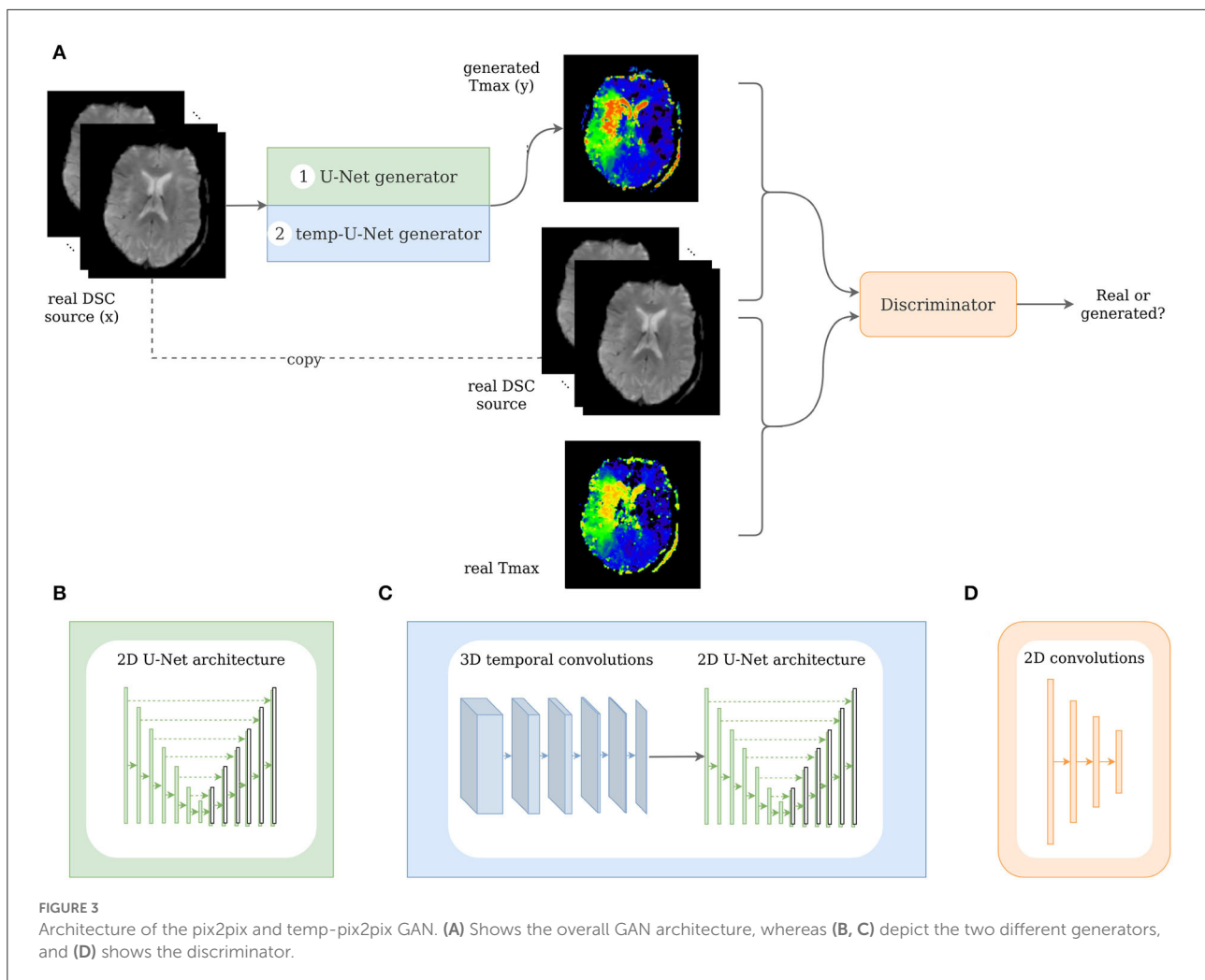
For the acute stroke data from Heidelberg, DSC data were post-processed with Olea Sphere[®] (Olea Medical, La Ciotat, France) in-house at the stroke center in Heidelberg, and automatic motion correction was applied. Raw DSC images were used to calculate perfusion maps of time-to-peak (TTP) from the tissue response curve. Maps of cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT), and time-to-maximum (Tmax) were created by deconvolution of a regional concentration-time curve with an arterial input function (AIF). Block-circulant singular value decomposition (cSVD) deconvolution was applied. The arterial input function (AIF) was detected automatically. All AIFs were visually inspected by a neuroradiology expert (MAM, with over 6 years of experience in perfusion imaging), and only in two cases, the automatically detected AIF needed to be manually corrected.

For PEGASUS patients, DSC data were post-processed with PGui software (version 1.0, provided for research purposes by the Center for Functional Neuroimaging, Aarhus University, Denmark). Motion correction was not available. Raw DSC images were used to calculate perfusion maps of TTP from the tissue response curve. Maps of CBF, CBV, MTT, and Tmax were

created by deconvolution of a regional concentration-time curve with an AIF. Parametric deconvolution was applied (28). For each patient, an AIF was determined by a junior rater (JB, with 2 years of experience in perfusion imaging) by manual selection of three or four intravascular voxels of the MCA M2 segment contralateral to the side of stenosis minimizing partial volume effects and bolus delay. The AIF shape was visually assessed for peak sharpness, bolus peak time, and amplitude width (12, 29). The AIFs were inspected by a senior rater (VIM, with over 12 years of experience in perfusion imaging).

The acute stroke data were resized to 21 horizontal slices each containing 128 x 128 voxels. To make the DSC source images suitable for our machine learning models, they were rescaled to 80 time points by copying the last slice until there were 80 slices in total (see Figure 2). All images of one parameter map and the DSC source images were normalized between -1 and $+1$ to ensure the stability of the GAN and then split into horizontal slices.

The post-processed data were split into training (acute stroke data: 142 patients, PEGASUS: 50 patients), validation (acute stroke data: 20 patients, PEGASUS: eight patients), and test (acute stroke data: 41 patients, PEGASUS: 12 patients) sets. Since the patient cohorts were different for the acute stroke dataset and the PEGASUS dataset consisting of patients with steno-occlusive disease, separate models were trained on the datasets. For optimization, hyperparameters such as the learning rate were selected based on visual inspection of the generated



images and the performance on the validation test (for details, see Section 2.4). The generalizable performance was estimated by the performance of the test set.

2.2. General methodological approach

We utilized a special type of AI model, which was developed for generating an image based on the input of another image: the pix2pix GAN (23). A pix2pix GAN consists of two neural networks that try to mislead each other. The first network, the generator, aims to produce realistic looking images based on another image (e.g., produce a CT based on an MR image), whereas, the second network, the discriminator, tries to distinguish between the generated image and real image. Based on the discriminator's feedback, both networks get better in their respective tasks.

Typically, the input and output to a pix2pix GAN generator are a 2D image. For this use case, we modified the pix2pix GAN to take a 3D image (time sequence of the 2D DSC

source image) as an input and synthesize the corresponding 2D perfusion map slice (e.g., Tmax). In this work, we implemented two different generator architectures. The first architecture, the classical pix2pix GAN, took in the 3D input image without accounting for the temporal relation between the images. In contrast to that, the second architecture, the temp-pix2pix GAN, was designed to first extract the temporal relation between the images followed by the transformation to the output image (see Figure 3). Both architectures were first trained on the acute stroke dataset. The resulting models served as weight initialization for the models trained on the PEGASUS dataset. In the following, the technical details of the two approaches are described in depth.

2.3. Network architecture

The GAN architecture was adapted from the pix2pix GAN (23). In our first architecture, we utilized the original U-Net generator as proposed in the paper with the time steps being

represented in the channels. For the second architecture, we modified the U-Net by adding 3D temporal convolutions before feeding the result into the U-Net in the generator (see [Figure 3](#)).

Both GAN architectures consisted of two neural networks: the generator G and the discriminator D . On the one hand, the generator's task was to synthesize perfusion parameter maps such as Tmax or CBF from the DSC source image. The discriminator, on the other hand, learned to distinguish between the real DSC source image together with the real perfusion parameter map and the real DSC with the generated perfusion parameter map.

In general, the objective function of a conditional GAN such as the pix2pix GAN is as follows:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))] \quad (1)$$

Where x is the input image (DSC source in our case) and y is the output image (for example, Tmax) and z is a noise vector, which is implicitly implemented as a dropout in the generator architecture (23). The generator tries to maximize the objective which is achieved when the discriminator outputs a high probability of the generated image pair being real and a low probability for the real image pair, respectively. In contrast to that, the discriminator tries to minimize this objective, and identify the real input images. The pix2pix GAN does not directly incorporate the noise vector z but introduces noise in the network using dropout in the generator.

The loss of the generator consisted of two parts. The first part was the adversarial loss, which took into account the feedback of the discriminator as described earlier. In addition, a reconstruction loss directly penalized deviation from the original image using the L1 norm. This second loss was added to the adversarial loss, and both were weighted by 1 after testing different weightings.

Two different GAN variants, the pix2pix GAN and the temp-pix2pix GAN, were implemented. They both differed in the generator's architecture and the generator's input representation.

The pix2pix generator was a 2D U-Net with six downsampling and upsampling layers (see [Figure 3B](#)). One DSC source slice at a time was fed as an input to the generator. The different time points of the DSC were concatenated in the channel dimension leading to 3D input data (channel, image height, and image width). Each downsampling layer consisted of a convolutional layer, batch normalization layer, and a LeakyReLU with slope 0.2, and the upsampling layers of ConvTranspose-layers, batch normalization, and a ReLU activation. After the last convolution, a tanh was applied.

In contrast to that, the generator of the temp-pix2pix GAN took one slice of the DSC source at all time points as an input. Here, the time points were represented in a separate time dimension leading to a 4D input (channel, image height, image width, and time). The time sequence of slices was then fed through six 3D convolutions over the time dimension iteratively

reducing this dimension to 1. Each convolutional layer was followed by a batch normalization layer and a LeakyReLU with slope 0.2. After the temporal path, the output is fed into a 2D U-Net with convolutions over the spatial dimensions with six downsampling and upsampling layers in an early fusion approach as shown in [Figure 3C](#).

The discriminator adapted the architecture of the discriminator from the PatchGAN as suggested by Isola et al. (23). It consisted of three convolutional layers with batch normalization and a LeakyReLU activation function followed by another convolutional layer and a sigmoid activation function (see [Figure 3D](#)). For both the generator and discriminator, the kernel size was four with strides of two.

2.4. Training

For each architecture, five GANs were trained on the acute stroke dataset from Heidelberg for each of the five parameter maps (CBF, CBV, MTT, Tmax, and TTP). The models were trained with a learning rate of 0.0001 for both the generator and discriminator using the Adam optimizer with $\beta_1 = 0.5$ and $\beta_2 = 0.999$. The networks were trained for 100 epochs at which point convergence of the two networks was achieved. The batch size was 4 and the dropout was 0. As the PEGASUS dataset was smaller, the models trained on the acute stroke dataset served as a weight initialization for the PEGASUS models and were then only fine-tuned for additional 50 epochs. Thus, in total, 10 models were trained per architecture.

All hyperparameters mentioned earlier were tuned and selected according to visual inspection of the generated images and the performance on the validation set. Due to the computational limitations, an automated search was not feasible. The code was implemented in PyTorch and is publicly available². The models were trained on a TESLA V100 GPU (NVIDIA Corporation, Santa Clara, CA, USA).

2.5. Performance evaluation

The generated images were first visually inspected. In addition, four metrics were applied: the mean absolute error (MAE) or L1 norm of the error, the normalized root mean squared error (NRMSE), the structural similarity index measure (SSIM), and the peak-signal-to-noise-ratio (PSNR). These metrics are standard performance measures for pairwise image comparison and were selected to better compare with existing studies (30).

The MAE is defined voxel-wise and measures the average absolute of the error between the real image y and the generated

² <https://github.com/prediction2020/DSC-to-perfusion>

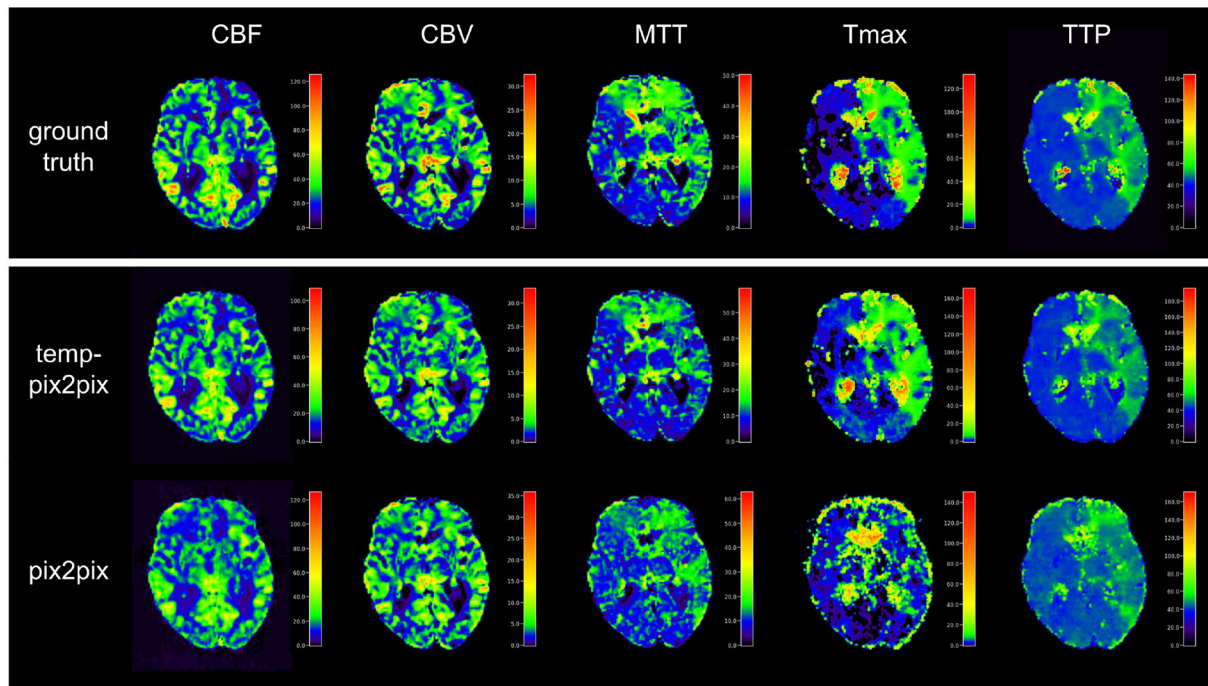


FIGURE 4 Synthesized perfusion parameter maps (middle, bottom rows) compared with the ground truth reviewed by an expert (top row) for one representative patient showing average performances from the acute stroke test dataset. The perfusion parameter maps generated by the temp-pix2pix all look similar to the ground truth, whereas the time-dependent parameters (Tmax and TTP) are not well captured by the pix2pix GAN.

image \hat{y} :

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{2}$$

The NRMSE is defined as the root mean squared error normalized by average euclidean norm of the true image y :

$$NRMSE = \frac{RMSE}{\sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2}} \tag{3}$$

with

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{4}$$

The SSIM is defined as a combination of luminance, contrast, and structure and can be summed up as follows:

$$SSIM(y, \hat{y}) = \frac{(2\mu_y\mu_{\hat{y}} + c_1)(2\sigma_{y\hat{y}} + c_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + c_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + c_2)}, \tag{5}$$

Where μ_y and $\mu_{\hat{y}}$ are the average values of y and \hat{y} , respectively, σ_y is the variance, and $\sigma_{y\hat{y}}$ is the covariance. c_1 and c_2 are constants for stabilization and defined as $c_1 = (k_1L)^2$ and $c_2 = (k_2L)^2$, with L being the dynamic range of the pixel values

and $k_1, k_2 \ll 1$ small constants. The higher the SSIM, the more similar are the two images with 1 denoting the highest similarity. The PSNR is defined as follows:

$$PSNR = 10 \log \left(\frac{MAX_I}{MSE} \right) \tag{6}$$

with

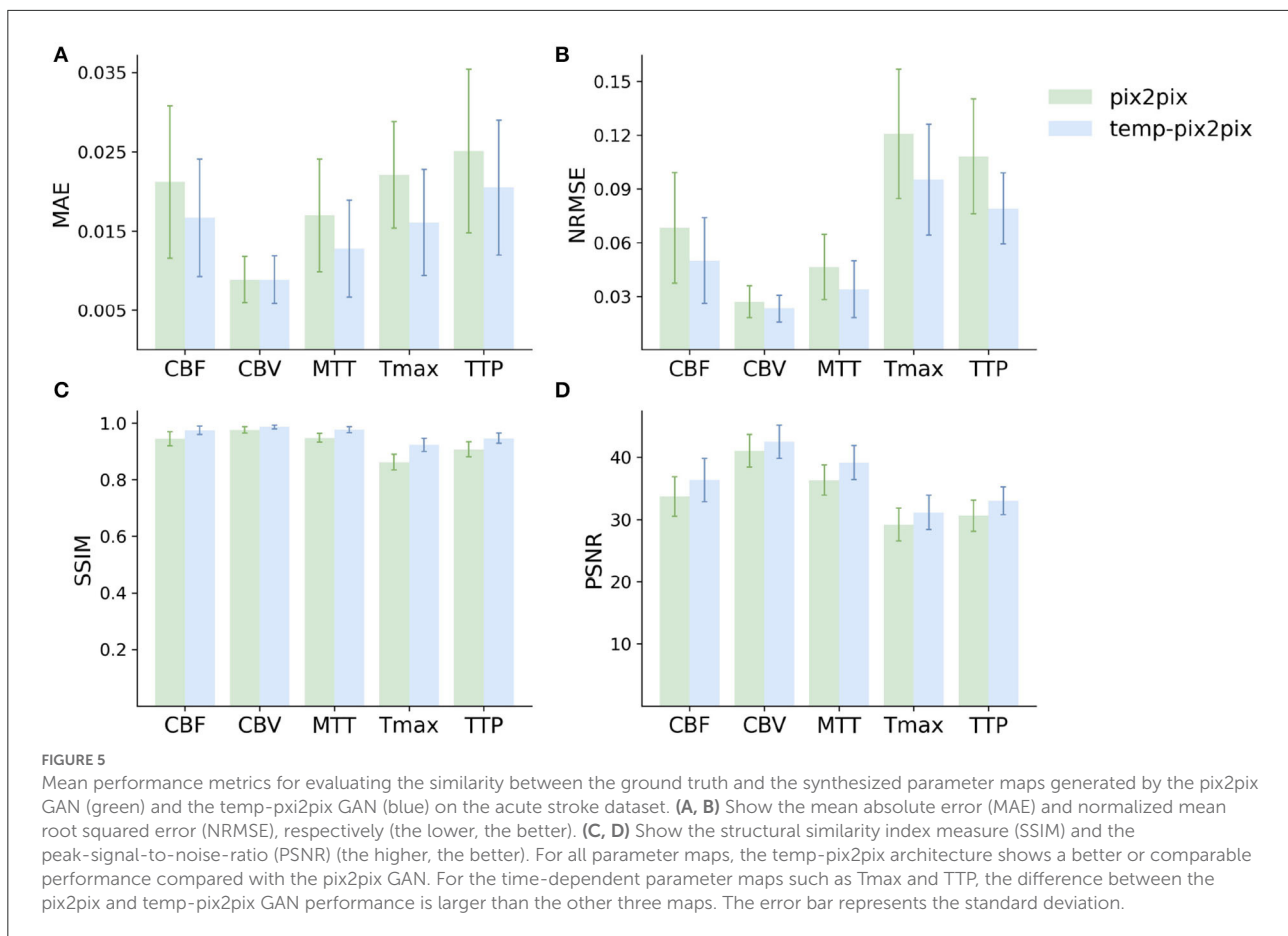
$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{7}$$

MAX_I is the maximal possible pixel/voxel value. It describes the ratio between the maximal possible signal power and noise power contained in the sample.

The performance metrics of each trained temp-pix2pix GAN were compared to the pix2pix GAN performance using the paired Wilcoxon signed-rank test. For $p < 0.05$, the difference between performances was considered statistically significant.

3. Results

Visual inspection of the results of the acute stroke dataset showed that the perfusion parameter maps generated by the temp-pix2pix GAN looked similar to the ground truth (see Figure 4). For the pix2pix model, on the contrary, only CBF,



CBV, and MTT were of sufficient quality, whereas the time-dependent parameters TTP and Tmax did not consistently resemble the ground truth (also Figure 4).

The quantitative analysis in the acute stroke dataset revealed for all parameter maps a high SSIM ranging from 0.92 to 0.99 for the temp-pix2pix model (Figure 5). In contrast to this, the pix2pix GAN showed a comparable or worse SSIM ranging from 0.86 to 0.98. A performance difference between the pix2pix and temp-pix2pix models was especially prominent for Tmax and TTP (SSIM 0.92 vs. 0.86 and 0.95 vs. 0.91, respectively). The temp-pix2pix model showed a significantly better performance than the pix2pix model ($p < 0.05$) for all metrics and perfusion maps, except for the MAE of CBV (MAE of 0.009 for both models, $p = 0.53$).

For the PEGASUS dataset, the perfusion maps generated by both the fine-tuned pix2pix and temp-pix2pix GAN look similar to the ground truth (see Figure 6). For both networks, MTT appeared to be the least well-reconstructed parameter map, which is also reflected in the metrics (Figure 7). Furthermore, the high intensities of Tmax were not well-captured by the pix2pix GAN (Figure 6). The performance metrics of the pix2pix and temp-pix2pix GAN and the ground truth for the PEGASUS

dataset showed a low error and high SSIM and PSNR for CBF, CBV, and Tmax. Here, for most metrics, the temp-pix2pix GAN achieved a slightly better performance in contrast to the pix2pix GAN. For MTT and TTP, the temp-pix2pix showed a better performance compared with the pix2pix GAN (SSIM 0.84 vs. 0.78 and 0.86 vs. 0.82, respectively). Overall, the metrics of the synthesized MTT and TTP maps obtained a worst performance compared with the other parameter maps. Again, the temp-pix2pix model showed a significantly better performance than the pix2pix model ($p < 0.05$) for most metrics and perfusion maps, except for the PSNR and NRMSE of Tmax (PSNR of 37.28 vs. 37.83, $p = 0.29$; NRMSE of 0.032 vs. 0.031, $p = 0.10$).

Figure 8A shows two patients whose generated parameters showed the worst performance. For the acute stroke dataset, these are two Tmax maps (Figure 8A, first and second columns), whereas, the generated Tmax in the first column did not capture the high intensities well, the generated map in the second column visually looked well. For the PEGASUS models, MTT performed the worst (Figure 8A, third and fourth columns). In the third column, the generated MTT appears less noisy than the ground truth. In contrast to that, in the fourth column, the generated MTT map looked noisier compared with the ground

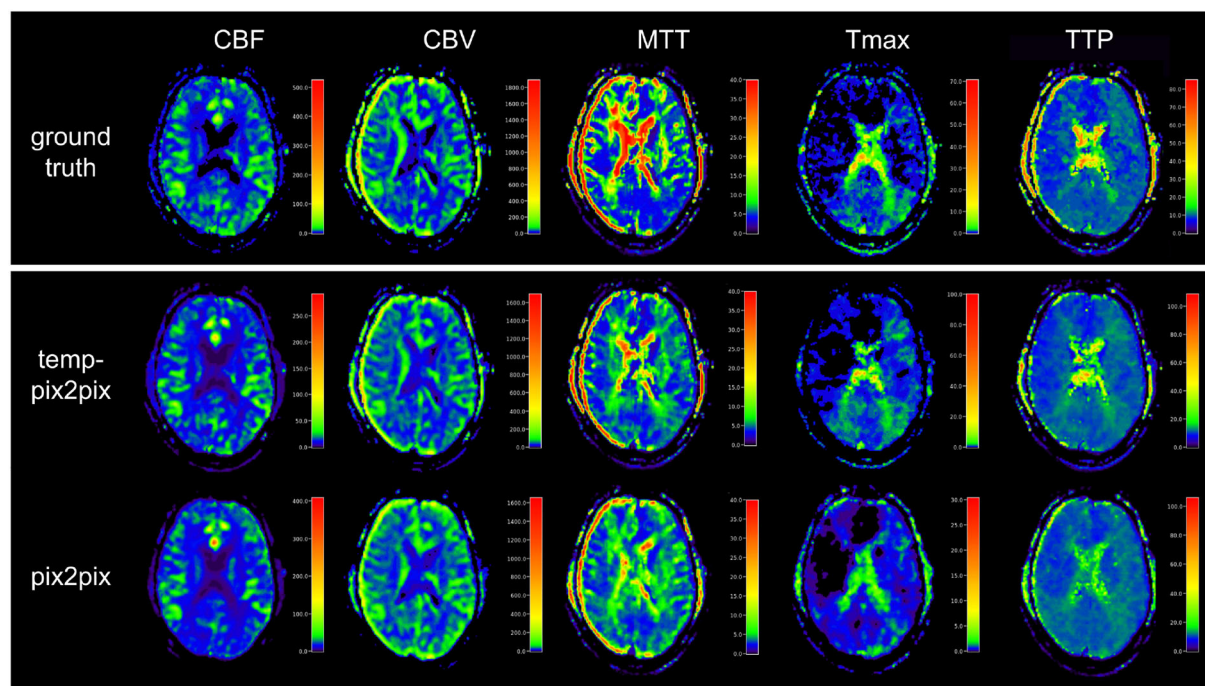


FIGURE 6
Synthesized perfusion parameter maps (**middle, bottom rows**) compared with the ground truth reviewed by an expert (**top row**) for one representative patient showing average performances from the PEGASUS test dataset. Both pix2pix and temp-pix2pix GAN synthesized most parameter maps that resemble the ground truth. Parts of MTT were not entirely captured by pix2pix and temp-pix2pix. Moreover, the pix2pix GAN did not synthesize the higher intensities of Tmax well. For MTT and Tmax, the temp-pix2pix GAN showed better performance in all metrics compared with the pix2pix GAN.

truth. [Figure 8B](#) shows that the Tmax maps generated by the temp-pix2pix and pix2pix GAN for four patients for which an AIF could not be placed.

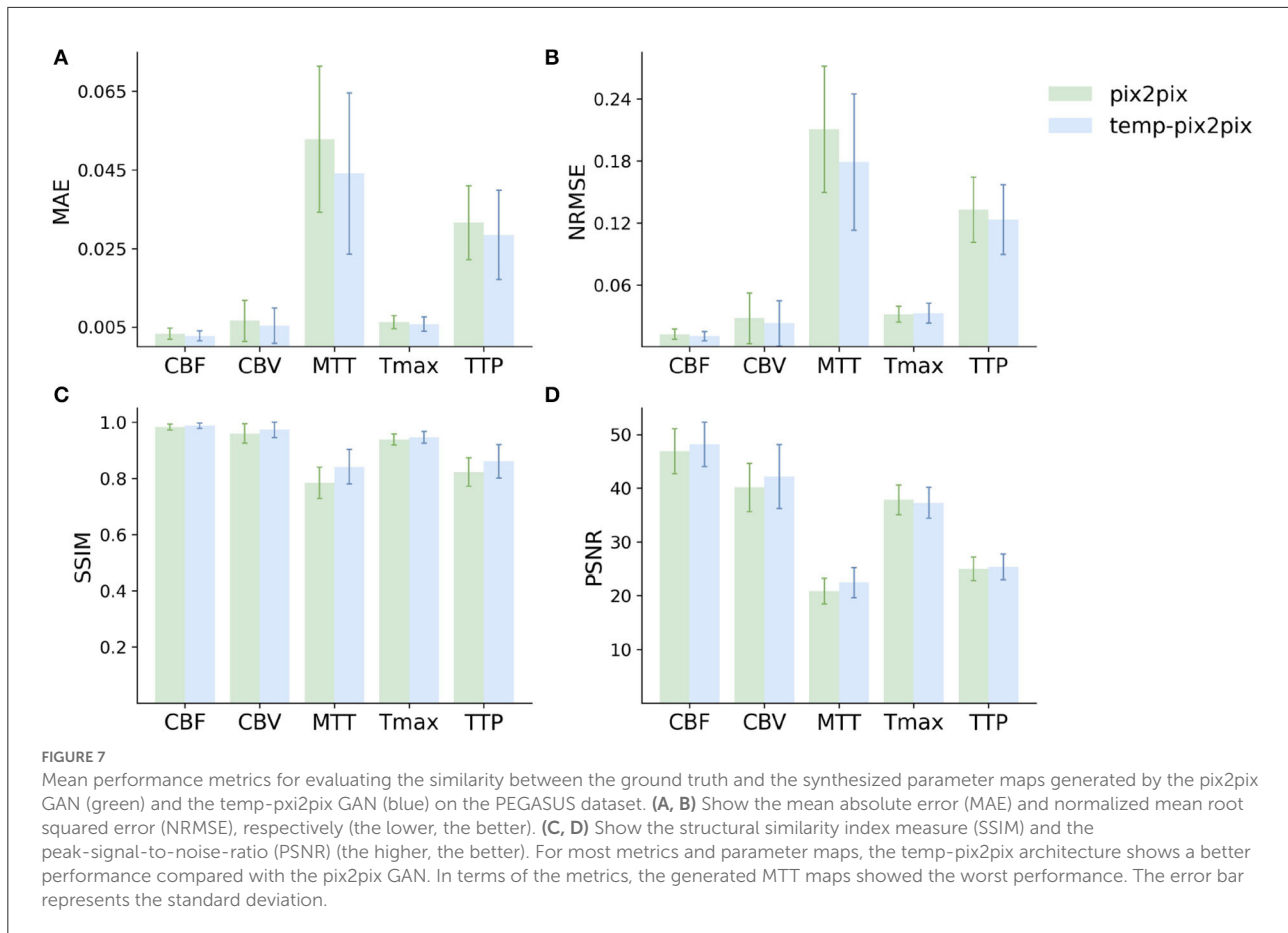
4. Discussion

In the present study, we propose a novel pix2pix GAN variant with temporal convolutions-coined temp-pix2pix-to generate expert-level perfusion parameter maps from DSC-MR images in an end-to-end fashion for the first time. The temp-pix2pix architecture showed high performance in a dataset of patients with acute stroke and good performance on data of patients with the chronic steno-occlusive disease. Our results mark a decisive step toward the automated generation of expert-level DSC perfusion maps for acute stroke and their application in the clinical setting.

This requires rapid decision-making in the clinical setting to ensure an optimal outcome for an affected patient. While automated methods for parameter map generation have shown inconclusive results in the literature (26, 31–35), they are successfully used in acute stroke to identify stroke-affected

tissue. In our sample, this was confirmed as the automatically derived arterial input functions only required expert adjustment in 2 out of 204 patients in the acute stroke set. Nevertheless, this approach still requires a manual check resulting in a time delay of a few minutes per patient before patient stratification. As a consequence, there is a major clinical need for automated methods that provide final perfusion parameter maps without any manual input. Here, we chose a GAN AI approach, as presenting this methodology expert-level perfusion maps would lead to a model after training that could then *generate* expert-level perfusion maps, implicitly encoding the choice of AIFs within ~ 1.8 s per patient (speed of our GAN). Our exploratory results show that this approach was successful.

This may have a positive impact on the clinical setting. First, it would eliminate the need for a manual review of AIFs. This would reduce the time needed to calculate perfusion parameter maps and also reduce resource requirements as radiologists and neurologists would no longer need to be trained on how to identify optimal AIFs. Second, as we have shown, it is even possible to calculate parameter maps for patients who currently have to be excluded due to motion artifacts that make it impossible for the standard software to calculate the parameter



maps. At this point, it is important to emphasize that our study is exploratory, and the generated model is only used for internal research purposes. This is due to the fact that the generative AI has fundamentally learned to approximate the non-AI algorithm that was originally used to calculate the perfusion parameter maps. To maximize clinical impact, we thus encourage the developers and vendors of relevant clinically used perfusion software to consider adding GAN-based automated perfusion calculation modules to their products. To facilitate this process, we have made our code publicly available.

One of the most important contributions of our approach was the consideration of the temporal dimension of the time series input. Not surprisingly, the temp-pix2pix architecture performed better than the pix2pix GAN without a temporal component in both datasets. This was particularly noticeable in the acute stroke dataset for parameters directly related to the correct order of the time-intensity curve, namely TTP and Tmax. Maps of CBF, CBV, and MTT (derived by the central volume theorem as CBV/CBF) also performed quite well in the baseline architecture without a temporal component, as for these maps, the order of input is not relevant. This is because CBV corresponds to the area under the time-intensity curve and

CBF is calculated based on the height of the slope, which are indifferent to the order. In the chronic stroke dataset, the temp-pix2pix also outperformed the baseline GAN without a temporal component. However, the difference in performance was not as pronounced as in the acute stroke dataset. This could be due to the fact that patients with acute vascular obstruction usually have significantly higher delays than patients with chronic stenocclusive disease, and the performance advantage of temp-pix2pix increases with increasing delay. It is noteworthy that in contrast to the patients with acute stroke in the chronic stenocclusive cohort, MTT and TTP maps performed worse than the other parameter maps. This might be related to the more complex perfusion pathophysiology in chronic stenocclusive disease. Whereas in acute stroke, delay is the main contributor to blood flow abnormalities, and in chronic stenocclusive disease, it is the sum of delay and considerable dispersion due to vessel abnormalities (36). This could pose particular difficulties for neural networks to learn the relationships required to create parameter maps: MTT is a parameter that depends on two other parameters (CBV and CBF) in the original software solutions, which are likely to have greater variability in chronic stenocclusive disease. In addition, TTP delays are attributable to

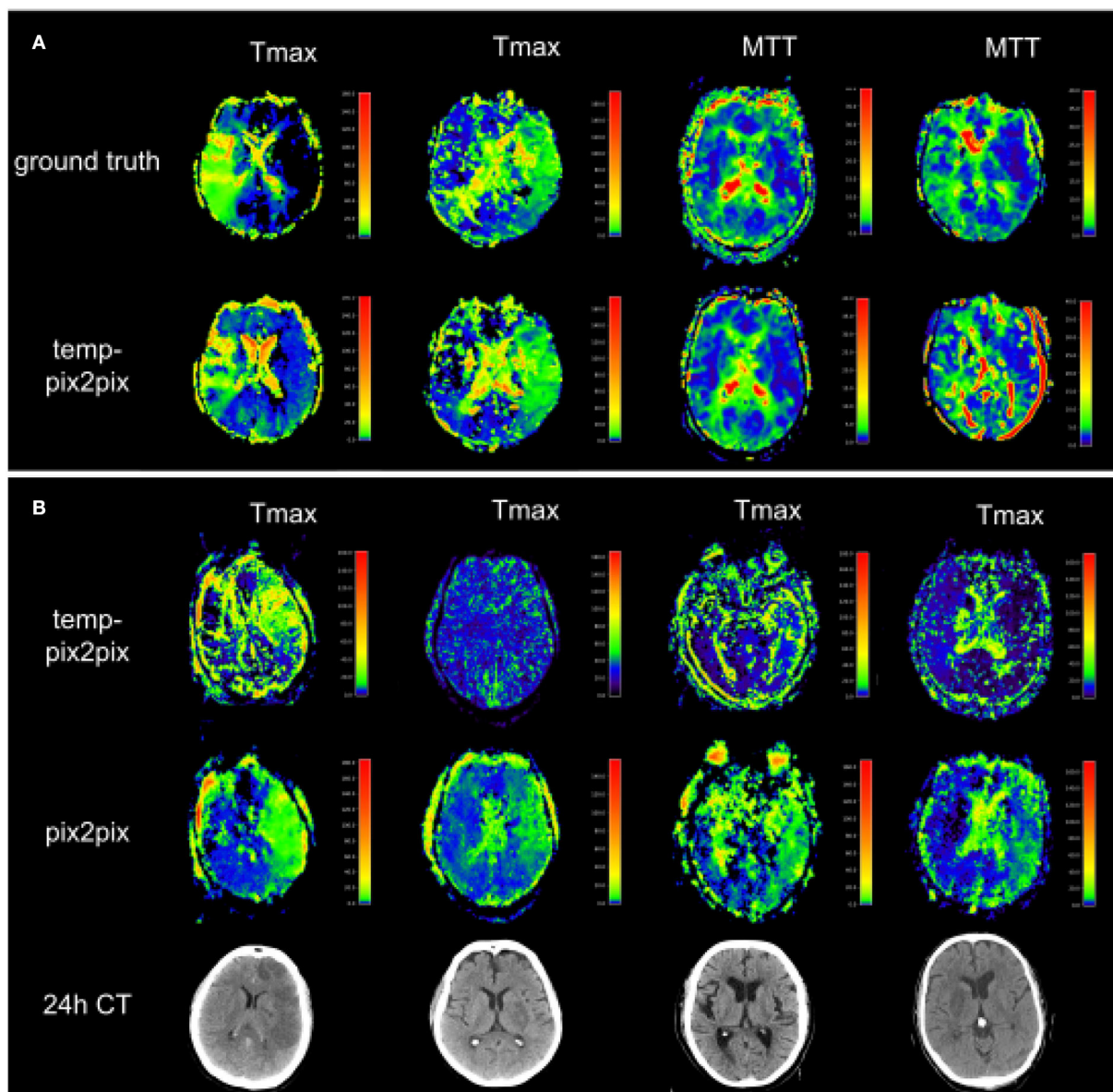


FIGURE 8

Two patients with the poorest performance according to the metrics for each of the two datasets (A) and patients for which no AIF could be computed (B). (A) The first and second columns show Tmax for two patients with acute stroke. Whereas the synthesized image in the first column does not fully capture the hypoperfused areas, the generated image in the second column looks quite close to the ground truth. Columns three and four show MTT for two PEGASUS patients. While the generated image in the third column shows less noise than the ground truth, the GAN introduced noise in the fourth column in the synthesized image. (B) Four Tmax maps generated by temp-pix2pix (upper row) and pix2pix (lower row) for cases from the acute stroke data for which no AIF could be computed, and thus with conventional methods not imaging would be available. To get an estimate of the true lesion, 24 h CT is plotted in the lower row. Note that for the CT images, only 2D visualizations were available. Thus, they are not aligned with the perfusion parameter map. Since motion artifacts affect the quality of the time series, in these cases the baseline pix2pix performs better than the temp-pix2pix.

both delay and dispersion, with varying weights in individual patients leading again to a larger variability (this effect is much less pronounced in Tmax parameter maps due to the deconvolution procedure). Such increased variability might lead to less stable models and thus increased noise in the generated maps.

Our work is the first work to utilize GANs to create perfusion parameter maps in DSC imaging. A few works exist that used different machine learning and deep learning methods to generate parameter perfusion maps from the DSC source image. For instance, McKinley et al. (37) used several classical voxel-wise machine learning approaches to generate manually

validated perfusion parameter maps and identified a tree-based algorithm as the best performing model. Their best results for Tmax achieved a lower performance with an NRMSE of 0.113 compared with our best model with an NRMSE of 0.095. Vialard et al. (38) suggested a deep learning-based spatiotemporal U-net approach for translating DSC-MR patches to CBV maps in patients with brain tumors. With an SSIM of 0.821, their generated CBV maps obtained a poor performance compared to our CBV generated by the temp-pix2pix model, with a SSIM of 0.986. In the field of stroke, Ho et al. (39) proposed a patch-based deep learning approach to generate CBE, CBV, MTT, and Tmax. The average RMSE for their generated Tmax showed a higher error of 1.33 compared with ours with 0.06. Hess et al. (40) utilized a different voxel-wise deep learning approach to approximate Tmax from DSC-MR. This approach was clinically evaluated in another study (41). In a study by Hess et al. (40), they reported the performance in terms of MAE with clipping to not account for noise. The generated Tmax achieved an MAE with clipping of 0.524 compared to our approach showing an MSE of 0.016. These differences compared to our study might be due to the novel use of the GAN method and the fact that our model considered whole slices instead of patches to better account for the spatial dimension. Of course, these comparisons need to be additionally evaluated as they might depend on differences in training and test data.

It is also worth noting that AIFs are known to exhibit considerable variation (42, 43) and the predictive performance of perfusion image maps can be influenced by the AIF shape (44). We performed a cursory visual inspection of the AIFs of the patients in the test set of our study but could not find any striking correlation between the AIF shape and the performance metrics of a generated image. However, all AIFs in our study were selected and reviewed by well-trained and experienced staff, so differences between AIFs could be rather subtle. Therefore, a potential relationship between the shape of the AIF and the performance of the generated image can probably only be detected in a quantitative analysis where the AIFs are parameterized and numerically compared with performance metrics. This is beyond the scope of our study but is a very interesting approach for further research.

Our study has several limitations. First, our network was based on 2D slices instead of the full 3D volumes due to computational restrictions. It is likely that the results could be improved further using the full 3D images. Second, our study is an exploratory hypothesis generating study. Its results need to be clinically validated in a future study before integrating into clinical practice would be possible. This includes clinical evaluation metrics beyond voxel-wise comparison such as comparing manually segmented lesion volumes. Furthermore, due to data availability, this analysis was only performed on MR data and did not include CT perfusion images. Lastly, our approach so far is a black-box approach. It could be extended with explainable AI to generate insights into which areas in

the source images are particularly relevant for the creation of different perfusion parameter maps. This could further elucidate the causes of the performance differences between the maps that we identified and could guide the way for further improvements.

5. Conclusion

We generated expert-level perfusion parameter maps using a novel GAN approach showcasing that AI approaches might have the ability to overcome the need for oversight by medical experts. Our exploratory study paves the way for fully-automated DSC-MR processing for faster patient stratification in acute stroke. In the clinical setting where time is crucial for patient outcome, this could have a big impact on standardized patient care in acute stroke.

Data availability statement

The datasets presented in this article are not readily available because data protection laws prohibit sharing of the PEGASUS and acute stroke datasets at the current time point. Requests to access these datasets should be directed to the Ethical Review Committee of Charite Universitätsmedizin Berlin, ethikkommission@charite.de.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Heidelberg University and Ethical Review Committee of Charité-Universitätsmedizin Berlin. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TK designed and conceptualized study, implemented the networks and ran the experiments, interpreted the results, and drafted the manuscript. VM collected data, designed and conceptualized study, interpreted the results, and drafted the manuscript. MM collected and processed data, interpreted the results, and revised the manuscript. AHe and KH advised on methods and revised the manuscript. JB processed data and revised the manuscript. CA advised on conceptualization and revised the manuscript. AHi revised the manuscript. JS and MB collected data and revised the manuscript. DF acquired funding, advised on conceptualization, and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work has received funding from the European Commission through a Horizon2020 grant (PRECISE4Q Grant No. 777 107, coordinator: DF) and the German Federal Ministry of Education and Research through a Go-Bio grant (PREDICTioN2020 Grant No. 031B0154 lead: DF).

Acknowledgments

Computation has been performed on the HPC for Research cluster of the Berlin Institute of Health.

Conflict of interest

TK, VM, and AH reported receiving personal fees from ai4medicine outside the submitted work. While not related

to this work, JS reports receipt of speakers' honoraria from Pfizer, Boehringer Ingelheim, and Daiichi Sankyo. DF reported receiving grants from the European Commission, reported receiving personal fees from and holding an equity interest in ai4medicine outside the submitted work.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Saver JL. Time is brain—quantified. *Stroke*. (2006) 37:263–6. doi: 10.1161/01.STR.0000196957.55928.ab
- Berge E, Whiteley W, Audebert H, De Marchis G, Fonseca AC, Padiglioni C, et al. European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute Ischaemic stroke. *Eur Stroke J*. (2021) 6:1–LXII. doi: 10.1177/2396987321989865
- Turc G, Bhogal P, Fischer U, Khatri P, Lobotesis K, Mazighi M, et al. European stroke organisation (ESO)-European Society for Minimally Invasive Neurological Therapy (ESMINT) Guidelines on Mechanical Thrombectomy in Acute Ischaemic Stroke Endorsed by Stroke Alliance for Europe (SAFE). *Eur Stroke J*. (2019) 4:6–12. doi: 10.1177/2396987319832140
- Lin CH, Saver JL, Ovbiagele B, Huang WY, Lee M. Endovascular thrombectomy without versus with intravenous thrombolysis in acute ischemic stroke: a non-inferiority meta-analysis of randomized clinical trials. *J Neurointervent Surg*. (2022) 14:227–32. doi: 10.1136/neurintsurg-2021-017667
- McDermott M, Skolarus LE, Burke JF. A systematic review and meta-analysis of interventions to increase stroke thrombolysis. *BMC Neurol*. (2019) 19:86. doi: 10.1186/s12883-019-1298-2
- Goyal M, Menon BK, van Zwam WH, Dippel DWJ, Mitchell PJ, Demchuk AM, et al. Endovascular thrombectomy after large-vessel Ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet*. (2016) 387:1723–31. doi: 10.1016/S0140-6736(16)00163-X
- Rehani B, Ammanuel SG, Zhang Y, Smith W, Cooke DL, Hetts SW, et al. A new era of extended time window acute stroke interventions guided by imaging. *Neurohospitalist*. (2020) 10:29–37. doi: 10.1177/1941874419870701
- Sharobeam A, Yan B. Advanced imaging in acute ischemic stroke: an updated guide to the hub-and-spoke hospitals. *Curr Opin Neurol*. (2022) 35:24–30. doi: 10.1097/WCO.0000000000001020
- Wintermark M, Albers GW, Broderick JP, Demchuk AM, Fiebach JB, Fiehler J, et al. Acute stroke imaging research roadmap II. *Stroke*. (2013) 44:2628–39. doi: 10.1161/STROKEAHA.113.002015
- Copen WA, Schaefer PW, Wu O. MR perfusion imaging in acute ischemic stroke. *Neuroimaging Clin North Am*. (2011) 21:259–83. doi: 10.1016/j.nic.2011.02.007
- Jahng GH, Li KL, Ostergaard L, Calamante F. Perfusion magnetic resonance imaging: a comprehensive update on principles and techniques. *Korean J Radiol*. (2014) 15:554. doi: 10.3348/kjr.2014.15.5.554
- Calamante F. Arterial input function in perfusion MRI: A comprehensive review. *Progr Nuclear Magn Reson Spectrosc*. (2013) 74:1–32. doi: 10.1016/j.pnmrs.2013.04.002
- Potreck A, Seker F, Mutke MA, Weyland CS, Herweh C, Heiland S, et al. What is the impact of head movement on automated CT perfusion mismatch evaluation in acute ischemic stroke? *J Neurointervent Surg*. (2022) 14:628–33. doi: 10.1136/neurintsurg-2021-017510
- Wernick MN, Yang Y, Brankov JG, Yourganov G, Strother SC. Machine learning in medical imaging. *IEEE Signal Process Mag*. (2010) 27:25–38. doi: 10.1109/MSP.2010.936730
- Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*. (2019) 29:102–27. doi: 10.1016/j.zemedi.2018.11.002
- Ho KC, Speier W, El-Saden S, Arnold CW. Classifying acute ischemic stroke onset time using deep imaging features. *AMIA Ann Symp Proc*. (2018) 2017:892–901.
- Chen L, Bentley P, Rückert D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. *Neuroimage Clin*. (2017) 15:633–43. doi: 10.1016/j.nicl.2017.06.016
- Nielsen A, Hansen MB, Tietze A, Mouridsen K. Prediction of tissue outcome and assessment of treatment effect in acute ischemic stroke using deep learning. *Stroke*. (2018) 49:1394–401. doi: 10.1161/STROKEAHA.117.019740
- Kamal H, Lopez V, Sheth SA. Machine learning in acute ischemic stroke neuroimaging. *Front Neurol*. (2018) 9:945. doi: 10.3389/fneur.2018.00945
- Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. *arXiv:1406.2661 [cs, stat]*. (2014). doi: 10.48550/arXiv.1406.2661
- Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal*. (2019) 58:101552. doi: 10.1016/j.media.2019.101552
- Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv:170310593 [cs]*. (2020). doi: 10.48550/arXiv.1703.10593
- Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. *arXiv:161107004 [cs]*. (2018). doi: 10.1109/CVPR.2017.632
- Brou Boni KND, Klein J, Vanquin L, Wagner A, Lacornerie T, Pasquier D, et al. MR to CT synthesis with multicenter data in the pelvic area using a conditional generative adversarial network. *Phys Med Biol*. (2020) 65:075002. doi: 10.1088/1361-6560/ab7633

25. Nie D, Trullo R, Lian J, Wang L, Petitjean C, Ruan S, et al. Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans Biomed Eng.* (2018) 65:2720–30. doi: 10.1109/TBME.2018.2814538
26. Ghodrati V, Bydder M, Bedayat A, Prosper A, Yoshida T, Nguyen KL, et al. Temporally aware volumetric generative adversarial network-based MR image reconstruction with simultaneous respiratory motion compensation: initial feasibility in 3D dynamic cine cardiac MRI. *Magn Reson Med.* (2021) 86:2666–83. doi: 10.1002/mrm.28912
27. Mutke MA, Madai VI, von Samson-Himmelstjerna FC, Zaro Weber O, Revankar GS, Martin SZ, et al. Clinical evaluation of an arterial-spin-labeling product sequence in steno-occlusive disease of the brain. *PLoS ONE.* (2014) 9:e87143. doi: 10.1371/journal.pone.0087143
28. Mouridsen K, Hansen MB, Østergaard L, Jespersen SN. Reliable estimation of capillary transit time distributions using DSC-MRI. *J Cereb Blood Flow Metab.* (2014) 34:1511–21. doi: 10.1038/jcbfm.2014.111
29. Thijs VN, Somford DM, Bammer R, Robberecht W, Moseley ME, Albers GW. Influence of arterial input function on hypoperfusion volumes measured with perfusion-weighted imaging. *Stroke.* (2004) 35:94–98. doi: 10.1161/01.STR.0000106136.15163.73
30. Nečasová T, Burgos N, Svoboda D. Chapter 25 - Validation and evaluation metrics for medical and biomedical image synthesis. In: Burgos N, Svoboda D, editors. *Biomedical Image Synthesis and Simulation. The MICCAI Society Book Series.* Academic Press (2022). p. 573–600. Available online at: <https://www.sciencedirect.com/science/article/pii/B9780128243497000323>
31. Hansen MB, Nagenthiraja K, Ribe LR, Dupont KH, Østergaard L, Mouridsen K. Automated estimation of salvageable tissue: comparison with expert readers. *J Magn Reson Imaging.* (2016) 43:220–8. doi: 10.1002/jmri.24963
32. Galinovic I, Ostwaldt AC, Soemmer C, Bros H, Hotter B, Brunecker P, et al. Automated vs manual delineations of regions of interest- a comparison in commercially available perfusion MRI software. *BMC Med Imaging.* (2012) 12:16. doi: 10.1186/1471-2342-12-16
33. Pistocchi S, Strambo D, Bartolini B, Maeder P, Meuli R, Michel P, et al. MRI software for diffusion-perfusion mismatch analysis may impact on patients' selection and clinical outcome. *Eur Radiol.* (2022) 32:1144–53. doi: 10.1007/s00330-021-08211-2
34. Deutschmann H, Hinteregger N, Wiefßpeiner U, Kneihsl M, Fandler-Höfler S, Michenthaler M, et al. Automated MRI perfusion-diffusion mismatch estimation may be significantly different in individual patients when using different software packages. *Eur Radiol.* (2021) 31:658–65. doi: 10.1007/s00330-020-07150-8
35. Ben Alaya I, Limam H, Kraiem T. Applications of artificial intelligence for DWI and PWI data processing in acute ischemic stroke: current practices and future directions. *Clin Imaging.* (2022) 81:79–86. doi: 10.1016/j.clinimag.2021.09.015
36. Calamante F, Willats L, Gadian DG, Connelly A. Bolus delay and dispersion in perfusion MRI: implications for tissue predictor models in stroke. *Magn Reson Med.* (2006) 55:1180–5. doi: 10.1002/mrm.20873
37. McKinley R, Hung F, Wiest R, Liebeskind DS, Scalzo F. A machine learning approach to perfusion imaging with dynamic susceptibility contrast MR. *Front Neurol.* (2018) 9:717. doi: 10.3389/fneur.2018.00717
38. Vialard JV, Rohé MM, Robert P, Nicolas F, Bône A. *Going Beyond Voxelise Deconvolution in Perfusion MRI: Learning and Leveraging Spatio-Temporal Regularities With the stU-Net.* (2021). p. 6.
39. Ho KC, Scalzo F, Sarma KV, El-Saden S, Arnold CW. A temporal deep learning approach for MR perfusion parameter estimation in stroke. In: *2016 23rd International Conference on Pattern Recognition (ICPR).* Cancun: IEEE (2016). p. 1315–20.
40. Hess A, Meier R, Kaesmacher J, Jung S, Scalzo F, Liebeskind D, et al. Synthetic perfusion maps: imaging perfusion deficits in DSC-MRI with deep learning. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Lecture Notes in Computer Science.* Cham: Springer International Publishing (2019). p. 447–55.
41. Meier R, Lux P, Med B, Jung S, Fischer U, Gralla J, et al. Neural network-derived perfusion maps for the assessment of lesions in patients with acute ischemic stroke. *Radiol Artif Intell.* (2019) 1:e190019. doi: 10.1148/ryai.2019190019
42. Meijs M, LMAGCF Christensen S. Simulated perfusion MRI data to boost training of convolutional neural networks for lesion fate prediction in acute stroke. *Magn Reson Med.* (2015) 76:1282–90.
43. Livne M, Madai VI, Brunecker P, Zaro-Weber O, Moeller-Hartmann W, Heiss WD, et al. A PET-guided framework supports a multiple arterial input functions approach in DSC-MRI in acute stroke. *J Neuroimaging.* (2017) 27:486–92. doi: 10.1111/jon.12428
44. Debs N, Rasti P, Victor L, Cho TH, Frindel C, Rousseau D. Simulated perfusion MRI data to boost training of convolutional neural networks for lesion fate prediction in acute stroke. *Comput Biol Med.* (2020) 116:103579. doi: 10.1016/j.combiomed.2019.103579