This study explores the process of generating perfusion parameter maps from time series of DSC-MRI images of the brain. The task addressed relates to the inaccuracy and inefficiency of perfusion map generation due to insufficient analysis of temporal features, dependence on the arterial inflow function, as well as excessive computational complexity of the models.

This paper proposes an approach to generating perfusion parameter maps using neural networks that directly process time series of perfusion images. Three deep learning architectures have been designed and experimentally investigated, differing in the way in which temporal information is taken into account and the presence of recurrent layers. The evaluation was performed on an open medical data set with patient-level sample distribution to prevent information leakage between the training and test parts.

A controlled comparison of models with and without recurrent layers was performed to determine the impact of explicitly taking into account temporal dependences on the accuracy of map generation. The normalized root mean square error decreased from 0.027 to 0.016 and 0.015. The structural similarity index increased from 0.826 to 0.957 and 0.973. The peak signal-to-noise ratio increased from 31.493 to 36.095 and 36.412.

Additional comparison to approaches reported in other studies showed that the proposed architecture with recurrent layers demonstrates competitive or higher values of image quality metrics. The results confirm the feasibility of using neural networks with recurrent layers for more accurate generation of perfusion parameter maps.

The practical significance is the possibility of integrating the approach into automated perfusion analysis systems and clinical decision support in the diagnosis of stroke and brain tumors

Keywords: perfusion parameters, perfusion with dynamic susceptibility contrast, magnetic resonance imaging

UDC 004.932:616-073.756.8
DOI: 10.15587/1729-4061.2026.360606

IMPROVING THE ACCURACY OF PERFUSION MAP GENERATION USING DYNAMIC SUSCEPTIBILITY CONTRAST MAGNETIC RESONANCE IMAGING DATA BASED ON RECURRENT NEURAL NETWORKS

Oleksii Diumin*
Corresponding author
E-mail: o.dyumin-fbmi@lll.kpi.ua
ORCID: https://orcid.org/0000-0002-0196-005X
Svitlana Alkhimova
Candidate of Technical Sciences, Associate Professor
ORCID: https://orcid.org/0000-0002-9749-7388
*Department of Biomedical Cybernetics
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"
Beresteiskyi ave., 37, Kyiv, Ukraine, 03056

Received 19.02.2026
Received in revised form 28.04.2026
Accepted 08.05.2026
Published 30.06.2026

How to Cite: *Diumin, O., Alkhimova, S. (2026). Improving the accuracy of perfusion map generation using dynamic susceptibility contrast magnetic resonance imaging data based on recurrent neural networks. Eastern-European Journal of Enterprise Technologies, 3 (9 (141)), 38–48. https://doi.org/10.15587/1729-4061.2026.360606*

1. Introduction

Strokes and cancer are common causes of disability and mortality worldwide [1]. Timely and high-quality diagnostics are essential for improving patient outcomes, which requires improvements in existing diagnostic imaging analysis and visualization technologies. Diagnostics often includes assessment of regional blood flow by calculating and analyzing perfusion parameters. Perfusion tomography methods are aimed at analyzing the amount, time, and velocity of blood flow through the tissues under study, usually using computed tomography (CT) or magnetic resonance imaging (MRI). Dynamic susceptibility contrast magnetic resonance imaging (DSC-MRI) is mainly used to study brain tissue. It allows for assessment of cerebral perfusion without ionizing radiation, unlike CT [2]. In clinical practice, interpretation of DSC-MRI perfusion data includes two basic approaches:

1) quantitative assessment of perfusion parameters;

2) visual analysis of perfusion maps (perfusion parameters coded by color scales) [3].

Analysis of DSC-MRI data is based on obtaining time series of T2-/T2*-weighted MRI images, which are characterized by loss of signal intensity caused by the introduction of a contrast agent. According to the change in the pixel signal over time, time-concentration curves are formed, which are a source of information for obtaining perfusion parameters: cerebral blood volume (CBV), blood flow (CBF), time to maximum of the residual impulse function (Tmax), mean transit time (MTT), etc. CBV plays a particularly important role, as it reflects the density of micro vessels and angiogenic activity in brain tissues and is closely correlated with the degree of tumor malignancy, its aggressiveness and patient prognosis in gliomas, in particular glioblastoma. Relative CBV (rCBV) is used to characterize tumors, assess response to therapy, and differentiate relapse from treatment-related changes. Compared to flow-based parameters (CBF) and time-based metrics (MTT and Tmax), CBV estimation is generally less sensitive to the choice of arterial inflow function (AIF), bolus delivery delays, and noise.

Accurate assessment of cerebral perfusion is essential for the diagnosis of stroke and cancer. It is usually performed us-

ing DSC-MRI. Traditional analysis involves AIF estimation and deconvolution, which are sensitive to noise and artifacts in the data. Recent deep learning approaches [4] aim to generate maps of perfusion parameters directly from DSC-MRI data. However, most existing models are characterized by high computational complexity and focus mainly on spatial features and do not sufficiently take into account temporal dynamics.

Therefore, improving the accuracy of perfusion analysis without excessive computation remains a pressing scientific challenge.

2. Literature review and problem statement

A general scheme of the stages in operation of modern software for processing and analyzing perfusion data is given in [5]. It is shown that modern software uses complex and time-consuming algorithms and multi-step calculations to obtain perfusion parameters. The conventional approach to DSC-MRI analysis includes AIF deconvolution for quantitative assessment of blood flow (CBF) and the impulse function of the residual R. It is based on the convolution model $C(t) = \text{CBF} * R(t)$, where $*$ denotes deconvolution [6]. However, two significant problems of this method remain unresolved.

One of the problems is the definition of AIF, which reduces the accuracy of calculating perfusion parameters. The reason for this is the following problems that inevitably arise regardless of the chosen option for determining AIF. Usually, either a global AIF for all image pixels is defined, or local AIFs are defined that minimize the distance between the blood source and the studied tissues. Global AIF assumes no arterial-tissue delays. Local AIF is affected by partial volume effects due to the limited spatial resolution of MRI, which results in the averaging of several different signals in a single pixel. This introduces inaccuracies, making AIF determination a significant issue in perfusion studies.

A second problem is associated with the significant influence of noise and motion artifacts in the initial data on the accuracy of blood flow estimation (maximum CBF value $* R(t)$). The reason for this is that conventionally, singular value decomposition (SVD)-based deconvolution is used for calculations [6]. However, SVD strongly distorts the result even in the presence of the smallest noise and artifacts in the data. An option to overcome this problem may be modernized SVD methods that are less susceptible to noise and artifacts in the data. Such methods were used in work [7], in which modernized deconvolution methods, in particular cSVD and oSVD, were proposed to improve CBF estimation. However, these methods give significantly different CBF estimates, which often contain erroneous values. This is confirmed in work [8], in which the results of various modernized meconvolutions on synthetic data are compared.

However, synthetic data may not contain all possible variants of noise and artifacts that occur in real data. An option to overcome these difficulties may be to increase the spatiotemporal coherence of MRI slices obtained during the studies. For this purpose, various algorithms or their combinations are used. These include, in particular, motion artifact correction [9] to reduce the influence of motion during scanning. However, motion artifact correction algorithms can distort the signal and negatively affect the determination of AIF. Data optimization is also used, in particular, manual or automated selection of the region of interest on MRI slices [10]. However, manual selection of the region of interest is time-consuming, requires an experienced radiologist and is not protected from the human factor. Therefore, preference is given to automated selection of the region of interest, taking into account its inaccuracy and susceptibility to noise and motion artifacts in MRI slices.

In [11], a method of automated selection of the region of interest by removing pixels of low and high intensity is considered. However, the method was found to be ineffective for determining the area of brain perfusion and may cause deterioration of the quality of perfusion maps. The reason for this may be the incorrect way of determining the filtering threshold of MRI image pixels. A more reasonable automatic approach for determining the filtering threshold may be an option to overcome this problem.

Such approaches are considered in [12], in particular, global search using Otsu's method, local search using Niblack's method, search in the anatomical location of the brain. However, such approaches are also susceptible to noise and artifacts in the data, as well as to anomalous data, which can lead to inaccuracies, especially at the brain boundary. An option to overcome this problem may be to use a special filter that is resistant to data anomalies, which will more accurately determine the brain boundaries. Such an approach is considered in [13], which involves using the CUSUM filter, including for anomalous data [14]. However, the CUSUM filter remains susceptible to noise and artifacts in the data, although it is resistant to data anomalies and determines the boundaries of the region of interest more accurately than threshold filters. An option to overcome the problem of noise and artifacts in the data may be to use deep learning methods to determine the region of interest.

That is the approach used in [15], in which a convolutional neural network of the UNet type is used to define the region of interest. However, the disadvantage of this neural network model is its homogeneity and insufficient amount of data for training. Other methods, such as noise reduction, automated AIF estimation, and correction of partial volume effects [16], significantly increase the data processing time and do not provide complete elimination of errors in the perfusion parameter maps.

An option to overcome the above problems can be machine learning methods, which demonstrate high efficiency in solving a wide range of scientific problems in the field of informatics and automation using medical and biomedical data of various modalities. In particular, work [17] investigates the diagnosis of COVID-19-associated cardiopulmonary pathology using CT data using machine learning methods. Among the disadvantages of using the methods in the study are the limited size of the data set, computational complexity, and lack of explanation of the solution. In addition, the authors of [18] devised a deep learning method for non-invasive diagnosis of Bankart lesions. The disadvantages of the study are the limited size of the data set, class imbalance, and also the lack of explanation of the solution.

Work [19] investigates the use of a diffusion model for noise reduction of 3D PET images. The limitations of the study are the insufficient sample size for testing and the lack of a gold standard. In [20], the use of machine learning methods for blood group recognition and classification is investigated. Among the shortcomings, the dependence on image quality and the insufficient sample size for testing can be noted.

The above gives grounds to argue that the application of machine learning methods for medical data processing is a

promising area of research. Studies of calculating perfusion parameters from DSC-MRI data are increasingly using machine learning methods [4]. Their potential advantages include independence from AIF and relative resistance to noise and motion artifacts. Convolutional neural networks (CNNs) are widely used to generate maps of perfusion parameters from time series of DSC-MRI data due to their efficiency in image processing.

A number of papers have proposed CNN architectures specifically designed for this task. In [21], a two-input convolutional neural network with special convolutional filters for noise and motion artifact suppression was proposed. However, the time required to apply the model significantly exceeds the time required to apply the conventional deconvolution method. This may be due to the complexity of the model, which is associated with the presence of two separate convolutional chains, as well as additional convolutional filters for noise reduction. This can be overcome by optimizing the model, in particular, independence from the definition of AIF. This is the approach used in [22, 23], in which maps of perfusion parameters are generated without explicit definition of AIF.

Independence from AIF reduces the amount of preprocessing and model complexity. However, an unresolved issue in those models is the insufficient amount of data for training and testing. This may be due to the high cost of MRI and the reluctance of patients to provide their data. The solution to this may be to attract funding for data collection and anonymize patient data. Also insufficient is the consideration of the temporal component in the data, which can be solved by using convolutional or recurrent layers for the temporal dimension.

Other studies are aimed at explicitly using the temporal dynamics of DSC-MRI sequences during the generation of perfusion maps. For example, in [24], an architecture based on a generative adversarial neural network (GAN) is used, in which temporal information is processed either by combining time points in the channel dimension or using specialized 3D temporal convolutions before extracting spatial features. An unresolved issue is the weak validity of the method, due to the insufficient amount of data, as well as the lack of explanation of how the result was obtained, which is inherent in most deep learning methods. The solution to the latter problem can be the use of explanatory artificial intelligence, or visualization and analysis of convolutional filters.

In [25], a 1D convolutional encoder (part of the encoding) is used to analyze signal changes over time, followed by the use of a 2D U-Net architecture to generate maps of perfusion parameters. The issue of homogeneity of the data set remains unresolved. A larger and more diverse data set should be the solution. Similarly, in [26], a two-level architecture combining a 3D CNN for spatial feature extraction with a 1D CNN-based temporal network to account for both global and local temporal dependencies is proposed. Another problem with the study is the insufficient amount of data. In addition to CNN and GAN-based approaches, conventional feedforward neural networks (FNNs) were also considered as a simpler alternative [27]. The problem with the use of FNNs is their low accuracy and efficiency. This is likely due to the insufficient capabilities for analyzing spatial and temporal features. Overcoming this limitation is the use of convolutional neural networks, or others that are capable of qualitatively detecting spatial and temporal features.

Studies have mainly focused on spatial features, sometimes taking into account temporal features only through convolutional layers. Papers [28, 29] confirm the increase in accuracy when using recurrent neural networks (RNNs), including GRU and LSTM, applied to temporal data. However, the issue of using recurrent layers for analyzing the temporal component specifically in DSC-MRI data and for the task of generating perfusion maps is not sufficiently studied. Thus, the use of recurrent neural networks for processing time series of MRI brain images is a promising direction, which potentially makes it possible to increase the accuracy of perfusion parameter maps.

Despite promising results, the approaches proposed in the above reviewed studies have a number of limitations. The effectiveness of deep learning models depends largely on the availability of large training samples, and limited data can reduce the generalization ability and accuracy of the models. Some approaches are based on simplified assumptions about voxel independence, which may not be correct in the presence of patient motion or spatial correlations. Some methods still require manual pre-processing steps, such as AIF determination, which reduces the level of automation of the analysis.

In addition, complex models often require significant computational resources, which makes their implementation in clinical practice difficult. Also, insufficient analysis of temporal information can reduce the accuracy of the obtained perfusion maps since the signal dynamics (time-concentration curves) play a key role in the assessment of perfusion. Therefore, the main unsolved issue is the inaccuracy and inefficiency of perfusion map generation, which is caused by insufficient analysis of the temporal component in the data, dependence on AIF, and high computational complexity of neural network models.

The above allows us to argue that it is advisable to conduct a study aimed at improving the accuracy of perfusion map generation using a convolutional neural network, which is independent of AIF, has moderate computational complexity, and effectively processes temporal features through the use of recurrent layers.

3. The aim and objectives of the study

The aim of our study is to improve the accuracy of generating maps of perfusion parameters from time series of DSC-MRI images of the brain by using neural networks with recurrent layers for effective processing of temporal features. The proposed approach is aimed at increasing the accuracy, reliability, efficiency, and level of automation of perfusion analysis in the diagnosis of stroke and brain tumors. At the same time, excessive complexity of the model is avoided to ensure practical applicability and faster application of the model in clinical conditions.

To achieve this aim, the following objectives were accomplished:

– to design convolutional neural networks for generating maps of perfusion parameters from time series of DSC-MRI images of the brain both with and without the use of recurrent layers;

– to investigate the effect of recurrent layers on the accuracy of generating perfusion maps by comparing the proposed convolutional neural networks with and without recurrent components;

– to evaluate the designed neural network by comparing it with existing similar neural networks.

4. The study materials and methods

4. 1. The object and hypothesis of the study

The object of our study is the process of generating perfusion parameter maps from time series of DSC-MRI images of the brain.

The basic hypothesis of the study assumes increasing the accuracy of perfusion map generation by using recurrent layers to analyze the temporal component in the data. The study assumes that recurrent layers are better than convolutional layers at analyzing the temporal components in DSC-MRI data. The study adopts simplifications, which consist in using only one CBV map for training and testing the models, as well as using only three metrics to assess the accuracy of the generated perfusion maps.

Our study uses data from the open UPENN-GBM dataset [30]. The dataset was used to train and evaluate the proposed deep learning model for generating perfusion parameter maps from DSC-MRI time series. This dataset contains dynamic susceptibility contrast MRI images together with the corresponding perfusion maps. A detailed description of the dataset is given below.

The UPENN-GBM dataset contains multiparametric brain MRI images from 630 patients with newly diagnosed glioblastoma (GBM) acquired at the University of Pennsylvania Health System. The collection includes structural MRI sequences (T1, T1-Gd, T2, T2-FLAIR), diffusion-weighted imaging, and DSC-MRI perfusion imaging, all of which are co-registered, cleared of non-brain tissue, and resampled to an isotropic resolution of 1 mm^3.

Although DSC-MRI allows for the assessment of various perfusion parameters, only CBV maps were used in our study. This choice was driven by both clinical relevance and data reliability. The UPENN-GBM dataset contains standardized relative CBV (rCBV) maps obtained using validated preprocessing pipelines that account for contrast leakage and bolus quality [30]. These maps are a consistent and widely accepted benchmark for perfusion analysis in glioblastoma. Focusing on a single perfusion parameter has reduced methodological variability and ensured that differences in model performance are driven primarily by architectural features, particularly the efficiency of temporal feature processing. Extending the analysis to include parameters such as CBF or MTT would introduce additional factors related to AIF definition and deconvolution methods. This is beyond the scope of this study and could be addressed in future work.

The following data were used in our study: time series of DSC-MRI perfusion images (T2-weighted) and rCBV maps obtained using standardized preprocessing procedures.

The rCBV maps were calculated from 4D DSC-MRI data using validated postprocessing methods. These include contrast pre-injection and bolus quality control to reduce leakage effects. These maps are used as voxel-based ground truth labels to train deep learning models aimed at generating perfusion maps from raw DSC-MRI data. The full UPENN-GBM dataset also contains tumor masks, diffusion indices, radiomic features, and extended clinical and molecular metadata. However, in this work, only the DSC-MRI time series and rCBV maps were used.

4. 2. Data preprocessing

To ensure consistency and robustness of model training, a single preprocessing pipeline was applied to all DSC-MRI volumes and the corresponding CBV parametric maps. It standardized the data in spatial and temporal dimensions and normalized the intensity values. It also ensured strict separation of data at the patient level between training, validation, and test sets to prevent data leakage and ensure correct evaluation.

Each 4D DSC-MRI scan was first downloaded in its original form, after which axial slices were selected for each patient. Specifically, every 15th slice in the central anatomical region was used – from the 25th to the 130th slice. This allowed us to exclude the first and last 25 slices, which usually contain low-informative areas or partial brain coverage and may be prone to artifacts. A slice interval of 15 was chosen to reduce redundancy between similar neighboring slices while preserving anatomical diversity and representation of lesions within the brain volume. Each selected 2D slice was scaled to 128×128 pixels using nearest neighbor interpolation, which ensured preservation of anatomical boundaries and compatibility with both intensity maps and discrete labels. The full temporal sequence of each slice was preserved without interpolation, and the corresponding CBV map was co-registered and resampled at the same spatial resolution. To ensure consistency of intensity values across patients and scanners, each DSC volume was normalized to the maximum intensity value of the voxels in the training set.

The dataset was divided at the patient level into three subsets: training (68%), validation (12%), and test (20%). This ratio corresponds to the common practice, in which most of the data is used for training, and a sufficient amount is reserved for unbiased evaluation. Importantly, all data from one patient belonged to only one subset, which excluded the intersection of information between samples. As a result, the training sample contained 2,471 examples, the validation sample contained 441 examples, and the test sample contained 721 examples. Each example was a spatiotemporal input volume combined with a corresponding perfusion map, prepared for training with a teacher. Such preprocessing ensured anatomical diversity, temporal consistency, and methodological correctness. That allowed the development and evaluation of deep learning models under conditions close to real clinical variability.

4. 3. Performance evaluation

Three traditional metrics were used to evaluate the generated perfusion maps: normalized root mean square error (NRMSE), structural similarity index (SSIM), and peak signal-to-noise ratio (PSNR).

NRMSE measures the normalized deviation between the generated image y^* and the reference (true) image y, and is calculated from the following formula

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^*\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^{\,2}}}, \quad (1)$$

where n is the total number of voxels in the image; y_i and y_i^* are the true and generated values in voxel i.

The SSIM metric estimates the similarity between two images based on local brightness, contrast, and structure characteristics. It is defined as

$$SSIM\left(y, y^*\right) = \frac{\left(2m_y m_{y^*} + c_1\right)\left(2s_{yy^*} + c_2\right)}{\left(m_y^2 + m_{y^*}^2 + c_1\right)\left(s_y^2 + s_{y^*}^2 + c_2\right)}, \quad (2)$$

where m_y, m_{y^*} – mean values of intensity of reference and generated images; s_y^2, $s_{y^*}^2$ – variances; s_{yy^*} – covariance; $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ – constants to ensure numerical stability, L – dynamic range of image intensity (e.g., 1 for normalized data), typical values: $k_1 = 0.01$, $k_2 = 0.03$. The SSIM value lies in the range from 0 to 1, where 1 corresponds to perfect structural similarity.

The PSNR metric evaluates the reconstruction quality as the ratio of the maximum possible signal to the reconstruction error

$$PSNR = 10 \cdot \left(\frac{MAX_I^2}{MSE} \right), \quad (3)$$

where MAX_I is the maximum possible value of image intensity (usually 1 for normalized data); MSE is the mean square error between the generated and reference images

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - y_i^* \right)^2. \quad (4)$$

Higher PSNR values correspond to better image quality and lower levels of distortion.

The study used the same dataset, preprocessing pipeline, and evaluation protocol to ensure methodological consistency.

5. Results of investigating perfusion map generation by the proposed neural network models

5. 1. The designed convolutional neural networks with and without recurrent layers for perfusion map generation

To evaluate CBV maps from DSC-MRI time series, three deep learning architectures were designed and compared, each of which implements different approaches to spatiotemporal feature learning. These architectures include three options: the first is a fully convolutional 3D model based on spatiotemporal convolutions, the second is a model using ConvLSTM2D blocks, and the third is a U-Net-type architecture with a bottleneck based on GRU. The purpose of this comparison was to evaluate the contribution of explicit estimation of temporal dynamics using recurrent layers compared to an approach based solely on convolutional compression of temporal information. The purely convolutional architecture was used as a baseline model to assess the feasibility of using recurrent layers in this task.

The Conv3D architecture is a fully convolutional model that uses 3D convolutions to simultaneously learn spatial and temporal features from input DSC-MRI time series. The model takes a four-dimensional tensor representing a single axial slice in time, of size (128, 128, 45, 1), and generates the corresponding two-dimensional CBV map.

The encoder consists of four dimensionality reduction blocks. Each block starts with a 3D convolution that spans the entire time axis and reduces the time dimension to unity. The resulting output is converted into a 2D spatial feature map that is used for bandwidth connections. A second 3D convolution is then applied with a smaller kernel size and steps that reduce the spatial resolution, providing more abstract spatiotemporal features. Hierarchical coding reduces the dimensionality from 128 × 128 to 8 × 8 pixels along the encoder. The number of filters is reduced from 16 to 1 to ensure the compactness of the model.

At the bottleneck, a final 3D convolution with a temporal kernel is applied, which completely eliminates the temporal dimension and forms a compact spatial representation of size 8 × 8. This architecture deliberately does not use recurrent blocks – instead, large temporal kernels are used to take into account dynamic information. Thus, this model serves as a baseline for assessing the need and effectiveness of explicit modeling of temporal dynamics using recurrent layers.

The decoder (decoding part) restores spatial resolution using a sequence of 2D transposed convolutions. At each stage, the feature maps are combined with the encoder bandwidth connections, which allows the details lost during dimensionality reduction to be restored. The final output layer applies a 1 × 1 convolution with a ReLU activation function, forming a 128 × 128 CBV map with non-negative values at the voxel level.

The model is optimized using a mean squared error (MSE) loss function, which estimates the difference between the generated and reference perfusion maps at the voxel level. The mean absolute error (MAE) is used as an additional evaluation metric for a more interpretable analysis of the generation accuracy. The general structure of the Conv3D model is in in Fig. 1, which shows the components of spatiotemporal encoding, temporal information convolution, and multi-level 2D decoding used to generate CBV maps.

Our architecture illustrates a basic spatiotemporal convolutional strategy used to generate CBV maps without focusing on temporal dynamics.

The ConvLSTM2D architecture introduces temporal dynamics processing using recurrent layers to better account for the dynamics of the passage of contrast agent over time. Unlike the Conv3D model, which takes into account temporal dynamics implicitly through the use of large 3D kernels, this architecture uses ConvLSTM2D blocks to learn the temporal dependencies at each spatial point, taking into account the sequential data structure. The model accepts an input tensor of size (45, 1, 128, 128), where the value 45 corresponds to the number of time points, and the spatial dimensions correspond to one axial slice. This data representation format is convenient for sequential processing using ConvLSTM2D layers.

The encoder consists of four dimensionality reduction blocks, each containing two consecutive ConvLSTM2D layers. The first layer in each block reduces the spatial resolution using stepwise convolutions and returns only the final hidden state, which is stored for use in forward connections. The second ConvLSTM2D layer in each block preserves the temporal dimension and further refines the formed representation. After each recurrent layer, batch normalization is applied to stabilize training and improve convergence. After encoding, the tensor has a spatial dimension of 8 × 8 and a minimum feature depth. At the bottleneck, the final ConvLSTM2D layer processes the full time sequence and forms a time-aggregated representation of size 8 × 8 × 1. This recurrent bottleneck layer allows for the training of long-term temporal dependencies critical for accurate perfusion assessment, such as contrast arrival time (TTA), curve shape, and contrast washout process. The temporal dimension is treated as an ordered sequence and is not compressed prematurely.

The decoder replicates the encoder structure and gradually recovers the spatial resolution using 2D transposed convolutions. At each stage, the decoder merges the recovered feature map with the corresponding bandwidth connection from the encoder. This multi-level merging allows for effi-

cient recovery of fine spatial details and boundaries, which are critical for accurate estimation of CBV maps.

As in the case of the Conv3D model, the result is a 128 × 128 CBV map obtained using a final tensor transformation and a ReLU activation function that ensures non-negativity at the voxel level. The model is trained using an MSE loss function, while MAE is used for estimation. The overall architecture is shown in Fig. 2, which demonstrates the integration of recurrent layers for temporal dynamics processing and the use of bandwidth connections in a multi-level decoder.

This architecture demonstrates the integration of recurrent layers to explicitly process temporal dependences in DSC-MRI time series during CBV map generation.

The UNet architecture with GRU is based on the widely used UNet paradigm, extended by a recurrent bottleneck to explicitly process global temporal dynamics in DSC-MRI time series. This approach combines the advantages of spatial feature-rich convolutional encoding and decoding with the sensitivity of recurrent blocks to temporal dependencies, providing a balanced combination of anatomical feature extraction and temporal information processing. The model accepts a five-dimensional input tensor of size (128, 128, 45, 1), which is transformed into a two-dimensional spatial grid with 45 channels corresponding to time points. The initial convolutional layer reduces the spatial resolution and increases the feature depth, forming the input to the encoder.

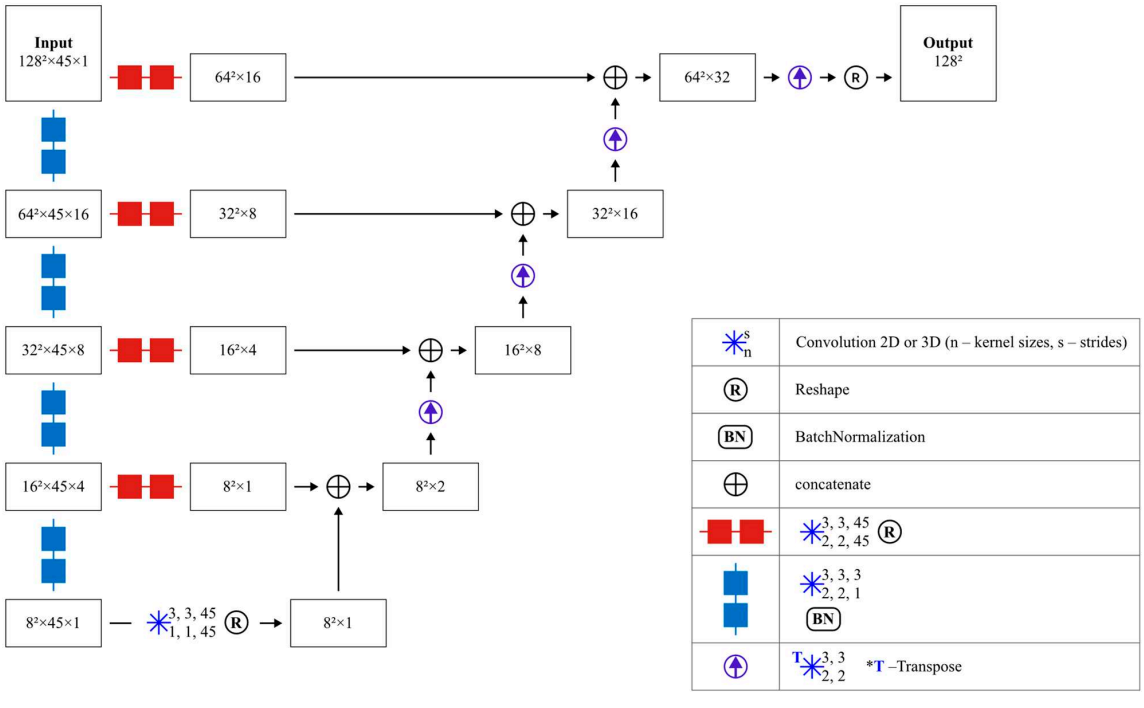

Fig. 1. Proposed architecture of the Conv3D model for generating perfusion maps from dynamic susceptibility contrast magnetic resonance imaging data

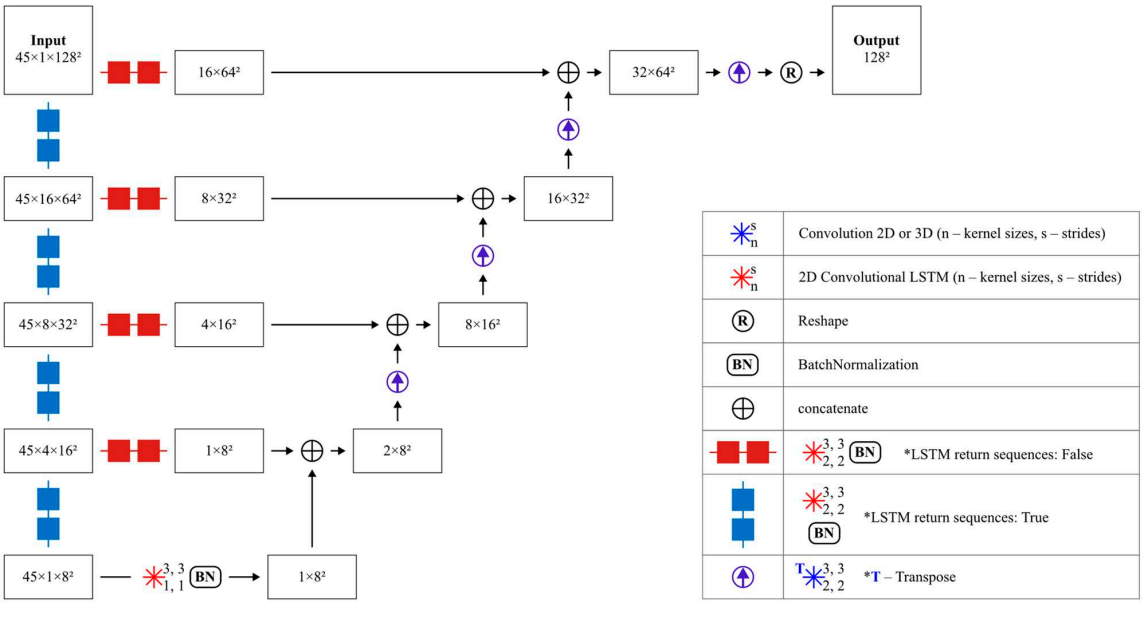

Fig. 2. Proposed architecture of the ConvLSTM2D model for generating perfusion maps

The encoder has a classic UNet structure and consists of four blocks with dimensionality reduction. Each block includes a sequence of convolutional layers with ReLU activation functions and residual connections, which allows for the formation of increasingly abstract spatial features while simultaneously reducing the spatial resolution by half at each level. The architecture supports compression and excitation modules by channels and spatial dimensions (scSE). However, in the studied variant, attention blocks are not used, since the main emphasis is on processing temporal dynamics using recurrent layers.

At the bottleneck, the encoder output is transformed into a sequence of 16 time steps, each with a 512-dimensional representation. A GRU is applied to this sequence along the time axis, which allows long-term dependencies to be captured. These features are critical for accurate estimation of CBV maps. Placing a recurrent block at the bottleneck ensures that temporal aggregation is performed after spatial generalization, preserving anatomical structure and maximizing temporal context.

The decoder is symmetric to the encoder and gradually restores spatial resolution using transposed convolutions. Each level uses feedforward connections from the encoder, allowing for the recovery of details lost during dimensionality reduction. The final convolutional layers restore the full 128×128 spatial resolution and form a CBV map using a 1×1 convolution with a ReLU activation function to ensure non-negative values.

The model is trained using the mean square error (MSE) loss function, while the mean absolute error (MAE) is used for evaluation as an interpreted voxel-level error metric. The complete architecture is shown in Fig. 3, which depicts the UNet structure with integrated temporal dynamics processing using GRUs at the bottleneck.

The proposed UNet architecture with GRU emphasizes the role of recurrent temporal aggregation at the bottleneck stage to improve the accuracy of CBV map generation based on DSC-MRI time series.

5. 2. Comparing the designed convolutional neural networks with and without recurrent layers

Our comparison included the convolutional model Conv3D without recurrent components and two recurrent models – ConvLSTM2D and UNet with GRU. The quantitative results are in Table 1, which gives values for the NRMSE, PSNR, and SSIM metrics for the models built.

As shown in Table 1, the Conv3D model achieved NRMSE of 0.027, PSNR of 31.493, and SSIM of 0.826. The ConvLSTM2D model reduced NRMSE to 0.016, increased PSNR to 36.095, and improved SSIM to 0.957. The UNet architecture with GRU showed the best results: NRMSE of 0.015, PSNR of 36.412, and SSIM of 0.973. Compared with Conv3D, the recurrent architectures reduced NRMSE by about 40–44% and increased SSIM by 0.131–0.147 in absolute terms. The PSNR value increased by about 4.6–4.9 dB after introducing recurrent layers. The difference between the two recurrent models is insignificant: NRMSE differs by 0.001, PSNR by 0.317, and SSIM by 0.016.

Table 1

Comparative efficiency of proposed architectures for generating CBV maps

Model	NRMSE↓	PSNR↑	SSIM↑
Conv3D (proposed)	0.027	31.493	0.826
ConvLSTM2D (proposed)	0.016	36.095	0.957
UNet with GRU (proposed)	**0.015**	**36.412**	**0.973**

Note: Arrows indicate the desired direction of metric optimization, and the best values are highlighted in bold.

5. 3. Evaluating the designed neural network with recurrent layers by comparison with analogs

To solve the third research task, the best proposed architecture – UNet with GRU – was compared with representative approaches described in previous studies [23, 24, 26]. The comparative results are given in Table 2, which contains the quantitative values of the metrics reported in the relevant publications, together with the results obtained by the proposed architecture.

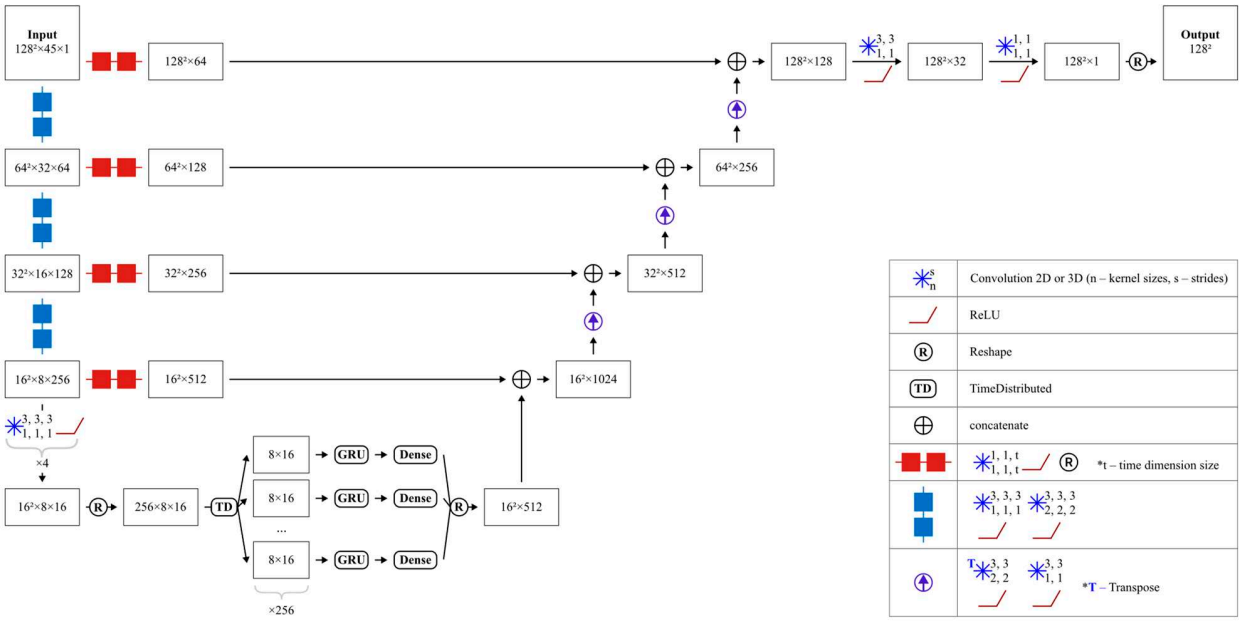

Fig. 3. Proposed UNet model architecture with recurrent layers in the bottleneck for generating perfusion maps

Table 2

Comparative evaluation of the proposed UNet model with GRU relative to existing approaches

Model	NRMSE↓	PSNR↑	SSIM↑
L UNet [23]	0.067	24.732	0.886
GAN [24]	0.025	**42.5**	0.955
ST-Net [26]	0.252	31.644	0.952
UNet with GRU (proposed)	**0.015**	36.412	**0.973**

Note: Arrows indicate the desired direction of metric optimization, and the best values are highlighted in bold.

As shown in Table 2, the proposed UNet approach with GRU achieved the lowest NRMSE value (0.015) among all the compared methods, outperforming the GAN-based approach (0.025), the L UNet variant (0.067) and ST-Net (0.252). Compared to the GAN model, this corresponds to a reduction in the normalized error of approximately 40%, while the improvement is even more pronounced for L UNet and ST-Net. The proposed approach also achieved the highest SSIM value (0.973), outperforming the corresponding values for GAN (0.955), ST-Net (0.952) and L UNet (0.886), indicating the best structural agreement with the CBV reference maps. The obtained PSNR value (36.412) further confirms the high quality of the reconstruction of the generated maps. At the same time, the GAN-based approach demonstrated the highest PSNR value (42.5). To further refine the comparison with the GAN approach, for which a PSNR of 42.5 was obtained, the mean absolute error (MAE) was additionally estimated for the same normalization range [−1, +1] [24]. With such consistent intensity scaling, the proposed UNet approach with GRU achieved an MAE of 0.0078, while the GAN approach achieved an MAE of 0.0091. Since MAE characterizes the absolute deviation of intensity, lower values correspond to higher generation accuracy. Thus, the proposed approach provided a reduction in the mean absolute error by approximately 14% compared to the GAN-based approach.

6. Discussion of results related to using recurrent neural networks for generating perfusion maps

The neural network architectures shown in Fig. 1–3 have been designed. All of them have convolutional layers and a UNet-type structure, which is explained by the need to take into account spatial features, as well as by optimizing the computational complexity by repeatedly reducing the dimensionality of the convolutional layers. The last two models shown in Fig. 2, 3 also have recurrent layers, which is explained by the need for high-quality processing of the temporal component in the data. The designed convolutional neural networks with and without recurrent layers were compared by three metrics: NRMSE, PSNR, and SSIM. As can be seen from the results given in Table 1, both architectures with recurrent layers outperform the basic convolutional model without recurrent layers by all evaluated metrics. Our results show that the use of recurrent layers consistently increases the accuracy and structural consistency of the generated CBV maps from the DSC-MRI time series. Such results are explained by a more efficient analysis of the temporal information in the DSC-MRI data due to the use of recurrent layers and confirm the feasibility of their application. The best developed model in all metrics is compared with existing approaches from the considered studies, which had the least shortcomings and were evaluated on the same metrics. As can be seen from the results given in Table 2, the proposed UNet approach with GRU outperformed existing approaches [23, 24, 26] in two out of three metrics, which is explained by the use of recurrent layers in the proposed model in contrast to existing methods. The value of the PSNR metric of the developed model was competitive, as it exceeded the values of the other two existing approaches [23, 26], but turned out to be worse than the GAN-based approach [24].

That can be explained by the fact that the PSNR metric mainly reflects the global accuracy of intensity reproduction and is sensitive to peak signal values. It does not provide a direct assessment of structural similarity or the distribution of normalized error across the image. Therefore, the proposed model was additionally compared with the GAN-based approach by an additional metric, the value of which turned out to be better for the proposed developed model. The better values of the vast majority of metrics of the proposed UNet-type model with GRU are explained by a better analysis of the temporal component in the DSC-MRI data due to the optimized use of recurrent layers in combination with convolutional layers.

Unlike [22, 24–27], in which models for generating perfusion maps were built exclusively on the basis of CNN, GAN, and FNN architectures, our study also used RNN architectures, namely recurrent LSTM and GRU layers. The proposed architectures with and without recurrent layers are compared both with each other and with existing approaches, which allows a broader assessment of the impact of using recurrent layers for the analysis of temporal information in DSC-MRI. Unlike [21], in which the definition of AIF is required, the proposed approach is independent of AIF, due to the fact that it takes as input and analyzes the entire DSC-MRI time series. Unlike [6], in which the conventional deconvolution method can strongly distort the result even with the smallest errors in the input data, the proposed approach is less susceptible to noise and artifacts in the data. This becomes possible due to the use of kernels in the convolutional layers, which analyze not only each pixel, but also its neighborhood. Unlike [21, 23], in which complex time and computational complexity models were developed, the proposed model optimizes the use of convolutional layers by multiple dimensionality reduction and optimizes the use of recurrent layers by using them in the bottleneck with the smallest dimensions. This allows us to minimize the time and computational complexity, as well as achieve a result acceptable for clinical use.

The issue of insufficient data volume was partially solved by using a dataset with the largest number of patients compared to the datasets used in other reviewed studies on the application of machine learning methods to generate perfusion maps. The focus was also focused on CBV perfusion maps to reduce methodological variability. The task of the simplified assumption of voxel independence was solved by using convolutional layers to analyze spatial dependencies. The kernels of convolutional layers take into account the voxel together with the surrounding area, rather than separately, which allows for qualitative detection of spatial features. The task of determining AIF is solved by abandoning the use of AIF for generating perfusion maps. The proposed approach is independent of AIF, instead accepting the full DSC-MRI time series as input. The issue of high time and computational complexity is overcome in the proposed model due to the optimal use of convolutional and recurrent layers. Convolutional layers reduce the dimensionality many times, and recurrent layers are used in a bottleneck, which makes it possi-

ble to minimize the time and computational complexity of the model. The issue of insufficient consideration of the time component was solved by using recurrent layers, which confirmed their effectiveness according to the results given in Tables 1, 2.

The obtained better metric values in the models with recurrent layers, compared to the models without them, demonstrate that the use of recurrent layers for processing temporal dynamics significantly improves the accuracy of CBV map generation from DSC-MRI time series. The UNet architecture with GRU demonstrated the best overall efficiency among the studied models. This creates a practical basis for using the proposed approach as an automated tool for CBV map generation. Such an approach could reduce the dependence on sensitive deconvolution steps and provide more stable, operator-independent processing of perfusion data. In clinical workflows, such models can act as a fast additional tool for DSC-MRI-based perfusion analysis. They can be useful for tumor assessment and therapy monitoring, as well as for generating standardized CBV maps for further tasks such as segmentation or radiomic analysis.

Our results indicate that the use of recurrent layers increases quantitative accuracy while maintaining the compactness of the architecture, which confirms the practical feasibility of applying the proposed approach in automated perfusion analysis systems.

The dataset, although it includes the largest number of patients compared to other studies reviewed, is homogeneous. Direct comparisons with previously reported approaches [23, 24, 26] should be interpreted with caution, as these models were trained and evaluated on different datasets. Furthermore, only the generation of CBV maps is evaluated in our work; the generalization of the results to other perfusion parameters (e.g., CBF, MTT, Tmax) was not tested in this study. Although the use of recurrent layers improves accuracy, it involves sequential processing of time points, which may lead to increased model application time and memory requirements compared to convolutional models. This should be taken into account in time-sensitive clinical scenarios.

Future studies should aim to extend the approach to other perfusion parameters, such as CBF and MTT, and to test the generalization ability of the model on different datasets and across different medical institutions. In addition, it is advisable to optimize the architecture to reduce computational costs when applying the model without compromising accuracy.

7. Conclusions

1. Three convolutional neural networks have been designed and experimentally verified for quantitative CBV map generation from DSC-MRI time series. The neural networks included a basic convolutional model without recurrent components and two models using recurrent layers. The proposed models provide direct generation of perfusion maps from dynamic data using a standardized preprocessing pipeline and data partitioning at the patient level.

2. The impact of recurrent layers on the accuracy of map generation was systematically investigated by a controlled comparison of architectures with and without recurrent components. The recurrent approaches demonstrated improvements in all evaluation metrics. The value of the normalized root mean square error (NRMSE) decreased from 0.027 to 0.016 and 0.015. The structural similarity index (SSIM) increased from 0.826 to 0.957 and 0.973. The peak signal-to-noise ratio (PSNR) increased from 31.493 to 36.095 and 36.412. Our quantitative results confirm the positive effect of recurrent layers on increasing the accuracy of perfusion map generation from time series of magnetic resonance images.

3. Comparative evaluation of the designed neural network with existing similar neural networks showed that the best proposed recurrent architecture provides competitive or better results on key quantitative metrics. Under the conditions of consistent intensity normalization, which corresponds to GAN-based approaches, the proposed model achieved a lower MAE value of 0.0078 compared to 0.0091. This decrease corresponds to the improved consistency between the generated and reference CBV maps.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study and the results reported in this paper.

Funding

The study was conducted without financial support.

Data availability

The data will be provided upon reasonable request.

Use of artificial intelligence

ChatGPT 5.2 (OpenAI) was used to check grammar in this work. Each suggested correction was checked for appropriateness and, where necessary, for correctness of content. The results do not affect the conclusions of the study.

Authors' contributions

Oleksii Diumin: Methodology, Software, Validation, Formal analysis, writing – original draft, Writing – review & editing, Visualization; **Svitlana Alkhimova**: Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration.

References

1. Feigin, V. L., Stark, B. A., Johnson, C. O., Roth, G. A., Bisignano, C., Abady, G. G. et al. (2021). Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet Neurology, 20 (10), 795–820. https://doi.org/10.1016/s1474-4422(21)00252-0
2. O'Connor, J. P. B., Tofts, P. S., Miles, K. A., Parkes, L. M., Thompson, G., Jackson, A. (2011). Dynamic contrast-enhanced imaging techniques: CT and MRI. The British Journal of Radiology, 84, S112–S120. https://doi.org/10.1259/bjr/55166688

3. Alkhimova, S., Sorokina, V., Kabala, I. (2026). Comparative assessment of commonly used color lookup tables to determine key performance indicators for perfusion map data visualization. Technology Audit and Production Reserves, 1 (2 (87)), 85–92. https://doi.org/10.15587/2706-5448.2026.352787
4. Diumin, O. D., Alkhimova, S. M. (2024). Generation of perfusion map images using deep learning methods. Nauka i tekhnika sohodni, 11 (39), 849–866. https://doi.org/10.52058/2786-6025-2024-10(38)-849-866
5. Alkhimova, S. M., Diumin, O. D. (2024). Issues of automated perfusion analysis based on dynamic susceptibility contrast MRI. Nauka i tekhnika sohodni, 8 (36), 859–875. https://doi.org/10.52058/2786-6025-2024-8(36)-859-875
6. Østergaard, L., Weisskoff, R. M., Chesler, D. A., Gyldensted, C., Rosen, B. R. (1996). High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. Magnetic Resonance in Medicine, 36 (5), 715–725. https://doi.org/10.1002/mrm.1910360510
7. Wu, O., Østergaard, L., Weisskoff, R. M., Benner, T., Rosen, B. R., Sorensen, A. G. (2003). Tracer arrival timing-insensitive technique for estimating flow in MR perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix. Magnetic Resonance in Medicine, 50 (1), 164–174. https://doi.org/10.1002/mrm.10522
8. Kudo, K., Christensen, S., Sasaki, M., Straka, M., Fujiwara, S., Ishizaka, K. et al. (2010). Acuracy and reliability of post-processing software for DSC MR perfusion: Quantitative analysis by digital phantom data. In Proceedings of the 18th Annual Meetings of the ISMRM. Available at: https://archive.ismrm.org/2010/1791.html
9. Gautam, R., Sivaswamy, J., Varma, R. (2012). An efficient, bolus-stage based method for motion correction in perfusion weighted MRI. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), 145–148. IEEE. Available at: https://ieeexplore.ieee.org/abstract/document/6460093
10. Alkhimova, S. (2018). Detection of perfusion ROI as a quality control in perfusion analysis. In Science, research, development. arXiv. https://doi.org/10.48550/arXiv.1902.01855
11. Alkhimova, S. (2019). Impact of perfusion ROI detection to the quality of CBV perfusion map. Technology Audit and Production Reserves, 5 (2 (49)), 27–30. https://doi.org/10.15587/2312-8372.2019.182789
12. Alkhimova, S. M., Sliusar, S. V. (2019). Analysis of effectiveness of thresholding in perfusion roi detection on T2-weighted MR images with abnormal brain anatomy. KPI Science News, 4, 35–43. https://doi.org/10.20535/kpi-sn.2019.4.180237
13. Alkhimova, S., Krenevych, A. (2019). Brain tissues segmentation on mr perfusion images using CUSUM filter for boundary pixels. International Journal of Computing, 18 (2), 127–134. https://doi.org/10.47839/ijc.18.2.1411
14. Alkhimova, S. (2019). CUSUM Filter for Brain Segmentation on DSC Perfusion MR Head Scans with Abnormal Brain Anatomy. Proceedings of the 2019 International Conference on Intelligent Medicine and Image Processing, 43–47. https://doi.org/10.1145/3332340.3332357
15. Alkhimova, S., Diumin, O. (2023). Deep learning-based neural network for regions of interest retrieval in T2*-weighted brain perfusion MRI. Nauka I Tekhnika Sohodni, 14 (14). https://doi.org/10.52058/2786-6025-2022-14(14)-301-314
16. Fieselmann, A., Kowarschik, M., Ganguly, A., Hornegger, J., Fahrig, R. (2011). Deconvolution-Based CT and MR Brain Perfusion Measurement: Theoretical Model Revisited and Practical Implementation Details. International Journal of Biomedical Imaging, 2011, 467563. https://doi.org/10.1155/2011/467563
17. Nastenko, I., Linnik, M., Honcharuk, M., Davydovych, I., Lutchenko, V., Babenko, V., Dolinchuk, L. (2025). Results of Machine Learning Applications in the Study of COVID-19 Associated Cardiopulmonary Pathology Using Computed Tomography Data. Innovative Biosystems and Bioengineering, 9 (4), 16–27. https://doi.org/10.20535/ibb.2025.9.4.325335
18. Sethi, S., Reddy, S., Sakarvadia, M., Serotte, J., Nwaudo, D., Maassen, N., Shi, L. (2025). Toward non-invasive diagnosis of Bankart lesions with deep learning. Medical Imaging 2025: Computer-Aided Diagnosis, 134. https://doi.org/10.1117/12.3046251
19. Xie, H., Gan, W., Bayerlein, R., Zhou, B., Chen, M.-K., Kulon, M. et al. (2026). Dose-aware diffusion model for 3D PET image denoising: Multi-institutional validation with reader study and real low-dose data. Medical Image Analysis, 111, 104039. https://doi.org/10.1016/j.media.2026.104039
20. Mahmood, M. F. (2024). Recognition and Categorization of Blood Groups by Machine Learning and Image Processing Method. Innovative Biosystems and Bioengineering, 8 (2), 53–68. https://doi.org/10.20535/ibb.2024.8.2.298201
21. Ho, K. C., Scalzo, F., Sarma, K. V., El-Saden, S., Arnold, C. W. (2016). A temporal deep learning approach for MR perfusion parameter estimation in stroke. 2016 23rd International Conference on Pattern Recognition (ICPR), 1315–1320. https://doi.org/10.1109/icpr.2016.7899819
22. Hess, A., Meier, R., Kaesmacher, J., Jung, S., Scalzo, F., Liebeskind, D. et al. (2019). Synthetic Perfusion Maps: Imaging Perfusion Deficits in DSC-MRI with Deep Learning. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, 447–455. https://doi.org/10.1007/978-3-030-11723-8_45
23. Asaduddin, M., Roh, H. G., Kim, H. J., Kim, E. Y., Park, S.-H. (2022). Perfusion Maps Acquired From Dynamic Angiography MRI Using Deep Learning Approaches. Journal of Magnetic Resonance Imaging, 57 (2), 456–469. https://doi.org/10.1002/jmri.28315

24. Kossen, T., Madai, V. I., Mutke, M. A., Hennemuth, A., Hildebrand, K., Behland, J. et al. (2023). Image-to-image generative adversarial networks for synthesizing perfusion parameter maps from DSC-MR images in cerebrovascular disease. Frontiers in Neurology, 13. https://doi.org/10.3389/fneur.2022.1051397
25. Talebi, S., Gai, S., Sossin, A., Zhu, V., Tong, E., Mofrad, M. R. K. (2024). Deep Learning for Perfusion Cerebral Blood Flow (CBF) and Volume (CBV) Predictions and Diagnostics. Annals of Biomedical Engineering, 52 (6), 1568–1575. https://doi.org/10.1007/s10439-024-03471-7
26. Cao, A., Le, P.-Y., Qie, Z., Guo, Y., Zaman, A., Lu, J. et al. (2024). Quantitative Perfusion Maps Using a Novelty Spatiotemporal Convolutional Neural Network. 2024 IEEE International Symposium on Biomedical Imaging (ISBI), 1–5. https://doi.org/10.1109/isbi56570.2024.10635136
27. McKinley, R., Hung, F., Wiest, R., Liebeskind, D. S., Scalzo, F. (2018). A Machine Learning Approach to Perfusion Imaging With Dynamic Susceptibility Contrast MR. Frontiers in Neurology, 9. https://doi.org/10.3389/fneur.2018.00717
28. Ahmed, Md. M., Hossain, Md. M., Islam, Md. R., Ali, Md. S., Nafi, A. A. N., Ahmed, Md. F. et al. (2024). Brain tumor detection and classification in MRI using hybrid ViT and GRU model with explainable AI in Southern Bangladesh. Scientific Reports, 14 (1). https://doi.org/10.1038/s41598-024-71893-3
29. Murase, K., Nakamoto, A., Tomiyama, N. (2024). Performance of recurrent neural networks with Monte Carlo dropout for predicting pharmacokinetic parameters from dynamic contrast-enhanced magnetic resonance imaging data. Journal of Applied Clinical Medical Physics, 26 (2). https://doi.org/10.1002/acm2.14586
30. Bakas, S., Sako, C., Akbari, H., Bilello, M., Sotiras, A., Shukla, G. et al. (2022). The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics. Scientific Data, 9 (1). https://doi.org/10.1038/s41597-022-01560-7