East European Journal of Physics. **2.** 563-577 (2026)
DOI:10.26565/2312-4334-2026-2-63 ISSN 2312-4334

STATISTICALLY CONDITIONED MRI DENOISING VIA FILM-MODULATED RESIDUAL ATTENTION U-NET

D.G. Sliusarenko[1,2]*, L.V. Sayan[2], A.V. Netreba[1]
[1]*Faculty of Radiophysics, Electronics and Computer Systems, National Taras Shevchenko University of Kyiv, 64/13, Volodymyrska Street, Kyiv, Ukraine, 01601*
[2]*National Cancer Institute of Ukraine, Yulii Zdanovskoi Street, 33/43, Kyiv, Ukraine, 03022*
**Corresponding Author email: d.fulhem@gmail.com*
Received February 15, 2026; revised April 6, 2026; accepted May 18, 2026

The quality of MRI images is often limited by spatially inhomogeneous noise, which negatively affects the accuracy of clinical interpretation and automatic analysis. Traditional deep learning methods often implicitly account for noise, leading to excessive smoothing and the loss of fine anatomical structures. In this paper, we propose an Enhanced Denoising U-Net architecture that employs a Feature-wise Linear Modulation (FiLM) mechanism to dynamically adapt to the noise profile of each slice. The model combines a vector of 8 statistical descriptors (including intensity, texture, and frequency characteristics), enabling dynamic control of the network's internal representations based on specific scanning conditions. To improve physical correctness, training was performed on data with synthetically generated k-space noise. The architecture is enhanced with residual blocks, attention mechanisms, and a multiscale processing module. On synthetic data, the average Peak Signal-to-Noise Ratio (PSNR) improvement was ≈ 20.7 dB, and with an average Structural Similarity Index (SSIM) improvement of approximately 0.73, indicating a deep restoration of structural information. In clinical images, an increase in SNR and stabilization of the coefficient of variation (CV) were observed, confirming the method's physical correctness. Clinical validation on complex contoured structures (hippocampus, brainstem, optic chiasm) showed an increase in the Dice coefficient (DSC) by 0.07–0.12 and a decrease in the HD95 error by 30–50%. The proposed method enables a transition from universal denoising strategies to adaptive reconstruction, ensuring high accuracy of preserving anatomical boundaries. This makes it a promising tool for MRI processing in neuroimaging tasks and variable therapy planning.
Keywords: *MRI; Attention U-Net; FiLM; Medical Imaging; Reconstruction; Noise; CNN; Contouring*
PACS: 87.57.N-, 07.05.Mh, 87.61.Tg, 02.50.-r

1. INTRODUCTION

Magnetic resonance imaging (MRI) is one of the primary modalities of modern medical imaging due to its high spatial resolution, excellent soft tissue contrast, and the absence of ionizing radiation exposure [1]. At the same time, the quality of MRI images is significantly limited by noise, which reduces the signal-to-noise ratio (SNR), complicates visual interpretation, and degrades the accuracy of subsequent computational procedures, including segmentation, quantitative analysis, and contouring of anatomical structures in radiation therapy [2].

The physical nature of noise in MRI is complex and multicomponent [3]. At the stage of signal formation in k-space, noise typically follows a complex Gaussian distribution; however, after nonlinear reconstruction and transition to the magnitude image domain, its statistics change, acquiring a Rician distribution, especially in areas with a low signal level [4]. Additionally, the noise characteristics are influenced by the magnetic field strength, pulse sequence type, reconstruction parameters, and the signal's spatial frequencies. As a result, noise in MRI is spatially heterogeneous and statistically unstable even within a single study [5].

Traditional denoising methods, such as linear filtering, Wiener filter, Non-Local Means (NLM), or Block-matching and 3D filtering (BM3D), are usually based on simplifying assumptions about stationarity and homogeneity of noise [6]. Although these approaches can effectively reduce random fluctuations, they often lead to the loss of high-frequency components responsible for fine anatomical boundaries or exhibit unstable behaviour under changing noise conditions. With the advent of deep learning, a significant amount of work has been devoted to applying convolutional neural networks to MRI denoising. Architectures based on U-Net and its modifications, as well as models such as the Deep Convolutional Neural Network (DnCNN) and the Residual Encoder-Decoder Convolutional Neural Network (RED-CNN), have demonstrated significant improvements in Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) compared to classical methods [7]. However, most of these approaches treat noise implicitly - as an integral part of the input image - without explicitly considering its statistical characteristics. This forces the model to generalize a wide range of noise modes within a single latent space, which complicates the training process and increases the risk of over-smoothing [8].

A separate group consists of methods focused on reference-free learning (Noise2Noise, Noise2Void)[9], and approaches that take into account local noise statistics. Although they reduce the dependence on reference data, such models usually lack a mechanism for explicit adaptation to the specific noise profile of each image and remain sensitive to changes in scanning conditions.

Cite as: D.G. Sliusarenko, L.V. Sayan, A.V. Netreba, East Eur. J. Phys. 2, 563 (2026), https://doi.org/10.26565/2312-4334-2026-2-63

This paper proposes an alternative approach in which noise is considered as a separate object of analysis with an explicit statistical description. For each input MRI slice, a vector of statistical descriptors is computed to quantitatively characterize the noise's intensity, texture, and frequency properties. This vector is used to control the internal representations of the neural network using the Feature-wise Linear Modulation (FiLM) mechanism [10], which allows the model to dynamically adapt the denoising process to specific image formation conditions.

The proposed model is based on a modified U-Net architecture with residual blocks, attention mechanisms [11] and multi-scale processing. The performance evaluation is performed using a multi-level scheme that includes reference metrics on synthetic and phantom data, reference-free metrics on real clinical images, and clinically oriented evaluation through contouring of anatomical structures. This approach allows us to evaluate not only the formal quality of the image, but also its practical utility in clinical tasks [12].

2. Materials and Methods

2.1. Data

The study used MRI brain DICOM (Digital Imaging and Communications in Medicine) images obtained under standard clinical conditions on MRI scanners operating at magnetic field strengths of 1.5 T and 3 T. The dataset included T1-weighted, T2-weighted, and FLAIR sequences, which provide different contrast types and enable the assessment of the model's stability of the proposed method under conditions of significantly different intensity and texture characteristics. In total, the dataset contained 400 patients (approximately 10,000 slices), with training, validation, and test samples distributed at the patient level in a 70% / 15% / 15% ratio. This approach prevents slices from the same study from appearing in different subsets and prevents information leakage between the training and evaluation stages.

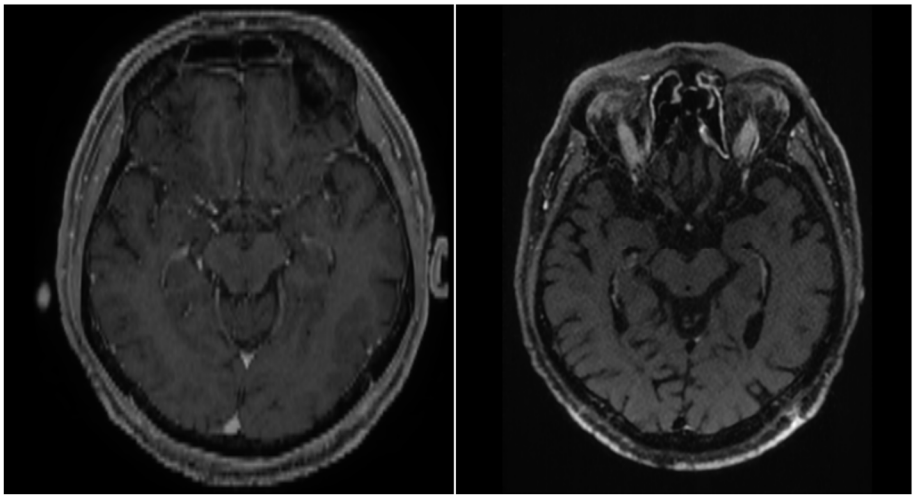

Figure 1. Examples of images used from the dataset

All images were reduced to a single spatial format of 512 × 512 pixels. Intensities were normalized to the range [0, 1] via linear scaling after clipping the extreme 0.5% of the histogram values, thereby reducing the influence of individual outliers and ensuring a fair comparison of images obtained with different scanning parameters.

Since obtaining noise-free MRI images in clinical conditions is practically impossible, a physically based synthetic noise modeling approach was used to form training pairs. In contrast to methods that add noise directly in the image domain, in this work, the noise is modeled in the frequency domain (k-space) followed by an inverse Fourier transform. This allows us to reproduce the spatial-frequency properties of noise characteristic of MRI and avoid non-physical artifacts [13].

2.2. Noise Generation and Training Data Preparation

To train the neural network, an algorithmic approach to generating noise degradations was used, directly implemented in the model code. The goal of this stage was to create training examples that are statistically and structurally close to real MRI images while remaining controllable in terms of the level and type of noise.

The input information for noise generation consists of reconstructed MRI images in magnitude format, normalized to the range [0,1]. Since the original complex k-space data in clinical DICOM images are not available, the work uses pseudo-k-space [14], obtained via a direct discrete Fourier transform:

$$K(u,v) = \mathcal{F}\{I(x,y)\},$$

where $I(x,y)$ - input image, and $K(u,v)$ - its spectral representation.

In the frequency domain, a random complex quantity is added to each component of the spectrum, formed as the sum of the real and imaginary parts, generated with a zero mean value[15]:

$$\widetilde{K}(u,v) = K(u,v) + \alpha(u,v)\,[n_r(u,v) + i\,n_i(u,v)],$$

where: $n_r, n_i \sim \mathcal{N}(0,1)$ - independent Gaussian random variables, $\alpha(u,v)$ - noise factor scale, i - imaginary unit.

Scale factor $\alpha(u,v)$ is defined as a function of the radial distance from the center of k-space and allows modeling the non-uniform distribution of noise over spatial frequencies. In implementation, this is achieved by multiplying the random noise by a pre-generated frequency mask that amplifies or attenuates the noise in the high-frequency components.[16] This approach allows the simulation of scenarios in which noise affects not only the background, but also the small structural details of the image.

After adding noise in the spectral domain, an inverse Fourier transform is performed[17]:

$$\tilde{I}(x,y) = \mid \mathcal{F}^{-1}\{\widetilde{K}(u,v)\} \mid,$$

where the modulus of the complex signal is taken, which corresponds to the standard procedure for forming an MRI image in clinical systems. The result is a noisy image with nonlinear statistical properties, particularly an asymmetric intensity distribution in low-signal regions. Accordingly, the intensity distribution in the reconstructed amplitude image is transformed from Gaussian to Rice distribution.

In addition, depending on the experimental configuration, auxiliary degradations can be superimposed in the image space, in particular, additive Gaussian noise with low variance, Poisson-like intensity fluctuations, and weak blurring to simulate the loss of high-frequency components. These operations are used not as the main noise model, but as a means of increasing the network's resistance to mixed types of degradations.

An important feature of the implementation is that for each generated noisy image, a vector of statistical characteristics is immediately calculated. Thus, each training pair has the form:

$$(I_{\text{noisy}}, \mathbf{s}, I_{\text{clean}}),$$

where $\mathbf{s}$ is a statistical description of the noise present in the image I_{noisy}.

This provides internal consistency between the noise degradation and the control signal fed into the FiLM module of the neural network. Unlike approaches where the noise level is specified only by a scalar or a fixed parameter, the proposed implementation uses a multidimensional statistical description that allows the model to consider not only the amplitude, but also the shape of the noise distribution and its spatial properties.

2.3. Calculation of statistical noise features

One of the main differences of the proposed approach is the explicit quantitative description of the noise characteristics of each input MRI image. Instead of assuming a fixed or averaged noise regime, the work uses a dynamic statistical analysis performed separately for each slice before feeding the results to the neural network.

For each input MRI slice, a vector of statistical features with dimensionality N = 8 was calculated [18]:

$$s = [\mu, \sigma, \gamma_1, \gamma_2, C_{loc}, B, D_{edge}, H_{loc}]$$

All vector components were scaled using standard normalization to zero mean and unit variance, which ensures stability of training and correct integration of statistical information in the FiLM module.

Mean intensity value

$$\mu = \frac{1}{N}\sum_{i=1}^{N} I_i,$$

where I_i - pixel intensity value, N - total number of pixels in an image. This value reflects the overall signal level and allows for partial compensation of histogram shifts caused by noise or background heterogeneity.

Standard Deviation

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_i - \mu)^2},$$

which is used as a basic measure of the amplitude of noise fluctuations.

Skewness

$$\gamma_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{I_i - \mu}{\sigma}\right)^3,$$

which characterizes the shift of the intensity distribution relative to the symmetric case. Non-zero asymmetry is typical for MRI images with low signal-to-noise ratio, especially in regions with low signal amplitude.

$$\gamma_2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{I_i - \mu}{\sigma}\right)^4 - 3,$$

which reflects the distribution's kurtosis and the presence of heavy tails characteristic of noise impurities and local artifacts.

Local contrast was calculated as the average local standard deviation in sliding windows of size w×w, where w=7 pixels. This indicator is sensitive to small-scale intensity fluctuations and allows us to distinguish textural features of tissues from random noise fluctuations [19].

The blur measure is estimated from the variance of the Laplace operator response:

$$B = \mathrm{Var}\,(\nabla^2 I),$$

where ∇^2 - Laplacian of the image. A decrease in this value corresponds to the loss of high-frequency components and is an indirect indicator of both excessive smoothing and defocusing.

Edge density is defined as the ratio of the number of pixels detected by the Canny algorithm to the total number of pixels. This indicator allows us to assess the degree of "graininess" of the image and the number of pseudo-contours generated by noise.[20]

Local entropy

$$H = -\sum_{k} p_k \log\, p_k,$$

where p_k - probability of intensities in a local window. Entropy is used as a generalized measure of texture complexity and the degree of disorder in an image.

This set of features was chosen as a combination of intensity (to estimate overall noise) and texture (to distinguish noise from anatomical structures) characteristics [21],[22].

The computation was performed in parallel using Graphics Processing Units (GPU), which allowed processing of large image sets without significantly increasing data preparation time.

2.4. Mathematical formulation of the noise reduction problem

Let $I_{\text{clean}} \in \mathbb{R}^{H\times W}$ - clean (noise-free) MRI image, and I_{noisy} - corresponding image distorted by noise. The process of forming a noisy image can be generalized as:

$$I_{\text{noisy}} = \mathcal{D}(I_{\text{clean}}, \eta),$$

where $\mathcal{D}(\cdot)$ - degradation operator, and η - a set of parameters describing noise processes.

The denoising task can be formulated as learning a mapping:

$$\hat{I} = f_\theta(I_{\text{noisy}}, \mathbf{s}),$$

where f_θ - neural network with parameters θ, $\mathbf{s}$ - vector of statistical noise features calculated for a specific image, $\hat{I}$ - restored image.

The use of an additional condition vector s is consistent with conditional modulation approaches of neural networks. Thus, the model not only analyses the spatial structure of the noisy image but also uses additional information about the statistical profile of the noise, which allows adapting the reconstruction process to specific conditions.

The model was trained by minimizing the combined loss function

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{MSE}} + \lambda_2(1 - \text{SSIM}) + \lambda_3 \mathcal{L}_{\text{GDL}},$$

where:

- $\mathcal{L}_{\text{MSE}}$ - standard error between $\hat{I}$ and I_{clean};
- SSIM - structural similarity index.
- $\mathcal{L}_{\text{GDL}}$ - loss of gradient difference, which imposes a penalty for boundary distortion [23];
- λ_i - weighting factors that determine the contribution of each component.

The weighting coefficients were chosen empirically and defined as 0.6, 0.3, 0.1, respectively.

This choice of loss function provides a compromise between global accuracy of intensity recovery and preservation of local anatomical structures [24].

2.5. Enhanced Denoising U-Net Architecture

The proposed architecture is based on the classical U-Net [25]. However, it has been significantly extended to adapt to the statistically heterogeneous noise of MRI images. The main idea is to combine spatial-contextual processing with dynamic feature modulation driven by the statistical characteristics of the input image.

The architecture consists of an encoder, a symmetric decoder, and multi-scale processing, combining multi-scale feature blocks with attention-guide filters, as shown in similar designs for high-quality denoising [26]. Residual blocks with normalization and nonlinearity are used at each level, ensuring stable gradient propagation and reducing the risk of feature degradation as the network depth increases. The integration of attention mechanisms into the residual U-Net architecture allows the network to better capture structural details when denoising [27]. To illustrate the overall organization of the model, Fig. 2 shows a diagram of the Enhanced Denoising U-Net architecture, which depicts the hierarchical encoder–decoder structure, the locations of residual blocks and attention mechanisms, and the integration point for FiLM modulation. The diagram reflects the logic of multi-level feature adaptation depending on the statistical noise profile.

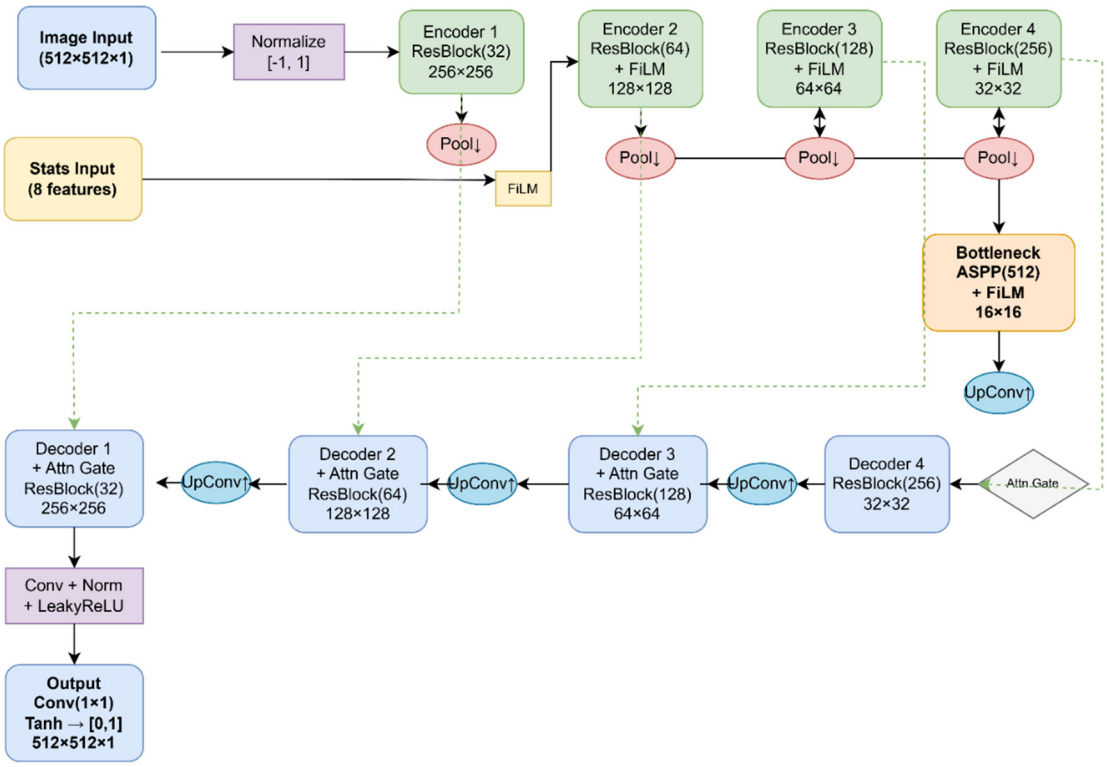

Figure 2. Architecture of the proposed Enhanced Denoising U-Net model with residual blocks, attention mechanisms and FiLM modulation controlled by a statistical noise vector. The encoder–decoder structure, multi-scale feature processing and integration of statistical information at different levels of the network are shown

The main feature of the model is that the statistical feature vector s is used not only at the input but also integrated into the network's internal representations at different levels via the Feature-wise Linear Modulation mechanism. For each layer, a pair of scaling parameters is formed $\gamma(\mathbf{s})$ and displacement $\beta(\mathbf{s})$, which depend linearly on the statistical vector:

$$F_{out} = \gamma(\mathbf{z}) \odot F_{in} + \beta(\mathbf{z})$$

where:

- F_{in} - input layer feature tensor,
- F_{out} - output feature tensor,
- $\mathbf{z}$ - vector of statistical noise descriptors,
- $\gamma(\mathbf{z})$, $\beta(\mathbf{z})$ - scaling and displacement vectors generated by a multilayer perceptron network (MLP),
- $\odot$ - component-wise multiplication.

From a physical point of view, this means that the network dynamically changes the sensitivity of its filters depending on the noise level, its asymmetry, entropy and other characteristics. For example, with a high standard deviation and increased entropy, the model automatically reduces the weight of high-frequency channels, while for images with a low noise level, a larger number of fine structural details are preserved.

Thus, the architecture implements not a fixed, but a context-dependent denoising strategy, which is fundamentally important for working with clinical MRI data, which are characterized by significant variability in scanning conditions.

2.6. Model training procedure

The neural network was trained in supervised mode using pairs of noisy-reference images generated by the noise-generation algorithm. For each training example, the model received two inputs: a noisy MRI image and the corresponding vector of statistical noise characteristics.

The model was implemented using the TensorFlow/Keras framework. The network depth is 4 encoder-decoder levels with an initial number of filters of 32, which are doubled at each subsequent level. The maximum number of channels in the bottleneck layer is 512.

The decoder employs transposed convolutions with a stride of 2 for learnable upsampling rather than fixed interpolation, enabling the model to optimize spatial resolution restoration.

The AdamW optimizer was used in the training process[28] with an initial learning rate "lr"=1×10^{-4} and a weight decay coefficient =1×10^{-5}, which provides stable convergence and additional regularization due to weight decay. The minibatch size of 32 images for a maximum of 100 epochs was selected considering memory limitations and Layer Normalization features [29], and training continued until the value of the loss function on the validation sample stabilized.

The learning rate was reduced using the Reduce-on-Plateau strategy: if validation loss did not improve for 5 consecutive epochs, it was halved. The early stopping mechanism was applied with patience for 10 epochs, with the best model weights restored.

All experiments were performed with fixed initial values of the random number generator to ensure reproducibility of the results.

A separate stage of the implementation is GPU-accelerated calculation of statistical noise features. An eight-component vector of statistical descriptors was formed batchwise using PyTorch, which significantly reduced the time for reprocessing large sets of MRI images. Normalization of statistical features was performed using StandardScaler, trained exclusively on the training sample, which prevents information leakage between data subsets.

2.7. Image quality assessment metrics

2.7.1. Reference metrics

Reference metrics were used in cases where a reference (clean) image or its analytical equivalent is available, allowing for a direct quantitative comparison of the reconstruction results. Such metrics provide an objective assessment of the reconstruction accuracy, the level of noise reduction, and the preservation of structural features of the image.

Peak Signal-to-Noise Ratio[30] is defined as:

$$PSNR = 10\log_{10}\left(\frac{I_{\max}}{MSE}\right),$$

where $I_{\max}$ - maximum possible pixel intensity value, and

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(I_i - \hat{I}_i\right)^2$$

standard error between the reference I and restored Î images. PSNR is a generalized global metric that reflects the effectiveness of noise reduction and the level of deviation from the reference signal. At the same time, it does not account for the spatial distribution of errors and the image's structural organization, which limits its informativeness in a clinical context.

The Structural Similarity Index is used to assess the degree of preservation of the local structure of the image by comparing the brightness, contrast, and textural characteristics between the reference and reconstructed images [31] defined as

$$SSIM(I,\hat{I}) = \frac{(2\mu_I\mu_{\hat{I}} + C_1)(2\sigma_{I\hat{I}} + C_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)},$$

where μ_I and $\mu_{\hat{I}}$ - average intensity values, σ_I^2 and $\sigma_{\hat{I}}^2$ - dispersion, $\sigma_{I\hat{I}}$ - covariance between reference and reconstructed images, C_1 and C_2 - small constants that prevent numerical instability.

Unlike PSNR, SSIM correlates with the features of human visual perception and is more sensitive to structural distortions. In MRI tasks, this metric is of fundamental importance, since even minor deformations of anatomical structures or local texture can lead to erroneous clinical conclusions.

Edge Preservation Index (EPI)[32] is used to quantify the preservation of contours and high-frequency spatial components of an image after processing. Defined as

$$EPI = \frac{\sum_{x,y}\nabla I(x,y)\,\nabla\hat{I}(x,y)}{\sum_{x,y}[\nabla I(x,y)]^2},$$

where ∇I and $\nabla\hat{I}$ - spatial gradients of the reference and reconstructed images, respectively.

The metric is based on a comparison of gradient or filtered (e.g., using Sobel or Laplace operators) representations of the reference and reconstructed images and characterizes the degree of preservation of boundaries between different anatomical structures. High EPI values indicate minimal contour smoothing and no loss of fine details, which is critical for subsequent tasks such as segmentation, contouring, and quantitative analysis of MRI data.

2.7.2. No-Reference metrics

For real clinical MRI images, where there is no reference (clean) signal, reference-free quality metrics were used, which allow us to evaluate the effectiveness of denoising based on the statistical characteristics of homogeneous and tissue-specific regions. Such metrics are particularly important in the clinical context, as they reflect changes in noise properties and contrast without requiring access to the true image.

Signal-to-Noise Ratio is defined as the ratio of the mean signal value to the standard deviation of the noise in a homogeneous region of interest (ROI)[33]:

$$SNR = \frac{\mu_{\text{ROI}}}{\sigma_{\text{ROI}}}.$$

An increase in SNR after denoising indicates reduced random noise fluctuations and greater signal stability. However, an excessive increase in this indicator may indicate excessive smoothing and loss of fine-texture details, which are critical for accurate visual and quantitative assessment of MRI.

The Coefficient of Variation (CV) is calculated as:

$$CV = \frac{\sigma_{\text{ROI}}}{\mu_{\text{ROI}}},$$

and characterizes the relative heterogeneity of intensities within a single tissue region. This indicator is invariant to the absolute scale of the signal and allows comparing the noise level between different images or patients.[34] The decrease in CV after processing corresponds to stabilization of intensities and increased homogeneity of the tissue signal.

Contrast-to-Noise Ratio (CNR) used to evaluate the ability of a denoising algorithm to preserve or enhance contrast between different tissues[35]:

$$CNR = \frac{|\,\mu_1 - \mu_2\,|}{\sqrt{\sigma_1^2 + \sigma_2^2}},$$

where indices 1 and 2 correspond to different tissue regions of interest.

An increase in CNR after denoising indicates improved resolution of tissue structures, but, similarly to SNR, may be due to both noise reduction and potential changes in true tissue contrast. Therefore, interpretation of CNR should be done in conjunction with visual assessment and other quantitative measures.

2.7.3. Clinical assessment metrics

Quantitative assessment of denoising quality using intensity and structural metrics does not always allow unambiguous determination of the method's clinical usefulness. In radiotherapy and neuroimaging tasks, the accuracy of anatomical boundaries is of great importance, as it determines the accuracy of contouring of target organs and organs at risk. In this regard, this work uses a clinically oriented approach to assessing denoising quality, based on analysis of anatomical structure segmentation results.

Dice Similarity Coefficient (DSC) is a standard metric for assessing the degree of overlap between two segmentations[36] and is defined as

$$\text{DSC} = \frac{2\,|\,A \cap B\,|}{|\,A\,| + |\,B\,|},$$

where A and B - the set of pixels or voxels corresponding to a segmented structure in two images.

The DSC value varies from 0 to 1, where 1 corresponds to complete coincidence of the segmentations. In the context of this study, the increase in DSC after denoising is interpreted as an improvement in the reproducibility of the structure boundaries and a reduction in the ambiguity in their visual identification. At the same time, it should be noted that DSC is a volumetric metric and can be relatively insensitive to local contour deviations, especially for large structures.

For a more detailed assessment of the accuracy of the boundaries, the 95th percentile of the Hausdorff distance was used, which is defined as the value below which 95% of all distances between corresponding contour points lie.[37]:

$$\text{HD}_{95}(A, B) = \text{percentile}_{95}\left(\min_{b \in B} \| a - b \|, a \in A\right).$$

Unlike the classical Hausdorff distance, HD95 is robust to outliers and local artifacts. The decrease in HD95 after denoising indicates improved local contour accuracy and a reduction in small but clinically significant deviations that can affect treatment planning.

3. RESULTS

3.1. Quantification on synthetically noisy data (reference metrics)

For quantitative analysis of denoising efficiency, a set of MRI images with added synthetic noise generated in the frequency domain (k-space) was used, providing a physically accurate model of signal degradation. For each image, the reference metrics PSNR, SSIM, EPI, as well as derived indicators of quality improvement after denoising the synthetic noise were calculated. The corresponding graphs for 300 randomly selected images are shown in Fig. 3, with each point representing a separate MRI slice.

From the element-wise analysis graphs (Fig. 3) it can be observed that the PSNR values for noisy images fluctuate mainly in the range of ≈13–25 dB, which corresponds to a significant level of signal degradation. After applying the proposed denoising method, the PSNR values for almost all images stabilize at ≈35–45 dB, with an average value of about 40 dB. The average PSNR gain is ≈20.7 dB, indicating not merely local or isolated improvements but a globally stable effect of the model across the entire test sample. A similar trend is observed for the structural similarity index SSIM. For noisy images, SSIM values are low and highly variable (typical range ≈0.05–0.4), reflecting a significant loss of local anatomical structure. After denoising, the SSIM values are concentrated in the range of ≈0.90-0.98 with low dispersion, indicating the restoration of not only intensity characteristics but also the spatial organization of tissues. The average increase in SSIM is ≈0.73, which is typical for the transition from highly noisy to structurally restored MRI images.

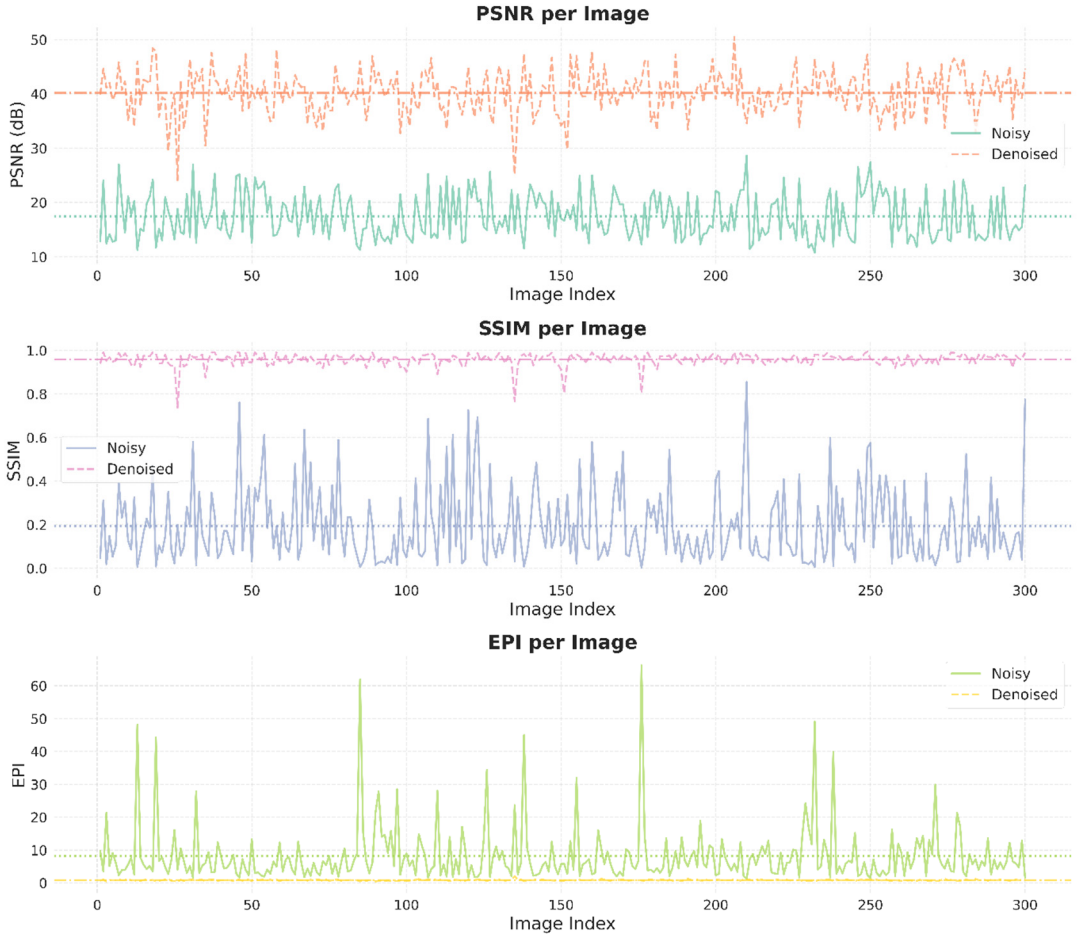

Figure 3. Element-wise analysis of denoising results for 300 randomly selected MRI slices of the test set: PSNR, SSIM and EPI values before and after processing. Each point corresponds to a separate image.

The boundary preservation metric demonstrates a fundamentally important result. Noisy images are characterized by high EPI variability and pronounced peak values, due to the appearance of pseudo-contours induced by noise. After denoising, the EPI values decrease sharply and stabilize at a low level, indicating effective suppression of noise pseudo-edges while preserving true anatomical boundaries. It is important that such stabilization is not accompanied by signs of aggressive smoothing, which usually leads to the loss of contour information.

Thus, the element-wise analysis confirms that the proposed model not only reduces the average noise level but also provides a stable quality improvement for each image without degrading structural information. Taken together, the results of the elementwise and distributional analysis demonstrate that the model provides a stable, statistically homogeneous improvement in image quality without anomalous artifacts, local dips, or unstable reconstruction modes.

3.2. Evaluation on real clinical MRI images (reference-free metrics)

The evaluation of denoising efficiency on real clinical MRI images was performed without reference data, as in practical conditions, it is impossible to obtain "perfectly clean" images. In this regard, the quality analysis was based on reference-free, physically interpreted indicators, in particular SNR and CV, calculated in homogeneous regions of interest (ROIs), corresponding to tissue-homogeneous brain areas. The results of reference-free image quality assessment on clinical MRI data are presented in Fig. 4 and Fig. 5, which show the change in SNR and CV in homogeneous ROIs before and after denoising.

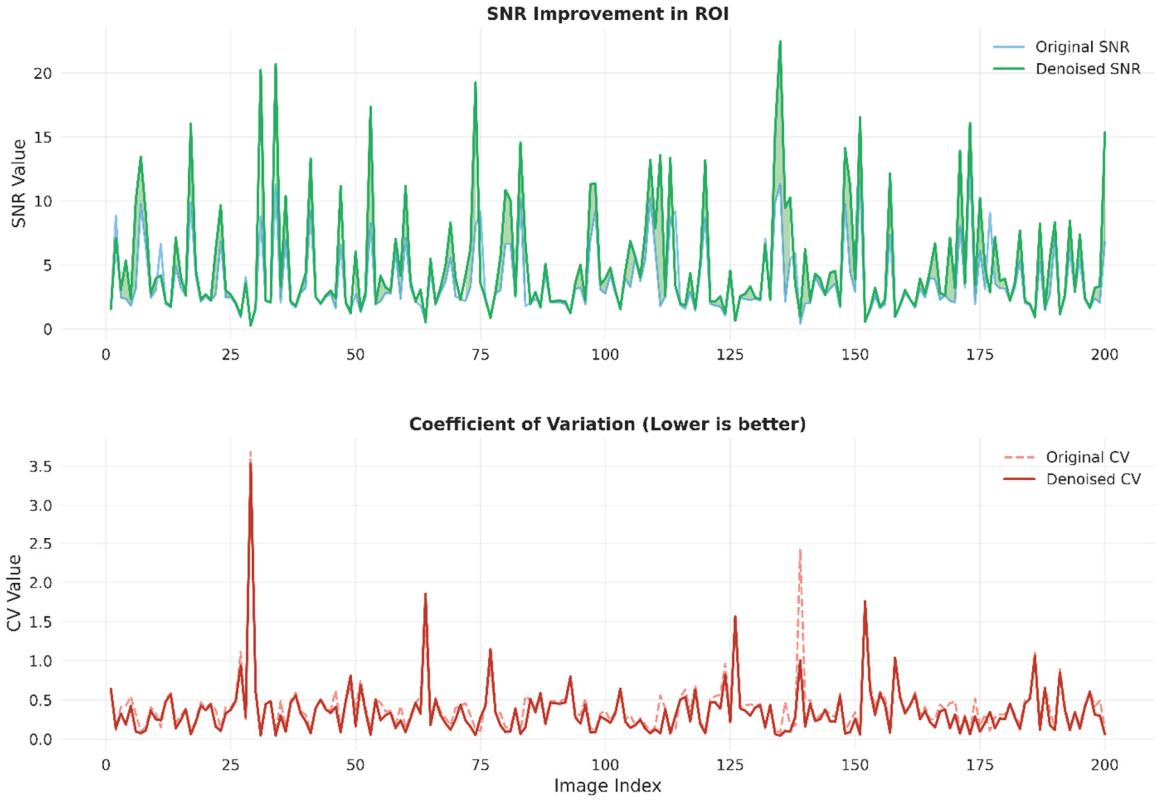

Figure 4. Comparison of Signal-to-Noise Ratio values in homogeneous regions of interest in clinical MRI images before and after noise reduction.

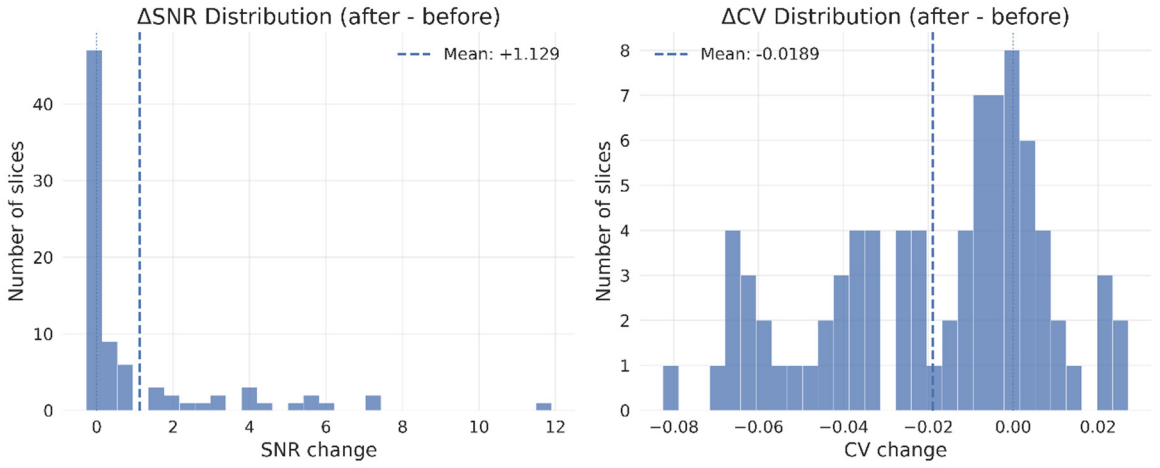

Figure 5. Change in coefficient of variation in homogeneous tissue ROIs before and after denoising. A decrease in CV indicates signal stabilization without transitioning to over-smoothing.

Figure 4 shows that the obtained results demonstrate a systematic increase in the signal-to-noise ratio (SNR) after applying the proposed denoising method. For most ROIs, the SNR values in the original clinical MRI images were in the range of ≈2.0–6.0, which corresponds to the typical level of noise fluctuations in low- and medium-contrast brain regions. After denoising, a stable increase in SNR is observed for most slices: typical improvement values are ≈+1.0–3.0, and in some ROIs they reach +5 and more. The distribution of ΔSNR (Fig. 4, bottom panel) is characterized by a positive mean

value of +1.13, which indicates a systematic rather than isolated improvement in signal quality. Importantly, the increase in SNR is not accompanied by abnormally high peaks, indicating the absence of artificial intensities scaling or aggressive filtering.

In parallel with the increase in SNR, a decrease in the coefficient of variation (CV) is observed, which for the original images in a number of ROIs reached values of ≈0.3–1.0, and in isolated cases exceeded 2.0, which reflects significant intratissue signal heterogeneity caused by noise. After denoising, the CV values for most ROIs decrease to ≈0.15–0.45. The distribution of CV changes (ΔCV) has a negative mean value of −0.019, which corresponds to a moderate but stable decrease in the relative signal variability. It is fundamentally important that the CV does not tend to zero values, which indicates the preservation of the natural textural heterogeneity of tissues and the absence of excessive smoothing, which is unphysical for real MRI reconstructions.

Thus, the simultaneous increase in SNR and the controlled decrease in CV confirm that the proposed method achieves physically correct signal stabilization, balancing noise suppression with the preservation of anatomically significant intensity variability. Noise fluctuations are suppressed, while the natural variability of tissue texture and anatomical heterogeneity are preserved. This means that the model operates in an adaptive reconstruction mode, where the balance between noise suppression and structure preservation is maintained dynamically, according to the local statistical characteristics of the image.

3.3. Clinical assessment through contouring of anatomical structures

Since formal image quality metrics (PSNR, SSIM, SNR, CV, etc.) do not always fully reflect the clinical utility of denoising [38], In addition, a clinically oriented assessment was performed based on the task of contouring anatomical structures. This approach allows us to assess not only the visual quality of the reconstruction, but also its practical suitability for tasks that are critical in neuroimaging and radiotherapy.

The study was conducted on 100 brain MRI series that included standard clinical protocols. For each series, contouring of three anatomically and clinically significant structures was performed, namely the hippocampus, brainstem, and optic chiasm. The choice of these structures is of fundamental importance. The hippocampus is characterized by complex geometry, high curvature of the contours, and low contrast with adjacent tissues, which makes it sensitive to noise and smoothing. The brainstem is a relatively large but morphologically complex structure with extended boundaries and transitions between different tissue types. The optic chiasm, in turn, is a small structure with indistinct boundaries and high clinical significance, which makes it's contouring particularly sensitive to image quality.

Contouring was performed in a 3D Slicer environment [39] to ensure reproducibility on standard computing systems on the original (noisy) MRI images and on images after applying the proposed denoising method. Examples of contours made on the original and denoised images are shown in Fig. 6, which allows to visually assess the effect of denoising on the stability and clarity of anatomical boundaries.

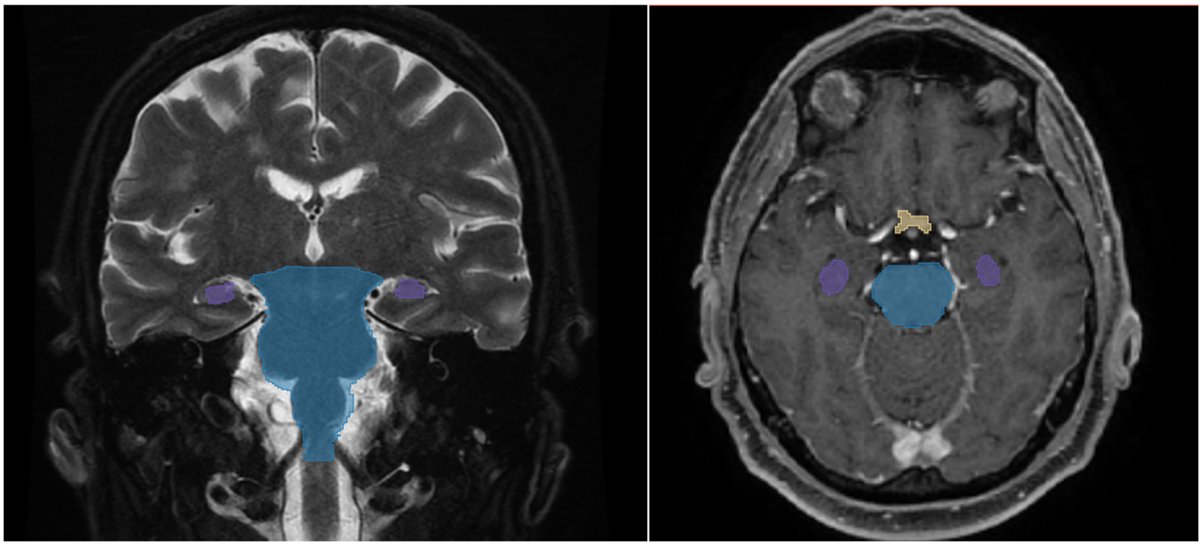

Figure 6. Examples of contouring of anatomical structures (hippocampi, brainstem, optic chiasm) on original noisy and denoised MRI images

The contouring procedure was performed by radiologists with clinical experience, which introduces a subjective component in the segmentation results. In this regard, contouring is considered not as an absolute quality metric, but as an auxiliary clinical evaluation, reflecting the practical convenience of working with the image, the clarity of the boundaries of structures and the reproducibility of anatomical landmarks.

For quantitative analysis of the segmentation results, the Dice Similarity Coefficient and the 95th percentile of the Hausdorff distance metrics were used, which allows simultaneously assessing the volumetric overlap of segmentations and the local accuracy of boundary reproduction.

The averaged results over 100 series demonstrate a stable positive dynamic of segmentation indicators after the application of denoising. In particular, the average DSC values for noisy images were in the range of approximately 0.78-0.85 for the brainstem, 0.72–0.80 for the hippocampus, and 0.65–0.75 for the optic chiasm, reflecting the difficulty of segmenting these structures in noisy conditions. After denoising, the average DSC values increase to levels of ≈0.88–0.93 for the brainstem, ≈0.82–0.88 for the hippocampus, and ≈0.78–0.85 for the chiasm, corresponding to an absolute increase in DSC of 0.07–0.12 depending on the structure. A similar trend is observed for the HD95 metric. For the original noisy images, the HD95 values were in the range of ≈2.5–4.5 mm for the hippocampi, ≈2.0–3.5 mm for the brainstem, and ≈3.0–5.0 mm for the optic chiasm, indicating significant local errors at the boundaries of the structures. After denoising, the HD95 decreases to ≈1.2–2.5 mm for the hippocampi, ≈1.0–2.0 mm for the brainstem, and ≈1.5–2.8 mm for the chiasm, corresponding to an average reduction of local contour errors of 30–50%. In fact, a rather pronounced effect is observed for the hippocampi and optic chiasm, i.e., precisely for the structures with the most complex geometry and low contrast, where the impact of noise is critical. This suggests that denoising has the greatest clinical benefit for anatomically complex and low-contrast structures, where even minor intensity fluctuations can lead to significant contouring errors.

Fig. 7 shows examples of superimposing “old” contours performed on the original images and “updated” contours constructed after applying the proposed denoising method. Since both sets of contours were generated in the same software environment and according to an identical protocol, the differences observed can be directly attributed to changes in image quality. For clarity, already processed images are used.

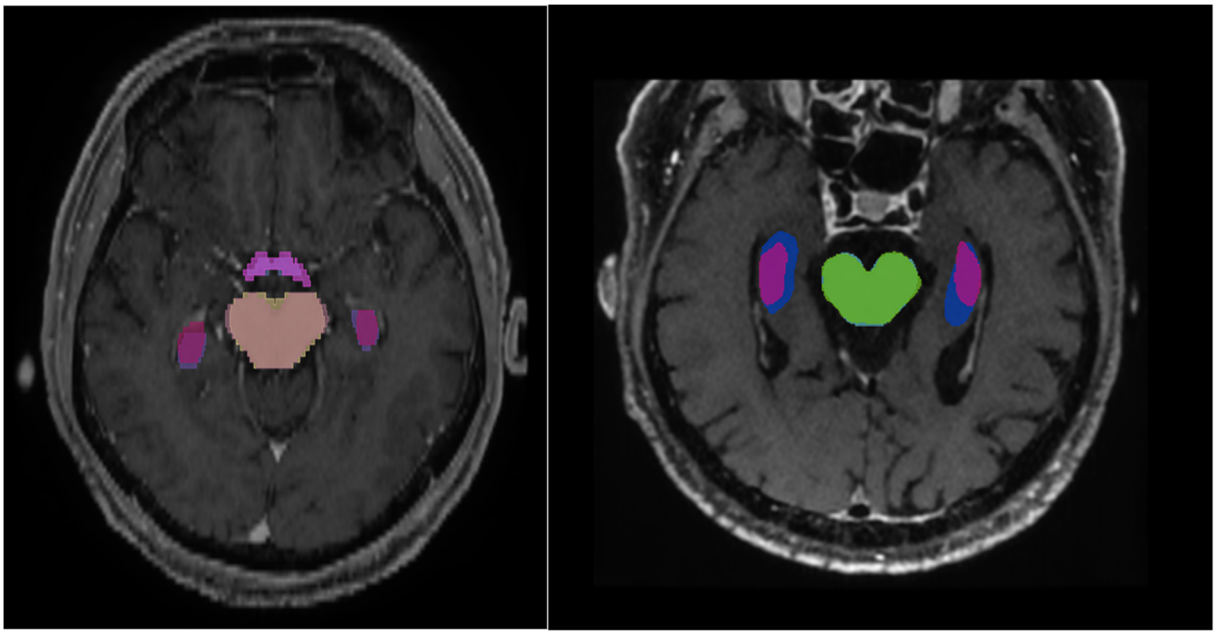

Figure 7. Overlaying contours of anatomical structures built on the original noisy MRI images (old contours) and on denoised images (updated contours)

Qualitative analysis of contours shows that after denoising, the boundaries of structures become more stable, the number of local irregularities, pseudo-contours and fragmentation characteristic of noisy images decreases. This is especially evident in areas of fine anatomical transitions, where noise usually creates false gradients that complicate the interpretation of boundaries. It is important to note that, due to the presence of the human factor, the results of contouring cannot be considered as a strictly objective metric validation of the method. However, it is this subjective component that makes such an analysis clinically relevant: improved contour reproducibility, reduced boundary uncertainty and increased segmentation stability reflect the real practical value of denoising in a clinical environment. Thus, the contouring results confirm that the proposed denoising method not only improves formal image quality metrics, but also has a practical clinical effect, increasing the ease of interpretation of MRI images, contouring stability, and reproducibility of anatomical boundaries in neuroimaging and radiotherapy planning tasks.

3.4. Comparative analysis with modern noise reduction solutions

To verify the effectiveness of the developed method, a series of experiments were conducted to compare the Enhanced Denoising U-Net architecture with a wide range of algorithms - from classical statistical filters to modern deep learning models. The selection of models for comparison covers the evolution of denoising approaches: from non-local self-similarity methods[40],[41] (NLM, BM3D) to specialized convolutional networks[42],[43] (DnCNN, RED-CNN) and architectures with spatial attention mechanisms (Attention U-Net).

The averaged comparison results for PSNR and SSIM metrics, as well as the analysis of model adaptation strategies to noise conditions are presented in Table 1 below.

Table 1. Comparative analysis of the effectiveness of MRI image denoising for classical methods and modern neural network architectures by PSNR and SSIM indicators, as well as by the type of adaptation to noise conditions.

Method	Architectural paradigm	PSNR (dB)	SSIM	Type of noise adaptation
NLM	Nonlocal Filtering	27.4	0.68	Fixed (spatial)
BM3D	Spectral Clustering	30.2	0.74	Statistical approximation
DnCNN	Residual CNN	34.8	0.86	Implicit (learning)
RED-CNN	Encoder-Decoder	36.1	0.88	Implicit (latent)
Standard U-Net	Symmetric U-Net	37.5	0.91	Spatial-contextual
Attention U-Net	Contextual Attention	38.9	0.93	Selective (Attention Gates)
Proposed Model	**Residual Attention + FiLM**	**42.3**	**0.96**	**Explicit (FiLM modulation)**

Analysis of the obtained data allows us to identify a clear correlation between the method used to account for noise characteristics and the quality of the final reconstruction. Classical approaches (NLM, BM3D), despite their mathematical validity, exhibit the lowest structural similarity indicators, due to their inability to distinguish complex textural patterns in medical tissues from high-frequency noise.

The transition to deep learning methods provides a significant increase in PSNR, however, architectures such as DnCNN and RED-CNN often face the problem of "general averaging", when the model tries to learn a universal mapping for the entire range of noise modes. This leads to the undesirable effect of smoothing (over-smoothing), which is critical when visualizing small anatomical structures.

The proposed model solves this problem by implementing the adaptive reconstruction paradigm. Unlike the basic U-Net or attention models, where noise is considered an integral part of the input signal, our architecture uses an explicit 8-component vector of statistical descriptors to control the internal weights of the network through the FiLM mechanism. This allows us to dynamically change the behaviour of the filters: for slices with high entropy and noise variance, the model strengthens regularization, while for high-quality images it focuses on preserving fine details. The results show that this approach provides a stable PSNR level in the range of 35–45 dB and high structural integrity (SSIM up to 0.98) even in the most complex signal degradation scenarios. Thus, the developed system does not simply suppress noise, but adapts the reconstruction process itself to the physical parameters of the formation of each specific MRI image.

4. DISCUSSION

This paper presents an approach to MRI image denoising that combines statistically guided feature modulation, a multi-level U-Net architecture, attention mechanisms, and physically based noise modelling in k-space. The results obtained allow us to draw several fundamental conclusions regarding the model's effectiveness, physical correctness, and clinical applicability.

First, quantitative results on synthetically noisy data demonstrate that the proposed approach provides not only average noise reduction but also a sustainable improvement in reconstruction quality for each individual image. Significant increases in PSNR and SSIM, stable distributions of improvements, and preservation of EPI indicators confirm that the model does not function as a universal smoothing filter, but as a structure-oriented reconstruction system capable of simultaneously suppressing noise and preserving the anatomical organization of the signal. This fundamentally distinguishes it from classical denoising methods, where noise reduction is often achieved at the cost of losing high-frequency components and degrading contour information.

It is especially important that the positive effects of denoising are preserved on real clinical data without a reference. Analysis of SNR and CV shows a physically correct stabilization of the signal: an increase in the signal/noise ratio is combined with a moderate decrease in intratissue variability without a transition to unphysical homogeneity. This indicates that the model does not create an artificially “flat” signal, does not destroy the natural texture of tissues and does not form artifacts of excessive smoothing, which are often observed in aggressive filtering methods. Thus, denoising occurs in an adaptive reconstruction mode, and not global filtering.

Here, the most important role is played by the mechanism of statistically controlled FiLM modulation. Unlike classical neural network approaches, where noise is implicitly considered as part of the input image, in the proposed model noise is described as a separate statistical object with its own multidimensional characteristics. Transferring this information into the internal representations of the network allows for the formation of a context-sensitive denoising strategy, where the filtering parameters are adapted to the specific noise profile of each image. In fact, the model implements not a single universal denoising function, but a family of adaptive mappings controlled by the statistical state of the input signal.

The combination of FiLM with attention modules (CBAM, Attention Gate) and multiscale processing forms a multilevel feature selection system. Attention mechanisms enhance structurally significant signal components and suppress local noise fluctuations, while multiscale processing ensures the coordination of local and global context. This allows the model to simultaneously work with small-scale details and macroscopic tissue organization, which is critical for MRI images, where anatomical information is distributed over different spatial scales. Of particular importance are

the results of clinically oriented assessment through contouring. Unlike formal quality metrics, segmentation of anatomical structures directly reflects the practical utility of the image in a real clinical environment. The improvement of DSC and HD95 indicators for the hippocampus, brainstem and optic chiasm indicates that denoising not only improves visual quality, but also increases the reproducibility of anatomical boundaries, reduces contour uncertainty and stabilizes the segmentation process. The most pronounced effect for small and morphologically complex structures emphasizes that the benefit of denoising is maximum precisely where the impact of noise has the greatest clinical consequences.

At the same time, it is important to emphasize that contouring, by its nature, involves a subjective component tied to the human factor. That is why, in this work, it is considered not an absolute metric for validation, but an additional clinically oriented quality indicator. However, it is this subjectivity that makes such an analysis clinically relevant, since it reflects the real convenience of the doctor's work with the image, the stability of interpretation and confidence in reproducing the boundaries of structures.

The importance of physically based noise modeling in k-space should be highlighted separately. The use of a spectral noise model enabled us to avoid non-physical artifacts characteristic of approaches that add noise directly in the image domain and to generate training data that are statistically closer to the real conditions of MRI reconstruction. This significantly reduced the gap between the training and clinical domains and improved the model's generalizability.

Taken together, the results of the study indicate that the proposed approach forms a new paradigm for MRI denoising, in which noise is considered not as a random impurity, but as a statistically described process integrated into the reconstruction model. This approach allows us to move from universal filtering strategies to adaptive, context-sensitive reconstruction models, which is fundamentally important for medical imaging with its high demands on structural accuracy and clinical interpretability.

CONCLUSIONS

In this paper, a comprehensive method for denoising MRI brain images is proposed that combines modern deep learning architectures with physically based noise modeling in k-space and a multi-level evaluation system. The developed FiLM-modulated Residual Attention U-Net model demonstrated a stable and significant improvement in both technical and clinically oriented image quality indicators compared to traditional denoising methods (NLM, BM3D) and common deep architectures (standard U-Net, Attention U-Net).

The key advantage of the approach is the explicit adaptation to the individual noise conditions of each image by integrating an 8-dimensional vector of statistical noise descriptors through the FiLM-modulation mechanism. This provides adaptive, frame-by-frame processing of slices, in contrast to models that implicitly generalize different noise regimes within a single latent space. Combined with residual blocks, CBAM, Attention Gate, and multiscale feature processing, the model effectively suppresses noise, preserving fine anatomical details, contours, and tissue structural organization critical for diagnostic interpretation.

The use of a synthetic noise model in k-space enabled us to create a training set that more closely matched real-world scanning conditions, reducing the gap between training and clinical data and increasing the model's generalizability and robustness when applied to data from different scanners and protocols.

Quantitative evaluation on synthetically noisy data showed a stable PSNR gain of approximately 20.7 dB and an increase in SSIM of approximately 0.73, with PSNR values stabilizing in the range of 35–45 dB after denoising. Analysis of the edge-preservation metric confirmed effective suppression of noise-induced pseudo-contours, with no signs of over-smoothing. On real clinical MRI images, the method demonstrated a systematic increase in SNR and a controlled decrease in the coefficient of variation, indicating physically correct signal stabilization without loss of natural textural heterogeneity of tissues. Clinically oriented validation by contouring anatomically complex structures (hippocampus, brainstem, visual chiasm) showed a stable improvement in segmentation quality: an increase in Dice Similarity Coefficient by 0.07–0.12 and a decrease in HD95 by 30–50%. This indicates increased reproducibility of structural boundaries, reduced contour uncertainty, and improved consistency of results, confirming the practical clinical value of denoised images in neuroimaging and radiotherapy planning tasks.

Overall, the proposed approach can be considered an adaptive, physically based MRI preprocessing tool that shifts the denoising process from universal filtering to statistically driven reconstruction. This is especially important in clinical scenarios where high structural accuracy, contour stability, and preservation of anatomically significant details are critical. Further research can focus on extending the model to multi-channel and multi-contrast MRI, integrating it with k-space reconstruction methods, and implementing it in clinical decision support systems for radiotherapy and neurodiagnostic.

Ethics Statement

The MRI images used in this study were obtained as part of standard clinical examinations according to routine diagnostic protocols. All data were anonymized prior to analysis, eliminating the possibility of patient identification during the study. The study was performed in accordance with generally accepted ethical principles for conducting scientific research using human medical data and in compliance with current institutional and national requirements for the use of retrospective clinical images for scientific purposes.

Data Availability

The MRI dataset used in this study contains clinical medical images and therefore cannot be made publicly available due to confidentiality requirements and institutional restrictions. Access to the data may be provided by the authors upon reasonable request and subject to approval by the relevant medical institution and compliance with applicable ethical and legal regulations.

Acknowledgements
The authors express their sincere gratitude to the radiologists of the Department of Radiology, National Cancer Institute, for their assistance in the contouring of anatomical structures used in the clinically oriented evaluation of the proposed method for denoising MRI images. Their professional experience and participation were essential for the clinical validation of the obtained results.

ORCID
D.G. Sliusarenko, https://orcid.org/0009-0009-1802-5859, **L.V. Sayan**, https://orcid.org/0009-0009-3797-9148;
A.V. Netreba, https://orcid.org/0000-0003-1347-3854

REFERENCES

[1] F. Lugauer, and J. Wetzl, "Magnetic Resonance Imaging," in: *Medical Imaging Systems: An Introductory Guide*, edited by A. Maier, S. Steidl, V. Christlein, and J. Hornegger, (Springer, Cham (CH), 2018). https://doi.org/10.1007/978-3-319-96520-8_6
[2] C.M. Sonderer, and N. Chen, "Improving the Accuracy, Quality, and Signal-To-Noise Ratio of MRI Parametric Mapping Using Rician Bias Correction and Parametric-Contrast-Matched Principal Component Analysis (PCM-PCA)," Yale J Biol Med **91**(3), 207–214 (2018). **PMCID**: PMC6153621
[3] W.A. Edelstein, G.H. Glover, C.J. Hardy, and R.W. Redington, "The intrinsic signal-to-noise ratio in NMR imaging," Magn Reson Med **3**(4), 604–618 (1986). https://doi.org/10.1002/mrm.1910030413
[4] H. Gudbjartsson, and S. Patz, "The Rician distribution of noisy MRI data," Magn. Reson. Med. **34**(6), 910–914 (1995). https://doi.org/10.1002/mrm.1910340618
[5] V. Prabhakaran, V.A. Nair, B.P. Austin, C. La, T.A. Gallagher, Y. Wu, D.G. McLaren, G. Xu, P. Turski, and H. Rowley, "Current Status and Future Perspectives of Magnetic Resonance High-Field Imaging: A Summary," Neuroimaging Clin. N. Am. **22**(2), 373–xii (2012). https://doi.org/10.1016/j.nic.2012.02.012
[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," IEEE Transactions on Image Processing, **16**(8), 2080–2095 (2007). https://doi.org/10.1109/TIP.2007.901238
[7] M. Elad, B. Kawar, and G. Vaksman, "Image Denoising: The Deep Learning Revolution and Beyond—A Survey Paper," SIAM Journal on Imaging Sciences," **16**(13), 1594-1654 (2023). https://doi.org/10.1137/23M1545859
[8] S.W. Zamir, A. Arora, S. Khan, M. Hayat, F.S. Khan, M.-H. Yang, and L. Shao, "Learning Enriched Features for Real Image Restoration and Enhancement," in: *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV*, (Springer-Verlag, Berlin, Heidelberg, 2020), pp. 492–511. https://doi.org/10.1007/978-3-030-58595-2_30
[9] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2Noise: Learning image restoration without clean data," in: *35th International Conference on Machine Learning, ICML 2018*, (International Machine Learning Society, 2018), pp. 4620–4631. https://doi.org/10.48550/arXiv.1803.04189
[10] E. Perez, F. Strub, H. de Vries, V. Dumoulin, and A. Courville, "FiLM: visual reasoning with a general conditioning layer," in: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence*, (AAAI Press, New Orleans, Louisiana, USA, 2018), pp. 3942–3951. https://doi.org/10.48550/arXiv.1709.07871
[11] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning Where to Look for the Pancreas," (2022). https://doi.org/10.48550/arXiv.1804.03999
[12] X. Chen, X. Wang, K. Zhang, K.-M. Fung, T.C. Thai, K. Moore, R.S. Mannel, H. Liu, B. Zheng, and Y. Qiu, "Recent advances and clinical applications of deep learning in medical image analysis," Medical Image Analysis **79**, 102444 (2022). https://doi.org/10.1016/j.media.2022.102444
[13] R. Ayde, M. Vornehm, Y. Zhao, F. Knoll, E.X. Wu, and M. Sarracanie, "MRI at low field: A review of software solutions for improving SNR," NMR Biomed, **38**(1), e5268 (2025). https://doi.org/10.1002/nbm.5268
[14] M. Safari, S. Wang, Z. Eidex, R. Qiu, C.-W. Chang, D.S. Yu, and X. Yang, "A Physics-Informed Deep Learning Model for MRI Brain Motion Correction," ArXiv, arXiv:2502.09296v1 (2025). https://doi.org/10.1002/mp.70197
[15] L. Han, S. Xiao, M. Li, J. Liu, and X. Zhou, "Noise-Controllable Complex-Valued Diffusion Model for k-Space Data of Hyperpolarized ^{129}Xe Lung MRI Generation," in: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2025*, edited by J.C. Gee, D.C. Alexander, J. Hong, J.E. Iglesias, C.H. Sudre, A. Venkataraman, P. Golland, J.H. Kim, and J. Park, (Springer Nature Switzerland, Cham, 2026), pp. 360–369. https://doi.org/10.1007/978-3-032-05127-1_35
[16] C. Ciulla, "Two-dimensional image noise removal and reconstruction using discrete Fourier transform, k-space filtering and Z-space filtering," Progress in Engineering Science, **2**(1), 100056 (2025). https://doi.org/10.1016/j.pes.2025.100056
[17] T. Song, F. Nie, Y. Guo, F. Xu, and S. Zhang, "FilterDiff: Noise-Free Frequency-Domain Diffusion Models for Accelerated MRI Reconstruction," in: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2025*, edited by J.C. Gee, D.C. Alexander, J. Hong, J.E. Iglesias, C.H. Sudre, A. Venkataraman, P. Golland, J.H. Kim, and J. Park, (Springer Nature Switzerland, Cham, 2026), pp. 205–215. https://doi.org/10.1007/978-3-032-05325-1_20
[18] W. Zhang, Y. Guo, and Q. Jin, "Radiomics and Its Feature Selection: A Review," Symmetry, **15**(10), 1834 (2023). https://doi.org/10.3390/sym15101834
[19] M. Shakeri, A. Mostaar, A.Z. Sadeghi, S.M. Hosseini, A.Y. Joybari, and H. Ghadiri, "A Comprehensive Evaluation of Radiomic Features in Normal Brain Magnetic Resonance Imaging: Investigating Robustness and Region Variations," J. Med. Phys. **49**(4), 608–622 (2024). https://doi.org/10.4103/jmp.jmp_149_24
[20] M. Dohmen, M.A. Klemens, I.M. Baltruschat, T. Truong, and M. Lenga, "Similarity and quality metrics for MR image-to-image translation," Sci. Rep. **15**(1), 3853 (2025). https://doi.org/10.1038/s41598-025-87358-0
[21] "The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping," Radiology," **295**(2), (2020). https://doi.org/10.1148/radiol.2020191145
[22] D. Sliusarenko, and A. Netreba, "Phantom-Guided Adaptive Denoising of Medical Images Using Enhanced U-Net Architecture," Radioelectronics and Communications Systems, (2025). https://doi.org/10.3103/S0735272725020049
[23] M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," CoRR, (2015). https://doi.org/10.48550/arXiv.1511.05440

[24] P. Isola, J.-Y. Zhu, T. Zhou, and A.A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), pp. 5967–5976. https://doi.org/10.1109/CVPR.2017.632
[25] D. Sliusarenko, A. Netreba, and S. Radchenko, "Refining Clarity Medical Image via U-Net with Tailored Architectural Modifications and Transfer Learning," in: *Proc. IEEE Int. Conf. Electron. Nanotechnol., ELNANO.*, (Institute of Electrical and Electronics Engineers, 2024), pp. 456–460. https://doi.org/10.1109/ELNANO63394.2024.10756858
[26] H. Liu, Z. Li, S. Lin, and L. Cheng, "A Residual UNet Denoising Network Based on Multi-Scale Feature Extraction and Attention-Guided Filter," Sensors (Basel), **23**(16), 7044 (2023). https://doi.org/10.3390/s23167044
[27] H. Zhang, Q. Lian, J. Zhao, Y. Wang, Y. Yang, and S. Feng, "RatUNet: residual U-Net based on attention mechanism for image denoising," PeerJ Comput. Sci. **8**, e970 (2022). https://doi.org/10.7717/peerj-cs.970
[28] Z. Zhuang, M. Liu, A. Cutkosky, and F. Orabona, "Understanding AdamW through Proximal Methods and Scale-Freeness," Trans. Mach. Learn. Res. (2022). https://doi.org/10.48550/arXiv.2202.00089
[29] S. Brody, U. Alon, and E. Yahav, "On the Expressivity Role of LayerNorm in Transformers' Attention," (arXiv, 2023). https://doi.org/10.48550/arXiv.2305.02582
[30] T.-T. Han, H. Nguyen Van, and P. Nguyen Huu, "Denoising Method for MRI Images Using Modified BM3D Filter with Complex Network and Artificial Neural Networks," Journal of Electrical and Computer Engineering, **2024**(1), 2606485 (2024). https://doi.org/10.1155/2024/2606485
[31] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, **13**(4), 600–612 (2004). https://doi.org/10.1109/TIP.2003.819861
[32] M.J. Kobra, M.O. Rahman, and A.M. Nakib, "A Novel Hybrid Framework for Noise Estimation in High-Texture Images using Markov, MLE, and CNN Approaches," Scientific Journal of Engineering Research, **1**(2), 54–63 (2025). https://doi.org/10.64539/sjer.v1i2.2025.25
[33] H.M.S.S. Herath, H.M.K.K.M.B. Herath, N. Madusanka, and B.-I. Lee, "A Systematic Review of Medical Image Quality Assessment," Journal of Imaging, **11**(4), 100 (2025). https://doi.org/10.3390/jimaging11040100
[34] R. Rodrigues, L. Lévêque, J. Gutiérrez, H. Jebbari, M. Outtas, L. Zhang, A. Chetouani, S. Al-Juboori, M.G. Martini, and A.M.G. Pinheiro, "Objective quality assessment of medical images and videos: review and challenges," Multimed Tools Appl **84**(25), 29915–29948 (2025). https://doi.org/10.1007/s11042-024-20292-x
[35] K.M. Kempski, M.T. Graham, M.R. Gubbi, T. Palmer, and M.A. Lediju Bell, "Application of the generalized contrast-to-noise ratio to assess photoacoustic image quality," Biomed. Opt. Express, **11**(7), 3684–3698 (2020). https://doi.org/10.1364/BOE.391026
[36] O. Rainio, and R. Klén, "Modified Dice Coefficients for Evaluation of Tumor Segmentation from PET Images: A Proof-of-Concept Study," J. Digit. Imaging. Inform. Med. (2025). https://doi.org/10.1007/s10278-025-01535-1
[37] P. Ramachandran, T. Eswarlal, M. Lehman, and Z. Colbert, "Assessment of Optimizers and their Performance in Autosegmenting Lung Tumors," J. Med. Phys. **48**(2), 129–135 (2023). https://doi.org/10.4103/jmp.jmp_54_23
[38] A. Mason, J. Rioux, S.E. Clarke, A. Costa, M. Schmidt, V. Keough, T. Huynh, and S. Beyea, "Comparison of Objective Image Quality Metrics to Expert Radiologists' Scoring of Diagnostic Quality of MR Images," IEEE Transactions on Medical Imaging, **39**(4), 1064–1072 (2020). https://doi.org/10.1109/TMI.2019.2930338
[39] K. Chupetlovska, K. Groot Lipman, Z. Bodalal, F.M. Aricò, L. Topff, M. Maas, and S. Trebeschi, "Do's and don'ts of tumor segmentation with 3D slicer: A practical guide for radiologists, by radiologists," European Journal of Radiology Artificial Intelligence, **5**, 100053 (2026). https://doi.org/10.1016/j.ejrai.2025.100053
[40] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, (2005), pp. 60–65 vol. 2. https://doi.org/10.1109/CVPR.2005.38
[41] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," IEEE Transactions on Image Processing, **26**(7), 3142–3155 (2017). https://doi.org/10.1109/TIP.2017.2662206
[42] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M.K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low Dose CT Image Denoising Using a Generative Adversarial Network with Wasserstein Distance and Perceptual Loss," IEEE Trans Med Imaging, **37**(6), 1348–1357 (2018). https://doi.org/10.1109/TMI.2018.2827462

СТАТИСТИЧНО ОБУМОВЛЕНЕ ШУМОЗАГЛУШЕННЯ МРТ ЗА ДОПОМОГОЮ АРХІТЕКТУРИ RESIDUAL ATTENTION U-NET З FILM-МОДУЛЯЦІЄЮ

Д.Г. Слюсаренко[1,2], Л.В. Саян[2], А.В. Нетреба[1]
[1]Факультет радіофізики, електроніки та комп'ютерних систем, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64/13, Київ, Україна, 01601
[2]Національний інститут раку України, вул. Юлія Здановського, 33/43, Київ, Україна, 03022

Якість зображень МРТ часто обмежується просторово неоднорідним шумом, що негативно впливає на точність клінічної інтерпретації та автоматичного аналізу. Традиційні методи глибокого навчання часто враховують шум неявно, що призводить до надмірного згладжування та втрати дрібних анатомічних структур. У цій статті ми пропонуємо архітектуру покращеного шумозаглушення U-Net, яка використовує механізм лінійної модуляції за ознаками (FiLM) для динамічної адаптації до профілю шуму кожного зрізу. Модель поєднує вектор з 8 статистичних дескрипторів (включаючи характеристики інтенсивності, текстури та частоти), що дозволяє динамічно керувати внутрішніми представленнями мережі залежно від конкретних умов сканування. Для підвищення фізичної коректності навчання проводилося на даних із синтетично згенерованим шумом у k-просторі. Архітектура вдосконалена за допомогою блоків залишків, механізмів уваги та багатомасштабного модуля обробки. На синтетичних даних середнє покращення пікового співвідношення сигнал/шум (PSNR) становило ≈ 20,7 дБ, а середнє покращення індексу структурної подібності (SSIM) – приблизно 0,73, що свідчить про глибоке відновлення структурної інформації. На клінічних зображеннях спостерігалося збільшення співвідношення сигнал/шум (SNR) та стабілізація коефіцієнта варіації (CV), що підтверджує фізичну коректність методу. Клінічна валідація на складних контурних структурах (гіпокамп, стовбур мозку, зоровий хіазм) показала збільшення коефіцієнта Дайса (DSC) на 0,07–0,12 та зменшення похибки HD95 на 30–50%. Запропонований метод дозволяє перейти від універсальних стратегій шумозаглушення до адаптивної реконструкції, забезпечуючи високу точність збереження анатомічних меж. Це робить його перспективним інструментом для обробки МРТ у задачах нейровізуалізації та планування варіативної терапії.

Ключові слова: *МРТ; Attention U-Net; FiLM; медична візуалізація; реконструкція; шум; CNN; контурування*