# Texture Preserving Deep Learning-based Noise Reduction for Anatomical Magnetic Resonance Images and Its Impact on Imaging Features

ELEFTHERIOS TRIVIZAKIS
Computational BioMedicine Laboratory (CBML)
Foundation for Research and Technology, Hellas, Heraklion, GREECE
ORCID: 0000-0002-0316-5252
trivizakis@ics.forth.gr

EUGENIA MYLONA
Biomedical Research Institute (BRI)
Foundation for Research and Technology, Hellas, Ioannina GREECE
ORCID: 0000-0002-1275-6249

KATERINA NIKIFORAKI
Computational BioMedicine Laboratory (CBML)
Foundation for Research and Technology, Hellas, Heraklion GREECE
ORCID: 0000-0003-0792-9800

DIMITRIOS I. ZARIDIS
Biomedical Engineering Laboratory, School of Medicine, School of Electrical & Computer Engineering, National Technical University of Athens, Athens, GREECE
ORCID: 0000-0002-7362-5082

MANOLIS TSIKNAKIS
Computational BioMedicine Laboratory/ Dpt. Electrical & Computer Engineering
Foundation for Research and Technology, Hellas / Hellenic Mediterranean University, Heraklion
GREECE
ORCID: 0000-0001-8454-1450

DIMITRIOS I. FOTIADIS
Biomedical Research Institute (BRI)/ Unit of Medical Technology and Intelligent Information Systems,
Foundation for Research and Technology, Hellas, Ioannina/ University of Ioannina
GREECE
ORCID: 0000-0002-7362-5082

NIKOLAOS TACHOS
Biomedical Research Institute (BRI)/ Unit of Medical Technology and Intelligent Information Systems,
Foundation for Research and Technology, Hellas, Ioannina/ University of Ioannina
GREECE
ORCID: 0000-0002-8627-6352

DANIELE REGGE
Department of Radiology, Candiolo Cancer Institute, FPO-IRCCS, Candiolo, Turin,
ITALY
ORCID: 0000-0001-8267-5279

NIKOLAOS PAPANIKOLAOU
Computational Clinical Imaging Group, Centre of the Unknown, Champalimaud Foundation,
PORTUGAL
ORCID: 0000-0002-1028-1016

PROCANCER-I CONSORTIUM
www.procancer-i.eu
complete list of members in Supplementary Information

KOSTAS MARIAS
Computational BioMedicine Laboratory/ Dpt. Electrical & Computer Engineering
Foundation for Research and Technology, Hellas / Hellenic Mediterranean University, Heraklion
GREECE

*Abstract:* - Anatomical magnetic resonance images are affected by different types of noise, including thermal, motion, radio interference, and magnetic field inhomogeneities. In a clinical setting, acquiring MR images of the highest quality is not always feasible. Quantitative and artificial intelligence-based decision support tools require high-quality data to accurately differentiate among pathological conditions, avoiding diminishing the

clinical relevance of a diagnostic model. A fully convolutional model with no pooling layers was trained on a set of noisy images, with the ground truth being the original image without the noise. Different levels of noise were incorporated into the training set. The experiments showed a reduction in noise levels, but it can impact quantification tasks when T2ws without noise are provided to the model. Six types of pairs of original T2w image slices and the corresponding slices with synthetic noise were generated with various thresholds of Gaussian noise, spanning from 4% to 14%. In total, 38500 pairs were utilized for convergence and evaluation of the proposed denoising models. The examined deep denoiser reduced improved image quality by up to 18.2% peak signal-to-noise ratio (PSNR), overall across the aforementioned noise thresholds.

*Key-Words:* - Deep learning; Denoiser; Magnetic resonance imaging; Radiomics; Prostate;

Received: November 26, 2024. Revised: December 10, 2024. Accepted: December 10, 2024. Published: December 21, 2024

# 1 Introduction

### 1.1 Background for Denoising

Noise reduction, also referred to in image processing as denoising, is a key preprocessing step in image analysis, which is frequently a necessity for accurate quantitative and qualitative analyses. Ideally, the noisy information should be minimized without compromising the underlying texture component during denoising.

High-resolution magnetic resonance (MR) examinations with a high signal-to-noise ratio (SNR) allow imaging with more detailed anatomical features, boosting the diagnostic power of a model and aiding in the accurate detection of illnesses. In real-life clinical scenarios, often the SNR can be reduced due to the nature of MR imaging, protocols for faster image acquisition, and other environmental variables [1]. There are various other ways of achieving higher quality with high SNR in MR scans, including increasing the number of acquired slices, utilizing a strong magnetic field, and adjusting the acquisition bandwidth. However, some of those methods for increasing the scan quality also increase the acquisition time, thereby introducing significant costs in a clinical setting. Therefore, several denoising methods can be employed to further enhance the quality of MR scans in examinations with low SNR. In particular, deep learning (DL) techniques for noise reduction have gained popularity in recent studies [2-4] because of their texture reconstruction and edge-preserving properties.

Endorectal coil (ERC) MR imaging is used in prostate examinations because of its distinct advantage it has over surface array coils in terms of image quality, free from artifacts in the area of interest. This technique provides high-resolution scans, which are necessary for optimal visibility of the prostate anatomy and other surrounding regions of interest, as well as detecting cancerous tissue. However, using ERC has some key drawbacks, such as cost and a painful or inconvenient application. This may result in patients refusing future MRI examinations, which could lead to a late diagnosis or inefficient monitoring of the condition. Deep learning denoising can address this by enhancing surface prostate MRI, rendering unnecessary the need for an intrusive ERC and, therefore, changing the way prostate cancer scans are currently acquired.

A key advantage of using deep learning-based models over traditional image processing denoising methods such as average or median filtering is that the parameters of the deep model are optimized during training on the examined images, while with traditional denoising, a predefined algorithm is used. Unlike traditional methods, deep learning denoising preserves the edges and granular details in texture. This is not the case with traditional methods, which overly smooth the original signal, introducing blurry textures into the image. Despite the popularity of DL in many image processing and analysis applications, only a handful of studies have been published regarding denoising in medical imaging. Employing these types of commercial or open-source DL methods in computed tomography (CT) for several anatomical areas, including abdominal scans for renal cancer [5], liver [6], lung [7], and pelvis [8], led to significant image quality enhancements. Kidoh et al. [9] successfully applied three types of DL denoising architectures for brain MRI. The fully-convolutional architectures advanced denoising by featuring novel layers such as soft shrinkage, in which a custom activation function is trained to adapt to the noise levels, and discrete cosine transform layers, where the input examinations are processed in the frequency

domain. Wang et al. [10] implemented a diverse set of denoising models with prostate MR data from different acquisition protocols. Siemens Deep Resolve [11] employs a comparable methodology wherein a deep learning model is utilized to reduce the acquisition time of an MRI. In particular, the author found that the DL-based method for reconstructing non-ERC images performed the best out of the models that were tested. This could indicate that a simple protocol with DL denoising could be used in the clinic.

### 1.2 Study's Contribution

In this paper, a custom fully convolutional deep learning-based architecture was tested in the context of denoising for T2w magnetic resonance images of the prostate. To the best of our knowledge, this is the first MRI DL denoiser that makes use of a structural similarity-based loss function, which is proposed as an alternative to the widely used peak signal-to-noise ratio-based loss functions. Additionally, the impact of such a preprocessing task on radiomics was also examined.

## 2 Problem Formulation

The T2-weighted images of the ProstateX dataset [12] were used to train and evaluate the DL denoising models. The scans were produced using two Siemens 3T MRI scanners, the MAGNETOM Skyra and Trio. T2-weighted images with a resolution of roughly 0.5 mm in plane and a slice thickness of 3.6 mm were obtained using a turbo spin echo procedure. There are two patient cohorts available: a) 203 patients with their clinical data, gland and lesion annotations; and b) 143 patients with only the multi-parametric MRI available. In the context of this task, all 346 scans were used since only the imaging data without annotations was required for the convergence of the denoising models. The publicly accessible PI-CAI dataset [13] was used as an unseen external validation set to quantify the impact of DL denoising on the stability of radiomics. The overlapping MRI examinations of the ProstateX and PI-CAI were removed prior to the radiomic analysis.

## 3 Problem Solution

### 3.1 MR Images Stratification for Deep Learning Denoising

The examined dataset of 326 patients was split into three different sets on a patient basis. The training set consisted of 276 patients, and it was used for fitting the deep learning models. A validation set of 25 patients was used for tuning the parameters of the deep learning models, early stopping, and assessing the status of overfitting during training. Finally, an unseen testing set of 25 patient scans was utilized for the model evaluation protocol, providing a fair and robust assessment of the denoising effectiveness. The models were trained and evaluated on a slice-basis, yielding approximately: a) 5500 unique slices for the training set, b) 450 slices for the validation set, and c) 470 slices for the testing set.

### 3.2 Image Preprocessing

A protocol for selecting the highest quality of MRI examinations was established to ensure that the best slices are used for model convergence and evaluation. An experienced radiophysicist evaluated all 346 T2-weighted scans of the publicly accessible ProstateX dataset. As a result, 20 scans (approximately 6% of the dataset) were rejected due to severe noise, motion, and other types of artifacts. Additionally, to mitigate the variation in spacing across the MRI examinations, aspect ratio-preserving reshaping with zero-padding and interpolation was applied to the original scans. This resulted in a pixel array of 384 by 384 pixels across all the slices. Prior to the analysis, the pixel intensities of the MRI slices were normalized. A Gaussian noise pattern was assumed for generating the synthetic noisy slices. Six noisy images for each real slice were generated with noise thresholds spanning from 4% to 14%. Therefore, approximately 38,500 noisy slices were used for convergence and evaluation of the examined deep denoising models. A few noisy slices are depicted in Fig. 1.

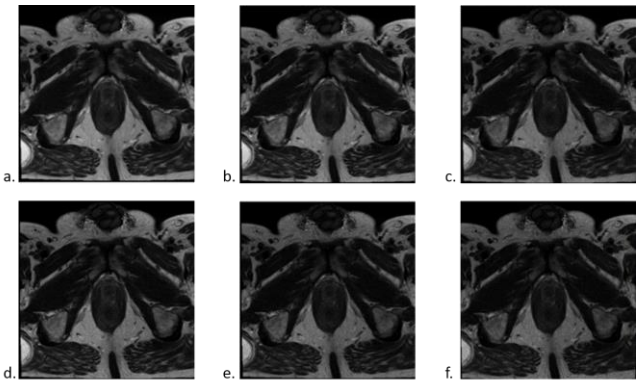


**Fig. 1** An original (a) prostate T2w slice with different levels of noise (4-12%, b-f) applied.

### 3.3 Image Augmentation

In the context of DL analysis, this step is essential for increasing the number of samples that are used during model fitting and also to minimize overfitting of the deep models. Aside from increasing the training sample count, data augmentation results in translation, perspective invariance, and artificially introduced variety in the examined dataset, which strengthens the generalization power of the deep models. Four types of transformations were performed: 1) pixel flipping from right to left, 2) pixel flipping from top to bottom, 3) 90-degree image rotation, and 4) 270-degree image rotation. The final training was comprised of approximately 132,000 slices. A sample of augmented data is presented in Fig. 2.

### 3.4 Deep Learning Architecture for Denoising

Four (4) fully-convolutional architectures were examined: a) a convolutional autoencoder with residual connections (CrAE), b) a denoising convolutional network (DnCNN [14]), c) a denoising convolutional network with residual connections (DrCNN), and d) a real image denoising network (RIDNet [15]). The fully-convolutional models were trained with a supervised learning strategy employing pairs of images; the high-quality ground truth image and the slice with synthetic noise. In most studies, mean squared error (MSE) is used as a loss function, despite the fact that this type of metric does not capture the statistical distribution of texture in an image. During hyperparameter optimization, the structural similarity index measure [16] (SSIM) was identified as a better method to formulate the cost function of the denoising task. The adaptive moment estimation (ADAM) was to minimize the proposed SDI loss between the ground truth and the noisy slice. An L1 penalty was applied to the kernels of each layer,

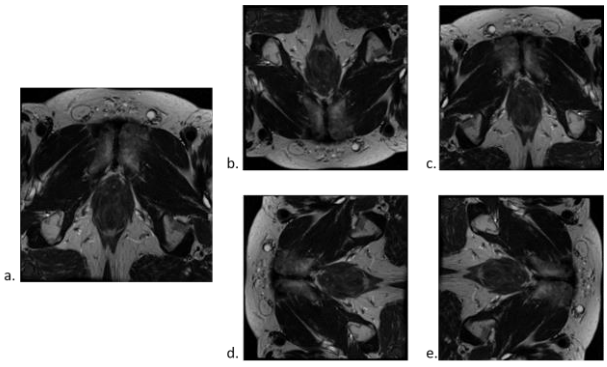


**Fig. 2** Data augmentation applied to a slice of the training cohort. This includes flipping the original image (a) from top to bottom (b) and right to left (c), rotating 270$^{o}$ (d) and 90$^{o}$ (e).

constraining the trainable weights of the model from taking outlier values and consequently preventing the model from learning noisy representations that can lead to overfitting. Residual connections were incorporated into the model's architecture to prevent the vanishing gradients [17] effect of the very deep convolutional networks. The integration of the soft-shrinkage activation function [18] was a key integration in the proposed fully convolutional architecture because it allowed the network to learn representations that were proportional to the noise power levels of the examined dataset.

### 3.5 Hyperparameter Optimization

This process was very important to the success of the denoising task because hyperparameters are the least reported information in the published studies, and their value is dependent on the dataset that is used. Using how well the model performed, the optimization was done on the validation set to find the best model parameters. These parameters are comprised of learnable elements of the architecture (number of modules, kernels, and neurons) as well as other fundamental factors such as the learning rate, optimizer, activation functions, kernel initializers, and regularization penalties. To minimize model overtraining, obtain the most optimal model, and prevent redundant training iterations, early-stopping was implemented with a threshold of 20 epochs after minimizing the validation loss function. Furthermore, comparing the learning curves for loss can reveal information about the fitting state of the deep model. Therefore, to assess the denoising performance and model generalization ability, the learning curves were examined by juxtaposing the minimum distance between the training and validation loss curves.

### 3.6 Imaging Feature Extraction

Radiomic features were extracted from the whole gland on T2w MRIs using the Pyradiomics open-source Python package [19]. These radiomic features consist of shape, first-order statistics, Gray Level Cooccurence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTDM), and Gray Level Dependence Matrix (GLDM) features. Apart from shape features, other texture features were also computed after applying wavelet and Laplacian transforms of Gaussian (LoG) transformations to the images, leading to a total of 1140 radiomic features per patient.

The distribution of these imaging features in the "excellent", "good", "modest" and "poor" stability

category, for each preprocessing pipeline is proposed.

### 3.7 Evaluation of Model's Performance

The evaluation of the proposed methodology was conducted exclusively on the unseen testing set as: a) qualitative assessment by experienced radiophysicists, and b) quantitative evaluation using the juxtaposition of noisy versus denoised images with metrics such as SSIM (Eq. 1), and PSNR (Eq. 2).

SSIM encapsulates three key factors for comparing the aforementioned pair of slices: a) luminance (captures the pixel distortions for brighter regions, Eq. 3), b) contrast (captures the pixel distortions of regions with high diversity, Eq. 4), and c) structure (a sliding window calculates the statistical local dependencies of texture regions, Eq. 5).

Therefore, the proposed loss integrates these three factors, and it is formulated as an index that captures the structural differences (SDI, Eq. 6) between two images.

$$SSIM(x,y) = l(x,y)^{\alpha} \cdot c(x,y)^{\beta} \cdot s(x,y)^{\gamma} \quad (1)$$

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

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (3)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (4)$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \quad (5)$$

$$SDI = 1 - SSIM(x,y) \quad (6)$$

## 4 Results

### 4.1 Convergence of Deep Models

The DL denoiser was trained on a workstation equipped with a thirty-two threaded AMD processor, two hundred and fifty four gigabytes of RAM, and a twenty-four gigabyte NVidia RTX graphics card. The T2w slices were split on a patient basis into the training set for model convergence, the validation set for early-stopping, and the unseen testing set for evaluating the model.

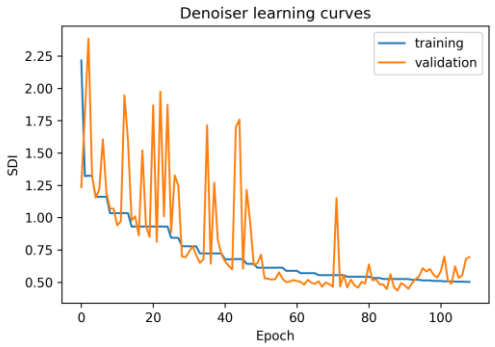


**Fig. 3** Learning curves of loss during the training phase.

The comparison of the training and validation loss, also known as learning curves, facilitated a thorough assessment of the degree of model convergence. In Fig. 3, the SDI validation loss was minimized at the 89th epoch, while early stopping kicked in at the 109th epoch as the deep model starts to overfit. The architecture that achieved the best convergence was the custom DnCNN with residual connections.

### 4.2 Texture Quality Evaluation – Structural Similarity

The internal validation of the denoising model was performed on the unseen testing set based on the ProstateX dataset. This enabled the identification of the best performing deep architecture out of ones that were tested for denoising in terms of image quality improvements. In particular, the denoised image was better up to a 20% in terms of SSIM compared to the noisy image, when both juxtaposed to the ground truth image. The quantitative analysis of the image quality improvements for different noise thresholds is depicted in Fig. 4.

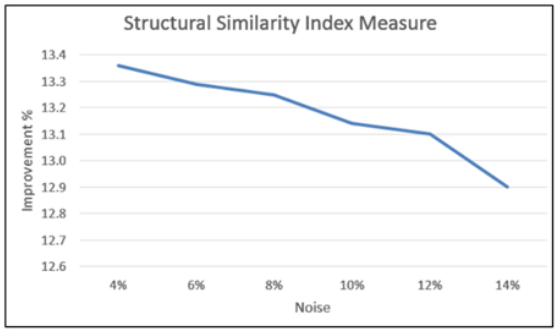


**Fig. 4** The improvement in image quality of the denoised, in terms of SSIM, versus the noisy image in different noise thresholds

### 4.3 Image Quality Evaluation – Signal-to-Noise Ratio

In conjunction with SSIM, PSNR was also calculated with the same methodology. The delta between the PSNR of the noisy scans and the

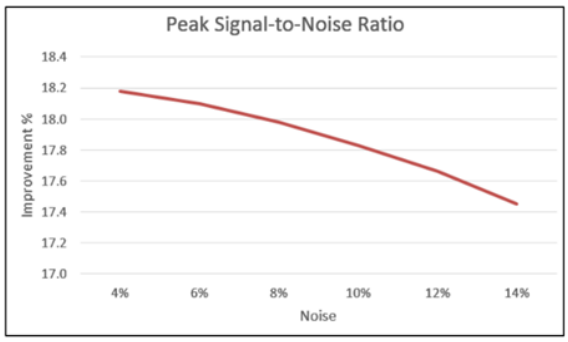


**Fig. 5** The improvement in image quality of the denoised, in terms of PSNR, versus the noisy image in different noise thresholds

denoised scans shows significant improvements of up to 22%. The mean PSNR improvements range from 17.5 to 18.2%, depending on the level of noise, as shown in Fig. 5.

### 4.4 Radiomic Stability Evaluation

An external validation approach was employed to quantify the impact of denoising on a texture level. The DL denoising model was tested on the unseen PI-CAI dataset to evaluate potential undesirable effects of the DL model, radiomic stability, and quantify the texture differences. In particular, no artifacts or other distortions were observed after denoising the T2w scans from the PI-CAI dataset. Since synthetic noise was not introduced in the dataset, a SSIM-based metric that estimates the differences between denoised and the original examinations was calculated. A difference mean of 2±4% spanning from 0.4 to 35%. This is to be expected because in such a dataset that has been curated for image analysis challenges. In particular, only thirty examinations were found to be visibly noisy by the experienced radiophysicist, while the SSIM difference calculated by the denoiser was more than 10%.

In Table I, the intraclass correlation coefficient (ICC) of various radiomic feature families is presented. The most heavily affected by the denoising is the 19.4% of first order (n=46) and the 19.2% of GLCM (n=60) features. The majority of

**Table I** Prostate gland radiomic feature stability, in terms of ICC, after applying the proposed DL denoising

| Feature Type | ≥ 90 | ≥ 75 | ≥ 50 | <50 |
|---|---|---|---|---|
| **First order (n=234)** | 84 | 81 | 23 | 46 |
| **GLCM (n=312)** | 128 | 81 | 59 | 60 |
| **GLDM (n=182)** | 71 | 64 | 37 | 10 |
| **GLRLM (n=208)** | 96 | 69 | 38 | 5 |
| **GLSZM (n=208)** | 68 | 89 | 36 | 15 |
| **Shape (n=14)** | 14 | 0 | 0 | 0 |

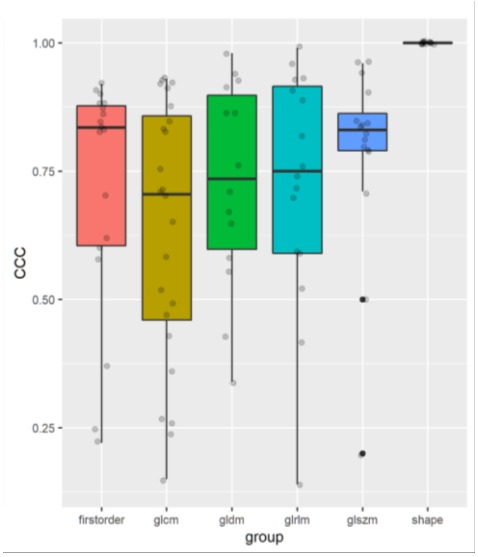


**Fig. 6** Boxplots of concordance correlation coefficient (CCC) for the different families of radiomic features, calculated between the original and the preprocessed images

the radiomics are highly stable, while shape features are completely unaffected by DL denoising. Additional context of the impact of denoising on radiomics is provided by the boxplots of Fig.6.

### 4.5 Assessing Denoising Qualitatively

An experienced radiophysicist review*ed* extensively the denoised T2w prostate slices from the internal (ProstateX) and external (PI-CAI) validation set. The bright regions of a T2w image were highly improved by the DL denoiser in terms of noisy pattern reduction and pixel intensity levels. In particular, it was observed that examinations with little textural enhancements depicted luminance and contrast levels similar to the original slice. This is crucial to clinicians and deep models for easier

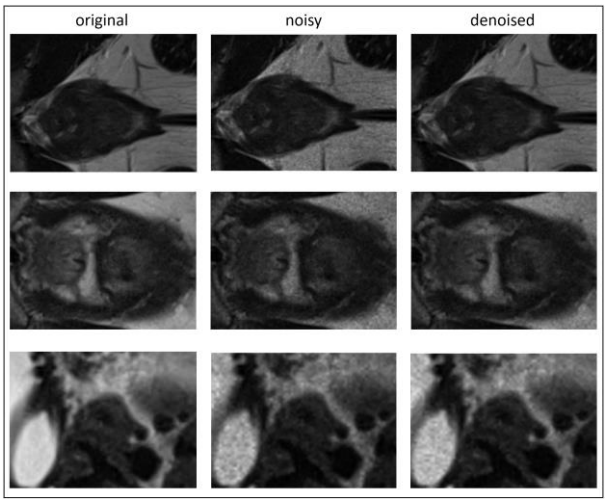


**Fig. 7** Comparison of the ground truth, synthetically noisy, and denoised slices.

differentiation of various tissue types of a prostate T2w examination. The results of DL denoising from three patients are presented in Fig 7.

# 5 Discussion

Denoising is a necessary step in datasets with noisy imaging data for alleviating potential risks of model overfitting on noise or reduction in robustness and reproducibility of AI-based decision support tools.

In our experiments, deep convolutional models with residual connections and no pooling layers produced denoising models that managed to reduce noise with no loss in texture complexity or image blurring, as shown in Fig. 7.

A modified version of the DrCNN performed best in terms of PSNR and SSIM on the unseen testing set. Images with high-noise thresholds were more challenging to restore, as presented in Figs 4 and 5.

Overall, the findings suggest that denoised scans from a deep model have higher image quality than images processed by traditional image processing techniques. According to the experienced radiophysicist, DL denoising delivered, in many cases, substantial noise reduction with no visible blurring or loss of image quality. Additionally, the edges were preserved, and in many cases, enhanced, the original texture distribution was partially restored, and pixel intensities were closer to the values of the original scan. The qualitative analysis on the external validation set (PI-CAI) resulted in similar findings for thirty noisy images.

Additionally, after denoising, the majority of radiomics feature types (67-100%, Table I) show relatively good stability (ICC $\geq$ 75) on the unseen external validation set. The boxplot in Fig. 6 shows that GLCM prostate imaging features were highly affected by DL denoising, although 67% of these features showed high stability. Shape features show resilience in denosing, which had no effect on them.

A limiting factor of this study is the relatively small patient cohort that the deep model was built upon. The Fig. 3 shows that there might be some room for better convergence (minimum loss was 0.43) with the additional T2w sequences in the training set. Moreover, the imaging data should be evaluated for high-quality and noise-free by a group of experts prior to model development, which requires a lot of resources, which is a major obstacle to further diversifying the dataset.

Nevertheless, the qualitative analysis showed that noise reduction should be treated with caution and not applied agnostically when developing artificial intelligence or radiomic-based models since they seem to have a greater impact on feature stability and repeatability.

In future iterations of the proposed deep denoiser a diverse set of data from multiple centers will be incorporated into the training and validation sets to enrich the data representation in the deep denoiser model.

Furthermore, varying types of noise could be approximated and incorporated in the convergence phase of the analysis. Therefore, this advancement in prostate MRI denoising, as suggested by the experienced radiophysicist, might be employed in retrospective sets of data for enhancing older MR sequences by the proposed DL model.

# 6 Conclusion

The proposed deep learning denoiser is a robust and viable option for enhancing prostate T2w examinations without the drawbacks of traditional noise reduction solutions. Radiomics can be negatively affected when a DL denoising model is applied agnostically across noisy and curated datasets.

## Author's contributions

E.T. conceived, designed the study, drafted the manuscript, and performed the analysis. E.M, D.I.Z. and K.N. provided assistance with data collection and curation. E.T., E.M., D.I.Z. and K.N performed the radiomic analysis. E.T., E.M., K.N., D.I.Z., N.T., M.T., D.I.F., D.R., N.P. and K.M. contributed to the literature research, interpretation of data and revised the manuscript. D.I.F., and N.P. contributed to the critical revision of the paper. K.M. contributed to the critical revision of the paper and was the guarantor of the integrity of the entire study. All authors have read and agreed to the published version of the manuscript.

### Data availability

The datasets analyzed during the current study are publicly accessible in the National Cancer Institute (NIH), Cancer Imaging Program (CIP) repository, under the title PROSTATEx (https://www.cancerimagingarchive.net/collection/prostatex/). Additionally, PI-CAI dataset is hosted by the Zenodo platform (https://zenodo.org/record/6624726).


## Funding

This work was supported by funding from the European Commission (EU Horizon 2020: ProCAncer-I project; grant number 952159). Additionally, this work was also supported by the

European Union's Horizon Europe digital simple grants under grant agreement no.101100633-EUCAIM.

## References

[1] Dimitriadis, A., Trivizakis, E., Papanikolaou, N., Tsiknakis, M. and Marias, K., 2022. Enhancing cancer differentiation with synthetic MRI examinations via generative models: a systematic review. Insights into Imaging, 13(1), p.188.

[2] Nakamura, Y., Higaki, T., Tatsugami, F., Honda, Y., Narita, K., Akagi, M. and Awai, K., 2020. Possibility of deep learning in medical imaging focusing improvement of computed tomography image quality. Journal of computer assisted tomography, 44(2), pp.161-167.

[3] Manjón, J.V. and Coupe, P., 2018. MRI denoising using deep learning. In Patch-Based Techniques in Medical Imaging: 4th International Workshop, Patch-MI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4 (pp. 12-19). Springer International Publishing.

[4] Koch, K.M., Sherafati, M., Arpinar, V.E., Bhave, S., Ausman, R., Nencka, A.S., Lebel, R.M., McKinnon, G., Kaushik, S.S., Vierck, D. and Stetz, M.R., 2021. Analysis and evaluation of a deep learning reconstruction approach with denoising for orthopedic MRI. Radiology: Artificial Intelligence, 3(6), p.e200278.

[5] Akagi, M., Nakamura, Y., Higaki, T., Narita, K., Honda, Y., Zhou, J., Yu, Z., Akino, N. and Awai, K., 2019. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *European radiology*, *29*, pp.6163-6171.

[6] Hur, B.Y., Lee, J.M., Joo, I., Yu, M.H., Yoon, J.H., Han, J.K. and Choi, B.I., 2014. Liver computed tomography with low tube voltage and model-based iterative reconstruction algorithm for hepatic vessel evaluation in living liver donor candidates. *Journal of computer assisted tomography*, *38*(3), pp.367-375.

[7] Kakinuma, R., Moriyama, N., Muramatsu, Y., Gomi, S., Suzuki, M., Nagasawa, H., Kusumoto, M., Aso, T., Muramatsu, Y., Tsuchida, T. and Tsuta, K., 2015. Ultra-high-resolution computed tomography of the lung: image quality of a prototype scanner. *PloS one*, *10*(9), p.e0137165.

[8] Tian, S.F., Liu, A.L., Liu, J.H., Liu, Y.J. and Pan, J.D., 2019. Potential value of the PixelShine deep learning algorithm for increasing quality of 70 kVp+ ASiR-V reconstruction pelvic arterial phase CT images. *Japanese journal of radiology*, *37*, pp.186-190.

[9] Kidoh, M., Shinoda, K., Kitajima, M., Isogawa, K., Nambu, M., Uetani, H., Morita, K., Nakaura, T., Tateishi, M., Yamashita, Y. and Yamashita, Y., 2020. Deep learning based noise reduction for brain MR imaging: tests on phantoms and healthy volunteers. *Magnetic resonance in medical sciences*, *19*(3), pp.195-206.

[10] Wang, X., Ma, J., Bhosale, P., Ibarra Rovira, J.J., Qayyum, A., Sun, J., Bayram, E. and Szklaruk, J., 2021. Novel deep learning-based noise reduction technique for prostate magnetic resonance imaging. *Abdominal Radiology*, *46*, pp.3378-3386.

[11] Behl, N., 2021. Deep resolve—mobilizing the power of networks. *MAGNETOM Flash (78)*, *1*, pp.29-35.

[12] Litjens, G., Debats, O., Barentsz, J., Karssemeijer, N., & Huisman, H. (2017). SPIE-AAPM PROSTATEx Challenge Data (Version 2) [dataset]. The Cancer Imaging Archive. https://doi.org/10.7937/K9TCIA.2017.MURS5CL

[13] Saha, A., Twilt, J.J., Bosma, J.S., van Ginneken, B., Yakar, D., Elschot, M., Veltman, J., Fütterer, J., de Rooij, M. and Huisman, H., 2022. The PI-CAI challenge: public training and development dataset. *Zenodo, Jun*.

[14] Zhang, K., Zuo, W., Chen, Y., Meng, D. and Zhang, L., 2017. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, *26*(7), pp.3142-3155.

[15] Anwar, S. and Barnes, N., 2019. Real image denoising with feature attention. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 3155-3164).

[16] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, *13*(4), pp.600-612.

[17] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[18] Isogawa, K., Ida, T., Shiodera, T. and Takeguchi, T., 2017. Deep shrinkage convolutional neural network for adaptive noise

reduction. *IEEE Signal Processing Letters*, *25*(2), pp.224-228.

[19] *Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 2017;77:e104–e107. doi: 10.1158/0008-5472.CAN-17-0339.*

## Appendix A. Supplementary information: ProCAncer-I members

• Manolis Tsiknakis (FORTH - Institute of Computer Science - Computational BioMedicine Lab, Greece)

• Kostas Marias (FORTH - Institute of Computer Science - Computational BioMedicine Lab, Greece)

• Stelios Sfakianakis (FORTH - Institute of Computer Science - Computational BioMedicine Lab, Greece)

• Varvara Kalokyri (FORTH - Institute of Computer Science - Computational BioMedicine Lab, Greece)

• Eleftherios Trivizakis (FORTH - Institute of Computer Science - Computational BioMedicine Lab, Greece)

• Grigorios Kalliatakis (FORTH - Institute of Computer Science - Computational BioMedicine Lab, Greece)

• Avtantil Dimitriadis (FORTH - Institute of Computer Science - Computational BioMedicine Lab, Greece)

• Dimitris Fotiadis (FORTH - Institute of Molecular Biology and Biotechnology (FORTH-IMBB/BR), Greece)

• Nikolaos Tachos (FORTH - Institute of Molecular Biology and Biotechnology (FORTH-IMBB/BR), Greece)

• Eugenia Mylona (FORTH - Institute of Molecular Biology and Biotechnology (FORTH-IMBB/BR), Greece)

• Dimitris Zaridis (FORTH - Institute of Molecular Biology and Biotechnology (FORTH-IMBB/BR), Greece)

• Charalampos Kalantzopoulos (FORTH - Institute of Molecular Biology and Biotechnology (FORTH-IMBB/BR), Greece)

• Nikolaos Papanikolaou (Champalimaud Foundation, Portugal)

• José Guilherme de Almeida (Champalimaud Foundation, Portugal)

• Ana Castro Verde (Champalimaud Foundation, Portugal)

• Ana Carolina Rodrigues (Champalimaud Foundation, Portugal)

• Nuno Rodrigues (Champalimaud Foundation, Portugal)

• Miguel Chambel (Champalimaud Foundation, Portugal)

• Henkjan Huisman (Radboud, Netherlands)

• Maarten de Rooij (Radboud, Netherlands)

• Anindo Saha (Radboud, Netherlands)

• Jasper J. Twilt (Radboud, Netherlands)

• Jurgen Futterer (Radboud, Netherlands)

• Luis Martí-Bonmatí(HULAFE - Biomedical Imaging Research Group, Instituto de Investigación Sanitaria La Fe; Medical Imaging Department, Hospital Universitari i Politècnic La Fe, Spain)

• Leonor Cerdá-Alberich (HULAFE - Biomedical Imaging Research Group. Instituto de Investigación Sanitaria La Fe, Spain)

• Gloria Ribas (HULAFE - Biomedical Imaging Research Group. Instituto de Investigación Sanitaria La Fe, Spain)

• Silvia Navarro (HULAFE - Biomedical Imaging Research Group. Instituto de Investigación Sanitaria La Fe, Spain)

• Manuel Marfil (HULAFE - Biomedical Imaging Research Group. Instituto de Investigación Sanitaria La Fe, Spain)

• Emanuele Neri (Academic Radiology, Department of Translational Research, University of Pisa, Italy)

• Giacomo Aringhieri (Academic Radiology, Department of Translational Research, University of Pisa, Italy)

• Lorenzo Tumminello (Academic Radiology, Department of Translational Research, University of Pisa, Italy)

• Vincenzo Mendola (Academic Radiology, Department of Translational Research, University of Pisa, Italy)

• Deniz Akata (Hacettepe - Department of Radiology, Turkey)

• Mustafa Özmen (Hacettepe - Department of Radiology, Turkey)

• Ali Devrim Karaosmanoglu (Hacettepe - Department of Radiology, Turkey)

• Firat Atak (Hacettepe - Department of Radiology, Turkey)

• Musturay Karcaaltincaba (Hacettepe - Department of Radiology, Turkey)

• Joan C. Vilanova (Institute of Biomedical Research of Girona Dr. Josep Trueta (IDIBGI), Department of Radiology (IDI), Girona, Spain)

• Jurgita Usinskiene (National cancer institute, Vilnius, Lithuania)

• Ruta Briediene (National cancer institute, Vilnius, Lithuania)

• Audrius Untanas (National cancer institute, Vilnius, Lithuania)

• Kristina Slidevska (National cancer institute, Vilnius, Lithuania)

• Katsaros Vasilis (General Anti-Cancer and Oncological Hospital of Athens, Greece)

• Georgiou Georgios (General Anti-Cancer and Oncological Hospital of Athens, Greece)

• Dow-Mu Koh (Radiology & AI Research Hub, The Royal Marsden NHS Foundation Trust, London; Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, UK, UK)

• Robby Emsley (Radiology & AI Research Hub, The Royal Marsden NHS Foundation Trust, London; Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, UK, UK)

• Sharon Vit (Radiology & AI Research Hub, The Royal Marsden NHS Foundation Trust, London; Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, UK, UK)

• Ana Ribeiro (Radiology & AI Research Hub, The Royal Marsden NHS Foundation Trust, London; Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, UK, UK)

• Simon Doran (Radiology & AI Research Hub, The Royal Marsden NHS Foundation Trust, London; Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, UK, UK)

• Tiaan Jacobs (Radiology & AI Research Hub, The Royal Marsden NHS Foundation Trust, London; Division of Radiotherapy and Imaging, The Institute of Cancer Research, London, UK, UK)

• Gracián García-Martí(Quirónsalud Hospital/CIBERSAM, Valencia, Spain)

• Daniele Regge (Candiolo Cancer Institute, FPO-IRCCS, Str. Prov.le 142 km 3.95, 10060 Candiolo, Turin, Italy)

• Valentina Giannini (Candiolo Cancer Institute, FPO-

IRCCS, Str. Prov.le 142 km 3.95, 10060 Candiolo, Turin, Italy)
• Simone Mazzetti (Candiolo Cancer Institute, FPO-IRCCS, Str. Prov.le 142 km 3.95, 10060 Candiolo, Turin, Italy)
• Giovanni Cappello (Candiolo Cancer Institute, FPO-IRCCS, Str. Prov.le 142 km 3.95, 10060 Candiolo, Turin, Italy)
• Giovanni Maimone (Candiolo Cancer Institute, FPO-IRCCS, Str. Prov.le 142 km 3.95, 10060 Candiolo, Turin, Italy)
• Valentina Napolitano (Candiolo Cancer Institute, FPO-IRCCS, Str. Prov.le 142 km 3.95, 10060 Candiolo, Turin, Italy)
• Sara Colantonio (Institute of Information Science and Technologies of the National Reserch Council of Italy, Italy)
• Maria Antonietta Pascali (Institute of Information Science and Technologies of the National Reserch Council of Italy, Italy)
• Eva Pachetti (Institute of Information Science and Technologies of the National Reserch Council of Italy, Italy)
• Giulio del Corso (Institute of Information Science and Technologies of the National Reserch Council of Italy, Italy)
• Danila Germanese (Institute of Information Science and Technologies of the National Reserch Council of Italy, Italy)
• Andrea Berti (Institute of Information Science and Technologies of the National Reserch Council of Italy, Italy)
• Gianluca Carloni (Institute of Information Science and Technologies of the National Reserch Council of Italy, Italy)
• Jayashree Kalpathy-Cramer (Mass General Hospital, Boston MA, USA)
• Christopher Bridge (Mass General Hospital, Boston MA, USA)
• Joao Correia (B3D, UK)
• Walter Hernandez (B3D, UK)
• Zoi Giavri (Advantis, Greece)
• Christos Pollalis (Advantis, Greece)
• Dimitrios Agraniotis (Advantis, Greece)
• Ana Jiménez Pastor (Quibim, S.L., Valencia, Spain)
• Jose Munuera Mora (Quibim, S.L., Valencia, Spain)
• Clara Saillant (Univie, Austria)
• Theresa Henne (Univie, Austria)
• Rodessa Marquez (Univie, Austria)

## Creative Commons Attribution License 4.0 (Attribution 4.0 International , CC BY 4.0)