Deep learning algorithms to improve bone scintigraphy

Akos Kovacs

Theses of the PhD Dissertation

Supervisors:
Dr. András Horváth,
Dr. Tamás Bükki

Roska Tamas Doctoral School of Sciences and Technology
Pázmány Péter Catholic University

Budapest, 2023

1. Introduction

Medical imaging devices are extremely important in today's patient-centered care based on scientific results. These appliances have completely changed the way medicine is practised over the last 30 years. They have made it possible to identify diseases at an early stage (especially in screening tests) and, thanks to early treatment, patients' chances of recovery have improved. The use of X-ray, ultrasound, CT, MRI and other imaging technologies has become commonplace in the clinic and thus widely known in society. Medical imaging is particularly beneficial in detecting and identifying cancer because early detection means better chances of cure. One of the tools routinely used for this purpose is the gamma camera, which uses a variety of radioactive isotopes to obtain information about physiological processes in the body.

I have been working on the improvement of one of the most commonly performed examinations, bone scintigraphy, a common, relatively inexpensive and widely available technique, which is invaluable in the diagnostic evaluation of many pathological conditions due to its sensitivity. In Hungary, an average of 5-600MBq of MDP isotope activity is administered to the patient and a 15-20 minute scan with the device is required. Gamma camera images typically have a low signal-to-noise ratio and are subject to significant Poisson noise. These circumstances motivated me to develop an image enhancement device during my doctoral studies that would

allow reducing the injected activity (radioactive dose) and shortening the imaging time.

Exceptionally high signal-to-noise ratios can be achieved by using noise filtering neural networks (NNs), which can be configured and trained to act as specialised noise filters. An important characteristic of NN-based noise filters is that they learn on noisy input data and also on reference or ground-truth image pairs with significant noise. It has been shown that, nevertheless, a properly constructed NN filter trained in this way can synthesize filtered images with a better signal-to-noise ratio than that of reference images, and outperforms conventional noise filters, e.g. BM3D method, for either Gaussian or Poisson noise [1]. Considering that the image database used for filter training usually consists of a highly limited number of images due to the difficult availability of real patient data, we consider it particularly important to investigate the robustness of the trained NN-based image processing algorithm, i.e., its sensitivity to the noise content of the images and their distribution according to various aspects, including patient age, gender, body mass index value, and the nature and distribution of characteristic pathological structures in the image. This analysis can reveal the robustness of such an image processing algorithm, either on its own or as part of a larger CADx (computer aided diagnosis) system deployed in clinics around the world.

A common indication for bone scintigraphy is to detect and track bone metastases of various tumours, so as a further step, we started to develop a software to search for pathological enrichment. This tool will also allow quantification of the impact of the image enhancement tool on diagnostics, so that an application-specific lesion-based evaluation can be performed in addition to image-based metrics. One method to create a software component for abnormal enrichment detection and prediction is the use of convolutional neural networks. Related to this, we developed a new segmentation metric, Wave Loss. I would like to show that topological

information incorporated in the loss function can be used to increase the accuracy of segmentation networks.

In light of the above, I seek to answer the following research questions:

- Can we use deep learning for high-quality, reliable noise filtering in planar bone scintigraphy?
- How robust is such a solution in real life?
- How should such a tool be evaluated where we do not have noise-free, perfect images as a basis for comparison?
- Is it possible to construct a loss function for neural network training that can take into account the topography of segmentations instead of just pixel-level comparisons?

2. New scientific contributions and thesispoints

Thesis I a: *I have developed a robust high-quality noise filtering method tailored for planar bone scintigraphy, showcasing the effectiveness of deep learning in this context. Through validation on a real patient database, isolated for accuracy assessment, the top-performing neural network achieved a mean RMSE of 1.15 under normal statistics, outperforming the best non-neural network solution, BM3D, which attained a mean error of 1.29. Furthermore, when tested under 1/3 statistics, the neural network yielded an average RMSE of 1.38, surpassing the best non-neural network-based solution, Gaussian 9mm, which attained an RMSE of 2.07. These findings demonstrate the superior noise filtering capabilities of neural networks compared to established non-neural network methods used in clinical practice.* Corresponding publication: [2]

In this thesis a neural network based noise filter is proposed that can be used with planar bone scintigraphy recordings at multiple noise levels, instead of developing a separate network for each noise level. The proposed denoising solution is a convolutional neural network (CNN) inspired by U-NET architecture. The noise-filtering network was trained using bone scintigraphy recordings with real statistics according to the standard protocol, without noise-free recordings.

The selection of data suitable for training the neural networks was done with the Q-Bot software. [3] For the development, 2430 anonymized recordings (from 1215 patients, anterior and posterior) were used, acquired by AnyScan® DUO/TRIO SPECT/CT (Mediso Ltd.) and InterView^{TM} processing SW (Mediso Ltd.). All patients were given 5-600MBq Tc-99m methylene diphosphonate (MDP) (Isotope Institute LTD, Budapest, Hungary) intravenously with 2-5 hour accumulation time. The matrix size was 256*1024 with 130mm/min scanning speed.

All recordings were resampled based on a binomial distribution, artificially generating realistically degraded recordings as if they had been taken at one-third, one-quarter, one-eighth, etc. recording times.

If we subtract this degraded image from the original measurement, we get an independent record with better statistics than the data on the input side. This is also a very powerful augmentation technique, whereby we change not only the input side, but also the output side.[4]

The task of the neural network is to estimate the transformation between the two generated recordings. We used Mean Absolute Error (MAE) as a function of the learning loss between the actual filtered and reference image, as was done in the [5] article. We used the commonly applied methodology: the backpropagation algorithm and gradient based optimization to train the neural networks [6].

The resulting solution proved to be robust to the noise level of the images within the examined limits. During the evaluation, the performance of the networks was compared to Gaussian and median filters and to the Block-matching and 3D filtering (BM3D) filter. It has been shown that particularly high signal-to-noise ratios can be achieved using noise-filtering neural networks (NNs), which are more robust than the traditional methods and can help diagnosis, especially for images with high noise content.

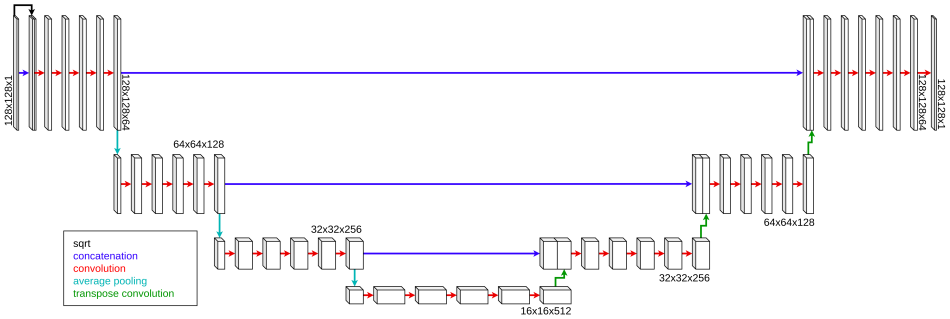


Figure 2.1: The network architectures used for noise filtering were all U-NET based. The differences between the architectures were in the number of convolutional layers per level, and the number of convolutional filters. The network, named *L-NN*, contained 4 levels and always had 5 convolutional layers following each other. The filter numbers of the convolutional blocks used at each level were 64, 128, 256 and 512. The neural network named *S-NN* also contained 4 levels, always with 3 convolutional layers following each other and the number of filters used at each level were 16, 32, 64 and 128.

From Table 2.1 showing the results by RMSE metric, it can be seen that for all statistics, the neural network based solutions achieved the best results. Note that under normal and 1/3 statistics, at this metric, the performance of the BM3D and Gaussian filters is comparable to the neural network, but with worse statistics, the performance of these solutions degrades to unusable levels.

Thesis I b: *My deep learning-based noise filtering solution for planar bone scintigraphy demonstrates robustness in real-life applications. Through a comprehensive investigation across various homogeneous and biased validation datasets (including diverse age groups, BMI ranges, and gender categories), I assessed the performance variability of the denoising algorithm. The evaluation revealed notable consistency and effectiveness, showcasing that for datasets with normal statistics, the average RMSE error ranged from 1.05 to 1.28 across different subsets, with the mixed dataset registering an average error of 1.15. Moreover, the standard*

	Statistics									
	normal		1/3		1/9		1/16		1/32	
RMSE:	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
BM3D	**1.29**	0.36	**2.07**	0.33	4.50	0.36	7.34	0.46	13.76	0.72
Gaussian 11mm	1.99	0.93	2.21	0.91	**2.76**	0.83	3.27	0.82	4.21	0.81
Gaussian 13mm	2.34	1.21	2.48	1.19	2.85	1.12	**3.23**	1.10	**3.94**	1.05
Gaussian 3mm	2.81	0.33	4.85	0.57	8.38	0.97	11.16	1.31	15.80	1.85
Gaussian 5mm	1.69	0.31	2.66	0.37	4.44	0.53	5.87	0.70	8.26	0.97
Gaussian 7mm	1.56	0.47	2.15	0.45	3.33	0.48	4.30	0.57	5.97	0.73
Gaussian 9mm	1.71	0.68	2.07	0.65	2.87	0.60	3.58	0.63	4.82	0.69
L-NN	**1.15**	0.40	**1.38**	0.41	**1.80**	0.47	**2.09**	0.54	**2.54**	0.63
Median 3px	1.79	0.47	2.48	0.45	4.36	0.42	6.81	0.37	11.70	0.82
Median 5px	2.64	0.91	3.06	0.87	4.50	0.72	6.71	0.55	11.51	0.82
Median 7px	3.57	1.30	3.91	1.26	5.15	1.08	7.17	0.83	11.82	0.93
Median 9px	4.57	1.67	4.90	1.63	6.01	1.45	7.85	1.17	12.32	1.13
S-NN	1.21	0.35	1.56	0.38	2.09	0.48	2.45	0.56	3.00	0.67

Table 2.1: Performance of different filters calculated by RMSE. From the table, it can be seen that for all statistics, the neural network based solutions achieved the best results (smallest Mean and SD). Note that under normal and 1/3 statistics, at this metric, the performance of the BM3D and Gaussian filters is comparable to the neural network, but with worse statistics, the performance of these solutions degrades to unusable levels.

deviation within each subset ranged from 0.4 to 0.48, highlighting the stability and reliability of the filtering algorithm. When evaluated under 1/3 statistics, the mean error exhibited a similar range, varying from 1.28 to 1.54, with standard deviations ranging from 0.41 to 0.49. These findings underscore the adaptability and consistent performance of our deep learning-based filter across diverse patient groups, reinforcing its robustness and applicability in real-world scenarios of planar bone scintigraphy imaging. Corresponding publication: [2]

The measurement results on which the thesis claims are based are given in Table 2.2. The trends in performance measured on the different sets as a function of the deterioration of the statistics are the same as those observed on the mixed set.

Thesis I c: *I have created an effective evaluation method for the deep learning-driven noise filtering tool in planar bone scintigraphy without the need for noise-free*

		Statistics									
		normal		1/3		1/9		1/16		1/32	
	RMSE:	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
L-NN	Age High	1.28	0.48	1.54	0.49	2.00	0.54	2.32	0.58	2.82	0.67
	Age Low	1.05	0.41	1.28	0.41	1.68	0.47	1.96	0.54	2.41	0.64
	BMI High	1.09	0.43	1.31	0.44	1.68	0.47	1.94	0.50	2.33	0.57
	BMI Low	1.28	0.48	1.54	0.48	2.02	0.54	2.35	0.61	2.89	0.72
	Female	1.20	0.43	1.43	0.44	1.83	0.48	2.12	0.53	2.55	0.61
	Male	1.20	0.45	1.46	0.46	1.91	0.54	2.23	0.61	2.71	0.73
	Mixed	1.15	0.40	1.38	0.41	1.80	0.47	2.09	0.54	2.54	0.63
S-NN	Age High	1.35	0.43	1.74	0.46	2.31	0.54	2.70	0.60	3.29	0.71
	Age Low	1.12	0.35	1.47	0.38	1.99	0.46	2.34	0.54	2.88	0.65
	BMI High	1.16	0.39	1.47	0.41	1.93	0.48	2.26	0.53	2.73	0.61
	BMI Low	1.34	0.42	1.76	0.45	2.36	0.55	2.78	0.62	3.42	0.74
	Female	1.27	0.38	1.62	0.41	2.14	0.48	2.50	0.55	3.02	0.64
	Male	1.27	0.41	1.65	0.44	2.21	0.56	2.59	0.63	3.17	0.77
	Mixed	1.21	0.35	1.56	0.38	2.09	0.48	2.45	0.56	3.00	0.67

Table 2.2: Performance of neural network-based filters computed by RMSE on different validation sets. The performance of the larger neural network is better than the smaller neural network for all sets and statistics. The trends in performance on different sets as a function of the degradation of the statistics are the same as those on the mixed set.

images as a reference. In the advanced stages of development, we identified a neural network exhibiting satisfactory performance in processing low noise content measurements. Using this selected neural network, we created a noise-free validation dataset of 544 measurements and then had these images analyzed by physicians to identify any abnormalities, unusual structures, accumulations or artifacts compared to the original images. Within our evaluation framework, these filtered images were considered as virtually noise-free, representing idealized images. We further employed these "noise-free" images, under normal statistical conditions, to generate Poisson noise-affected images, serving as inputs for our solutions. Additionally, leveraging these artificially created standard measurement like images, I have applied additional degradation by employing binomial sampling, thus creating lower-quality representations for comparative assessments. Corresponding publication: [2]

The whole pipeline and the examples of the images produced by the pipeline

are shown in Figure 2.2 and Figure 2.3.

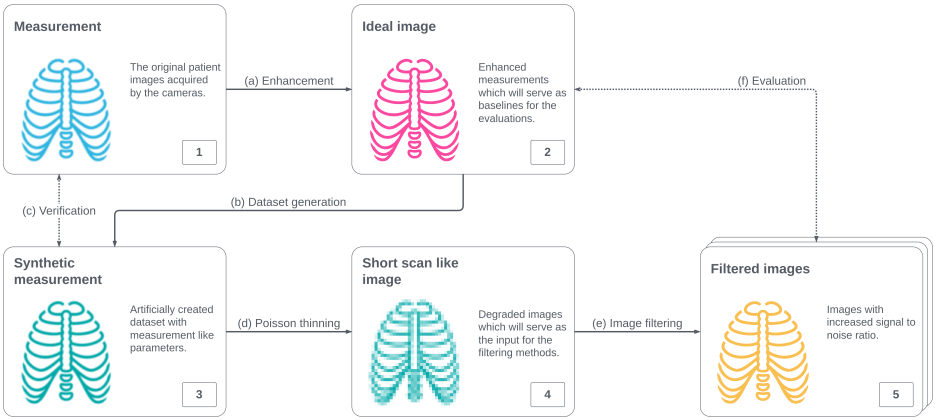
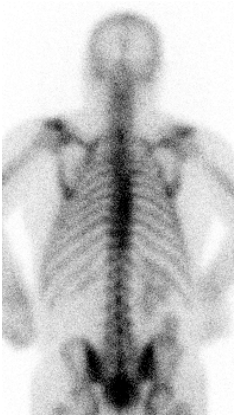
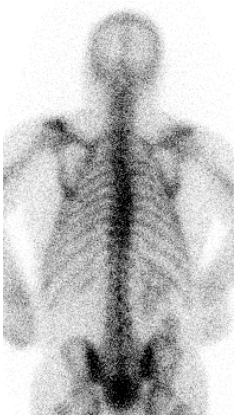


Figure 2.2: Evaluation pipeline: We start from the real measurements acquired by the scanner (1). The second step is to create a noise-free image (2) with a reference enhancement solution (a), which was a neural network based denoiser in our case [4]. The ideal image will be then examined by physicians to see if there was any unusual structure, accumulation or artifact in the image. From this noiseless ideal image we generate synthetic measurement (3) with adding poisson noise (b), which will be verified (c) by statistical tests. The next step is to construct the records with worse statistics (4) using Poisson thinning (d). Finally these images will be the inputs to the various filtering tools (e), which results' (5) will be compared (f) to the ideal images (2).



(a) Measurement

(b) Ideal



(c) Input (1/3 statistics) PSNR: 30.13 dB

(d) Median filter (3px) PSNR: 40.63 dB

(e) Gaussian filter (7mm) PSNR: 41.78 dB

(f) BM3D filter PSNR: 42.67 dB

(g) S-NN PSNR: 45.18 dB

(h) L-NN PSNR: 46.64 dB

Figure 2.3: Evaluation pipeline: We start from Measurement (a), from which we create a noise-free image with a reference filter (b). This will be then reviewed by doctors and taken as a benchmark. From this we generate an artificial degraded noisy image (c). Images (d), (e), (f), (g), (h) show the results of different filters. Since we have the noise-free reference image, we can correctly compute the errors of each method using the metrics.

Thesis II: *I have developed a training loss function for neural networks that considers the topographical structure of segmentations, as opposed to just pixel-level comparisons. The proposed method has increased segmentation accuracy by* 3% *on both the Cityscapes and MS-COCO datasets compared to cross entropy, using various network architectures.* Corresponding publication: [7]

The solution of segmentation problems with deep neural networks requires a well-defined loss function for comparison and network training. In most network training approaches, only area-based differences that are of differing pixel matter are considered; the distribution is not. Our brain can compare complex objects with ease and considers both pixel level and topological differences simultaneously. Comparison between objects requires a properly defined metric that determines similarity between them considering changes both in shape and values. In past years, topographic aspects were incorporated in loss functions where either boundary pixels or the ratio of the areas were employed in difference calculation. During our work we showed how the application of a topographic metric, called wave loss, can be applied in neural network training and increase the accuracy of traditional segmentation algorithms.

Algorithm 1 calculates wave loss for two grayscale images. The input values are $Img1$ and $Img2$ and the output of the algorithm is a scalar variable $WaveLoss$.

Algorithm 1: Calculation of wave loss.

Data: Img1, Img2
Parameters : ValInc, SpaInc, SpaW, ValW
Result: WaveLoss

```
Union ← max(Img1,Img2);
CurrentWave ← min(Img1,Img2);
NewWave ← min(Img1,Img2);
WaveLoss = 0;
i ← 0;
num_iter ← int(1/ValInc);
while i ≤ num_iter do
    /* Loss for intensity differences                    */
    NewWave += ValInc;
    NewWave = min(NewWave,Union);
    ValueChange = sum(NewWave-CurrentWave);
    WaveLoss += ValW[i]*ValueChange;
    CurrentWave = NewWave;
    /* Loss for spatial differences                      */
     NewWave = maxpool(CurrentWave,[SpaInc,SpaInc], [1,1]);
    NewWave = min(NewWave,Union);
    SpatialChange = sum(NewWave - CurrentWave);
    WaveLoss += SpaW[i] * SpatialChange;
    CurrentWave = NewWave;
    i += ValInc;
end
```

The parameters of the algorithm are the following:

- $ValInc$ will determine how fast the wave propagates along the intensity differences; every pixel's intensity will be increased by this amount in every iteration. This parameter will also determine the maximum number of required iterations and by this it will also determine the largest distance from the intersection where topological differences are considered. Having a larger distance than the maximal receptive field of a neuron in the network is illogical because this way the error could be derived back to a neuron which had no vote in the classification of that input pixel. In our experiments, this value was between 0.05 and 0.1, meaning that the wave from a selected point could propagate for 20 and 10 pixels.

- $SpaInc$ will determine the spatial propagation speed of the wave. Spatial propagation is implemented by a max pooling operation with window size $SpaInc$ and a stride of one. In our simulations, this value was always set to 3.

- $ValW$ is a vector of penalties for the intensity differences. If this value is constant, the weight differences will be linearly proportional to the penalties in the loss. If this is increasing, it means larger differences (where more iterations are required to reach the desired value) will have larger and larger penalties. In our simulations, we used constant values in $ValW$.

- $SpaW$ is a vector containing the penalties for topographical differences. $SpaW[0]$ will weight those points which can be reached in one spatial propagation and which are in the direct neighborhood of the intersection. $SpaW[k]$ will have a penalty for those values which will be reached at the k-th iteration. In our simulations, we applied linearly increasing values which were all lower

than the values of $ValW$. In most networks, we want to have good results on average, but minor mistakes about the shape of the object can be tolerated. Applying lower values than the intensity weights ($ValW$) means that the importance of the shape of the segmented object will become less important. Monotonically increasing $SpaW$ means that the further we are from the object, the higher the cost a misclassification will result. Applying higher weights than $ValW$, which are monotonically decreasing, would mean that the boundaries are really important and classifying a pixel around a boundary is a larger problem than misclassifying a pixel somewhere far from the object.

Table 2.3: This table contains the average accuracy results of five independent runs on the Cityscapes dataset using four different network architectures (rows) and six different loss functions for semantic segmentation.

Model	**L1 Loss**	**CrossEnt**	**Dice**	**Boundary**	**ShapeAware**	**Wave**
SegNet	54.2%	57.0%	57.3%	57.7	58.6%	59.5%
DeepLab	59.7%	63.1%	64.1%	64.3%	65.4%	66.7%
DeepLabv3	77.6%	81.3%	81.4%	81.5%	81.7%	82.2%
HRNET	77.4%	81.6%	81.8%	81.8%	82.1%	83.4%

Table 2.4: Average precision results on COCO 2017 validation set using the same network architectures with three different loss functions in different columns (ℓ_1, cross entropy, Dice loss, active boundary loss, shape aware loss and wave loss). Two different architectures (ResNet-50 and ResNet-101) can be found in the rows, with feature pyramid networks (FPNs) or when the activation of the fourth convolution layer (C4) was used for region proposals. The results display the mean average precision for all objects, except the last three rows, where the accuracy results for the best performing network are detailed for small-, medium- and large-sized objects as well.

Model	**L1**	**CrossEnt**	**Dice**	**Boundary**	**Shape**	**Wave**
R50-C4 mAP all	28.75%	32.2%	32.83%	32.9%	34.721%	35.93%
R50-FPN mAP all	29.43%	35.2%	36.14%	36.12%	37.53%	38.11%
R101-C4 mAP all	30.17%	36.7%	37.2%	37.4%	38.86%	38.23%
R101-FPN mAP all	31.67%	38.6%	38.8%	39.3%	40.25%	41.7%
R101-FPN mAP s	14.25%	17.37%	18.18%	18.35%	19.33%	22.24%
R101-FPN mAP m	37.53%	39.23%	39.74%	40.52%	41.27%	43.26%
R101-FPN mAP l	50.14%	51.64%	51.83%	52.17%	52.22%	53.27%

The proposed method has increased segmentation accuracy by 3% on both the Cityscapes and MS-COCO datasets compared to cross entropy, using various network architectures.

2.1 Perspectives of the bone scintigraphy noise filter

After the robustness test, a study of clinical pre-testing had been accomplished involving physicians (ScanoMed Ltd., Debrecen, Hungary). The aim of this study was to allow doctors who have worked with many similar images to point out possible defects, artificial products and to give their opinion on the usability of the device. The images of 412 routine bone scintigraphy whole-body examinations at ScanoMed were denoised using the AI-based application presented here. Patients routinely received 550-600 MBq of 99mTc-MDP intravenously, and whole-body images were acquired after 2 hour of accumulation time. Once the planar image was acquired, the filtered image was obtained within 1-2 minutes and helped physicians to decide on additional investigations such that if any image showed a lesion suspicious for metastasis, SPECT/CT was indicated. As we reported in the Reference [4], the doctors looked at the original unmodified image with normal statistics and the noise-filtered version of the image in parallel, and evaluated the images in this way. The physicians found that the neural network based filter did not delete or generate new lesions, and they don't identified artifacts on the pictures. They concluded, that it was easier to localize the abnormalities (count ribs, vertebrae), decide whether additional examinations (SPECT/CT) was needed, and all this accelerated the diagnosis itself.

This experiment suggests that the use of a noise filter is useful for images with normal statistics, but further studies are needed to see how much it is possible to reduce the measurement time or the activity administered preserving the original, reliable diagnostic capability. Therefore we have been working on the complex clinical evaluation of the given denoising algorithm integrated with lesion detection and classification software components in order to optimize the

performance regarding ROC (receiver operation curve) analysis. Our future aim is to ensure clinical diagnostic value regarding sensitivity and specificity even at significantly lower administered activities or measurement time using the presented denoising solution.

3. The Author's journal papers related to the thesis

[4] Á. Kovács, G. Légrádi, A. Wirth, F. Nagy, A. Forgács, S. Barna, I. Garai, and T. Bükki, "A mesterséges és emberi intelligencia értéke a csontszcintigráfia példáján keresztül," Magyar Onkológia, vol. 64, no. 2, pp. 153–158, 2020

[2] A. Kovacs, T. Bukki, G. Legradi, N. J. Meszaros, G. Z. Kovacs, P. Prajczer, I. Tamaga, Z. Seress, G. Kiszler, A. Forgacs, S. Barna, I. Garai, and A. Horvath, "Robustness analysis of denoising neural networks for bone scintigraphy," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 1039, p. 167003, sep 2022.

[7] Á. Kovács, J. Al-Afandi, C. Botos, and A. Horváth, "Wave Loss: A Topographic Metric for Image Segmentation," Mathematics 2022, Vol. 10, Page 1932, vol. 10, p. 1932, jun 2022.

References

[1] 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*, vol. 7, pp. 4620–4631, International Machine Learning Society (IMLS), 2018. 1

[2] A. Kovacs, T. Bukki, G. Legradi, N. J. Meszaros, G. Z. Kovacs, P. Prajczer, I. Tamaga, Z. Seress, G. Kiszler, A. Forgacs, S. Barna, I. Garai, and A. Horvath, "Robustness analysis of denoising neural networks for bone scintigraphy," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 1039, p. 167003, sep 2022. 2, 2, 2, 3

[3] F. Nagy, A. K. Krizsan, K. Kukuts, M. Szolikova, Z. Hascsi, S. Barna, A. Acs, P. Szabo, L. Tron, L. Balkay, M. Dahlbom, M. Zentai, A. Forgacs, and I. Garai, "Q-Bot: automatic DICOM metadata monitoring for the next level of quality management in nuclear medicine," *EJNMMI Physics*, vol. 8, pp. 1–13, dec 2021. 2

[4] Á. Kovács, G. Légrádi, A. Wirth, F. Nagy, A. Forgács, S. Barna, I. Garai, and T. Bükki, "A mesterséges és emberi intelligencia értéke a csontszcintigráfia

példáján keresztül," *Magyar Onkológia*, vol. 64, no. 2, pp. 153–158, 2020. 2, 2.2, 2.1, 3

[5] N. Yuan, J. Zhou, and J. Qi, "Half2Half: deep neural network based CT image denoising without independent reference data," *Physics in Medicine and Biology*, vol. 65, p. 215020, nov 2020. 2

[6] C. C. Aggarwal, *Neural Networks and Deep Learning*. Springer International Publishing, 2018. 2

[7] Á. Kovács, J. Al-Afandi, C. Botos, and A. Horváth, "Wave Loss: A Topographic Metric for Image Segmentation," *Mathematics 2022, Vol. 10, Page 1932*, vol. 10, p. 1932, jun 2022. 2, 3