ISSN 1870-4069

3D Convolutional Neural Network for Positron Emission Tomography Image Enhancement

Leandro José Rodríguez Hernández[1],
Humberto de Jesús Ochoa Domínguez[1],
Vianey Guadalupe Cruz Sánchez[1],
Osslan Osiris Vergara Villegas[1], Juan Humberto Sossa Azuela[2]

[1] Universidad Autónoma de Ciudad Juárez,
Instituto de Ingeniería y Tecnología,
Mexico

[2] Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

{hochoa,vianey.cruz,overgara}@uacj.mx,
hsossa@cic.ipn.mx, al194726@alumnos.uacj.mx,

Abstract. In this work, we propose a 3D convolutional neural network (CNN) for positron emission tomography (PET) image enhancement as an application of the artificial intelligence (AI) in the area of health. Our proposed network manages to increase the number of counts in the PET sinograms, thus, positively influencing the final quality of the reconstructed image. The enhanced sinogram, obtained by the network, is reconstructed using the ordered subset expectation maximization (OSEM) algorithm. The results show that the proposed network is able to increase the PSNR by 6% on average and the contrast almost twice.

Keywords: Positron emission tomography, convolutional neural network, sinogram, image enhancement.

1 Introduction

Industry 4.0 is changing the way medical devices are produced and delivered. Artificial intelligence (AI) has become a key element of this industry and it is knocking down barriers and forcing to assess the way the traditional work is done. In the medical area, the AI trends and solutions are transforming the medical imaging field to improve the diagnosis process [1].

Medical imaging is the set of techniques used to inspect the human body, with the objective of diagnosing, monitoring, or treating medical conditions [2]. Positron emission tomography (PET) is a technique to acquire images representative of the metabolic activity of the body. At the beginning of the PET scan, the patient is injected with a small dose of radioactive material called radiotracer.

For a cancer study, the radiotracer used is fluorodeoxyglucose (F^{18}). This substance decays by a neutrino and a positron (beta+) with a lifetime of about 109 minutes. Each

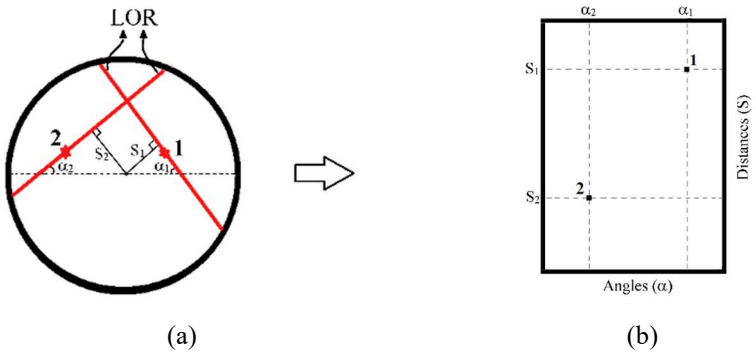


Fig. 1. Representation of (a) two LORs where two and one events have happened respectively and (b) resulting 2D sinogram with the events stored.

positron annihilates with an electron, producing two photons of high energy, traveling in opposite directions. This is known as coincidence or event. Each event is counted upon reaching the scanner's detectors within a time window. The point of annihilation is ideally located on a straight line connecting a pair of detectors.

This line is known as the line of response (LOR) [3]. Reconstruction can be performed using filtered back projection methods or iterative methods. In a PET study, the time is reduced, not all the events are counted as true coincidences, but they can also be random or scattered coincidences. In the last two cases, the coincidence detected by each of the detectors comes from different LORs.

Randoms are one of the main sources of degradation of the image, since they introduce noise, making difficult the quantification. Also, the acquisition time is reduced, resulting in noisy and low resolution images [2, 3]. The Poisson noise affects the image quality negatively, influencing the detectability of lesions and the medical diagnoses. It is the current interest of the PET community to find methods to recover and preserve the important information in the images. In this paper, we present an application of the AI in the health area.

Our proposal consists of a three-dimensional (3D) CNN to improve the PET sinogram, demonstrating that it is feasible and important to retrieve information between slices, but also intra slices. The paper is organized as follows. In section 2, the recent related works are presented. Section 3 presents the main concepts related to the topic and the proposed method. In section 4, we show the results of the experimentation and, finally in Section 5, conclusions are presented.

2 Related Work

A set of techniques have been proposed to solve the problem of poor quality in PET images. Conventional approaches include processing algorithms [4, 5], anatomically guided [6], and magnetic resonance imaging (MRI) guided algorithms with partial volume correction [7]. Although these methods try to minimize noise, loss of spatial resolution is still observed. The algorithms of artificial intelligence have been included

in the area of medical image reconstruction and enhancement. Most of the works focus on the reconstructed images.

They propose to use trained networks with pairs of low resolution and high resolution images [8,9]. The high-resolution images are obtained from an acquisition with a modern ultra-high definition scanner, and degraded to obtain the low-resolution version. Other authors incorporate into the network training, anatomical information obtained from a computed tomography (CT) or MRI scanner [10, 11], arguing that this information is useful for estimating a more robust model and higher quality images.

Recent studies have focused their efforts on improving the sinogram instead of the reconstructed image. For example, in [12], the authors use Monte Carlo simulations and CNN to recover improved 2D sinograms from the low-resolution originals produced by simulated tomographs with large and small crystals.

2.1 The Importance of the Third Dimension

With the increase in computing power, some researchers have explored the possibility of addressing the problem of the low resolution of PET images in a three-dimensional way using deep learning systems.

For example, in [13], a 3D variant of the U-Net network is proposed to improve reconstructed images, reducing noise in PET images of the brain and chest. We think that in the domain of the sinogram is possible to obtain advantages if the sinogram is processed as a volume instead of sequential two-dimensional (2D) slices.

Despite the more common view of sinograms, as 2D slices, however, in PET, the acquired sinogram are volumes. Sinograms are three-dimensional structures, where not only the intra slice information (x, y, axes) is important but also the inter slices information (z-axis). In this work, we propose a 3D CNN to improve the quality of the acquired 3D sinograms.

3 Materials and Methods

This section presents the main methods and concepts related to PET tomography. Terms related to CNNs are also explained, as well as the 3D CNN proposed in this work is detailed.

3.1 PET Sinograms

A 2D sinogram is a matrix whose axes correspond to the angles (α) versus the orthogonal distances (S) of the orthogonal lines from the center of the tomograph to the LOR. Figure 1(b) shows a 2D sinogram with two events from the LORs shown in Figure 1(a).

In this case, on one of the LORs, at an angle (α_1), occurred two events and on the LOR at an angle (α_2) occurred one event. The number of possible angles is determined by the number of crystals per ring and the number of possible distances.

Also, this number depends on the number of crystals, and the transaxial field of view of the scanner [2,3]. Figure 2(a) shows a scanner with the arrow pointing to the axial axis. Figure 2(b) shows the acquisition of 2D sinograms. In this case, the events whose

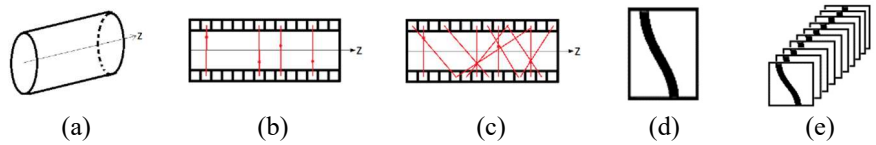


Fig. 2. Events acquisition from (a) a tomograph, (b) acquisition of 2D sinograms (one sinogram per ring detector), (c) acquisition of a 3D sinogram (only one volume), (d) 2D sinogram and (e) 3D sinogram.

LORs occur in the same axial plane of each detector ring are recorded in different 2D slices. Figure 2(c) shows that the events detected by crystals, from different rings, are stored in volumetric files, also called 3D sinograms.

This implies a substantial increase in the size and in the reconstruction time. The quality of reconstructed images is higher, due to the greater amount of information available. All the information detected in a 3D acquisition is stored in a 3D sinogram, where each corresponds to a 2D sinogram. Figures 2(d) and 2(e) shows the representation of 2D and 3D sinograms respectively.

3.2 PET Ordered Subset Expectation Maximization Algorithm

The maximum likelihood expectation-maximization algorithm (MLEM) was introduced in the field of image reconstruction in [14]. A variant of the MLEM method is the ordered subset expectation-maximization algorithm (OSEM) [15]. This method groups the PET scanner detectors into subsets to perform the processing of each subset in iterations, doing one subset at a time. OSEM reduces the reconstruction time relative to MLEM.

3.3 Scanner MicroPET FOCUS 220

We simulated the MicroPET FOCUS 220 preclinical scanner using the Gamos software [16] to perform the experiments. The scanner consists of four detector rings: each ring is made up of 42 detector blocks. Each detector block is composed of a matrix of 12 × 12 LSO crystals with dimensions of 1.5 mm × 1.5 mm × 10.0 mm. Its axial field of view is 7.6 cm and its transaxial field of view is 19 cm [17]. So that, one can acquire a 3D sinogram of size 252 × 287 × 2304.

3.4 Convolutional Filters

The convolution filters are two-dimensional or three-dimensional matrices to perform the convolution operation on the image to extract different characteristics. Typically, they are of size 3 × 3 or 5 × 5. The filter moves through the image from left to right, and from top to bottom, and in the case of 3D, from front to back, advancing a certain number of steps known as a stride.

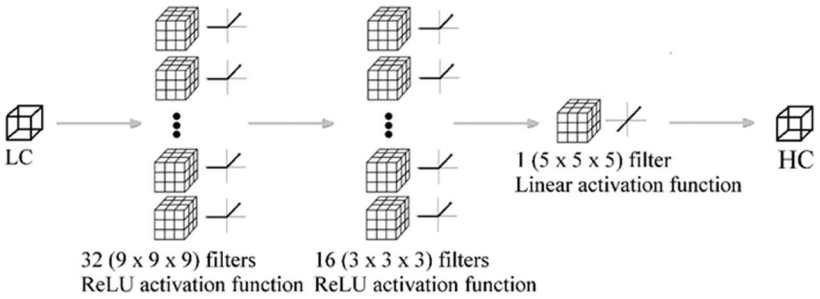


Fig. 3. Proposed 3D CNN.

Table 1. Resulting hyperparameters after tuning the proposed 3D CNN.

Parameter	Value
Loss function	Mean Squared Error (MSE)
Optimization algorithm	Adam
Learning rate	0.0003
Batch size	128
Epochs	200
Size of filters in the input layer	9 × 9 × 9
Size of filters in the mapping layer	3 × 3 × 3
Size of filters in the transpose convolution layer	5 × 5 × 5
Number of filters in the input layer	32
Number of filters in the mapping layer size	16
Number of filters in the transpose convolution layer	1

3.5 Convolutional Neural Networks

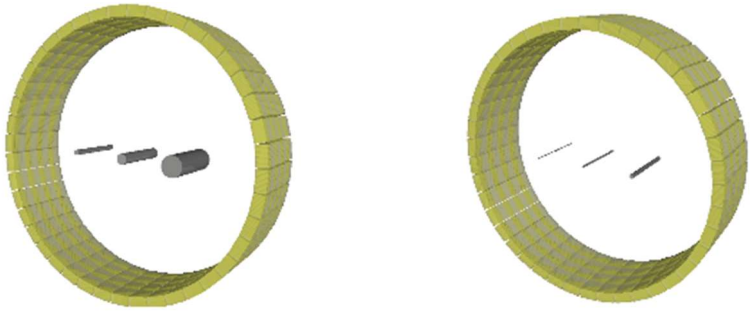
A CNN is a deep learning algorithm that can incorporate an input image, assign importance (weights and learnable biases) to various aspects or objects in the image and differentiate one from another, among other activities.

The preprocessing required in a CNN is much less compared to other classification algorithms. While in primitive methods, the filters are designed by hand, with enough training, CNNs can learn these filters.

Also, a CNN can capture spatial and temporal dependencies in an image through the application of filters.The architecture is better suited to the image data set due to the reduction in the number of parameters involved and the reuse of weights [18].

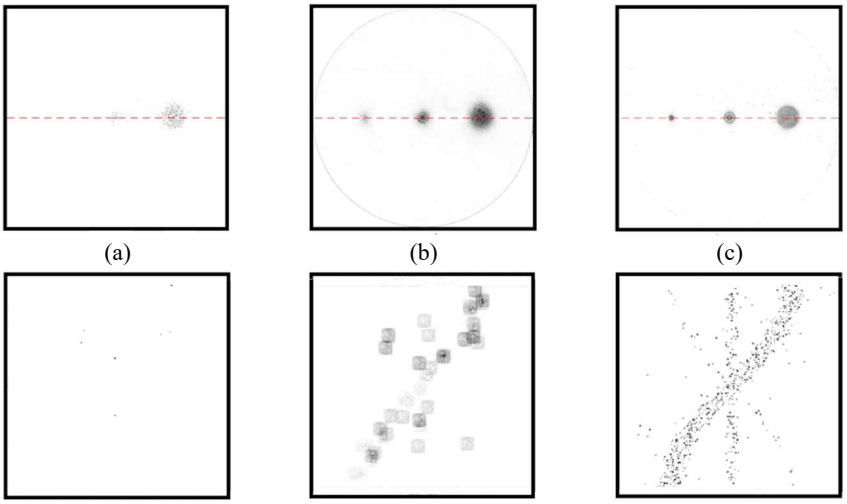
3.6 Deconvolution

Deconvolution is the inverse operation of convolution. It is used to recover data degraded by a physical process modeled as a convolution. If the degraded signal and



(a) (b)

Fig. 4. Simulated phantom of dimensions (a) 5 mm, 10 mm and 20 mm in diameter and 6 cm in length and (b) 1 mm, 2 mm, and 5 mm in diameter and 6 cm in length.



(a) (b) (c)

(d) (e) (f)

Fig. 5. MIPs of the transversal view of the simulated phantom with three cylinders of 5 mm, 10 mm and 20 mm in diameter. (a) LC (PSNR=26.33 dB, C=1.1680), (b) enhanced (PSNR=32.01 dB, C=2.5014) and (c) ground truth phantoms. Middle slice from the 3D (d) LC sinogram, (e) recovered sinogram from the proposed 3D CNN and (f) ground truth sinogram.

the system are known, it is possible to find the original signal by a deconvolution operation [19].

3.7 Dataset

High counts (HC) and low counts (LC) sinograms are generated to train the network. Five phantoms were simulated for the generation of the sinograms, containing spheres with diameters ranging from 0.5mm to 5mm. The spheres are randomly distributed within the simulated phantoms, so that the sinograms obtained contain enough useful

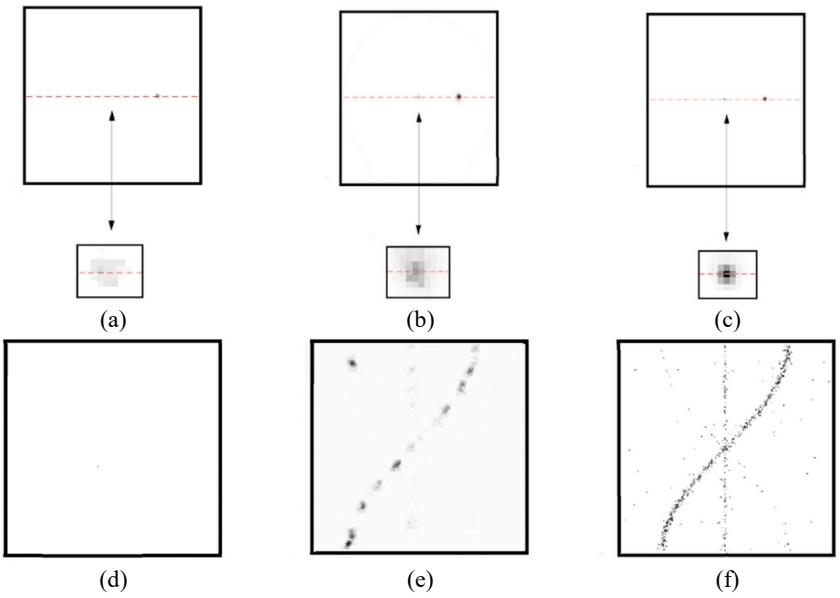


Fig. 6. MIPs of the transversal view of the simulated phantom with three cylinders of 1 mm, 2 mm and 5 mm in diameter. (a) LC (PSNR=25.12 dB, C=1.1521), (b) enhanced (PSNR=27.23 dB, C=2.4072) and (c) ground truth phantoms. Middle slice from the 3D (d) LC sinogram, (e) recovered sinogram from the proposed 3D CNN and (f) ground truth sinogram.

information for training. The LC sinogram is generated with 10 million events, of which only effective PET events are, about 5%.

One hundred million events were generated in the phantom to obtain the HC sinograms of which only 5% are true events. For network training, we extract 3D patches of size 32 × 32 × 32, selected randomly from the pairs of LC and HC sinograms. Our training set contains 15250 LC-HC patch pairs, while our validation set consists of 5000 patch pairs.

3.8 Proposed 3D CNN for PET Sinogram Enhancement

Figure 3 shows the diagram of the 3D CNN proposed to improve low resolution sinograms. The network has three layers. The input layer extracts the characteristics of the low-resolution sinograms, the mapping layer performs the mapping between the low resolution and high-resolution features, and the third layer or transpose convolution (deconvolution) layer performs the final reconstruction.

The input layer consists of 32 filters of size 9 × 9 × 9, with the ReLU [20]. The mapping layer has 16 filters of 3 × 3 × 3, with ReLU. The third layer has one filter of 5 × 5 × 5, with a linear activation function. Table 1 shows the hyperparameters adjusted by using the grid search method for the proposed 3D CNN. The proposed network has three layers only. Therefore, it is feasible and more efficient to tune the hyperparameters by the grid search method [21, 22].

4 Experiments

In this section, we present the methodology of evaluation and discuss the experimental results. We simulated six rods of 1 mm, 2 mm, 5 mm, 10 mm and 20 mm using the scanner MicroPET FOCUS 220. To examine the performance, we utilized the metrics peak signal to noise ratio (PSNR) and contrast (C) according to [23].

In our case, contrast indicates the difference between the lesions and the background in the image. If the difference is large, the lesions are more distinguishable.

4.1 Methodology of Evaluation

The maximum intensity projection (MIP) of the transversal views was used to obtain the quantitative results. A series of case studies were tested to verify the efficiency of the proposed method. In each experiment, the original low quality phantom was simulated with 500 thousand events and its corresponding ground-truth with 10 million events. For the simulation of the test phantoms, the Gamos software [16] was used. The OSEM algorithm [15] was used for the reconstruction.

4.2 Experimental Results

Figure 4(a) shows a phantom with three cylinders of 5 mm, 10 mm, and 20 mm in diameter and 6 cm in length, respectively. Figure 4(b) shows a phantom with three cylinders of 1 mm, 2 mm, and 5 mm in diameter and 6 cm in length. Both phantoms filled with F^{18} were used to acquire 3D sinograms for testing. Figure 5 shows the MIPs of the reconstructed images with OSEM from three cylinders of 5 mm, 10 mm, and 20 mm in diameter and their corresponding sinograms.

Figure 6 shows the MIPs of the reconstructed images with OSEM from three cylinders of 1 mm, 2 mm, and 5 mm in diameter and their corresponding sinograms. The proposed network increases the number of counts and the contrast, which positively affects the visualization. For example, Figure 7 (a) shows the MIPs of Figure 5. In this case, the proposed CNN recovers the 5mm rod. Figure 7 (b) shows the MIPs of Figure 6. In this case, the proposed CNN also recovers the 1mm rod.

The results obtained in the experiments allow us to affirm that the proposed method applied to the PET sinograms, in three dimensions, permits to increase the contrast of the rods in the reconstructed images.

A better definition of the limits between the area containing the radiotracer and the background area is achieved. Figures 5(a to c) and 6(a to c) visually show that when the LC sinograms are processed by our method, valuable information is recovered when reconstructed.

For example, in Figure 6(b), it is possible to observe how the proposed method recovers the intermediate lesion of 2mm in diameter, when it was originally almost invisible Figure 6(a). Table 2 compares the PSNR and contrast after recovering the low count phantom and the enhanced phantom. Notice that our network results (bold values) increase the PSNR by 6% on average and the contrast almost twice. Figures 5(d to f) and 6(d to f) clearly shows how the proposed method manages to recover effective counts in the sinograms, while the LC sinograms (Figures 5(d) and 6(d)) contain little

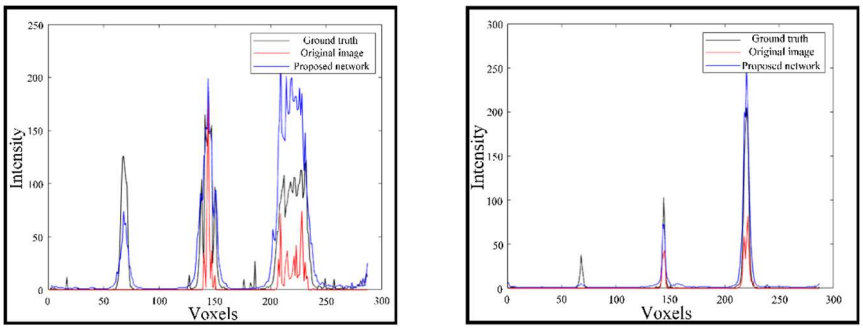


Fig. 7. Horizontal profiles of the (a) Figure 5 for the MIP of the phantom of dimensions 5 mm, 10 mm, and 20 mm in diameter and (b) Figure 6 for the MIP of the phantom of dimensions 1 mm, 2 mm, and 5 mm in diameter.

Table 2. Summary of results.

		LC phantom	Enhanced phantom
Case study 1	PSNR	26.33 dB	**32.01 dB**
	Contrast	1.1680	**2.5014**
Case study 2	PSNR	25.12 dB	**27.23 dB**
	Contrast	1.1521	**2.4072**

Note: Bold values indicate the best results.

information, after being processed by our method, useful information is recovered for reconstruction (Figures 5(e) and 6(e)).

5 Conclusions

In this article, we presented an application of the AI in the health area. A 3D CNN was proposed to enhance sinograms of PET images.

Sinograms with a high number of counts and their corresponding sinograms with few counts were generated 32 × 32 ×32 patches were extracted from both pairs of sinograms to form the training sets. An analysis of the profiles obtained from the MIP of the reconstructed images is carried out.

The results show that the proposed network can increase the number of counts in the sinogram, which positively influences the quality of the images. Our method demonstrates the importance of treating the sinogram as a three-dimensional structure to consider the intra-slice information.

The main advantages are the improvement of the 3D sinograms and the 3D CNN architecture with three layers only. However, the main drawback is that the method needs to be trained first, so that it is computationally expensive and time-consuming.

Acknowledgments. L. J. Rodríguez thanks the UACJ for the support provided and the CONACYT for the scholarship granted to carry out his doctoral studies. H. Sossa thanks the Instituto Politécnico Nacional for the support under projects SIP 20210788, and CONACYT under projects 65 (Fronteras de la Ciencia) and 6005 (FORDECYTPRONACES).

References

1. Al-Jaroodi, J., Mohamed, N., Abukhousa, E.: Health 4.0: On the way to realizing the healthcare of the future. IEEE Access, vol. 8, pp. 211189–211210 (2020) doi: 10.1109/ACCESS.2020.3038858
2. Saha, G. B.: Pet scanning systems. In: Basics of PET Imaging. Springer-Verlag New York (2010)
3. Cherry, S., Dahlbom, M.: PET Physics, Instrumentation, and Scanners. Springer-Verlag, New York (2006)
4. Arabi, H., Zaidi, H.: Improvement of image quality in pet using post-reconstruction hybrid spatial-frequency domain filtering. Physics in Medicine & Biology, vol. 63**,** pp. 215010 (2018) doi: 10.1088/1361–6560/aae573
5. Arabi, H., Zaidi, H.: Spatially guided nonlocal mean approach for denoising of pet images. Medical Physics, vol. 47, pp. 1656–1669 (2020) doi: 10.1002/mp.14024
6. Chan, C., Fulton, R., Barnett, R., Feng, D., Meikle, S.: Postreconstruction nonlocal means filtering of whole-body pet with an anatomical prior. IEEE Transactions on Medical Imaging, vol. 33**,** pp. 636–650 (2014) doi: 10.1109/TMI.2013.2292881
7. Yan, J., Lim, J., Townsend, D.: MRI-guided brain pet image filtering and partial volume correction. Physics in Medicine & Biology, vol. 60, no. 3, pp. 961–976 (2015)
8. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., van Ginneken, B., Sánchez, C. I.: A survey on deep learning in medical image analysis. Medical Image Analysis, vol. 42, pp. 60–88 (2017) doi: 10.1016/j.media.2017.07.005
9. Chen, K., Gong, E., de Carvalho, F., Xu, J., Boumis, A., Khalighi, M., Poston, K. L., Sha, S. J., Greicius, M. D., Mormino, E. Pauly, J. M., Srinivas, S., Zaharchuk G.: Ultra–Low-Dose (18) F-Florbetaben Amyloid PET Imaging Using Deep Learning with Multi-Contrast MRI Inputs. Radiology, vol. 290, 649–656 (2019) doi: 10.1148/radiol.2018180940
10. Xu, J., Gong, E., Pauly, J., Zaharchuk, G.: 200x Low-dose PET reconstruction using deep learning. ARXIV (2017) eprint arXiv:1712.04119
11. Liu, C., Qi, J.: Higher SNR PET image prediction using a deep learning model and MRI image. Physics in Medicine & Biology, vol. 64, no. 2019, pp. 115004, doi: 10.1088/1361–6560/ab0dc0
12. Hong, X., Zan, Y., Weng, F., Tao, W., Peng, Q., Huang, Q.: Enhancing the image quality via transferred deep residual learning of coarse pet sinograms. IEEE Transactions on Medical Imaging, vol. 37, pp. 2322–2332 (2018)
13. Lu, W., Onofrey, J. A., Lu, Y., Shi, L., Ma, T., Liu, Y., Liu, C.: An investigation of quantitative accuracy for deep learning based denoising in oncological PET. Physics in Medicine & Biology, vol. 64, pp. 5019 (2019)
14. Shepp, L., Vardi, Y.: Maximum likelihood reconstruction for emission tomography. IEEE Transactions on Medical Imaging, vol. 1, pp. 113–122 (1982)
15. Hudson, H., Larkin, R.: Accelerated image reconstruction using ordered subsets of projection data. IEEE Transactions on Medical Imaging, vol. 13, pp. 601–609 (1994)
16. Arce, P., Lagares, J., Harkness, L., Pérez-Astudillo, D., Cañadas, M., Rato, P., de Prado, M., Abreu, Y., de Lorenzo, G., Kolstein, M., Díaz, A.: Gamos: A framework to do Geant4 simulations in different physics fields with a user-friendly interface. Nuclear Instruments and Methods in Physics, vol. 735, pp. 304–313 (2014)
17. Tai, Y., Ruangma, A., Rowland, D., Siegel, S., Newport, D., Chow, P., Laforest, R.: Performance evaluation of the microPET focus: a third-generation microPET scanner dedicated to animal imaging. J Nucl Med., vol. 46, pp. 455–463 (2005)
18. Socher, R., Huval, B., Bath, B., Manning, C. D., Ng, A.: Convolutional-recursive deep learning for 3D object classification. Advances in Neural Information Processing Systems, vol. 25 (2012)

19. Zeiler, M., Krishnan, D., Taylor, G., Fergus, R.: Deconvolutional networks. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 2528–2535 (2010) doi: 10.1109/CVPR.2010.5539957
20. Nair, V., Hinton, G.: Rectified linear units improve restricted Boltzmann machines. In: Proceedings of ICML. vol. 27, pp. 807–814 (2010)
21. Bengio, Y.: Practical recommendations for gradient-based training of deep architectures. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 437–478 (2012) doi: 10.1007/978-3-642-35289-8_26
22. Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. Journal of Machine Learning Research, vol. 13, pp. 281–305 (2012)
23. Kennedy, J., Israel, O., Frenkel, A., Bar-Shalom, R., Azhari, H.: Super-resolution in PET imaging. IEEE Transactions on Medical Imaging, vol. 25, pp. 137–147 (2006) doi: 10.1109/TMI.2005.861705