# Deep Learning methodologies for brain image reconstruction in Positron Emission Tomography

N. Rufo Rafael de la Cruz[1], J. E. Ortuño Fisac[2,3] and G. Kontaxakis[3]

[1] ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain, n.rufo@alumnos.upm.es

[2] CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, Spain

[3] Biomedical Image Technologies Group, Information Processing and Telecommunications Center, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain, {je.ortuno, g.kontaxakis}@upm.es

## Abstract

*Positron emission tomography (PET) is a functional imaging modality that, by its nature, is very prone to noise. Most current image reconstruction methods in PET have dealt with noise, but not always satisfactorily. Recently, novel approaches have been proposed, using the properties of Deep Learning (DL) methodologies to address this problem. We have designed and implemented a DL-based image reconstruction method, directly obtaining images from sinograms using an encoder-decoder architecture. The proposed method was tested in geometric phantoms, brain phantoms, as well as real brain scans, and its performance was compared with the one obtained by analytical and iterative methods. Our results suggest that DL methodologies might constitute a valid alternative approach for PET image reconstruction, reducing the influence of noise compared to traditional methods and potentially achieving a more accurate image interpretation for diagnosis and treatment, in addition to a potential reduction of the injected radioactive dose.*

## 1. Introduction

Positron emission tomography (PET) [1] is a nuclear medicine imaging modality that provides functional information about the metabolic activity of the living organism. It has been proven especially useful in diseases, such as cancer, where it has important role for both diagnosis and treatment [2]. Data collected during image acquisition are stored in the form of sinograms that cannot be interpreted directly by an observer and must be first reconstructed into images. However, a solution to the inverse problem is not straightforward due to the nature of this imaging modality, and thus, the reconstructed images are very prone to noise.

Image reconstruction methods that have been used over the years are analytical techniques, such as filtered backprojection (FBP) [3], which uses projection operators that do not model the physics and noise of a PET scanner, or iterative methods, such as Maximum Likelihood Expectation Maximization (ML-EM) [3], which can model the Poisson nature of the data and include a more complex physical model of the acquisition process. In addition, regularization terms can be added in the iterative process to consider a priori information. However, iterative methods have high computational demands, which increases the more elaborate and accurate the physical model of the system used is. Moreover, the arbitrariness in selecting the proper number of iterations

or in defining the most appropriate parameters involved in the specific algorithms used do not always contribute to obtaining the most accurate image as a final result. Hence, with the current growth of Deep Learning (DL) [4], very innovative methods have been recently proposed, that can learn the physical process of PET data generation and the noise model from the data provided for training, based on learning mappings from noisy data to high-quality images. Depending on the number, form, and type of these mappings, different architectures can be achieved. Of particular importance are the encoders and decoders that can represent abstract maps of the essential features of the data introduced in the encoder stage to be later transformed into images of outstanding quality in the decoder stage. Since noise is not a relevant feature in data, it is not included in the feature maps and, thus, not represented in the final image.

Currently, the main approaches to the use of DL in PET image reconstruction are: synthesis, analysis, unfolded, and direct methods [5]. The first three maintain the mechanisms of traditional iterative reconstruction methods, except for the use of DL for learning the elements that need prior image knowledge, avoiding any analytic, intuitive, or handcrafted component of the reconstruction process. Synthesis and analysis approaches learn the regularization step, and the unfolded methods extend the iterative methods in a succession of reconstruction operators, partially or completely interleaving them in repeated blocks with deep denoising operators.

The fourth approach leaves behind everything done so far with iterative methods and learns the direct mapping between sinogram and reconstructed image. These methods have the advantage not to require any prior information from the image, but demand a large amount of training data. Among these methods are: AUTOMAP [5], initially developed for magnetic resonance imaging (MRI), which learns to perform the image reconstruction in a way analogue to an inverse Radon transform, with fully connected layers before an autoencoder; DeepPET [6], which learns entirely on an encoder-decoder architecture; DirectPET [7], which again uses an encoder-decoder but incorporates at the end of the encoder phase the idea of the inverse Radon transform with fully connected layers. We study here the viability of direct methods by implementing a modified version of DeepPET.
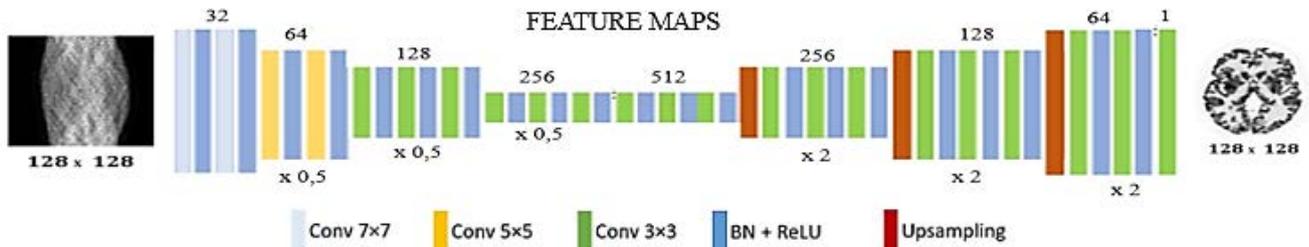
**Figure 1.** *Network architecture for our version of the DeepPET image reconstruction method.*

## 2. Methodology

### 2.1. Deep Learning method

We implemented a modified version of DeepPET, where another image format is given and thus the abstraction level that the feature maps reach differs from the original, resulting in a reduction of the total number of layers from 31 to 22, which was found to produce betters results and also reduces the memory requirements. We also changed the optimizer from Adam stochastic gradient descent to RMSProp from Keras, the Python DL Application Programming Interface (API), that we also found to produce better results. The mean squared error (MSE) was maintained as the loss function.

Fig. 1 depicts the resulting architecture, which consists of an encoding phase, where the sinogram input is compressed into feature maps through convolution layers, and a decoding phase, in which additional up-sampling layers are introduced to reach the final reconstructed image. After each convolution at encoding, a ReLU activation function and a batch normalization are used.

To further alleviate the massive memory requirements of this method, the network has been fed with a custom batch generator that loads GPU input data from the disk in a volatile and dynamic manner when needed.

### 2.2. Traditional methods

For comparison, the FBP with a Hamming filter and the ML-EM and its regularized version with a median filter, have been implemented using the ASTRA Toolbox library [8], as the most representative current image reconstruction methods in PET.

### 2.3. Datasets

To test the performance of the network the following datasets are used:

- Geometric phantom dataset: 100000 unique images, each of them with 20 ellipses of varying intensity, orientation and size simulated in MATLAB.
- Brain phantom dataset, containing the 4120 images (2D slices) from the 20 brain phantoms 3D volumes initially simulated for MRI and adapted to the PET domain, publicly available [9][10].
- Brain phantom lesions dataset: same as before, with hot and cold lesions introduced, publicly available [9][10].
- Real brain images dataset: 5230 PET images (2D slices) from 37 healthy patients 3D volumes, publicly available for research purposes [11].

For all datasets, 80% of the images are destined for training, 10% for validation, and 10% for testing, except for the geometric phantom, where only 100 images were kept for the tests.

For the brain datasets, data augmentation ($\times 50$) was performed to increase the variability during the training. Sinograms were simulated using the ASTRA Toolbox library, with 2D ideal geometry and parallel projectors. Finally, random noise with 160 mean Poisson distribution is added to the sinograms to mimic the statistics of the PET acquisition. Neither attenuation nor scattering effects were included. An example is shown in Fig. 2.
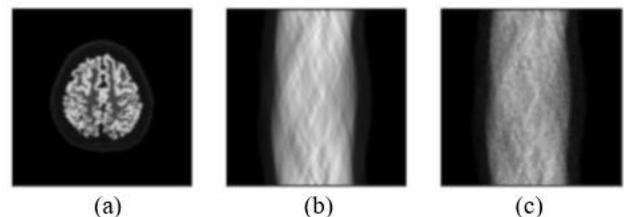


**Figure 2.** *(a) Sample from the brain phantom dataset, (b) ideal sinogram simulated, (c) Sinogram with Poisson noise.*

## 3. Results and discussion

The experiments have been divided into three phases according to the dataset used and whose results have been evaluated with the structural similarity metric (SSIM), peak signal-to-noise ratio (PSNR) and normalized root mean square error (NRMSE) [12]. Besides, the working environment implemented (traditional methods, network model and training, inference module) has been uploaded to a GitHub repository to allow the community to perform their own experiments and explore this field, given that such code was not available before, available at: *https://github.com/NuriaRufo/DeepPET*.

### 3.1. Geometric phantom dataset

Initially, a training of the network was conducted using the training set of geometric phantoms of 90000 images with size 128×128. The learning rate was set to $10^{-4}$ and the batch size to 32. The GPU used was PNY QUADRO RTX-A6000 48GB GDDR6.

The mean error of images in the test set reconstructed with the different methods can be seen in Table 1. The DL method outperforms the traditional methods, which can be also qualitatively verified in Fig. 3.

| Method | SSIM | PSNR | NRMSE |
|---|---|---|---|
| FBP | 0,4769 | 21,6767 | 0,5606 |
| ML-EM | 0,8002 | 28,1906 | 0,2664 |
| ML-EM regularized | 0,8875 | 31,1608 | 0,1892 |
| DL method | **0,9228** | **32,0470** | **0,1710** |

**Table 1.** *Average error for geometric phantom dataset reconstruction.*
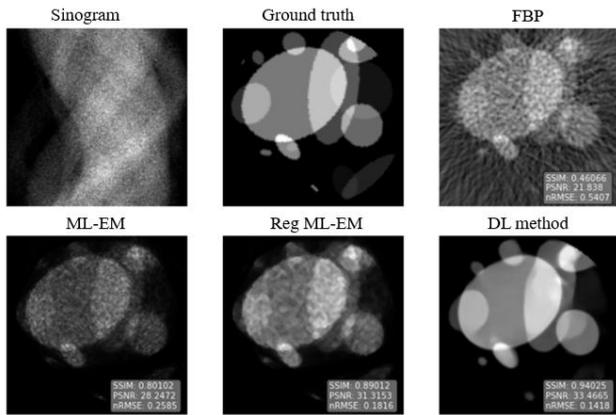


**Figure 3.** *Reconstructed images with traditional methods and the direct DL method.*

### 3.2. Brain phantom dataset

The first training was performed excluding lesions and considering only axial slices, to delimit the problem. The same hyperparameters and GPU of the previous experiment were maintained. Once the network was trained, the results showed that the DL method was able to reconstruct structures as delicate as the brain, although it did not surpass the quality of traditional methods.

To test the generalizability of the method, the model trained with the phantom without lesions and axial slices was confronted with the phantom test dataset with lesions. It was observed that the model could recognize lesions, although not with sufficiently high quality. Therefore, a new training was performed with the phantom training set with lesions, but still considering just the axial slices and where hyperparameters were maintained. The introduction of random lesions in the training set improved network performance by 4%, although it still did not outperform traditional methods.

In search of a more considerable amount of training data, sagittal and coronal slices were added to the axial ones for a new training with the same hyperparameters. However, the model trained with all slices was confronted with only axial slices from the test dataset to ensure homogeneity between brain phantom results. Fig. 4 shows how the qualitative results are finally favourable for the direct DL method. It is worth mentioning that the reconstruction times for this method are remarkably fast, as all images were reconstructed in approximately one second, compared to traditional reconstruction times that were 800 seconds using CPU and 20 seconds using GPU.

Despite the satisfactory results in single samples, the mean error yields a better reconstruction by the traditional
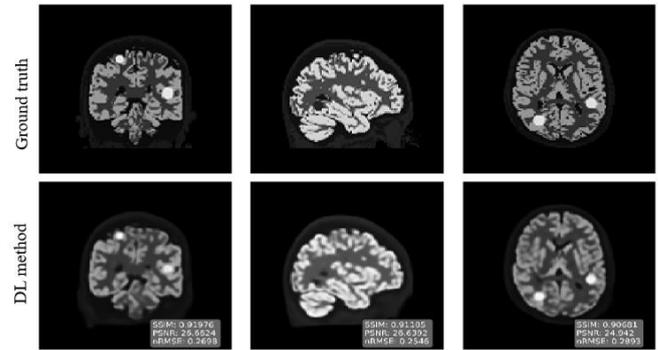


**Figure 4.** *Three slices of a brain phantom test set reconstructed using the DL method (lower row), compared with the ground truth (upper row).*

| Method | SSIM | PSNR | NRMSE |
|---|---|---|---|
| FBP | 0,3900 | 20,8424 | 0,6425 |
| ML-EM | 0,9223 | 27,6359 | 0,2934 |
| ML-EM regularized | **0,9335** | **28,2540** | **0,2733** |
| DL method | 0,8889 | 26,0090 | 0,4892 |
| DL method (lesions) | 0,9041 | 26,7298 | 0,4344 |
| DL method, all cuts | 0,9260 | 27,5362 | 0,3329 |

**Table 2.** *Average error for the brain phantom test dataset containing lesion reconstruction and only axial slices.*

ML-EM methods and its regularized version. However, a study of the distribution of these errors reveals that this is due to a few outliers with poor DL reconstructions, caused by the presence of structures adjacent to the brain, which are underrepresented in the training and, consequently, not learned well enough. Nevertheless, regarding the median in the error distribution, the DL method is placed ahead of the traditional methods. In addition, 74,4% of the cases were better reconstructed with the network than with the traditional methods.

A set of 5 random lesions was analysed in detail (Fig. 5). The DL method resulted in a better qualitative and quantitative reconstruction with average relative error at 0,2928, whereas this was 0,3443 for regularized ML-EM.
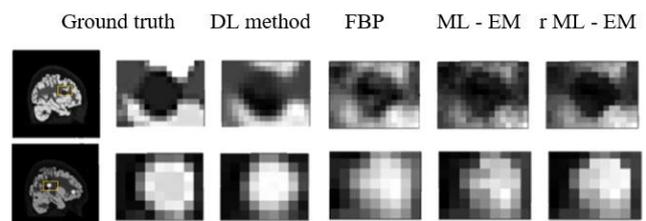


**Figure 5.** *Reconstructed lesions using FBP, ML-EM, regularized ML-EM and the DL method.*

### 3.3. Real brain images dataset

Regarding real brain PET images, the first experiment with the test dataset (all slices) was first performed with the best model trained with the brain phantom (trained with all slices). We observed that it tends to over-define the structures of the cerebellum, as it occurs normally in the brain phantom.

Ground truth      Reg ML-EM      DL method, All Cuts      DL method, Fine Tune
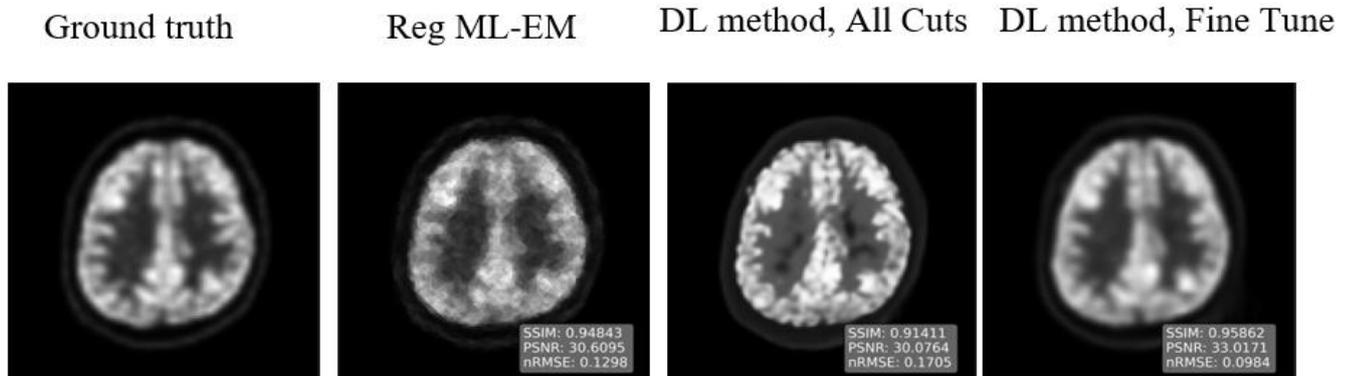


Figure 6. *Reconstruction results for a sample of the real PET brain images dataset using the DL method, compared with the ground truth.*

| Method | SSIM | PSNR | NRMSE |
|---|---|---|---|
| FBP | 0,4435 | 20,4825 | 0,5579 |
| ML-EM | 0,9241 | 30,1782 | 0,1900 |
| ML-EM regularized | **0,9464** | 31,9411 | 0,1530 |
| DL method, All Cuts | 0,8788 | 32,8847 | 0,2276 |
| DL met. Fine Tune | 0,9411 | **32,8847** | **0,1489** |

Table 3. *Average error for real PET dataset image reconstruction.*

Therefore, a new training was performed for fine-tuning, taking advantage of the prior knowledge of an already trained network to adapt it to the new problem. Based on the model trained with the phantom, training was performed with the real brain datasets to adapt the network to the domain of these images. The learning rate was lowered to 0,00005 and the number of epochs was set to only 8. In the training of one hour, the results of the tests yield that it both qualitatively (Figure 6) and quantitatively (Table 3) overcomes traditional methods, although the SSIM is again worsened by outliers containing structures other than the brain.

## 4. Conclusions

The current study focused on the case of brain image reconstruction for PET. In view of the results, we conclude that direct DL image reconstruction methods constitute a powerful tool for PET image reconstruction from sinograms. It has been confirmed that the modified version of DeepPET is able to produce real PET image reconstructions with very satisfactory quality in terms of noise characteristics from noisy sinograms. Moreover, it is much faster and overcomes in image quality traditional FBP, ML-EM and regularised ML-EM algorithms in 100%, 90,3% and 85,4 % of the cases, respectively.

However, we have also confirmed the method's robustness, and therefore, the outperformance of the traditional methods is due to the homogeneity and size of the training dataset used. This, in turn, highlights the major disadvantage of direct methods, which is the vast memory requirements to hold and handle the minimum 250000 image pairs necessary to learn the whole mapping from sinogram to image and thus to obtain a functional and competitive model. Consequently, the balance may

tip towards the use of synthesis or unfolded methods that take the basis of iterative image reconstruction, which, as we have seen, has not always been trivial to overcome in terms of image quality, therefore combining Deep Learning with traditional methods may be a promising approach.

This work implies, however, the validation of direct reconstruction and shows the potential that DL can have in the field of medical image reconstruction, providing us with higher quality images compared with traditional methods that can result both in better diagnosis and treatment for cancer patients as well as in opening the possibility that less doses of radioactive material may be needed to perform the image acquisition.

## References

[1] M. E. Phelps, *PET: physics, instrumentation, and scanners*. Springer, 2006.

[2] D. Delbeke, "Oncological applications of FDG PET imaging: Brain tumors, colorectal cancer lymphoma and melanoma," *J. Nucl. Med.*, 40(4): 591–603, 1999.

[3] D.W. Wilson and B.M.W. Tsui, "Noise properties of filtered-backprojection and ML-EM reconstructed emission tomographic images," *IEEE Trans. Nucl. Sci.*, 40(4): 1198-1203, 1993.

[4] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, *et al*, "Introduction to machine learning, neural networks, and deep learning," *Transl. Vis. Sci. Techn.*, 9(2), 2020.

[5] A.J. Reader, G. Corda, *et al*, "Deep Learning for PET Image Reconstruction," *IEEE Trans. Rad. Plasma Med. Sci.*, 5(1): 1–25, 2020.

[6] I. Häggström, C.R. Schmidtlein, *et al*, "DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem," *Med. Imag. Anal.*, 54: 253–262, May 2019.

[7] W. Whiteley, W. K. Luk, and J. Gregor, "DirectPET: full-size neural network PET reconstruction from sinogram data," *J. Med. Imag*, 7(3): 32503, 2020.

[8] W. van Aarle *et al.*, "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, 157: 35–47, Oct. 2015.

[9] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans, "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database," *Neuroimage*, 5(4): S425, 1997.

[10] A. Mehranian and A.J. Reader, "Model-Based Deep Learning PET Image Reconstruction Using Forward–Backward Splitting EM," *IEEE Trans. Rad. Plasma Med Sci*, 5(1): 54–64, 2021.

[11] I. Mérida, *et al*, "CERMEP-IDB-MRXFDG: a database of 37 normal adult human brain [18F]FDG PET, T1 and FLAIR MRI, and CT images available for research," *EJNMMI Res.*, 11(1):91, 2021.

[12] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Imag. Proc.*, 13(4): 600–612, April 2004.