Deep Learning methodologies for direct image reconstruction and integrated attenuation correction in brain PET/MRI

P. Díaz Serrano[1], J. E. Ortuño Fisac[2,3], J. López Santiago[1], F. Panetsos Petrova[4,5], G. Kontaxakis[3]

[1] Universidad Carlos III de Madrid (UC3M), Madrid, España; pameladiaz2318@gmail.com, jlopez@tsc.uc3m.es

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

[3] ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain, {je.ortuno, g.kontaxakis}@upm.es

[4] Neurocomputing and Neurorobotics Research Group, Faculty of Biology and Faculty of Optics, Universidad Complutense de Madrid (UCM), Madrid, Spain; fivos@ucm.es

[5] Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain

Abstract

This study proposes the ATTDeepPET model, a novel deep learning architecture crafted specifically for advancing positron emission tomography (PET) image reconstruction in PET/MR scanners. By incorporating magnetic resonance (MR) images into its learning process, ATTDeepPET addresses the persistent challenges associated with attenuation effects in PET/MR scanners, eliminating the need for simulated transmission scans. ATTDeepPET's performance is assessed alongside the deep learning model DeepPET, as well as established methods such as FBP, ML-EM, and ML-EMR for comparison. The findings reveal noteworthy achievements since ATTDeepPET accomplishes competitive image quality compared to FBP, ML-EM, and ML-EMR when applied to brain phantoms while also demonstrating a reduction in reconstruction times. Nevertheless, when dealing with real PET images, ATTDeepPET does exhibit some performance variability, underscoring the increased complexity of real-set scenarios and the importance of employing diverse datasets to enhance its robustness. Moreover, ATTDeepPET, despite inherent limitations, including heightened memory requirements and sensitivity to dataset variations, presents a promising path forward for PET image reconstruction. Its hallmark traits include exceptional execution speed, liberation from the prerequisite of prior physics knowledge, and the prospect of obviating the need for an additional CT scan for attenuation correction. These attributes hold transformative potential in terms of enhancing diagnostic precision and curtailing patient radiation exposure.

1. Introduction

Positron emission tomography (PET) is a valuable tool in medical imaging, providing insights into an organism's metabolic activity [1], particularly benefiting cancer diagnosis and treatment [2]. During PET image acquisition, data is recorded as sinograms, requiring reconstruction to produce interpretable images. However, solving the inverse problem for this imaging modality is complex, leading to reconstructed images that are highly susceptible to noise.

Analytical methods, like Filtered Backprojection (FBP), lack noise and physics modelling [3]. Conversely, iterative methods, such as Maximum Likelihood Expectation Maximization (ML-EM), offer the capability to model the Poisson statistical noise present in the acquired data and incorporate more intricate physical models [3]. Furthermore, these iterative methods allow for the inclusion of regularization terms during the iterative process, considering prior knowledge.

Nonetheless, it is essential to note that iterative methods come with significant computational demands, particularly as the complexity and accuracy of the employed physical model increase. Also, PET imaging suffers from attenuation artefacts, which are caused due to varying tissue photon absorption coefficients, leading to a distorted tracer distribution depiction.

Current attenuation correction methods have limitations, such as the practical challenges and regulatory complexities associated with gamma sources [4]. Alternatively, using attenuation maps from X-ray computerized tomography (CT) scans exposes patients to additional ionizing radiation.

As Deep Learning (DL) methodologies continue to expand, novel approaches have emerged within the field. These methods possess the capability to grasp the underlying physics of PET data generation, addressing noise and attenuation effects directly through the training data [5]. DL methods find applications in various aspects of image reconstruction, often replacing specific components, such as regularization, within conventional methods. However, this work focuses on DL architectures that aim for end-to-end mapping, involving direct reconstruction by deep networks from raw sinogram data to a reconstructed image. This approach requires assimilating all aspects of the image reconstruction process, including physics, imaging models, and statistics, demanding a substantial amount of training data. With this objective, encoders-decoders architectures are of relevance, capturing essential data features in the encoder stage and transforming them into high-quality images in the decoder stage [5].

Traditional practice involves the use of CT images for attenuation correction in PET scans. However, there is a growing interest in replacing CT with magnetic resonance (MR) images due to the advantages offered by PET/MR systems, such as improved soft tissue contrast and reduced ionizing radiation [6]. However, since MR images do not inherently provide attenuation information, generating attenuation maps from PET/MR scanners is challenging. Various methodologies, including machine learning, have been proposed to address this issue, such as creating synthetic CT (sCT) images from MR data [7].

ISBN: 978-84-17853-76-1

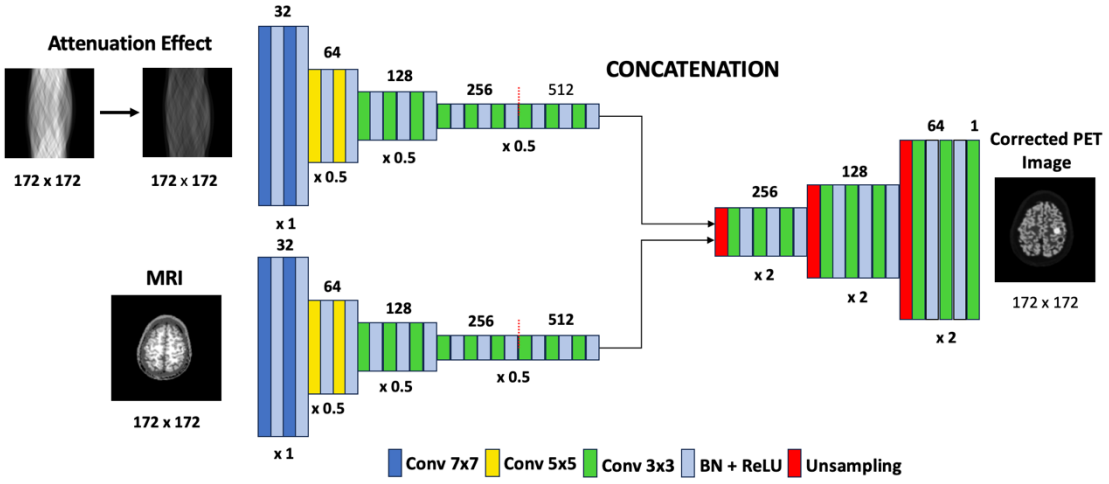

Figure 1. *Network ATTDeepPET image reconstruction method.*

This study introduces an alternative approach within the DeepPET [8] architecture. Instead of creating intermediate sCT images, this proposed method directly learns from MR images to perform both attenuation correction and PET image reconstruction. This approach eliminates the need to convert MR images to sCT, simplifying the process by enabling the network to learn the complex relationship between attenuation correction and image reconstruction directly from MR inputs.

2. Methodology

2.1. Deep Learning methods

A modified version of the DeepPET method [9], was implemented. Several improvements were made to the original architecture, including reducing the total number of layers from 31 to 22, which helped lower the memory requirements. Additionally, the optimizer was switched from Adam stochastic gradient descent to RMSProp from Keras[1]. The loss function remained as the mean squared error (MSE). The resulting architecture comprised an encoding phase where sinogram input is compressed into feature maps through convolution layers, and a decoding phase introducing up-sampling layers to produce the final reconstructed image. After each convolution in the encoding stage, a ReLU activation function and batch normalization are applied.

Then, a second architecture was implemented, where a second encoder was added to incorporate co-registered MR images (Fig. 1), introducing anatomical details to the reconstruction and attenuation correction process. We named this architecture ATTDeepPET. For both networks, sinograms with attenuation effects were introduced to evaluate the ability of the networks to reconstruct and correct the attenuation effect.

Furthermore, in order to reduce the memory demands of the method, a customized batch generator was employed. This batch generator dynamically loads GPU input data from the disk as required, operating in a volatile and adaptable manner.

2.2. Traditional methods

To evaluate the efficacy of the DL models in comparison to conventional approaches, a comparative analysis involving three key methodologies was implemented: FBP, ML-EM and its regularized variant (ML-EMR) were implemented in a code using the ASTRA Toolbox library.

2.3. Datasets

To test the performance of the network, the following datasets were used, all conformed by PET, CT, and MR co-registered images. The CT sinogram was used to generate the attenuation effect in the input PET sinograms, which were obtained using the ASTRA toolbox with 2D ideal geometry and parallel projectors. Lastly, random noise following a Poisson distribution was introduced into the sinograms to replicate the statistical characteristics of PET data acquisition. Importantly, this simulation does not incorporate attenuation or scattering effects:

- Brain phantom dataset consisting of 64596 2D axial slice images extracted from 20 brain phantom 3D volumes. These volumes were originally designed for MRI simulation and were subsequently adapted for PET imaging. This dataset is publicly accessible [10]. To broaden the analysis, some of the images were induced with artificial lesions.
- 15725 real PET images, including 2D axial, sagittal, and coronal slices, derived from the 3D volumes of 37 healthy patients [11].

For all datasets, 80% of the images are destined for training, 10% for validation, and 10% for testing.

[1] https://keras.io/

3. Results and discussion

The two distinct architectures were trained and tested with the same datasets. The evaluation phase involved the reconstruction of the test image set through the three distinct methodologies: FBP, ML-EM, ML-EMR, DeepPET, and the ATTDeepPET model. The resultant reconstructed images were subsequently subjected to a comparative analysis against a reference ground truth image with the structural similarity metric (SSIM), peak signal-to-noise ratio (PSNR) and normalized root mean square error (NRMSE) [12].

3.1. Brain phantom dataset

Initially, the DeepPET and ATTDeepPET networks were trained using the training set of brain phantoms of size 172x172. 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 results demonstrate that ATTDeepPET consistently outperforms other methods across all three metrics. DeepPET also exhibits strong performance, closely trailing behind ML-EMR and ML-EM, surpassing the performance of FBP, which consistently score lower across all metrics.

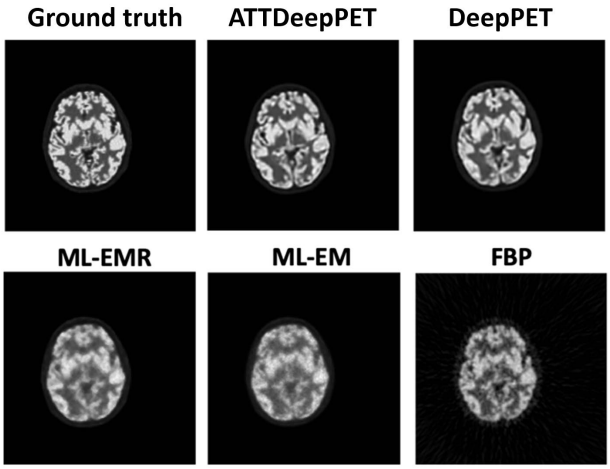

Figure 2. *Phantom reconstructed axial slices employing FBP, ML-EM, ML-EMR, DeepPET and ATTDeepPET*

A Tukey's Honestly Significant Difference (HSD) test was conducted, a statistical analysis that provided a nuanced understanding of the performance dynamics. In the context of brain phantoms, ATTDeepPET emerges as the standout performer, showcasing significantly higher SSIM and PSNR values, with statistical significance supported by the rejection of the null hypothesis ($p < 0.05$). Furthermore, ATTDeepPET demonstrates the lowest nRMSE, again with statistical significance compared to the other methods. Conversely, while DeepPET delivers commendable results, surpassing the performance of ML-EM and FBP methods, it's important to note that it does not exhibit a statistically significant improvement over the ML-EMR method.

Method	SSIM	PSNR	NRMSE
FBP	0,5461	26,4315	0,2554
ML-EM	0,9545	29,0161	0,2198
ML-EM regularized	0,9634	29,5834	0,2041
DeepPET	0,9677	30,2062	0,2082
ATTDeepPET	**0,9753**	**31,5333**	**0,1811**

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

3.2. Real brain images dataset

Subsequently, DeepPET and ATTDeepPET architectures were trained using the real PET images dataset. The learning rate was set to 10-4, and the batch size to 32. The same GPU was used. Obtained quantitative metrics SSIM, PSNR, and NRMSE were computed between each of the reconstructed images and the ground truth. The ensuing averages of these metrics are compiled in Table 2.

Both ATTDeepPET and DeepPET showed lower performance compared to ML-EMR and ML-EM when trained using real data. However, upon examining the reconstructed images (Fig. 3), it becomes apparent that the deep learning models produce smoother images with reduced noise compared to the ML-EMR and ML-EM approaches. This suggests that the reconstructions may bear a closer resemblance to the ground truth. However, there is a noticeable reduction in the definition of cerebral gyri in these reconstructions, which results in an overall decrease in image fidelity to the ground truth.

Method	SSIM	PSNR	NRMSE
FBP	0,3966	25,1741	0,2213
ML-EM	0,8939	30,3253	0,1358
ML-EM regularized	**0,9078**	**30,5839**	**0,1211**
DeepPET	0,8030	26,8114	0,1707
ATTDeepPET	0,8776	27,1281	0,1648

Table 2. *Average error for the brain phantom dataset containing lesion reconstruction.*

To further investigate this phenomenon, a region of interest (Fig. 4) was selected. It becomes apparent that ATTDeepPET struggles to precisely reconstruct the defining boundaries of the cerebral cortex. Measurements of SSIM, PSNR, and NRMSE were computed by comparing the encoder-decoder models with the ML-EMR, which showed the best accuracy (Table 3).

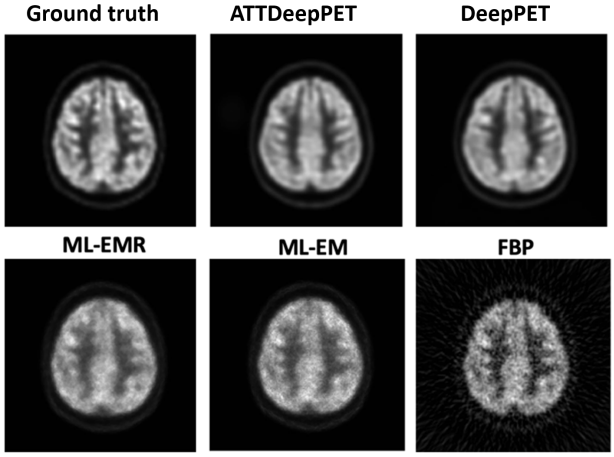

Figure 3. *Real PET reconstructed axial slices employing FBP, ML-EM, ML-EMR, DeepPET and ATTDeepPET.*

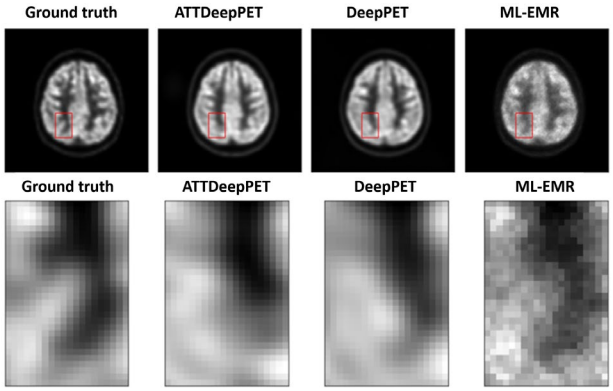

Figure 4. *Cerebral gyri comparison.*

Additionally, the analysis revealed the presence of notable outliers characterized by elevated nRMSE values. A closer examination of these outlier images uncovered a consistent pattern: they predominantly featured images with minimal brain content, particularly in the axial plane. These outlier images were primarily composed of slices located at the extremities of the head structure.

Method	SSIM	PSNR	NRMSE
ML-EMR	0,7539	17,4505	0,1350
DeepPET	0,5609	15,8724	0,1896
ATTDeepPET	0,5002	15,6831	0,1797

Table 3. *ROI's SSIM, PSNR and NRMSE averages.*

3.3. Reconstruction times

The temporal efficiency of each reconstruction method was quantified by measuring and calculating the average execution times across all images. The encoder-decoder architectures, DeepPET and ATTDeepPET, demonstrated significant reductions in reconstruction time, highlighting their notable computational efficiency. DeepPET exhibited a remarkable processing speed. Compared to FBP, DeepPET achieved a substantial speed enhancement of approximately 1.72 times. In comparison to ML-EM and ML-EMR, DeepPET demonstrated acceleration rates of approximately 735.13 times and 1299.23 times, respectively. Similarly, ATTDeepPET presented acceleration rates of approximately 1.20 times compared to FBP, 511.47 times compared to ML-EM, and 906.73 times compared to ML-EMR.

4. Conclusions

The results of this study emphasize the effectiveness of direct image reconstruction methods utilizing DL. These methods demonstrate the capability of performing attenuation correction directly from sinogram data, eliminating the need for synthetic CT images in favor of MRI. This transition is advantageous as it reduces patient exposure to ionizing radiation and significantly reduces reconstruction times, potentially enabling real-time studies. However, while these models perform well with simulated brain data, their performance diminishes when applied to real-world image datasets, highlighting the challenges posed by authentic data. Enhancing the robustness of these models for real-world scenarios should be a priority. On the downside, the computational complexity of CNNs requires substantial resources and time for network training. These models also rely heavily on extensive and diverse training datasets, limiting their applicability in scenarios with limited or suboptimal data. Traditional methods, grounded in established mathematical principles, offer stability and predictability across various contexts. However, the opacity of modern techniques using encoder-decoder raises interpretability concerns. Furthermore, there is a limitation in terms of model generalization. Models trained specifically with brain data may perform poorly when applied to other anatomical regions. A promising direction could involve hybrid approaches, merging deep learning with traditional methods. These hybrid methods, including unfolded techniques for iterative image reconstruction or adaptive learning of frequency filters, aim to balance the strengths of both paradigms and might serve to overcome the limitations inherent in singular methodologies.

Acknowledgments

This work has been partially funded by the project S2022/BMD-7236 MINA-CM of the Autonomous Community of Madrid and project IEL.GKA.100 DIANIA of FUNDETEL-UPM.

References

[1] P. Suetens, "Nuclear medicine imaging," in Fundamentals of Medical Imaging, 2nd ed. Cambridge University Press, 2009, pp. 105–127. doi: 10.1017/CBO9780511596803. 006.

[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, doi: 10.1109/23.256736.

[4] P. Kinahan, B. Hasegawa, and T. Beyer, "X-ray-based attenuation correction for positron emission tomography/computed tomography scanners," *Semin. Nucl. Med.*, 33(3): 166–179, 2003. doi: 10.1053/snuc.2003.127307.

[5] A. J. Reader, G. Corda, A. Mehranian, et al, "Deep Learning for PET Image Reconstruction," *IEEE Trans. Rad. Plasma Med. Sci.*, 5(1): 1–25, Aug. 2020, doi: 10.1109/trpms.2020.3014786.

[6] T. Wang *et al.*, "Machine learning in quantitative PET: A review of attenuation correction and low-count image reconstruction methods," *Phys. Med.*, 76: 294–306, 2020. doi: https://doi.org/10.1016/j.ejmp.2020.07.028.

[7] F. Liu, H. Jang, R. Kijowski, T. Bradshaw, and A. B. McMillan, "Deep learning MR imaging–based attenuation correction for PET/MR imaging," *Radiology*, 286(2): 676–684, 2018. doi: 10.1148/radiol.2017170700

[8] I. Häggström, C. R. Schmidtlein, G. Campanella, and T. J. Fuchs, "DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem," *Med. Imag. Anal.*, 54: 253–262, 2019, doi: 10.1016/j.media.2019.03.013.

[9] N. Rufo Rafael de la Cruz, J. Ortuño Fisac, and G. Kontaxakis, "Deep learning methodologies for brain image reconstruction in positron emission tomography," in Actas del XL Congreso Anual de la Soc. Esp. Ing. Biomédica, Valladolid, Nov. 2022, pp. 337–340.

[10] 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.

[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 Research*, 11(1): 91, 2021, doi: 10.1186/s13550-021-00830-6.

[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, doi: 10.1109/TIP.2003.81